ALLOSTERY AND HYSTERESIS ARE COUPLED IN UDP-GLUCOSE DEHYDROGENASE

by

NATHANIEL RICHARD BEATTIE

(Under the Direction of Zachary A. Wood)

ABSTRACT

UDP-glucose dehydrogenase (UDGH) catalyzes the NAD^+ dependent oxidation of UDP-glucose to UDP-glucuronic acid. In humans, (h)UGDH is regulated by the feedback inhibitor UDP-xylose using an atypical allosteric mechanism. Briefly, UDP-xylose competes with UDP-glucose for the active site. Binding UDP-xylose triggers a conformational change from an active 32 symmetry hexamer (E) to an inhibited horseshoe shaped complex (E^{Ω}). The ligand induced conformational change is the result of altering the affinity of the interface between subunits, i.e. allostery. hUGDH also displays hysteresis, the slow activation of the enzyme upon the addition of substrate. Hysteresis is caused by the slow isomerization from an inactive (E*) to the active state. Crystal structures of unliganded, as well as UDP-glucose or UDP-xylose bound hUGDH identified a buried structural element (the T131 loop-α6 helix) as the potential allosteric switch. The allosteric switch connects the active site to hexamer interfaces and adopts a unique conformation depending on the ligation state of the enzyme. Here, we have conducted a series of studies to elucidate the role of the switch in both allostery and hysteresis and established that these phenomena are not only coupled but a feature of many UGDH proteins. Introducing the A136M

substitution at the apex of the α6 helix trapped the allosteric switch in the E state. Restricting the movement of the allosteric switch abolished both allostery and hysteresis; providing the first evidence these phenomena are coupled. The movement of the allosteric switch between the E and E^{Ω} states requires a substantial repacking of the protein core. This repacking was hypothesized to be facilitated by large packing defects surrounding the allosteric switch. Filling a packing defect that only exists in the E state of hUGDH with the A104L substitution also abolished both hysteresis and allostery supporting this hypothesis. Analysis of UGDH primary sequences suggested the large to small substitutions that resulted in these packing defects may serve as a motif to predict allostery in UGDH proteins. *Caenorhabditis elegans* UGDH was identified using this motif, and was shown to be both allosteric and hysteretic, showing the motif can predict allostery in UGDH proteins.

INDEX WORDS: Allostery, Hysteresis, Protein Core Packing, UDP-Glucose Dehydrogenase, Atypical Allostery Motif.

ALLOSTERY AND HYSTERESIS ARE COUPLED IN UDP-GLUCOSE DEHYDROGENASE

by

NATHANIEL RICHARD BEATTIE

B.S., Washington State University, 2013

A Dissertation Submitted to Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

ALLOSTERY AND HYSTERESIS ARE COUPLED IN UDP-GLUCOSE DEHYDROGENASE

by

NATHANIEL RICHARD BEATTIE

Major Professor: Zachary A. Wood

Committee: William N. Lanzilotta
Robert S. Phillips
Kelly W. Moremen

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

DEDICATION

While there have been many people to have a tremendous impact on my life, all pale in comparison to that of my mother Tammy Standfier, my grandfather Jack Standfier and my Aunt Kimberely Beattie-Stokes. Without the constant support of these three people, none of this would have been even close to possible.

ACKNOWLEDGEMENTS

I must start by acknowledging my mentor Dr. Zachary Wood. I have never met anyone with your passion for science and the innate ability to spread that passion to others. You kept me going through long nights and early morning, both at the bench chasing down data and behind a keyboard writing the manuscripts that make the chapters in this dissertation. During the early days of my graduate career, I found your insanely high standards endlessly frustrating. Now, as this phase of my education ends, I understand why you set the bar so high and am proud to have inherited your standards. You were always there to keep both my ego and critical nature in check, with honest assessments of my strengths and weaknesses. I have always been curious, but it was you who made me a scientist and for that I will be forever grateful.

My graduate education would not have nearly as fruitful without my intense and dedicated Committee members. Dr. Kelley Moremen, you were by far my most challenging committee member, bringing the hard-hitting questions and never yielding an inch that wasn't earned. I often tell other that heat and pressure is how you make diamonds, you brought them both and I am better for it. Dr. William Lanzilotta "Wild Bill" if anyone can rival Zac's passion for science it would be you. You always encouraged me to work hard, think both critically and outside the box. Dr. Robert Phillips, I have lost count of the number of times you have helped me in both the technical and theoretical aspects of enzymology. I would not have made it without your help, and to no small degree, your patience.

I had some very challenging professional experiences early in my graduate career, that almost ended in me leaving the program. The guidance of Dr. Nancy Manley, Dr. Walter Schmidt

and Dr. Lance Wells led me through those hard times and to Zac's lab. As luck would have it, Zac shares a floor with Walter, so I was never far from great advice when the need arose.

Brittany (with two t's) J. Pioso aka "Minion of Darkness" of all my accomplishments in grad school, being your mentor is the one I am the proudest of. Without your incredible work ethic and endless resolve, Chapter three of this dissertation would not have been possible. I am insanely proud of the scientist you became and the small role I was able to play in it. I can't wait to see where your journey takes you. Always remember our motto!

The work presented here would not have been possible without my lab mates. The most impactful would have to be my friend and colleague Dr. Renu Kadirvelraj, the "Rage Raj" herself. We always got along well for two people with such hot tempers. Your dedication to teaching others, although you will never admit to having it, helped me become an excellent structural biologist. Tiffany Hicks Sirmans, aka "The Lost Child", I cannot describe the magnitude of your raw potential. You were a pleasure to mentor and work with. Your efforts resulted in a hard-earned publication that is presented here in Chapter four.

Reflecting on the path that lead me to this point, I must acknowledge three people who put a tremendous amount of faith in me when, to be honest, they had absolutely no reason to. My undergraduate mentor Dr. Cliff Berkman, he took a wild hothead into his lab and showed me I had the potential do more. Dr. Greg Crouch, you are without a doubt the single most important teacher I ever had. You made Organic Chemistry come alive and instilled the love for science that made this possible. When I needed a job prior to starting graduate school, you were able to convince the Chemistry Department to hire me. You and Cliff always had time for me, to offer sage advice and mentorship. Finally, I must acknowledge the biggest gambler of all, Dr. Consetta Helmick. You got departmental permission to allow me to TA your microbiology lab despite the fact I didn't

have the grades for it. You put faith in me, taught me how to be a teacher, and guided me along the way. I cannot thank you enough. The three of you wrote the letters and taught the lessons that got me into graduate school, without the impact you made on my life this would have never happened.

I must acknowledge the endless support of my family during the crazy journey that was graduate school. My mother Tammy, her fiancé Nick, my Grandfather Jack and My aunts Kim and Terry. No matter how stressful things became I always knew you were all just a phone call away and that got me through this.

Finally, I must acknowledge my girlfriend, teammate and partner in crime, Dr. Courtney Daczkowski. From the moment I met you, I was completely smitten. You have been my rock and inspiration to be the absolute best I can be. We went through some rough times and came out each one stronger. I would not have made it this far without you and look forward to the next chapter of our life together.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Glucuronidation is the process of covalently attaching the glucuronic acid moiety from UDP-Glucuronic acid (UDP-GlcA) to a target compound.[1-3] Glucuronidation has been shown to both alter the pharmacokinetics and solubility of the target compound.[2, 4, 5] Some lung, breast and colorectal cancers have been shown to use this pathway as a means of chemotherapeutic resistance.[6-9] Threating these cancers is complicated by the fact that the enzymes that transfer the glucuronic acid, glucuronosyltransferases, have over 28 isoforms with a broad range of substrate specificities.[3, 10-12] However, UDP-GlcA is only produced by a single enzyme, UDP-glucose dehydrogenase (UGDH).[13] Understanding the regulatory mechanism of UGDH would be a key step in regulating glucuronidation as whole. The work presented here will build upon our current understanding of how the structural features of UGDH impact the regulation this enzyme.

Before a detailed discussion on the regulation of human (h)UGDH can be had, it is important to establish clear definitions of specific terminology. Cooperativity is the thermodynamic coupling between any binding sites in a protein or oligomer.[14, 15] This definition is designed to separate thermodynamic cooperativity, which arises from a ligand stabilizing an alternate conformation altering the affinity or activity of a protein, from kinetic cooperativity which is conceptually distinct and will not be discussed.[14-18] Cooperativity can be either positive, where one binding increases the affinity of the next, or negative, where binding decreases the affinity of the subsequent events.[15, 19-21] Positive cooperativity is by far the most common and best

studied; however, both instances are observed as a non-hyperbolic response to ligand.[14, 18, 22-26] One of the best studied, examples of positive cooperativity is of oxygen binding to hemoglobin.[16, 22, 27-32] Hemoglobin is a 4-subunit, heme dependent, oxygen binding protein that consists of 2 αβ dimers (Figure1.1.). [29-31] The structure of hemoglobin has been described as existing in equilibrium between a high oxygen affinity R state and low affinity T state. [29-31, 33] When oxygen is low, the T state dominates the equilibrium. [29-31, 33] Binding of oxygen at one of the four sites triggers a conformational change to the high affinity R state, giving rise to positive cooperativity.[14, 16, 29, 30, 34-36] The positive cooperativity displayed by hemoglobin binding oxygen is often misidentified as allostery.[16] Allostery is a subtype of cooperativity and has specifically defined criteria. Allostery as defined by Monod requires two unique sites, one binding site and a distal, functionally distinct, allosteric site.[16, 17, 34-36] Thus, hemoglobin binding oxygen is not allosteric as the sites are not distinct, they are functionally identical symmetry related sites (Figure1.1.). In contrast, the small molecule 2,3-bisphosphoglycerate (2,3-BPG) is a true allosteric effector of hemoglobin with respect to oxygen binding.[37] 2,3-BPG binds to a unique site between the β subunits of hemoglobin and stabilizes the low affinity T state (Figure1.1.).[14, 37, 38] In this way, 2,3-BPG is an allosteric effector of oxygen binding, because the two sites and are unique but thermodynamically coupled. It is worth noting that no part of the definition of allostery prohibits the active site, from also serving as an allosteric site so long as binding there alters the function or affinity at a unique distal site.[16, 17, 34, 36] This distinction will be important when discussion atypical allostery in hUGDH as the active site doubles as an allosteric site.[39] Allosteric responses are also not limited to small molecule binding. The Lac repressor was an early example of allostery, where binding of lactose to the allosteric site altered the affinity at the DNA binding site.[16, 17, 35, 36, 40] Any

time there are distal, thermodynamically coupled and functionally distinct site on a protein they are allosteric.[16, 17, 34]

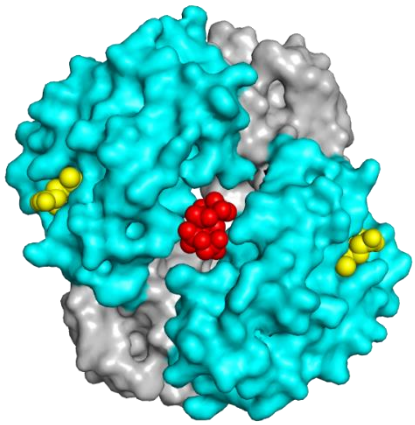


Figure 1.1: Crystal Structure of Hemoglobin with 2,3-Bisphosphoglycerate Bound (PDB ID:5KSI). The structure contains 2 α(grey)β(cyan) subunits. The heme at the oxygen binding sites are shown by as yellow spheres while 2,3-bisphosphoglycerate bound to the allosteric site is shown by red spheres.

hUGDH catalyzes the two-fold, NAD^+ dependent, oxidation of UDP-glucose (UDP-Glc) to UDP-GlcA.[13, 41-43] In the absence of substrate, hUGDH is an inactive and weakly associated *32* symmetry hexamer (Figure 1.2.A).[39, 44, 45] Each monomer is comprised of an NAD^+ and UDP-Sugar binding domain as well as a dimerization domain (known as NB, SB and DD respectively) with the smallest functional unit being a homodimer (Figure 1.2.A).[42, 46] Binding of substrate drives the formation of the active conformation, known as the E state (Figure 1.2.B).[45] In contrast, binding of the atypical allosteric inhibitor UDP-xylose (UDP-Xyl) drives the formation of the inhibited (E^{Ω}) complex (Figure 1.2.B).[39, 47] Here, allostery is atypical because the active site doubles as the allosteric site. Briefly, UDP-Xyl binding to the active site alters the conformation of the interface between adjoining dimers in the hUGDH hexamer; referred to as the hexamer building interface (HBI) (Figure 1.2.C).[39, 44, 45, 48] Binding of UDP-Xyl at the active site has been shown to select for the high affinity conformation of the HBI.[39, 44, 45, 48] In this respect, the active site and HBI are functionally distinct but thermodynamically coupled, resulting in allostery.

Analytical ultracentrifugation experiments on hUGDH have shown that UDP-Xyl binding drives the formation of the E^{Ω} hexamer while UDP-Glc binding has little impact on the subunit affinities.[39, 44, 45, 48] The distinction that UDP-Xyl binding to the active site has a unique effect is important, as allosteric responses have often been described as specific.[14, 16, 17, 34-36]

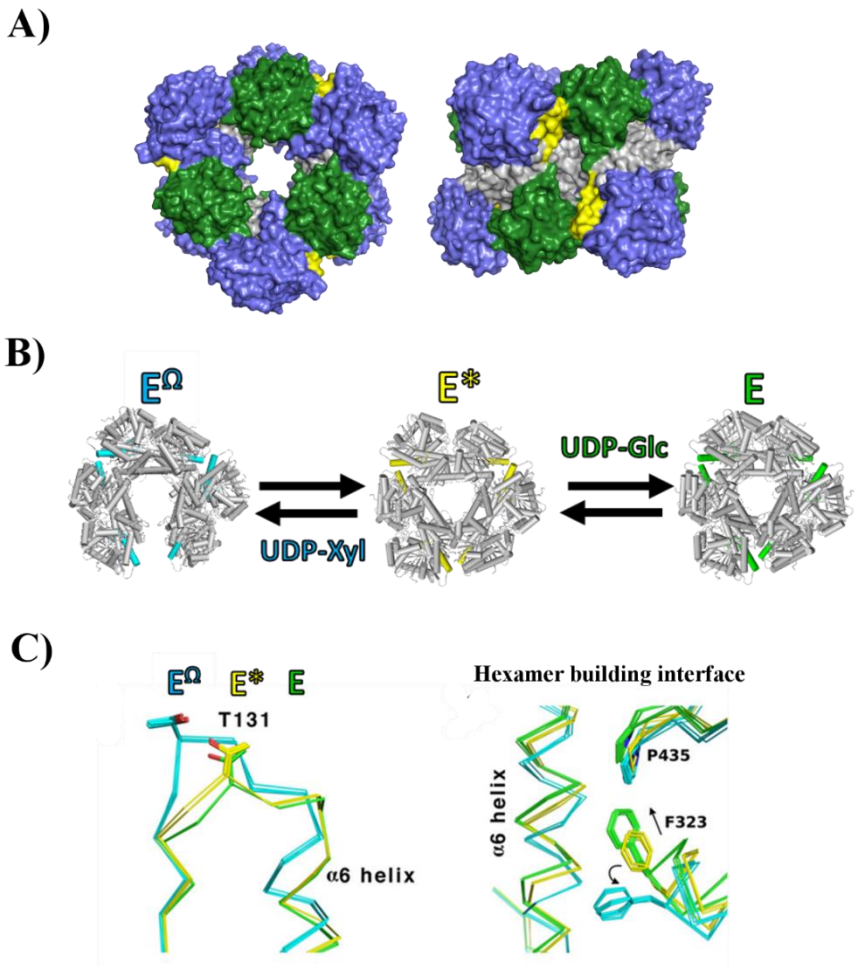


Figure 1.2: The Structure of hUGDH and Conformations of the Allosteric Switch. A) A surface rendering with a view from the top (left) and side (right) of the hUGDH 32 symmetry hexamer with the NB, Allosteric switch, DD and SB shown in slate, yellow, grey and green, respectively. B) The change in oligomeric structure of hUGDH depending on which ligand is bound to the active site. From left to right are the UDP-Xyl bound inhibited complex, unliganded E* and substrate bound E state, with the allosteric switch colored cyan, yellow and green, respectively. C) The change in position of the allosteric switch (left) and rearrangement of the hexamer building interface (right) depending on ligation state of the active site, colored as in B.

Crystallographic studies have identified three distinct conformations of hUGDH depending of what, if anything, is bound to the active site (Figure 1.2.B). In the absence of ligand, hUGDH exists in an inactive conformation known as E*.[39, 42, 44, 45, 48] When substrate binds, the enzyme isomerizes to the active E state (Figure 1.2.B).[45, 48, 49] In contrast, UDP-Xyl binding induces the transition to the inhibited E^{Ω} complex (Figure 1.2.B).[39] A comparison of the three structures has shown key differences in both domain rotation and the position of the T131 loop-α6 helix.[44, 48] In the absence of ligand, the NB domain is rotated away from the SB domain by approximately 10˚.[45] This has been called the "open" conformation of hUGDH and is believed to relieve packing constrains in the protein core.[39, 45, 48] In both the substrate and inhibitor bound structures the NB domain is rotated to the closed conformation.[39, 42, 45] The second key observation is the buried T131 loop-α6 helix, adopts a unique conformation depending on the ligation state of the enzyme (Figure 1.2.C). [39, 42, 45] This element was hypothesized to act as the "allosteric switch" as it connects the active/allosteric site to the hexamer building interface (Figure 1.2.B&C).[39, 44, 48]

hUGDH displays a pronounce lag, hysteresis, in progress curves (Figure 1.3.) [39, 49-51] It has been robustly shown that this lag is due to an isomerization from an inactive state to an active one.[48, 49] An isomerization is an observable structural change and to be hysteretic it must occur on the seconds timescale.[52] When combined with the structural data, it was hypothesized that hysteresis in hUGDH is the result of the slow isomerization of the allosteric switch from the unliganded E* to the E state.[44] Hysteresis can also be observed in the transition to the inhibited E^{Ω} complex upon UDP-Xyl binding.[48] This observation provided the first evidence that hysteresis and allostery are coupled but required a more robust interrogation before definitive conclusions could be made.

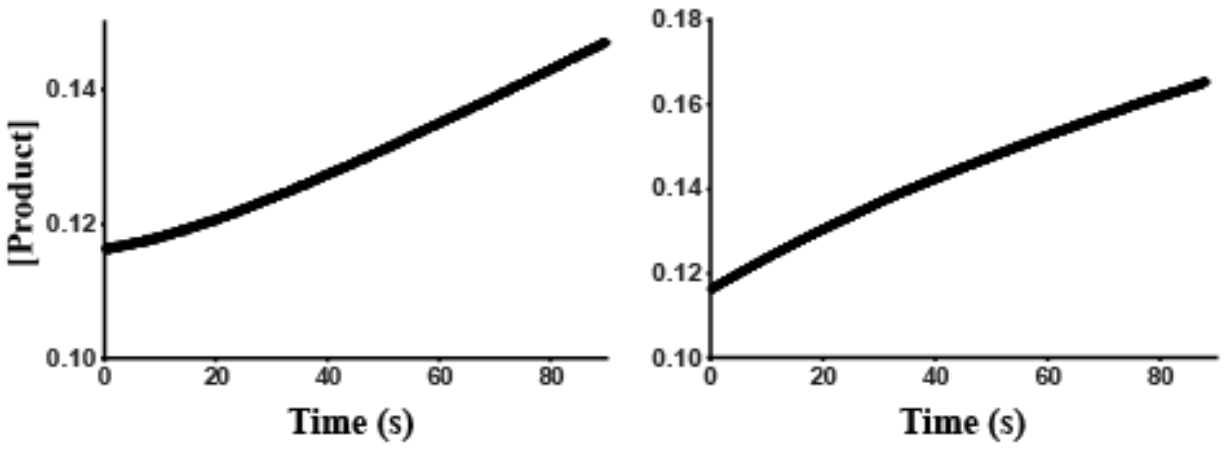


Figure 1.3: **Progress Curves for a Hysteretic and Non-Hysteretic Enzyme**. The left progress curve displays hysteresis; the right is non-hysteretic.

The second chapter of this dissertation provides direct evidence that the T131 loop-α6 helix is in fact the allosteric switch in hUGDH that controls conformation of the enzyme. It further shows that hysteresis in hUGDH is coupled to the movement of the allosteric switch. When hUGDH transitions from the E state to E^{Ω}, the α6 helix rotates 12˚ into the protein core.[39] Alanine 136 is located at the apex of the α6 helix and was targeted for substitution to methionine. The bulky substitution would be accommodated in the local packing environment of the E state but be sterically prohibited in the E^{Ω} conformation. Crystallographic studies showed that in the absence of ligand hUGDH with the A136M substitution ($hUGDH_{A136M}$) adopted the active E state. This was supported by the absence of a lag in $hUGDH_{A136M}$ progress curves, suggesting the E* conformation of the allosteric switch is important for hysteresis. As expected, the substitution also prevented the transition to the E^{Ω} complex as evident by the loss of inhibitor induced cooperativity. Briefly, the E^{Ω} conformation has a lower affinity for UDP-Glc than the E state.[48] When UDP-Xyl concentrations are high, the low UDP-Glc affinity E^{Ω} complex dominates the population. As the concentration of UDP-Glc is increased, the high affinity E state becomes the dominant species. While this is an indicator of the allosteric transition to the high HBI affinity E^{Ω} complex is

occurring, it should be made clear that UDP-Xyl is not an allosteric effector with respect to UDP-Glc as the ligands bind to the same site. UDP-Xyl is simply a competitive inhibitor of hUGDH that induces positive cooperativity with respect to UDP-Glc binding. Taken together this work was the first direct evidence that the movement of the allosteric switch was required for the isomerization between the E*, E and E^{Ω} states of hUGDH and suggested that the allosteric mechanism maybe coupled to hysteresis.

Establishing that the movement of the allosteric switch as the source of both allostery and hysteresis in hUGDH was big step forward; however, that work provided no mechanistic information on how the transition occurs within the confines of the protein core. To truly appreciate the significance of the isomerization of the allosteric switch, it is important to have a basic understanding of the nature of protein cores. The key driving force of protein folding is the entropy of water, that is to drive proteins to the smallest possible structure that requires the least number of ordered waters to solvate.[53-55] To accomplish this, the hydrophobic side chains of peptides are tightly packed together forming the protein core.[56] The tight packing of atoms in the protein core results in a large number of stabilizing Van der Waals contacts.[56, 57] Constraints on folding are to prevent the burial of unsatisfied polar or charged groups, and maintenance of the peptide plane and other stereo chemical restrains.[28, 58-60] There are two current models that attempt to explain the how hydrophobic sidechains pack together within the protein core. The "jigsaw puzzle" model describes the protein core as rigid, with the side chains of amino acids being packed together in fixed conformations.[61, 62] Support for this model came from the observations that 1) in crystal structures the packing of the core approached the closest packing of spheres 2) protein cores are typically highly ordered 3) evolution of the core is often slow and mutations in it tend to be destabilizing.[63-68] By this model, if a bulky phenylalanine were substituted to an alanine, the

structure would not relax to fill the space and a void volume would persist in the core where benzyl group should have been. This would result in a loss of favorable Van der Waals contacts that would destabilize the folded protein.[57, 69] The second model of the protein core is known as "nuts and bolts in a jar", and proposes that the interactions of amino acid side chains are not so rigidly defined.[64, 65, 70, 71]. This model takes into account the entropic cost of restricting side chain conformational freedom. [71]. It proposes that the small void spaces in the closets packing of spheres observed in protein cores allows side chains to move, reducing the entropic cost of confinement.[71] Using the previous phenylalanine to alanine substitution, the backbone and sidechain in a nuts and bolts core would move slightly to minimize the void space from the lost benzyl group and the small remaining space would be less entropically unfavorable than that in a jigsaw puzzle core. This model stresses the importance of Van der Waals contacts and side chain entropy over the maintenance of a predefined static interactions.[64, 65, 70, 71] The modern view of protein core packing is a mix of the two models, that is some areas of the core have more puzzle piece character while others are more nuts and bolts.[65, 72] This model explains why in some instances large to small substitutions generate packing defects of the predicted size, and in others the void is at least partially filled by the protein.[57, 64-68]

In hUGDH the transition of the allosteric switch between the E and E^{Ω} conformations requires a remarkable repacking of the protein core that includes 1) an approximately 5Å shift in the Cα of residues T131 into the protein core 2) eight buried residues to change rotamer and 3) the 180˚ inversion of the peptide backbone around P132.[39] Comparing crystal structures from UGDH proteins that are known to be non-allosteric, like *Streptococcus pyrogenese* (spUGDH), to hUGDH structures reviles stark differences in protein core packing (Figure 1.4.A). hUGDH has six large to small mutations that result in packing defects surround the allosteric switch which are not observed

in spUGDH (Figure 1.4.B). Packing defects are a general term for any area in the protein where a water molecule could fit between the Van der Waals radius of protein atoms.[63] There are two varieties of packing defect, declivities which have direct access to bulk solvent and cavities which are completely incased within the protein and do not access bulk solvent.[63, 73] Due to the six large to small substitutions, the active state of hUGDH has two large declivities and two cavities unique to the active state that surround the allosteric switch.[39] It was hypothesized that these packing defects provide the space required for hUGDH to undergo the allosteric transition.[39]

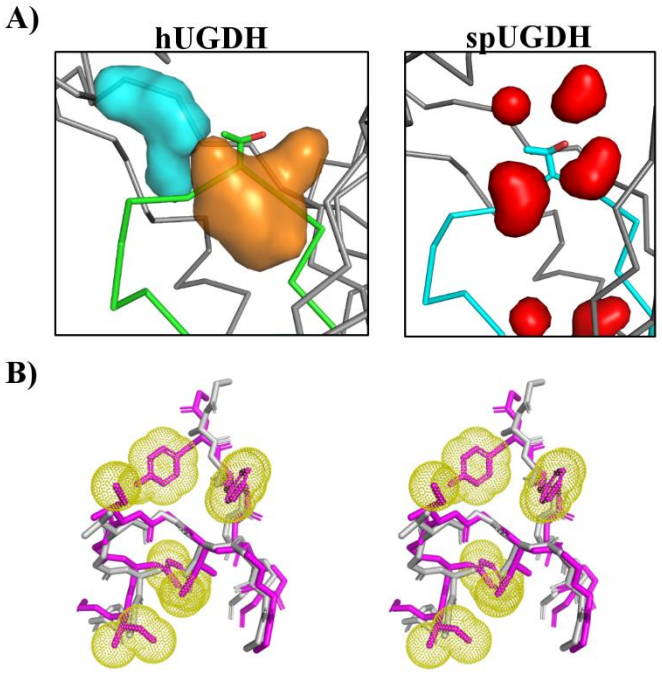


Figure 1.4: Evolution of Packing Defects in hUGDH. A) A comparison of protein core packing in hUGDH and spUGDH. Cavities are shown as colored surfaces and the allosteric switch is shown in green and cyan for hUGDH and spUGDH, respectively. Declivities in hUGDH are not shown for clarity. B) a cross-eye stereo view of the large to small substitutions in spUGDH (magenta) compared hUGDH (grey). Lost Van der Waals are shown as yellow dots surrounding the atoms in spUGDH.

Chapter three of this dissertation is aimed at understanding how the allosteric switch is able to isomerize between the E and E^{Ω} state within the confines of the protein core. The Rosetta Void Identification and Packing[74] ($Rosetta_{VIP}$) protocol was used to computationally design an ideal substitution to fill a packing defect found only in the E/E* state of hUGDH. The A104L

substitution ($hUGDH_{A104L}$) reduced the volume of a cavity in the E state by over 50%, while still allowing the formation of the E* state as the packing defects between these conformations are similar. The bulky substitution sterically prohibited $hUGDH_{A104L}$ from accessing the E^{Ω} state, as evident by the low affinity for UDP-Xyl and the absence of cooperativity in substrate saturation kinetics in excess inhibitor. Interestingly, $hUGDH_{A140L}$ did not show hysteresis even though crystal studies showed that it was able to adopt the E* conformation in the absence of ligand. This was attributed to the bulky A104L substitution reducing the conformational entropy of the protein core. In other words, the bulky leucine residue reduced the number of rotamers neighboring amino acids were able to adopt. With the reduced search space, the kinetic barrier separating the E* from the E state is lowered, eliminating hysteresis. This was supported by a B-factor analysis of hUGDH and $hUGDH_{A104L}$ crystal structures in E* conformations, which showed an increase in order surrounding the substitution. This data showed that the packing defects in the protein core of hUGDH were required for both hysteresis and allostery, further supporting the idea these features are coupled. The previous hypothesis for the origin of hysteresis, that focused exclusively on the movement of the allosteric switch, was expanded to take protein core entropy into account. This work showed the ability to adopt the E* state in crystal structures was not sufficient to induce hysteresis but is likely due to the increased conformational entropy of the protein core as result of the packing defects required for the allosteric transition.

Chapter four of this dissertation is aimed at both understanding the prevalence and identifying the predictive features of atypical allostery in homologous UGDH proteins. Using the six large to small substitutions previously identified as critical for hysteresis and allostery, a detailed phylogenetic analysis was conducted on non-redundant UGDH protein sequences form a wide range of organisms. This study predicted the UGDH from *Caenorhabditis elegans* (*C.*

elegans and cUGDH), which conserved 5 of the 6 large to small substitutions, would be an allosteric protein. Crystal structures and analytical ultracentrifugation data showed that like hUGDH, cUGDH was a weakly associated *32* symmetry hexamer that upon UDP-Xyl binding transitioned to a strongly associated horseshoe shaped complex. Stopped flow spectroscopy was used to show that cUGDH also displays a predominant hysteretic lag upon initiation of the reaction. Taken together this data shows that cUGDH conserves the atypical allosteric mechanism of hUGDH. This data showed that the six large to small substitutions could be used as a predictive "atypical allostery motif" for identifying other allosteric UGDH proteins. The most striking difference between hUGDH and cUGDH was the reduced affinity for the inhibitor UDP-Xyl. A comparison of both crystal structures bound to UDP-Xyl showed the active sites are identical. The most striking difference between the two structures is the degree of rotation in the α6 helix. The position of the α6 helix directly contributes to the packing interactions of the hexamer building interface, which is also unique between the two proteins. As the interface and the active site are thermodynamically coupled with respect to UDP-Xyl binding (i.e. allostery) the differences in the hexamer building interfaces could explain the altered affinity for inhibitor.

In summary, the work presented here answered critical questions about the atypical allosteric mechanism of hUGDH. It shows that both allostery and hysteresis are coupled via the movement of a buried allosteric switch within the protein core. The remarkable conformational changes the allosteric switch undergoes are made possible by the presence of large packing defects within the protein core. These packing defects are now believed to increase the conformational entropy of hUGDH, contributing to hysteresis. Six large to small hydrophobic mutations surrounding the allosteric switch were conserved in many UGDH proteins and have now been hypothesized to be a motif that can predict atypical allostery. It was shown that these substitutions

are sufficient to predict the conservation of atypical allostery and hysteresis, using cUGDH as an example protein. The broader impacts of this work as well as the future directions of these projects will be discussed in the fifth chapter of this dissertation

Finally, I would be remiss not to introduce the sole appendix of this dissertation, to which my contributions were small but sufficient to earn co-authorship. hUGDH has a 30 residue, intrinsically disordered C-terminal tail (ID-tail). These elements are often removed to aid in crystallization as they were believed to be nonfunctional.[75-77] Here, it was shown that the ID-tail works in concert with the allosteric switch and plays a critical role in the allosteric mechanism. The effect of the tail was found to be exclusively a property of its length with no requisite for folded structure or amino acid content. This was interpreted as an entropic effect of the ID-tail being exerted on the protein. Put simply, the tail acts as a polymer tethered to the protein. The protein surface restrains the conformational freedom of the protein exacting an entropic cost. The longer the polymer, to an extent, the greater the cost and force exerted on the protein to restrain it.[78, 79] HDX mass spectroscopy showed the presence of the tail increased the dynamics of the allosteric switch, supporting the idea of a distant regulatory effect. Strong orthogonal evidence came from both Monte Carlo simulations and a transient state kinetic model of UDP-Xyl binding. These techniques were used to quantify the energy required to restrain the conformations of the ID-tail. The energies predicted by both techniques agreed well and supported the hypothesis that the ID-tail served as an entropic rectifier, that tuned the affinity of hUGDH for UDP-Xyl in a length dependent manner. While this work provided no insight into the coupling of the allosteric mechanism with hysteresis, it was significant to both allosteric regulation in hUGDH as well as the function of ID segments in the proteome as a whole.

CHAPTER 2

ALLOSTERY AND HYSTERESIS ARE COUPLED IN HUMAN UDP-GLUCOSE DEHYDROGENASE

Reprinted here with permission from: **Beattie, N.R**., Keul, N.D., Sidlo, A.M. and Wood, Z.A., Allostery and Hysteresis are coupled in human UDP-glucose dehydrogenase. *Biochemistry*, 56(1), pp.202-211 (2016).

2.1 Abstract

Human UDP-glucose dehydrogenase (hUGDH) is regulated by an atypical allosteric mechanism in which the feedback inhibitor UDP-xylose (UDP-Xyl) competes with the substrate for the active site. Binding of UDP-Xyl triggers the T131-loop/α6 allosteric switch, which converts the hexameric structure of hUGDH into an inactive, horseshoe-shaped complex (E^{Ω}). This allosteric transition buries residue A136 in the protein core to produce a subunit interface that favors the E^{Ω} structure. Here we use a methionine substitution to prevent the burial of A136 and trap the T131-loop/α6 switch in the active conformation. We show that hUGDHA136M does not exhibit substrate cooperativity, which is strong evidence that the methionine substitution prevents the formation of the low-UDP-Glc-affinity E^{Ω} state. In addition, the inhibitor affinity of $hUGDH_{A136M}$ is reduced 14-fold, which most likely represents the K_i for competitive inhibition in the absence of the allosteric transition to the higher-affinity E^{Ω} state. hUGDH also displays a lag in progress curves, which is caused by a slow, substrate-induced isomerization that activates the enzyme. Stopped-flow analysis shows that $hUGDH_{A136M}$ does not exhibit hysteresis, which suggests that the T131-loop/α6 switch is the source of the slow isomerization. This interpretation is supported by the 2.05 Å resolution crystal structure of $hUGDH_{A136M}$, which shows that the A136M substitution has stabilized the active conformation of the T131-loop/α6 allosteric switch. This work shows that the T131-loop/α6 allosteric switch couples allostery and hysteresis in hUGDH.

2.2 Introduction

Glucuronidation is a major component of the phase II metabolism of drugs. [11, 80, 81] This pathway has been shown to be upregulated in some lung, breast, and colorectal cancer cells, where it functions as an intrinsic drug resistance mechanism.[6-9] Thus, controlling glucuronidation is a promising strategy for sensitizing this class of tumors to existing chemotherapeutics. It is possible to inhibit glucuronidation by limiting the availability of the essential substrate, UDP-glucuronic acid.[82-84] UDP-glucose dehydrogenase (UGDH) catalyzes the NAD^+-dependent oxidation of UDP-glucose (UDP-Glc) to produce UDP-glucuronic acid.[41, 42, 51] Understanding how UGDH activity is regulated is an important goal in developing new strategies for controlling glucuronidation-dependent drug resistance.

Human UGDH (hUGDH) forms a hexamer that is regulated by an atypical allosteric mechanism in which the feedback inhibitor UDP-xylose (UDP-Xyl) competes with the substrate for the active site.[39, 44, 45, 48] Here, allostery arises from the distinct conformational changes that are induced by the binding of the substrate or inhibitor (Figure 2.1.A). The binding of UDP-Glc favors the formation of an active 32 symmetry hexamer called the E state. In contrast, UDP-Xyl binding produces an inactive, horseshoe-shaped hexamer (E^{Ω}). These effector specific transitions are controlled by a buried allosteric switch called the T131-loop/α6 helix (Figure 2.1.A–C). To select for a specific hexamer conformation, the NAD^+ binding domain (NB) of hUGDH must adopt an "open" conformation (the E* or E* bound state) that exposes the allosteric switch (Figure 2.1.D). Once exposed, the T131-loop and α6 helix are free to change conformation in response to UDP-Glc or UDP-Xyl binding to produce the E or E^{Ω} hexamer, respectively (Figure 2.1.D).[39, 44, 48] Thus, the active site of hUGDH also functions as an allosteric site in that it controls the structure of the

hexamer-building interface and the affinity between subunits.[16, 17] To the best of our knowledge, the only other enzyme known to have a similar bifunctional active/allosteric site is dCTP-deaminase.[85, 86]

hUGDH also displays hysteresis, which can be observed as a lag in progress curves.[39, 44, 45, 48-50] In the absence of any effector, hUGDH favors the inactive E* conformation (Figure 2.1.D).[48] Hysteresis is caused by coenzyme and substrate binding, which induces the E* state to slowly isomerize to the active E conformation.[49] Still, the molecular basis of the hysteretic transition was not known. E* adopts an open domain structure that stabilizes the allosteric switch and the hexamer-building interfaces in an intermediate conformation between the E and E^{Ω} states (Figure B1.B–D).[42, 45, 48] When an effector binds to E*, the T131-loop/α6 helix and eight surrounding residues repack into the appropriate E or E^{Ω} conformation (Figure 2.1.D).[48] We have proposed that the repacking of the allosteric switch and adjacent residues from the E* to E conformation is the source of hysteresis observed in hUGDH.[49] Here we have used an amino acid substitution (A136M) to stabilize the E state and test our hypothesis.

2.3 Materials and Methods

Protein Expression, Crystallization, and Structure Solution

$hUGDH_{A136M}$ and wild-type hUDGH were recombinantly expressed in Escherichia coli as previously described.[39, 44, 45, 48, 49] Following purification, the His tags were removed from both proteins using TEV protease. The purified protein was dialyzed into storage buffer [25 mM Tris (pH 8.0) and 50 mM NaCl] and concentrated using a Millipore Amicon Ultra-15 10K centrifugal filter unit to ~20 mg/mL. Proteins were quantified using an Agilent 8453 UV/vis instrument with an ε280 of 48360 M^{-1} cm^{-1} and an ε280 of 49850 M^{-1} cm^{-1} for hUGDH and $hUGDH_{A136M}$, respectively. The molar absorptivities for these proteins were calculated using the amino acid

sequence and Protparam.[87] hUGDH$_{A136M}$ was crystallized at 20 °C using a hanging drop vapor diffusion method with a 2 μL drop mixed in a 1:1 ratio of protein to reservoir (final protein concentration of 10 mg/mL). The reservoir was comprised of 0.2 M NaCl, 12% PEG 3350, and 0.1 M Tris buffer (pH 7.6). Crystals were cryoprotected using a solution matching the reservoir supplemented with 20% of a cryoprotectant mixture (a 1:1:1 dimethyl sulfoxide:ethylene glycol:glycerol ratio) and then cooled by being plunged into liquid nitrogen. A 2.05 Å resolution data set was collected on the 21-ID beamline (SER-CAT) at the Argonne National Laboratory (Argonne, IL) using a Rayonix MX300HS CCD detector. The data set was processed using XDS[88] and 5% of the data set aside for cross validation.[89] The data collection statistics are listed in Table 2.1.

The structure of hUGDH$_{A136M}$ was solved by molecular replacement using Protein Data Bank (PDB) entry 4RJT as a search model in the PHENIX software suite.[90] The hUGDH$_{A136M}$ model was subjected to iterative cycles of manual rebuilding using COOT[91] followed by automated refinement with NCS restraints. *B* factors were refined using TLS as implemented in PHENIX.[90, 92] We used the following rationale in modeling the occupancy of the copurified UDP-Glc. At resolutions lower than 1.5 Å, B factors and occupancy are strongly correlated, which makes it impossible to refine both of these parameters. Still, it is possible to approximate occupancy because ligands and interacting residues are similarly ordered. Briefly, the occupancy of the ligand is fixed at different values, and only the *B* factors are refined. The approximate occupancy is the value at which the ligand *B* factors converge to those of the interacting residues. Final model refinement statistics are listed in Table 2.1.

Sedimentation Velocity

hUGDH and $hUGDH_{A136M}$ were dialyzed into buffers containing 150 mM KCl and 25 mM buffer [HEPES (pH 7.5) or TRIS (pH 8.5)], quantified (as described above), and then diluted to a final protein concentration of 9 μM. Samples were loaded into 12 mm double-sector Epon centerpieces equipped with quartz windows and equilibrated for 1 h at 20 °C in an An60 Ti rotor. Sedimentation velocity data were collected in an Optima XLA analytical ultracentrifuge using a rotor speed of 50000 rpm at 20 °C. Sedimentation data were recorded at 280 nm in radial step sizes of 0.003 cm. SEDNTERP[93] was used to estimate the partial specific volume of hUGDH (0.73840 mL/g) and $hUGDH_{A136M}$ (0.73842 mL/g), and the densities of both the pH 7.5 (1.00726 g/mL) and pH 8.5 (1.00603 g/mL) buffers. Viscosities for both pH 7.5 and 8.5 buffers were calculated to be 0.01018 and 0.01007 P, respectively. SEDFIT[94] was used to analyze raw sedimentation data. Data were modeled as a continuous sedimentation coefficient, *c(s)*, distribution and were fit using the baseline, meniscus, frictional coefficient, and systematic time-invariant and radial-invariant noise. Theoretical sedimentation coefficient (s) values were calculated from hUGDH atomic coordinates under standard conditions using HYDROPRO20 (www.bbri.org/RASMB).

Stopped-Flow Analysis of Hysteresis

NAD^+ and UDP-glucose were purchased from Sigma, and UDP-xylose was purchased from Carbosource (Complex Carbohydrate Research Center, University of Georgia, Athens, GA). Enzyme hysteresis was monitored at 25 °C using an Olis RSM 1000 Rapid-scanning Absorbance and Fluorescence Spectrophotometer with a stopped-flow assembly and a 0.4 mm path length. The 480 nM enzyme solution in 50 mM HEPES (pH 7.5), 50 mM NaCl, and 5 mM EDTA was rapidly mixed with an equal volume of 1 mM UDP-Glc and 10 mM NAD^+ in the same buffer [final concentrations of 240 nM enzyme, 0.5 mM UDP-Glc, 5 mM NAD^+, 50 mM HEPES (pH 7.5), 50

mM NaCl, and 5 mM EDTA]. Absorbance readings at 340 nm were taken every 0.1 s for a total of 120 s to monitor NADH production. hUGDH progress curves were fit to Frieden's equation[52] describing enzyme hysteresis:

$$P(t) = v_{ss}t - \tau(v_{ss} - v_i)(1 - e^{-t/\tau}) \qquad \text{(equation 1)}$$

where P is the concentration of the product at time t and τ is equal to $1/k_{obs}$, where k_{obs} is the apparent rate constant for the transition between the initial velocity (v_i) and the steady state velocity (v_{ss}). Here, the initial velocity, v_i, in eq 1 is the pre-steady state velocity and does not obey the steady state approximation. The length of the lag in seconds is calculated as the product of Euler's number and τ ($e\tau$). Data were fit using PRISM (GraphPad Software Inc., San Diego, CA), and all fits were analyzed using residual plot analysis.[95]

Steady State Kinetics

The assay conditions for the steady state analysis of hUGDH have been described previously.[39, 44, 45, 48, 49] Briefly, assays contained 100 nM hUDGH or 230 nM $hUGDH_{A136M}$ in a standard reaction buffer of 50 mM HEPES (pH 7.5), 50 mM NaCl, and 5 mM EDTA with a saturating concentration of either 2.5 mM NAD^+ or 1 mM UDP-glucose. Prior to the reaction, the substrate and enzyme were separately preincubated at 25 °C for 5 min, and the reaction was initiated by rapidly mixing the enzyme and substrate solutions. Progress curves were measured by monitoring NADH production at 340 nm (molar absorptivity coefficient of 6220 M^{-1} cm^{-1}) with 0.5 s data points using an Agilent 8453 UV/vis spectrometer at 25 °C. All data were fit using nonlinear regression from PRISM. hUDGH steady state velocities were calculated by fitting progress curves to eq 1, as we have previously described.[39, 44, 45, 48, 49] For a hysteretic enzyme, the initial velocity does not satisfy the steady state approximation. Thus, the v_{ss} prior to the depletion of 10% of the substrate represents the initial steady state velocity. $hUGDH_{A136M}$ did not display

hysteresis; thus, the initial velocities satisfy the steady state approximation and were determined using the linear portion of the progress curves. Initial steady state velocities were fit to a sigmoidal rate equation and analyzed with residual analysis: [95]

$$v = \frac{k_{cat}[E_t][S]^h}{K_M{}^h + [S]^h} \qquad \text{(equation 2)}$$

where h is the Hill coefficient. Negative cooperativity in the NAD^+ saturation curve of hUGDH was evaluated using Kurganov's analysis of a concave-up Eadie–Hofstee plot.[23] Briefly, data were fit to the equation:

$$\frac{v}{[S]} = \frac{V_{max} - v}{K_m^{eff}}, \text{ where } K_m^{eff} = K_0 + (K_{lim} - K_0)(v/V_{max}) \qquad \text{(equation 3)}$$

where K_0 is the estimated K_M for the high-affinity binding sites and K_{lim} is the average of the lower-affinity binding sites.

We determined the K_i for the inhibitor UDP-Xyl as previously described.[48] Briefly, UDP-Xyl competes with UDP-Glc for the active site and induces hUGDH to form the E^{Ω} state.[39, 48] Because the E and E^{Ω} states have different affinities for UDP-Glc and UDP-Xyl, the inhibition studies display cooperativity. To determine the K_i, the substrate saturation curves from different concentrations of UDP-Xyl were fit simultaneously to the equation for competitive inhibition corrected for substrate cooperativity using global analysis in PRISM:

$$v_o = \frac{k_{cat}[E_t][S]^h}{(K_M^{app})^h + [S]^h}, \text{ where } K_M^{app} = K_M\left(1 + \frac{[I]}{K_i}\right) \qquad \text{(equation 4)}$$

2.4 Results

Design and Rationale for the A136M Substitution

The A136M substitution was designed to prevent the isomerization of the T131-loop/α6 allosteric switch. In the active conformation of hUGDH, the Cβ atom of A136 is solvent-exposed (Figure 2.2.A). To form the inhibited E^{Ω} state, the α6 helix tilts ~10.6° and then rotates about the

helix axis ~12°, burying A136 near a cavity in the core of the protein (Figure 2.2.B).[39] We hypothesized that an amino acid substitution could prevent the burial of A136 and effectively lock the α6 helix in the active conformation. We excluded polar and charged residues from consideration because A136 is located in the hexamer-building interface, and the burial of an unsatisfied electrostatic group might destabilize the hexamer. We also excluded β-branched amino acids based on a modeling experiment with valine at position 136; each preferred rotamer of valine introduced at least one steric clash between a γ-methyl and the surrounding residues of the A136 pocket (Figure 2.2.C). Similarly, leucine, phenylalanine, and tryptophan are also expected to introduce steric clashes. The structure of the A136 surface pocket permits only an unbranched amino acid substitution with a χ1 torsion angle of 180° (Figure 2.2.C). On the basis of these observations, we chose the A136M substitution; methionine not only satisfies the criteria mentioned above but also is flexible enough to conform to the local packing constraints in the A136 pocket.

Structural Analysis of hUGDH$_{A136M}$

The crystal structure of hUDGH$_{A136M}$ was solved in space group P1 and refined to a resolution of 2.05 Å (Table 2.1). The unit cell contains a single 32 symmetry hexamer (Figure 2.3.A). The residues in loop 385–387 and C-terminal residues 467–494 are disordered in each monomer and are not included in the final model. All 466 Cα atoms in each of the six monomers superimpose with root-mean-square deviations ranging from 0.344 to 0.834 Å. The largest structural differences between the chains involve variability in the amount of hinge-bending motion of the NB (residues 1–212) and SB (residues 323–466) domains about an axis located between residues 219 and 220 (Figure 2.3.B). A DynDom[96] comparison of hUDGH$_{A136M}$ to the closed conformation of hUGDH (represented by PDB entry 2Q3E) shows that the NB and SB

domains of chains A–F are rotated open by 11.2°, 9.4°, 9.1°, 10°, 12.4°, and 5.9°, respectively. In the SB domain of chain C, we observed electron density consistent with a weakly ordered nucleotide sugar that has copurified with the enzyme (Figure 2.3.C). We have modeled the density as a UDP-Glc based on the observation that it binds in the same position and orientation as the substrate in the abortive ternary complex of UGDH (PDB entry 2Q3E). The occupancy of the UDP-Glc was adjusted to 50% to match the *B* factors of the interacting residues in the active site (see Methods). In the other five chains of $hUDGH_{A136M}$, the corresponding density is weaker and supports only our modeling of the diphosphate in UDP-Glc (not shown). The quality of the nucleotide sugar electron density appears to correlate with the average *B* factors of the chains; chain C has the lowest average *B* factor (44 Å2) and the best ordered UDP-Glc, while the average B factors for the remaining chains range from 51 to 60 Å2 with much weaker density for the nucleotide sugar. This is not the first observation of a nucleotide sugar copurifying with UGDH. Dickinson reported that bovine UGDH purified from liver contained a nucleotide sugar that he tentatively identified as UDP-Glc or UDP-Xyl.[50] We chose to model UDP-Glc based on the fact that *E. coli* does not produce UDP-Xyl.

The A136M substitution is well-ordered and fills the A136 surface pocket in the hexamer-building interface without making any van der Waals contacts with the adjacent monomer in the hexamer (Figure 2.4.A-B). As expected, M136 adopts the trans rotamer with the following angles: $\chi1 = 179.3°$, $\chi2 = 73.4°$, and $\chi3 = -111.0°$. The A136M substitution does not interfere with the lattice contacts in either of the two UDP-Xyl-bound E^{Ω} crystal forms we have previously described.[17] Still, our attempts to obtain X-ray diffraction quality crystals of $hUDGH_{A136M}$ using both previously published conditions failed. The fact that $hUDGH_{A136M}$ will not crystallize in the E^{Ω} state suggests that the substitution prevents the allosteric transition.

As in previous studies,[44] we analyzed the conformation of the allosteric switch by superimposing the Cα atoms of residues 1–124 in $hUDGH_{A136M}$ onto the active E, E*, and UDP-Xyl-inhibited E^{Ω} structures of hUGDH [PDB entries 2Q3E, 4QEJ, and 3PTZ, respectively (Figure 24.C)]. The T131-loops in all six monomers of $hUDGH_{A136M}$ are in a conformation between the active E and E* states, with the T131 Cα atom of $hUGDH_{A136M}$ being displaced 0.5–0.7 and 0.4–0.6 Å from the corresponding atoms in the E and E* structures, respectively. In contrast, the α6 helix and the hexamer building interface of $hUGDH_{A136M}$ are most similar to the E conformation of hUGDH (PDB entry 2Q3E) (Figure 2.4.D). Specifically, F323 in the hexamer-building interface adopts the same rotamer observed in the active state, with the Cα atoms of F323 in all chains of $hUDGH_{A136M}$ positioned 0.4–0.8 Å from the E state compared to being 1.5–1.9 Å from the E* state.

The A136M Substitution Stabilizes the Hexamer

We have previously shown that the ligand-free hUGDH hexamer is relatively unstable in solution and dissociates into a concentration-dependent distribution of dimers, tetramers, and hexamers in rapid equilibrium.[45] Because of its location in the hexamer-building interface, it is possible that the A136M substitution could alter the stability of the hexamer. We used sedimentation velocity studies to analyze the oligomeric structure of $hUGDH_{A136M}$ in solution. The *c(s)* distribution of 9 μM $hUGDH_{A136M}$ at pH 7.5 is dominated (94.1%) by an 11.5 S species corresponding to the hexamer, and a smaller (5.9%) amount of the 5.2 S dimer (Figure 2.5.A). For comparison, the c(s) distribution of 9 μM hUGDH reveals an 11.4 S hexamer (78.2%), an 8.3 S tetramer (4.2%), and a 5.5 S dimer (13.1%) (Figure 2.5.B). The distribution also contains a 3.2 S species that is most likely a small amount (4.5%) of misfolded monomer. The slight differences in S values for corresponding species in the hUDGH and $hUGDH_{A136M}$ distributions are not

significant as S values obtained from a system in rapid equilibrium are inherently biased by the mean of the *c(s)* distribution.[94, 97]

Next, we examined the oligomeric state of $hUGDH_{A136M}$ at the more alkaline pH of 8.5, which we have previously shown to weaken the hUGDH hexamer.[45] Again, the main sedimenting species in the $hUGDH_{A136M}$ *c(s)* distribution is an 11.85 S hexamer (83.7%), followed by a 5.49 S dimer (13.4%) and a 7.9 S tetramer (3.0%) (Figure 2.5.C). In contrast, the *c(s)* distribution of hUGDH reveals significantly less 10.6 S hexamer (64.4%) and a complementary increase in the 6.5 S dimer concentration (34.0%) (Figure 2.5.D). The broadening of the hexamer and dimer peaks in hUGDH is a characteristic of the rapid equilibrium between species.[45, 48, 98] This likely explains why we do not observe a quantifiable peak for the tetramer, which represents a transient in the formation and dissociation of the hexamer.[48] These results show that the A136M substitution favors the hexameric state of hUGDH.

Hysteresis Is Not Observed in $hUGDH_{A136M}$

Hysteresis in hUGDH and $hUDGH_{A136M}$ was analyzed using stopped-flow absorbance spectroscopy. hUGDH was injected into a rapid mixing cell with saturating substrate and NAD^+, and the characteristic lag in enzyme activity was measured by fitting the data to eq 1 as described in Methods (Table 2.2). The relaxation of hUGDH has a τ of 16.30 ± 0.24 s, which corresponds to a lag of 44.3 ± 0.66 s before reaching the steady state velocity of 1.0 ± 0.01 nM NADH s^{-1} (Figure 2.6.A). In contrast, we observed no lag for $hUDGH_{A136M}$ under these conditions (Figure 2.6.B). In fact, $hUGDH_{A136M}$ does not display hysteresis even at saturating concentrations of the feedback inhibitor UDP-Xyl (Table 2.2). This is not true of hUGDH, where the lag has been shown[48] to increase with UDP-Xyl concentration (Table 2.2). Finally, the steady state velocity of

hUDGH$_{A136M}$ appears to be 30% faster than that of hUGDH, but this is most likely due to negative cooperativity in the latter enzyme (discussed below) (Table 2.2).

The A136M Substitution Disrupts NAD$^+$-Induced Negative Cooperativity

The NAD$^+$ saturation curve of hUGDH displays negative cooperativity (Hill coefficient of 0.73 ± 0.07), which we have previously shown[49] to be an intrinsic property of the substrate-bound complex (in the absence of UDP-Glc, the binding of NAD$^+$ is noncooperative) (Figure 2.6.C and Table 2.3). In contrast, the NAD+ saturation curve of hUGDH$_{A136M}$ is hyperbolic (Figure 2.6.D and Table 2.3). Negative cooperativity indicates an asymmetry in an enzyme that results in a mixture of high-affinity and low-affinity binding sites. It is possible to estimate the K_M for high-affinity and low-affinity NAD$^+$ binding (K_0 and K_{lim}, respectively) using Kurgonov's methodology[23] (see Methods) (Figure 2.6.E). This analysis of hUGDH NAD$^+$ saturation curves yields a K_0 of 88 ± 20 μM, which is similar to the K_M (90 ± 6 μM) observed for hUGDH$_{A136M}$ (Figure 2.6.D and Table 2.3).

With respect to turnover, the k_{cat} values for the NAD$^+$ saturation curves of both hUGDH and hUGDH$_{A136M}$ are comparable at 0.76 ± 0.05 and 0.83 ± 0.02 s^{-1}, respectively (Table 2.3). The UDP-Glc saturation curves for hUGDH and hUGDH$_{A136M}$ are both hyperbolic, with similar K_M values of 9.7 ± 0.8 and 7.26 ± 0.67 μM, respectively (Table 2.3) However, there is a significant discrepancy in the k_{cat} values observed in the UDP-Glc saturation curves; while the hUGDH$_{A136M}$ k_{cat} (0.76 ± 0.02 s^{-1}) is similar to that measured in the NAD$^+$ saturation curves, the k_{cat} for hUGDH is significantly lower (0.55 ± 0.01 s^{-1}) (Table 2.3). The difference in turnover numbers is due to the coenzyme-dependent negative cooperativity of hUGDH, which makes it difficult to saturate the enzyme with NAD$^+$ (K_{lim} = 1800 μM) to achieve the pseudo-first-order conditions necessary for steady state analysis of UDP-Glc kinetics.

hUDGH$_{A136M}$ Binding of UDP-Xylose Is Noncooperative

Next we examined the impact of the A136M substitution on the allosteric inhibition mechanism of hUGDH. UDP-Xyl binding stabilizes the E^{Ω} conformation of hUGDH (Figure 2.1).[39, 44, 48] Competition with UDP-Glc induces the low-substrate-affinity E^{Ω} conformation of hUGDH to undergo a cooperative conformational change to the high-affinity E state.[48, 99] The allosteric transition can be observed as sigmoidicity in the UDP-Glc saturation curves during steady state analysis of UDP-Xyl inhibition (Figure 2.6.F). Global analysis of the hUGDH steady state inhibition curves yields a K_i of 0.32 ± 0.05 μM for UDP-Xyl, which is consistent with our earlier work.[48] The UDP-Glc saturation curve in the presence of UDP-Xyl at a concentration that is ~10 times the K_i is strongly cooperative, with a Hill coefficient of 2.1 ± 0.2 (Figure 2.6.F and Table B4). This agrees with our previous work, which showed that sigmoidicity increases with UDP-Xyl concentration.[48] In contrast, hUGDH$_{A136M}$ has a much lower affinity for UDP-Xyl (Ki of 4.2 ± 0.4 μM) and displays no cooperativity (Figure 2.6.G and Table 2.4). In fact, the UDP-Glc saturation curves are hyperbolic even at inhibitor concentrations 14 times the K_i (60 μM UDP-Xyl), which is strong evidence that hUGDH$_{A136M}$ does not undergo a conformational change between inhibited and uninhibited states (Figure 2.6.F).

2.5 Discussion

We have previously shown that binding of the feedback inhibitor UDP-Xyl to hUGDH induces the T131-loop/α6 allosteric switch to repack and form the inactive E^{Ω} state (Figure 2.1).[39] This remarkable allosteric transition requires the α6 helix to rotate, which buries A136 in the protein core (Figure 2.2.A-B). Here we have tested the allosteric mechanism by trapping the α6 helix in the active conformation using the A136M substitution (Figure 2.4). The burial of the methionine would be unlikely, as it would require a significant local unfolding of the protein and

repacking of the core to accommodate the bulky side chain (Figure 2.2). This rationale is supported by the substrate saturation kinetics of the UDP-Xyl-inhibited enzyme (Table 2.4). The E^{Ω} state has a lower affinity for UDP-Glc, which results in positive cooperativity in substrate saturation curves of hUGDH (Figure 2.6.F).[48] The fact that UDP-Xyl-saturated $hUGDH_{A136M}$ does not display substrate cooperativity is strong evidence that the A136M substitution prevents the formation of the E^{Ω} state (Figure 2.6.F and Table 2.4). This interpretation implies that the lower affinity for UDP-Xyl (K_i = 4.2 μM) observed in $hUGDH_{A136M}$ corresponds to the competitive inhibition constant for the E state only (Figure 2.6.G and Table 2.4). This is the first report of the E state affinity constant for UDP-Xyl and suggests the higher affinity (K_i = 0.32 μM) observed for hUGDH is a property of the E^{Ω} state (Table 2.4).

The specific activity of hUGDH shows a hyperbolic dependency on protein concentration that can be modeled as three low-activity dimers associating to form a higher-activity hexamer.[48] It is possible for enzyme association to result in hysteresis[100], but this is not the case for hUGDH. Instead, hysteresis in hUGDH is caused by the slow isomerization from the inactive E* state to the active E state upon binding of NAD^+ and substrate.[45, 49] While the increased stability of the $hUGDH_{A136M}$ hexamer does not explain the absence of a lag in progress curves, it does focus attention on the conformational flexibility of the allosteric switch in the E* state (Figures 2.5 and 2.6.A). The intermediate conformation T131-loop/α6 allosteric switch and the hexamer-building interface in E* led us to hypothesize that the hysteresis was the result of NAD^+ and substrate inducing the allosteric switch and surrounding residues to slowly repack into the E conformation (Figure 2.1.B-C).[45, 48] The A136M substitution allows us to test this model. The crystal structure of unliganded $hUGDH_{A136M}$ shows that both the allosteric switch and the hexamer-building interface favor the E state, despite the open domain conformation of the enzyme (Figure 2.4.C-D).

By stabilizing the allosteric switch in the E state, the A136M substitution would prevent repacking of the core and interface. This model accurately predicts the absence of a lag in the $hUGDH_{A136M}$ progress curves (Figure 2.6.B). Additional evidence comes from UDP-Xyl inhibition studies. We have previously shown[48] that the lag in hUGDH increases with UDP-Xyl concentration (Table 2.2). This increase is most likely caused by the inhibitor stabilizing the E^{Ω} conformation, which must first convert to E^{*Bound} before it can slowly isomerize into the active E state (Figure 2.1.D). Thus, the absence of hysteresis in UDP-Xyl-saturated $hUGDH_{A136M}$ is strong evidence that the A136M substitution prevents the formation of both E* and E^{Ω} states (Table 2.2). Together, these results identify the isomerization of the allosteric switch as the source of the hysteresis observed in hUGDH.

Negative cooperativity can arise from enzyme heterogeneity that results in a mixed population with different affinity constants, but this model is not supported by ligand binding studies of hUGDH.[49] Briefly, the binding of UDP-Glc or NAD^+ to hUGDH is noncooperative, indicating that the enzyme is homogeneous, with single affinity constants for substrate and coenzyme.[49] However, the binding of NAD^+ to the binary hUGDH–UDP-Glc complex displays negativity cooperativity.[49] The simplest interpretation of these results is that the binding of UDP-Glc induces an asymmetry in the enzyme that produces two distinct affinities for the coenzyme. The molecular basis of negative cooperativity in hUGDH is not known, but the fact that the T131-loop contributes to both substrate and cofactor binding[39] suggests that it may play a role. This assumption is supported by the observation that $hUGDH_{A136M}$ constrains the T131-loop and does not display negative cooperativity (Figure 2.6.C-D and Table 2.3). A comparison of hUGDH and $hUGDH_{A136M}$ suggests that the A136M substitution selects for the high-affinity conformation of the NAD^+ binding sites. Using Kurgonov's approach,[23] we estimated the high- and low-affinity

K_M values for negative cooperativity in binding of the coenzyme to hUGDH (Figure 2.6.E and Table 2.3). The apparent high affinity K_M (K_0) for hUGDH is very similar to the noncooperative K_M we observe for $hUGDH_{A136M}$ (Table 2.3). This suggests that the A136M substitution prevents the asymmetry induced by UDP-Glc binding. This suggests that the flexibility of the allosteric switch is linked to the negative cooperativity observed in NAD^+ substrate saturation curves.

This work shows that targeting the allosteric switch is a practical method of selecting specific conformations of hUGDH and, by extension, regulating enzyme activity. Regulating hUGDH is an important milestone in our long-term goal of controlling the pharmacokinetics of chemotherapeutics. Future work will focus on inhibiting hUGDH by trapping the allosteric switch in the E^{Ω} conformation.

2.6 Acknowledgements

We thank Dr. Renuka Kadirvelraj for helpful discussions and assistance with designing the A136M substitution. We also thank the SER-CAT beamline personnel (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) for assistance with X-ray data collection.

2.7 Figures and Tables

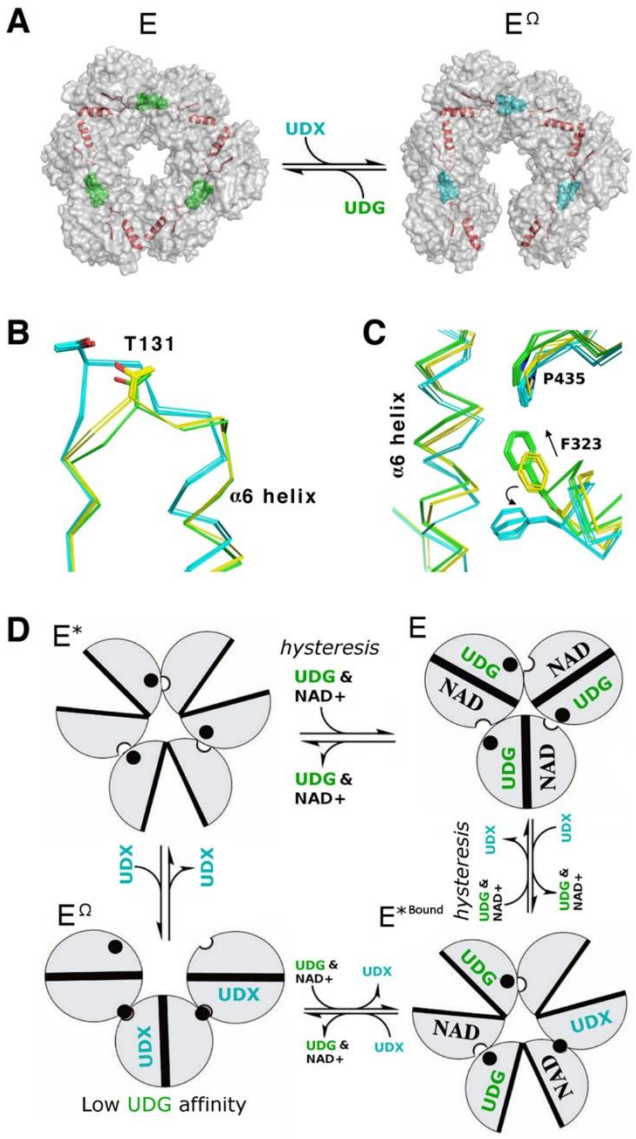


Figure 2.1: hUDGH Undergoes Conformational Changes During Allostery and Hysteresis. **(A)** Semitransparent surface rendering showing the buried T131-loop/α6 helix (red) in the E and E^{Ω} states. UDP-Glc (green spheres, UDG) binds to the active site and favors the 32 symmetry E state. UDP-Xyl (teal spheres, UDX) competes for the active site and triggers the allosteric switch to rotate the α6 helix and favor the E^{Ω} state. **(B)** Position of the T131-loop/α6 helix in the E (green), E* (yellow), and E^{Ω} (cyan) states. All of the unique chains in the crystal structures of E, E*, and E^{Ω} are superimposed (Protein Data Bank entries 2Q3E, 4RJT, and 3PTZ, respectively). **(C)** Hexamer-building interfaces of the E, E*, and E^{Ω} states depicted as in panel B. **(D)** Relationship between domain rotations (clamshell circles) and the α6 helix (black filled circle) conformation in the E (active), E^{Ω} (inhibited), E* (inactive), and E*Bound (inactive, inhibitor- and substrate-bound) states. In the absence of ligand, the inactive E* state is favored. The binding of the substrate and NAD^+ to E* induces a slow isomerization (hysteresis) to produce the active E state.[49] The

E*Bound state forms as UDG and UDX compete to induce the transition between the E and E^{Ω} states (the ligand stoichiometry that induces the transition is unknown). UDP-Xyl binding shifts the α6 helix to strengthen the packing interactions between adjacent subunits (semicircle cutouts) and form the E^{Ω} hexamer-building interface.[45, 48] The E^{Ω} conformation of the allosteric switch reduces the affinity for the substrate. For the sake of simplicity, only one trimer of the hUGDH hexamer is depicted.

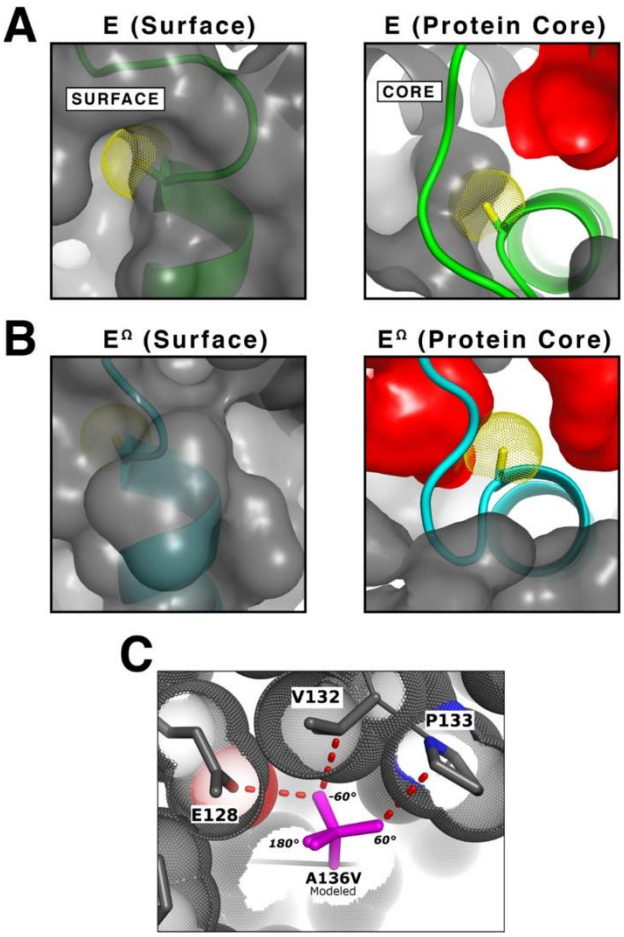


Figure 2.2: Rationale for the Design of the A136M Substitution. (A) The left panel is a semitransparent surface rendering of the E state (PDB entry 2Q3E) showing the α6 helix (green) with the Cβ atom of A136 (yellow van der Waals dots) in a solvent accessible pocket. The right panel depicts a "cutaway" of the protein surface looking down the axis of the α6 helix to show conformational changes that take place in the protein core, specifically, the position of the α6 helix and the solvent-exposed A136 in the E state. Cavities buried in the protein core (red surfaces) were identified with a 1.2 Å radius probe. The contribution of the A136 Cβ atom to the surface was removed prior to rendering. **(B)** The E^{Ω} conformation (cyan) of hUGDH rotates the α6 helix to bury A136 in the protein core (PDB entry 3PTZ). The right panel depicts the rotation of the α6 helix to bury A136 in the E^{Ω} state. The left and right panels are illustrated as in panel A. Note that the cavities in the core change shape and number because of the repacking of the allosteric switch. **(C)** To investigate the steric constraints (dotted van der Waals surfaces) of a β-branched amino acid substitution in the A136 pocket, we modeled A136V with its three preferred rotamers superimposed (purple sticks). Each rotamer introduces a bad contact (red dashed lines) between a γ-methyl and adjacent amino acids. Thus, the pocket will accommodate only an unbranched amino acid with a 180° χ1 torsion angle (60°, 180°, and −60° χ1 torsion angles are labeled).

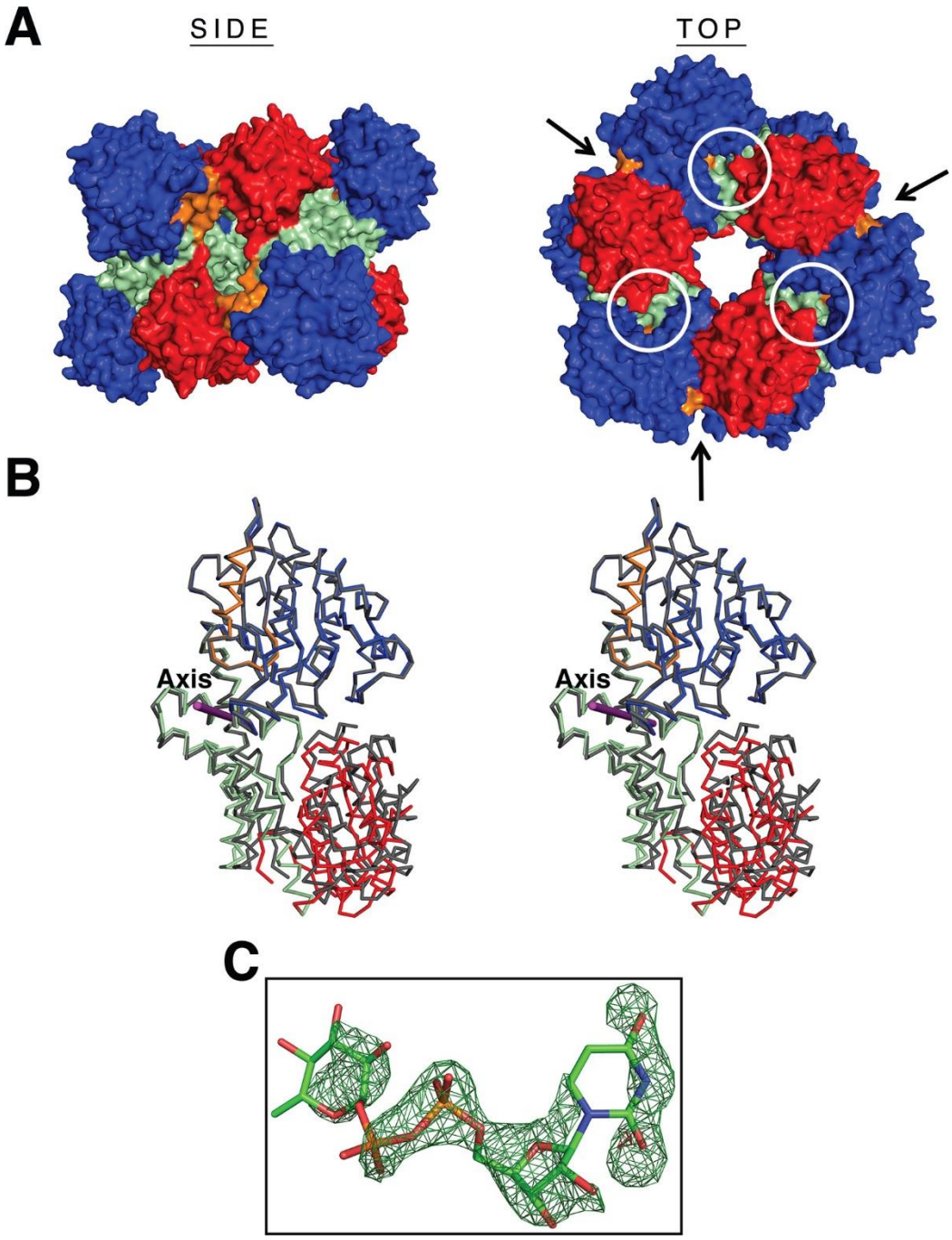


Figure 2.3: Crystal Structure of hUGDH$_{A136M}$ in the Open Domain Conformation. (A) Side and top views of the UDGH$_{A136M}$ 32 symmetry hexamer. The surfaces of the six monomers are colored by domain: NB (blue), dimerization (light green), and SB (red). The allosteric switch is colored orange. The active sites and hexamer-building interfaces are identified with circles and arrows, respectively. **(B)** Stereoview of a Cα trace of hUDGH$_{A136M}$ chain C (colored as in panel A) superimposed onto hUGDH in the closed conformation (gray, PDB entry 2Q3E) to illustrate the open domain conformation. The hinge-bending axis is depicted as a purple rod. **(C)** Difference (F_o – F_c) electron density map contoured at 3σ for the partially occupied UDP-Glc (sticks) calculated after the nucleotide sugar was omitted and the resulting model subjected to simulated annealing.

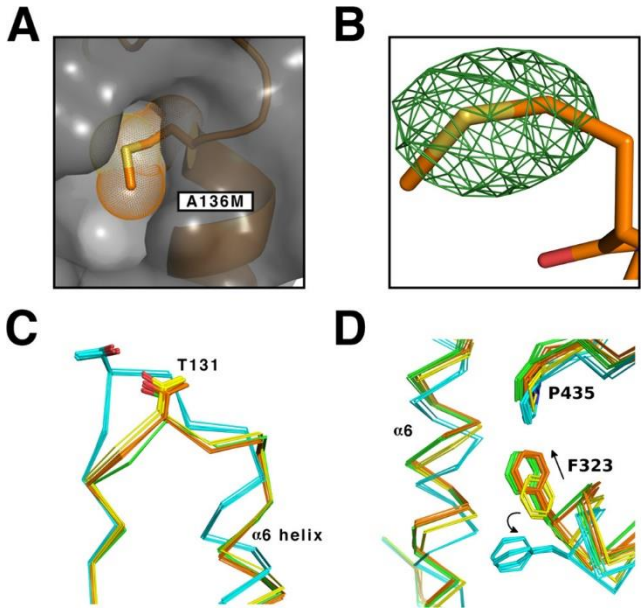


Figure 2.4: A136M Substitution that Stabilizes the E State of hUGDH. (A) The difference (F_o – F_c) electron density map contoured at 4σ for A136 calculated prior to modeling the methionine substitution. **(B)** Surface rendering (gray) of $hUDGH_{A136M}$ showing the A136M substitution (orange stick with van der Waals surface dots) and the α6 helix (orange). As expected, the A136M substitution fills the A136 pocket. **(C)** Position of the T131 loop in the inactive (cyan), active (green), E* (yellow), and $hUDGH_{A136M}$ (orange) states. **(D)** Hexamer-building interfaces of the inactive (cyan), active (green), E* (yellow), and $hUDGH_{A136M}$ (orange) states. Key residues are labeled with arrows used to depict the direction of motion between states.

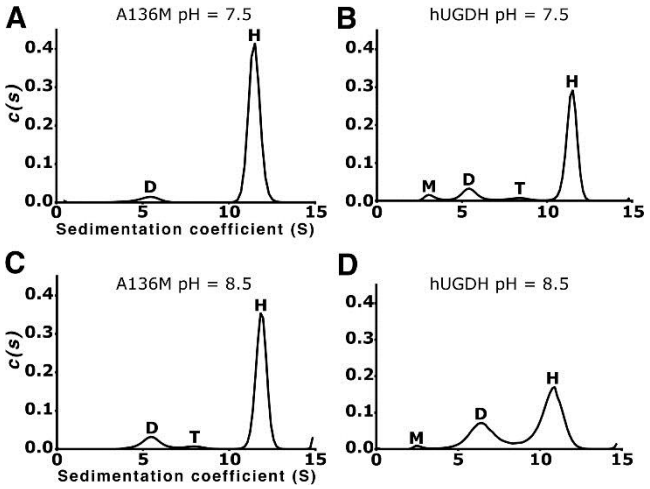


Figure 2.5: A136M Substitution that Stabilizes the Hexamer. (A) Sedimentation velocity c(s) distribution of hUDGH$_{A136M}$ at pH 7.5 showing the hexamer (H) and dimer (D). **(B)** hUDGH c(s) distribution at pH 7.5 showing the hexamer (H), tetramer (T), dimer (D), and a peak that we interpret as misfolded monomer (M). **(C)** hUDGH$_{A136M}$ at pH 8.5. **(D)** hUDGH at pH 8.5.

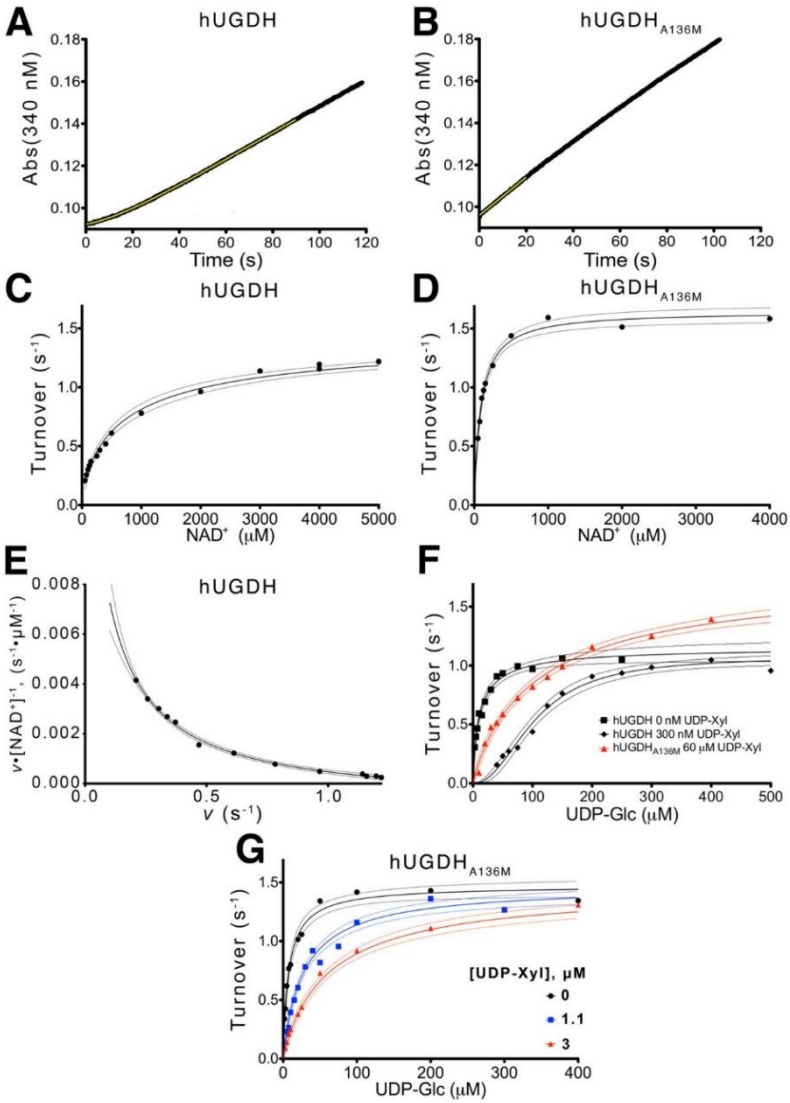


Figure 2.6: A136M Substitution that Disrupts Hysteresis and Cooperativity. (A) Stopped-flow analysis of hysteresis in the hUDGH progress curve under saturating concentrations of NAD^+ (5 mM) and UDP-Glc (0.5 mM). The data points (black line) are fit (yellow line) to eq 1 as described in Methods. **(B)** The $hUDGH_{A136M}$ progress curve was analyzed as described for panel A, shows no lag, and is fit to a linear equation (yellow line) to obtain the initial steady state velocity. The yellow line represents a linear fit for the first 20 s of the reaction. **(C and D)** NAD^+ saturation curves for hUGDH and $hUGDH_{A136M}$, respectively, fit to eq 3. The rates are normalized to turnover (nanomolar NADH per nanomolar enzyme per second). Dashed lines represent the 95% confidence interval of the fit in all panels. **(E)** An Eadie–Hofstee plot of the hUGDH NAD+ saturation data in panel C is concave-up (negative cooperativity) and was fit to eq 3 to determine the K_0 and K_{lim} values in Table 3. Velocity (v) was normalized to turnover (inverse seconds) as in panels C and D. **(F)** hUGDH UDP-glucose saturation curves (black lines) with 0 μM (■) and 3 μM (◆) UDP-xylose were globally fit to eq 4 to determine the K_i and the Hill coefficient. For comparison, data from a $hUGDH_{A136M}$ UDP-glucose saturation curve with 60 μM UDP-Xyl (red triangles) are plotted. **(G)** $hUGDH_{A136M}$ UDP-glucose saturation curves with 0 μM (black circles), 11 μM (blue squares), and 30 μM (red triangles) UDP-xylose were globally fit to eq 4 to determine K_i.

Table 2.1.: Data Collection and Refinement Statistics for hUDGH$_{A136M}$

Data Collection	
Protein Data Bank entry	5TJH
space group	*P*1
unit cell dimensions	
a, *b*, *c* (Å)	92.35, 104.59, 107.71
α, β, γ (deg)	64.82, 68.35, 73.72
completeness (%)	96.6 (97.1)^{a} [91.3]^{b}
redundancy	3.9 (4.0) [3.8]
no. of reflections	987044
$I/\sigma(I)$	7.82 (1.80) [1.03]
$CC_{1/2}$^{c}	99.6 (70.3) [45.6]
R_{meas}^{d} (%)	11.7 (92.8) [154.1]
Refinement	
resolution (Å)	2.05
R_{work}/R_{free}	0.171/0.203
R_{free} in the highest-resolution shell	0.345
no. of atoms	
protein	21736
ligand	89
water	799

B factor (Å^2)	
protein	53.5
ligand	58.6
water	47.7
Stereochemical Ideality	
bond lengths (Å^2)	0.008
bond angles (deg)	0.89
φ, ψ preferred region (%)	98.3
φ, ψ additionally allowed region (%)	1.7
φ, ψ disallowed region (%)	0.0

a Values in parentheses are for the highest-resolution shell (2.16–2.22 Å) based on an I/σ cutoff of 1.8.
b Values in brackets are for the highest-resolution shell (2.05–2.10 Å) based on the $CC_{1/2}$ cutoff defined by Diederichs and Karplus.[101]
c CC1/2 is the percentage of correlation between intensities from random half-data sets.[102]
d R_{meas} is the redundancy-independent merging *R* factor.[101]

Table 2.2: Effect of UDP-Xyl on Hysteresis

enzyme	[UDP-Xyl] (μM)	τ (s^{-1})	lag (s)	v_i	v_{ss}
hUGDH	0	16.3 ± 0.24	44.3 ± 0.66	0.43 ± 0.02[a]	1.0 ± 0.01[a]
	3[b]	38.3 ± 2.7	104 ± 7.3	0.35 ± 0.03	1.1 ± 0.01
$hUGDH_{A136M}$	0	ND[c]	ND[c]	ND[c]	1.3 ± 0.03
	60[b]	ND[c]	ND[c]	ND[c]	1.4 ± 0.03

[a] Units of nanomolar NADH per nanomolar enzyme per second.
[b] These concentrations of UDP-Xyl are ≥10 times the K_i listed in Table B4.
[c] Not detected.

Table 2.3: Steady State Parameters

enzyme	ligand	K_M (μM)	Hill	k_{cat}[a] (s^{-1})
hUGDH	UDP-Glc	9.7 ± 0.8	1	0.55 ± 0.01
	NAD^+	780 ± 180	0.73 ± 0.07	0.76 ± 0.05
$hUGDH_{A136M}$	UDP-Glc	7.26 ± 0.67	1	0.73 ± 0.02
	NAD^+	90 ± 6	1	0.83 ± 0.02

[a] One complete catalytic turnover produces two molecules of NADH.

Table 2.4: Global Analysis of Competitive Inhibition

enzyme	[UDP-Xyl] (μM)	UDP-Glc K_M (μM)	k_{cat}[a] (s^{-1})	UDP-Xyl K_i (μM)	Hill
hUGDH	0	9.5 ± 1.0[b]	0.55 ± 0.02[b]	0.30 ± 0.03	1
	3				2.1 ± 0.2
$hUGDH_{A136M}$	0	7.34 ± 0.80[b]	0.74 ± 0.02[b]	4.2 ± 0.4	1
	11				1
	30				1
$hUGDH_{A136M}$[c]	60	6.3 ± 0.5	0.85 ± 0.02		1

[a] Turnover produces two molecules of NADH per cycle.
[b] For global analysis, this parameter was refined as a shared value for all UDP-Xyl concentrations.
[c] The 60 μM UDP-Xyl experiment was conducted separately and excluded from the global fit.

CHAPTER 3

HYSTERESIS AND ALLOSTERY IN HUMAN UDP-GLUCOSE DEHYDROGENASE REQUIRE A FLEXIBLE PROTEIN CORE

Reprinted here with permission from: **Beattie, N.R.**, Pioso, B.J., Sidlo, A.M., Keul, N.D**.** and Wood, Z.A., Hysteresis and Allostery in Human UDP-Glucose Dehydrogenase Require a Flexible Protein Core. *Biochemistry*, 57(50), pp.6848-6859 (2018).

3.1 Abstract

Human UDP-glucose dehydrogenase (hUGDH) oxidizes UDP-glucose to UDP-glucuronic acid, an essential substrate in the phase II metabolism of drugs. The activity of hUGDH is regulated by the conformation of a buried allosteric switch (T131 loop/α6 helix). Substrate binding induces the allosteric switch to slowly isomerize from an inactive E* conformation to the active E state, which can be observed as enzyme hysteresis. When the feedback inhibitor UDP-xylose binds, the allosteric switch and surrounding residues in the protein core repack, converting the hexamer into an inactive, horseshoe-shaped complex (E^{Ω}). This allosteric transition is facilitated by large cavities and declivities in the protein core that provide the space required to accommodate the alternate packing arrangements. Here, we have used the A104L substitution to fill a cavity in the E state and sterically prevent repacking of the core into the E^{Ω} state. Steady state analysis shows that $hUGDH_{A104L}$ binds UDP-xylose with lower affinity and that the inhibition is no longer cooperative. This means that the allosteric transition to the high-UDP-xylose affinity E^{Ω} state is blocked by the substitution. The crystal structures of $hUGDH_{A104L}$ show that the allosteric switch still adopts the E and E* states, albeit with a more rigid protein core. However, the progress curves of $hUGDH_{A104L}$ do not show hysteresis, which suggests that the E* and E states are now in rapid equilibrium. Our data suggest that hysteresis in native hUGDH originates from the conformational entropy of the E* state protein core.

3.2 Introduction

UDP-glucuronic acid (UDP-GlcA) is the essential substrate in glucuronidation, a major component of phase II metabolism in mammalian cells.[80, 81, 103] Glucuronosyltransferases catalyze the nucleophilic addition of glucuronic acid to drugs or toxins, which are then rapidly excreted from the body.[12] Glucuronidation is exploited by some cancers as a mechanism of chemotherapeutic resistance.[6-9] Thus, controlling glucuronidation is a potential strategy for combating drug resistance in these cancers. Because there are 28 glucuronosyltransferase isozymes, each with broad substrate specificity, designing specific inhibitors is challenging.[12] Alternatively, it is possible to inhibit glucuronidation by reducing the availability of the essential substrate UDP-GlcA.[82-84] UDP-GlcA is produced by UDP-glucose dehydrogenase (UGDH), which uses two molecules of NAD^+ to oxidize the C6 hydroxyl of UDP-glucose (UDP-Glc).[13] The downstream metabolite UDP-xylose (UDP-Xyl) acts as an allosteric feedback inhibitor to control the activity of hUGDH.[47, 104] Understanding the allosteric mechanism of UGDH is an important goal in developing strategies to control glucuronidation.

Human UGDH (hUGDH) forms a hexamer, with each chain containing an NAD^+ binding (NB), dimerization (DD), and a nucleotide–sugar binding (SB) domain (Figure 3.1.A). The activity of hUGDH is controlled by the feedback inhibitor UDP-Xyl, which converts the enzyme into an inactive horseshoe-shaped hexamer (E^{Ω}) (Figure 3.1.A).[39, 47] This conformational change is controlled by a buried allosteric switch that links the active site to the α6 helix in the hexamer-building interface (Figure 3.1.B-C).[45, 48] The allosteric switch (T131 loop/α6 helix) can adopt one of three distinct conformations: the active E state, the allosterically inhibited E^{Ω}, and an inactive E* conformation that represents a structural intermediate between the E and E^{Ω} states (Figure 3.1.C). [39, 42, 45, 48]In the absence of a ligand, the enzyme favors the E* state, which resembles the E

state with the exception that the NAD^+ binding domains are rotated into an "open" conformation that exposes the T131 loop and surrounding residues to the solvent (Figure 3.1.B).[39, 48] When UDP-Glc binds, the allosteric switch and surrounding residues in the protein core adopt the active E conformation (Figure 3.1.B-C). The activity of hUGDH is regulated by an atypical allosteric mechanism. Briefly, UDP-Xyl competes with the substrate for the active site and upon binding induces the allosteric switch to adopt the E^{Ω} conformation (Figure 3.1.A&C).[39, 48] This remarkable transition occurs because UDP-Xyl lacks the C5 hydroxymethyl of UDP-Glc, which allows the allosteric switch to move ~5 Å into the active site (Figure 3.1.B).[39] What makes this allosteric is that the binding of UDP-Xyl in the active site increases the affinity between subunits in the E^{Ω} hexamer.[44] In the inhibited state, the T131 loop occludes the coenzyme binding site and reduces the affinity for UDP-Glc (Figure 3.1.B).[39, 48] Because the E^{Ω} and E states have different affinities for UDP-Glc, the allosteric transition can be observed as positive cooperativity in steady state studies of UDP-Xyl inhibition.[39, 48]

The substrate-induced transition from the inactive E* to the active E hexamer is slow and can be observed as hysteresis in progress curves.[39, 48, 105] It has been proposed that the source of this slow enzyme isomerization is the repacking of the allosteric switch and surrounding residues into the E state.[39, 48, 105] This hypothesis is supported by a previous study that showed trapping the allosteric switch in the active conformation abolishes hysteresis.[105] The remarkable plasticity exhibited by the hUGDH protein core likely originates from large cavities and deep surface pockets (declivities) near the allosteric switch (Figure 3.1.D).[39] Here, we have examined how the flexibility of the hUGDH protein core contributes to both hysteresis and allostery. Using a small to large amino acid substitution (A104L), we have partially filled a cavity to stabilize the E state. Our results show that these packing defects are potential targets for regulating hUGDH activity.

3.3 Materials and Methods

Cavity Characterization

Caver Analysis 1.0[106] was used to characterize the packing defects observed in crystal structures of hUGDH (UniProtKB entry O60701). Protein Data Bank (PDB) entries 2Q3E, 4RJT, and 3PTZ were used as the models for the E, E*, and E^{Ω} states, respectively. Each chain of the structures was converted to an individual PDB file prior to analysis, which includes 2Q3E chains A–F, 4RJT chains A–C, and 3PTZ chains A–F. The volumes of cavities were quantified using an inner probe radius of 1.2 Å and an outer probe radius of 1.9 Å. The inner probe radius was selected to approximate the radius of water. The outer probe radius was selected to maximize the volume while preventing the cavity from opening to the surface. Declivities were defined using an inner probe radius of 1.2 Å and an outer probe radius of 2.3 Å. For the chains of 3PTZ, the outer probe radius was varied from 1.9 to 2.3 Å to minimize extraneous surface inclusion in the declivity volume calculation. The volumes were calculated after maximizing tessellation by using the "increase precision" option that comes as part of Caver Analysis 1.0. Because the size and locations of the packing defects are nearly identical in all chains, the contributing atoms of each packing defect were identified using chain "A" of each hUGDH structure.

$Rosetta_{VIP}$ Redesign

Chain A of hUGDH in the E state (PDB entry 2Q3E) was used as the template for RosettaVIP[74] core redesign. Because the initial relaxation step in the RosettaVIP protocol introduced small shifts that decreased the size of the packing defects observed in the native crystal structures, we did not include it in the redesign effort. Packing defects were detected using a probe with a radius of 1.2 Å. See results for residues targeted in the redesign. The protein structure was subject to iterative rounds of substitution, fixed backbone relaxation, and then scoring.

Protein Expression, Crystallization, and Structure Solution

Both hUGDH and $hUGDH_{A104L}$ were recombinantly expressed in Escherichia coli and purified as previously described.[39, 44, 45, 48, 49, 105] Following purification, TEV protease was used to remove the His tag, and another round of purification was conducted to remove the TEV. Proteins were quantified using an Agilent 8453 UV/vis instrument and an ε280 of 49850 M^{-1} cm^{-1} for hUGDH and $hUGDH_{A104L}$ (calculated using Protparam[87]). The proteins were dialyzed into a storage buffer [25 mM Tris (pH 8.0) and 50 mM NaCl] and concentrated to ∼20 mg/mL using a Millipore Amicon Ultra-15 10K centrifugal filter unit. Aliquots of protein were flash-frozen in liquid nitrogen and stored at −80 °C. All proteins were crystallized at 20 °C using the hanging drop vapor diffusion method, with a 2 μL drop consisting of a 1:1 mixture of protein (10 mg/mL) and the reservoir solution. The E complex of $hUGDH_{A104L}$ ($hUGDH_{A104L}$:UDP-Glc:NADH) was formed with 3 mM UDP-Glc and 5 mM NADH in the protein solution over a reservoir of 16% polyethylene glycol 3350, 0.2 M ammonium sulfate, 0.3 M NaCl, 1% 1,2-propanediol, and 0.1 M Tris (pH 8.6). Unliganded $hUGDH_{A104L}$ (E*) was crystallized at 20 °C over a reservoir containing 16% polyethylene glycol 3350, 0.5 M NaCl, and 0.1 M Tris (pH 7.6). Crystals were cryoprotected using the reservoir conditions supplemented with 15% EDG (1:1:1 ethylene glycol:dimethyl sulfoxide:glycerol ratio) and then rapidly plunged into liquid nitrogen. All X-ray diffraction data were collected using the SER-CAT beamlines at the Advanced Photon Source (Argonne, IL). For the E complex ($hUGDH_{A104L}$:UDP-Glc:NADH), a 2.5 Å resolution data set was collected on the 22-ID beamline using a Rayonix MX300HS CCD detector. For the E* structure, a 2.65 Å resolution data set was collected on the 22-BM beamline using a MAR MX300 CCD detector. XDS[88] was used for data processing, and 5% of the data was set aside for cross validation.[89] The structure of the E complex ($hUGDH_{A104L}$:UDP-Glc:NADH) was determined by molecular

replacement using PDB entry 2Q3E as a search model in Phaser-MR of the PHENIX[90] software package. Because the $hUGDH_{A104L}$ E* crystal was isomorphous to the native E* structure (PDB entry 4RJT), the native R_{free} flags were preserved. Both structures were subjected to iterative rounds of manual rebuilding in COOT[91] followed by automated refinement in PHENIX.[90] The data collection and refinement statistics are listed in Table 3.1.

Normalized B-Factor Analysis

The *B*-factors of the isomorphous $hUGDH_{A104L}$ and native hUGDH (PDB entry 4RJT) E* crystal structures were set to 20 and then refined until convergence (approximately seven cycles) without TLS[92] or NCS restraints using PHENIX.[90] Because the NAD domain (residues 1–212), dimerization domain (residues 213–322), and sugar binding domain (residues 323–466) of each chain behave as separate rigid bodies in the crystal, the *B*-factor of each Cα atom was normalized relative to the other Cα atoms within the same domain using the equation[107]:

$$B_{normalized} = (B - B_{mean})/B_{stdev} \qquad \text{(equation 1)}$$

where *B* is the *B*-factor of the specific Cα atom, B_{mean} is the average *B*-factor of the domain in question, and B_{stdev} is the standard deviation of the Cα atom *B*-factors within the domain. Next, the normalized *B*-factors for each domain in a crystal were averaged. Because the crystals are isomorphous, the differences in the average normalized *B*-factors of $hUGDH_{A104L}$ and native hUGDH will reflect only changes in flexibility due to the A104L substitution.

Sedimentation Velocity

hUGDH and $hUGDH_{A104L}$ were dialyzed for 24 h into 25 mM HEPES (pH 7.5) and 150 mM KCl and then quantified as described above. The samples were then diluted with the dialysis buffer to a final protein concentration of 9 μM, loaded into 12 mm double-sector Epon centerpieces, and equilibrated for 1 h at 20 °C. Sedimentation velocity data were collected at 50000

rpm and 20 °C in an Optima XLA analytical ultracentrifuge. Data were recorded at 280 nm in radial step sizes of 0.003 cm. SEDNTERP[93] was used to estimate the partial specific volumes of hUGDH (0.73840 mL/g) and $hUGDH_{A104L}$ (0.73872 mL/g) and the density (1.00726 g/mL) and viscosity (0.01018) of the buffer. SEDFIT[94, 97] was used to analyze the raw sedimentation data. Data were modeled as a continuous sedimentation coefficient distribution [*c(s)*] and were fit using the baseline, meniscus, frictional coefficient, and systematic time-invariant and radial-invariant noise. Fit data for both experiments had root-mean-square deviations (rmsd's) of <0.004 AU.

Progress Curve Analysis of Hysteresis

UDP-glucose and NAD^+ were purchased from Sigma, and UDP-xylose was purchased from Carbosource (Complex Carbohydrate Research Center, University of Georgia, Athens, GA). Reactions were monitored at 25 °C using an Olis RSM 1000 spectrophotometer equipped with a stopped-flow assembly and a 0.4 mm path length. Separate enzyme (460 nM) and substrate (1 mM UDP-glucose and 10 or 5 mM NAD^+) solutions were made in reaction buffer [50 mM HEPES (pH 7.5), 50 mM NaCl, and 5 mM EDTA] and equilibrated at 25 °C for 5 min. We used the 5 mM NAD^+ concentration for $hUGDH_{A104L}$ because the A104L substitution increased the apparent affinity for the coenzyme (Table C3). Equal volumes of enzyme and substrate solutions were then rapidly mixed using the stopped flow. Absorbance readings at 340 nm were taken every 0.1 s for a total of 200 s for native hUGDH and 90 s for $hUGDH_{A104L}$ to monitor NADH production. Native hUGDH progress curves were fit to Frieden's equation for enzyme hysteresis:[52]

$$P(t) = v_{ss}t - \tau(v_{ss} - v_i)(1 - e^{-t/\tau}) \qquad \text{(equation 2)}$$

where P is the concentration of the product at time t and τ is equal to $1/k_{obs}$, where k_{obs} is the apparent rate constant for the transition between the initial rate (v_i) and the steady state rate (v_{ss}). For hysteretic enzymes, the initial velocity defined by eq 2 is the pre-steady state rate and does not

follow the steady state approximation. v_{ss} represents the initial steady state velocity and is calculated before 10% of the substrate is converted to product. The length of the lag is calculated as eτ. The $hUGDH_{A104L}$ progress curves did not display hysteresis, and the initial steady state velocity was derived from the linear portion of the progress curves prior to depleting 10% of the substrate. Data were fit using PRISM (GraphPad Software Inc., San Diego, CA), and all fits were analyzed using residual plots.[95]

Steady State Kinetics

The assay conditions were as previously described,[39, 44, 45, 48, 49, 105] with the exception that the enzyme dilution buffer was changed to 50 mM HEPES (pH 7.0), 50 mM NaCl, and 5 mM EDTA. Briefly, for NAD^+ saturation curves, the assays contained 100 nM enzyme in 50 mM HEPES (pH 7.5), 50 mM NaCl, and 5 mM EDTA with a saturating concentration of 1 mM UDP-glucose. The same conditions were used for UDP-Glc saturation curves, except that the NAD^+ concentration was fixed at 5 mM. Because hUGDH displays negative cooperativity with respect to NAD^+ binding, it is difficult to saturate the low-affinity sites with the coenzyme (see Results).[49] $hUGDH_{A104L}$ does not exhibit negative cooperativity with respect to NAD^+ binding and is saturated with 2.5 mM coenzyme. Prior to being mixed, the enzyme and substrate were incubated separately at 25 °C for 5 min. The reactions were initiated by rapidly mixing enzyme and substrate solutions, and the progress curves were measured by continuously monitoring the change in 340 nm absorbance every 0.5 s using an Agilent 8453 UV/vis spectrometer. The NADH concentration was determined at 340 nm using a molar absorptivity coefficient of 6220 M^{-1} cm^{-1}. Because native hUGDH is hysteretic, the steady state velocities were calculated by fitting progress curves to eq 2 as described above.[39, 44, 45, 48, 49, 105] $hUGDH_{A104L}$ does not display hysteresis; the steady state velocity was determined by fitting the initial liner portion of the progress curves to a straight line.

Initial steady state velocities from progress curves were fit to a sigmoidal rate equation and analyzed with residual plots:[95]

$$v = \frac{k_{cat}[E_t][S]^h}{K_M{}^h + [S]^h} \qquad \text{(equation 3)}$$

where *h* is the Hill coefficient. The negative cooperativity observed in hUGDH NAD^+ substrate saturation curves was evaluated using Kurganov's analysis as described previously.[23, 105] Briefly, the data were fit to the equation:

$$\frac{v}{[S]} = \frac{V_{max} - v}{K_M^{eff}}, \text{ where } K_M^{eff} = K_0 + (K_{lim} - K_0)(v/V_{max}) \qquad \text{(equation 4)}$$

where K_0 is the approximate K_M for the high-affinity binding sites and K_{lim} is the average apparent K_M for the low-affinity sites.[23]

Inhibition Kinetics

The K_i for the inhibitor UDP-Xyl was determined as previously described.[39, 44, 45, 48, 49, 105] Briefly, UDP-Xyl competes with UDP-Glc for the active site and induces a conformational change in hUGDH to the inhibited E^Ω state. Because the E and E^Ω states have different affinities for UDP-Glc and UDP-Xyl, the inhibition studies show cooperativity.[48] To determine the K_i, the substrate saturation curves at different concentrations of UDP-Xyl were fit simultaneously to an equation for competitive inhibition corrected for cooperativity using global analysis in Prism:

$$v_o = \frac{k_{cat}[E_t][S]^h}{(K_M^{app})^h + [S]^h}, \text{ where } K_M^{app} = K_M\left(1 + \frac{[I]}{K_i}\right) \qquad \text{(equation 5)}$$

3.4 Results

Rational Design of the A104L Cavity Filling Substitution

The cavities and declivities in the E state protein core are believed to provide space and accommodate the repacking of the allosteric switch and surrounding residues into the E^Ω conformation (Figure 3.1.D). Our goal was to design a cavity filling amino acid substitution that

would stabilize the E state and prevent the repacking of the protein core to the E^{Ω} state. We chose cavity C-I (109.9 ± 6.0 $Å^3$) because it includes the Cβ atom of A104, which is a good target for small-to-large amino acid substitutions (Figure 3.2.A and Table S3.1). In fact, A104 is conserved as a phenylalanine in the nonallosteric UGDH from Streptococcus pyogenes, which does not have a cavity corresponding to C-I.[39] The C-I cavity in hUGDH will not accommodate an A104F substitution without significant changes in the protein core. In fact, we are not able to express soluble hUGDH with the A104F substitution. This is most likely due to the low level of sequence identity (23%) between SpUGDH and hUGDH, which has resulted in numerous changes in the protein core packing.[39] As an alternative, we used RosettaVIP to predict cavity filling amino acid substitutions that would preserve the integrity of the protein core. The following cavity C-I residues were included in the automated redesign routine of RosettaVIP: T91, T93, R102, A103, A104, L106, V132, K279, D280, and N283. Cavity residues P92 and P133 were excluded because prolines reduce the configurational entropy of the peptide backbone and are often important for protein folding and stability.[67] Additionally, P133 is conserved in nonallosteric UGDH from S. pyogenes, which suggests that it may be significant to the structure or function of the enzyme.[39] T131 was also excluded from the redesign because it contributes to the catalytic mechanism by coordinating a nucleophilic water.[46] RosettaVIP identified A104L as the optimal cavity filling substitution; no other substitution was identified in multiple (<10) stochastic design simulations. The predicted A104L substitution fills approximately 61% of cavity C-I and will favor the E state by introducing steric clashes into the core of the E^{Ω} state (Figure 3.2.B-C).

We also examined the A104L model for unsatisfied polar groups contributing to cavity C-I. Native cavity C-I contains three ordered water molecules that form several hydrogen bonds with polar groups in the protein core (Figure 3.2.A). The first water (W1) forms hydrogen bonds with

the carbonyl oxygens of R102 and K279 as well as Oγ of T93 and the second water molecule (W2) (Figure 3.2.A). Water (W2) also forms hydrogen bonds with the carbonyl oxygen of T91 and the third water molecule (W3), which also accepts a hydrogen from Nδ of N283 (Figure 3.2.A). The RosettaVIP A104L model conserves a small cavity that can accommodate W1 and its hydrogen bonds; however, the leucine substitution will displace W2 and W3 (Figure 3.2.A-B). In the absence of W2 and W3, the carbonyl oxygen of T91 can still be satisfied by Nζ of K279, and the caboxamide of N283 can be exposed to solvent by a small rotation about the χ1 torsion angle.

The $hUGDH_{A104L}$ Crystal Structure Reveals a Reduced Cavity C-I Volume

hUGDH with the A104L substitution ($hUGDH_{A104L}$) was co-crystallized with UDP-Glc and NADH to obtain the E state conformation ($hUGDH_{A104L}$:UDP-Glc:NADH). The structure was determined to a resolution of 2.53 Å and revealed two 32 symmetry hexamers in the asymmetric unit (Table 3.1). All residues are modeled with the exception of loop 385–387 and the C-terminus (residues 467–494), which are disordered in all 12 chains. Residues 23–30, 34–54, and 69–72 in the NB domain of chain A were poorly ordered and not modeled. Each monomer contains unambiguous electron density for the bound NADH and UDP-Glc (not shown). In all chains, the allosteric switch and hexamer-building interface are in the E state. The largest structural difference between the chains involves the rotation of the NAD^+ domain; chains G and J adopt the closed conformation, while the 10 remaining chains have NAD^+ domains that are rotated between 6.8° and 10.6° toward the open conformation. In chains G and J, the allosteric switch and hexamer-building interface are in the E state (Figure S3.1). Because the open domain conformation exposes the C-I cavity to solvent, we will limit our analysis of the A104L-induced changes in the cavity volume to the closed domain conformations in chains G and J.

The crystal structure of $hUGDH_{A104L}$ agrees well with the $Rosetta_{VIP}$ model (Figure 3.2.B). The A104L substitution has split cavity C-I into two, smaller cavities with volumes of 20.7 ± 2.2 and 33.3 ± 2.9 $Å^3$ (Figure 3.2.B). The smaller cavity contains weak electron density for a water corresponding to W1 in chains B–F, K, and L. The electron density in the remaining chains is weak and did not support modeling the water. Because the absence of water W1 would leave unsatisfied polar groups, it is likely that there is a poorly ordered water in the remaining cavities (Figure 3.2.B).[66, 108] As expected, the L104 side chain has displaced waters W2 and W3 (Figure 3.2.B). As a result, the Cε and Nζ atoms of K293 are disordered, which exposes the carbonyl oxygen of T91 to solvent. To compensate for the loss of the hydrogen bond with W3, the χ2 torsion angle of N283 has rotated from −10.3 ± 6° to the more favorable rotomer value of 49.4 ± 0.5°, which exposes the carboxamide to solvent. The only significant difference between the $hUGDH_{A104L}$ structure and the $Rosetta_{VIP}$ model involves the χ2 torsion angle of N283, which is rotated 180° to flip the amide and carbonyl oxygen of the carboxamide group. Because it is not possible to assign the absolute conformation of the carboxamide group in the absence of known hydrogen bond acceptors or donors at this resolution, we chose to model the χ2 rotamer in the conformation originally reported in the 2Q3E crystal structure.[42]

To see if the A104L substitution destabilized the hexamer, we examined the oligomeric structure of native hUGDH and $hUGDH_{A104L}$ using sedimentation velocity. Consistent with previous reports,[39, 44, 45, 48, 49, 105] the sedimentation of 9 μM native hUGDH reveals an 11.4 S hexamer (78.0%), an 8.2 S tetramer (3.8%), a 5.5 S dimer (13.3%), and a 3.1 S peak that likely represents a monomer (4.9%) (Figure 3.2.D). Similarly, the *c(s)* distribution of 9 μM $hUGDH_{A104L}$ consists of the 11.5 S hexamer (83.5%), a small amount of tetramer (<1%), the 5.5 S dimer (9.6%), and the 3.0 S monomer (6.9%) (Figure 3.2.D). The *c(s)* distribution of native hUGDH is in rapid

equilibrium, with the tetramer representing a transient between the more stable hexamer and dimer states.[39, 45, 48, 98, 105] Because of this, it was difficult to assign an accurate S value for the tetramer in the $hUGDH_{A104L}$ distribution.

$hUGDH_{A104L}$ Adopts the E State*

$hUGDH_{A104L}$ was also crystallized under conditions known to stabilize the native E* state.[48] The resulting crystals diffract to a resolution of 2.65 Å and are isomorphous with the native E* structure (PDB entry 4RJT) (Table 3.1). Briefly, the asymmetric unit contains three chains, with the complete hexamer being formed by crystallographic symmetry. All residues are modeled except for a surface loop (residues 385–387) and the C-terminus (residues 467–494), which are disordered in every chain. Additionally, there was incomplete main chain electron density for residue 72 in chain A as well as residues 13, 14, 49, and 154 in chain C, so they were not modeled. As in the native E* crystal structure,[48] the NB domains of E* $hUGDH_{A104L}$ are rotated to the open conformation. The only significant conformational change in the hexamer involves the NB domain of $hUGDH_{A104L}$ chain C, which is rotated ~0.9° less open than the corresponding chain in the native E* structure. The allosteric switch and hexamer-building interface of $hUGDH_{A104L}$ are also in the E* conformation (Figure 3.3.A). In native hUGDH, four residues in the protein core adopt $\chi 1$ torsion angles that that are unique to the E* state (Figure 3.3.B-C). Residues S88 and E161 adopt unique $\chi 1$ angles in the open domain E* state that would introduce steric clashes if the domain rotated to the closed conformation (Figure 3.3.B-C). Residue E128 changes conformation with the allosteric switch and adopts a unique rotamer conformation in the E, E*, and E^{Ω} states (Figure 3.3.C). Finally, S120 adopts a $\chi 1$ rotamer angle of 67 ± 1°, which is observed only in the E* conformation. S120 is located at the C-terminus of the α5 helix, which is adjacent to the allosteric α6 helix in the hexamer-building interface (Figure 3.3.B-C). It is likely that the change

in the S120 χ1 rotamer is correlated with the movement of the α6 helix. In both χ1 rotamers, Oγ of S120 caps the carboxy terminus of the α5 helix by contributing a hydrogen bond to the carbonyl oxygen of I116. In the E* conformation of $hUGDH_{A104L}$, the χ1 rotamers of S88, E128, and E161 are conserved. However, Oγ of S120 is disordered and could not be modeled unambiguously.

The open domain conformation of E* exposes the allosteric switch and surrounding residues to solvent, which relaxes the steric packing constraints that would otherwise resist the transition between the E and E^{Ω} states.[44] A comparison of the $hUGDH_{A104L}$ and native E* normalized B-factors shows that the A104L substitution decreases the flexibility of several residues in cavity C-I (Figure 3.4.A). These changes originate from the A104L substitution, which packs against the allosteric switch residues V132 and P133, and the carboxamide of N283 (Figure 3.4.B-C). These stabilizing interactions are propagated to the surrounding second-shell residues V285, Y286, and L287, which become significantly more rigid (Figure 3.4.A-B). In contrast, the residues that are unique to cavity C-2 and declivities D-1 and D-2 show no significant change in flexibility (Figure 3.4.A).

$hUGDH_{A104L}$ Does Not Display Hysteresis or Negative Cooperativity

Unlike those of native hUGDH, the progress curves of $hUGDH_{A104L}$ do not exhibit hysteresis (Figure 3.5.A-B and Table 3.2). In fact, we observe no lag in the progress curves of $hUGDH_{A104L}$ even in the presence of UDP-Xyl, which is known[48, 49] to increase the lag in native hUGDH progress curves (Table 3.2). Next, we examined the impact of the A104L substitution on steady state kinetics. The NAD^+ saturation curves of $hUGDH_{A104L}$ are hyperbolic (Hill = 1), with a K_M of 48 ± 3 μM (Figure 3.5.C and Table 3.3). In contrast, the NAD^+ saturation curves of native hUGDH are known to display negative cooperativity[49] (Figure 3.5.D and Table 3.3). Earlier work showed that UDP-Glc binding induces an asymmetry that results in a mixture of high- and low-

affinity NAD^+ binding sites.[49] Using Kurganov's analysis[23] (eq 4), the high (K_o) and low (K_{lim}) NAD^+ binding affinities for native hUGDH can be estimated (157 ± 13 and 2300 ± 500 μM, respectively) (Figure 3.5.E and Table 3.3). Both native hUGDH and $hUGDH_{A104L}$ have similar k_{cat} values with respect to NAD^+ (0.94 ± 0.07 and 0.79 ± 0.01 s^{-1}, respectively). The UDP-Glc saturation curves for both $hUGDH_{A104L}$ and native hUGDH are hyperbolic, with similar values for K_M (9.4 ± 0.5 and 7.6 ± 0.6 μM, respectively) (Table 3.3). The k_{cat} values observed in the UDP-Glc and NAD^+ saturation curves of $hUGDH_{A104L}$ are also similar (0.82 ± 0.01 and 0.79 ± 0.01 s^{-1}, respectively) (Table 3.3). The difference between the kcat values measured in the UDP-Glc and NAD^+ saturation curves of native hUGDH (0.59 ± 0.01 and 0.94 ± 0.07 s^{-1}, respectively) is a consequence of the negative cooperativity in NAD^+ binding, which makes it difficult to saturate hUGDH with coenzyme during the pseudo-first-order steady state analysis.[49]

$hUGDH_{A104L}$ Does Not Form the E^{Ω} State

Because the A104L substitution will introduce steric clashes into the E^{Ω} state, $hUGDH_{A104L}$ should favor the E and E* states. Attempts to crystallize $hUGDH_{A104L}$ in the E^{Ω} state using both of the previously published conditions[39] were not successful. However, the allosteric transition from the E^{Ω} to the E conformation can be observed in steady state inhibition studies of native hUGDH.[48, 105] Briefly, UDP-Xyl binding favors the E^{Ω} state, which has a low affinity for UDP-Glc. As UDP-Glc outcompetes UDP-Xyl, the allosteric transition to the high-UDP-Glc affinity E state can be observed as substrate cooperativity (Figure 3.5.F-G and Table 3.4). As previously shown,[48] the degree of cooperativity increases with inhibitor concentration (Figure 3.5.F-G and Table 3.4). Global analysis of $hUGDH_{A104L}$ steady state inhibition curves shows that the UDP-Xyl affinity is 13-fold lower (K_i = 2.3 ± 0.2 μM) than that of native hUGDH (K_i = 0.18 ± 0.01 μM) (Table 3.4). In addition, none of the steady state inhibition curves for $hUGDH_{A104L}$ show evidence

of substrate cooperativity (Figure 3.5.H). Even at a saturating UDP-Xyl concentration, 15-fold higher than the K_i, the substrate saturation curves of $hUGDH_{A104L}$ are hyperbolic (Figure 3.5.I). This suggests that $hUGDH_{A104L}$ does not undergo the transition to the low-substrate affinity E^{Ω} state upon inhibitor binding.

The A104G Substitution Has No Significant Effect on Hysteresis or Allostery

The A104L substitution reduces the volume of cavity C-I and produces a more rigid protein core (Figures 3.2 and 3.4). We also investigated the effect of increasing the cavity volume and core flexibility using the large-to-small A104G substitution. The $hUGDH_{A104G}$ construct is expressed at a level (~2.5 mg/L) much lower than that of the native enzyme (~20 mg/L), which suggests that the substitution has negatively impacted folding. A decrease in stability may also explain the higher UDP-Glc K_M observed in $hUGDH_{A104G}$ (17.2 ± 1.1 compared to 7.4 ± 0.6 in the native form) (Table 3.4). Global analysis of UDP-Xyl inhibition shows that the UDP-Xyl affinity (K_i) of $hUGDH_{A104G}$ and substrate cooperativity are similar to that of the native enzyme (Figure 3.5.J and Table 3.4). Finally, the A104G substitution does not appear to have a significant effect on enzyme hysteresis; the progress curves of $hUGDH_{A104G}$ display lags that are similar to that observed in native hUGDH (Figure 3.5.K and Table 3.2).

3.5 Discussion

The allosteric inhibition mechanism in hUGDH involves the ligand-induced isomerization of the T131 loop from the E or E* state to the E^{Ω} conformation (Figure 3.1.A–C).[39, 42, 44] In the absence of ligand, the NB domain is rotated open to expose the allosteric switch and surrounding core residues to solvent (Figure 3.6.Ai). This open conformation relaxes the core packing constraints, which allows the allosteric switch and some core residues to adopt the inactive E* state (Figure 3.3.A&C). The binding of UDP-Glc and NAD^+ triggers the allosteric switch and core

residues to adopt the active E state conformation, which allows the domains to rotate closed (Figure 3.6.Aii&iii). In contrast, UDP-Xyl binding induces the allosteric switch and the protein core to repack into the E^{Ω} state (Figure 3.6.Aiv&v). This repacking is facilitated by the remarkable plasticity of the protein core, which originates from the large cavities and declivities surrounding the T131 loop (Figures 3.1.D and 3.6.A).[39] In previous work,[105] we tested the hypothesis that the repacking of the allosteric switch is the source of hysteresis using the A136M substitution ($hUGDH_{A136M}$). We showed that A136M traps the allosteric switch in the active E state and abolished ligand-induced hysteresis (Figure 3.6.Bi → ii). However, it was not clear if the hysteresis was a feature of the inactive E* state, the inhibited E^{Ω} complex, or the repacking of the protein core.

Here, we have examined the contribution of protein core flexibility to both allostery and hysteresis by filling cavity C-I with the A104L substitution (Figures 3.1.D and 3.2.A). This cavity is present only in the E state, and our modeling studies showed that the A104L substitution would sterically prevent the repacking of the protein core into the E^{Ω} conformation (Figures 3.2.C and 3.6.C). We tested this using steady state inhibition kinetics; the allosteric transition in native hUGDH can be observed as substrate-induced cooperativity in the presence of a saturating level of UDP-Xyl[48] (Figure 3.5.F&G and Table 3.4). Thus, the absence of cooperativity in UDP-Xyl inhibition studies is strong evidence that the A104L substitution prevents the formation of the E^{Ω} state (Figure 3.5.H&I and Table 3.4). This interpretation is also consistent with the reduced UDP-Xyl affinity of $hUGDH_{A104L}$, which has been shown[105] to be a property of a constitutive E state $hUGDH_{A136M}$ (Table 3.4).

The progress curves of $hUGDH_{A104L}$ do not display hysteresis, despite the fact that the crystal structures show that the enzyme can adopt the E and E* states (Figures 3.3.A, 3.5.A, and

3.6.C, i). This is the first evidence that the source of hysteresis in hUGDH is not simply the movement of the allosteric switch from the E* to E conformation.[39, 44, 48] Our data suggest that the conformational entropy of the native hUGDH protein core is an important factor in enzyme hysteresis. In this model, the large packing defects allow residues in the protein core to sample nonproductive conformations as they converge to the correct E state rotamers. By filling cavity C-I with the A104L substitution, we sterically constrain the conformational freedom of the core, which should increase the rate at which the core residues converge to the E state conformation (Figure 3.6.C, i → ii). This interpretation is supported by our normalized *B*-factor analysis of the native and $hUGDH_{A104L}$ crystal structures, which shows $hUGDH_{A104L}$ has a more rigid protein core surrounding the substitution (Figure 3.4.A). To see if increasing the flexibility of the protein core would increase hysteresis, we introduced the A104G substitution. $hUGDH_{A104G}$ displays a rate of hysteresis and an inhibitor affinity that are similar to those of the native enzyme (Tables 3.2 and 3.4). This suggests that the main chain flexibility of A104 does not contribute to hysteresis. It is also likely that the small increase in cavity C-I volume due to the A104G substitution does not significantly increase the conformational entropy of the protein core residues. Taken together, our data show for the first time that the packing defects in the core of hUGDH play a significant role in both the allosteric and hysteretic mechanisms of hUGDH.

3.6 Acknowledgments

The authors thank Phillip G. Gross for helpful discussions. The authors thank the staff at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, which is supported in part by the National Institutes of Health (S10 RR25528 and S10 RR028976). The authors also acknowledge the National Institute of General Medical Sciences for an equipment grant supporting our in-house X-ray facility (S10 OD021762).

3.7 Figures and Tables

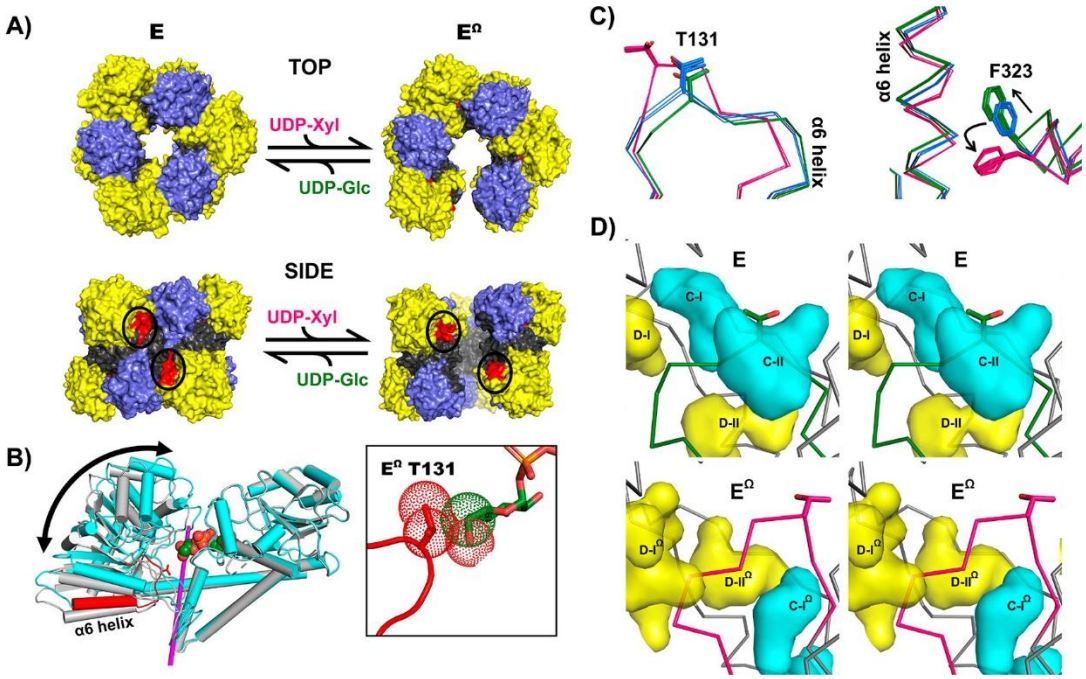


Figure 3.1: Allostery and Hysteresis in hUGDH. (A) Active E and allosterically inhibited E^{Ω} states of hUGDH. The allosteric switch (red), the NB domain (yellow), the dimerization domain (gray), and the SB domain (purple) are identified. The α6 helix of the allosteric switch helix (red) in the hexamer-building interface is identified with black circles. **(B)** Cartoon depicting the rotation of the NB domain (curved arrow) about the hinge bending axis (magenta rod) from the open (gray) to the closed (cyan) state. The allosteric switch (red) and UDP-Glc (spheres) in the active site are shown. The inset box shows a close-up of the active site and depicts the van der Waals clash (dots) that would occur between T131 of the allosteric switch in the E^{Ω} state (red) and the C5 hydroxymethyl of UDP-Glc (green). UDP-Xyl lacks the C5 hydroxymethyl. **(C)** The left panel depicts the E (green), E* (blue), and E^{Ω} (pink) conformations of the T131 loop. The right panel shows the corresponding conformations of the α6 helix and F323 of the adjacent subunit in the hexamer-building interface. The arrows depict the directions that F323 in the E* state (blue) moves to adopt the E (green) or E^{Ω} state (pink). **(D)** Stereo diagrams of the allosteric switch in the E (top panel, green sticks) and E^{Ω} (bottom panel, pink sticks) states. Cavities (blue surfaces) and declivities (yellow surface) are identified. Surfaces were calculated using a 1.2 Å probe. The residues lining each cavity are listed in Table S1.

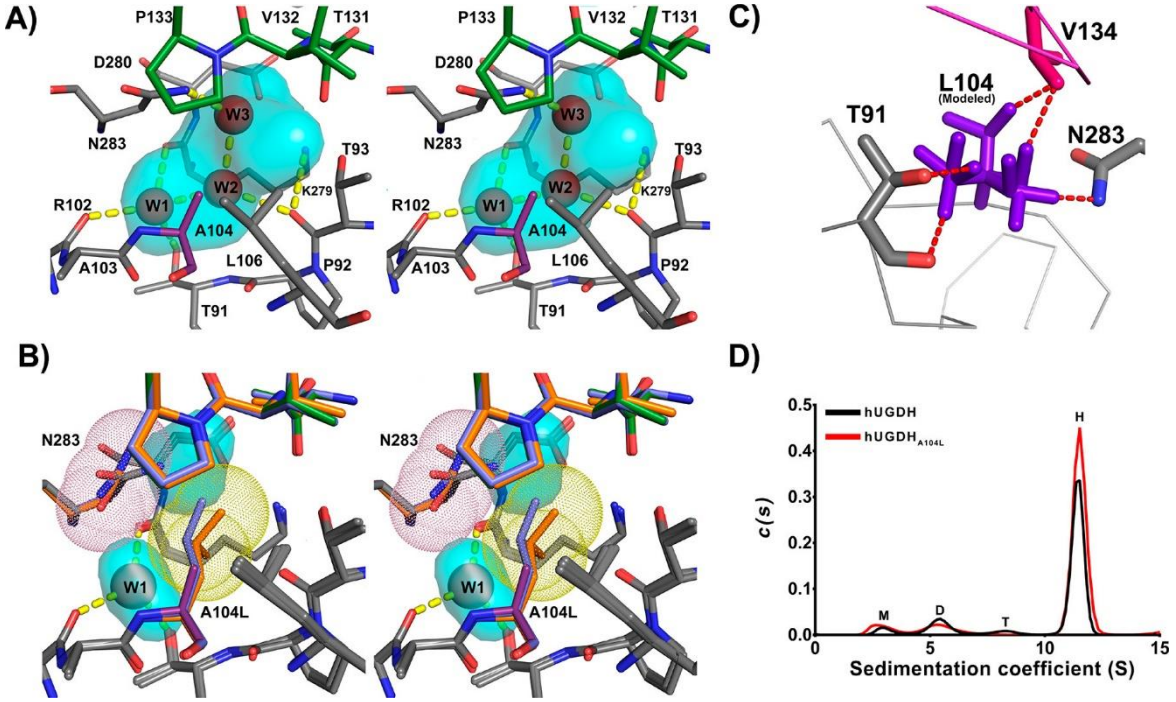


Figure 3.2: RosettaVIP Design of the A10L Cavity Filling Substitution in the E State. (A) Stereoview of the residues that line E state cavity C-I (cyan surface) and share hydrogen bonds (yellow dashes) with three buried waters (red spheres). A104L and the allosteric switch residues are colored purple and green, respectively. **(B)** Residues that line cavity C-I in the native hUGDH structure superimposed onto the RosettaVIP model and the $hUGDH_{A104L}$ crystal structure. The native A104 is colored purple (native), and all other residues are colored gray. The A104L substitution and allosteric switch are colored slate (RosettaVIP model), orange ($hUGDH_{A104L}$), or green (native hUGDH). The residual surfaces (cyan) of cavity C-I in $hUGDH_{A104L}$ and a single buried water molecule (gray sphere) are also shown. Surfaces were calculated with a 1.2 Å probe. The van der Waals radii of the $hUGDH_{A104L}$ L104 and the N283 carboxamide are depicted as yellow and pink dots, respectively. **(C)** Modeling L104 into the E protein core shows the substitution will introduce steric clashes in all rotameric states. In the E^{Ω} state (PDB entry 3PTZ), all preferred rotamers of L104 (purple) introduce steric clashes (red dashes). Even strained rotamers are sterically excluded (not shown). The E^{Ω} state of the allosteric switch is colored pink. **(D)** *c(s)* distribution of $hUGDH_{A104L}$ (red) and native hUGDH (black) as determined by sedimentation velocity analysis. M, D, T, and H correspond to the predicted S value of the monomer, dimer, tetramer, and hexamer, respectively.

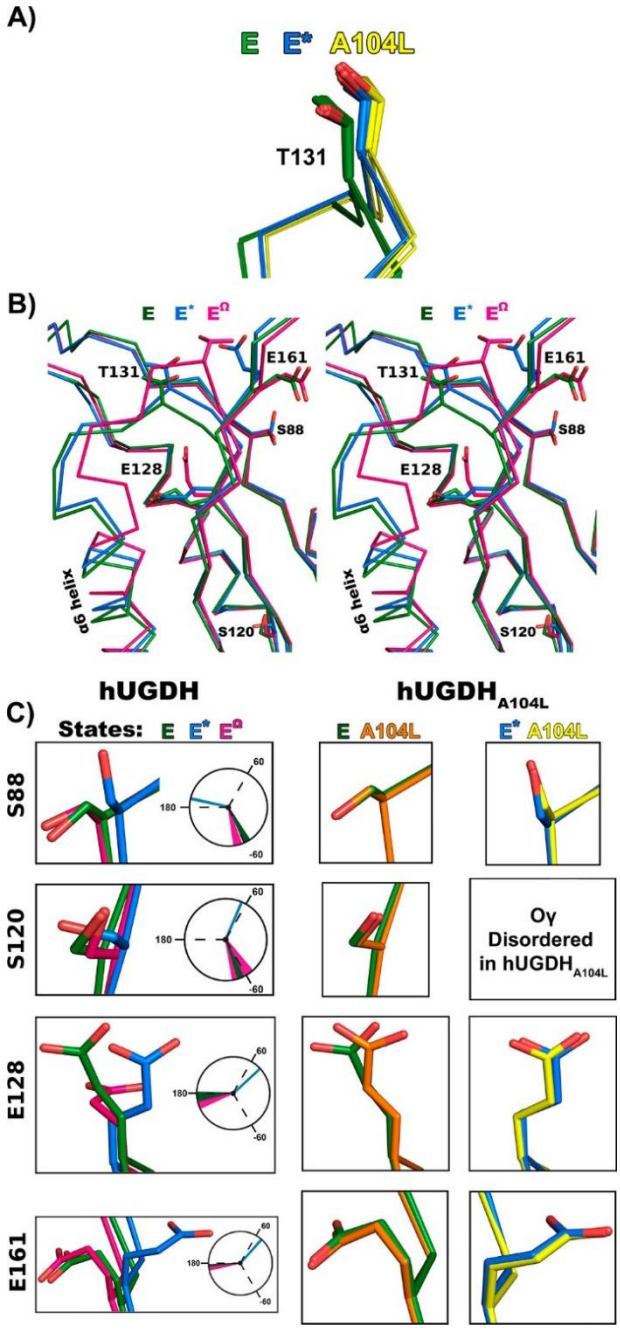


Figure 3.3: E* State of hUGDH$_{A104L}$. (A) The E* state of the hUGDH$_{A104L}$ allosteric switch (yellow) superimposes onto the native E (green) and E*(blue) conformations. **(B)** Stereodiagram of the Cα traces of the hUGDH protein core surrounding the allosteric switch in the E (green), E* (blue), and E$^{\Omega}$ (pink) states. Residue T131 is shown as sticks, as are residues that adopt rotamers unique to the E* state. The figure was created by superimposing residues 1–124. **(C)** Residues that adopt rotamers unique to the E* state. The column labeled hUGDH shows representative rotamers from the E (green), E* (blue), and E$^{\Omega}$ (pink) states as sticks with the corresponding χ1 angles observed in all chains in the native structures represented in a Newman projection. The last two columns labeled hUGDH$_{A104L}$ show the E (orange) and E* (yellow) conformations superimposed onto the corresponding native conformation.

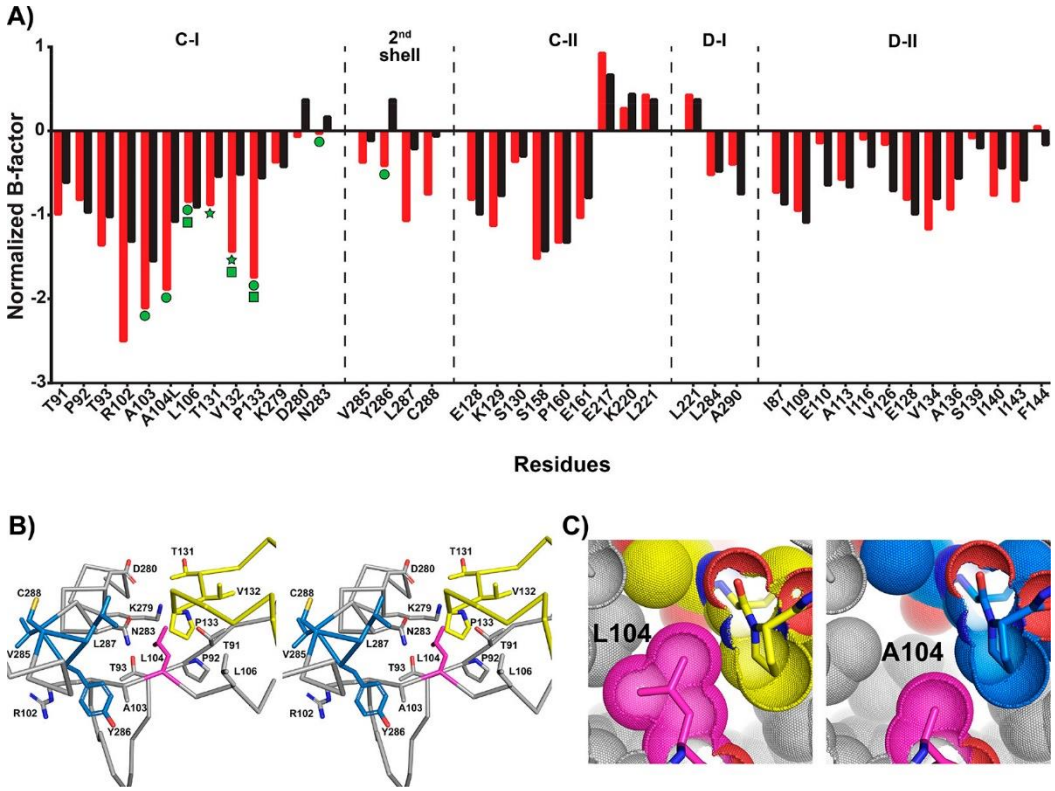


Figure 3.4.: A104L Substitution Decreases the Flexibility of Cavity C-I and Surrounding Residues. (A) Histogram of the normalized *B*-factors (eq 1) for the Cα atoms of residues lining the packing defects (C-I, C-II, D-I, and D-II) of native hUGDH (black) and hUGDH$_{A104L}$ (red) structures. The second shell identifies residues that are adjacent to cavity C-I that show a significant decrease in flexibility. Residues that are shared between cavity C-I and the other packing defects are identified in the C-I plot with green symbols (stars, circles, and squares for C-II, D-I, and D-II, respectively). **(B)** Stereoview of the A104L form (magenta), allosteric switch (yellow), remaining cavity C-I residues (gray), and second-shell residues (blue). **(C)** The left panel depicts new van der Waals contacts between A104L (magenta) and the E* allosteric switch (yellow). The right panel is the same view with the native A104 (magenta) and the native E* allosteric switch (blue).

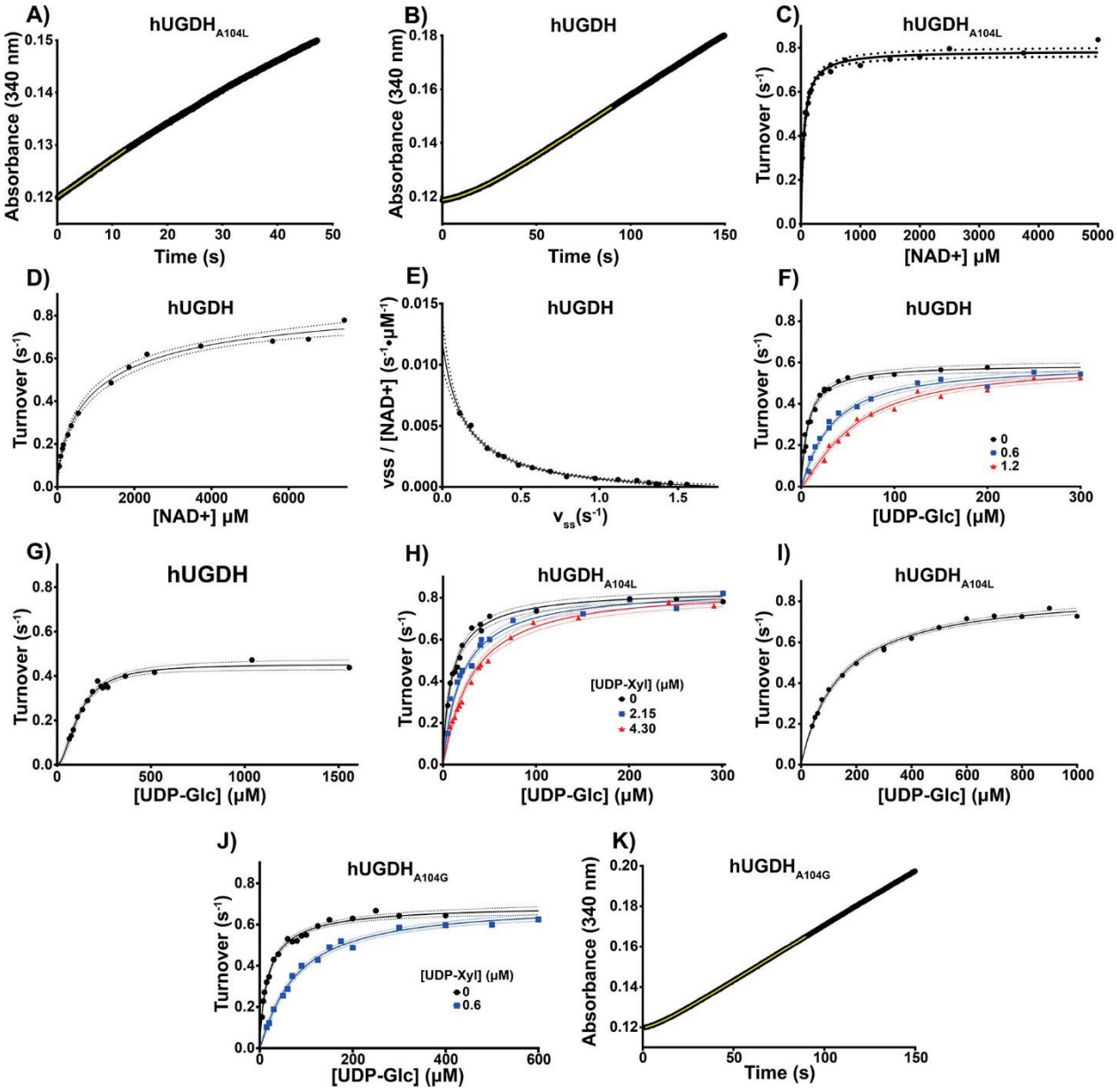


Figure 3.5: Kinetic Analysis of Native hUGDH and hUGDH$_{A104L}$. (A) Stopped-flow absorbance spectroscopy of an hUGDH$_{A104L}$ progress curve with the saturating substrate fit to a straight line (yellow). **(B)** Same as panel A but of a native hUGDH fit (yellow line) to Frieden's equation for hysteresis (eq 2). **(C)** NAD^+ substrate saturation curve with hUGDH$_{A104L}$ fit to eq 3. Rates are normalized by enzyme concentration to give turnover, and dashed lines represent the 95% confidence interval of the fit in all panels. **(D)** NAD^+ substrate saturation curve with native hUGDH as in panel C. For native hUGDH, the initial steady state velocities were obtained by fitting progress curves to Frieden's equation for hysteresis (eq 2), as described in Methods. **(E)** Eadie–Hofstee plot of the hUGDH NAD^+ saturation data that is concave up (negative cooperativity) and was fit to eq 4 to determine the K_0 and K_{lim} values in Table C4 (described in Methods). **(F)** hUGDH UDP-glucose saturation curves with 0 µM (black circles), 0.6 µM (blue squares), and 1.2 µM (red triangles) UDP-Xyl were globally fit to eq 5. **(G)** hUGDH UDP-Glc

saturation curves with 2.7 μM UDP-Xyl were fit to eq 3. **(H)** $hUGDH_{A104L}$ UDP-Glc saturation curves with 0 μM (black circles), 2.2 μM (blue squares), and 4.3 μM (red triangles) UDP-Xyl were globally fit to eq 5. **(I)** $hUGDH_{A104L}$ UDP-Glc saturation curves with 34 μM UDP-Xyl were fit to eq 3. **(J)** $hUGDH_{A104G}$ UDP-glucose saturation curves with 0 μM (black circles) and 0.6 μM (blue squares) UDP-Xyl were globally fit to eq 5. **(K)** The $hUGDH_{A104G}$ progress curve was fit to eq 2 (yellow).

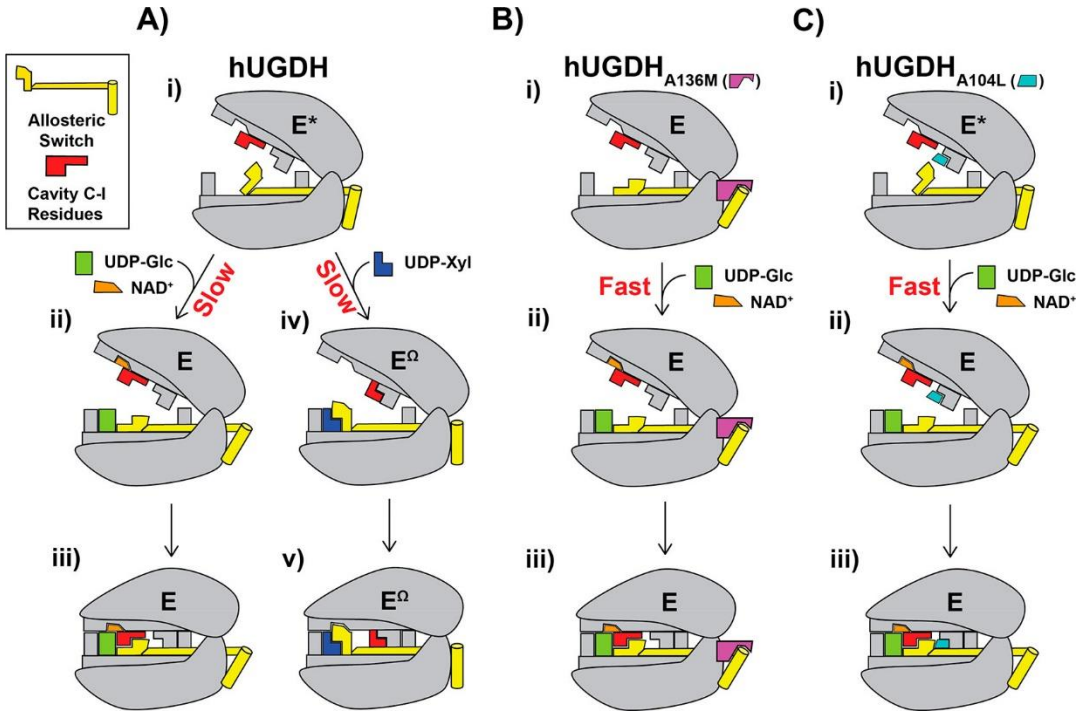


Figure 3.6: Proposed Mechanism for Allostery and Hysteresis in hUGDH. (A) (i) A single monomer of hUGDH is represented by a gray clamshell with the allosteric switch (yellow) in the E* state and cavity C-I residues depicted as a red L-shape. The cavity C-I residues are depicted in the E state in the open domain conformation (as observed in the crystal structures), but it is likely that they can sample alternate rotamers, including those of the E^{Ω} state. (ii) Binding of UDP-Glc (green rectangle) and NAD^+ (orange trapezium) triggers the allosteric switch to adopt the active E state, and (iii) the domains rotate closed with the cavities and declivities represented as white spaces between the domains. In contrast, (iv) binding of UDP-Xyl (blue L-shape) triggers an allosteric switch; the cavity C-I residues adopt the E^{Ω} conformation, and then, (v) the domains rotate closed. In the E^{Ω} conformation, the allosteric switch occludes the NAD^+ binding site. Previously, we proposed[15] that the slow isomerization of E* to E (i → ii) or EΩ (i → iv) originated from the repacking of the allosteric switch and core residues. **(B)** (i) The A136M substitution (purple polygon) traps the allosteric switch in the E state. (ii) Here, binding of UDP-Glc and NAD^+ does not induce a slow isomerization, and (iii) the domains rotate closed. **(C)** (i) The A104L substitution (cyan trapezoid) prevents the cavity C-I residues from sampling the EΩ state. However, the allosteric switch is free to adopt the E* state, as seen in Figure C3.A. (ii) Upon binding of UDP-Glc and NAD^+, the allosteric switch rapidly adopts the E state and (iii) the domains rotate closed.

Table 3.1: Data Collection and Refinement Statistics

	hUGDH$_{A104L}$:UDP-Glc:NADH	**hUGDH$_{A104L}$**
Protein Data Bank entry	6C4J	6C4K
Data collection		
Space Group	C121	C121
Unit cell dimensions (*a, b, c, β*)	403.5, 112.6, 184.7, 98.5°	178.5, 114.5, 96.9, 116.5°
Resolution Range	(48.63-2.53)	(48.16-2.65)
Completeness (%)	99.9 (99.9)[a]	99.7 (97.8)[a]
Redundancy	7.5 (7.5)	7.7 (7.7)
No. of reflections	2056278	389870
$I / \sigma(I)$	9.8 (1.0)	13.7 (1.1)
$CC_{1/2}$[b]	99.8 (53.3)	99.9 (52.8)
R_{meas}[c] (%)	14.8 (222.2)	12.5 (228.5)
Refinement		
Resolution (Å)	2.53 (2.56-2.53)	2.65 (2.71-2.65)
R_{work} / R_{free}	0.181/0.214	0.222/0.255
R_{free} in highest resolution shell	40.33	0.425
No. atoms Protein / Ligand / Water	43090 / 1043 / 594	10810 / 16 / 69
Wilson *B*-factor	62.3	71.5
B-factors (Å^2) Protein / Ligand / Water	73.2 / 72.6 / 59.1	105.3 / 77.5 / 63.5
Stereochemical Ideality		
Bond lengths (Å^2)	0.004	0.002
Bond angles (°)	0.64°	0.43°
φ,ψ Preferred region (%)	97.6	96.2
φ,ψ Additionally allowed (%)	2.4	3.8
φ,ψ Disallowed region (%)	0.0	0.0

[a] Values in parentheses are for the highest-resolution shell (2.559-2.530 Å and 2.712-2.650 Å for 6C4J and 6C4K, respectively).
[b] $CC_{1/2}$ is correlation between intensities from random half-data sets.[102]
[c] R_{meas} is the redundancy-independent merging *R* factor.[101]

Table 3.2: Effect of UDP-Xyl on Hysteresis

Enzyme	UDP-Xyl (μM)	τ (s^{-1})	Lag (*s*)
hUGDH	0	17.7 ± 1.4	48.2 ± 3.7
	1[a]	23.9 ± 2.0	65.0 ± 5.3
$hUGDH_{A104L}$	0	N.D.[b]	N.D.
	11.3[a]	N.D.	N.D.
$hUGDH_{A104G}$	0	16.2 ± 2.0	44.0 ±5.3
	1[a]	22.9 ± 2.2	62.2 ± 5.7

[a] These concentrations of UDP-Xyl are 5x K_i listed in Table 4
[b]N.D.= not detected

Table 3.3: Steady State Parameters

Enzyme	Substrate	K_M (μM)	Hill	k_{cat} (s^{-1})[a]	K_0[b]	K_{lim}[b]
hUGDH	UDP-Glc	7.6 ± 0.6	1	0.59±0.01[c]		
	NAD^+	1160 ± 290 [d]	0.69	0.94±0.07	157±13[b]	2300±500 [b]
A104L	UDP-Glc	9.4 ± 0.5	1	0.82±0.01		
	NAD^+	48 ± 3	1	0.79±0.01		

[a] One complete catalytic turnover produces 2 NADH.
[b] K_0 and K_{lim} represent the high and low affinity sites derived from Kurganov analysis, respectively.
[c] The k_{cat} for the UDP-Glc saturation curve is underestimated due to the negative cooperativity of NAD^+ binding (see text).
[d] Because of negative cooperativity, this value represents the apparent half-saturation of the ES complex $K_{0.5}$.

Table 3.4: Global Analysis of Competitive Inhibition

Enzyme	UDP-Xyl (μM)	UDP-Glc K_M (μM)	k_{cat} (s^{-1})[a]	K_i UDP-Xyl (μM)	Hill
hUGDH	0	7.4±0.6 [b]	0.59±0.01 [b]	0.18±0.01	1
	0.6				1.1±0.1
	1.2				1.3±0.1
	2.7[c]	7.4±0.1	0.45±0.01		1.85±0.14
$hUGDH_{A104L}$	0	10.3±0.6[b]	0.84±0.01 [b]	2.3±0.2	1
	2.2				1
	4.3				1
	34[c]	8.8±0.4	0.85±0.01		1
$hUGDH_{A104G}$	0	17.2 ± 1.1	0.67 ± 0.02	0.19±0.2	1
	0.6				1.2±0.1

[a] One complete catalytic turnover produces 2 NADH.

[b] For global analysis, this parameter was refined as a shared value.

[c] This experiment was excluded from global analysis because at long lags induced at high UDP-Xyl concentrations, native hUGDH is subject to NADH inhibition before the enzyme reaches steady state velocity.

3.8 Supplementary Figures and Tables

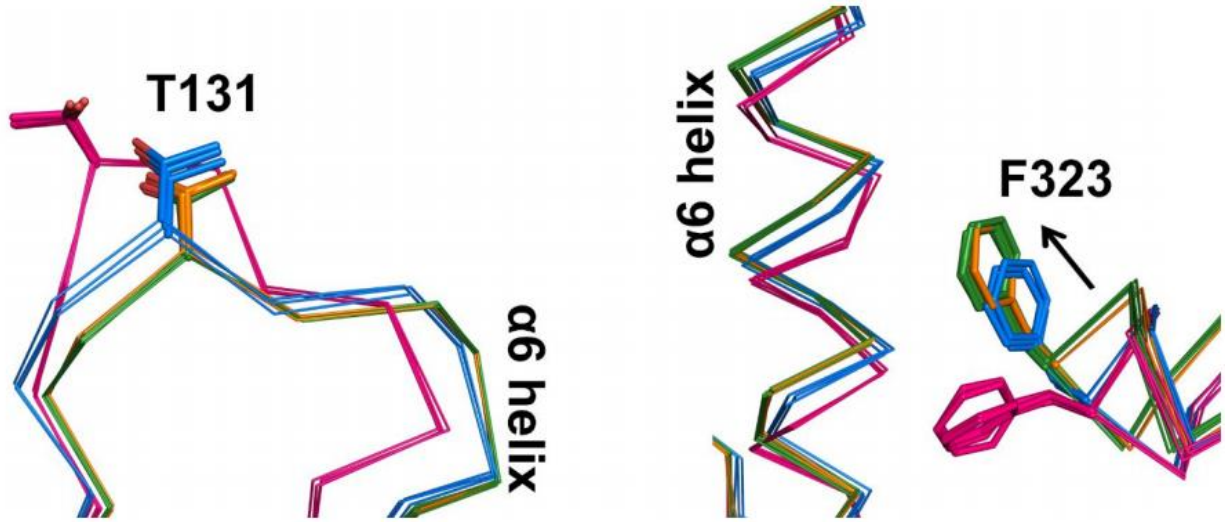


Figure S3.1.: When Substrate is Bound hUGDH$_{A104L}$ Adopts the E State. The hUDGH$_{A104L}$ (orange) structure superimposed onto the native structures in the E (green), E* (blue) and E$^{\Omega}$ (pink) conformations of the T131-loop (left). The α6 helix in the hexamer building interface and corresponding conformations of the F323 in the adjacent subunit (right). The arrows depict the directions that the F323 in the E*state moves to adopt the E or E$^{\Omega}$ states.

Table S3.1 Packing Defects and Associated Volumes

Packing Defect	Volume (Å^3)[a]	Waters[b]	Residue (Atoms)
C-1	109.9 ± 6.0	3	T91(O,Cβ,Cγ2,Oγ1), P92(C), T93(N,Cα,Oγ1), R102(O), A103(C,O), A104(N,Cα,Cβ), L106(Cδ2), T131(O,Cγ2), V132(Cα), P133(Cδ), K279(C,O,Cβ,Cγ,Cδ,Nζ), D280(Cα,Oδ1), N283(Cβ,Nδ2)
C-2	157.5 ± 16.8	4-5	E128(O,Cγ,Oε1), K129(O), S130(O), T131(Cα,Cγ2,Oγ1), V132(N,O,Cγ2), S158(Oγ), P160(N,Cα,Cβ,Cγ,Cδ) E161(N,O,Cβ,Cγ), E217(Cα,O,Cβ,Cγ), K220(Cβ,Cγ,Cδ,Cε), L221(N,Cγ,Cδ)
D-1	69.2 ± 23.6	4-6	A103(Cα,Cβ), A104(N,O,Cβ), L106(Cγ,Cδ1), P133(Cα,Cβ,Cγ), V134(N), L221(Cδ1), N283(Cα,O,Oδ1), L284(Cα), Y286(Cβ,Cδ,Cε2), A290(Cβ)
D-2	103.0 ± 26.4	1-4	I87(Cδ1), L106(Cβ,Cδ1,Cδ2), I109(C,O,Cβ,Cγ1,Cγ2), E110(Cα,Oε1) A113(N,Cα,Cβ), I116(Cδ1), V126(Cγ1), E128(Oε2), V132(Cγ1), P133(O,Cβ,Cγ,Cδ), V134(O), A136(N,Cα,O,Cβ), S139(C,Cβ,Oγ), I140(N,Cα,Cγ1,Cγ2,Cδ) I143(Cδ1), F144(Cζ)
C-1$^{\Omega}$	106.6 ± 10.6	1-2	I87(CβCγ1Cδ1), V126(C,O,Cγ1), T127(C), E128(N,Cα,O,Cβ,Cε1), S130(N,Cα,O,Cβ), T131(Cα,C,Oγ1), V132(Cγ1), P133(Cβ,Cγ), A137(Cα,Cβ), I140(Cγ2,Cδ2), V156(Cγ1,Cγ2), L157(O), S158(Cα,Cβ,Oγ), N159(N), P160(N,Cα,Cγ,Cδ), E217(Cγ, Oε2)
D-1$^{\Omega}$	130.3± 25.6	1-5	P133(O), V134(Cα,Cγ1,Cγ2), R135(N,Cα,C,Cγ,Cδ,Cε), E138(N,Cβ,Oε1), W214(Cγ,Cε,Cε,Cζ,Cη), E217(Cβ,Oε1), L221(Cδ1,Cδ2), N283(Cα,C,O,Cβ,Cδ), L284(N,Cα,Cδ1), Y286(O,Cβ,Cδ), L287(Cα,Cβ,Cδ1,Cδ2) A290(Cβ)
D-2$^{\Omega}$	157 ± 14.4	2-3	I87(C,O,Cγ2), S88(Cα), V89(N,O,Cγ2), N90(Cα,C,O,Cβ), T91(N,C,O,Cβ,Cγ2,Oγ1), A104(Cβ), L106(Cδ2), I109(Cγ1,Oδ1), E128(Oε2), K129(O), S130(Cβ,Oγ), T131(Oγ1), V132(N,O,Cβ,Cγ2), P133(Cα), V134(N,O,Cβ,Cγ2), A136(Cβ), S275(O), K279(Cε,Nζ)
C-1*	147.4 ± 8.6	0	E128(O,Cβ,Cγ,Cδ,Oε1), K129(C,O), S130(C,O), T131(N,Cα,Oγ1,Cγ2), V132(N,O,Cγ1,Cγ2), S158(Oγ), P160(N,Cα,Cβ,Cγ,Cδ), E161(N,O,Cγ,Cδ,Oε2), S216(O), E217(Cα,O,Cγ), K220(Cβ,Cγ,Cδ,Cε), L221(Cγ,Cδ2)

D-1*	100.5 ± 10.9	0	A103(Cα,Cβ), A104(N,O,Cβ), P133(Cα,Cβ,Cγ), V134(N,O,Cγ2) R135(Cζ,Nη1,Nη1), L221(Cδ2),N283(O,Cγ,Oδ1,Nδ2), L284(Cα), Y286(Cδ2,Cε2,Oη), L287(N,Cα,O,Cβ,Cδ2), A290(Cβ), L291(Cγ)
D-2*	136.0 ± 22.1	0	T91(O,Oγ1), P92(C), T93(N,Cα,Oγ1,Cγ2), G101(O), R102(Cα,O), A103(C,O), A104(N,Cα,Cβ), V132(C,O), P133(N,Cδ), L221(Cδ2), K279(Cα,Cγ,Cε,Nζ) D280(Oδ1), L282(Cβ,Cδ1), T131(O,Cγ2), N283(Cβ,Cγ,Oδ1,Nδ2)
D-3*	195.4 ± 30.5	0	I87(O,Cβ,Cγ1,Cγ2,Cδ1), L106(C,O,Cβ,Cδ1,Cδ2), K107(N,Cα,O,Cβ), I109(O,Cγ1,Cγ2,Cδ), A113(Cβ),V126(C,O,Cγ1), T127(C), E128(N,Cα,Cβ,Cγ,Cδ,Oε1), S130(Oγ),V132(Cγ1), E110(N,Cα,Cβ,Cγ,Cδ,Oε1) A136(C,O,Cβ), S139(C,Cβ,Oγ), I140(N,Cα,Cγ1,Cγ2,Cδ1), I143(Cδ1), V156(Cγ1,Cγ2), L157(O)

a calculated with w probe radius of 1.2 Å

b the number of waters is listed as a range where there is variability per chain in the observed ordered waters.

CHAPTER 4

THE CONSERVATION OF ATYPICAL ALLOSTERY IN *C. ELEGANS* UDP-GLUCOSE DEHYDROGENASE

Reprinted here with permission from: **Beattie, N.R.**, Keul, N.D., Hicks Sirmans, T.N., McDonald, W.E., Talmadge, T.M., Taujale, R., Kannan, N., and Wood, Z.A., The Conservation of Atypical Allostery in C. Elegans UDP-Glucose Dehydrogenase. *ACS Omega*, 4(15), pp.16318-16329 (2019). https://doi.org/10.1021/acsomega.9b01565.

4.1 Abstract

Human UDP-glucose dehydrogenase (hUGDH) oxidizes UDP-glucose to UDP-glucuronic acid, an essential substrate in the phase II metabolism of drugs. The activity of hUGDH is controlled by an atypical allosteric mechanism in which the feedback inhibitor UDP-xylose competes with substrate for the active site and triggers a buried allosteric switch to produce an inactive complex (E^{Ω}). Previous comparisons with a non-allosteric UGDH identified six large-to-small substitutions that produce packing defects in the protein core and provide the conformational flexibility necessary for the allosteric transition. Here we test the hypothesis that these large-to-small substitutions form a motif that can used to identify allosteric UGDHs. *Caenorhabditis elegans* UGDH (cUGDH) conserves this motif with the exception of an Ala-to-Pro substitution in position 109. The crystal structures of unliganded and UDP-xylose bound cUGDH show that the A109P substitution is accommodated by an Asn-to-Ser substitution at position 290. Steady-state analysis and sedimentation velocity studies show that the allosteric transition is conserved in cUGDH. The enzyme also exhibits hysteresis in progress curves, and negative cooperativity with respect to NAD^+ binding. Both of these phenomena are conserved in the human enzyme, which is strong evidence that these represent fundamental features of atypical allostery in UGDH. A phylogenetic analysis of UGDH shows that the atypical allostery motif is ancient, and identifies a potential transition point in the evolution of the UGDH family.

4.2 Introduction

The nucleotide sugar uridine diphosphate-α-D-glucuronic acid (UDP-GlcA) is the essential substrate of glucuronidation, a major mechanism of drug metabolism in mammals.[80, 81, 103] Some cancers have been shown to exploit glucuronidation as a means of chemotherapeutic resistance.[6-9] Limiting the availability of UDP-GlcA represents a novel strategy for combatting this drug-resistance mechanism.[82-84] UDP-GlcA is produced by UDP-α-D-glucose 6-dehydrogenase (UGDH), which catalyzes the NAD^+-dependent oxidation of UDP-α-D-glucose (UDP-Glc). In humans, hUGDH is a hexamer that is allosterically regulated by the downstream metabolite UDP-α-D-xylose (UDP-Xyl) (Figure 1A).[47, 51] Understanding the allosteric mechanism of hUGDH is an important step towards the design of therapeutics that can reduce the cellular levels of UDP-GlcA.

In the absence of any ligand, the three dimers of hUGDH form a weakly associated hexamer called E* (Figure 4.1.A-B).[39, 44, 45, 48, 49, 105, 109, 110] The E* hexamer slowly isomerizes to the active E state upon binding substrate, which produces a lag in progress curves (hysteresis).[48] The activity of hUGDH is regulated by an atypical allosteric mechanism[44, 48, 105, 109, 110]. What makes this atypical is that the effector (UDP-Xyl) competes with substrate for the active site. Upon binding, the UDP-Xyl triggers the Thr131-loop/α6-helix (the allosteric switch) to change conformation, which allosterically increases the affinity of adjacent dimers and produces an inactive horseshoe-shaped complex (E^{Ω}) (Figure 4.1.A-B).[39, 44, 105, 109, 110] This allosteric transition can also be observed as positive cooperativity in inhibition studies.[48, 105, 109, 110]

The Thr131-loop of the allosteric switch is buried in the protein core, and must change conformation for the enzyme to transition between the E and E^{Ω} states (Figure 4.1.C). The remarkable plasticity of the protein core originates from packing defects (cavities and deep surface pockets) that provide flexibility and space for the allosteric transition (Figure 4.1.C).[39, 105, 109]

Comparing hUGDH to the non-allosteric homolog from *Streptococcus pyogenes* (spUGDH), it was shown that the packing defects evolved from six large-to-small residue substitutions in the protein core (Figure 4.1.D). It was hypothesized[39] that these substitutions could be used as a sequence motif to identify other UGDHs that exhibit atypical allostery. Here we have tested that hypothesis using *Caenorhabditis elegans* UGDH (cUGDH), which shares 61% sequence identity with hUGDH, but has a single substitution in one of the allosteric motif residues (A109P) (Figure 4.1.D and 4.2). In the human enzyme, we have previously shown that the A109L substitution in this motif abolishes allostery.[109] Here we that show that cUGDH conserves the allosteric transition. The crystal structure shows that the A109P substitution is accommodated by a second N290S mutation. In addition, we observe both hysteresis and negative cooperativity with respect to NAD^+ binding (which has also been reported in the human enzyme).[49, 105, 109, 110] We suggest that the conservation of both hysteresis and negative cooperativity is good evidence that these phenomena are defining characteristics of atypical allostery in UGDH. Prior to this work, atypical allostery had only identified in the human enzyme. By combining structural and kinetic studies with a phyologenetic analysis, we shed light on the evolution of atypical allostery in UGDH.

4.3 Results

cUGDH Forms the E^{Ω} State

The defining feature of the allosteric transition is the formation of the horseshoe-shaped complex in the presence of UDP-Xyl (Figure 4.1.B).[39] The crystal structure of cUGDH in complex with UDP-Xyl (cUGDH:UDP-Xyl) was solved and refined to a resolution of 2.45 Å (Figure 4.3.A-B and Table 4.1). The asymmetric unit contains 12 chains that are structurally equivalent, with a small number of disordered residues at the N- and C-termini (≤7 and ≤11 residues, respectively). Each chain of cUGDH also contains a well-ordered molecule of UDP-Xyl in the active site, which

conserves all of the same ligand interactions previously described for the human enzyme (Figure 4.3.C- D). In addition, there is second molecule of UDP-Xyl bound in the NAD^+ binding site that is most likely an artefact of the high UDP-Xyl concentrations (1 mM) in the crystallization buffer (not shown); at high concentrations, UDP-Xyl is known to bind promiscuously to the NAD^+ site in hUGDH.

The cUGDH structure superimposes 453 corresponding Cα atoms onto hUGDH (PDB entry 3PTZ) with an rmsd of 0.7 Å and sequence identity of 61%. The domain structure is assigned based on homology to the human enzyme: NAD^+ binding domain (NB; res. 1-219), dimerization domain (DD; res. 220-329), and sugar-binding domain (SB; res. 330-481) (Figure 4.2). The application of crystallographic symmetry shows that all twelve chains of cUGDH assemble into horseshoe-shaped hexamers, similar to the UDP-Xyl bound E^{Ω} state[39] observed in the human enzyme (Figure 4.3.B). The largest difference between the cUGDH and hUGDH E^{Ω} hexamers involves a 5.6° larger rigid body rotation of adjacent dimers, which produces a wider opening in the cUGDH 'horseshoe' complex.

UDP-Xyl binding induces formation of the E^{Ω} hexamer

The UDP-Xyl binding allosterically stabilizes the UGDH hexamer in solution. [39, 44, 45] Sedimentation velocity analysis of 9 μM unliganded cUGDH revels a *c(s)* distribution consisting of an 11.2 S (68.5%), 8.3 S (8.7%), 5.6 S (21.6%) and 3.3 S (1.2%) species (Figure 4.4.A). This is consistent with the distribution of the hUGDH unliganded E* state, which exists as a mixture of hexamers, tetramers and dimers in rapid equilibrium under identical conditions.[39, 44, 45, 105, 109, 110] The 3.3 S species likely represents a small amount of misfolded monomer, which is also observed in sedimentation velocity studies with hUGDH.[39, 44, 45, 105, 109, 110] The cUGDH hexamer (11.2 S) appears to be slightly smaller than the predicted value (11.5 S), which is likely a result of the fact

that in a rapid-equilibrium system the modeled S-values are biased by the mean of the distribution.[94, 98] The addition of 80 μM UDP-Xyl shifts the distribution to a single 11.6 S species, which is nearly identical to the expected S value (11.7 S) of the E^{Ω} state, horseshoe-shaped hexameric complex (Figure 4.4.B).

To understand how the A109P substitution was accommodated during the allosteric transition, we conducted a detailed structural analysis on a 1.88 Å resolution unliganded crystal structure of the cUGDH *32* symmetry hexamer that was deposited by the New York SGX Research Center for Structural Genomics (unpublished, PDB:2O3J) (Figure 4.4.C). The NB and SB domains of the unliganded cUGDH subunits have rotated to an open conformation (~11° and ~5°, respectively) relative to the E^{Ω} hexamer (Figure 4.4.D). This open domain conformation resembles the E* state of hUGDH, which represents an intermediate in the allosteric transition from E to E^{Ω} (Figure 4.1.B).[48, 105, 109] The Thr136-Loop/α6-helix (allosteric switch) in cUGDH was identified based on homology with the human enzyme (Figure 4.2). A comparison of the unliganded cUGDH structure with the human enzyme confirms that the allosteric switch is in the E* conformation (Figure 4.4.E). This analysis also shows that cUGDH and hUGDH undergo a similar conformational change in the transition from the E* and E^{Ω} state (Figure 4.4.E).[39, 44] The only significant difference in the Thr136-Loop portion of the allosteric switch involves the conserved residue Val139 (cUGDH numbering) (Figure 4.5.A). In cUGDH, the side chain of Val139 is weakly ordered, and is modeled in the *gauche+*, *gauche-*, and *trans* rotamers in the different peptides of the crystal structure, while in hUGDH the side chain is in the *trans* conformation (-170°). The rotational disorder of the valine appears to be in response to the bulky pyrrolidine ring of Pro109, which displaces the Val139 Cα atom by 1.4 Å relative to the human enzyme. This displacement also induces the conserved residue Leu228 to change rotamers (Figure 4.5.A). The

largest conformational difference between the enzymes involves the α6-helix; in the allosteric transition from the E* hexamer to the E^{Ω} complex the cUGDH helix rotates by 14 ± 3°, but only 6 ± 2° in the human enzyme (Figure 4.5.B). The difference in rotation originates from several changes in the packing interactions between the α6-helix and the groove formed between the α5-helix and the β6-strand. First, the Cβ atom Ala141 and the Cγ atom of Val114 are ~4.7 Å apart, which creates a small void that weakens the packing of the α6-helix in cUGDH (Figure 4.5.C). In hUGDH, Val114 is replaced with an isoleucine which fills the void with a Cδ atom and packs tightly against the alanine. Another significant change involves an alanine-to-serine substitution at position 118 in cUGDH, which is buried in a hydrophobic pocket where it weakens a hydrogen bond in the α5-helix by competing for the main chain carbonyl oxygen of Val114 (Figure 4.5.D). The most significant change in the α6-helix involves Ala152 in cUGDH, which packs into a buried hydrophobic cluster formed by Ala122, Leu149 and Phe161 (Figure 4.5.E). In the human enzyme, the hydrophobic cluster is maintained with conservative substitutions, but Ala152 is replaced by an asparagine (Asn147 in hUGDH, Figure 4.5.E). As a result, the C-terminus of the α6-helix in hUGDH has locally unfolded to expose the Asn to solvent. There is also a two-residue insertion in the loop following the α6-helix in cUGDH, but it does not appear to contribute directly to the change in helix rotation (Figure 4.2 and 4.5.E).

The hexamer-building interface was examined to determine the origin of the increased rotation of the α6-helix in cUGDH (Figure 4.5F and S4.1). The most significant differences involve the conserved residues Phe330 and Pro440, both of which pack directly against the α6-helix. In cUGDH, Phe330 adopts a χ_1 torsion angle of -69° and packs against the α6-helix residue Cys147 (Figure 4.5.F). In contrast, Cys147 is replaced with an arginine in the human enzyme, which forces the corresponding Phe into a strained, weakly ordered rotamer of -102°. The

conserved residues Glu143 and Ser144 change conformations to produce a complementary packing surface for Phe330 in each enzyme. The conserved residue Pro440 is also displaced by ~1.9 Å relative to the human enzyme (Figure 4.5.F). This shift is a consequence of the altered α6-helix conformation and the replacement of alanine by Glu151 in cUGDH.

The Atypical Allostery Motif in cUGDH

With the exception of the alanine-to-proline substitution at position 109, hUGDH and cUGDH conserve the atypical allostery motif and most of the associated packing defects in the E^{Ω} hexamer (cavity C-1^{Ω} and deep surface pockets D-1^{Ω} and D-2^{Ω}) (Figure 4.2 and 4.6.A-C). In hUGDH, the main-chain amide of the alanine forms a hydrogen bond with the carboxamide of an asparagine (Asn283, human numbering) (Figure 4.6.D). The proline substitution in cUGDH removes the amide, which would break the hydrogen bond (Figure 4.6.E). As a result, the selective pressure on the asparagine is relaxed, and it has been replaced with Ser290 which forms a hydrogen bond with an ordered water molecule in the deep surface pocket D-2^{Ω}. The pyrrolidine ring of Pro109 also alters the structure of deep surface pocket D-1^{Ω} by introducing a steric clash that induces the rotameric changes in Val139 and Leu228 that were described in the previous section (Figure 4.5.A). The only other change in D-1^{Ω} involves Val114, which replaces an isoleucine in the human enzyme.

cUGDH Conserves Hysteresis and Cooperativity

The atypical allosteric hUGDH undergoes a slow, substrate-induced isomerization from the inactive E* to the active E state, which can be observed as a lag in progress curves.[105] The cUGDH progress curves also display a similar hysteresis, with a lag that is ~30% shorter than that observed in hUGDH (Figure 4.7.A-B). The kinetic parameters for cUGDH were determined using the steady-state velocities derived from Frieden's equation for enzyme hysteresis.[52] The NAD^{+}

saturation curve exhibits negative cooperativity (Hill Coefficient = 0.82), with a $K_{0.5}$ and k_{cat} of 333 μM and 1.7 s^{-1}, respectively (Figure 4.7.C and Table 4.2). This is comparable to the NAD^+ negative cooperativity in saturation curves that has previously been reported[45, 48, 49] for the human enzyme (Figure 4.7.D and Table 4.2). In contrast, the UDP-Glc saturation curve is hyperbolic, with a K_M and k_{cat} of 55 μM and 1.2 s^{-1}, respectively (Figure 4.7.E and Table 4.2). The lower k_{cat} is a result of the negative cooperativity associated with NAD^+ binding, which makes it difficult to saturate cUGDH with co-enzyme in order to achieve pseudo-first order conditions.[49, 105, 109]

Substrate cooperativity in the presence of saturating UDP-Xyl has been observed in hUGDH and is strong evidence of the allosteric transition from a low UDP-Glc affinity E^{Ω} complex to a higher UDP-Glc affinity E hexamer. [48, 105, 109, 110] Steady-state analysis of UDP-Xyl inhibition yields a K_i of 6.9 μM, which corresponds to a ~16-fold weaker affinity for the inhibitor than that seen in hUGDH (Figure 4.7.E-F and Table 4.3). The cUGDH UDP-Glc saturation curves also exhibit positive cooperativity in the presence of 15 μM UDP-Xyl (Hill coefficient of 1.4) (Figure 4.7.E and Table 4.3).

Evolutionary Analysis of Atypical Allostery in UGDH

The evolution of atypical allostery in UGDH was examined using over 70,000 sequences. A phylogenetic tree of 31 representative UGDH sequences from diverse eukaryotic and prokaryotic phyla is shown in Figure 8. Based on overall sequence similarity and taxonomic representation, these 31 sequences can be broadly classified into 8 clades. Since the dendrogram shown in Figure 8 is a gene tree and not a species tree, taxonomically diverse species such as bacteria and protists often fall in different clades.[111, 112] Here, our definition of clades is based on the overall similarity of full length UGDH sequences. The nature of amino acids present in the atypical allostery motif (including the newly defined S/N290 position) was compared across phyla (Figure

4.8). While both the atypical allosteric motif residues and the packing defect residues are generally conserved across the eukaryotic and prokaryotic phyla (clades 1-7), they are strikingly different in the non-allosteric *S. pyogenes* clade (clade 8), which was used as an outgroup (see methods) (Figures 4.8 and 4.9). If a clade contains a single, conservative variation, then it is identified as likely allosteric, but if there are significant changes, then it is assumed to be non-allosteric (Figure 4.9). Sequences in clades 1-2 strictly conserve the allosteric motif residues, while notable variations are observed in clades 3-7. In clade 3, the Capsaspora phylum absolutely conserves the motif, while the Platyhelminthes substitute a serine for the alanine at position 141. The impact of the serine is likely to be negligible since the Oγ would point into the deep surface pocket D-2^{Ω}, which is large enough to accomadate a water molecule for hydrogen bonding. In clades 5-8, we observe the persistent conservation of a Thr substitution in position 141. This is most pronounced in the Stramenopiles, Bacillariophyta and Phaeophyceae phyla of clade 5, where the Thr141 substitution represents the only significant variation in either the atypical motif or packing defect residues. In contrast, the remaining phyla of clade 5 also contain a tryptophan substituion in the packing defect residue position 114, which will likely introduce significant clashes in the inhibited state of the enzyme. In clade 6, the protist Parasabalia clusters with bacterial phyla Lentisphaerae, Cyanobacteria and Verrucomicrobia. The grouping of protists and bacteria in gene dendrograms has previously been observed in other gene families and is often attributed to lateral gene transfer between protists and bacterial.[113]The Ascomycota (clade 7) contain a leucine substitution at position 109, which is interesting because the equivalent substitution in the human enzyme has been shown to eliminate allostery by introducing steric clashes in the deep surface pocket D-2^{Ω}. In general, the overall conservation of the motif residues in fungal and cyanobacterial clades (clades 6-7) suggests that the atypical allosteric mechanism is likely ancient and evolved

progressively in higher eukaryotes (clades 1-2) (Figure 4.9). Because of the specificity for UDP-Xyl, it is unlikely that the atypical allosteric inhibition mechanism would be present in the UGDHs of organisms that do not produce this nucleotide-sugar. Unfortunately, data concerning the distribution of UDP-Xyl throughout various phyla is poorly characterized and incomplete.

4.4 Discussion

What makes allostery in UGDH atypical is that the feedback inhibitor competes with substrate for the active site to trigger a transition to the E^{Ω} state (Figure 4.1.B).[39, 48] Key to this remarkable transition is the atypical allostery motif, which identifies a series of large-to-small amino acid substitutions that produce the packing defects in the protein core that are necessary to accommodate both the E and E^{Ω} conformations of the buried allosteric switch (Figures 4.4.E and 4.6.A-C).[39, 109] This motif is conserved in cUGDH with the exception of an alanine to proline substitution in position 109 (Figure 4.2). In an early study with the human enzyme, it was shown that a leucine substitution at the equivalent position abolished allostery.[109] Here, we show that despite the A109P substitution, the crystal structure of cUGDH:UDP-Xyl still forms the allosterically inhibited E^{Ω} hexamer (Figure 4.3.B). The structure also suggests a reason for why Ser290 is replaced with an asparagine in hUGDH (Figures 4.6.D-E). There, the asparagine forms a hydrogen bond with the main-chain amide of the alanine in position 109 of the atypical motif (Ala104 in hUGDH) (Figures 4.6.D-E). The Pro109 substitution in cUGDH removes the amide, which relaxes the selective pressure on the asparagine and allows the Ser290 substitution.

A comparison of the E* and E^{Ω} cUGDH structures shows that the allosteric switch undergoes a conformational change similar to that observed in the human enzyme, but with a larger rotation of α6-helix (Figure 4.4.E and 4.5.B). It has been shown that changes in the conformation of the α6 helix can affect the stability of the hUGDH hexamer and the affinity for UDP-Xyl.[39, 105,

[109] Here, we have shown that the E* state of cUGDH is a weakly associated hexamer, while the allosterically inhibited enzyme forms a stable E^{Ω} complex, similar to human enzyme (Figure 4.4.A-B).[39, 110] However, the UDP-Xyl affinity of cUGDH is ~16-fold weaker than that of the human enzyme, despite the fact that both enzymes conserve the same interactions in the binding site (Table 4.3 and Figure 4.3.D). Since UDP-Xyl is coupled to the hexamer-building interface through the allosteric switch, it is likely that the changes in the rotation of the α6-helix are responsible for the reduced affinity (Figure 4.4.B).

Hysteresis and allostery have been shown to be coupled in hUGDH.[39, 48, 105, 109] Our observation of hysteresis in cUGDH suggests that this is a defining characteristic of the atypical allosteric mechanism (Figure 4.7.A-B). The allosteric transition can observed as positive cooperativity in steady-state inhibition studies (Figure 4.7.E-F, Table 4.3). Briefly, the E^{Ω} conformation has low affinity for the substrate UDP-Glc (Figure 4.1.D).[39, 48] As UDP-Glc outcompetes the inhibitor, the enzyme allosterically switches to the high substrate affinity E state, resulting in a sigmoidal substrate saturation curve. Both hUGDH and cUGDH also display negative cooperativity with respect to NAD^+ binding (Figure 4.7.C-D and Table 4.2). The negative cooperativity in hUGDH is poorly understood at a structural level, but it is known to originate from a substrate-induced asymmetry in the enzyme. Negative cooperativity is believed to be linked to the allosteric mechanism, because amino acid substitutions that disrupt allostery also abolish negative cooperativity.[105, 109] The conservation of negative cooperativity in cUGDH strongly suggests that this phenomenon is also a defining characteristic of the atypical allosteric mechanism.

To obtain new insights into the evolution of allostery in UGDH, we have examined the patterns of conservation and variation in the atypical allostery motif residues across diverse species

(Figures 4.8 and 4.9). Phyla that conserve Ala109 also conserve the Asn290, supporting our hypothesis that these residues may have co-evolved. Because the dendrogram shown in Figure 4.8 is a gene tree and not a species tree, taxonomically diverse species such as bacteria and protists often fall in different clades, which are based only on the overall similarity of full length UGDH sequences.[111, 112] UGDHs in higher eukaryotes (clades 1-3) generally conserve the motif residues and are likely to be allosteric, while clade 8, which is predominantly bacterial, is unlikely to be nonallosteric as they display striking divergence in motif residues (Figures 4.8 and 4.9). Specifically, phyla in clade 8 show no conservation of the amino acids corresponding to the motif residues A108, A141, A142 and N290. Notably, some bacterial sequences (Cyanobacteria, Lentisphaerae and Verrucomicrobia) fall in a different clade (clade 6) and are more similar to eukaryotic UGDHs (49-60% similarity to sequences in clades 1-3) than they are to other bacterial sequences (23-37% sequence similarity to sequences in clade 8). Interestingly, clade 6 bacterial sequences also share some of the allosteric motif residues (A108, A142, N290) suggesting an ancient origin for the proposed allosteric mechanism. We also note the diversity of protist UGDH sequences as they also fall into distinct clades in the phylogenetic tree. For example, Capsaspora, a protist most closely related to animals, clusters with UGDH sequences in clades 1 and 2 and conserve nearly all of the allosteric motif residues, whereas other ciliated protozoa such as Intermacronucleata lack the allosteric motif residues and groups with bacterial sequences in clade 8. It is also interesting that some protists such as Parabasalia group with bacterial sequences in clade 6 and share some of the allosteric motif residues (A108, N290). Grouping of protists and bacteria in gene dendrograms has previously been observed in other gene families and is often attributed to lateral gene transfer between protists and bacterial..[113] While we have focused our analysis on the allosteric motif residues, it is likely that variations beyond the allosteric motifs

contribute to the placement of protist and bacterial sequences in different clades. These variations, along with those in the remaining clades (clades 4-7) will require additional analysis to fully map the allosteric diversity of the UGDH super-family. In particular, we are intrigued by the selective conservation of threonine at position 141 in clades 4-7 (Figures 4.9). If that substitution prevents the formation of the E^{Ω} complex, then it may be that we have identified a key step in the evolution atypical allostery in UGDH. Future work will focus on mapping these evolutionary steps through the experimental characterization of UGDHs from the Stramenopiles, Bacillariophyta and Phaeophyceae phyla in clade 5. Because these sequences contain the Thr141 substitution and conserve most of the packing defect residues, they appear to be important transition points between non-allosteric bacterial UGDH and allosteric human UGDH.

4.5 Materials and Methods

Protein expression and purification

The enzymes cUGDH (UniProt ID: Q19905-1) and hUGDH (UniProt ID: O60701-1) were recombinantly expressed and purified as previously described.[39, 44, 45, 48, 49, 105, 114, 115] Briefly, cUGDH was cloned into a modified pET-15b vectors (Norclone) with N-terminal hexahistidine affinity tags adjacent to a tobacco etch virus (TEV) cleavage site. Recombinant proteins were purified using a Talon® immobilized metal affinity column, and the affinity tags were removed with 5 μM TEV protease for ~16 hrs at 20 °C. Proteins were dialyzed into a storage buffer (25 mM TRIS pH 8.0 and 50 mM NaCl) and then concentrated to 20 mg/mL using an Ultra-15 10K centrifugal filter. Proteins were quantified from dilution replicates ($N \geq 6$) using the molar extinction coefficients of 53806 M^{-1} cm^{-1} and 49850 M^{-1} cm^{-1} for cUGDH and hUGDH, respectively (calculated with the program PROTPARAM).[87] Proteins were aliquoted and flash frozen in liquid nitrogen, and then stored at -80 °C.

Protein Crystallization and Structure Determination

The cUGDH protein with bound UDP-Xyl was crystallized at 20 °C using the hanging drop vapor diffusion method, with a 2 μL drop that consisted of a 1:1 mixture of protein (10 mg/mL) and reservoir solution. Optimized crystals were obtained using a reservoir solution that consisted of citric acid buffer pH 5.0, 200 mM LiCl, 1 mM UDP-Xyl and 4% PEG 8000. These crystals were soaked in a cryoprotectant mixture composed of the same reservoir solution supplemented with 15% *v*/*v* of cryoprotectant solution (1:1:1 ethylene glycol:DMSO:glycerol ratio) and then rapidly plunged into liquid nitrogen.

Diffraction data were collected on the 22-ID beamline (SER-CAT) at Argonne National Laboratory (Argonne, IL) using a MAR 300 mm CCD detector. The 2.45 Å data set was processed with XDS in the space group $P22_12_1$, setting aside 5% of the data for cross-validation.[88, 89] The structure was solved by molecular replacement using PHENIX[90] and a single dimer of the unliganded cUGDH crystal structure (PDB: 2O3J; unpublished) as a search model. During refinement, the structure was subjected to iterative cycles of manual model rebuilding using COOT and the automated refinement procedure that is implemented within the PHENIX software suite.[90-92] The R_{free} test set was corrupted during refinement, so a posterior R-free was calculated by selecting 7% of the reflections for a new cross-validation data set and subjecting the final model to Cartesian dynamics at 4000º to decouple the test set.[89] Data collection and refinement statistics can be found in Table 4.1.

Structural Analyses

The program DynDom was used to identify the hinge bending axes for the domain rotation between the open and closed states and the rigid body rotation axis that relates the *32* symmetry hexamer to the horseshoe-shaped conformation.[96,116] PISA[117] and DIMPLOT[118] were used to

identify interacting residues at the hexamer building interfaces in the cUGDH and hUGDH structures. The α6-helix rotation angle was calculated using PyMOL (https://pymolwiki.org/index.php/AngleBetweenHelices). Briefly, the monomers from the unliganded and UDP-Xyl bound structures were superimposed via least-squares-refinement using the Cα residues 1-132, and the rotation of the α6-helix was measured by using the Cα residues 141-151. Both PyMOL[119] and UCSF Chimera 1.13[120] were used to generate figures.

Sedimentation Velocity

Sedimentation velocity experiments were conducted as previously described.[39, 44, 45, 48, 105, 114, 115] Briefly, 9 μM cUGDH was dialyzed >12 hr at 4 °C into a buffer with or without 80 μM of UDP-Xyl containing 25 mM HEPES pH 7.5 and 150 mM KCl. Samples were loaded into 12 mm double-sector Epon centerpieces equipped with quartz window cells. Cells were then loaded into an An60 Ti rotor and equilibrated to 20 °C for 1.5 hrs. Sedimentation velocity data were collected using an Optima XLA analytical ultracentrifuge at 50,000 rpm for 8-10 hrs. Absorbance data were recorded at 280 nm in radial step sizes of 0.003 cm. SEDNTERP was used to estimate the partial specific volume of cUGDH (0.73895 mL/g) and the buffer density (1.00726 g/mL) and viscosity (0.01018 P).[93] Data were modelled as a continuous sedimentation coefficient (*c(s)*) distribution using SEDFIT.[97] The baseline, meniscus, frictional coefficient, and systematic time-invariant, and radial invariant noise were fit.[94] Theoretical *s*-values were calculated using HYDROPRO[121] and the available crystal structures (PDB: 2O3J and 6OM8).

Steady-State and Transient-State Kinetics

All steady-state kinetic assays were performed as previously described.[39, 44, 45, 48, 49, 105, 114, 115] Briefly, assays contained 100 nM enzyme and either saturating amounts of UDP-Glc or NAD^+ in the reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA). Solutions

containing enzyme and substrate/cofactor were incubated separately at 25 °C for 5 mins, and then the reaction was initiated by rapidly mixing the two solutions. The progress of the reaction monitored NADH production ($\varepsilon_{340\,nm}$ = 6220 $M^{-1}cm^{-1}$) on an Agilent 8453 UV/vis spectrometer at 25 °C, with absorbance readings collected every 0.5 s. Because the progress curves for both cUGDH and hUGDH display hysteresis, the initial velocity (v_i) represents a transient that violates the steady-state approximation.[48, 105, 114, 115] Thus, the initial steady-state velocities (v_{ss}) were derived from fitting progress curves to Frieden's equation for hysteresis before the depletion of 10% substrate:[48, 52, 105, 114, 115]:

$$P(t) = v_{ss}t - \tau(v_{ss} - v_i)(1 - e^{\frac{-t}{\tau}}) \qquad (1)$$

Where P is the product produced at time t, and τ is the relaxation time of the lag. The length of the lag is equal to $e\tau$. All data were fit using nonlinear regression analysis in PRISM (GraphPad Software). UDP-Glucose substrate saturation curves were fit to equation 2:

$$v_o = \frac{k_{cat}[E_t][S]}{K_M + [S]} \qquad (2)$$

Where v_o is the initial steady-state velocity (v_{ss} in equation 1), [E_t] is the enzyme concentration (100 nM), and [S] is the variable substrate concentration. Both enzymes display negative cooperativity in NAD^+ saturation curves and were fit to the sigmoidal rate equation 3:

$$v_o = \frac{k_{cat}[E_t][S]^h}{(K_{0.5})^h + [S]^h} \qquad (3)$$

Where $K_{0.5}$ is the half saturation point and h is the Hill coefficient. The affinity (K_i) for UDP-Xyl was determined by globally fitting substrate saturation curves with and without inhibitor to a competitive inhibition model with cooperativity (equation 4) using PRISM:[48, 105, 114]

$$v_\circ = \frac{k_{cat}[E_t][S]^h}{(K_M^{app})^h + [S]^h} \text{ where } K_M^{app} = K_M\left(1 + \frac{[I]}{K_i}\right) \qquad (4)$$

The hysteresis (lag in activity) in progress curves was measured at 25 °C using an Applied Photophysics SX20 stopped-flow spectrophotometer as previously described.[114] Briefly, 500 nM cUGDH or hUGDH in the reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA) was rapidly mixed with an equal volume of the same buffer supplemented with saturating amounts of both substrate and cofactor. Progress curves were monitored by NADH production (ε_{340} = 6220 M^{-1} cm^{-1}) and fit to equation 1.

Phylogenetic Analysis

Human UGDH orthologs were identified by running BLAST on the NCBI non-redundant database. A subset of 31 sequences from a total of over 70,000 UDGH related sequences were selected that represented the overall taxonomic diversity of UGDH sequences from bacteria to humans. The selected sequences were aligned using the MAPGAPS scheme.[122] The alignment and taxonomic information for the 31 sequences is provided in supplementary material (Supplementary Table 3.1). The aligned sequences were used for phylogenetic tree generation using IQTree v1.6.1 with the following options: -m MFP -bb 10000.[123] Bootstrap support values were calculated using UltraFast bootstrap approximation (UFBoot) implemented in IQ-Tree. Independent trees generated using FastTree showed topologies similar to IQTree.[124]

4.6 Acknowledgements

We thank Dr. Renuka Kadirvelraj for advice in the refinement of the cUGDH crystal structure

4.7 Figures

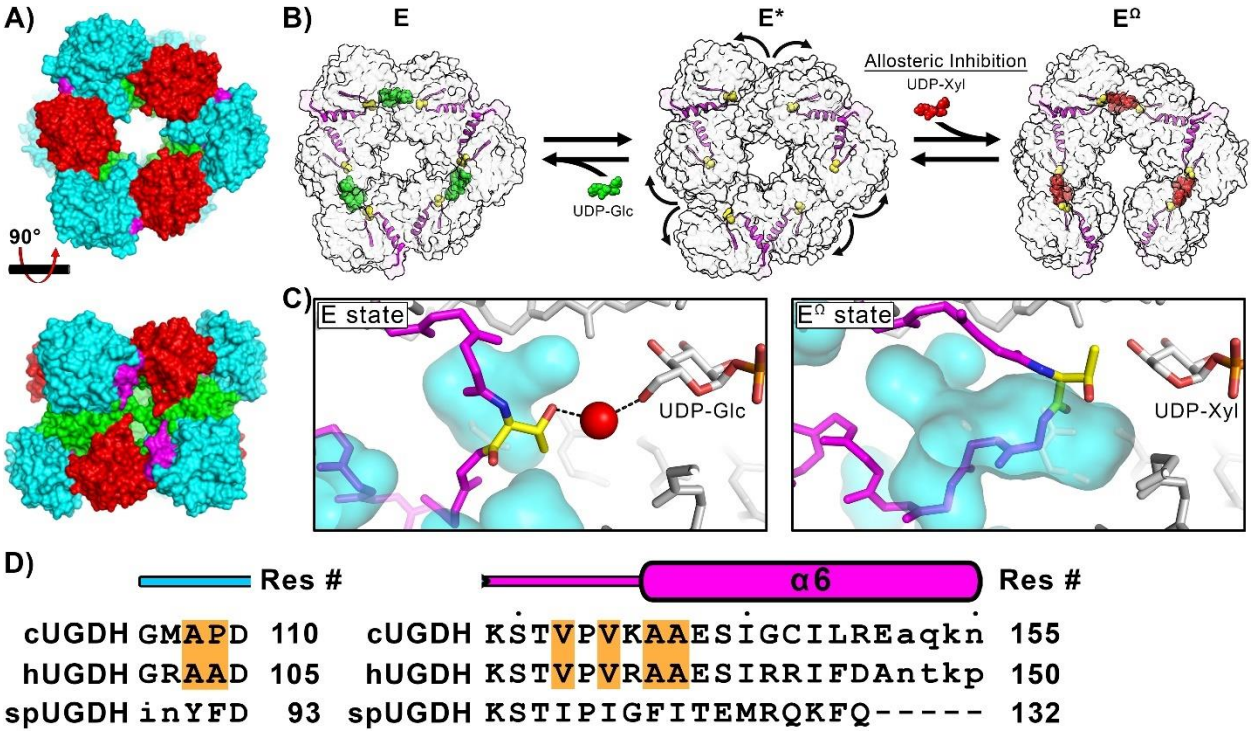
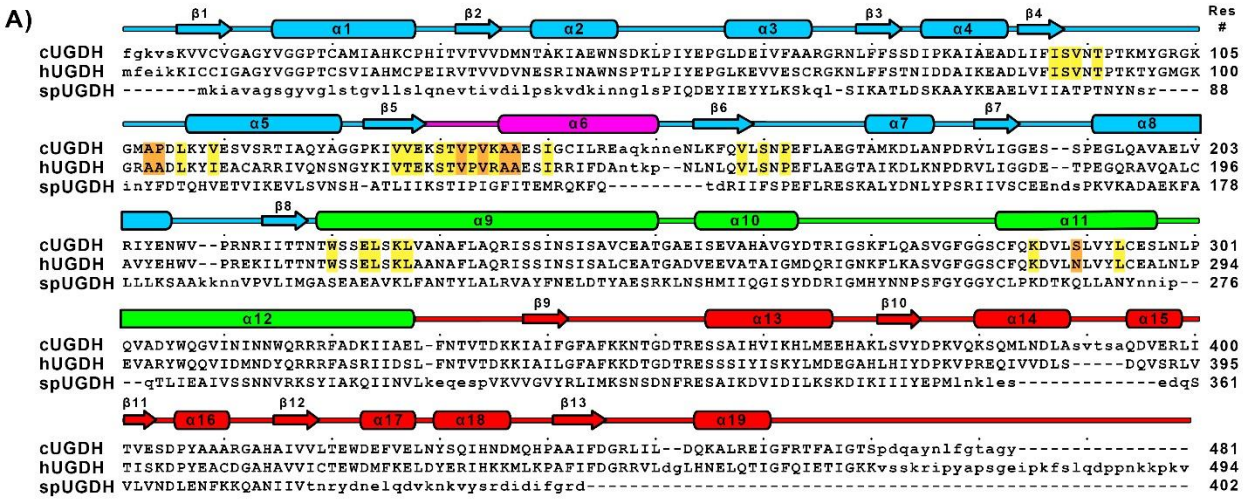


Figure 4.1: The Allosteric Transition in hUGDH. (A) Hexameric hUGDH forms from the association of three dimers. The allosteric switch (magenta), NAD^+ binding (NB) domain (cyan), dimerization domain (green), and nucleotide-sugar binding (SB) domain (red) are identified. **(B)** In the inactive E* state, the NB and SB domains adopt an 'open' conformation (curved arrows). The binding of substrate (green spheres) induces the formation of the active, E state, while UDP-Xyl (red spheres) produces the inactive, horseshoe-shaped E^{Ω} complex. **(C)** The E and E^{Ω} conformations of the allosteric switch (magenta). The packing defects in the protein core are colored cyan. In the E state, a water molecule (red sphere) forms hydrogen bonds with the C5'CH_2OH of UDP-Glc and Thr131 of the allosteric switch. UDP-Xyl lacks the C5'CH_2OH, which triggers the allosteric switch to produce the E^{Ω} state. **(D)** The atypical allostery motif residues (orange boxes) hypothesized to facilitate allostery in UGDH.[39]



B) **C-1$^{\Omega}$: I92, V131, V132, E133, S135, T136, V137, P138, A142, I145, V163, S165, P167, E224, K227, L228**
D-1$^{\Omega}$: A108, P109, P138, V139, K140, W221, E224, L225, L228, S290, L294
D-2$^{\Omega}$: I92, S93, V94, T96, A108, P109, L111, V114, E133, S135, T136, V137, P138, V139, A141, K286, S290

Figure 4.2: Structure-Based Sequence Alignment. (A) The structure-based sequence alignment of cUGDH (residues 5-481), hUGDH (residues 1-494), and *S. pyogenes* UGDH **(**spUGDH; residues 1-402) constructed by superimposing the corresponding crystal structures (PDB entries 2O3J, 2Q3E and 1DLI, respectively) using Sequoia[125]. Uppercase letters identify structurally equivalent residues. The structural elements are colored according domain: NB domain (cyan), dimerization domain (green), and SB domain (red). The allosteric switch (magenta), atypical allostery motif (orange boxes), and packing defect residues (yellow boxes) are depicted. **(B)** The packing defect residues in cUGDH that form cavity C-1$^{\Omega}$ and the deep surface pockets D-1$^{\Omega}$ and D-2$^{\Omega}$, with the atypical allostery motif residues highlighted in orange.

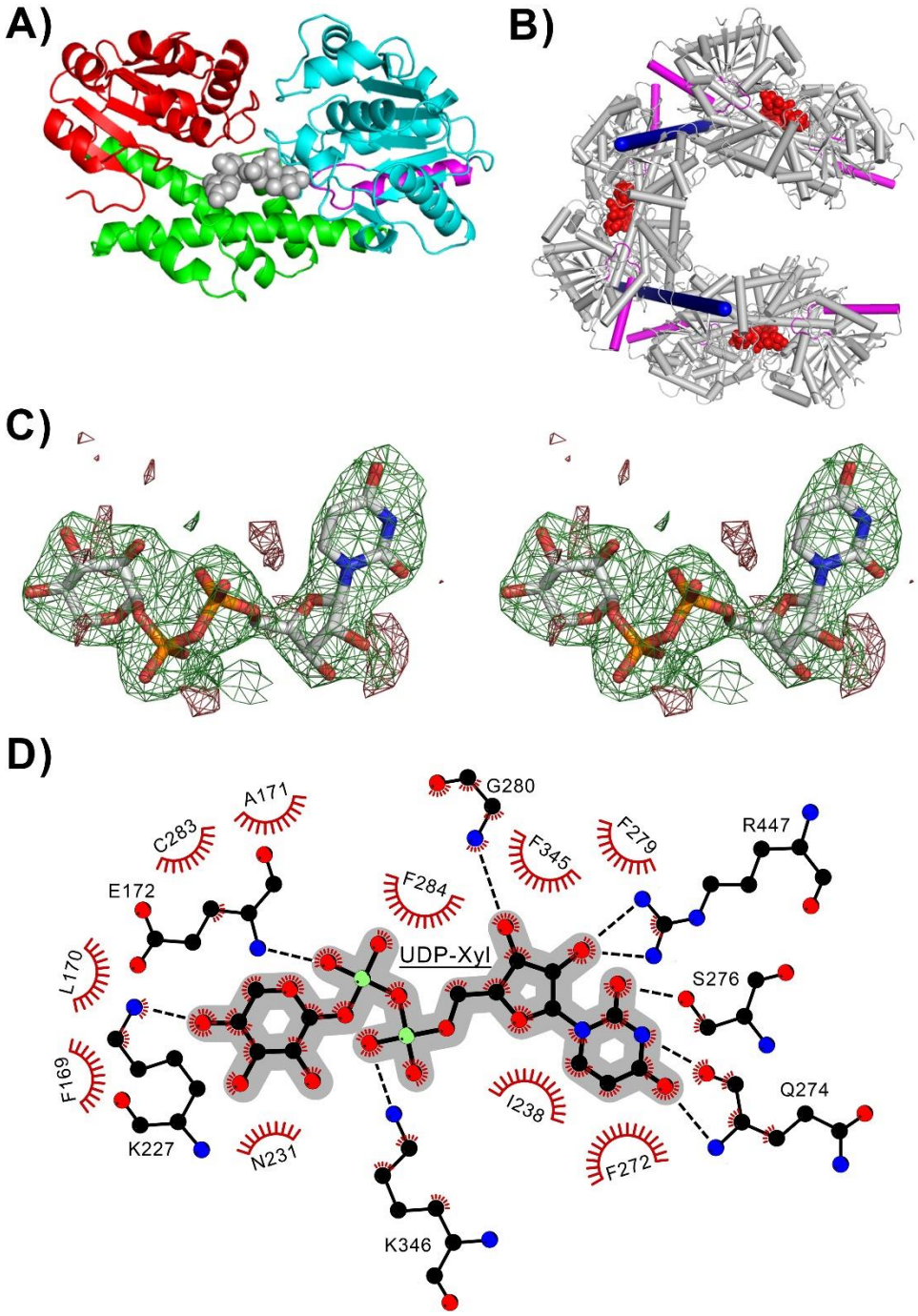


Figure 4.3: The Inhibited E^{Ω} State of cUGDH. (A) The NB domain (cyan), dimerization domain (green), SB domain (red) and allosteric switch (magenta) are shown in single chain of cUGDH in complex with UDP-Xyl (grey spheres) **(B)** The cUGDH horseshoe-shaped hexameric complex (E^{Ω}), depicted with rigid body rotation axes (blue rods) in the hexamer interfaces. The active sites with UDP-Xyl (red spheres) and allosteric switches (magenta) are also identified. **(C)** A stereodiagram of the F_o - F_c density map for UDP-Xyl contoured at 3σ and calculated by omitting UDP-Xyl subjecting the model to simulated annealing. **(D)** LigPlot[118] displaying hydrogen bonds (dashed lines) and van der Waals (red feathers) interactions with UDP-Xyl (grey highlighting).

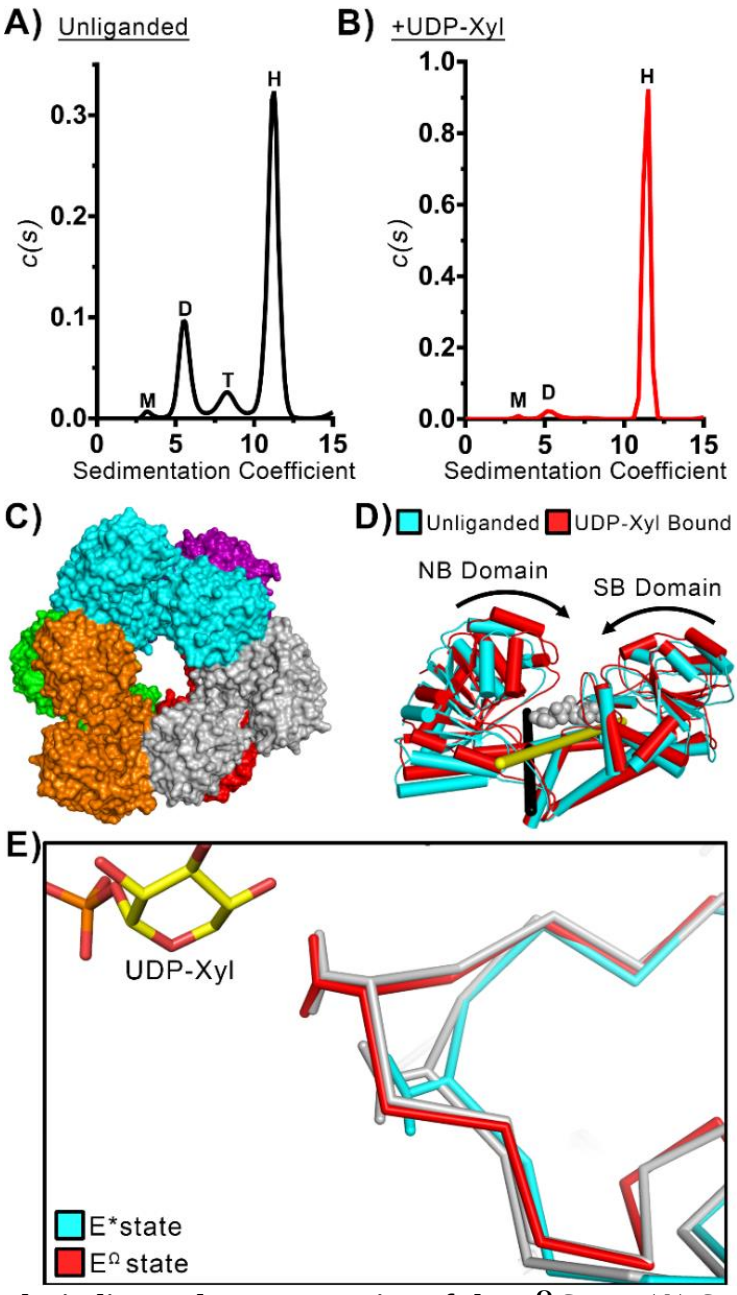


Figure 4.4: UDP-Xyl Binding Induces Formation of the E^{Ω} State. (A) Sedimentation velocity analysis of unliganded cUGDH reveals a distribution corresponding to a (M)onomer, (D)imer, (T)etramer, and (H)examer species. **(B)** UDP-Xyl shifts the distribution to a single 11.6 S species consistent with the E^{Ω} hexamer. **(C)** Unliganded cUGDH structure forms a 32 symmetry hexamer (PDB: 2O3J; unpublished). **(D)** A superposition of an unliganded (cyan) and UDP-Xyl bound (red) monomer of cUGDH illustrates the rotation of the NB and SB domains. DynDom[96] detects hinge-bending axes in the NB (black rod) and SB (yellow rod) domains. The SB hinge-bending axis is only detected in 6 of the 12 chains. **(E)** The cUGDH allosteric switch (Cα trace with Thr136 in sticks) in the E* (cyan) and E^{Ω} (red) states are similar to the equivalent structural states of hUGDH (grey).

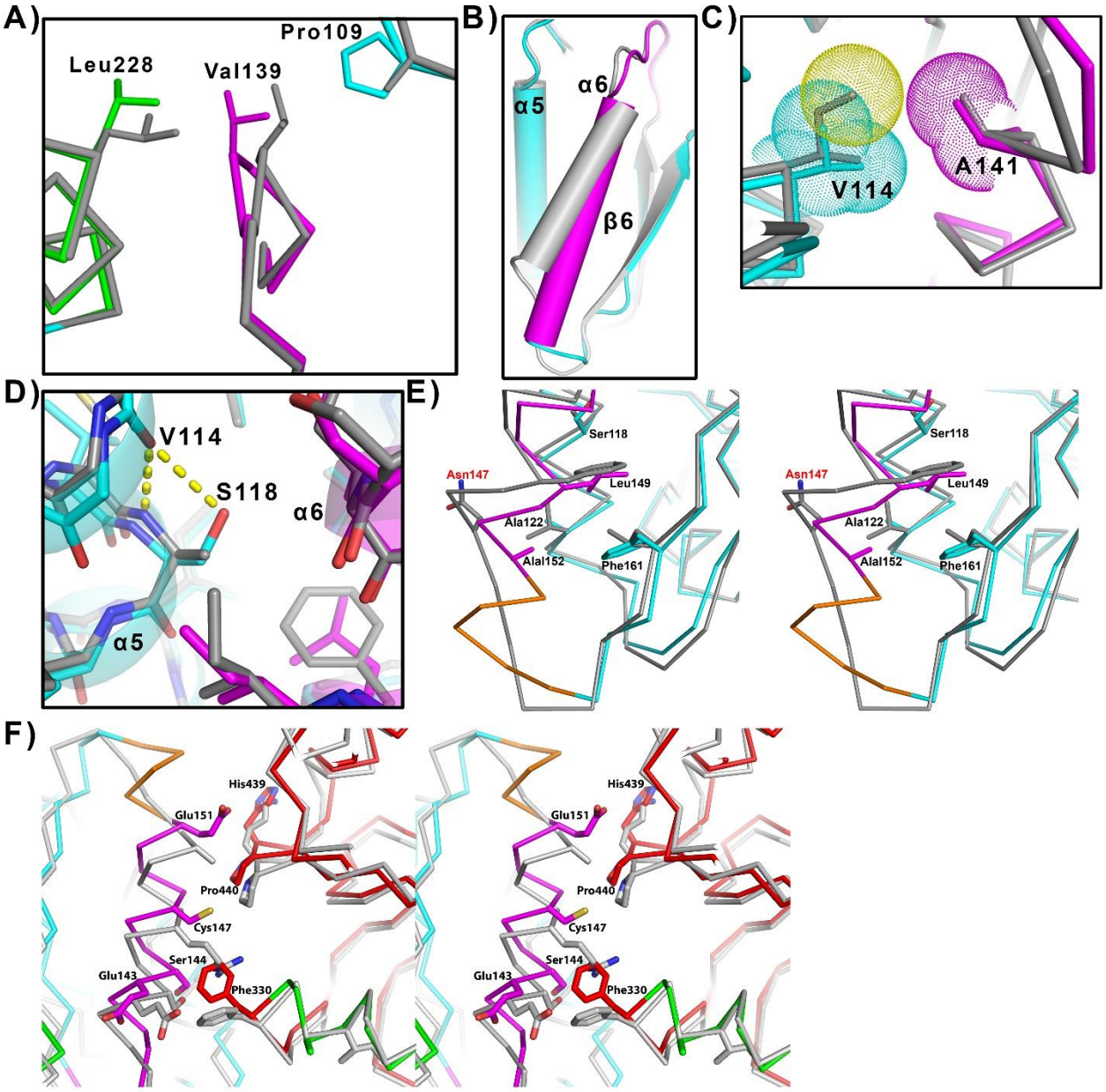


Figure 4.5: Structural Divergence in the cUGDH Allosteric Switch. (A) The Pro109 substitution in cUGDH displaces the Val139 (magenta) which causes Leu228 to change rotamer relative to hUGDH (grey). **(B)** The α6-helix in the E^{Ω} conformation of cUGDH (magenta, cyan) is rotated more than that observed in hUGDH. **(C)** Val114 (cyan) replaces an isoleucine in hUGDH (grey). VDWs distances (dots) and the α6-helix (magenta) are identified. **(D)** Ser118 of cUGDH (cyan) replaces an alanine in hUGDH (grey). **(E)** Stereoview depicting significant changes in the packing of the α6-helix of cUGDH (magenta) against the α5-helix and β6-strand (cyan) compared to the hUGDH (grey). cUGDH contains an insertion that lengthens the α6-helix (orange). **(F)** Stereoview of changes in the packing interactions that form the hexamer building interface of cUGDH (colored by domain; NB-cyan, allosteric switch-magenta, dimerization domain-green, and SB-red). hUGDH is colored grey.

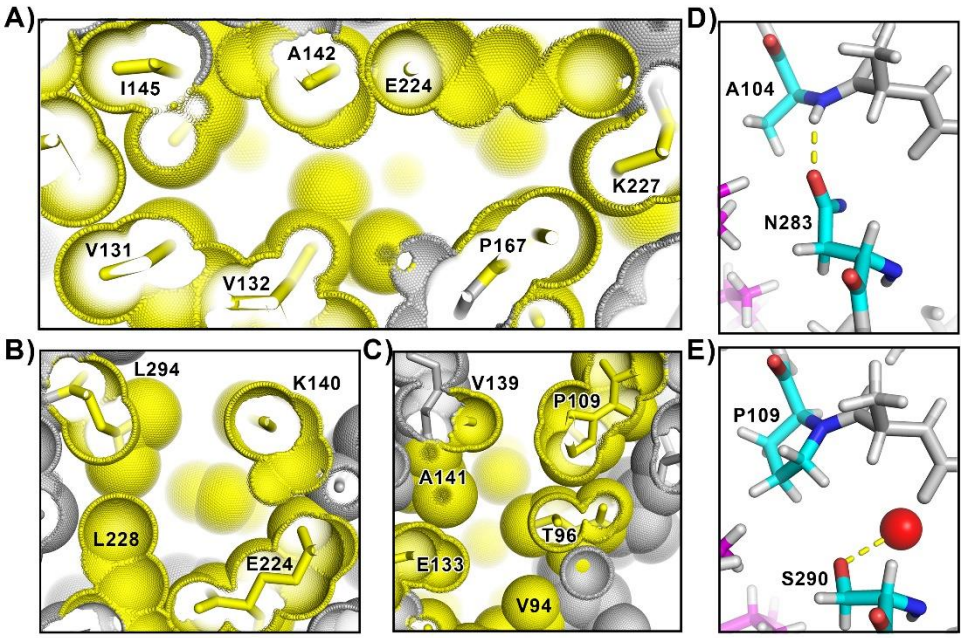


Figure 4.6: The Packing Defect Residues (PD Residues) are Conserved in cUGDH. (A) A cutaway of the cUGDH E^{Ω} structure protein core depicting cavity C-1^{Ω}. **(B, C)** Same as in (A), depicting deep surface pockets D-1^{Ω} and D-2^{Ω}, respectively. **(D)** In hUGDH, Asn283 in D-2^{Ω} forms a hydrogen bond with the backbone amide of Ala104 (hUGDH numbering). **(E)** In cUGDH, the alanine and asparagine are replaced with a Pro109 and Ser290, respectively. Ser290 forms a hydrogen bond with a water molecule (red sphere).

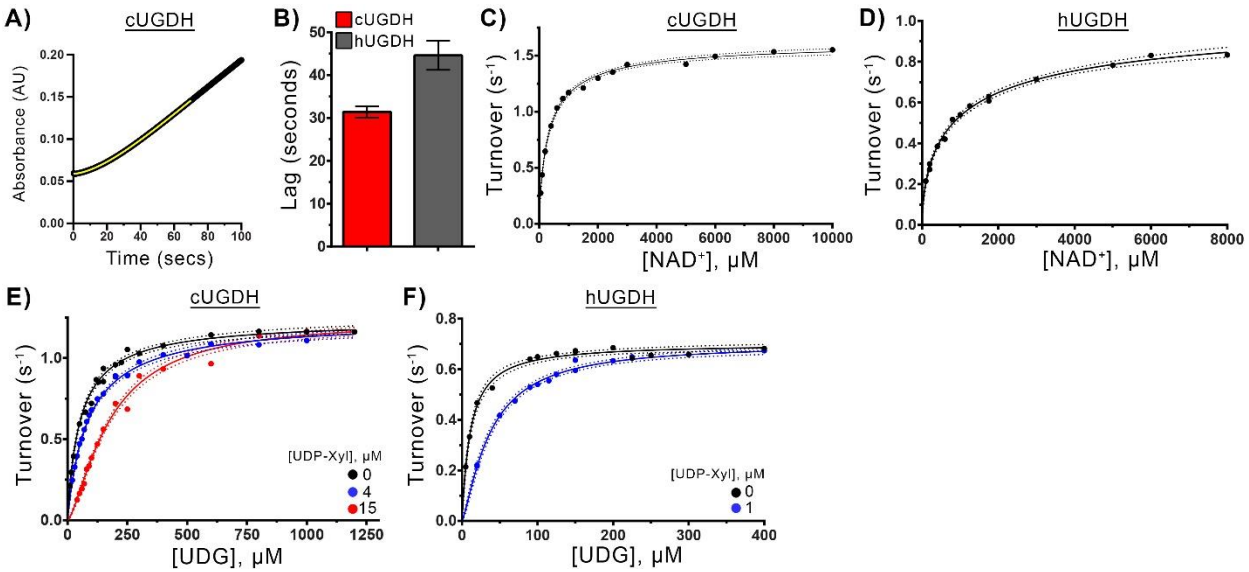


Figure 4.7: Kinetic Analyses of cUGDH and hUGDH. (A) Representative cUGDH progress curve (black line) from stopped-flow absorbance spectroscopy, which is fit to equation 1 (yellow line) (eq 1). **(B)** The length of the progress curve lag under the same conditions for cUGDH and hUGDH. **(C)** NAD^+ saturation curve with cUGDH fit to eq 3 (thin line). Rates (black dots) have been normalized by enzyme concentration to give turnover, and dashed lines represent 95% confidence intervals. **(D)** NAD^+ saturation curve with hUGDH. **(E)** cUGDH substrate saturation curves with 0 μM (black), 4 μM (blue), and 15 μM UDP-Xyl (red), which were globally fit to eq. 4. **(F)** hUGDH substrate saturation curves with 0 μM (black) and 1 μM (blue) UDP-Xyl, which were globally fit to eq. 4. Kinetic parameters derived from all fits can be found in Tables 3.2 and 3.3.

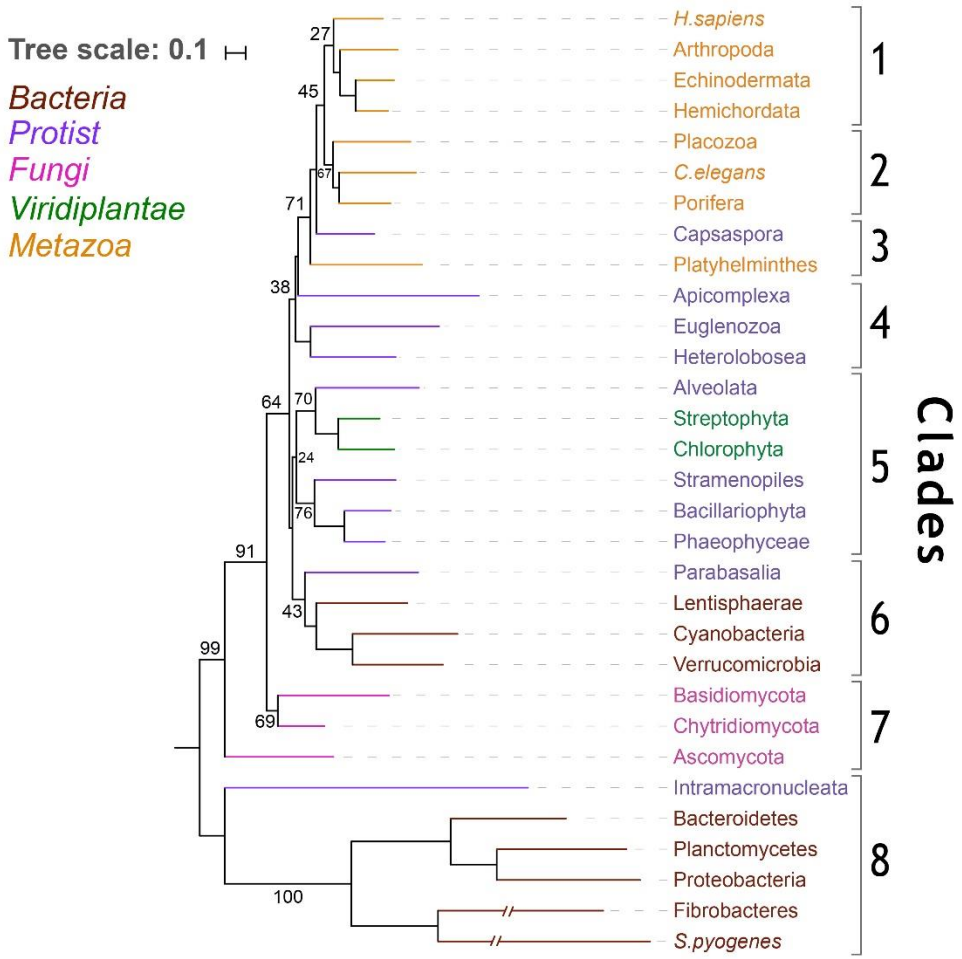


Figure 4.8: Phylogenetic Analysis of UGDH. Using taxonomically diverse UGDH sequences, the evolution of UGDH was followed. This analysis distinguishes eight distinct clades of UGDH. The long branch lengths of the non-allosteric outgroup (clade 8) indicates that these phyla contain highly divergent UGDHs. Bootstrap values are indicated.

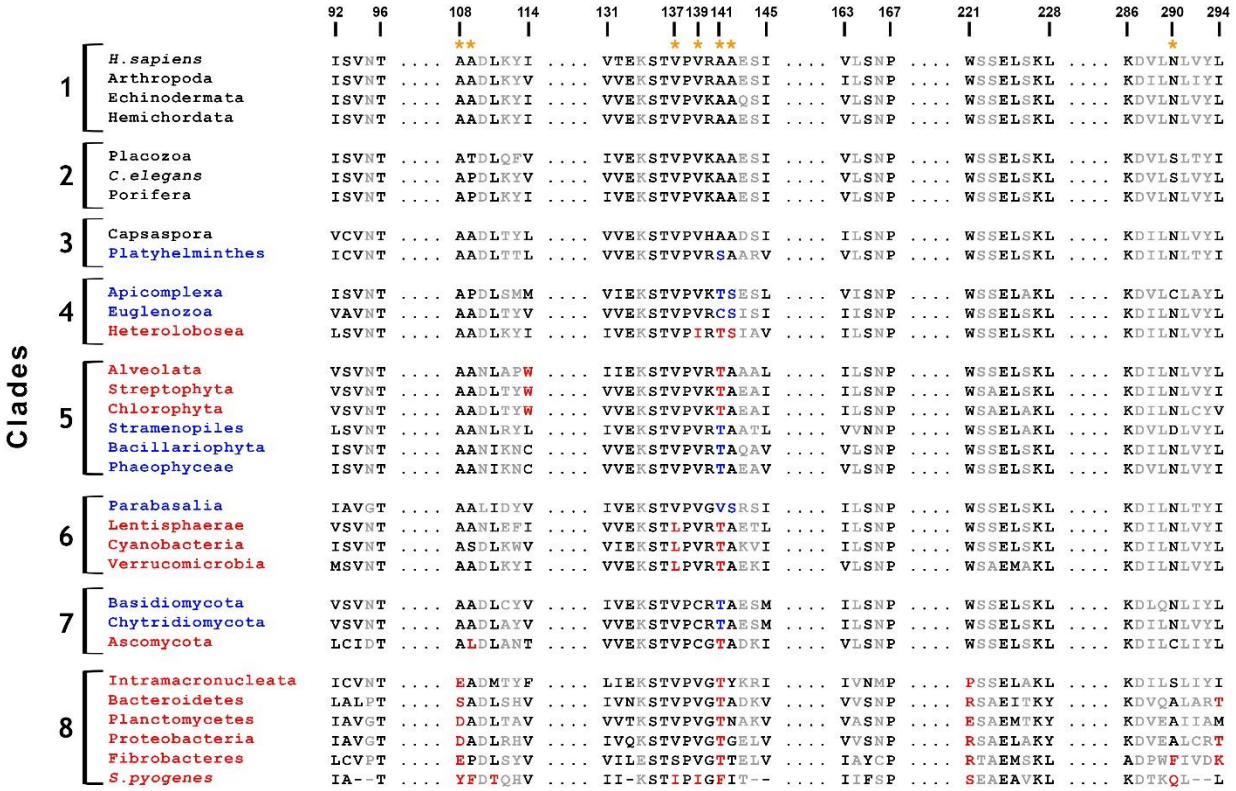


Figure 4.9: Representive Sequences from Each Phyla in the Evolutionary Analysis of UGDH. Residues that make up the allosteric sequence motif (orange stars) and the packing defects (bold black lettering), or simply adjacent in sequence (light grey). The prediction of allostery in the UGDHs that comprise clades 3-7 is purely speculative (see text for details). Within each clade, the phyla that are expected to be allosteric are shown in black. If it is unknown whether or not a specific phylum conserves the allosteric mechanism, it is colored blue along with the questionable substitutions. If a specific phylum is expected to be non-allosteric, it is colored red along with the questionable substitutions that are unlikely to be accommodated in the E^{Ω} structure.

4.8 Tables

Table 4.1: Data Collection and Refinement Statistics for UDP-Xyl Bound cUGDH

Data Collection	
Protein Data Bank Entry	6OM8
Space group	P 2 2_1 2_1
Unit cell dimensions (Å)	157.7, 168.2, 279.6
Completeness (%)	98.1 (98.4)[a]
No. reflections	266577 (19605)
Redundancy	6.3 (6.1)
$I / \sigma(I)$	9.1 (1.3)[a]
$CC_{1/2}$[b]	99.6 (45.5)[a]
R_{meas} (%)[c]	15.3 (164.6)
R_{sym} (%)	14.1 (150.6)
Refinement	
Resolution (Å)	2.45
R_{work} / R_{free}	0.19 / 0.22[d]
No. atoms: Protein / Ligand / Water	42921 / 816 / 1194
B-factors (Å^2): Protein / Ligand / Water	53.3 / 49.1 / 49.0
Stereochemical Ideality	
Bond lengths (Å^2)	0.008
Bond angles (°)	0.905
φ,ψ Most favored (%)[d]	96.6
φ,ψ Additionally allowed (%)	3.4

[a]Values in parentheses are for highest-resolution shell (2.51-2.45)
[b]$CC_{1/2}$ is the percentage of correlation between intensities from random half-data sets[102]
[c]R_{meas} is the redundancy-independent merging *R* factor[101]
[d]Posterior R_{free} (see methods for details)

Table 4.2: Kinetic Parameters for cUGDH and hUGDH

Enzyme	Ligand	K_M (μM)	k_{cat}[a] (s^{-1})	$K_{0.5}$ (μM)	k_{cat}[a] (s^{-1})	Hill (*h*)
cUGDH	UDP-Glc	55 ± 3	1.23 ± 0.03			1.0
	NAD^+			333 ± 18	1.7 ± 0.1	0.8 ± 0.1
hUGDH	UDP-Glc	11.0 ± 0.8	0.71 ± 0.02			1.0
	NAD^+			1094 ± 244	1.1 ± 0.1	0.6 ± 0.1

[a]One complete catalytic turnover produces two molecules of NADH.

Table 4.3: Global Analysis of UDP-Xylose Inhibition

Enzyme	UDP-Xyl (μM)	K_M (μM)[a]	k_{cat} (s^{-1})[a]	K_i (μM)[a]	Hill (*h*)
cUGDH	0	54 ± 3	1.23 ± 0.03	6.9 ± 0.6	1.0
	4	-	-	-	1.0 ± 0.1
	15	-	-	-	1.4 ± 0.1
hUGDH	0	11.0 ± 0.8	0.70 ± 0.01	0.42 ± 0.04	1.0
	1	-	-	-	1.3 ± 0.1

[a]Shared parameter in global fitting.

4.9 Supplementary Figures and Tables

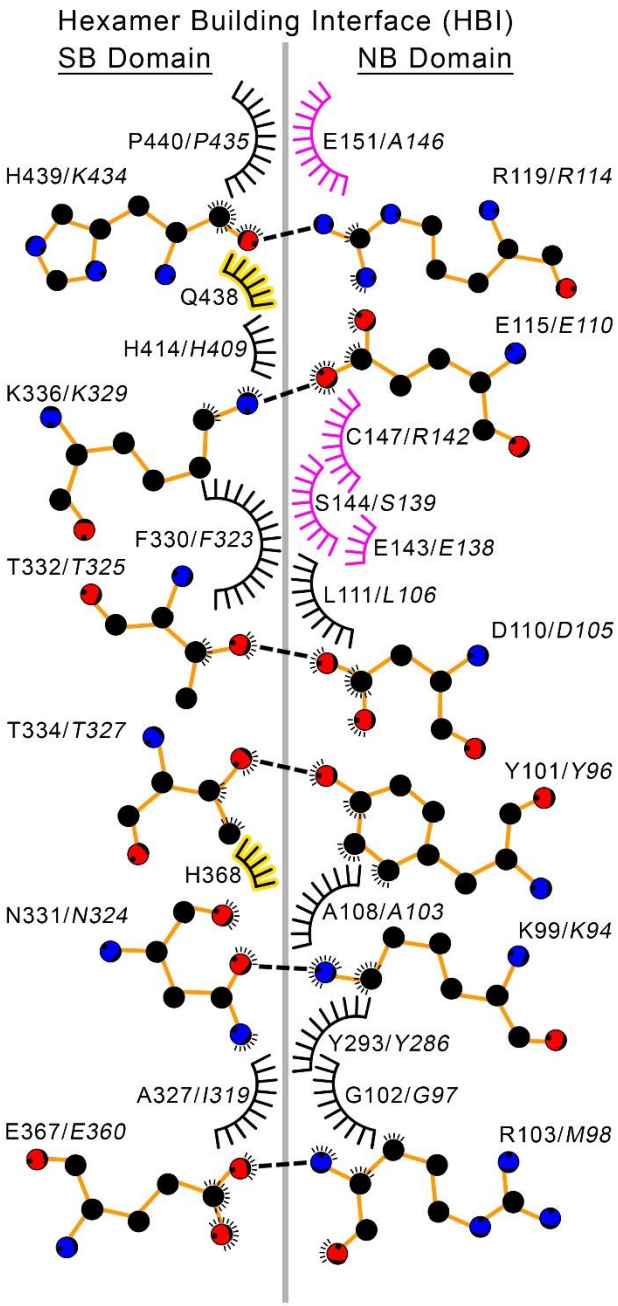


Figure S4.1: The Hexamer Building Interface Residues of cUGDH and hUGDH. The hexamer building interface (grey line) formed from the packing of SB and NB domains from adjoining dimers. The residues that pack at the interface for cUGDH are shown first, with the equivalent residues in hUGDH shown second and italicized. Hydrogen bonds (dashed lines) and van der Waals interactions (feathered lines) are depicted. Allosteric switch residues that interact in the interface are colored magenta. Yellow highlighting identifies residues that are present in cUGDH but missing in hUGDH.

Table S4.1: UGDH Phyla and Associated Genbank/Uniprot ID's that were Used in the Phylogenetic Analysis.

Phylum	Genbank/Uniprot ID	Species	Kingdom	Clade Number
Chordata	O60701.1	Homo sapiens	Metazoa	1
Arthropoda	AAC97125.1	Drosophila melanogaster	Metazoa	1
Echinodermata	XP_011683830.1	Strongylocentrotus purpuratus	Metazoa	1
Hemichordata	XP_002741382.1	Saccoglossus kowalevskii	Metazoa	1
Nematoda	NP_505730.1	Caenorhabditis elegans	Metazoa	2
Placozoa	XP_002113339.1	Trichoplax adhaerens	Metazoa	2
Porifera	XP_003383497.1	Amphimedon queenslandica	Metazoa	2
Platyhelminthes	CDS17287.1	Echinococcus granulosus	Metazoa	3
Capsaspora	XP_004365811.1	Capsaspora owczarzaki ATCC 30864	Protozoa	3
Euglenozoa	XP_813451.1	Trypanosoma cruzi strain CL	Protozoa	4
Heterolobosea	XP_002676546.1	Naegleria gruberi	Protozoa	4
Apicomplexa	XP_627000.1	Cryptosporidium parvum Iowa II	Protozoa	4
Streptophyta	NP_197053.1	Arabidopsis thaliana	Viridiplantae	5
Chlorophyta	XP_001703656.1	Chlamydomonas reinhardtii	Viridiplantae	5
Alveolata	XP_002766755.1	Perkinsus marinus ATCC 50983	Protozoa	5
Bacillariophyta	GAX09934.1	Fistulifera solaris	Protozoa	5
Phaeophyceae	CBJ28343.1	Ectocarpus siliculosus	Protozoa	5
Stramenopiles	XP_012897639.1	Blastocystis hominis	Protozoa	5
Parabasalia	XP_001310931.1	Trichomonas vaginalis G3	Protozoa	6
Cyanobacteria	WP_011818811.1	Prochlorococcus marinus	Bacteria	6
Lentisphaerae	WP_007280104.1	Lentisphaera araneosa	Bacteria	6
Verrucomicrobia	PHX69165.1	Opitutae bacterium	Bacteria	6
Ascomycota	RAQ69836.1	Aspergillus flavus	Fungi	7
Basidiomycota	XP_025347162.1	Pseudomicrostroma glucosiphilum	Fungi	7
Chytridiomycota	EGF83877.1	Batrachochytrium dendrobatidis JAM81	Fungi	7
Intramacronucleata	CDW89424.1	Stylonychia lemnae	Protozoa	8
Bacteroidetes	WP_014682690.1	Solitalea canadensis	Bacteria	8
Fibrobacteres	WP_022637184.1	Chitinivibrio alkaliphilus	Bacteria	8
Firmicutes	P0C0F4.1	Streptococcus pyogenes	Bacteria	8
Planctomycetes	WP_082839084.1	Gemmata sp. SH-PL17	Bacteria	8
Proteobacteria	WP_013900935.1	Ramlibacter tataouinensis	Bacteria	8

CHAPTER 5

CONCLUSIONS

The function of the T131loop-α6 helix as an allosteric switch that regulates the activity of hUGDH was a long-standing hypothesis.[39, 48, 49] Supporting evidence for this came predominantly form X-ray crystal studies that showed the switch adopted unique conformations depending on which, if any, ligand was bound to the active site.[39, 42, 44, 45] The movement of the switch was shown to alter the packing of the hexamer building interface; tuning the affinity between the subunits resulting in a allosteric response.[39] Analytical ultracentrifugation studies showed the allosteric effect is specific to UDP-Xyl binding the active site.[39, 44, 45] The position of the allosteric switch was also implicated in hysteresis, as crystal structures showed it adopted a unique intermediate conformation in the absence of ligand.[44, 45, 49] In Chapter two, the first direct evidence that the movement of the allosteric switch couples both allostery and hysteresis is presented.[105] The A136M substitution ($hUGDH_{A136M}$) was designed to sterically restrict the conformation of the allosteric switch to the active state. Kinetic, crystal, and analytical ultracentrifugation studies were conducted to verify the effects of the substitution. Restricting the allosteric switch to the E state completely abolished both the allosteric response, as inferred by both the loss of inhibitor induced cooperativity and sedimentation velocity studies, and hysteresis in progress curves. This work not only provided the first direct evidence that supported a long-standing hypothesis on hUGDH regulation, it also generated a new model protein that can potentially be used to study dysregulation of hUGDH *in vivo*.

In metabolism, committed steps are typically under tight regulation.[16, 17, 34, 36, 126, 127] Oxidation of UDP-Glc to UDP-GlcA is not a committed step, as UDP-GlcA is the substrate of many pathways.[1, 128-131] This begs the question; why does hUGDH have such a complex regulatory mechanism and is it even biologically relevant? While the work presented here is sheerly a biophysical study and cannot address this question directly, it has generated $hUGDH_{A136M}$ that could potentially be used to probe the biological relevance of this mechanism. This construct has the advantage of only requiring a single point mutation to the native protein to completely abolish allostery and greatly diminish UDP-Xyl sensitivity. Introducing the A136M substitution into native protein that is expressed in a human cell line could provide the first evidence that hUGDH is sensitive to UDP-Xyl inhibition *in vivo*. This has been a topic of debate as hUGDH is a cytosolic protein and UDP-Xyl is produced in the endoplasmic reticulum.[13, 131] The existence of an UDP-GlcA:UDP-Xyl antiporter has been proposed in support of this mechanism.[104, 132] However, more evidence is required to show it is sufficient to regulate hUGDH. A glycomics study monitoring glycan levels in cells with constitutively active hUGDH protein would go a long way in testing the model of UDP-Xyl inhibition. One would expect that if this mechanism of feedback inhibition is biologically relevant, these cells would have reduced levels of UDP-Glc and increased levels of UDP-GlcA as well as other more complex GlcA containing glycans.[104] While the study described here is preliminary, it would provide a great deal of insight into the physiological relevance of UDP-Xyl inhibition of UGDH.

The first study presented here provided strong evidence that the isomerization of the allosteric switch was responsible for both allostery and hysteresis. However, it provided little insight as to how these remarkable transitions occur within the confines of a protein core. Crystal structures of non-allosteric, bacterial, UGDH and human protein with either UDP-Glc of UDP-

Xyl bound provided the first tentative explanations for how the allosteric transition is possible.[39, 42, 45] Large packing defects that are unique to the E/E* and the inhibited E^{Ω} conformation in human enzyme were observed surround the allosteric switch.[39] These packing defects were not observed in the bacterial protein, which does not undergo the allosteric transition.[39] It was previously hypothesized that these packing defects evolved to provide the space in the protein core required for the allosteric transition.[39] In Chapter three, strong evidence has been provided to support this hypothesis. $Rosetta_{VIP}$ was used to design the A104L substitution that fills a packing defect unique to the E/E* conformation. Modeling of the A104L substitution into the E^{Ω} state core showed the substitution would not be sterically permitted. Kinetic studies were used to verify that without the additional space provided by the cavity, the allosteric transition was abolished. This result suggests packing defects evolved to sterically permit both conformations of the allosteric switch, and the transition between them, within the confines of the protein core. Although the E and E* states share similar packing defects, hysteresis was also abolished in $hUGDH_{A104L}$. Crystal studies confirmed the substituted protein still adopted the inactive E* state in the absence of ligand, albeit with a more ordered protein core. This result challenged the previous hypothesis that hysteresis in hUGDH is simply the movement of the allosteric switch. It instead suggests that the flexibility of the protein core required for the allosteric transition comes, with an increase in conformational entropy that results in hysteresis. Put simply, packing defects increase the number of non-productive protein core conformations that must be sampled before converging on the active state. The increased number of conformations that must be sampled could translate to a slow activation time i.e. hysteresis. The data presented here supports the conformational entropy model of hysteresis in hUGDH.

Interest in protein core packing is not unique to hUGDH.[57, 63, 65, 66, 68, 133] In fact, the role of protein core packing in allostery has been observed in other proteins.[134] Packing defects can offer exciting alternatives as druggable targets for regulating enzyme activity.[135-140] This is desirable in protein like hUGDH as the substrates are used in many other reactions, thus design of competitive inhibitor would likely result in off target effects.[141, 142] This work has provided a proof of concept that packing defects in the protein core of hUGDH are targetable features of the enzyme that can be used to regulate function. While these first steps are promising, there is greater interest in showing these packing defects can be targeted to favor the inhibited conformation of hUGDH. As the packing defects in the E^{Ω} conformation are unique and required for the transition to the E state, it stands to reason that they could be targeted to select for the inhibit conformation the enzyme. Once demonstrated robustly, these cavities in the E^{Ω} become potentially druggable targets for downregulating UDP-GlcA levels and by extension glucuronidation.

Thus far, the work presented here has provided strong evidence that allostery and hysteresis in hUGDH are couple through the movement of the allosteric switch, and this movement is facilitated by large packing defects in the protein core. In contrast, it was known bacterial UGDH proteins were non-allosteric with well packed protein cores. Based on the limited data, it was unknown if these features are unique to the human enzyme or if they are more prevalent; if so, how can these proteins be quickly and robustly identified. In Chapter four, a bioinformatic study showed that the large to small mutations in the protein core which created the packing defects surrounding the allosteric switch in hUGDH are observed in many other UGDH proteins. UGDH from *C. elegans* was identified as one such protein, which contained five of the six large to small mutations. Crystallographic, kinetic and analytical ultracentrifugation studies were conducted on cUGDH and it was found that this protein conserves both allostery and hysteresis. Crystal and

solutions structure studies have shown the movement of the allosteric switch in the presences of UDP-Xyl, the adoption of the E^{Ω} complex, and a strengthening of the hexamer building interface in cUGDH is similar to hUGDH. This work not only provided additional evidence that allostery and hysteresis are coupled, but that this feature expands beyond hUGDH to other homologs. The atypical allostery motif identified here can serve as a critical tool for identifying other allosteric UGDH proteins for further study. It would be of great interest to characterize proteins that contain only a few of the large to small substitutions identified by this motif, without the required insertion to form a hexamer. A single point mutation could be sufficient to cause a packing defect that increases both the conformational entropy of the protein core and allows for alternate conformations of the allosteric switch, resulting in hysteresis. This protein, and by extension the organism it is found in, could serve as a model to understand the regulatory effects of hysteresis alone in UGDH proteins. It has also been hypothesized that the hexamer evolved to stabilize the protein enough to accept the large to small mutations.[39] A protein that contains some of the atypical allostery motif but lacks the structural elements required to oligomerize would be an interesting branch point and provide great insight to the evolution of UGDH.

While human and *C. elegans* UGDH both display hysteresis and allosteric regulation, there is a substantial difference in the affinity for UDP-Xyl between the two proteins. Crystal structures of both proteins bound to UDP-Xyl showed the active sites and interactions with the ligand were identical, offering no explanation for the difference in affinity. However, there was a significant divergence in the amino acid composition of the hexamer building interface between the proteins. Because the active site serves as an allosteric site for dimer affinity these sites are coupled; thus, there exists thermodynamic reciprocity between the hexamer building interface and active site. Put simply, altering the hexamer building interface would change the affinity for the inhibitor at the

active site. The exact mechanism for how these sites are coupled is a key question that arose from this work. Currently, chimeric cUGDH proteins are being designed that have a hexamer building interface more similar to the human enzyme. Upon the conclusion of this study, key residues and interactions required for high UDP-Xyl affinity will be identified and added to the atypical allostery motif. This will result in not only being able to quickly identify allosteric UGDH proteins but also estimate the intrinsic affinity for the inhibitor from primary sequence. This study also suggests the hexamer building interface is a potential target for regulating enzyme function. Because the interfaces are unique between the E, E* and E^{Ω} states, a biologic that stabilizes one form preferentially would alter the enzymes function. While intracellular peptide fragments are not currently a common therapeutic tool, this would show the viability of developing small regulatory peptides in lieu of traditional small molecule approaches.[143-145]

The work presented here has not only answered several outstanding questions in UDP-glucose dehydrogenase field, it has generated both tools and new hypotheses for further study. It has now been shown that allostery and hysteresis are coupled in UDP-glucose dehydrogenase by the movement of a buried allosteric switch. The plasticity of the protein core that facilitates this movement originates from large to small mutations in the protein core surrounding the allosteric switch. These substitutions make up a motif that can be used to identify potential allosteric UGDH proteins from primary sequence. Additional studies will be required to refine this motif. There exists the potential for this motif to be expanded to predict high or low inhibitor affinity. Two hUGDH constructs have been developed that are both non-hysteretic and non-allosteric. These enzymes have the potential to be of great use in studying the allosteric regulation of hUGDH in *in vivo* systems. The old model of hysteresis in hUGDH which focused solely on the movement of the allosteric switch has also been called into question by this work.[39, 44, 45, 49] Instead, it has been

proposed that hysteresis in hUGDH arises from the conformational entropy of the protein core. The packing defects that evolved to facilitate the transition between the active and inhibited states, also allow the protein core to adopt an increased number of non-productive states. Evidence has been provided that shows the expanded conformational landscape that must be searched to find the active state results in hysteresis. This feature of protein core packing has the potential to exist in a large number of yet uncharacterized UGDH proteins. While this work has provided the first evidence that the evolution of allostery in UGDH proteins requires hysteresis, this work does not preclude a hysteretic UGDH protein that does not display allostery. Further study of UGDH proteins will be required to identify the exact origins of hysteresis, and the subsequent origin of allostery.

REFERENCES

[1] de Wildt, S. N., Kearns, G. L., Leeder, J. S., and van den Anker, J. N. (1999) Glucuronidation in Humans, *Clinical Pharmacokinetics 36*, 439-452.

[2] Kroemer, H. K., and Klotz, U. (1992) Glucuronidation of Drugs, *Clinical Pharmacokinetics 23*, 292-310.

[3] Burchell, B., and Coughtrie, M. W. H. (1989) UDP-glucuronosyltransferases, *Pharmacology & Therapeutics 43*, 261-289.

[4] Faed, E. M. (1984) Properties of Acyl Glucuronides: Implications for Studies of the Pharmacokinetics and Metabolism of Acidic Drugs, *Drug Metabolism Reviews 15*, 1213-1249.

[5] Osborne, R., Joel, S., Grebenik, K., Trew, D., and Slevin, M. (1993) The pharmacokinetics of morphine and morphine glucuronides in kidney failure, *Clinical Pharmacology & Therapeutics 54*, 158-167.

[6] Cummings, J., Boyd, G., Ethell, B. T., Macpherson, J. S., Burchell, B., Smyth, J. F., and Jodrell, D. I. (2002) Enhanced clearance of topoisomerase I inhibitors from human colon cancer cells by glucuronidation, *Biochem Pharmacol 63*, 607-613.

[7] Cummings, J., Ethell, B. T., Jardine, L., Boyd, G., Macpherson, J. S., Burchell, B., Smyth, J. F., and Jodrell, D. I. (2003) Glucuronidation as a mechanism of intrinsic drug resistance in human colon cancer: Reversal of resistance by food additives, *Cancer Res 63*, 8443-8450.

[8] de Almagro, M. C., Selga, E., Thibaut, R., Porte, C., Noe, V., and Ciudad, C. J. (2011) UDP-glucuronosyltransferase 1A6 overexpression in breast cancer cells resistant to methotrexate, *Biochemical Pharmacology 81*, 60-70.

[9] Oguri, T., Takahashi, T., Miyazaki, M., Isobe, T., Kohn, N., Mackenzie, P. I., and Fujiwara, Y. (2004) UGT1A10 is responsible for SN-38 glucuronidation and its expression in human lung cancers, *Anticancer Res 24*, 2893-2896.

[10] King, C. D., Rios, G. R., Green, M. D., and Tephly, T. R. (2000) UDP-Glucuronosyltransferases, *Current Drug Metabolism 1*, 143-161.

[11] Tukey, R. H., and Strassburg, C. P. (2000) Human UDP-Glucuronosyltransferases: Metabolism, Expression, and Disease, *Annual Review of Pharmacology and Toxicology 40*, 581-616.

[12] Ritter, J. K. (2000) Roles of glucuronidation and UDP-glucuronosyltransferases in xenobiotic bioactivation reactions, *Chemico-Biological Interactions 129*, 171-193.

[13] Maxwell, E. S., Kalckar, H. M., and Strominger, J. L. (1956) Some Properties of Uridine Diphosphoglucose Dehydrogenase, *Arch Biochem Biophys 65*, 2-10.

[14] Perutz, M. F. (1989) Mechanisms of cooperativity and allosteric regulation in proteins, *Quarterly Reviews of Biophysics 22*, 139-237.

[15] Bohr, C., Hasselbalch, K., and Krogh, A. (1904) About a new biological relation of high importance that the blood carbonic acid tension exercises on its oxygen binding., *Skand Arch Physiol 16*, 402-412.

[16] Changeux, J. P. (2012) Allostery and the Monod-Wyman-Changeux model after 50 years, *Annu Rev Biophys 41*, 103-133.

[17] Monod, J., and Jacob, F. (1961) General Conclusions - Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation, *Cold Spring Harb Sym 26*, 389-&.

[18] Goldbeter, A. (1976) Kinetic cooperativity in the concerted model for allosteric enzymes, *Biophysical Chemistry 4*, 159-169.

[19] Hill, A. V. (1913) The Combinations of Haemoglobin with Oxygen and with Carbon Monoxide. I, *Biochem J 7*, 471-480.

[20] Weber, G. (1972) Ligand binding and internal equilibria in proteins, *Biochemistry 11*, 864-878.

[21] Fersht, A. (2017) *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, World Scientific, New Jersey.

[22] Cui, Q., and Karplus, M. (2008) Allostery and cooperativity revisited, *Protein Sci 17*, 1295-1307.

[23] Kurganov, B. I. (2000) Analysis of negative cooperativity for glutamate dehydrogenase, *Biophysical Chemistry 87*, 185-199.

[24] Levitzki, A., and Koshland, D. E., Jr. (1969) Negative cooperativity in regulatory enzymes, *Proc Natl Acad Sci U S A 62*, 1121-1128.

[25] Changeux, J. P., Thiery, J., Tung, Y., and Kittel, C. (1967) On the cooperativity of biological membranes, *Proc Natl Acad Sci U S A 57*, 335-341.

[26] Koshland, D. E., Jr. (1996) The structural basis of negative cooperativity: receptors and enzymes, *Curr Opin Struct Biol 6*, 757-761.

[27] Nussinov, R., and Tsai, C. J. (2015) Allostery without a conformational change? Revisiting the paradigm, *Curr Opin Struc Biol 30*, 17-24.

[28] Pauling, L. (1935) The Oxygen Equilibrium of Hemoglobin and Its Structural Interpretation, *Proc Natl Acad Sci U S A 21*, 186-191.

[29] Perutz, M. F. (1970) Stereochemistry of cooperative effects in haemoglobin, *Nature 228*, 726-739.

[30] Perutz, M. F., Bolton, W., Diamond, R., Muirhead, H., and Watson, H. C. (1964) Structure of Haemoglobin. An X-Ray Examination of Reduced Horse Haemoglobin, *Nature 203*, 687-690.

[31] Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. (1960) Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis, *Nature 185*, 416-422.

[32] Tsai, C. J., and Nussinov, R. (2014) A Unified View of "How Allostery Works", *Plos Comput Biol 10*.

[33] Koshland, D. E., Jr., Nemethy, G., and Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits, *Biochemistry 5*, 365-385.

[34] Monod, J., Changeux, J. P., and Jacob, F. (1963) Allosteric proteins and cellular control systems, *J Mol Biol 6*, 306-329.

[35] Monod, J. (1972) *Chance and necessity; an essay on the natural philosophy of modern biology*, Vintage Books, New York,.

[36] Monod, J., Wyman, J., and Changeux, J. P. (1965) On the Nature of Allosteric Transitions: A Plausible Model, *J Mol Biol 12*, 88-118.

[37] Benesch, R., and Benesch, R. E. (1967) The effect of organic phosphates from the human erythrocyte on the allosteric properties of hemoglobin, *Biochemical and Biophysical Research Communications 26*, 162-167.

[38] Mulquiney, P. J., Bubb, W. A., and Kuchel, P. W. (1999) Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: in vivo kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using 13C and 31P NMR, *The Biochemical journal 342 Pt 3*, 567-580.

[39] Kadirvelraj, R., Sennett, N. C., Polizzi, S. J., Weitzel, S., and Wood, Z. A. (2011) Role of packing defects in the evolution of allostery and induced fit in human UDP-glucose dehydrogenase, *Biochemistry 50*, 5780-5789.

[40] Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. Z. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer, *Science 271*, 1247-1254.

[41] Balduini, C., Brovelli, A., De Luca, G., Galligani, L., and Castellani, A. A. (1973) Uridine diphosphate glucose dehydrogenase from cornea and epiphysial-plate cartilage, *The Biochemical journal 133*, 243-249.

[42] Egger, S., Chaikuad, A., Kavanagh, K. L., Oppermann, U., and Nidetzky, B. (2011) Structure and mechanism of human UDP-glucose 6-dehydrogenase, *J Biol Chem 286*, 23877-23887.

[43] Gainey, P. A., and Phelps, C. F. (1972) Uridine diphosphate glucuronic acid production and utilization in various tissues actively synthesizing glycosaminoglycans, *The Biochemical journal 128*, 215-227.

[44] Sennett, N. C., Kadirvelraj, R., and Wood, Z. A. (2011) Conformational flexibility in the allosteric regulation of human UDP-alpha-D-glucose 6-dehydrogenase, *Biochemistry 50*, 9651-9663.

[45] Sennett, N. C., Kadirvelraj, R., and Wood, Z. A. (2012) Cofactor binding triggers a molecular switch to allosterically activate human UDP-alpha-D-glucose 6-dehydrogenase, *Biochemistry 51*, 9364-9374.

[46] Campbell, R. E., Mosimann, S. C., van de Rijn, I., Tanner, M. E., and Strynadka, N. C. J. (2000) The First Structure of UDP-Glucose Dehydrogenase Reveals the Catalytic Residues Necessary for the Two-fold Oxidation, *Biochemistry 39*, 7012-7023.

[47] Neufeld, E. F., and Hall, C. W. (1965) Inhibition of Udp-D-Glucose Dehydrogenase by Udp-D-Xylose - A Possible Regulatory Mechanism, *Biochemical and Biophysical Research Communications 19*, 456-&.

[48] Kadirvelraj, R., Custer, G. S., Keul, N. D., Sennett, N. C., Sidlo, A. M., Walsh, R. M., and Wood, Z. A. (2014) Hysteresis in Human UDP-Glucose Dehydrogenase Is Due to a Restrained Hexameric Structure That Favors Feedback Inhibition, *Biochemistry 53*, 8043-8051.

[49] Kadirvelraj, R., Sennett, N. C., Custer, G. S., Phillips, R. S., and Wood, Z. A. (2013) Hysteresis and negative cooperativity in human UDP-glucose dehydrogenase, *Biochemistry 52*, 1456-1465.

[50] Dickinson, F. M. (1988) Studies on the unusual behaviour of bovine liver UDP-glucose dehydrogenase in assays at acid and neutral pH and on the presence of tightly bound nucleotide material in purified preparations of this enzyme, *The Biochemical journal 255*, 775-780.

[51] Gainey, P. A., and Phelps, C. F. (1975) Interactions of Uridine-Diphosphate Glucose Dehydrogenase with Inhibitor Uridine-Diphosphate Xylose, *Biochemical Journal 145*, 129-134.

[52] Frieden, C. (1970) Kinetic aspects of regulation of metabolic processes. The hysteretic enzyme concept, *J Biol Chem 245*, 5788-5799.

[53] Chan, H. S., and Dill, K. A. (1991) Polymer principles in protein structure and stability, *Annu Rev Biophys Biophys Chem 20*, 447-490.

[54] Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry 29*, 7133-7155.

[55] Dill, K. A. (1985) Theory for the folding and stability of globular proteins, *Biochemistry 24*, 1501-1509.

[56] Tanford, C. (1962) Contribution of Hydrophobic Interactions to Stability of Globular Conformation of Proteins, *J Am Chem Soc 84*, 4240-&.

[57] Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene, *Nature 355*, 371-373.

[58] G D Rose, a., and Wolfenden, R. (1993) Hydrogen Bonding, Hydrophobicity, Packing, and Protein Folding, *Annual Review of Biophysics and Biomolecular Structure 22*, 381-415.

[59] Pauling, L., and Corey, R. B. (1951) Atomic Coordinates and Structure Factors for 2 Helical Configurations of Polypeptide Chains, *P Natl Acad Sci USA 37*, 235-240.

[60] Pauling, L., Corey, R. B., and Branson, H. R. (1951) The Structure of Proteins - 2 Hydrogen-Bonded Helical Configurations of the Polypeptide Chain, *P Natl Acad Sci USA 37*, 205-211.

[61] Crick, F. (1953) The packing of [alpha]-helices: simple coiled-coils, *Acta Crystallographica 6*, 689-697.

[62] Chothia, C., Levitt, M., and Richardson, D. (1981) Helix to helix packing in proteins, *Journal of Molecular Biology 145*, 215-250.

[63] Richards, F. M. (1974) The interpretation of protein structures: Total volume, group volume distributions and packing density, *Journal of Molecular Biology 82*, 1-14.

[64] Baldwin, E. P., Hajiseyedjavadi, O., Baase, W. A., and Matthews, B. W. (1993) The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme, *Science 262*, 1715.

[65] Gassner, N. C., Baase, W. A., and Matthews, B. W. (1996) A test of the "jigsaw puzzle" model for protein folding by multiple methionine substitutions within the core of T4 lysozyme, *P Natl Acad Sci USA 93*, 12155-12158.

[66] Matthews, B. W., and Liu, L. (2009) A review about nothing: are apolar cavities in proteins really empty?, *Protein Sci 18*, 494-502.

[67] Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *P Natl Acad Sci USA 84*, 6663-6667.

[68] Zhang, X. J., Wozniak, J. A., and Matthews, B. W. (1995) Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme, *J Mol Biol 250*, 527-552.

[69] Kellis, J. T., Nyberg, K., Sˇail, D. a., and Fersht, A. R. (1988) Contribution of hydrophobic interactions to protein stability, *Nature 333*, 784-786.

[70] Lim, W. A., and Sauer, R. T. (1991) The role of internal packing interactions in determining the structure and stability of a protein, *Journal of Molecular Biology 219*, 359-376.

[71] Bromberg, S., and Dill, K. A. (1994) Side-chain entropy and packing in proteins, *Protein science : a publication of the Protein Society 3*, 997-1009.

[72] Munson, M., Balasubramanian, S., Fleming, K. G., Nagi, A. D., O'Brien, R., Sturtevant, J. M., and Regan, L. (1996) What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties, *Protein science : a publication of the Protein Society 5*, 1584-1593.

[73] Connolly, M. L. (1983) Solvent-accessible surfaces of proteins and nucleic acids, *Science 221*, 709.

[74] Borgo, B., and Havranek, J. J. (2012) Automated selection of stabilizing mutations in designed and natural proteins, *Proceedings of the National Academy of Sciences 109*, 1494.

[75] van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., and Babu, M. M. (2014) Classification of Intrinsically Disordered Regions and Proteins, *Chem Rev 114*, 6589-6631.

[76] He, B., Wang, K., Liu, Y., Xue, B., Uversky, V. N., and Dunker, A. K. (2009) Predicting intrinsic disorder in proteins: an overview, *Cell Res 19*, 929-949.

[77] Papaleo, E., Saladino, G., Lambrughi, M., Lindorff-Larsen, K., Gervasio, F. L., and Nussinov, R. (2016) The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery, *Chem Rev 116*, 6391-6423.

[78] Bickel, T., Jeppesen, C., and Marques, C. M. (2001) Local entropic effects of polymers grafted to soft interfaces, *Eur Phys J E 4*, 33-43.

[79] Carmichael, S. P., and Shell, M. S. (2015) Entropic (de)stabilization of surface-bound peptides conjugated with polymers, *J Chem Phys 143*.

[80] Guillemette, C. (2003) Pharmacogenomics of human UDP-glucuronosyltransferase enzymes, *Pharmacogenomics J 3*, 136-158.

[81] Lin, J. H., and Lu, A. Y. (1997) Role of pharmacokinetics and metabolism in drug discovery and development, *Pharmacol Rev 49*, 403-449.

[82] Kakizaki, I., Kojima, K., Takagaki, K., Endo, M., Kannagi, R., Ito, M., Maruo, Y., Sato, H., Yasuda, T., Mita, S., Kimata, K., and Itano, N. (2004) A novel mechanism for the inhibition of hyaluronan biosynthesis by 4-methylumbelliferone, *J Biol Chem 279*, 33281-33289.

[83] Kultti, A., Pasonen-Seppanen, S., Jauhiainen, M., Rilla, K. J., Karna, R., Pyoria, E., Tammi, R. H., and Tammi, M. I. (2009) 4-Methylumbelliferone inhibits hyaluronan synthesis by depletion of cellular UDP-glucuronic acid and downregulation of hyaluronan synthase 2 and 3, *Exp Cell Res 315*, 1914-1923.

[84] Wei, Q., Galbenus, R., Raza, A., Cerny, R. L., and Simpson, M. A. (2009) Androgen-Stimulated UDP-Glucose Dehydrogenase Expression Limits Prostate Androgen Availability without Impacting Hyaluronan Levels, *Cancer Res 69*, 2332-2339.

[85] Helt, S. S., Thymark, M., Harris, P., Aagaard, C., Dietrich, J., Larsen, S., and Willemoes, M. (2008) Mechanism of dTTP Inhibition of the Bifunctional dCTP Deaminase:dUTPase Encoded by Mycobacterium tuberculosis, *Journal of Molecular Biology 376*, 554-569.

[86] Johansson, E., Thymark, M., Bynck, J. H., Fanø, M., Larsen, S., and Willemoës, M. (2007) Regulation of dCTP deaminase from Escherichia coli by nonallosteric dTTP binding to an inactive form of the enzyme, *The FEBS Journal 274*, 4188-4198.

[87] Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server, *Methods Mol Biol 112*, 531-552.

[88] Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.

[89] Brunger, A. T. (1997) Free R value: cross-validation in crystallography, *Methods Enzymol 277*, 366-396.

[90] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.

[91] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.

[92] Urzhumtsev, A., Afonine, P. V., and Adams, P. D. (2013) TLS from fundamentals to practice, *Crystallogr Rev 19*, 230-270.

[93] Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) *Analytical Ultracentrifugation*, The Royal Society of Chemistry, Cambridge.

[94] Schuck, P. (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation, *Anal Biochem 320*, 104-124.

[95] Cornish-Bowden, A. (2001) Detection of Errors of Interpretation in Experiments in Enzyme Kinetics, *Methods 24*, 181-190.

[96] Hayward, S., and Berendsen, H. J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme, *Proteins 30*, 144-154.

[97] Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling, *Biophys J 78*, 1606-1619.

[98] Zhao, H., Balbo, A., Brown, P. H., and Schuck, P. (2011) The boundary structure in the analysis of reversibly interacting systems by sedimentation velocity, *Methods 54*, 16-30.

[99] Kuo, L. C. (1983) Allosteric cofactor-mediated enzyme cooperativity: a theoretical treatment, *P Natl Acad Sci USA 80*, 5243-5247.

[100] Kurganov, B. I., Dorozhko, A. I., Kagan, Z. S., and Yakovlev, V. A. (1976) The theoretical analysis of kinetic behaviour of "hysteretic" allosteric enzymes III. Dissociating and associating enzyme systems in which the rate of installation of equilibrium between the oligomeric forms is comparable to that of enzymatic reaction, *Journal of Theoretical Biology 60*, 287-299.

[101] Diederichs, K., and Karplus, P. A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography, *Nat Struct Biol 4*, 269-275.

[102] Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality, *Science 336*, 1030-1033.

[103] Tukey, R. H., and Strassburg, C. P. (2000) Human UDP-glucuronosyltransferases: Metabolism, expression, and disease, *Annual Review of Pharmacology and Toxicology 40*, 581-616.

[104] Bakker, H., Oka, T., Ashikov, A., Yadav, A., Berger, M., Rana, N. A., Bai, X., Jigami, Y., Haltiwanger, R. S., Esko, J. D., and Gerardy-Schahn, R. (2009) Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase, *The Journal of biological chemistry 284*, 2576-2583.

[105] Beattie, N. R., Keul, N. D., Sidlo, A. M., and Wood, Z. A. (2017) Allostery and Hysteresis Are Coupled in Human UDP-Glucose Dehydrogenase, *Biochemistry 56*, 202-211.

[106] Kozlikova, B., Sebestova, E., Sustr, V., Brezovsky, J., Strnad, O., Daniel, L., Bednar, D., Pavelka, A., Manak, M., Bezdeka, M., Benes, P., Kotry, M., Gora, A., Damborsky, J., and Sochor, J. (2014) CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures, *Bioinformatics 30*, 2684-2685.

[107] Carugo, O., and Argos, P. (1997) Correlation between side chain mobility and conformation in protein structures, *Protein Eng 10*, 777-787.

[108] Levitt, M., and Park, B. H. (1993) Water: now you see it, now you don't, *Structure 1*, 223-226.

[109] Beattie, N. R., Pioso, B. J., Sidlo, A. M., Keul, N. D., and Wood, Z. A. (2018) Hysteresis and Allostery in Human UDP-Glucose Dehydrogenase Require a Flexible Protein Core, *Biochemistry 57*, 6848-6859.

[110] Keul, N. D., Oruganty, K., Schaper Bergman, E. T., Beattie, N. R., McDonald, W. E., Kadirvelraj, R., Gross, M. L., Phillips, R. S., Harvey, S. C., and Wood, Z. A. (2018) The entropic force generated by intrinsically disordered segments tunes protein function, *Nature*, 584-588.

[111] Maddison, W. P. (1997) Gene trees in species trees, *Syst Biol 46*, 523-536.

[112] Nichols, R. (2001) Gene trees and species trees are not the same, *Trends in Ecology & Evolution 16*, 358-364.

[113] Andersson, J. O., Hirt, R. P., Foster, P. G., and Roger, A. J. (2006) Evolution of four gene families with patchy phylogenetic distributions: influx of genes into protist genomes, *BMC Evolutionary Biology 6*, 27.

[114] Keul, N. D., Oruganty, K., Schaper Bergman, E. T., Beattie, N. R., McDonald, W. E., Kadirvelraj, R., Gross, M. L., Phillips, R. S., Harvey, S. C., and Wood, Z. A. (2018) The entropic force generated by intrinsically disordered segments tunes protein function, *Nature*.

[115] Beattie, N. R., Pioso, B., Sidlo, A. M., Keul, N. D., and Wood, Z. A. (2018) Hysteresis and Allostery in Human UDP-Glucose Dehydrogenase Require a Flexible Protein Core, *Biochemistry*.

[116] Girdlestone, C., and Hayward, S. (2016) The DynDom3D Webserver for the Analysis of Domain Movements in Multimeric Proteins, *J Comput Biol 23*, 21-26.

[117] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol 372*, 774-797.

[118] Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, *J Chem Inf Model 51*, 2778-2786.

[119] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.

[120] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem 25*, 1605-1612.

[121] Ortega, A., Amoros, D., and Garcia de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models, *Biophys J 101*, 892-898.

[122] Neuwald, A. F. (2009) Rapid detection, classification and accurate alignment of up to a million or more related protein sequences, *Bioinformatics 25*, 1869-1875.

[123] Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, *Mol Biol Evol 32*, 268-274.

[124] Price, M. N., Dehal, P. S., and Arkin, A. P. (2010) FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments, *Plos One 5*.

[125] Bruns, C. M., Hubatsch, I., Ridderstrom, M., Mannervik, B., and Tainer, J. A. (1999) Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products, *J Mol Biol 288*, 427-439.

[126] Pinon, V., Ravanel, S., Douce, R., and Alban, C. (2005) Biotin Synthesis in Plants. The First Committed Step of the Pathway Is Catalyzed by a Cytosolic 7-Keto-8-Aminopelargonic Acid Synthase, *Plant Physiology 139*, 1666.

[127] Rose, I. A., and Rose, Z. B. (1969) Chapter III - Glycolysis: Regulation and Mechanisms of the Enzymes* *The research of the authors referred to in this chapter has been sponsored by grants from the U.S. Public Health Service and the American Cancer Society, In *Comprehensive Biochemistry* (Florkin, M., and Stotz, E. H., Eds.), pp 93-161, Elsevier.

[128] Sugahara, K., Schwartz, N. B., and Dorfman, A. (1979) Biosynthesis of hyaluronic acid by Streptococcus, *J Biol Chem 254*, 6252-6261.

[129] Heleno, S. A., Martins, A., Queiroz, M. J. R. P., and Ferreira, I. C. F. R. (2015) Bioactivity of phenolic acids: Metabolites versus parent compounds: A review, *Food Chemistry 173*, 501-513.

[130] Kearns, A. E., Vertel, B. M., and Schwartz, N. B. (1993) Topography of glycosylation and UDP-xylose production, *Journal of Biological Chemistry 268*, 11097-11104.

[131] Vertel, B. M., Walters, L. M., Flay, N., Kearns, A. E., and Schwartz, N. B. (1993) Xylosylation is an endoplasmic reticulum to Golgi event, *Journal of Biological Chemistry 268*, 11105-11112.

[132] Bossuyt, X., and Blanckaert, N. (2001) Differential regulation of UDP-GlcUA transport in endoplasmic reticulum and in Golgi membranes, *J Hepatol 34*, 210-214.

[133] Sheffler, W., and Baker, D. (2009) RosettaHoles: rapid assessment of protein core packing for structure prediction, refinement, design, and validation, *Protein science : a publication of the Protein Society 18*, 229-239.

[134] Corrêa, F., Key, J., Kuhlman, B., and Gardner, Kevin H. (2016) Computational Repacking of HIF-2α Cavity Replaces Water-Based Stabilized Core, *Structure 24*, 1918-1927.

[135] Weisel, M., Proschak, E., Kriegl, J. M., and Schneider, G. (2009) Form follows function: Shape analysis of protein cavities for receptor-based drug design, *PROTEOMICS 9*, 451-459.

[136] Yuan, Y., Pei, J., and Lai, L. (2013) Binding Site Detection and Druggability Prediction of Protein Targets for Structure- Based Drug Design, *Current Pharmaceutical Design 19*, 2326-2333.

[137] Elliott, P. R., Pei, X. Y., Dafforn, T. R., and Lomas, D. A. (2000) Topography of a 2.0 Å structure of α1-antitrypsin reveals targets for rational drug design to prevent conformational disease, *Protein Science 9*, 1274-1281.

[138] Yuan, Y., Pei, J., and Lai, L. (2011) LigBuilder 2: A Practical de Novo Drug Design Approach, *Journal of Chemical Information and Modeling 51*, 1083-1091.

[139] Nayal, M., and Honig, B. (2006) On the nature of cavities on protein surfaces: Application to the identification of drug-binding sites, *Proteins: Structure, Function, and Bioinformatics 63*, 892-906.

[140] Zhang, Z., Li, Y., Lin, B., Schroeder, M., and Huang, B. (2011) Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction, *Bioinformatics 27*, 2083-2088.

[141] Campillos, M., Kuhn, M., Gavin, A.-C., Jensen, L. J., and Bork, P. (2008) Drug Target Identification Using Side-Effect Similarity, *Science 321*, 263-266.

[142] Brömme, D., and Lecaille, F. (2009) Cathepsin K inhibitors for osteoporosis and potential off-target effects, *Expert Opinion on Investigational Drugs 18*, 585-600.

[143] Chan, Y. P., Meyrueix, R., Kravtzoff, R., Nicolas, F., and Lundstrom, K. (2007) Review on Medusa®:a polymer-based sustained release technology for protein and peptide drugs, *Expert Opinion on Drug Delivery 4*, 441-451.

[144] George, M., and Abraham, T. E. (2006) Polyionic hydrocolloids for the intestinal delivery of protein drugs: Alginate and chitosan — a review, *Journal of Controlled Release 114*, 1-14.

[145] Humphrey, M. J., and Ringrose, P. S. (1986) Peptides and Related Drugs: A Review of Their Absorption, Metabolism, and Excretion, *Drug Metabolism Reviews 17*, 283-310.

APPENDEX A

THE ENTROPIC FORCE GENERATED BY INTRINSICALLY DISORDERED SEGMENTS TUNES PROTEIN FUNCTION

Nicholas D. Keul, Krishnadev Oruganty, Elizabeth T. Schaper Bergman, **Nathaniel R. Beattie**, Weston E. McDonald, Renuka Kadirvelraj, Michael L. Gross, Robert S. Phillips, Stephen C. Harvey, Zachary A. Wood. *Nature*, 563(7732), p.584 (2018). Reprinted here with permission of the publisher.

The contributions of Nathaniel R. Beattie to this manuscript include: refining, analyzing and depositing the crystal structures and design of the A136M substitution.

A1 Abstract

Protein structures are dynamic and can explore a large conformational landscape.[1,2] Only some of these structural substates are important for protein function (such as ligand binding, catalysis and regulation).[3-5] How evolution shapes the structural ensemble to optimize a specific function is poorly understood.[4,5] One of the constraints on the evolution of proteins is the stability of the folded 'native' state. Despite this, 44% of the human proteome contains intrinsically disordered peptide segments greater than 30 residues in length[6], the majority of which have no known function.[7-9] Here we show that the entropic force produced by an intrinsically disordered carboxy terminus (ID-tail) shifts the conformational ensemble of human UDP-α-D-glucose-6-dehydrogenase (UGDH) towards a substate with a high affinity for an allosteric inhibitor. The function of the ID-tail does not depend on its sequence or chemical composition. Instead, the affinity enhancement can be accurately predicted based on the length of the intrinsically disordered segment, and is consistent with the entropic force generated by an unstructured peptide attached to the protein surface.[10-13] Our data show that the unfolded state of the ID-tail rectifies the dynamics and structure of UGDH to favour inhibitor binding. Because this entropic rectifier does not have any sequence or structural constraints, it is an easily acquired adaptation. This model implies that evolution selects for disordered segments to tune the energy landscape of proteins, which may explain the persistence of intrinsic disorder in the proteome.

A2 Main Text (Introduction, Results, and Discussion)

Intrinsically disordered segments can exhibit complex functions such as ligand binding, scaffolding of multi-protein complexes and mediating allosteric regulation.[14-18] However, many intrinsically disordered segments are assumed to be nonfunctional and are often removed from proteins to facilitate structural studies. For example, the 30-residue disordered C terminus of UGDH (residues 465–494) is often removed with no apparent impact on kinetic parameters.[19] Here, we show that this C-terminal segment (called the ID-tail) plays a role in the allosteric mechanism of UGDH. UGDH catalyses the NAD^+-dependent oxidation of UDP-α-D-glucose (UDP-Glc) to UDP-α-D-glucuronic acid[19], and is regulated by the allosteric feedback inhibitor UDP-α-D-xylose (UDP-Xyl).[20,21] Three UGDH dimers associate to form an inactive hexamer (E*)[22-26] (Figure A1.a-b). The binding of substrate induces an allosteric switch (T131-loop–α6 helix) in the E* hexamer to produce the active state (E)[22,23,25,27] (Figure A1.a&c). The allosteric inhibitor UDP-Xyl competes with UDP-Glc for the active sites, and upon binding, triggers the allosteric switch to produce the inhibited state (E^{Ω}).[22,23,25,26] This inhibition mechanism is atypical in that the active site also functions as an allosteric site to control the structure and activity of the hexamer[22-27] (Figure A1.a&c). The E^{Ω} state has a high affinity for UDP-Xyl and a low affinity for UDP-Glc.[22,23] Therefore, the allosteric transition of the inhibited E^{Ω} hexamer to the E state can be observed as cooperativity in substrate saturation curves.[22,23] We compared the structure and activity of full-length UGDH (UGDH(FL)) to a construct lacking the ID-tail (UGDH(ΔID)). We solved the structures of E* states of UGDH(FL) and UGDH(ΔID) in isomorphous crystal lattices, showing that there were no substantial differences (Figure A.S1.a–c). UGDH(FL) and UGDH(ΔID) also have a similar catalytic rate constant (k_{cat}) and Michaelis constant (K_M) for both substrate and coenzyme, consistent with earlier reports[19] (Table A.S2). By contrast, the allosteric

response is sensitive to the ID-tail; deletion of the ID-tail reduces the affinity for UDP-Xyl by an order of magnitude (Figure A1.d). UGDH(ΔID) still binds UDP-Glc cooperatively, indicating that the deletion of the ID-tail reduces UDP-Xyl affinity but does not prevent the formation of the E^{Ω} hexamer (Figure A1.d, Figure A.S2.a&b).

Both the ID-tail and the α6 helix of the allosteric switch are located in the hexamer-building interface between adjacent dimers, suggesting that these two elements may work together to increase the affinity for UDP-Xyl (Figure A1.b). We used the allostery-quenching A136M substitution to determine whether the ID-tail functions independently of the allosteric switch. This substitution has been shown to lock the allosteric switch and the hexamer in the low UDP-Xyl-affinity E state.[22] Inhibition studies show no marked difference between the UDP-Xyl affinities of UGDH(FL/A136M) and UGDH(ΔID/A136M), which suggests that the ID-tail requires both a functional allosteric switch and the E^{Ω} state to enhance the affinity for UDP-Xyl (Figure A1.d).

The location of the α6 helix in the hexamer-building interface suggests that the oligomeric structure might be important for the function of the ID-tail (Figure A1.b). Sedimentation velocity studies show that the UGDH(ΔID) E* hexamer is slightly less stable than the UGDH(FL) E* hexamer, perhaps explaining its reduced affinity for UDP-Xyl (Figure A.S3.a). We tested the role of the hexamer with the M11 interfacial loop substitution, which prevents hexamer formation and stabilizes the dimer (UGDH(FL-dimer) and UGDH(ΔID-dimer)).[23] UDP-Xyl binds to the UGDH(FL-dimer) with sevenfold higher affinity than UGDH(ΔID-dimer), demonstrating that the ID-tail does not require the hexamer to enhance the affinity for UDP-Xyl (Figure A1.d and Supplementary Information Section 1).

The ID-tail is highly conserved in vertebrate UGDHs (Figure A2.a). We examined the importance of primary structure in the ID-tail by randomizing the native sequence to create two

distinct ID-tails (UGDH(R1) and UGDH(R2)) (Figure A2.b). Surprisingly, the UGDH(FL), UGDH(R1) and UGDH(R2) constructs have similar affinities for UDP-Xyl (Figure A2.c). Next, all six prolines in the ID-tail were substituted with serine (UGDH(−Pro)) (Figure A2.b). Because serine and proline both promote disorder[28,29], this substitution conserves the unfolded state and disrupts any possible proline-specific interactions. Analysis of UGDH(−Pro) shows that the prolines do not contribute to UDP-Xyl affinity (Figure A2.c). Because all of the above constructs conserve the positive charge of the native ID-tail ($pI = 10.1$), we created a negatively charged ID-tail ($pI = 4.4$) using lysine-to-serine substitutions (UGDH(−Lys)) (Figure A2.b). Despite the charge switch, there was still no substantial change in UDP-Xyl affinity (Figure A2.c). Finally, we replaced the ID-tail with polyserine (UGDH(Ser)) without causing a marked change in UDP-Xyl affinity (Figure A2.b&c). Therefore, the conserved primary structure is not required for UDP-Xyl affinity, but may have been selected for because of an additional, unrelated function in vivo (Figure A2.a). The absence of sequence constraints argues against any mechanism in which the ID-tail specifically interacts with the inhibitor or the protein.

Next, we considered the possibility that the ID-tail might enhance UDP-Xyl affinity through a sequence-independent interaction involving the polypeptide backbone. Because the six prolines in the UGDH(FL), UGDH(R1) and UGDH(R2) ID-tails sample 16 unique positions throughout the sequence without altering UDP-Xyl affinity, it is unlikely that a backbone-specific interaction is important for function (Figure A2.b&c). Nevertheless, if there was a backbone-specific interaction, a plot of affinity versus ID-tail length would reveal a discontinuity when the critical segment was removed. Inhibition studies comparing UGDH(FL), UGDH(ΔID) and three new constructs with ID-tails of varying length (UGDH(2×FL), UGDH(0.5×FL) UGDH(0.26×FL) and UGDH(0.13×FL), shown in Figure A2.b) show that the affinity can be modelled as a simple

exponential decay (Figure A2.d). We confirmed that this saturable effect is independent of sequence by using polyserine ID-tails of corresponding lengths (UGDH(Ser), UGDH(0.5×Ser), UGDH(0.26×Ser) and UGDH(0.13×Ser)) and similarly, using corresponding lengths of the scrambled R1 construct (UGDH(R1), UGDH(0.5×R1), UGDH(0.26×R1) and UGDH(0.13×R1)) (Figure 2.2d). It is notable that UGDH(0.13×FL), UGDH(0.13×Ser) and UGDH(0.13×R1) still enhance UDP-Xyl binding affinity; the conformations of these short, four-residue ID-tails are tightly constrained within a surface pocket, which should stabilize any weak structure (Figure A3.a). Nevertheless, none of the E, E* and E^{Ω} UGDH(FL) crystal structures (42 unique chains) show evidence of an ordered interaction within the pocket (Figure A.S1 and refs[22,25-27]).

The data presented so far provide strong evidence that the high-affinity binding of UDP-Xyl is a function of the unfolded state of the ID-tail. An unstructured polymer tethered to a surface generates an entropic force at the point of attachment[10,11,13], which can be strong enough to distort lipid bilayers[30] and alter protein stability[12]. This force originates from the volume exclusion effects of the surface, which reduce the conformational entropy of the attached polymer (Figure A3.b). Because the entropy of the polymer increases with distance from the surface, the entropic force converges to a maximum value as the chain length increases.[10,11,13] The unfavourable change in free energy produced by constraining an unstructured, non-interacting peptide ($\Delta G_{\text{constrained}}$) is:

$$\Delta G_{\text{constrained}} = -RT\, ln\left(\frac{\Omega_2}{\Omega_1}\right) \qquad \text{(equation 1)}$$

where Ω_1 is the sum of all possible states of an unconstrained peptide and Ω_2 is the subset of states constrained by the protein surface and the adjacent ID-tail (RT is the product of the molar gas constant, R, and the temperature, T). Using Monte Carlo sampling of coarse-grained, sterically allowed bins of ϕ and ψ torsion angles we calculated the fraction of constrained conformations for various ID-tail lengths (see Methods, Figure A3.b, Figure A.S5). For this simulation, the adjacent

ID-tail was held in a fixed conformation (Figure A.S5). If the conformational entropy of the ID-tail contributes to the change in UDP-Xyl affinity, then we would expect Ω_2/Ω_1 and the affinity constant K_i to display similar behaviour with increasing tail length. Despite the simplicity of the Monte Carlo model, the simulations confirm that Ω_2/Ω_1 converges as the ID-tail length increases (Figure A3.c).

Studies have shown that the entropic force generated by a tethered polymer can alter protein stability.[12] We carried out thermal denaturation studies of UGDH dimers (chosen to avoid complications arising from hexamer dissociation), and found that the high-affinity UGDH(FL-dimer) ($K_i = 0.17$ μM) is less stable than the low-affinity UGDH(ΔID-dimer) ($K_i = 1.23$ μM) (Figure 2.3d). The destabilizing effect of the ID-tail should also be reflected in the structure and dynamics of UGDH. To examine these changes at the peptide level, we compared the hydrogen–deuterium exchange (HDX) rates of UGDH(FL-dimer) and UGDH(ΔID-dimer) using mass spectrometry. As expected, the fragment corresponding to the ID-tail is fully exchanged in less than 120 s, which is consistent with a disordered peptide[31] (Figure AS.6.a). The ID-tail increases the HDX rates of several segments in the NAD^+ binding domain, with the largest increases occurring in the allosteric switch and an adjacent peptide (Figure A3.e–g). An increase in HDX rates for a buried peptide such as the allosteric switch and the surrounding segments indicates an increase in the overall dynamics of the domain. This is notable, because the binding of UDP-Xyl induces the allosteric switch and surrounding core residues to change conformation and repack into the high affinity E^{Ω} state[22,23] (Figure A1.a&c). The ID-tail also decreases the HDX rates of several segments in the dimerization and sugar-binding domains, suggesting that these areas become more structured (Figure A3.e&g). The largest decrease is observed in the α9 helix of the dimerization domain (residues 222–240). This helix is largely inaccessible to solvent in crystal

structures, which suggests that the ID-tail reduces the overall dynamics of the dimer interface (Figure A3.g). Overall, the data show that the cost of constraining the ID-tail destabilizes a low-affinity substate, which biases the conformational ensemble towards a structurally and dynamically distinct high-affinity substate. A simple exponential fit of Ω_2/Ω_1 (Figure A3.c) shows that the energetic cost of constraining the ID-tail converges to approximately 2.4 kcal•mol^{-1} (Equation (1)). Therefore, our simple Monte Carlo model supports the argument that entropic confinement effects generate sufficiently strong forces to explain the maximum expected gain in UDP-Xyl binding affinity of −1.45 kcal•mol^{-1} (Figures A2.d, A3.c, Figure A.S5). More rigorous calculations on other systems using simpler polymer models (and simpler confinement geometries) also find confinement free-energy costs of the same magnitude.[32,33]

If the ID-tail favours the dynamics associated with the repacking of the allosteric switch into the E^{Ω} state, then we would expect to see a difference in the activation (E* to E) and inhibition kinetics (E* to E^{Ω}) (Figure A1.a). Pre-steady-state analysis of progress curves shows that the ID-tail slows the rate of activation hysteresis (E* to E) by 39% (Figure A4.a). Next, we examined the UDP-Xyl-induced isomerization of UGDH to the EΩ state. Transient-state analysis of UDP-Xyl binding kinetics revealed a three-phase exponential decay of UGDH time-resolved tryptophan fluorescence, and the data were globally fit by computer simulation (see Methods and Figure A.S7.a–e). The same kinetic model produced the best fit for both UGDH(FL) and UGDH(ΔID) and predicts UDP-Xyl affinities that are consistent with our steady-state inhibition studies (Figure A.S7):

$$E^{*} + UDX \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} E^{*}{:}UDX \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} E^{\dagger}{:}UDX \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E^{\Omega}{:}UDX \quad \text{(equation 2)}$$

Where k_n is the rate constant for reaction n. According to this model, UDP-Xyl binds to the E* state and induces two sequential isomerizations. On the basis of the allosteric model, we had expected a single isomerization from E* to the E^{Ω} state (Figure A1.a). We call the additional transient $E^{\dagger}$; it represents an intermediate between the E* and E^{Ω} states. The ID-tail changes the kinetic parameters of each transient observed in the time-resolved fluorescence (Figure A.S7.e). The largest effect of the ID-tail is a 4.4-fold enhancement of the initial UDP-Xyl binding step, corresponding to a −0.9 kcal•mol^{-1} gain in affinity (Figure A4.b). The kinetic model predicts an overall favourable gain in binding affinity of −1.3 kcal•mol^{-1}, which agrees well with the observed gain of −1.39 kcal•mol^{-1} (Figure A4.b, Table A.S2). The different stabilities of the corresponding UGDH(FL) and UGDH(ΔID) transients, combined with the fact that the ID-tail slows activation hysteresis and accelerates inhibition kinetics, supports our conclusion that the ID-tail alters the energy landscape to favour inhibition by UDP-Xyl (Figure A4.c).

Collectively, our data supports a model in which the entropic force of the ID-tail rectifies the energy landscape of UGDH to favour a substate with a high affinity for UDP-Xyl. We can now interpret the exponential curve in Figure A2.d as follows:

$$K_i(l) = K_i^{\text{unbiased}} e^{-kl} + K_i^{\text{biased}}(1 - e^{-kl}) \qquad \text{(equation 3)}$$

This implies that: *(i)* UGDH exists as an ensemble of low-affinity (K_i unbiased) and high-affinity (K_i biased) substates; *(ii)* the ID-tail functions as a length (l)-dependent entropic rectifier that shifts (with bias k) the distribution towards the high affinity substate; and *(iii)* the observed UDP-Xyl affinity results from a fractional summation of the low and high affinity substates at a given ID-tail length (Figure A4.c). The fit to equation (3) produces a K_i biased of 0.46 ± 0.18 μM UDP-Xyl, which corresponds to a maximum favourable gain in binding energy of approximately −1.45 kcal•mol^{-1}. The lack of sequence constraints implies that the entropic force of any intrinsically

disordered segment is capable of shaping the conformational ensemble of a protein. In fact, an N-terminal hexahistidine affinity tag has been shown to alter the internal dynamics of a myoglobin.[34] Thus, the persistence of low-complexity intrinsically disordered segments in the proteome may reflect the selection for entropic rectifiers that can tune the function of a protein by shaping the native-state ensemble.

A3 Materials and Methods

Protein expression, purification, and quantification of UGDH constructs

All UGDH coding sequences were synthesized and cloned into pET-15b vectors (Norclone). Sequences contained an N-terminal hexahistidine affinity tag adjacent to a tobacco etch virus (TEV) cleavage site. The expression and purification of UGDH constructs were conducted under identical conditions as previously described.[22-27] Following purification, the N-terminal hexahistidine tag was cleaved with TEV protease. An additional immobilized metal affinity column (IMAC) was used to obtain the pure, His-tag-free protein. Unless otherwise noted, all proteins were dialysed into a storage buffer (25 mM Tris pH 8.0 and 50 mM NaCl) and concentrated to ≥20 mg/ml. Proteins were quantified in dilution replicates ($n \geq 6$) using their respective molar extinction coefficients, based on their specific amino acid composition.[35]

Protein crystallization, data collection, and structure solution

To crystallize the E* conformation of UGDH(ΔID), the protein (10.4 mg/ml) was dialysed into 20 mM MES pH 5.6, 150 mM NaCl and crystallized at 20 °C using free interface diffusion in a 1.0-mm capillary containing 5 µl of 10.4 mg/ml enzyme and 200 µl of precipitant solution (100 mM MES pH 6.2, 100 mM MgCl2, and 16% PEG 3350). Crystals were cryoprotected in the precipitant solution supplemented with 18% glycerol and then plunged into liquid nitrogen. A 2.64 Å resolution dataset was collected on the 22-ID beamline (SER-CAT) at the Argonne National

Laboratory using an MAR 300-mm CCD detector. The data were processed in space group C2 using XDS[36], and 5% of the data were set aside for cross-validation[37]. The crystal parameters and data collection statistics are summarized in Table A.S1. The structure was solved by molecular replacement using the PHENIX software suite[38] and human UGDH (Protein Data Bank (PDB) entry: 3TF5) as a search model. The structure was then subjected to iterative cycles of manual rebuilding using Coot[39] and automated refinement using PHENIX with both NCS restraints.[38,40] *B*-factors were refined using TLS as implemented in PHENIX. Refinement statistics[41,42] are summarized in Table A.S1.

The E^{Ω} UGDH(FL) was crystallized in the presence of 5 mM UDP-Xyl and 10 mM adenosine diphosphate at 25 °C using the hanging drop vapour diffusion method. One microlitre of protein was mixed in a 1:1 ratio with reservoir solution (0.1 M HEPES pH 7.2, 14% 1,6-hexanediol, and 10% PEG 3350). Crystals were cryoprotected in the precipitant solution supplemented with 20% glycerol and then plunged into liquid nitrogen. A 2.0 Å resolution dataset was collected on the 21-ID beamline (SER-CAT) at the Argonne National Laboratory using a MAR 300-mm CCD detector. The dataset was processed using XDS[36] and 5% of the data were set aside for cross validation[37]. The data collection statistics are listed in Table A.S1. The E^{Ω} UGDH(FL) structure was solved by molecular replacement using the PDB entry 2Q3E as a search model in PHENIX[38], and refined as described above. Refinement statistics[41,42] are summarized in Table A.S1.

Steady-state kinetics

All steady-state kinetic assays were conducted as previously described.[22-27] In brief, assays contained either 100 nM UGDH (FL, FL-A136M, ΔID, ΔID-A136M, R1, R2, −Pro, −Lys, 0.13×FL, 0.26×FL, 0.5×FL, 2×FL, 0.13×Ser, 0.26×Ser, 0.5×Ser, Ser, 0.13×R1, 0.26×R1 or

0.5×R1) or 500 nM UGDH (FL-dimer, ΔID-dimer) in a standard reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA) with either saturating amounts of NAD^+ or UDP-Glc (Sigma). Substrate and enzyme were incubated separately at 25 °C for 5 min, and then reactions were initiated by rapid mixing of both solutions. Progress curves were obtained by continuously monitoring NADH production at 340 nm (molar absorptivity coefficient of 6,220 M^{-1}•cm^{-1}) on an Agilent 8453 UV/Vis spectrometer equipped with a Peltier temperature controller (25 °C). UGDH progress curves display hysteresis, thus the observed initial velocity (v_i) represents a transient and does not satisfy steady-state conditions. To obtain steady-state initial velocities (v_{ss}), progress curves before the depletion of 10% substrate were fit to Frieden's equation[43] as in previous studies[22,23,43]:

$$P(t) = v_{ss}t - \tau(v_{ss} - v_i)\left(1 - e^{-\frac{t}{\tau}}\right) \qquad \text{(equation 4)}$$

where P is the product produced at time t, τ is the relaxation time of the lag, and the length of the lag is $e\tau$. The v_{ss} was used for determination of UGDH steady-state kinetic parameters. Data were fit using nonlinear regression analysis in PRISM (GraphPad Software).

Because the UGDH(FL-A136M), UGDH(ΔID-A136M), UGDH(FL-dimer) and UGDH(ΔID-dimer) constructs do not exhibit hysteresis, the observed initial velocity was used for the determination of steady-state parameters as previously described.[22] UDP-Glc substrate saturation curves were fit to equation (5).

$$v_\circ = \frac{k_{cat}[E_t][S]}{K_M + [S]} \qquad \text{(equation 5)}$$

where v_0 is the initial steady state velocity (vss in equation (4)), E_t and S are the enzyme and substrate concentrations, respectively. As previously reported[22-24], the NAD^+ saturation curves of the UGDH hexameric enzyme display negative cooperativity and were fit to the sigmoidal rate equation (equation (6)):

$$v_\circ = \frac{k_{cat}[E_t][S]^h}{(K_{0.5})^h + [S]^h} \qquad \text{(equation 6)}$$

where $K_{0.5}$ is the half saturation point and h represents the Hill coefficient. The determination of the K_i for the allosteric inhibitor UDP-Xyl has been previously described.[22,23] In brief, data were globally fit to the model for competitive inhibition with cooperativity (equation (7)) using PRISM.

$$v_\circ = \frac{k_{cat}[E_t][S]^h}{(K_M^{app})^h + [S]^h} \text{ where } K_M^{app} = K_M\left(1 + \frac{[I]}{K_i}\right) \qquad \text{(equation 7)}$$

K_M, k_{cat}, and K_i were shared parameters in global fitting, whereas h was fit locally to each curve. The UGDH dimers (UGDH(FL-dimer) and UGDH(ΔID-dimer)) exhibited mixed inhibition with respect to both UDP-Glc and NAD^+, and were globally fit to equation (8).

$$v_\circ = \frac{(k_{cat}^{app})[E_t][S]}{(K_M^{app}) + [S]} \text{ where } k_{cat}^{app} = \frac{k_{cat}}{\left(1 + \frac{[I]}{(\alpha K_I)}\right)} \text{ and } K_M^{app} = K_M \frac{\left(1 + \frac{[I]}{K_i}\right)}{\left(1 + \frac{[I]}{\alpha K_I}\right)} \qquad \text{(equation 8)}$$

Here, K_i refers the competitive inhibition component, and αK_i gives the noncompetitive contribution. K_M, kcat, α and K_i were shared parameters for global fitting.

Sedimentation velocity

Sedimentation velocity analysis was conducted as previously described.[22-27] In brief, UGDH constructs were dialysed for >12 h at 4 °C into 25 mM HEPES pH 7.5 and 150 mM KCl and diluted to a final concentration of 9 μM. In ligand-bound studies, UGDH constructs were dialysed with comparable amounts of either substrate (UDP-Glc) or allosteric inhibitor (UDP-Xyl) for >24 h. Samples were loaded into cells equipped with 12-mm double-sector Epon centrepieces and quartz windows. The cells were then loaded into an An60 Ti rotor and equilibrated to 20 °C for 1 h. Sedimentation velocity data were collected at 50,000 r.p.m. in an Optima XLA analytical ultracentrifuge for 8–12 h. Data were recorded at 280 nm in radial step sizes of 0.003 cm. SEDNTERP[44] was used to estimate the partial specific volume of all UGDH constructs, and the buffer density (1.00726 g/ml) and viscosity (0.01018 P). SEDFIT[45] was used to model and fit all

data. Data were modelled as a continuous sedimentation coefficient (*c(s)*) distribution. The baseline, meniscus, frictional coefficient, and systematic time-invariant, and radial invariant noise were fit.[46] HYDROPRO[47] was used to predict s values based on crystal structures. The expected drag from the ID-tail was estimated by calculating the expected s values from crystal structures with and without modelled, energy minimized ID-tails. The data fits for all experiments can be found in Figure S2.3.

Evolutionary rate analysis

Seventy-nine UGDH sequences from vertebrates were used for analysis after removing redundancy at the organism level (only one UGDH sequence used per organism). The protein sequences were aligned using MUSCLE[48], and rates of evolution at each alignment position was calculated under the JTT model[49] using MEGA7 (log-likelihood method)[50]. The rates were normalized such that the average rate of evolution was 1.0 across the entire protein. Residue positions evolving faster than average show a rate greater than 1.0. In Figure S2.4, only the rates at alignment positions where the human UGDH did not have an indel were used.

Monte Carlo sampling

The free-energy cost of tethering an unstructured, non-interacting peptide to an impenetrable surface depends on the ratio of all constrained and unconstrained states:

$$\Delta G_{\text{constrained}} = -RT\, ln\left(\frac{\Omega_2}{\Omega_1}\right) \qquad \text{(equation 1)}$$

Where R is the gas constant, T is temperature, Ω_1 is the number of all possible states of an unconstrained, self-avoiding peptide and Ω_2 is the number of the Ω_1 states that do not conflict with the constraint imposed by the protein surface. To simplify, we used polyserine peptides, ignored side-chain entropy and used a hard sphere potential along with 166 coarse-grained ϕ, ψ bins to calculate Ω_1 and Ω_2. Each bin represents a $10 \times 10°$ range of ϕ, ψ values of peptide conformations

in the 'allowed' region of the original Ramachandran map (Figure A.S5.a&b). This calculation is nontrivial for large polymers, and an exhaustive grid search of all conformations was only conducted for the 3- and 4-residue ID-tails (Figure A.S5.c). We used the following Monte Carlo procedure to estimate the fraction of surface-constrained conformations (Ω_2/Ω_1) for each ID-tail. To determine the self-avoiding Ω_1 mesostates, we randomly assigned one of the 166 ϕ, ψ bins to each ϕ, ψ torsion angle in the ID-tail and then looked for steric clashes within the conformer using the 'outer limit' for atomic clashes as described in the original Ramachandran map.[51] Next, each of Ω1 mesostates was analysed for steric clashes with the surface or the adjacent ID-tail (Figure A.S5.d–l). Prior to the simulation, hydrogens were added to the hexamer structure using the 'reduce' program[52], and an adjacent ID-tail was modelled in an extended conformation and fixed during the simulation (Figure A.S5.d–f). The simulation was stopped when a minimum of 124,000 self-avoiding conformers were analysed and the ratio of surface-constrained conformations (Ω_2/Ω_1) reached convergence (Figure A.S5.c). The convergence threshold was defined as a change in the cumulative ratio of less than 10−5 within a window of 5,000 trials. All runs reached convergence except for the 10-mer simulations, which only converged to 2 decimal places (Figure A.S5.c–l). We estimated the accuracy in our Monte Carlo simulations by comparing the results to the full grid search of the 3- and 4-residue ID-tails (Figure A.S5.c).

Thermodynamic shift assay

Solutions of UGDH (FL-dimer or ΔID-dimer) at 0.1 mg/ml were prepared with 5× SYPRO Orange ThermoFluor (Thermo Fisher) in the standard reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA). Samples were then briefly spun and allowed to equilibrate for 20 min. The thermal denaturation experiments were conducted in replicates ($n \geq 3$) and data were acquired using a Bio-Rad MiniOpticon Real-Time qPCR machine. A fluorescence excitation

spectrum wavelength between 470–505 nm and an emission spectrum between 540–570 nm were used. The fluorescence emission for each solution was recorded every 30 s as the temperature was increased from 25 to 80 °C (ramp speed of 0.5 °C/s). Baselines were subtracted from the raw data using the buffer control experiments. The baseline, plateau and slope of the denaturation curve were fit to equation (9) to obtain the apparent T_m (melting temperature) values.[53]

$$Y = Baseline + \frac{Plateau - Baseline}{1 + 10^{\frac{T_m - X}{Slope}}} \qquad \text{(equation 9)}$$

where Y represents the fluorescence signal at temperature X.

Hydrogen–deuterium exchange–mass spectrometry

Studies have shown that hydrogen–deuterium exchange (HDX) is an appropriate probe for protein dynamics and can illuminate differences between wild-type and mutant proteins.[54,55] HDX is a powerful tool for footprinting the solvent-accessible regions of a protein[56], and was used in this study to compare structural and dynamic changes between the dimerized versions of UGDH (UGDH(FL-dimer) and UGDH(ΔID-dimer)).

Proteins were expressed and purified in the Wood laboratory as previously described.[22-27] Proteins were then flash-frozen and shipped overnight on dry ice to the Gross laboratory at Washington University in St. Louis for hydrogen–deuterium exchange–mass spectrometry (HDX–MS) analysis. Protein solutions (2 µl) were continuously labelled at 25 °C by adding 20 µl of 10 mM HEPES buffer containing 99.9% deuterium oxide (pD = 7.4). Samples were quenched by adding 33 µl of 8 M guanidine hydrochloride and 100 mM TCEP (final pH = 3.0) at 30 s, 1 min, 2 min, 15 min, 1 h, and 2 h time points[57,58]. One minute after quenching, samples were flash-frozen in liquid nitrogen and stored for less than 36 h at −80 °C. Control samples contained 10 mM HEPES in water rather than deuterium oxide. Each sample was thawed seconds before liquid chromatography followed by mass spectrometry (LC–MS). On-line protein digestion was

performed with a custom-packed pepsin column (2 mm × 20 mm) at a flow rate of 200 µl/min in 0.1% trifluoroacetic acid. For desalting, a Zorbax Eclipse XDB-C-8 trap column (2.1 × 15mm, 3.5 µm) was used to trap peptic peptides for 3 min. Following this, peptides were separated using a Hypersil Gold C-18 analytical column (2.1 × 50 mm, 2.5 µm), 4–80% gradient of acetonitrile with 0.1% formic acid (B), and a 100 µl/min flow rate. Peptides were detected using a LTQ XL Orbitrap mass spectrometer (Thermo Fisher Scientific), with a mass resolving power of 50,000, m/z 400. Additional parameters were spray voltage of 5 kV, capillary temperature of 275 °C, capillary voltage of 49 V, and a tube lens of 163 V. All experiments were conducted in quadruplicate.

As a prelude to HDX, protein mapping was performed by identifying pepsin-digested peptides. Product-ion mass spectra were collected in the data-dependent mode, picking the six most abundant ions from selected MS/MS. Peptides were identified using Mascot (Matrix Science). Following HDX, mass spectra were analysed with HDX Examiner (Sierra Analytics). The per cent deuterium uptake was plotted against time for UGDH(FL-dimer) and UGDH(ΔID-dimer). To magnify slight, yet significant changes in uptake, the cumulative differences in HDX for UGDH(FL-dimer) versus UGDH(ΔID-dimer) were calculated. These values were plotted alongside 3 times the error propagation for all measurements of both variants for each peptide, after the data and error were normalized—divided by the number of time points considered for each data point (Figure A.S6). The propagation error for each peptide is equal to the square root of the sum of all squared standard deviation values for collective time-dependent measurements of UGDH(FL-dimer) and UGDH(ΔID-dimer). The cumulative per cent deuterium uptake was compared to 3 times the propagation error. Differences that were greater than 3 times the propagation error were noted as regions of change affected by the presence of the ID-tail. We chose to normalize the data to be more inclusive of peptides with low intensity that are found at

most time points. In a similar manner, we have excluded those peptides that have avoided detection for more than two time points.

Stopped-flow analysis of UGDH hysteresis

The allosteric activation (E* to E) of UGDH can be observed as a lag (hysteresis) in progress curves[22,23] (see Figure S2.7f for examples). The allosteric activation rates for UGDH(FL) ($n \geq 6$) and UGDH(ΔID) ($n \geq 6$) were monitored at 25 °C using an Applied Photophysics SX20 stopped-flow spectrophotometer. Enzyme solutions contained 500 nM UGDH(FL) or UGDH(ΔID) in the standard reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA). This solution was rapidly mixed with an equal volume of standard reaction buffer that contained both substrate and cofactor. The mixed solution contained 250 nM UGDH(FL) or UGDH(ΔID), with saturating amounts of both substrate and cofactor. The progress of the reaction was monitored by NADH production, with the absorbance reading at 340 nm being acquired every 10–15 ms. Progress curves were fit to equation (4) to determine the length of the lag in enzyme activation (E* to E). The mean and standard deviation of the hysteretic lags were derived from 6 or more progress curves.

Transient-state kinetics of UDP-Xyl binding

Stopped-flow fluorescence studies were conducted at 25 °C using an Applied Photophysics SX20 stopped-flow spectrophotometer with a dead time of ~1.2 ms. Syringes were loaded with 500 nM UGDH(FL) or UGDH(ΔID) and variable concentrations of UDP-Xyl, and then rapidly mixed. The change in intrinsic tryptophan fluorescence was continuously monitored using an excitation wavelength of 290 nm and an emission filter with a cut-off below 320 nm (Figure A.S7). Fluorescence decay curves were averaged from experimental replicates ($n \geq 4$) for each concentration in the series. Raw data was corrected for the inner-filter effect using the molar

absorptivity at both the excitation and emission of UDP-Xyl.[59] Data were globally fit using computer simulation with KinTek Global Kinetic Explorer program[60,61] (KinTek). Multiple input models based on the known structural states were tested, and the best fit model was determined using confidence contour analysis[62]. Microscopic rate constants and errors are reported in Figure A.S7.e. Fit data and confidence contours can be found in Figure A.S7.a–d.

A4 Supplementary Information

Section 1. hUGDH Dimer [FL and ΔID] Mixed Inhibition

In contrast to hexameric hUGDH, both hUGDHFL-dimer and hUGDHΔID-dimer display mixed inhibition kinetics with UDP-Xyl (Figure A.S2). Briefly, the mixed inhibition occurs because UDP-Xyl binds strongly to the allosteric/substrate-binding site, but can also bind weakly to the coenzyme-binding site. This occurs in both the dimers and the native hexamer (see Figure A.S1.e-l). However, we do not observe mixed inhibition in the native hexamer under steady state conditions because the coenzyme binds with a K_d of 30 uM[24] and our inhibition studies contain a saturating concentration of NAD^+ (6 mM). This means that the low concentration of UDP-Xyl (<5 μM) in these assays cannot outcompete NAD^+ binding in the coenzyme site. In contrast, the NAD^+ K_M of the dimers is roughly two orders of magnitude weaker than that of the hexamer, which makes it difficult to saturate the enzyme (Table A.S3). Thus, in the dimer we observe mixed inhibition, as we are able to detect both the tight binding of UDP-Xyl at the allosteric site and the much weaker binding at the coenzyme site (Tables A.S2 and A.S3). The $\alpha > 1$ shows that the mixed inhibition is dominated by UDP-Xyl competing with UDP-Glc for the allosteric site. Using NAD^+ as the competing substrate, we confirmed that the mixed inhibition was a result of UDP-Xyl weakly binding to the coenzyme binding site in the $hUGDH_{dimers}$ (Figure A.S2 and Table A.S3).

A5 Acknowledgements

The authors thank A. P. Karplus, B. W. Matthews, S. N. Savvides and the members of the Z.A.W. laboratory for helpful discussions. We also thank R. Wang of Norclone for producing the R1 truncation constructs. This work was supported by the NIH National Institute of General Medicine grants R01GM114298 awarded to Z.A.W. and P41GM103422 awarded to M.L.G.

A6 Main Text Figures

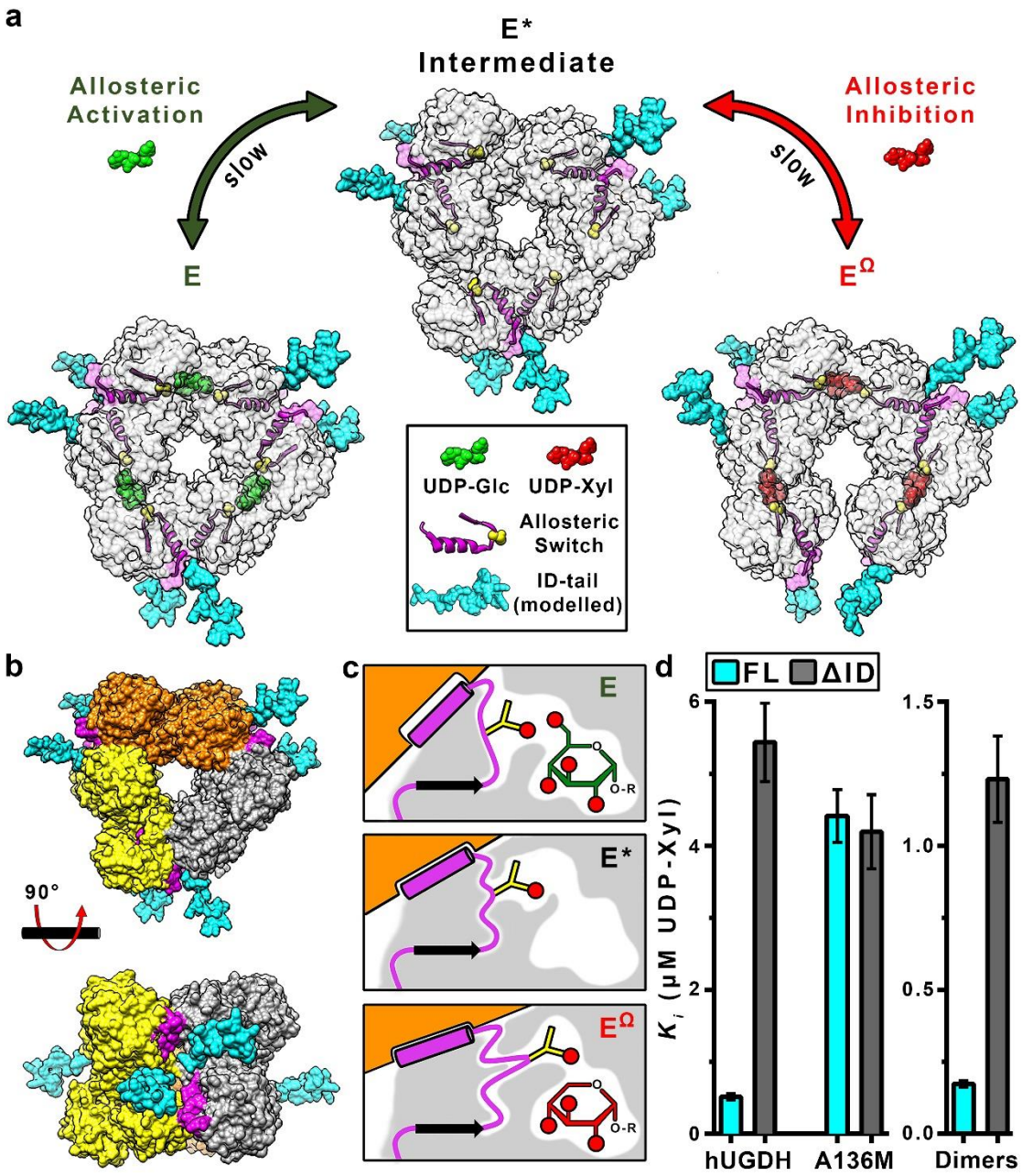


Figure A1: The Role of the ID-Tail in Allosteric Inhibition of UGDH. a, Unliganded UGDH forms an inactive (E*) hexamer. UDP-Glc (green) induces the Thr131-loop–α6 allosteric switch (yellow spheres and magenta ribbons and surface) to slowly isomerize into the active (E) state. UDP-Xyl (red) competes with UDP-Glc for the active site, and induces the allosteric switch to slowly isomerize into the inhibited (E^{Ω}) state. The slow isomerizations are due to the repacking of the allosteric switch in the protein core.[22,25-27] Because the ID-tail is disordered in the E, E* and E^{Ω} states (Figure A.S1 and refs[22,23,25,26]), we have modelled energy-minimized conformations of

the ID-tail (cyan) onto the structures of UGDH to depict the proximity to the active site, hexamer-building interface and the allosteric switch. **b,** Top and side view of the UGDH E* hexamer that forms from the association of three dimers (orange, grey and yellow).[22-27] The ID-tail of each dimer is located near two allosteric switches in the hexamer-building interface. **c,** The allosteric switch (magenta) is buried in the protein core (grey shading), which changes conformation in the E, E* and E^{Ω} states. Thr131 (yellow sticks) responds to the presence or absence of the C6′ OH in UDP-Glc (green) or UDP-Xyl (red), respectively. This response shifts the α6 helix (magenta cylinder) in the hexamer-building interface, which rotates the adjacent subunit (orange) to produce the E or E^{Ω} hexamer, as appropriate. Red circles depict hydroxyl (OH) groups. **d,** The UDP-Xyl affinity depends on the ID-tail and allostery. Data are the globally fit $K_i \pm$ s.e.m. derived from two or three independent rate curves with varying amounts of inhibitor ($n \geq 31$ independent data points; see Table S2.2 for specific values).

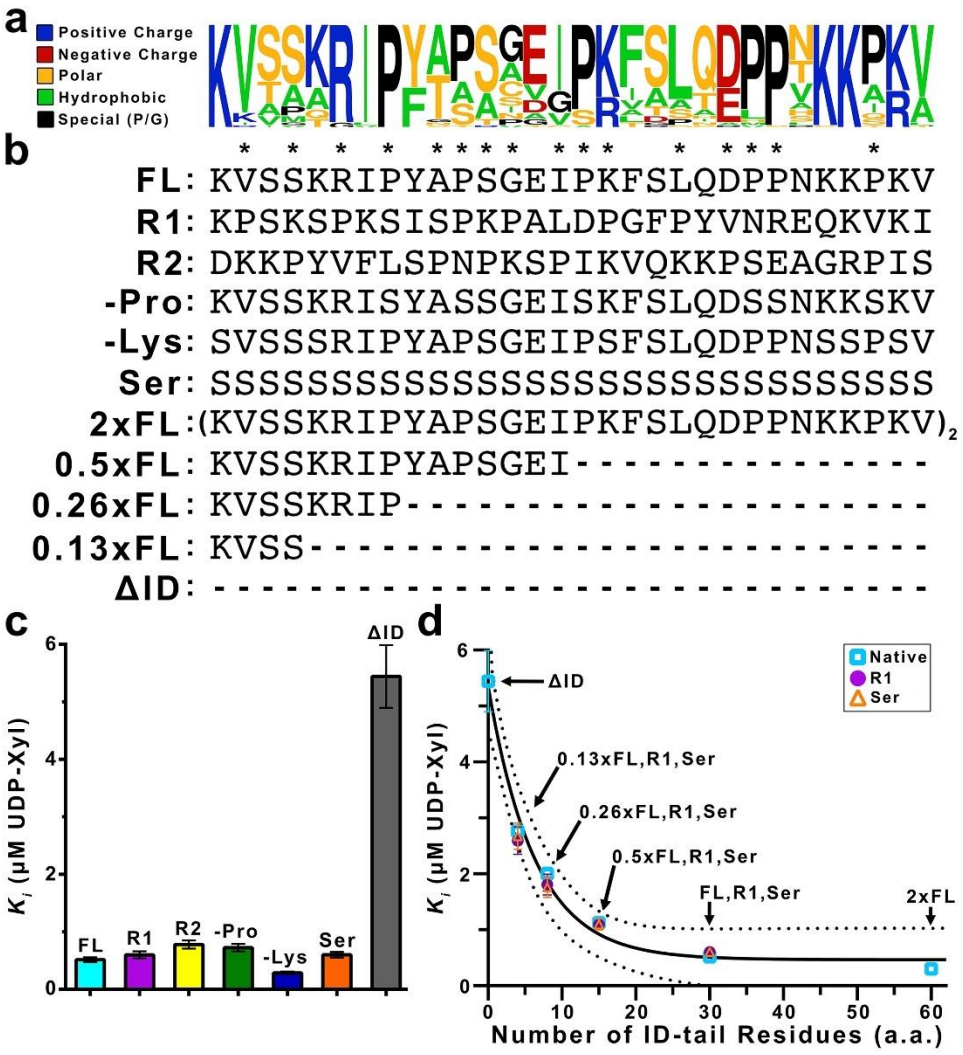


Figure A2: Structural Constraints of the ID-Tail. a, Alignments of the 30-residue ID-tail sequence (residues 465–494) from 79 vertebrate UGDHs (Figure A.S4.a&b). Residues are coloured according to type, and the height of each residue represents the relative frequency. Alignments were generated using the WebLogo server (http://weblogo.berkeley.edu). **b,** Sequence modifications made to the primary structure of the ID-tail (Figure A.S4.b). Asterisks indicate positions in the sequence that are sampled with a proline residue in either UGDH(FL), UGDH(R1) or UGDH(R2). **c,** UDP-Xyl affinity is independent of the ID-tail sequence. Data are the globally fit $K_i \pm$ s.e.m. derived from two or three independent rate curves with varying amounts of inhibitor. See Table A.S2 for the specific number of independent data points ($n \geq 27$). **d,** The affinity for UDP-Xyl depends on the length of the ID-tail. Data are the globally fit $K_i \pm$ s.e.m. derived from three independent rate curves with varying amounts of inhibitor ($n \geq 38$ independent data points; see Table A.S2 for specific values). For some points the s.e.m. is smaller than the data label. The data were fit to equation (3) (solid line) with 95% confidence intervals indicated (dashed lines). The fit predicts a maximum affinity of 0.46 ± 0.18 μM, corresponding to a free-energy change of -1.45 kcal•mol^{-1}.

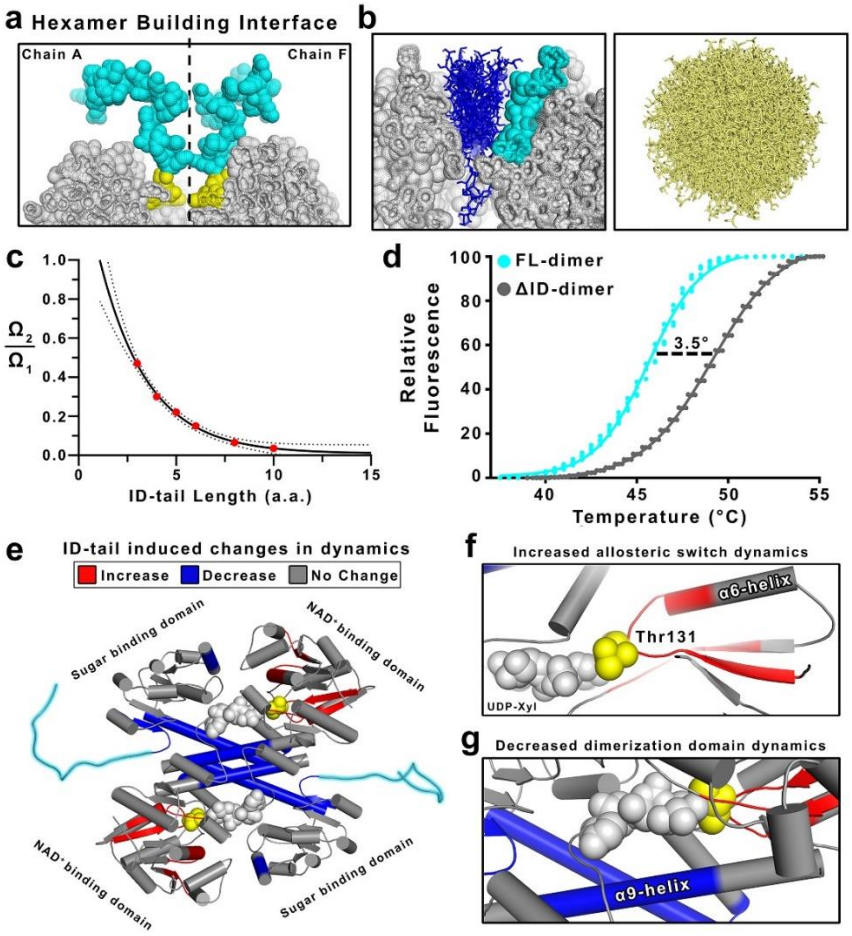


Figure A3: The Entropic Force of the ID-Tail Alters the Structure of UGDH. a, Cut-away of the UGDH surface (grey spheres) at the hexamer-building interface (dashed lines), depicting the modelled ID-tails (cyan and yellow spheres) from adjacent subunits (grey, chains A and F). The volume-exclusion effect of the hexamer-building interface tightly constrains the conformations of the first four disordered residues (465–468) of the ID-tail (yellow). **b,** Left, a representative subset of the surface-constrained conformations of a 10-residue ID-tail (blue sticks) from Monte Carlo sampling (see Methods for details). The adjacent ID-tail is shown as cyan spheres. Right, a representative sampling (tan sticks) of accessible conformations without surface constraints (see also Figure A.S5). **c,** The fraction of constrained ID-tail conformations (Ω_2) over the possible conformations of a free ID-tail (Ω_1) exponentially converges with increasing ID-tail length. The data were fit to an exponential decay (Figure A.S5.c). **d,** The ID-tail destabilizes UGDH by 3.5 °C. **e,** Comparing HDX rates of UGDH(FL-dimer) and UGDH(ΔID-dimer) shows that the ID-tail (cyan) alters the structure and dynamics of UGDH. Peptides displaying increases (red), decreases (blue) and no change (grey) in HDX rates are mapped to the structure. UDP-Xyl (grey spheres) was not used in the assay but is modelled in the active site. Thr131 of the allosteric switch is shown as yellow spheres. **f,** Close-up view of the allosteric switch (Thr131–α6 helix), which shows an increase in HDX rates. **g,** Close-up view of the of the dimerization domain, which is largely inaccessible to solvent. Data shown in e–g were derived from the normalized cumulative per cent deuterium uptake (%D) comparing UGDH(FL-dimer) and UGDH(ΔID-dimer) (Figure A.S6).

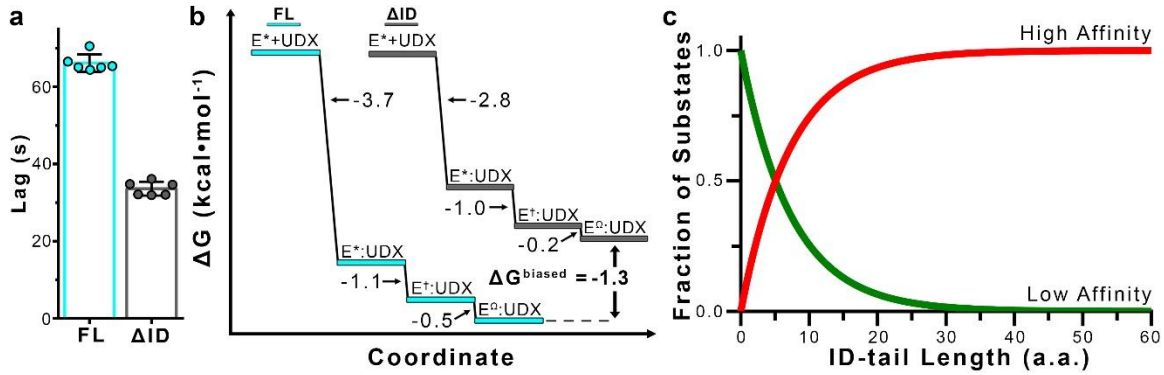


Figure A4: The ID-Tail Shapes the Conformational Landscape of UGDH. a, The ID-tail increases the hysteresis of allosteric activation (E* to E). Data are mean ± s.d. (n = 6 independent experiments). **b,** Free-energy plot of the transient-state kinetic model for the allosteric inhibition with UDP-Xyl (UDX) (equation (2)) of UGDH(FL) (E*FL) and UGDH(ΔID) (E*ΔID). The free energy for the initial binding step (second order) was determined using the standard state defined in units of mM, and the two isomerization steps (first order) were determined from the equilibrium constants K_1, K_2 and K_3 defined in Figure A.S7.e. **c,** Plot of the high- and low-affinity fractional components of equation (3), showing that the shift in states is a function of length of the ID-tail.

A7 Supplementary Figures and Tables

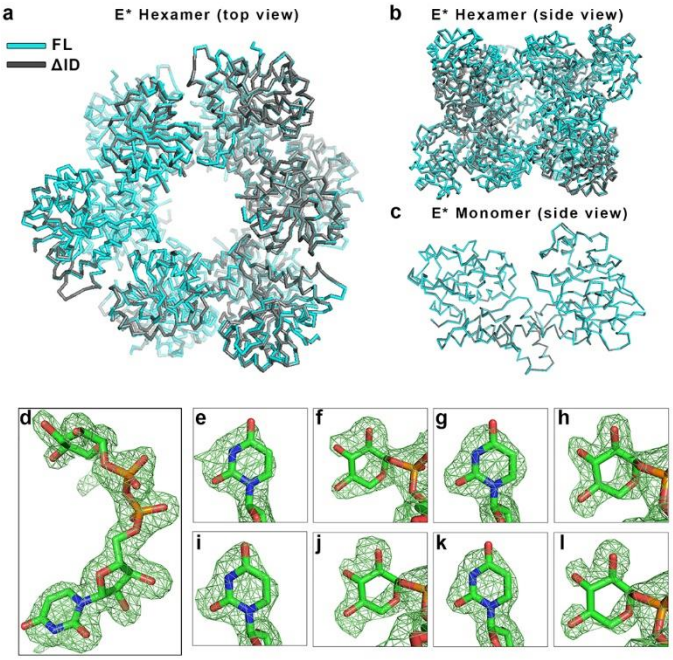


Figure A.S1: The Crystal Structures of UGDH(FL) and UGDH(ΔID) Show no Significant Structural Deviations, and Structural Evidence for UDP-Xyl Binding in the NAD$^+$ Site. a–c, Structural overlay (root mean square deviation (r.m.s.d) = 0.385 Å), comparing the UGDH(FL) (cyan) and UGDH(ΔID) (grey) E* hexamers (a, b) and monomers (c). PDB entries for UGDH(FL) and UGDH(ΔID) are 4RJT and 5W4X, respectively (Table S2.1). **d,** Crystal structure of native UGDH with UDP-Xyl bound in the active site. Difference density map ($F_0 - F_c$) of UDP-Xyl (chain B) calculated at 2.0 Å resolution and contoured at 3.5σ. The map was calculated after omitting the UDP-Xyl and subjecting the model to simulated annealing. **e–l,** UDP-Xyl can also bind weakly to the NAD$^+$-binding site of native UGDH. Difference electron density maps ($F_0 - F_c$) were calculated as in d. The uracil and xylose in the NAD$^+$-binding sites were contoured at 3.5 and 3σ for chain A (e and f, respectively), chain B (g and h, respectively), chain D (i and j, respectively) and chain E (k and l, respectively). Chains C and F do not contain UDP-Xyl in the NAD$^+$-binding site. UDP-Xyl binding in the NAD$^+$ site is the source of mixed inhibition observed in the UGDH(FL-dimer) and UGDH(ΔID-dimer) constructs. (see Supplementary Information, Section 1). PDB entry: 5VR8 (this work, Table A.S1).

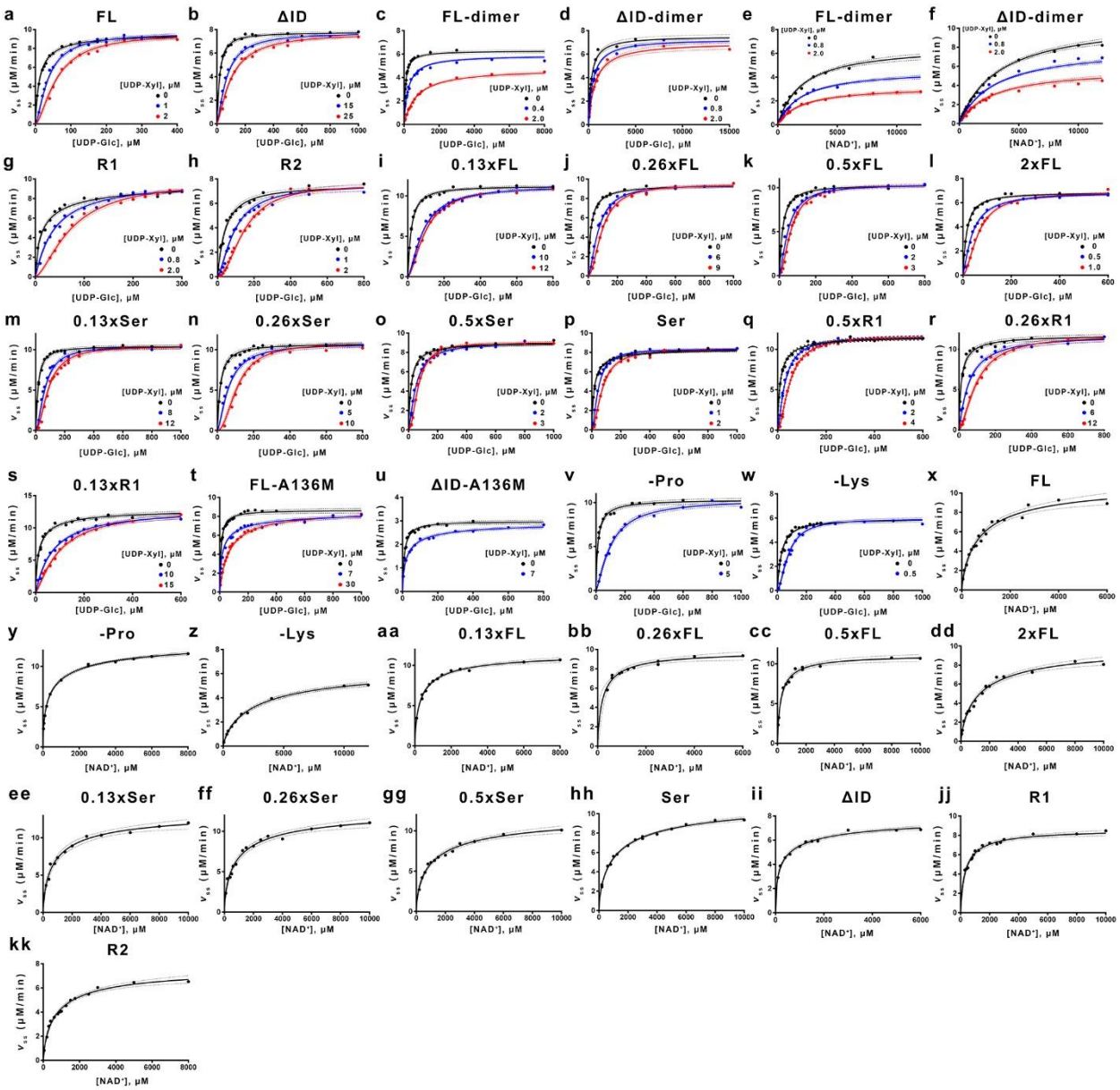
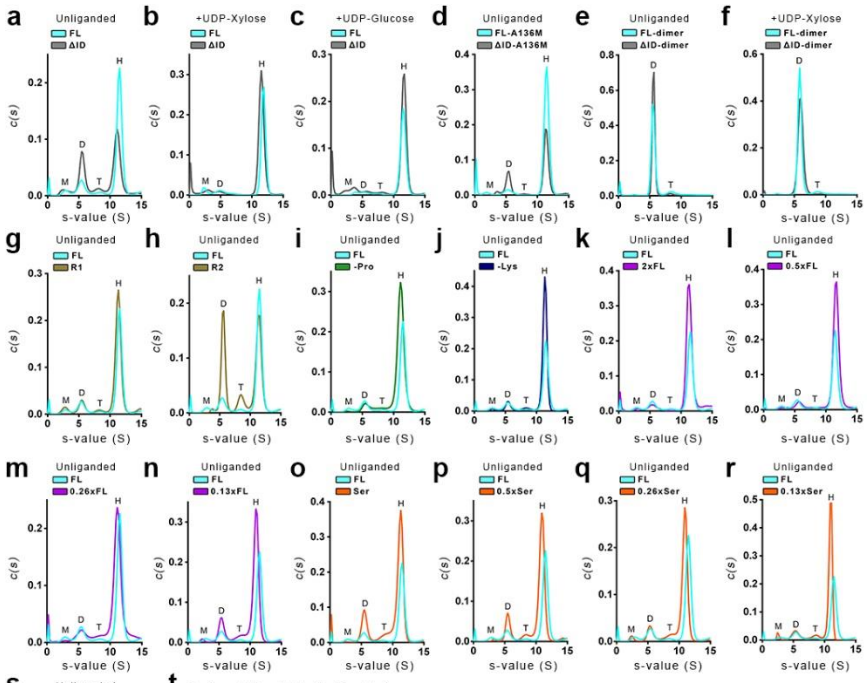


Figure A.S2: Steady-State Kinetic Analysis of All UGDH Constructs. a–w, Inhibition studies with the allosteric inhibitor UDP-Xyl. Data from two or three independent rate curves were globally fit to equation (7) (or equation (8) for dimers c–f) using nonlinear regression ($n \geq 26$ data points). See Table S2.2 for the specific number of data points and fit parameters. Dashed lines indicate 95% confidence intervals. **x–kk,** NAD^+ substrate-saturation curves fit to equation (6) using nonlinear regression ($n \geq 10$ independent data points). See Table S2.3 for the specific number of data points used in global fitting.



t Sedimentation Velocity Results[a]

hUGDH	Ligand	M[b] S (%)	D[b] S (%)	T[b] S (%)	H[b] S (%)	Predicted S-Value	R.M.S.D[a]
FL	--------	3.0 (4.1)	5.5 (12.6)	8.5 (3.4)	11.5 (69.2)	----------	0.005113
FL	UDX	2.5 (5.5)	5.4 (8.6)	0.0 (0.0)	11.9 (70.9)	11.2[c]	0.003801
FL	UDG	0.0 (0.0)	5.7 (3.7)	0.0 (0.0)	11.5 (68.2)	11.7[c]	0.005448
ΔID	--------	2.9 (6.4)	5.6 (29.4)	8.3 (7.7)	11.1 (50.9)	----------	0.003784
ΔID	UDX	2.9 (6.0)	5.6 (5.0)	0.0 (0.0)	11.6 (78.3)	11.5[c]	0.008909
ΔID	UDG	3.3 (9.6)	5.9 (3.9)	8.3 (3.0)	11.6 (67.8)	12.0[c]	0.004379
FL-A136M	--------	0.0 (0.0)	5.4 (6.1)	0.0 (0.0)	11.5 (80.9)	----------	0.009808
ΔID-A136M	--------	3.8 (3.6)	5.4 (22.9)	8.0 (2.0)	11.4 (63.6)	----------	0.003788
FL-dimer	--------	2.5 (1.1)	5.6 (78.9)	8.8 (9.7)	0.0 (0.0)	5.6[d]	0.009960
FL-dimer	UDX	0.0 (0.0)	5.9 (88.6)	9.0 (5.8)	0.0 (0.0)	----------	0.004547
ΔID-dimer	--------	0.0 (0.0)	5.6 (90.9)	8.5 (4.0)	0.0 (0.0)	5.7[d]	0.004564
ΔID-dimer	UDX	0.0 (0.0)	6.0 (91.3)	0.0 (0.0)	0.0 (0.0)	----------	0.008205
R1	--------	2.9 (4.9)	5.6 (9.7)	8.3 (3.8)	11.3 (70.5)	----------	0.003979
R2	--------	3.9 (1.3)	5.6 (33.9)	8.6 (10.3)	11.4 (47.2)	----------	0.005785
-Pro	--------	0.0 (0.0)	5.8 (7.6)	8.0 (3.6)	11.1 (86.4)	----------	0.013871
-Lys	--------	3.0 (1.2)	5.5 (8.0)	8.4 (4.3)	11.3 (84.8)	----------	0.004136
2×FL	--------	3.1 (1.6)	5.7 (6.8)	0.0 (0.0)	11.4 (75.7)	----------	0.011126
0.5×FL	--------	3.4 (2.2)	5.8 (7.2)	8.4 (2.6)	11.7 (84.4)	----------	0.005328
0.26×FL	--------	0.0 (0.0)	5.5 (10.0)	0.0 (0.0)	11.3 (84.1)	----------	0.000772
0.13×FL	--------	2.4 (1.4)	5.5 (14.8)	0.0 (0.0)	10.9 (75.6)	----------	0.003486
Ser	--------	0.0 (0.0)	5.5 (17.4)	0.0 (0.0)	10.9 (77.7)	----------	0.004350
0.5×Ser	--------	3.4 (1.0)	5.6 (14.2)	8.3 (5.9)	10.9 (73.6)	----------	0.003647
0.26×Ser	--------	2.5 (2.4)	5.5 (10.8)	0.0 (0.0)	10.9 (77.8)	----------	0.003143
0.13×Ser	--------	2.5 (2.8)	5.5 (11.1)	8.5 (6.4)	10.9 (80.2)	----------	0.003144
K465D	--------	2.5 (4.1)	5.4 (21.7)	8.3 (2.8)	11.6 (64.5)	----------	0.008206

[a] Sedimentation values, relative distributions, and root-mean-square deviation (R.M.S.D) were derived from fit data to greater than 65 independent sedimentation profiles.
[b] M = Monomer; D = Dimer; T = Tetramer; H = Hexamer
[c,d] HYDROPRO predicted S value for the hexamer[c] and the dimers[d]

Figure A.S3: Sedimentation Velocity Studies of the UGDH Constructs. a–s, Plots of the c(s) distributions with oligomeric species labelled as H (hexamer), T (tetramer), D (dimer) or M (monomer). The R2 mutant (panel h) shows no change in UDP-Xyl affinity (Figure A2.c and Table A.S2), yet shows evidence of a less stable hexamer. Panel s was included to show that the hexamer in h is less stable partly owing to the K465D substitution in the UGDH(R2) construct. The K465D substitution introduces an unfavourable negative charge near E460 in the hexamer interface, which may reduce the stability. **t,** Relative distributions, s values (S) and r.m.s.d. values for all sedimentation velocity experiments.

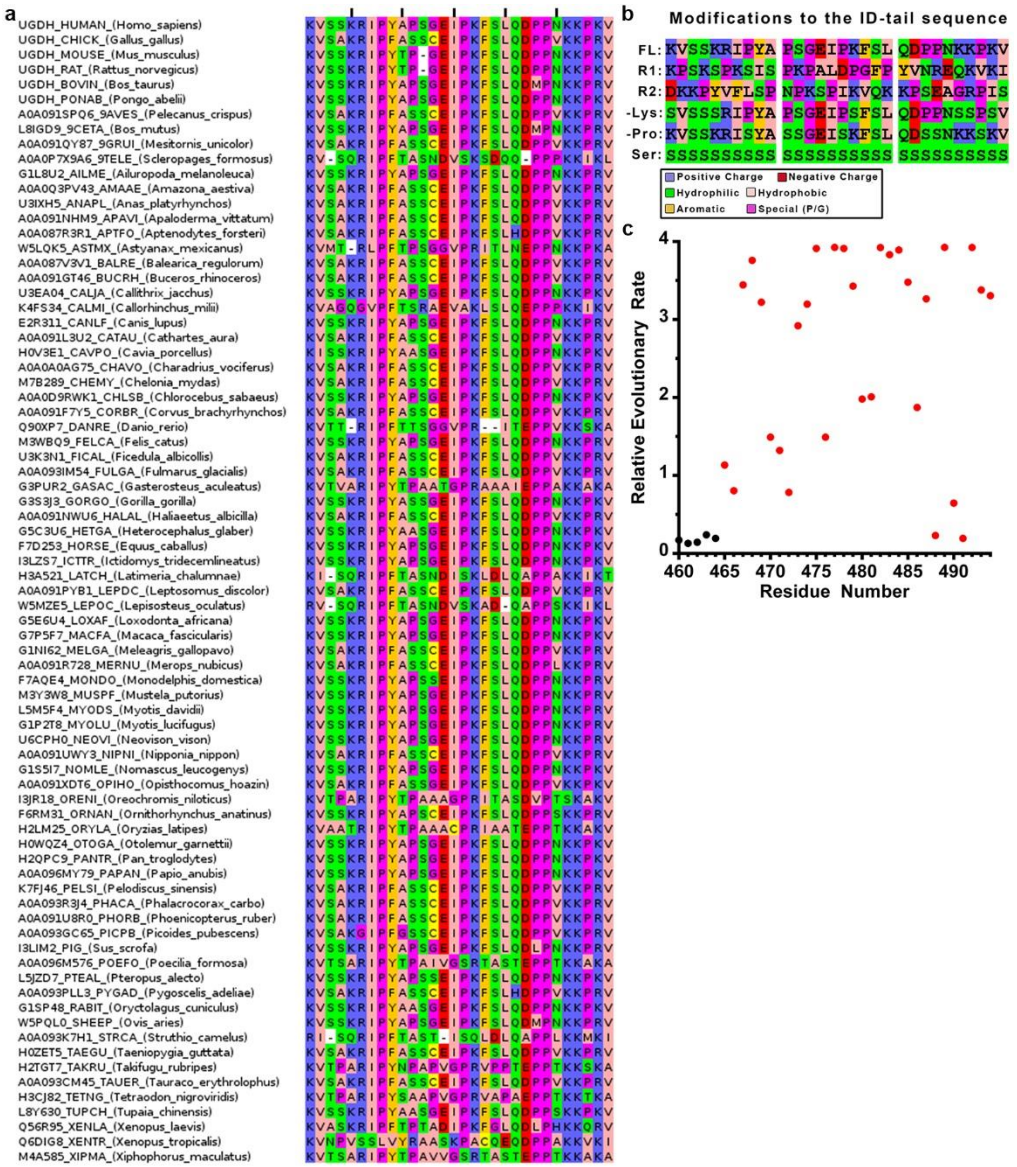
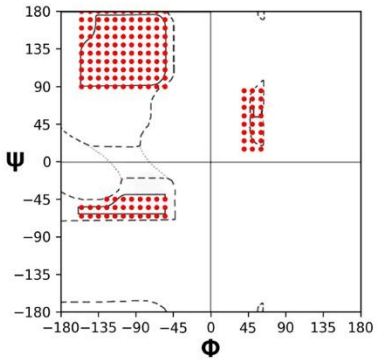


Figure A.S4: The ID-Tail is Conserved in Vertebrates. a, ClustalO sequence alignment of all vertebrate UGDH ID-tail regions (79 total). Residues are coloured by type, where blue is positive charge (K, R, H), red is negative charge (D, E), peach is hydrophobic (A, V, L, I, M), orange is aromatic (F, W, Y), green is hydrophilic (S, T, N, Q), yellow is cysteine (C), and magenta is special (P, G). **b,** The ID-tail was extensively randomized and modified. Sequences of UGDH (FL, R1, R2, −Lys, −Pro, and Ser), aligned by position and coloured by residue type. **c,** Relative evolutionary rate of UGDH residues from the alignment of 79 vertebrate sequences. The ID-tail (red dots) begins at residue 465 and displays an approximately threefold higher rate of divergence than the folded portion of the protein (black dots). For clarity, only a small, representative segment of the folded protein is shown (residues 460–464). All rates were scaled such that the average rate is 1.0 across the entire dataset.

a



b Values for the Φ,Ψ ranges used in Monte Carlo Simulations

Allowed Regions	ϕ-range	ψ-range	Number of Conformations
I	-155° to -55°	90° to 180°	110
II	-125° to -55°	-65° to -45°	24
III	-155° to -125°	-65° to -55°	8
IV	40° to 60°	15° to 85°	24
			166 (total)

c Monte Carlo Simulation Results

Tail	Ω_2/Ω_1 (M.C.) [1]	Ω_1	Ω_2/Ω_1 (Exhaustive Search)[2]	$\Delta G_{constrained}$ (kcal•mol^{-1}) [3]
3	0.47	124075	0.52	0.446
4	0.30	124876	0.29	0.709
5	0.22	125697	-	0.905
6	0.15	124838	-	1.134
8	0.06	479271	-	1.612
10	0.03	150112	-	1.98

[1] Monte Carlo calculation (see text and methods). Ω_1 and Ω_2 are the numbers of unconstrained and surface-constrained conformations, respectively.
[2] The difference between the grid search and the MC results are a measure of accuracy of the latter
[3] *Equation 1* (see text)

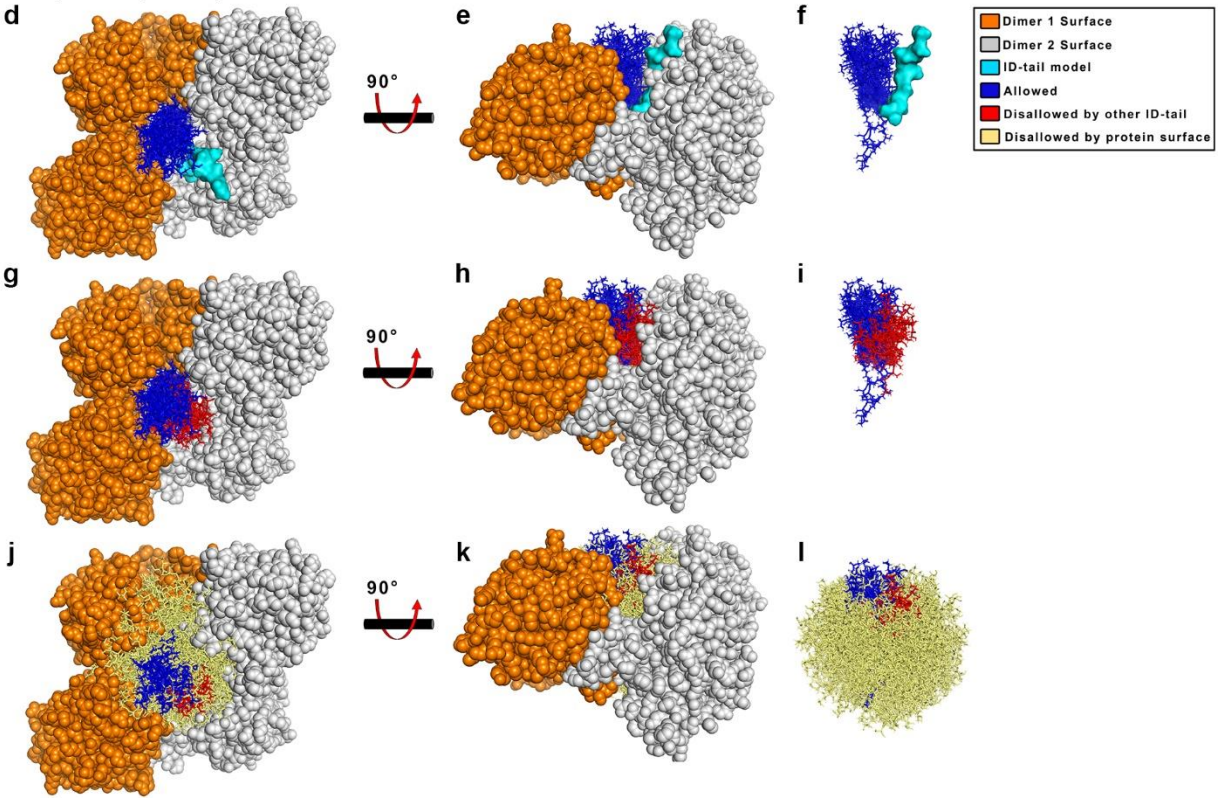


Figure A.S5: Exhaustive Monte Carlo Simulations Constraining the ID-Tail. a, Dashed lines outline the traditional, generously allowed regions of the Ramachandran plot, whereas the red circles identify the conformations used in the Monte Carlo simulations. **b,** The ranges of ϕ and ψ

angles depicted in a. The 10 × 10° bins are centred on the first and last numbers in the range. For example, in region 1, the first ϕ, ψ bin (−155°, 90°) represents the ϕ range −155° to −145° and the ψ range 85° to 95°. **c,** Ratio of ID-tail conformations constrained (Ω_2) to the number of conformations when the ID-tail is unconstrained (Ω_1). The entropic costs of confining tails of each length were calculated using equation (1). **d-e** The results of the 10-residue ID-tail simulations, shown in a surface representing the hexamer-building interface (orange and grey dimers) with the adjacent ID-tail (cyan) that was fixed during simulations. Also depicted is a representative sampling of 20 allowed Ω_2 conformations (blue sticks) from the 4,503 identified in the Monte Carlo simulation. **f,** The same view as in e, but without the protein surface. **g–i,** Same as in d–f, but now including a sampling of 20 of the 3,002 Ω_1 conformations (red sticks) that clash with the fixed adjacent ID-tail (not depicted for clarity). **j–l**, Same as g–i, but including 750 of the 142,607 Ω_1 conformations (tan sticks) that clash with the protein surface.

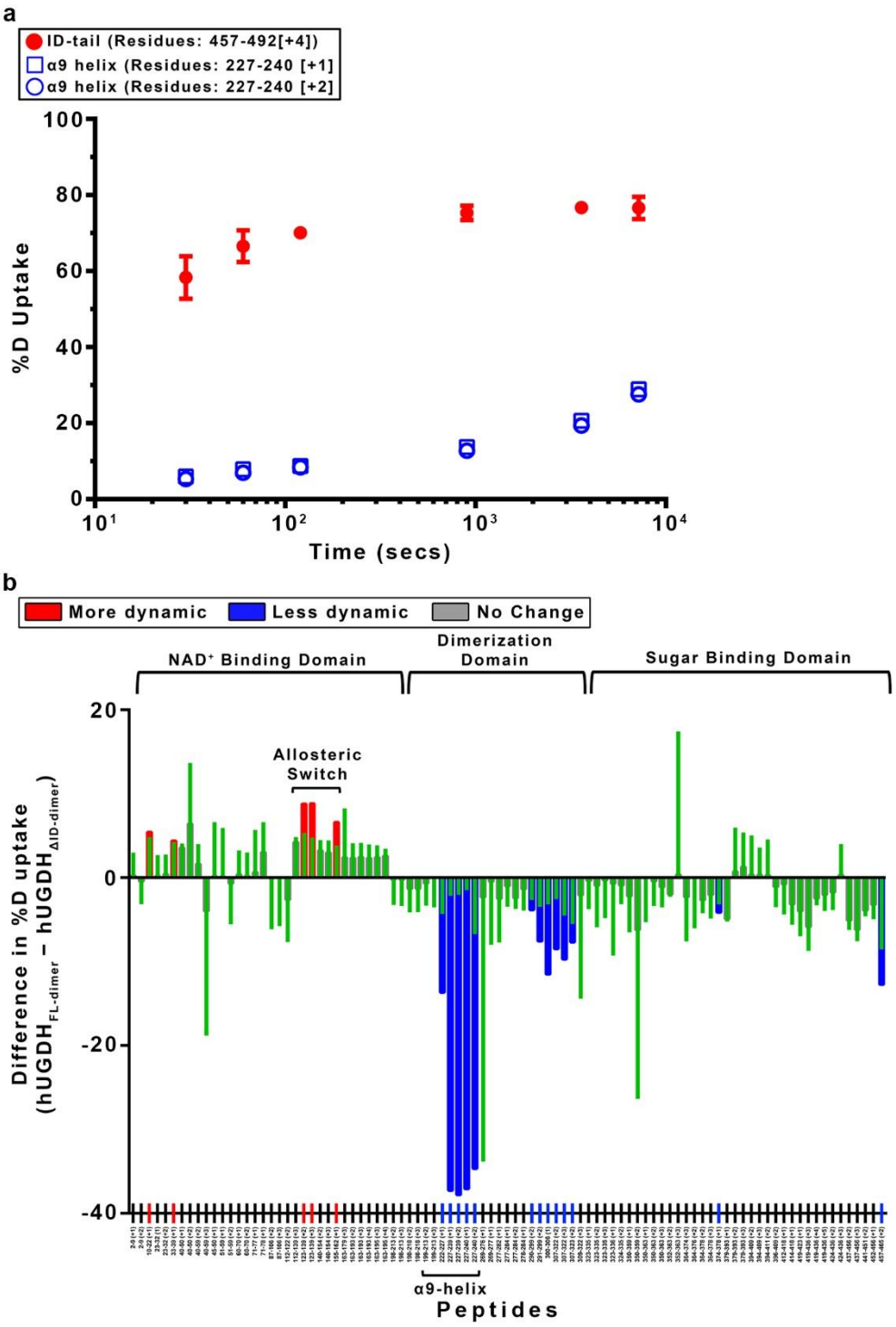
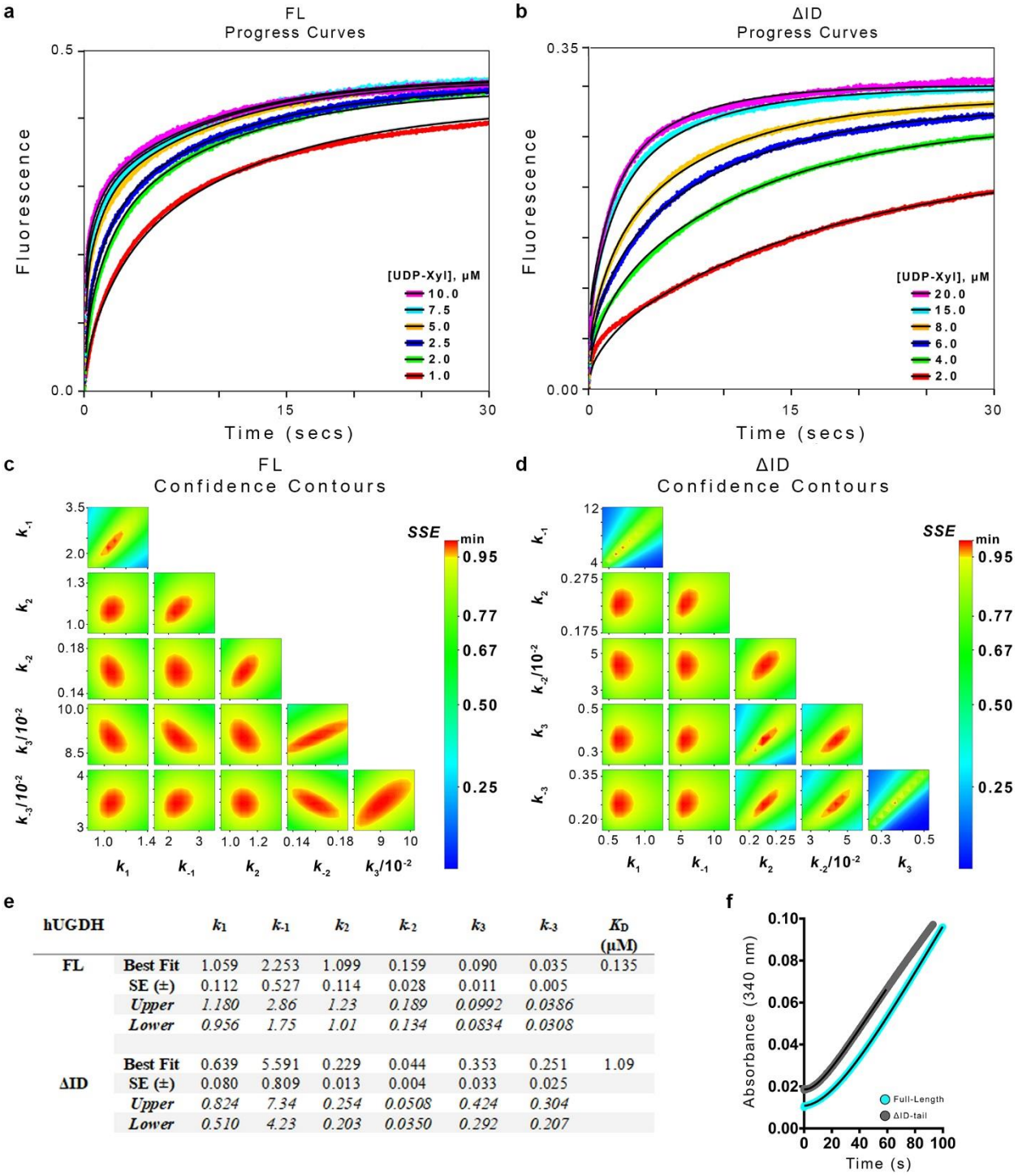


Figure A.S6: The ID-Tail Induces Global Changes in the Structure and Dynamics of UGDH. **a,** The per cent deuterium uptake of the ID-tail peptide region (residues 457–492; red closed circles) saturates rapidly, consistent with an unfolded peptide[31]. For comparison, two peptides corresponding to the well-ordered α9 helix region (open blue squares and circles) saturate slowly. Data are mean ± s.d. of independently replicated time points ($n = 4$). For some points, the standard deviation is less than the dimensions of the data symbol. **b,** The normalized cumulative changes in

the hydrogen–deuterium exchange rates (UGDH(FL-dimer) –UGDH(ΔID-dimer)). Most of the kinetics measurements consisted of six independently replicated time points (n = 4), processed to give the mean exchange (red, blue or green bars). Approximately 5% of the data displayed low signal:noise or was missing, and in those cases the means were derived from four or more time points. Results were normalized by dividing by the number of measurements. The propagation error for each peptide is equal to the square root of the sum of all squared standard deviation values for the collective measurement of UGDH(FL-dimer) and UGDH(ΔID-dimer).



e

hUGDH		k_1	k_{-1}	k_2	k_{-2}	k_3	k_{-3}	K_D (µM)
FL	Best Fit	1.059	2.253	1.099	0.159	0.090	0.035	0.135
	SE (±)	0.112	0.527	0.114	0.028	0.011	0.005	
	Upper	*1.180*	*2.86*	*1.23*	*0.189*	*0.0992*	*0.0386*	
	Lower	*0.956*	*1.75*	*1.01*	*0.134*	*0.0834*	*0.0308*	
ΔID	Best Fit	0.639	5.591	0.229	0.044	0.353	0.251	1.09
	SE (±)	0.080	0.809	0.013	0.004	0.033	0.025	
	Upper	*0.824*	*7.34*	*0.254*	*0.0508*	*0.424*	*0.304*	
	Lower	*0.510*	*4.23*	*0.203*	*0.0350*	*0.292*	*0.207*	

f

Figure A.S7: Transient-State Analysis of UGDH(FL) and UGDH(ΔID). a, b, Transient-state analysis of UDP-Xyl binding kinetics using intrinsic protein fluorescence. Six independent progress curves (coloured traces) at different inhibitor concentrations were globally fit (black line) to the allosteric inhibition model (see Figure 2.4b) for UGDH(FL) and UGDH(ΔID). Each progress curve was replicated ($n \geq 4$) with similar results, and the final kinetic model was refined against the averaged progress curves (see e for fit parameters). **c, d,** Confidence contour plots

depicting how constrained each globally fit parameter is relative to the others, for all progress curves in a and b (parameters are listed in e). **e,** Table of the microscopic rate constants from global fitting of the progress curves described in a and b. The best fit and s.e.m. were obtained from global nonlinear regression based on the numerical integration of rate equations for the described model (see main text and Methods). Upper and lower limits were obtained from the confidence contour analysis. $K_d = (K_1K_2K_3)^{-1}$, where $K_n = k_n/k_{-n}$. **f,** Enzyme hysteresis is observed as a lag in progress curves. Representative progress curves (of $n = 6$ independent measurements) for both UGDH(FL) (cyan) and UGDH(ΔID) (grey) are fit to equation (4) (black line). Curves are displayed with the y axis offset for clarity. Final results for all replicate curves are displayed in Figure A4.a.

Table A.S1: Data Collection and Refinement Statistics

Data collection		
Protein Data Bank Entry	5W4X	5VR8
	E* $hUGDH_{\Delta ID}$	E^{Ω} $hUGDH_{FL}$
Space group	C2	$P12_11$
Unit cell dimensions a,b,c (β)	178.19, 114.07, 97.24 (116.9°)	89.08, 196.49, 111.26 (111.9°)
Completeness (%)	99.9 (91.1)[a]	93.2 (60.0)[a]
No. reflections	324,675	2,730,154
Redundancy	6.4 (6.1)	12.3 (10.3)
***I* / σ(*I*)**	21.9 (1.5)	14.9 (2.5)
$CC_{1/2}$[b]	99.9 (64.9)	99.7 (79.3)
R_{meas} (%)[c]	6.5 (122.5)	13.2 (89.3)
Refinement		
Resolution (Å)	2.65	2.00
***R*work / *R*free**	0.19 / 0.23	0.16 / 0.19
No. atoms: Protein / Ligand / Water	10887 / 33 / 36	21584 / 394 / 1097
***B*-factors (Å²): Protein / Ligand / Water**	89.9 / 97.4 / 64.3	33.2 / 27.1 / 32.3
Stereochemical Ideality		
Bond lengths (Å²)	0.004	0.008
Bond angles (°)	0.75	0.91
φ,ψ Preferred (%)[d]	98.98	97.8
φ,ψ Additionally allowed (%)	1.02	2.2
φ,ψ Disallowed region (%)	0.0	0.0

[a] Values in parenthesis are for the highest-resolution shell (2.71–2.64 and 2.0221–1.9994 for 5W4X and 5VR8, respectively).
[b] $CC_{1/2}$ is the percentage of correlation between intensities from random half-data sets.[42]
[c] R_{meas} is the redundancy-independent merging R factor.[41]

Table A.S2: Kinetic Parameters of All UGDH Constructs[a]

hUGDH	K_M (UDP-Glc, μM)	k_{cat}[b] (s^{-1})	K_I^{UDX} (UDP-Xyl, μM)	α_{UDG}[c]	Δ ΔG[e] (kcal•mol^{-1})	# of Data Points[h]
ΔID	17.8 ± 0.9	0.7 ± 0.01	5.44 ± 0.55	------	0.00	42
FL	12.7 ± 0.6	0.8 ± 0.01	0.52 ± 0.04	------	-1.39	38
R1	12.9 ± 1.0	0.8 ± 0.01	0.60 ± 0.06	------	-1.31	59
0.13×R1	12.8 ± 1.2	1.0 ± 0.01	2.59 ± 0.24	------	-0.44	40
0.26xR1	12.4 ± 1.0	1.0 ± 0.01	1.81 ± 0.18	------	-0.65	42
0.5xR1	11.1 ± 0.8	1.0 ± 0.01	1.09 ± 0.08	------	-0.95	47
R2	43.7 ± 3.6	0.7 ± 0.01	0.78 ± 0.07	------	-1.15	50
-Lys	30.1 ± 1.9	0.5 ± 0.01	0.29 ± 0.03	------	-1.73	39
-Pro	13.1 ± 0.9	0.9 ± 0.01	0.72 ± 0.07	------	-1.20	26
0.13×FL	18.8 ± 0.9	1.0 ± 0.01	2.76 ± 0.15	------	-0.40	46
0.26×FL	18.3 ± 0.7	0.8 ± 0.01	1.99 ± 0.12	------	-0.60	42
0.5×FL	18.8 ± 0.9	0.9 ± 0.01	1.12 ± 0.08	------	-0.94	50
2×FL	15.2 ± 0.7	0.6 ± 0.01	0.30 ± 0.02	------	-1.72	43
0.13×Ser	16.9 ± 1.0	0.9 ± 0.01	2.67 ± 0.24	------	-0.42	49
0.26×Ser	18.4 ± 1.0	0.9 ± 0.01	1.76 ± 0.18	------	-0.67	43
0.5×Ser	17.4 ± 1.3	0.8 ± 0.01	1.09 ± 0.10	------	-0.95	49
Ser	17.8 ± 1.0	0.7 ± 0.01	0.60 ± 0.05	------	-1.31	53
ΔID-dimer	286 ± 27	0.1 ± 0.01	1.23 ± 0.15[d]	22 ± 12	0.00[f]	36
FL-dimer	83.2 ± 2.2	0.1 ± 0.01	0.17 ± 0.01[d]	36 ± 5	-1.17[f]	50
ΔID-A136M	9.9 ± 0.6	0.3 ± 0.01	4.20 ± 0.51	------	0.00[g]	30
FL-A136M	8.5 ± 0.6	0.7 ± 0.01	4.41 ± 0.37	------	0.03[g]	55

[a] Kinetic parameters and associated s.e.m. for all constructs were derived from global analyses of data in Figure A.S2.
[b] One catalytic turnover of UDP-GlcA produces two molecules of NADH per cycle.

[c] α describes the mode of mixed inhibition (equation (8)). An α >1 in the UDP-Glc saturation curves shows that UDP-Xyl binds preferentially to the allosteric binding site, and secondarily to the coenzyme-binding site.
[d] Competitive K_i from the fit to the mixed inhibition (equation (8)).
[e] Change in UDP-Xyl binding free energy (kcal•mol^{-1}) of UGDH constructs relative to UGDH(ΔID): $\Delta\Delta G = R \bullet T \bullet \ln((K_i \text{ Construct})/(K_i \Delta ID))$.
[f] Change in UDP-Xyl binding free energy relative to the UGDH(ΔID-dimer).
[g] Change in UDP-Xyl binding free energy relative to the UGDH(ΔID-A136M).
[h] The number of independent data points used in global analysis (see Methods).

Table A.S3: NAD^+ Kinetic Parameters for UGDH[a]

hUGDH	K_M (NAD^+, mM)	$K_{0.5}$[b] (NAD^+, mM)	Hill (*h*)	k_{cat}[c] (s^{-1})	UDX (K_i, μM)	α_{NAD}[d]	# of Data Points[e]
FL	----------	0.8 ± 0.20	0.8 ± 0.1	0.9 ± 0.08	----------	----------	18
ΔID	----------	0.3 ± 0.06	0.6 ± 0.1	0.7 ± 0.03	----------	----------	12
FL-dimer	2.0 ± 0.26	----------	----------	0.1 ± 0.01	2.1 ± 0.4	0.9 ± 0.2	37
ΔID-dimer	3.2 ± 0.10	----------	----------	0.2 ± 0.01	3.6 ± 0.8	0.6 ± 0.2	47
R1	----------	0.4 ± 0.03	0.9 ± 0.1	0.7 ± 0.01	----------	----------	17
R2	----------	0.8 ± 0.14	0.7 ± 0.1	0.7 ± 0.01	----------	----------	15
-Lys	----------	2.9 ± 0.61	0.8 ± 0.1	0.6 ± 0.04	----------	----------	10
-Pro	----------	0.5 ± 0.06	0.6 ± 0.1	1.2 ± 0.03	----------	----------	12
0.13×FL	----------	0.4 ± 0.03	0.7 ± 0.1	1.0 ± 0.03	----------	----------	13
0.26×FL	----------	0.2 ± 0.03	0.8 ± 0.2	0.8 ± 0.05	----------	----------	11
0.5×FL	----------	0.3 ± 0.03	0.9 ± 0.1	0.9 ± 0.02	----------	----------	12
2×FL	----------	1.4 ± 0.31	0.8 ± 0.1	0.9 ± 0.01	----------	----------	18
0.13×Ser	----------	0.9 ± 0.24	0.7 ± 0.1	1.2 ± 0.10	----------	----------	12
0.26×Ser	----------	1.0 ± 0.27	0.7 ± 0.1	1.1 ± 0.09	----------	----------	15
0.5×Ser	----------	1.2 ± 0.34	0.7 ± 0.1	1.1 ± 0.09	----------	----------	13
Ser	----------	1.3 ± 0.19	0.7 ± 0.1	1.0 ± 0.04	----------	----------	15

[a] Kinetic parameters and associated s.e.m. for all constructs were derived from global analyses of data in Figure A.S2.

[b] Hexameric UGDH displays negative cooperativity with NAD^+ binding, which indicates a mix of high-affinity and low-affinity sites.[22-27] In previous work, we showed that the native UGDH(FL) $K_{0.5}$ of 0.8 mM NAD^+ corresponds to a mix of high-affinity and low-affinity sites with K_M of 88 μM and 1.8 mM, respectively. This is consistent with the published K_d of 30 μM for the coenzyme.[24]

[c] One catalytic turnover of UDP-GlcA produces two molecules of NADH per cycle.

[d] α describes the mode of mixed inhibition (equation (8)). An α <1 in the NAD^+ saturation curves show that UDP-Xyl binds preferentially to the allosteric binding site, and secondarily to the coenzyme-binding site.

[e] The number of independent data points used in nonlinear regression (see Methods).

A8 REFERENCES

1 Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972, doi:10.1038/nature06522 (2007).

2 Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science* **254**, 1598-1603, doi:DOI 10.1126/science.1749933 (1991).

3 Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* **9**, 10-19, doi:10.1110/ps.9.1.10 (2000).

4 Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* **5**, 789-796, doi:10.1038/nchembio.232 (2009).

5 Campbell, E. *et al.* The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **12**, 944-950, doi:10.1038/nchembio.2175 (2016).

6 Oates, M. E. *et al.* D(2)P(2): database of disordered protein predictions. *Nucleic Acids Res* **41**, D508-516, doi:10.1093/nar/gks1226 (2013).

7 van der Lee, R. *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem Rev* **114**, 6589-6631, doi:10.1021/cr400525m (2014).

8 He, B. *et al.* Predicting intrinsic disorder in proteins: an overview. *Cell Res* **19**, 929-949, doi:10.1038/cr.2009.87 (2009).

9 Papaleo, E. *et al.* The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem Rev* **116**, 6391-6423, doi:10.1021/acs.chemrev.5b00623 (2016).

10 Bickel, T., Marques, C. & Jeppesen, C. Pressure patches for membranes: the induced pinch of a grafted polymer. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* **62**, 1124-1127 (2000).

11 Bickel, T., Jeppesen, C. & Marques, C. M. Local entropic effects of polymers grafted to soft interfaces. *Eur Phys J E* **4**, 33-43, doi:10.1007/s101890170140 (2001).

12 Carmichael, S. P. & Shell, M. S. Entropic (de)stabilization of surface-bound peptides conjugated with polymers. *J Chem Phys* **143**, doi:10.1063/1.4929592 (2015).

13 Waters, J. T. & Kim, H. D. Calculation of a fluctuating entropic force by phase space sampling. *Phys Rev E* **92**, doi:10.1103/PhysRevE.92.013308 (2015).

14 Hilser, V. J. An Ensemble View of Allostery. *Science* **327**, 653-654, doi:10.1126/science.1186121 (2010).

15 Li, J. *et al.* Genetically tunable frustration controls allostery in an intrinsically disordered transcription factor. *Elife* **6**, doi:10.7554/eLife.30688 (2017).

16 Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *P Natl Acad Sci USA* **104**, 8311-8315, doi:10.1073/pnas.0700323104 (2007).

17 Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025, doi:10.1038/nature05858 (2007).

18 Ferreon, A. C., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390-394, doi:10.1038/nature12294 (2013).

19 Egger, S., Chaikuad, A., Kavanagh, K. L., Oppermann, U. & Nidetzky, B. Structure and mechanism of human UDP-glucose 6-dehydrogenase. *J Biol Chem* **286**, 23877-23887, doi:10.1074/jbc.M111.234682 (2011).

20 Gainey, P. A. & Phelps, C. F. Interactions of Uridine-Diphosphate Glucose Dehydrogenase with Inhibitor Uridine-Diphosphate Xylose. *Biochemical Journal* **145**, 129-134, doi:10.1042/bj1450129 (1975).

21 Neufeld, E. F. & Hall, C. W. Inhibition of Udp-D-Glucose Dehydrogenase by Udp-D-Xylose - A Possible Regulatory Mechanism. *Biochem Bioph Res Co* **19**, 456-&, doi:10.1016/0006-291x(65)90146-4 (1965).

22 Beattie, N. R., Keul, N. D., Sidlo, A. M. & Wood, Z. A. Allostery and Hysteresis Are Coupled in Human UDP-Glucose Dehydrogenase. *Biochemistry-Us* **56**, 202-211, doi:10.1021/acs.biochem.6b01044 (2017).

23 Kadirvelraj, R. *et al.* Hysteresis in Human UDP-Glucose Dehydrogenase Is Due to a Restrained Hexameric Structure That Favors Feedback Inhibition. *Biochemistry-Us* **53**, 8043-8051, doi:10.1021/bi500594x (2014).

24 Kadirvelraj, R., Sennett, N. C., Custer, G. S., Phillips, R. S. & Wood, Z. A. Hysteresis and negative cooperativity in human UDP-glucose dehydrogenase. *Biochemistry-Us* **52**, 1456-1465, doi:10.1021/bi301593c (2013).

25 Kadirvelraj, R., Sennett, N. C., Polizzi, S. J., Weitzel, S. & Wood, Z. A. Role of packing defects in the evolution of allostery and induced fit in human UDP-glucose dehydrogenase. *Biochemistry-Us* **50**, 5780-5789, doi:10.1021/bi2005637 (2011).

26 Sennett, N. C., Kadirvelraj, R. & Wood, Z. A. Conformational flexibility in the allosteric regulation of human UDP-alpha-D-glucose 6-dehydrogenase. *Biochemistry-Us* **50**, 9651-9663, doi:10.1021/bi201381e (2011).

27 Sennett, N. C., Kadirvelraj, R. & Wood, Z. A. Cofactor binding triggers a molecular switch to allosterically activate human UDP-alpha-D-glucose 6-dehydrogenase. *Biochemistry-Us* **51**, 9364-9374, doi:10.1021/bi301067w (2012).

28 Theillet, F. X. *et al.* The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord Proteins* **1**, e24360, doi:10.4161/idp.24360 (2013).

29 Uversky, V. N. The intrinsic disorder alphabet. III. Dual personality of serine. *Intrinsically Disord Proteins* **3**, e1027032, doi:10.1080/21690707.2015.1027032 (2015).

30 Busch, D. J. *et al.* Intrinsically disordered proteins drive membrane curvature. *Nat Commun* **6**, 7875, doi:10.1038/ncomms8875 (2015).

31 Balasubramaniam, D. & Komives, E. A. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim Biophys Acta* **1834**, 1202-1209, doi:10.1016/j.bbapap.2012.10.009 (2013).

32 Chen, J. Z. Y. Theory of wormlike polymer chains in confinement. *Prog Polym Sci* **54-55**, 3-46, doi:10.1016/j.progpolymsci.2015.09.002 (2016).

33 Smyda, M. R. & Harvey, S. C. The entropic cost of polymer confinement. *J Phys Chem B* **116**, 10928-10934, doi:10.1021/jp302807r (2012).

34 Thielges, M. C., Chung, J. K., Axup, J. Y. & Fayer, M. D. Influence of histidine tag attachment on picosecond protein dynamics. *Biochemistry-Us* **50**, 5799-5805, doi:10.1021/bi2003923 (2011).

35 Wilkins, M. R. *et al.* Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552 (1999).

36 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).

37 Brunger, A. T. Free R value: cross-validation in crystallography. *Methods Enzymol* **277**, 366-396 (1997).

38 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

39 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

40 Urzhumtsev, A., Afonine, P. V. & Adams, P. D. TLS from fundamentals to practice. *Crystallogr Rev* **19**, 230-270, doi:10.1080/0889311X.2013.835806 (2013).

41 Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-275 (1997).

42 Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030-1033, doi:10.1126/science.1218231 (2012).

43 Frieden, C. Kinetic aspects of regulation of metabolic processes. The hysteretic enzyme concept. *J Biol Chem* **245**, 5788-5799 (1970).

44 Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. *Analytical Ultracentrifugation*. (The Royal Society of Chemistry, 1992).

45 Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619, doi:10.1016/S0006-3495(00)76713-0 (2000).

46 Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**, 104-124 (2003).

47 Ortega, A., Amoros, D. & Garcia de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J* **101**, 892-898, doi:10.1016/j.bpj.2011.06.046 (2011).

48 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

49 Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275-282 (1992).

50 Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874, doi:10.1093/molbev/msw054 (2016).

51 Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J Mol Biol* **7**, 95-99 (1963).

52 Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol* **285**, 1735-1747, doi:DOI 10.1006/jmbi.1998.2401 (1999).

53 Huynh, K. & Partch, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. *Curr Protoc Protein Sci* **79**, 28 29 21-14, doi:10.1002/0471140864.ps2809s79 (2015).

54 Fang, J. *et al.* Conformational dynamics of the Escherichia coli DNA polymerase manager proteins UmuD and UmuD'. *J Mol Biol* **398**, 40-53, doi:10.1016/j.jmb.2010.02.040 (2010).

55 Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* **25**, 158-170, doi:10.1002/mas.20064 (2006).

56 Johnson, B. *et al.* Dimerization Controls Marburg Virus VP24-dependent Modulation of Host Antioxidative Stress Responses. *J Mol Biol* **428**, 3483-3494, doi:10.1016/j.jmb.2016.07.020 (2016).

57 Chen, E. *et al.* Broadly neutralizing epitopes in the Plasmodium vivax vaccine candidate Duffy Binding Protein. *Proc Natl Acad Sci U S A* **113**, 6277-6282, doi:10.1073/pnas.1600488113 (2016).

58 Yan, Y., Grant, G. A. & Gross, M. L. Hydrogen-Deuterium Exchange Mass Spectrometry Reveals Unique Conformational and Chemical Transformations Occurring upon [4Fe-4S] Cluster Binding in the Type 2 L-Serine Dehydratase from Legionella pneumophila. *Biochemistry-Us* **54**, 5322-5328, doi:10.1021/acs.biochem.5b00761 (2015).

59 Palmier, M. O. & Van Doren, S. R. Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal Biochem* **371**, 43-51, doi:10.1016/j.ab.2007.07.008 (2007).

60 Johnson, K. A. Fitting enzyme kinetic data with KinTek Global Kinetic Explorer. *Methods Enzymol* **467**, 601-626, doi:10.1016/S0076-6879(09)67023-3 (2009).

61 Johnson, K. A., Simpson, Z. B. & Blom, T. Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. *Analytical Biochemistry* **387**, 20-29, doi:10.1016/j.ab.2008.12.024 (2009).

62 Johnson, K. A., Simpson, Z. B. & Blom, T. FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal Biochem* **387**, 30-41, doi:10.1016/j.ab.2008.12.025 (2009).