Title

Rare Genetic Variants in Nutrient Transporters: Impact on drug response and human disease

Permalink

https://escholarship.org/uc/item/28q4m6hq

Author

Enogieru, Osatohanmwen Jessica

Publication Date

2020

Supplemental Material

https://escholarship.org/uc/item/28q4m6hq#supplemental

Peer reviewed|Thesis/dissertation

Rare Genetic Variants in Nutrient Transporters: Impact on drug response and human disease

by
Osatohanmwen Enogieru

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Kathleen Giacomini
Kathleen Giacomini
6C466CD5B84548C...
Chair

DocuSigned by:
Rada Savic
Rada Savic

DocuSigned by:
Nadav Ahituv
Nadav Ahituv

DocuSigned by:
Kathleen Giacomini
Kathleen Giacomini
6C466CD5B84548C...

Committee Members

Acknowledgements

This dissertation is the culmination of almost 6 years of research that would not have reached fruition without the constant love, encouragement, guidance and support I have received from family, friends and colleagues.

First, I would like to thank Dr. Kathleen Giacomini, my mentor and advisor, for her guidance and support. Kathy is a premier example of scientific expertise, rigor, leadership and creativity. I am quite fortunate to have spent my formative years as a scientist learning not only from her scientific expertise, but her professional expertise as a leader. Kathy's mentorship style has encouraged my scientific growth, independence and creativity; skills that build the foundation for any career in science. Second, I would like to thank members of my dissertation committee Dr. Nadav Ahituv and Dr. Rada Savic, and members of the qualifying exam committee Dr. Pui Kwok, Dr. Deanna Kroetz, and Dr. Lisa Kroon for their important feedback and advice regarding my proposed research projects and dissertation. Third, I would like to acknowledge and thank Dr. SookWah Yee for her guidance, encouragement and significant contribution to my scientific development and technical expertise.

Fourth, I would like to thank all members of the Giacomini lab, both past and present, who have contributed to my scientific development and education (Huan-Chieh Chien (James), Elizabeth Green, Ling Zou, Xiaomin Liang, Adrian Stecula, Shylaja Srinivasan, Deanna Brackman, Bianca Vora, Huy Ngo, Renata Gallagher) and making the Giacomini lab a welcoming learning environment. Fifth, I would like to thank Dr. Daniel L Minor for providing access to his lab and its equipment, which was critical in the execution of my research.

Contributions

Enogieru OJ, Ung PMU, Yee SW, Schlessinger A, Giacomini KM. Functional and structural analysis of rare *SLC2A2* variants associated with Fanconi-Bickel syndrome and metabolic traits. Hum Mutat. 2019 Jul;40(7):983-995.

Rare Genetic Variants in Nutrient Transporters: Impact on drug response and human disease

Osatohanmwen Jessica Enogieru

Abstract

Translational research, often referred to as "bench-to-bedside" science, involves converting novel biological, chemical or other scientific discoveries into clinically relevant medical treatments, devices and/or applications. Reverse translational research uses basic science tools and experimental models to investigate the biological, physiological and molecular mechanisms behind observations in human clinical data.

In this dissertation project, we use both approaches to investigate the influence of nutrient membrane transporters on clinical drug response and human disease.

Metformin is first-line treatment for Type 2 Diabetes (T2D), a metabolic disorder characterized by excessive glycemia and reduced insulin production and sensitivity. The prevalence of Type 2 Diabetes is growing both in the U.S. and globally. Therefore, a larger and increasingly diverse population of individuals will be treated with metformin. The heritability of metformin response is substantial (about 33%), but common genetic variants previously identified to associate with metformin response (*SLC2A2,* solute carrier family 2 member 2; *ATM,* ataxia telangiectasia mutated serine/threonine kinase) do not fully account for all the heritability. Our translational approach starts with conducting a metformin sequencing association study in a cohort of patients who responded very well (extreme responders) or poorly (extreme non-responders) to metformin to identify rare variants in the glucose transporter, GLUT2 (*SLC2A2*) that modulate metformin response. In this dissertation, I describe the study design and various data analytical approaches that will be used to discover rare variants in *SLC2A2* that associate with metformin response.

In our reverse translational approach we focus on rare variants in the same gene (*SLC2A2*), which are causal for Fanconi-Bickel syndrome (FBS), an autosomal recessive glycogen storage disease. Mild and severe forms of FBS have been observed and reported in the clinical literature, but the question of whether residual function of mutant GLUT2 associates with disease severity has not been examined. So, I functionally characterized the impact of GLUT2 mutations causal for FBS on both transporter function and membrane expression. In the study, I identified a variant (p.Leu153_Ile154del), found in patients with a mild form of FBS that retained residual glucose transport activity, suggesting that small increases in GLUT2 activity could substantially reduce the severity of FBS symptoms.

In our final study I conducted a screen of prescription drugs as inhibitors of THTR-1 (thiamine transporter encoded by *SLC19A2,* solute carrier family 19 member 2). The results of that screen identified several prescription and one over-the-counter drugs that inhibit THTR-1. Using data from electronic medical records, we observed significantly reduced thiamine pyrophosphate levels in blood samples from patients exposed to two THTR-1 inhibitors (erythromycin, omeprazole) versus samples from age and sex- matched individuals not on the drugs, suggesting inhibition of THTR-1 *in vivo*. Both drugs are associated with megaloblastic anemia, which is also caused by genetic mutations in THTR-1. The results suggest the possibility of a novel mechanism for drug-induced megaloblastic anemia.

Our overall findings of the dissertation have important implications for understanding the role of micro- and macro- nutrient transporters in drug response and human disease. Our studies continue to examine the role of genetic variants in the glucose transporter gene, *SLC2A2* and metformin response, extending the application of precision medicine to the pharmacological

treatment of T2D. Importantly, our studies suggest novel treatment strategies for rare diseases caused by genetic mutations in nutrient transporter, like FBS, and novel mechanisms for drug-induced disorders, such as drug-induced megaloblastic anemia.

Table of Contents

List of Figures

List of Tables

Chapter 1: The role of *SLC2A2* (GLUT2) in human disease and drug response

INTRODUCTION

Transporters are integral membrane proteins that mediate the translocation of substrates across biological membranes. SLC (solute carrier) proteins are encoded by genes in a large superfamily, which are expressed in various biological membranes (i.e. extracellular plasma membrane, intracellular organelle membranes) in cells.[1] In humans, SLC transporters are responsible for the absorption, distribution and elimination of nutrients, vitamins and minerals required for the metabolic and physiological function of cells, tissues and the body as a whole[1] as well as for exogenous compounds including prescription drugs and herbal medications.

Glucose is a necessary nutrient for nearly all forms of life; glucose and other simple monosaccharides are used for critical functions such as energy production/storage (i.e. glucose, fructose) structural integrity (i.e. xylose) and genetic replication (i.e. ribose). However, because of their polar nature and large size, monosaccharides such as glucose poorly traverse biological membranes via passive diffusion.[2] Within the SLC super-family, there are three eukaryotic glucose transporter sub-families that facilitate transport of glucose across biological membranes, which are named GLUTs (encoded by *SLC2A* genes), SGLTs (encoded by *SLC5A* genes) and SWEETs (encoded by *SLC50A1,* solute carrier family 50 member 1).[3] GLUT transporters, with one exception (HMIT, a H+/myo-inositol symporter), are facilitative which means they do not use energy from ATP or a secondary source (electrolyte gradients) to translocate glucose across membranes; instead, they move their substrates across the membrane in accordance with the concentration gradient.[4]

In this introductory chapter, I present a brief overview of some of the transporters in the GLUT family, in particular, GLUT1, GLUT4 and GLUT2. A large component of this dissertation focuses on GLUT2; therefore more introductory information on GLUT2 is included. This is followed by a brief overview of the micronutrient transporter, THTR1, encoded by *SLC19A2*. Following this overview section, I then present goals of my dissertation research and briefly describe the individual chapters.

Overview of Glucose Transporters in the SLC2A Family

Two of the most widely studied glucose transporters are members of the GLUT family: GLUT1 and GLUT4. GLUT1, encoded by *SLC2A1* (solute carrier family 2 member 1) and discovered in 1985,[5] has a high Michaelis-Menten constant for glucose (K_m = 3 mM).[6] The Michaelis-Menten constant (K_m) is the concentration at which a transporter reaches 50% of its activity (i.e. rate of transport) and represents the affinity of a substrate towards a transporter. In general, GLUTs have high Michaelis-Menten constant (i.e. a low transport affinity for its substrate) that are in the millimolar range for glucose. But considering the average blood glucose concentration in humans is 5 mM, GLUT1 efficiently transports glucose even in fasted (3.9 – 5.5 mM) or hypoglycemic (< 3.9 mM) conditions. It is ubiquitously expressed throughout different human tissues, the most critical being tissues within the brain. During normal conditions, the brain is completely reliant on glucose as an energy source; glucose transport into the brain is the rate-limiting step for brain glucose metabolism.[4,6] GLUT1 is highly expressed on the luminal membrane of brain endothelial cells and in astrocytes.[4] Homozygous knockdown of GLUT1 in mice leads to embryonic lethality. In humans, heterozygous knockdown of GLUT1 leads to autosomal dominant Mendelian disease called GLUT1 deficiency syndrome (GDS) that is characterized by seizures in early infancy, developmental delay, ataxia, and neurobehavioral

symptoms.[4] Intronic variants in GLUT1 significantly associate with hematological traits (e.g. monocyte count). Interestingly, there are no published studies that report a genome-wide association between genetic variants in GLUT1 and a common disease or drug response phenotype[7] consistent with high selective pressure on *SLC2A1*.

GLUT4, encoded by *SLC2A4* (solute carrier family 2 member 4) and discovered in 1988,[8] also has a high K_m for glucose (K_m = 5 mM) and is mainly expressed in skeletal muscle cells, cardiomyocytes and adipocytes. GLUT4 functions as an insulin-responsive glucose transporter. At low insulin concentrations, GLUT4 resides primarily in intracellular membrane compartments. When insulin concentrations increase after a carbohydrate meal and interact with insulin receptors, intracellular GLUT4 transporters redistribute to the plasma membrane. The insulin-stimulated plasma membrane redistribution of GLUT4 results in increases in glucose uptake and metabolism in these tissues, preventing chronic elevations in blood glucose levels. Mice born without functional GLUT4 in any tissue develop cardiac hypertrophy, exhibit growth retardation and die prematurely.[9] To date, there are no known human Mendelian diseases or syndromes linked specifically to GLUT4 deleterious variants. Furthermore, variants in GLUT4 significantly associate with blood pressure traits, but notably, there are no published studies that report a genome-wide association between GLUT4 variants and a common disease or drug response phenotype.[7]

GLUT2's physiological role

In contrast to the plethora of manuscripts focused on studies of GLUT1 and GLUT4, considerably fewer studies focus on GLUT2, the second glucose transporter in the GLUT subfamily. Compared with the other four major GLUT proteins, GLUT2 (encoded by *SLC2A2,*

solute carrier family 2 member 2) exhibits a unique substrate specificity and transporter kinetic characteristics that underlie its specific physiological roles. First characterized in 1988,[8,10] human GLUT2 is mainly expressed in four cell types: hepatocytes, enterocytes, renal proximal tubule cells and pancreatic beta cells (Figure 1.1A). In body tissues, GLUT2 expression is highest in the liver. It is the primary glucose transporter in the liver and is the rate-limiting step for glucose flux in and out of the liver (Figure 1.1B).
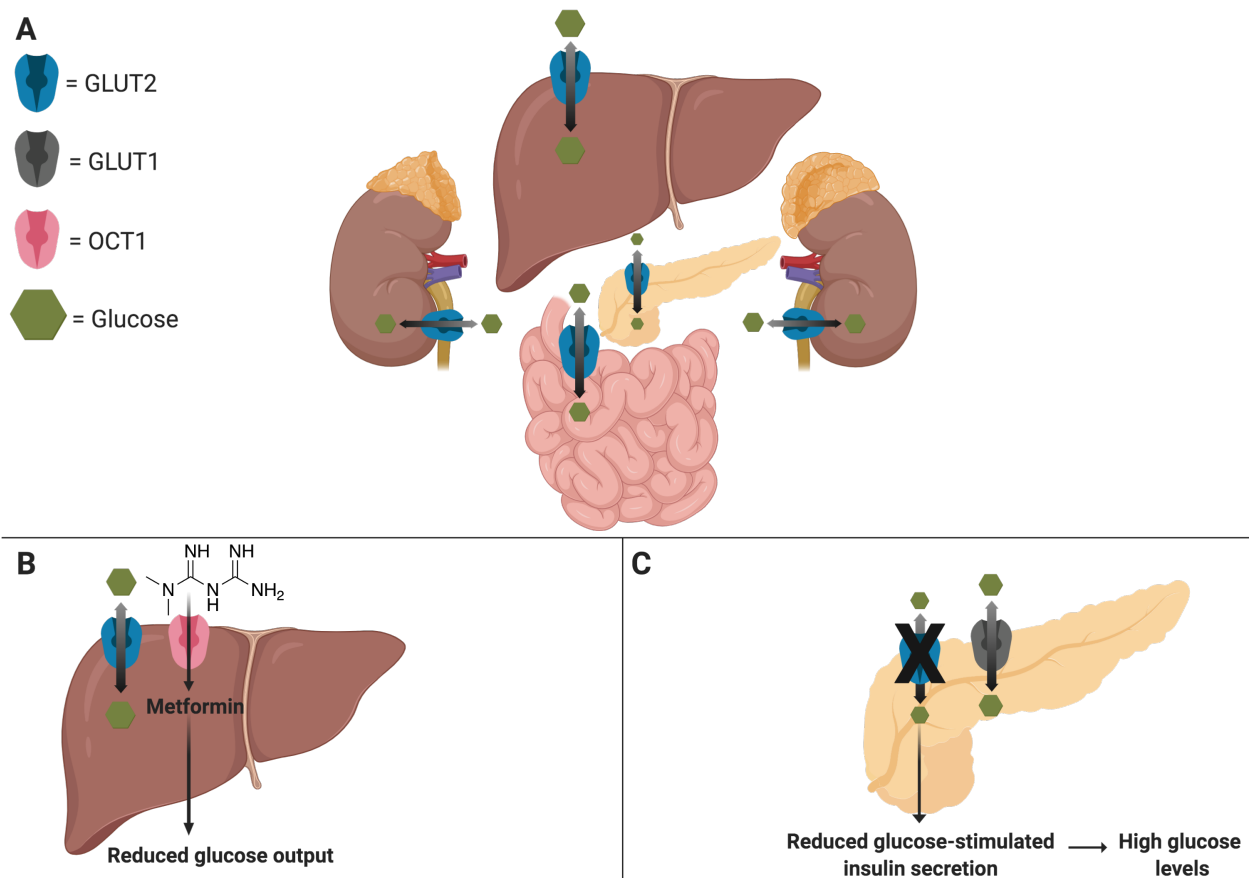
Figure 1.1 GLUT2 transporter tissue localization.
Human GLUT2 is mainly expressed in four cell types: hepatocytes, enterocytes, renal proximal tubule cells and pancreatic beta cells (A). In hepatocytes, GLUT2 is responsible for the bidirectional transport of glucose in and out of the liver. The primary mechanism of action of biguanide compounds, like metformin, occurs in the liver and results in decreased glucose synthesis and glucose output into the bloodstream (B). In the pancreatic beta cells, GLUT2 acts as a glucose sensor. Inactivation of GLUT2 in pancreatic beta cells leads to a significant reduction in glucose-stimulated insulin secretion (C).

GLUT2 expression is second highest in the enterocytes of the small intestine. However, in the intestine, other glucose transporters, which have higher affinities (low K_m) are also expressed (i.e. SGLT1 (encoded by *SLC5A1*, solute carrier family 5 member 1), GLUT1, GLUT3 (encoded by *SLC2A3*, solute carrier family 2 member 3)), GLUT2 is also expressed in the renal proximal tubule of the kidney, where is participates in glucose reabsorption in partnership with SGLT2 (encoded by *SLC5A2*, solute carrier family 5 member 2). GLUT2 is expressed in low but notable levels in pancreatic beta cells that are solely responsible for insulin synthesis. Studies that evaluated glucose-stimulated insulin secretion (GSIS) of islet cells in GLUT2 $^{-/-}$ mice found that the number of beta cells was decreased, gene transcription and mRNA translation of insulin was suppressed and GSIS was severely reduced (Figure 1.1C). Once GLUT2 cDNA was reintroduced back into GLUT2-null islet cells, GSIS was normalized.[11,12,13] These studies suggest that GLUT2 is not only a glucose transporter, but based on its unique kinetic parameters, is a glucose sensor and detector for specific tissues.

Compared with other GLUT and non-GLUT glucose transporters, GLUT2 has the highest K_m for glucose (K_m = 17 mM, Table 1.1). GLUT2 is defined as a low affinity, high capacity glucose transporter that can transport large amounts of substrate during periods of high glucose concentrations (i.e. post-prandial carbohydrate ingestion).[14, 15] Additionally, GLUT2 has a broad substrate range compared with other GLUTs. For example, GLUT2 is one of the few fructose transporters in the body. GLUT2 in the basolateral membrane of enterocytes works in concert with GLUT5 on the apical membrane to transport fructose (K_m = 76 mM), a monosaccharide found primarily fruits and plants, into the bloodstream from the small intestine.

Table 1.1 The affinity (K_m) of glucose transporters for D-glucose.

Glucose transporter	K_m
GLUT1	3 mM
GLUT2	17 mM
GLUT3	1.4 mM*
GLUT4	5 mM
GLUT5	--
SGLT1	0.4 mM
SGLT2	≤ 6 mM
SGLT3	-- (glucose sensor)
SWEET1	Stimulates glucose efflux

The K_m of glucose for the main, human glucose transporters in each class of glucose transporters (GLUTs, SGLTs, SWEET1) is listed. * = K_m only available for 2-deoxyglucose, not glucose. -- = does not transport glucose
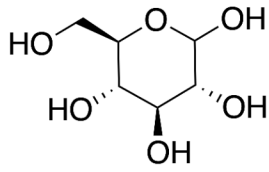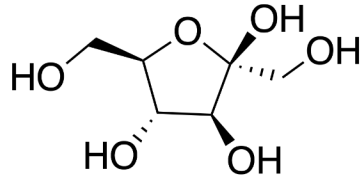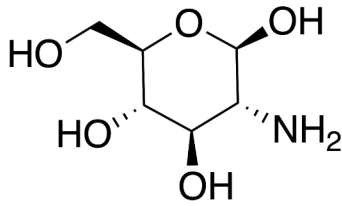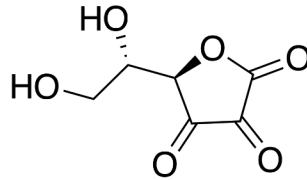
Figure 1.2 Chemical structures of GLUT2 substrates
The structure of GLUT2 substrates has either a five-membered ring or a six-membered ring. GLUT2 substrates are monosaccharides that are hexoses (six carbon atoms) or pentoses (five carbon atoms).

Interestingly, GLUT2 has a lower K_m for glucosamine (K_m = 0.8 mM) and dehydroascorbic acid (K_m = 2.33 mM) than glucose (17 mM) (Figure 1.2). In addition to GLUT2's broad list of known substrates, I identified three novel substrates of human GLUT2, as described in Chapter 3 of this dissertation, which were previously unknown: ribose, 1,5 anhydroglucitol, and xylose.

GLUT2's unique affinity for glucose and its broad substrate range has been linked to its specific amino acid sequence in key regions of the protein. GLUT2's amino acid sequence is only 39-53% identical to the amino acid sequence of GLUTs 1-5, with GLUT1 having the highest percentage (53%) and GLUT5 having the lowest (39%).[16] Several protein chimera studies have implicated trans membrane helix 7 as the region responsible for GLUT2's high Km for glucose and its ability to transport fructose.[17,18,19,20]

GLUT2's role in disease genetics and drug response

Genome-wide association studies (GWAS) are a powerful and agnostic approach used to identify causal variants and genes that influence development of a phenotype like a disease or response to a drug. GWAS associate specific genetic variations with particular diseases by scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease. Notably, GLUT2 has been implicated in more GWAS studies of common diseases (i.e. Type 2 Diabetes)[21,22] and metabolic traits (i.e. fasting blood glucose, body mass index, serum cholesterol)[23,24,25,26] then any other GLUT transporter.[7] In addition to disease phenotypes, variants in *SLC2A2* have also been associated with response to drugs, specifically response to the highly prescribed anti-diabetic drug, metformin (Figure 1.1B). In 2016, in collaboration with Ewan Pearson and others, our laboratory performed a large GWAS with replication (N = 10,298 from several cohorts) with the focus on

initial glycemic response to metformin.[27] The GWAS, conducted initially in patients of European

ancestry, identified a genome-wide level significant association between rs8192675 (C-allele) in

the *SLC2A2* gene (encodes for the GLUT2 transporter) and greater response to metformin (p =

5.9×10^{-15}). Also, rs8192675, a common variant with a minor allele frequency (MAF) of 30% in

Europeans, significantly associated with increased metformin response in three other ethnic

groups (African-Americans, Latinos, East Asians), though p values were weaker likely due to

smaller sample sizes.[28] Subsequently, a second independent study published in 2019 confirmed

rs8192675 association with metformin response[29] and collectively these studies provide strong

support for a role of GLUT2 in initial glycemic response to metformin. Unfortunately, due to the

inherent limitations of GWAS analysis, we do not know if or how rare (MAF < 5%) *SLC2A2*

variants influence glycemic response to metformin.

It is also known that rare, inactivating variants in *SLC2A2* cause a Mendelian glycogen storage

disease known as Fanconi-Bickel syndrome (FBS, OMIM# 227810; FBS disease characteristics

are described in Chapter 3). Over 40 FBS variants have been identified, but the impact of these

variants on GLUT2 transporter function, expression and FBS phenotype were rarely, if ever,

studied.[30] Unfortunately, this is a common issue with Mendelian diseases caused by a SLC

transporter, with a few exceptions (i.e. Carnitine Transporter Deficiency syndrome). The lack of

functional studies likely contributes to the unavailability of FDA-approved drugs for genetic

diseases caused by SLC transporters.

THTR-1's role in disease and adverse drug reactions

Thiamine Human Transporter 1 (THTR-1), which is encoded by *SLC19A2*, is a thiamine

transporter expressed ubiquitously in many different tissues. However, THTR-1 is the main

thiamine transporter in pancreatic islet tissue and hematopoietic stem cells. Rare, deleterious variants in *SLC19A2* cause thiamine-responsive megaloblastic anemia (TRMA), a syndrome characterized by deafness, diabetes and megaloblastic anemia (MA).[31] MA is generally caused by folic acid or vitamin B_{12} deficiencies or drug-induced. Prescription drugs can induce MA through a variety of mechanisms.[32] But, to date, the impact of these prescription drugs on THTR-1 transporter activity and thiamine levels was unknown.

Summary of Dissertation Chapters

The research described in this dissertation focuses mainly on GLUT2 and to a lesser extent, THTR-1. The overall goal of this dissertation is to understand the influence of rare variants in nutrient transporter genes (i.e. *SLC2A2*, *SLC19A2*) on drug response, disease expression and phenotype. In Chapter 2, the goal of the research is to determine if rare, *SLC2A2* variants modulate initial therapeutic response to metformin. In Chapter 3, the aim of the study is to functionally characterize the impact of FBS-causal *SLC2A2* variants on GLUT2 transporter function, structure and expression and relate that information to FBS phenotype. Here, we identified an in-frame deletion variant (p.Leu153_Ile154del) with residual transport activity that associated with a mild form of FBS. In Chapter 4, the aim of the study is to identify prescription drugs that can phenocopy or mimic the effects of rare, deleterious *SLC19A2* variants (i.e. thiamine-responsive megaloblastic anemia). Two drugs known to induce megaloblastic anemia (omeprazole, erythromycin) were found to inhibit THTR-1 thiamine transport *in vitro* and associate with lower levels of thiamine pyrophosphate in a general patient population. In Chapter 5, I summarize the results of the studies in this dissertation, highlight our major findings and discuss future studies needed to identify treatments that correct dysfunctional expression or

activity of SLC transporters and uncover the biological and physiological mechanisms that

determine GLUT2's relationship with metformin response.

REFERENCES

1.   Lin L, Yee SW, Kim RB, Giacomini KM. SLC transporters as therapeutic targets: emerging opportunities. *Nat Rev Drug Discov*. 2015;14(8):543-560. doi:10.1038/nrd4626

2.   Navale AM, Paranjape AN. Glucose transporters: physiological and pathological roles. *Biophys Rev*. 2016;8(1):5-9. doi:10.1007/s12551-015-0186-2

3.   Chen L-Q, Cheung LS, Feng L, Tanner W, Frommer WB. Transport of sugars. *Annu Rev Biochem*. 2015;84:865-894. doi:10.1146/annurev-biochem-060614-033904

4.   Mueckler M, Thorens B. The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med*. 2013;34(2-3):121-138. doi:10.1016/j.mam.2012.07.001

5.   Mueckler M, Caruso C, Baldwin SA, et al. Sequence and structure of a human glucose transporter. *Science*. 1985;229(4717):941-945. doi:10.1126/science.3839598

6.   Barron CC, Bilan PJ, Tsakiridis T, Tsiani E. Facilitative glucose transporters: Implications for cancer detection, prognosis and treatment. *Metabolism*. 2016;65(2):124-139. doi:10.1016/j.metabol.2015.10.007

7.   Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005-D1012. doi:10.1093/nar/gky1120

8.   James DE, Brown R, Navarro J, Pilch PF. Insulin-regulatable tissues express a unique insulin-sensitive glucose transport protein. *Nature*. 1988;333(6169):183-185. doi:10.1038/333183a0

9.   Abel ED, Kaulbach HC, Tian R, et al. Cardiac hypertrophy with preserved contractile function after selective deletion of GLUT4 from the heart. *J Clin Invest*. 1999;104(12):1703-1714. doi:10.1172/JCI7605

10. Kayano T, Fukumoto H, Eddy RL, et al. Evidence for a family of human glucose transporter-like proteins. Sequence and gene localization of a protein expressed in fetal skeletal muscle and other tissues. *J Biol Chem*. 1988;263(30):15245-15248.

11. Guillam MT, Dupraz P, Thorens B. Glucose uptake, utilization, and signaling in GLUT2-null islets. *Diabetes*. 2000;49(9):1485-1491.

12. Thorens B. GLUT2 in pancreatic and extra-pancreatic gluco-detection (review). *Mol Membr Biol*. 2001;18(4):265-273. doi:10.1080/09687680110100995

13. Thorens B. GLUT2, glucose sensing and glucose homeostasis. *Diabetologia*. 2015;58(2):221-232. doi:10.1007/s00125-014-3451-1

14. Mueckler M. Facilitative glucose transporters. *Eur J Biochem*. 1994;219(3):713-725. doi:10.1111/j.1432-1033.1994.tb18550.x

15. Leturque A, Brot-Laroche E, Le Gall M. GLUT2 mutations, translocation, and receptor function in diet sugar managing. *Am J Physiol Endocrinol Metab*. 2009;296(5):E985-92. doi:10.1152/ajpendo.00004.2009

16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2

17. Buchs A, Wu L, Morita H, Whitesell RR, Powers AC. Two regions of GLUT 2 glucose transporter protein are responsible for its distinctive affinity for glucose. *Endocrinology*. 1995;136(10):4224-4230. doi:10.1210/endo.136.10.7664639

18. Arbuckle MI, Kane S, Porter LM, Seatter MJ, Gould GW. Structure-function analysis of liver-type (GLUT2) and brain-type (GLUT3) glucose transporters: expression of chimeric transporters in Xenopus oocytes suggests an important role for putative transmembrane helix 7 in determining substrate selectivity. *Biochemistry*. 1996;35(51):16519-16527.

doi:10.1021/bi962210n

19. Wu L, Fritz JD, Powers AC. Different functional domains of GLUT2 glucose transporter are required for glucose affinity and substrate specificity. *Endocrinology*. 1998;139(10):4205-4212. doi:10.1210/endo.139.10.6245

20. Seatter MJ, De la Rue SA, Porter LM, Gould GW. QLS motif in transmembrane helix VII of the glucose transporter family interacts with the C-1 position of D-glucose and is involved in substrate selection at the exofacial binding site. *Biochemistry*. 1998;37(5):1322-1326. doi:10.1021/bi972322u

21. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. 2018;50(11):1505-1513. doi:10.1038/s41588-018-0241-6

22. Type 2 Diabetes Knowledge Portal. http://www.type2diabetesgenetics.org. Accessed December 30, 2019.

23. Igl W, Johansson A, Wilson JF, et al. Modeling of environmental effects in genome-wide association studies identifies SLC2A2 and HP as novel loci influencing serum cholesterol levels. *PLoS Genet*. 2010;6(1):e1000798. doi:10.1371/journal.pgen.1000798

24. Wessel J, Chu AY, Willems SM, et al. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat Commun*. 2015;6:5897. doi:10.1038/ncomms6897

25. Manning AK, Hivert M-F, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 2012;44(6):659-669. doi:10.1038/ng.2274

26. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting

glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2):105-116. doi:10.1038/ng.520

27. Zhou K, Yee SW, MetGen coauthors, Giacomini K PE. *Variation in the Glucose Transporter Gene GLUT2 (SLC2A2) Is Associated with Glycaeic Response to Metformin: A MetGen Study*.; 2016.

28. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet*. 2016;48(9):1055-1059. doi:10.1038/ng.3632

29. Rathmann W, Strassburger K, Bongaerts B, et al. A variant of the glucose transporter gene SLC2A2 modifies the glycaemic response to metformin therapy in recently diagnosed type 2 diabetes. *Diabetologia*. 2019;62(2):286-291. doi:10.1007/s00125-018-4759-z

30. Enogieru OJ, Ung PMU, Yee SW, Schlessinger A, Giacomini KM. Functional and structural analysis of rare SLC2A2 variants associated with Fanconi-Bickel syndrome and metabolic traits. *Hum Mutat*. 2019;40(7):983-995. doi:10.1002/humu.23758

31. Habeb AM, Flanagan SE, Zulali MA, et al. Pharmacogenomics in diabetes: outcomes of thiamine therapy in TRMA syndrome. *Diabetologia*. 2018;61(5):1027-1036. doi:10.1007/s00125-018-4554-x

32. Hesdorffer CS, Longo DL. Drug-Induced Megaloblastic Anemia. *N Engl J Med*. 2015;373(17):1649-1658. doi:10.1056/NEJMra1508861

Chapter 2: Design of a Massively Parallel Sequencing Study to Discover Rare Variants in *SLC2A2* that Associate with Metformin Response.

INTRODUCTION

Type 2 Diabetes (T2D) is a growing epidemic in the United States and globally. As of 2015, over

30 million individuals in the U.S.[1] and 1 in 11 adults worldwide have diabetes, most of whom

have T2D.[2] Heritability studies have estimated that 33% of the variance observed in this disease

can be attributed to genetic factors[3,4].   Thus, the genetic architecture of T2D has been under

investigation for over two decades using powerful methods like genome-wide association studies

(GWAS).[5] A meta-analysis of thirty-two GWAS studies focused on T2D has demonstrated that

various genes and single variants (i.e. at least 243 genome-wide significant loci) influence risk of

developing T2D.[6] The majority of these genome-wide significant loci each have small effect

sizes (i.e. less than 10% change in risk) and do not drastically increase or decrease risk of

developing T2D individually. Interestingly, common variants (minor allele frequency ≥5%) that

associated with T2D development had risk ratios between 1.03-1.37 whereas low frequency and

rare variants (minor allele frequency (MAF) < 5%) had higher risk ratios between 1.08 to 8.05,

demonstrating that rare variants often have larger influences (or effect sizes) on T2D risk than

common variants. Rare variants can reveal novel biological pathways involved with T2D

development.[6]

According to the American Diabetes Association, first line treatment for T2D is metformin, a

biguanide molecule.[7] In 2017, 79 million prescriptions of metformin were generated in the U.S.

In fact, globally, the use of metformin is so common that it has been detected in lakes and rivers

used to produce potable water in Germany.[8] In 2011, the first GWAS of metformin response was

conducted, which found a genome-wide significant signal near *ATM* (common variants near

ATM, 2011).[9] In 2014, with others, our laboratory, conducted the largest GWAS of metformin response in over 20,000 T2D patients taking metformin.[10] Rs8192675, an intronic SNP in GLUT2, was the only variant to reached genome-wide significance; it was associated with improved response to metformin using change in hemoglobin A1c (HbA1c) as the pharmacologic endpoint.[11]

Specifically, rs8192675 (C allele) was associated with a 0.17% (P = 6.6×10^{-14}) greater reduction in hemoglobin A1c (HbA1c) versus the reference allele (T allele); this reduction was even larger in obese patients (0.33% reduction in HbA1c in obese patients who harbored the C allele). With an average percent reduction in HbA1c for patients initiating metformin of approximately 1%, the polymorphism accounted for 33% of the variance in response in obese patients and 17% of the variance in non-obese patients[12] For other anti-hyperglycemic drug classes, except for insulin and sulfonylureas, the average percent reduction in HbA1c is approximately 0.5%.[13,14] Thus, the presence of rs8192675 in a patient on metformin can provide over a third of the HbA1c reduction that most other anti-hyperglycemic drugs classes provide and two-thirds if the patient is obese. Importantly, the effect of rs8192675 on HbA1c reduction was specific to metformin. In contrast to the effect observed in patients using metformin, rs8192675 exhibited no net pharmacogenetic impact (beta = 0.04%, P = 0.44) on sulfonylurea-induced HbA1c reduction.

Consistent with the genetic analysis, *in vitro* reporter studies revealed a reduced expression of luciferase using an upstream region that contained the rs8192675 C-allele compared with the wild-type upstream region, suggesting this variant may influence *SLC2A2* expression in-vivo.

The evaluation of 1,226 liver samples for expression quantitative trait locus (eQTL) analysis, revealed rs8192675 as the top cis-eQTL for *SLC2A2*, with the C allele associated with decreased ($P = 4.2 \times 10^{-12}$) *SLC2A2* expression.[11] A second, independent GWAS conducted by the German Diabetes Study group revealed that rs8192675 associated with metformin response in newly diagnosed Type 2 Diabetics using change in fasting blood glucose levels.[15] Overall, these two GWAS provide substantial evidence that genetic polymorphisms in *SLC2A2* (i.e. rs8192675) contribute to differing responses to metformin treatment.

Despite the fact that rare variants generally have a larger effect on a phenotype versus common variants,[6] the overwhelming majority of GWAS studies (disease or drug response) have not incorporated rare variants, to date. This may be due to limitations in budget as sequencing may be costly, genotype calling, analysis software[16,17] or standard statistical tests (i.e. linear or logistic regression) that are not powered to analyze rare variants.[18,19,20,21] Several pharmacogenomic sequence association studies, which capture rare variants in the genome via next generation sequence technologies have found statistically significant signals from rare variants in genes identified from previous candidate gene or genome-wide association studies.[22,23,24,25,26]

For example, a GWAS of children with acute lymphoblastic leukemia (ALL) identified several *SLCO1B1* common variants that significantly associated with methotrexate clearance at a genome-wide level (p-value < 1.7×10^{-9}).[27] Three years later, Ramsey et al and colleagues then re-sequenced *SLCO1B1* exons with the aim of identifying rare *SLCO1B1* variants that influence methotrexate disposition. They found eleven non-synonymous rare *SLCO1B1* variants and one

non-synonymous less common variant that associated with metformin clearance. They also

found that, after adjusting for genetic and non-genetic covariates, variance in methotrexate

clearance attributable to *SLCO1B1* rare variants accounted for 1.9% of the 10.7% variance in

methotrexate clearance due to *SLCO1B1*.[22] This example underscores the issue of "missing

heritability" in GWAS, which tend to underestimate the variance explained by genetic variants

compared to family-based heritability estimates.[28]   Though controversial, it has been proposed

that this "missing heritability" in GWAS could be, in part, explained by rare variants, which are

not evaluated in GWAS studies.

Based on this, we hypothesize that rare variants (individually or in combination) in *SLC2A2* will

associate with and influence response to metformin. Our first objective was to design a study that

uses next generation sequencing of DNA from a cohort of extreme metformin responders and

non-responders to capture and analyze rare variants that would not otherwise be captured by

genotyping platforms. Our second objective was to statistically analyze the influence of these

rare variants and functionally characterize the impact of these rare variants on GLUT2 function

and/or expression. Here I describe the study design, discuss various statistical analyses that

could be applied and highlight the general limitations of pharmacogenomic sequencing

association studies. Actual data analyses are forthcoming and will be reported at a later date.

CLINICAL STUDY DESIGN

For this study, we used the METRO cohort (METformin RespOnse, N ~ 1000). METRO patients

were selected from a large consortium of subjects with Type 2 Diabetes who have been on

metformin for at least 3 months (total of 20,000 subjects) that exhibited the best (reduction in

HbA1c > 1%) or the worst (no change or increase in HbA1c level) response to metformin

treatment. The premise behind this type of clinical study design, known as extreme trait design (ETD), is the observation that rare variants are generally enriched in individuals exhibiting extreme versions of a trait (disease or drug response). (Figure 2.1).[29] For example, studies of human disease using ETD have identified novel genes and/or rare variants that associate with the presence or severity of complex diseases such as diabetic retinopathy, atrial fibrillation and Alzheimer's disease.[30,31,32] Additionally, studies of adverse drug reactions have employed ETD to identify variants that increase risk of serious adverse reactions, identifying a genetic marker that can be used proactively to prevent adverse reactions.[23]
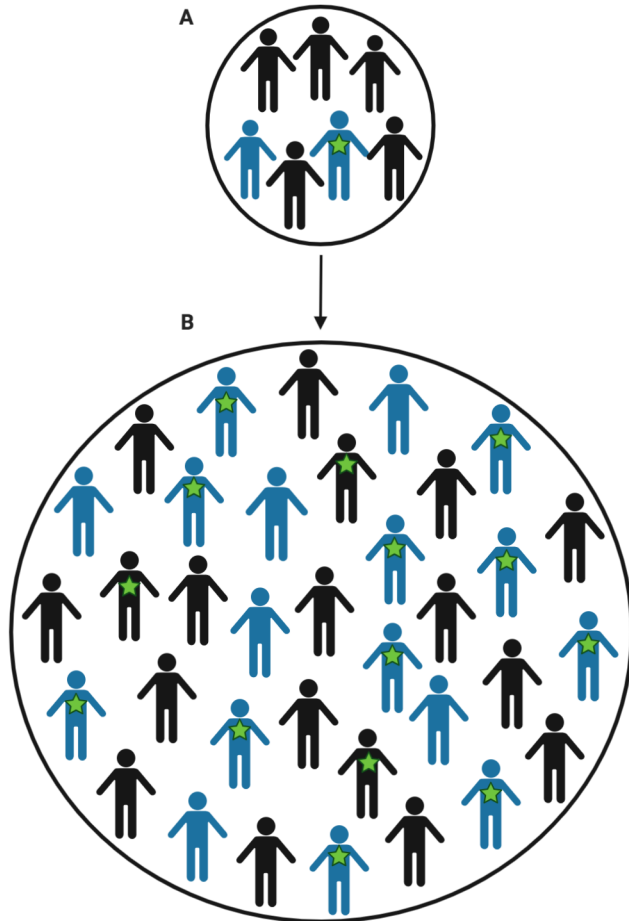
Figure 2.1 Rare variant enrichment in extreme phenotypes of a trait.
The figure shows the distribution of individuals with an extreme phenotype of a trait (blue person), the rare variant of large effect that associates with the extreme phenotype (green star) and the individuals with out that extreme phenotype (black person). In a small sample of the general population, the rare variant of large effect is at a low frequency (A). But in the larger general population, the rare variant of large effect associates with the extreme phenotype (B).

Patient selection for the METRO cohort was determined as follows. First, to generate a

phenotype distribution, the absolute change in HbA1c was plotted for each of 20,000 patients in

the Kaiser Permanente Northern California database with a diagnosis of T2D who are initially

placed on metformin. Next, the top and bottom 2.5% (N ~ 500 per group) was selected from the

distribution; the top 2.5% being extreme responders to metformin and the bottom 2.5% being

extreme non-responders to metformin (Figure 2.2). Then these individuals were contacted via

letters and/or telephone and recruited to enroll into the METRO study with a target goal of 500

extreme non-responders and 500 extreme responders. Primary component analysis (PCA) was

used to verify the genetic ancestry of each subject and to confirm a homogenous genetic

population (European).
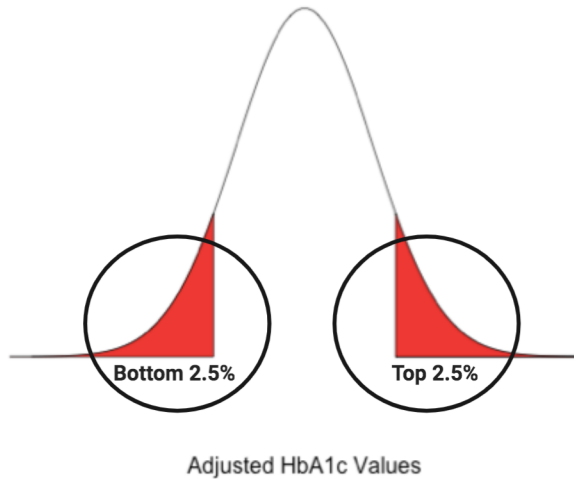
Figure 2.2 Phenotype (metformin response) distribution curve.
The figure represents the phenotype distribution curve for 20,000 T2D patients who initiated metformin for at least 3 months. The number of patients in each "bin" is plotted on the y-axis and the change in HbA1c is plotted on x-axis. Five hundred patients with the largest change in HbA1c (top 2.5%) were selected as extreme responders and five-hundred patients with the smallest change in HbA1c (bottom 2.5%) were selected as extreme non-responders.

After recruitment of subjects and extraction of DNA, the stored genomic DNA samples of the

1000 METRO patients were sequenced using next generation sequencing technology (GenScript

Inc). Specifically, we conducted targeted sequencing of *SLC2A2* by sequencing the eleven exons

of *SLC2A2* plus the 5' and 3' untranslated regions (UTR) of each sample. Raw sequence reads

(fastq files) were processed, indexed, aligned, duplicates removed and variants called using the

Sentieon pipeline ®.[33] We targeted an average depth of coverage per base of at least 50X.

Variants were stringently filtered then compiled into the finalized vcf file. Rare non-synonymous

variants (missense, nonsense, in-frame indels) that are identified in the vcf file will be verified

using PCR amplification and sanger sequencing (MCLAB Inc). 1000 Genomes has identified

5154 genetic variants in *SLC2A2*; less than half of those variants (2258) have a MAF ≥ 0.001.[34]

Table 2.2 lists the sample size needed to detect variants with each respective MAF assuming a

99.9% probability of detection. [34,35] Once sequencing of *SLC2A2* was completed and the genetic

data processed and compiled, we detected 326 variants.

Table 2.1 Detection of low frequency variants varies by sample size.

MAF	0.05	0.03	0.01	0.005	0.0043	0.0034	0.001
Minimum n	68	114	344	690	802	1000	3453

The table shows the minimum number of samples/patients needed to detect (at a 99.9% probability) a mutation at various minor allele frequencies (MAF). As the MAF of variants decreases, the required number of samples needed to detect those variants increases. At a sample size of 1000, there is a 99.9% probability to detect variants with a MAF $\geq$ 0.0034.

EVALUATION OF PROPOSED ANALYSES METHODS

The METRO cohort is a deeply phenotyped cohort that has collected dozens of clinical and laboratory traits such as HbA1c (pre and post-metformin treatment), serum creatinine, age, weight, and metformin dose. Because the ETD divides the METRO cohort into two groups (i.e. cases and controls), we can conduct different types of statistical analyses. Below I describe the potential analyses that can be conducted.

- First, we can use simple statistical analyses for categorical variables (chi-squared or fisher exact test) to count the number of mutations in cases and controls and evaluate if the differences in those counts are statistically significant.

- Second, we can use simple statistical analyses for numerical variables (student's t-test) by determining the average number of variants per person in cases and controls and testing if the difference in averages is statistically significant. Conveniently, the same analysis can be repeated using different definitions of "variants" (i.e. common and rare variants, rare variants only, non-synonymous variants only, rare and non-synonymous variants only).

- Technically, we could also conduct a logistic regression analysis and calculate the odds ratio (OR) of each *SLC2A2* mutation being in case versus control patients. Logistic regression is often used in case-control GWAS studies. For instance, if a mutation has a MAF = 0.025 (i.e. a less common variant), a sample size of 1,352 would provide 80% power to detect the less common mutation with an odds ratio of 3 at an alpha level = 10^{-6}.[18] However, if a mutation with an odds ratio = 3 has a MAF = 0.01 (i.e. a rare variant), a sample size of 3,236 would be required to have 80% power to detect the mutation at an alpha level = 10^{-6}.[18] At a MAF = 0.001, 31,510 individuals are needed to detect an odds ratio of 3 at 80% power.[18] Thus we

are too underpowered to successfully identify rare variant associations with logistic regression.[18,36,37]

- Additionally, because change in HbA1c is a numeric variable (and many other clinical phenotypes are numeric values), we can also conduct more complex statistical analyses to determine if the presence of these rare variants influences the direction or magnitude of HbA1c reduction. Linear regression is often used in GWAS studies evaluating a numerical phenotype (i.e. blood pressure, height). But similar to the inherent issue of using logistic regression, linear regression is underpowered to detect rare variant associations. Essentially, a common variant with MAF = 0.1 would need 967 patients to detect change of 0.17% in HbA1c at a power = 0.8 and alpha = 0.05, but a rare variant with MAF = 0.01 would need 8832 patients to detect change of 0.17% in HbA1c at a power = 0.8 and alpha = 0.05.[38]

With the explosion of massively parallel sequencing data, newer statistical tests were developed in the last decade to detect rare variant associations that could not be feasibly detected with single marker tests (tests that analyze individual variants separately, e.g. logistic and linear regression).[18] Multi-marker statistical tests, also known as aggregation association tests, jointly test the effect of all variants as a whole and determine if cumulative effect of the aggregated rare variants associates with the phenotype.[39] Each type of aggregation association test provides advantages and disadvantages based on what is known about the underlying biology of the phenotype being evaluated.[39]

- Burden test: The most basic of these aggregation association tests is the burden test, which assesses "whether there is an excess of rare variants in cases or controls."[39] The approach aggregates multiple rare variants in a gene or region into a single "super" variant then

performs the association test on the single "super" variant. Burden tests achieve the greatest power when the combined rare variants all have the same directional effect on the phenotype (Figure 2.3). However, the burden test becomes a disadvantage if the grouped rare variants individually affect the phenotype in different directions; the sums of the negative and positive effects neutralize and result in no association being detected.

- Sequence kernel association test (SKAT): SKAT is a variance component test that was developed in 2011 to address this problem. Unlike burden tests, the SKAT test statistic is robust to aggregation of rare variants with negative and positive affects on the trait because it is the weighted sums of the squared variance [21,39] Both negative and positive variances are squared thus converting all the effects to positive before being summed together (Figure 2.3). Additional tests have been created which combine the power of burden tests (rare variants with the same effect) with the robustness of SKAT to the presence of variants with opposing effects. These tests (CMC, SKAT-O) can be especially effective in the presence of many non-causal variants.[21] Because we cannot predict the direction of each rare variant's effect on metformin response, the most fitting test to select for statistical analysis is SKAT (or SKAT-O).

$$Q_{burden} = \left(\sum_{j=1}^{m} w_j S_j \right)^2$$

$$Q_{SKAT} = \sum_{j=1}^{m} w_j^2 S_j^2$$

Figure 2.3 Burden and SKAT test statistic equations.

The figure shows the equation for the test statistic of the burden test and SKAT test. In the burden test, Wj*Sj is summed for all rare variants then squared; thus positive Wj*Sj and negative Wj*Sj can neutralize each other when summed together. In the SKAT test, Wj and Sj are each squared before being summed creating a positive magnitude of effect for each variant before being combined.

LIMITATIONS OF STUDY DESIGN AND PROPOSED ANALYSES

Note that sequencing association studies that analyze rare variants have limitations that should be addressed before study initiation. The first involves the cost of NGS sequencing large numbers of DNA samples. The cost of NGS will vary depending upon the type of sequencing platform selected (Illumina vs Pacific Biosciences; whole genome versus whole exome versus targeted region sequencing) and number of samples to be sequenced. Generally, the cost of whole genome sequencing per sample is more expensive compared to whole exome sequencing per sample or targeted region sequencing per sample.[40] Budget restraints may require selection of a cheaper NGS platform to maximize the number of samples sequenced or selection of a smaller cohort in order to use an optimal, but expensive NGS platform. ETD may enrich for rare variants and improve detection power while keeping the total number of samples at an affordable level.

The second limitation involves identifying replication cohorts to validate significant rare variant associations from the pharmacogenomic study. Since there are no more than ten published pharmacogenomic sequence association studies, a replication cohort is unlikely to exist. If a replication cohort does exist, a meta-analysis (seqMeta, R package) can be conducted to combine the two cohorts and provide increased power.[21] If a replication cohort does not exist, the effect of the rare variant(s) on the phenotype can be validated in-vivo using transgenic animal models or clinical trials of humans that express the rare variant(s).

CONCLUSIONS AND NEXT STEPS

To summarize, we have designed metformin sequencing association study to discover and characterize rare *SLC2A2* variants that contribute to metformin response. DNA samples from the METRO cohort, T2D patients who experienced a great (extreme responders) or poor (extreme non-responders) response to metformin, were sequenced and the genetic data processed and compiled into a vcf file. We explored several statistical tests, which can be used to analyze our data. The simplest methods involve counting the total number of variants in cases versus controls using a 2x2 contingency table (chi-square) or obtaining the average number of variants per person in cases versus controls (student's t-test) then testing if the differences are statistically different.

Complex aggregation tests (burden test, SKAT), which increase power to detect rare variant associations, have been used successfully for numerical phenotypes in normally distributed data. But because we employed ETD, our data are not normally distributed (top and bottom 2.5% of the distribution); thus these aggregation tests may not be ideal. Therefore, we will execute the simplest statistical analyses (chi-square, student's t-test) and consider more complex analyses if needed. The next steps of this project involve the execution of the aforementioned data analyses and functional evaluation of relevant variants. The results will be reported at a later date.

REFERENCES

1. Centers for Disease Control and P. *Diabetes 2017 Report Card*. Atlanta, Georgia; 2017.

2. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol*. 2018;14(2):88-98. doi:10.1038/nrendo.2017.151

3. Zhou K, Donnelly L, Yang J, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol*. 2014;2(6):481-487. doi:10.1016/S2213-8587(14)70050-6

4. Meigs JB, Cupples LA, Wilson PW. Parental transmission of type 2 diabetes: the Framingham Offspring Study. *Diabetes*. 2000;49(12):2201-2207. doi:10.2337/diabetes.49.12.2201

5. Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann N Y Acad Sci*. 2010;1212:59-77. doi:10.1111/j.1749-6632.2010.05838.x

6. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. 2018;50(11):1505-1513. doi:10.1038/s41588-018-0241-6

7. 15. Diabetes Care in the Hospital: Standards of Medical Care in Diabetes-2020. *Diabetes Care*. 2020;43(Suppl 1):S193-S202. doi:10.2337/dc20-S015

8. Trautwein C, Berset J-D, Wolschke H, Kummerer K. Occurrence of the antidiabetic drug Metformin and its ultimate transformation product Guanylurea in several compartments of the aquatic cycle. *Environ Int*. 2014;70:203-212. doi:10.1016/j.envint.2014.05.008

9. The GoDARTS and UKPDS Diabetes Pharmacogenetics Study G, The Wellcome Trust Case Control Consortium 2. Common variants near ATM are associated with glycemic

response to metformin in type 2 diabetes. *Nat Genet*. 2011;43(2):117-120. doi:10.1038/ng.735

10.   Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet*. 2016;48(9):1055-1059. doi:10.1038/ng.3632

11.   Zhou K, Yee SW, MetGen coauthors, Giacomini K PE. *Variation in the Glucose Transporter Gene GLUT2 (SLC2A2) Is Associated with Glycaeic Response to Metformin: A MetGen Study*.; 2016.

12.   Hirst JA, Farmer AJ, Ali R, Roberts NW, Stevens RJ. Quantifying the effect of metformin treatment and dose on glycemic control. *Diabetes Care*. 2012;35(2):446-454. doi:10.2337/dc11-1465

13.   Buse JB, Henry RR, Han J, Kim DD, Fineman MS, Baron AD. Effects of exenatide (exendin-4) on glycemic control over 30 weeks in sulfonylurea-treated patients with type 2 diabetes. *Diabetes Care*. 2004;27(11):2628-2635. doi:10.2337/diacare.27.11.2628

14.   Hirst JA, Farmer AJ, Dyar A, Lung TWC, Stevens RJ. Estimating the effect of sulfonylurea on HbA1c in diabetes: a systematic review and meta-analysis. *Diabetologia*. 2013;56(5):973-984. doi:10.1007/s00125-013-2856-6

15.   Rathmann W, Strassburger K, Bongaerts B, et al. A variant of the glucose transporter gene SLC2A2 modifies the glycaemic response to metformin therapy in recently diagnosed type 2 diabetes. *Diabetologia*. 2019;62(2):286-291. doi:10.1007/s00125-018-4759-z

16.   Wright CF, West B, Tuke M, et al. Assessing the Pathogenicity, Penetrance, and Expressivity of Putative Disease-Causing Variants in a Population Setting. *Am J Hum Genet*. 2019;104(2):275-286. doi:10.1016/j.ajhg.2018.12.015

17. Weedon MN, Jackson L, Harrison JW, et al. Very rare pathogenic genetic variants detected by SNP-chips are usually false positives: implications for direct-to-consumer genetic testing. *bioRxiv*. January 2019:696799. doi:10.1101/696799

18. Witte JS. Rare Genetic Variants and Treatment Response: Sample Size and Analysis Issues. *Stat Med*. 2012;31(25):3041-3050. doi:10.1002/sim.5428

19. Dutta D. Multi - SKAT : General framework to test for rare - variant association with multiple phenotypes. 2019;(July 2018):4-23. doi:10.1002/gepi.22156

20. Wheeler E, Leong A, Liu C-T, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med*. 2017;14(9):e1002383. doi:10.1371/journal.pmed.1002383

21. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *Am J Hum Genet*. 2014;95(1):5-23. doi:10.1016/j.ajhg.2014.06.009

22. Ramsey LB, Bruun GH, Yang W, et al. Rare versus common variants in pharmacogenetics: SLCO1B1 variation and methotrexate disposition. *Genome Res*. 2012;22(1):1-8. doi:10.1101/gr.129668.111

23. Weeke P, Mosley JD, Hanna D, et al. Exome sequencing implicates an increased burden of rare potassium channel variants in the risk of drug-induced long QT interval syndrome. *J Am Coll Cardiol*. 2014;63(14):1430-1437. doi:10.1016/j.jacc.2014.01.031

24. Chua EW, Cree S, Barclay ML, et al. Exome sequencing and array-based comparative genomic hybridisation analysis of preferential 6-methylmercaptopurine producers. *Pharmacogenomics J*. 2015;15(5):414-421. doi:10.1038/tpj.2015.9

25. Li M, Mulkey F, Jiang C, et al. Identification of a Genomic Region between SLC29A1 and HSP90AB1 Associated with Risk of Bevacizumab-Induced Hypertension: CALGB 80405 (Alliance). *Clin Cancer Res*. 2018;24(19):4734-4744. doi:10.1158/1078-0432.CCR-17-1523

26. Mak ACY, White MJ, Eckalbar WL, et al. Whole-Genome Sequencing of Pharmacogenetic Drug Response in Racially Diverse Children with Asthma. *Am J Respir Crit Care Med*. 2018;197(12):1552-1564. doi:10.1164/rccm.201712-2529OC

27. Trevino LR, Shimasaki N, Yang W, et al. Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. *J Clin Oncol*. 2009;27(35):5972-5978. doi:10.1200/JCO.2008.20.4156

28. van der Sluis S, Verhage M, Posthuma D, Dolan C V. Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS One*. 2010;5(11):e13929. doi:10.1371/journal.pone.0013929

29. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010;11(6):415-425. http://dx.doi.org/10.1038/nrg2779.

30. Shtir C, Aldahmesh MA, Al-Dahmash S, et al. Exome-based case-control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum Genet*. 2016;135(2):193-200. doi:10.1007/s00439-015-1624-8

31. Tsai C-T, Hsieh C-S, Chang S-N, et al. Next-generation sequencing of nine atrial fibrillation candidate genes identified novel de novo mutations in patients with extreme trait of atrial fibrillation. *J Med Genet* . 2015;52(1):28-36. doi:10.1136/jmedgenet-2014-

102618

32.   Nho K, Corneveaux JJ, Kim S, et al. Whole-exome sequencing and imaging genetics
       identify functional variants for rate of change in hippocampal volume in mild cognitive
       impairment. *Mol Psychiatry*. 2013;18(7):781-787. doi:10.1038/mp.2013.24

33.   Freed D, Aldana R, Weber JA, Edwards JS. The Sentieon Genomics Tools - A fast and
       accurate solution to variant calling from next-generation sequence data. *bioRxiv*. January
       2017:115717. doi:10.1101/115717

34.   Consortium T 1000 GP. A global reference for human genetic variation. *Nature*.
       2015;526(7571):68-74. doi:10.1038/nature15393

35.   Lee, S., Wu, M. C., Cai, T., Li, Y., Boehnke, M., & Lin X. *Power and Sample Size
       Calculations for Designing Rare Variant Sequencing Association Studies.*; 2011.

36.   Demidenko E. Sample size determination for logistic regression revisited. *Stat Med*.
       2007;26(18):3385-3397. doi:10.1002/sim.2771

37.   Demidenko E. Sample size and optimal design for logistic regression with binary
       interaction. *Stat Med*. 2008;27(1):36-46. doi:10.1002/sim.2980

38.   Dupont WD, Plummer WDJ. Power and sample size calculations for studies involving
       linear regression. *Control Clin Trials*. 1998;19(6):589-601. doi:10.1016/s0197-
       2456(98)00037-3

39.   Li B, Liu DJ, Leal SM. Identifying rare variants associated with complex traits via
       sequencing. *Curr Protoc Hum Genet*. 2013;Chapter 1:Unit 1.26.
       doi:10.1002/0471142905.hg0126s78

40.   Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*.
       2010;11(1):31-46. doi:10.1038/nrg2626

Chapter 3: Functional and Structural Analysis of Rare *SLC2A2* Variants Associated with Fanconi-Bickel Syndrome and Metabolic Traits

ABSTRACT

Deleterious variants in *SLC2A2* cause Fanconi-Bickel Syndrome (FBS), a glycogen storage disorder, whereas less common variants in *SLC2A2* associate with numerous metabolic diseases. Phenotypic heterogeneity in FBS has been observed, but its causes remain unknown. Our goal was to functionally characterize rare *SLC2A2* variants found in FBS and metabolic disease-associated variants to understand the impact of these variants on GLUT2 activity and expression and establish genotype-phenotype correlations. Complementary RNA-injected Xenopus laevis oocytes were used to study mutant transporter activity and membrane expression. GLUT2 homology models were constructed for mutation analysis using GLUT1, GLUT3, and XylE as templates. Seventeen FBS variants were characterized. Only c.457_462delCTTATA (p.Leu153_Ile154del) exhibited residual glucose uptake. Functional characterization revealed that only half of the variants were expressed on the plasma membrane. Most less common variants (except c.593 C>A (p.Thr198Lys) and c.1087 G>T (p.Ala363Ser)) exhibited similar GLUT2 transport activity as the wild type. Structural analysis of GLUT2 revealed that variants affect substrate-binding, steric hindrance, or overall transporter structure. The mutant transporter that is associated with a milder FBS phenotype, p.Leu153_Ile154del, retained transport activity. These results improve our overall understanding of the underlying causes of FBS and impact of GLUT2 function on various clinical phenotypes ranging from rare to common disease.

INTRODUCTION

Pathogenic variants in *SLC2A2*, the gene that encodes for the glucose transporter, GLUT2, have been associated with Fanconi-Bickel Syndrome (FBS; OMIM# 227810), a rare glycogen storage disorder that has over 100 documented cases published.[1,2] These deleterious, rare variants (minor allele frequency < 1%) effectively abolish GLUT2 transporter function, which leads to significant physiological sequelae. Currently, forty-six *SLC2A2* variants have been identified and associated with FBS; these include missense, nonsense, in-frame indels (insertions/deletions), frame-shift indels and splice site variants. FBS is an autosomal recessive disease; in order for FBS to manifest in an individual, the person must be homozygous or compound heterozygous with two pathogenic *SLC2A2* variants. Figure 3.1 displays the location of many of the GLUT2 variants that are causal for FBS; these variants are scattered throughout GLUT2.
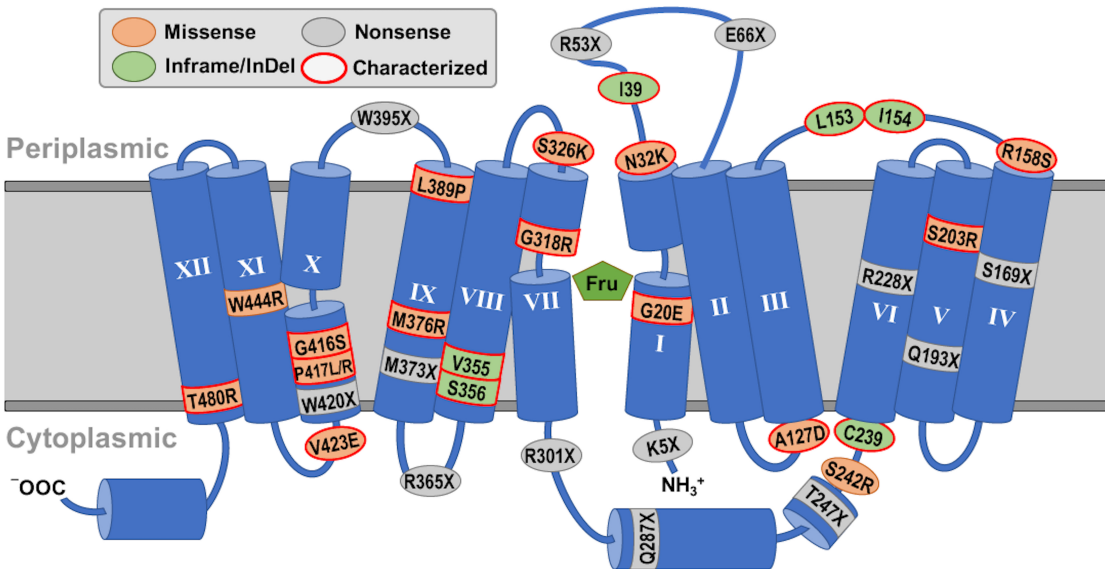
Figure 3.1 Fanconi-Bickel syndrome (FBS) variant location within GLUT2 transporter secondary structure.

An illustration of the secondary structure of GLUT2 that displays the location of all published missense (orange), nonsense (grey), and in-frame indel (green) *SLC2A2* variants associated with FBS. The majority of these FBS variants had not been evaluated in-vitro to test their impact on GLUT2 membrane expression or function. We selected seventeen variants (red border) to functionally characterize in our study. Roman numerals I-XII represent each transmembrane helix in the GLUT2. The green pentagon labeled "Fru" represents fructose.

Common signs and symptoms of the syndrome include hepatomegaly, stunted growth and rickets

due to renal losses of calcium and phosphate, high post-prandial and low pre-prandial blood

glucose levels. These typical manifestations are directly related to the expression pattern of

GLUT2 in particular body tissues (liver, kidney, intestine, and to a lesser extent pancreatic beta

cells) [3]. As the primary glucose transporter in hepatocytes, GLUT2 regulates the normal flux of

glucose into and out of the liver, and its loss leads to excessive glycogen storage (resulting in

hepatomegaly). In the kidney, GLUT2 is located in the proximal tubule and is involved in the

reabsorption of glucose from urine into the blood. The absence of GLUT2 in the kidney causes

massive glucosuria plus significant loss of minerals required for bone growth and strength

(calcium, phosphate) leading to bone deformities.[4] However, as with many diseases, FBS

presents with a spectrum of phenotypes. Whereas some patients develop severe symptoms that

are recalcitrant to treatment such as bone fractures[5,6], hepatocellular carcinoma[7], liver failure and

premature death[8,9], others present with a "milder" clinical picture (relatively normal growth, no

hepatomegaly, minor kidney dysfunction) that can be corrected with conservative measures[10].

Although GLUT2 variants associated with published cases of FBS have been identified, the

majority have not been evaluated for their impact on GLUT2 function. To date, only five FBS

variants have been functionally characterized (c.724A>C, p.Ser242Arg; c.1330T>C,

p.Trp444Arg; c.1250C>T, p.Pro417Leu; c.59G>A, p.Gly20Asp; c.372A>C, p.Arg124Ser)[11,12].

Furthermore, the relationship between the functional effects of any of the mutant transporters and

the phenotypic presentation of the disease has not been established. Since 2000, many new cases

of FBS from around the world have been identified revealing a host of novel variants, but these

new variants have not been functionally evaluated or characterized.[7,10,13,14,15, 16,17,18,19,20] Cases of

heterozygote carriers of FBS variants, i.e. parents or siblings of FBS patients, that experience glycemic perturbations (i.e. glycosuria, impaired glucose tolerance) have been documented [21]and have led to speculation that GLUT2 proteins could exhibit a dominant negative effect (i.e. haploinsufficiency) in individuals heterozygous for an FBS GLUT2 missense variant.

In addition to individuals who are homozygous for rare, deleterious variants in GLUT2 that are causal for FBS, those who harbor at least one less common (minor allele frequency between 1 - 5%) or common (minor allele frequency > 5%) *SLC2A2* allele may be at risk for various metabolic traits. For example, genome-wide association studies (GWAS) have identified common and less common variants in *SLC2A2* that associate with an increased risk of Type 2 diabetes (T2D)[22], as well as increases in fasting glucose[23,24], glycated hemoglobin levels (HbA1c)[25,26] and even differential response to metformin.[27] Furthermore, several less common *SLC2A2* variants that are only found in specific populations significantly associate with increased risk of T2D, elevated blood glucose, and cholesterol levels and other metabolic perturbations (Type 2 Diabetes Knowledge Portal: http://www.type2diabetesgenetics.org). Pharmacochaperones are drugs that improve the expression of misfolded mutant proteins. Several pharmacochaperones are FDA-approved to treat mutated enzyme diseases (i.e. migalastat for Fabry disease) [29] and most recently, ionchannelopathies (i.e. tezacaftor or lumacaftor for cystic fibrosis).[30]. However, there are no FDA-approved small molecule drugs for genetic diseases caused by a mutant or misfolded SLC transporter protein. Additionally, premature termination codon (PTC) modulating agents are a structurally diverse group of compounds that interferes with the recognition of nonsense codons promoting the complete translation of protein that would have otherwise been truncated then degraded via nonsense

mediated decay. Various PTC-modulating agents (i.e. erythromycin, gentamicin, ataluren) have been clinically-evaluated for different genetic diseases (familial adenomatous polyposis (FAP), duchenne muscular dystrophy (DMD), cystic fibrosis) generating mixed results in efficacy. To date, there are no published data demonstrating improved readthrough of PTCs in genes encoding for SLC transporters.

In addition to glucose, GLUT2 transports and distributes other important monosaccharides throughout the body (i.e. fructose, glucosamine)[31,32] and has a broader substrate range than other GLUT transporters.[33] Studies have identified other hexoses and pentoses (i.e. 1,5 anhydroglucitol, ribose) as substrates of a carrier-mediated process in tissues that express GLUT2 (liver, kidney);[34,35,36] using in-vitro studies, ribose was found to be substrate of mouse GLUT2.[37] But, to date, these monosaccharides have not been directly tested to determine if they are substrates of human GLUT2.

In this study, I performed detailed functional analyses of all reported, but uncharacterized missense and in-frame indel FBS variants, as well as less common variants that associate with metabolic perturbations. In order to determine the impact of these variants on GLUT2 transporter function, our studies focused on assessing the function of mutant or polymorphic transporters, as well as plasma membrane expression. A GLUT2 homology model was used to propose the molecular mechanisms by which particular FBS variants cause GLUT2 structural instability and abrogate GLUT2 transporter function. Subsequently, I tested the effectiveness of several pharmacochaperones and premature termination codon (PTC) modulating agents in specific FBS-associated variants (missense, indel and nonsense). Finally, we tested a variety of

monosaccharides (e.g. pentoses, hexoses) to determine if any are substrates of human GLUT2. Overall, this study creates a comprehensive catalog of the functional characteristics of all reported missense and in-frame indel variants associated with FBS and metabolic perturbations, along with a GLUT2 structure-function map. *In vitro* function for some of the GLUT2 variants is related to clinical phenotypic characteristics suggesting that understanding the functional effects of variants in GLUT2 may help explain the phenotypic heterogeneity within the FBS patient population, as well as the influence of *SLC2A2* polymorphisms on metabolic traits.

METHODS

Chemicals and Reagents

The following radionuclides were obtained from American Radiolabeled Corporation, Inc (ARC) and Roche then tested as substrates of human GLUT2: [^{3}H]-ribose (ARC), [^{3}H]-xylose (ARC), [^{3}H]-1,5 anhydroglucitol (Roche). The following pharmacochaperones compounds were obtained from Sigma Aldrich: Alvespimycin (17-DMAG), 4-phenylbutyric acid (4-PBA), Corr-4a, vorinostat, luminespib, ursodeoxycholic acid (URDA). The following sterile-filtered PTC agents were obtained from Sigma Aldrich: G418, Ataluren.

Mutation selection for functional and structural analysis

The approach for variant selection was divided into two categories: variants associated with FBS and variants associated with common metabolic traits. Published FBS case reports, case series and reviews were acquired via PubMed and various databases. Information on each FBS patient such as demographics, clinical signs and symptoms and genotype, if available, was collected. Cases that did not have clinical data or did not validate clinical FBS diagnosis with a genotype were excluded (approximately twenty cases). The compiled list of variants and associated

phenotypes (87 cases) is available online in the published supplemental (Supplemental Table 3.1, PMID: 30950137). We used the Type 2 Diabetes knowledge portal to only select missense variants associated with metabolic traits (e.g. T2D, HbA1c, fasting glucose, fasting insulin and body mass index; p-value < 0.05). Then we used 1000 Genomes Browser and gnomAD browser to determine the allele frequency of the variants in each population (1000 Genomes Consortium, 2015). Based on this criteria, we selected five *SLC2A2* variants that are rare in the general population, but found in less common frequencies in specific populations and ethnicities.

Mutant plasmid generation using site-directed mutagenesis

For *X. laevis* oocyte studies, site-directed mutagenesis service, performed by GenScript®, was used to create 4 FBS in-frame deletions, 13 FBS missense, and 5 less common missense variants in the human GLUT2 expression vector (pSP64T-GLUT2) provided by Dr. Graeme I. Bell (University of Chicago). All capped cRNA (wild-type and mutant GLUT2) were synthesized *in vitro* from each respective pSP64T-GLUT2 plasmid using the mMessage mMachine SP6 kit (Thermo Fisher Scientific).

For studies in mammalian cells, GLUT2 cDNA (RefSeq NM_000340.1, NP_000331.1) was cloned into the pcDNA5™/FRT mammalian expression vector (Thermo Fisher Scientific) to produce the pcDNA5-GLUT2 plasmid. Site-directed mutagenesis service, performed by GenScript®, was used to introduce the previously mentioned variants plus four nonsense mutations (p.Arg301Ter, p.Arg365Ter, p.Trp395Ter, p.Trp420Ter) into the expression vector (pcDNA5).

Radiolabeled substrate uptake assays in X. laevis oocytes

Ecocyte Bio Science LLC (Austin, TX) provided stage VI, defollicuted oocytes along with oocyte injection services. Six to ten oocytes were selected per construct and each oocyte was

injected with 50nL of saline, wild-type or mutant GLUT2 cRNA (40ng per oocyte). Oocytes

were incubated on a shaker at 18°C in Modified Barth's solution with gentamicin (NaCl 88mM,

KCl 1mM, MgSO$_4$ 1mM, HEPES 5mM, NaHCO$_3$ 2.5mM, CaCl$_2$ 0.7mM, Gentamicin

100µg/mL; Sigma-Aldrich) that was replenished every 24 hours. Forty-eight to seventy-two

hours post-injection, thirty-minute uptake assays were performed using labeled [^{3}H]-2-deoxy-

glucose (10µM total of 2-deoxy-glucose, PerkinElmer). Modified Barth's solution was used as

the uptake buffer. Previous time range experiments revealed that 30 minutes was within the

linear range of GLUT2 uptake. After uptake, oocytes were washed 5 times in cold uptake buffer

then lysed in 0.5mL of 1% SDS (sodium dodecyl sulfate) solution (VWR International) for at

least 1 hour. Radioactivity (disintegrations per minute) was measured using liquid scintillation

counter machine. Radiolabeled substrate uptake rates were calculated by converting the DPM

(disintegrations per minute) values of each oocyte to picomoles of substrate. Using the same

process as mentioned, several radioactive monosaccharides were evaluated to identify novel

substrates of human GLUT2, including [3H]-xylose, [3H]-1,5 anhydroglucitol (1,5 AG), [3H]-

ribose. Additionally, using the same aforementioned process, GLUT2 mutant oocytes were

exposed to pharmacochaperones for 24 hours and their uptake of [^{3}H]-2-deoxy-glucose (10µm)

was measured. Figures are representative of 3-5 experiments, except Figures 3.6-3.9.

Transfection and establishment of mammalian cell line transiently expressing GLUT2 reference and mutant cDNA

Flp-In™ 293 cells (human embryonic kidney cells, Thermo Fisher Scientific) were transfected

with select pcDNA5/FRT-GLUT2 mutant constructs using Lipofectamine® LTX with Plus™

reagent according to the manufacturer's instructions (Life Technologies) to confirm uptake and

membrane expression results found in oocyte assays. Approximately 200,000 Flp-In™ 293 cells

were plated in 24-well format; once each well reached 90% confluence, transient transfection

was performed (4 replicates per construct). Transfected cells were maintained in DMEM

containing 10% (v/v) FBS, 100 units/mL penicillin, 100 μg/mL streptomycin, and 100 μg/mL

Zeocin (Thermo Fisher Scientific) at 37 °C and 5% CO_2 for 48 hours. Thereafter, thirty-minute

uptake assays were performed using [^{3}H]-fructose (100 μM total of fructose). Cells were lysed

for one hour using 800 μL of lysis buffer (0.1N NaOH, 0.1% SDS) and the radioactivity of

500μL of lysate was measured using a liquid scintillation counter machine. The DPM values

were converted to picomoles of substrate. Total protein concentrations were determined using

the bicinchoninic acid (BCA) assay using 25 μL of lysate per well (Thermo Fisher Scientific).

Substrate uptake (in picomoles) was normalized by protein amount in each respective well.

Using the same aforementioned process, cells were transfected with pcDNA5/FRT-GLUT2

nonsense variant plasmids, exposed to PTC agents (G418, ataluren) for 24-48 hours, then their

uptake of [^{3}H]-ribose (1μM) was measured.

Plasma membrane expression assays

Forty-eight to seventy-two hours post oocyte injection, GLUT2 plasma membrane expression

experiments were conducted for each FBS and population-specific mutant (6-10 oocytes).

Oocytes were placed in blocking buffer (1% bovine serum albumin solution) for 30 minutes at

4°C. Next, oocytes were incubated in primary antibody solution (anti-GLUT2 mouse IgG

antibody; MAB 1414, R&D systems) for 30 minutes at 4°C and then washed in blocking buffer

six times. After washing, oocytes were transferred into secondary antibody solution (anti-mouse

IgG VisUCyte HRP conjugate, R&D systems) for 30 minutes at 4°C then washed 6 times in

blocking buffer and 6 times in ND96 buffer (NaCl 96mM, KCl 2mM, $CaCl_2$ 1.5mM, $MgCl_2$

1mM, HEPES 5mM; Sigma-Aldrich). Each oocyte was transferred to the well of a white 96-well

plate (catalog #3912, Corning Inc); 100 μL of SuperSignal ELISA Femto (Fisher Scientific) mix

was added to each oocyte. One minute later, luminescence was measured using Glomax 96-well plate luminometer (Promega). All blocking and antibody solutions were produced using ND96 buffer. Plasma membrane expression was defined as a mutant transporter whose average luminescence's measurement was statistically similar to the average luminescence measurement of wild-type GLUT2. Using the same aforementioned process, GLUT2 mutant oocytes were exposed to pharmacochaperones for 24 hours and mutant transporter expression on the plasma membrane was measured via luminescence. Figures are representative of 3-5 experiments, except Figures 3.6-3.9.

Kinetic studies of 2-deoxyglucose transport by GLUT2 and mutants

The transporter kinetics (V_{max} and K_m) of functional mutants were generated and compared with wild-type GLUT2. Uptake was linear between 0 – 40 minutes (0-60 minutes tested) so 30 minutes was selected to conduct uptake and kinetic studies. The thirty-minute uptake of [^{3}H]-2DG in a range of 2DG concentrations (0-100mM) was measured as previously described and plotted as rate of uptake versus concentration to generate non-linear regression curves. The K_m and V_{max} were calculated by fitting the curves to a Michaelis–Menten equation using GraphPad Prism 7 (La Jolla, CA).

Statistical analysis

For comparison in radiolabeled uptake assays and plasma membrane expression assays, Student's t-test and one-way ANOVA statistical tests were performed. Statistical analyses, as specified in the legends of the figures and tables, were performed to determine significance, and a p-value < 0.05 was considered significant. Unless specified, data in figures and tables were expressed as mean ± standard error of the mean (SEM). All experiments (except for experiments

represented by Figures 3.6-3.9) were performed at least 3 to 5 times and in 6 to 10 oocytes or in 4 replicates in mammalian cell lines.

Homology model construction for structural analysis

Homology models of GLUT2 were constructed with MODELLER v9.14 for mapping *SLC2A2* variants [39]. Specifically, three conformational states of GLUT2 were modeled: an inward-open state based on the human GLUT1 (PDB identifier 4PYP; sequence identity of 51.7%) [40]; an occluded state based on a *Escherichia coli* XylE transporter structure (4GC0; 23.8%) [41]; and an outward-open state based on the human GLUT3 (4ZWC; 48.1%) [42]. The alignment between GLUT2 and the template was generated using Promals3D [43]. The models were evaluated and ranked with Z-DOPE [44] obtaining scores of -0.70, -1.20, and -0.85, suggesting they are sufficiently accurate for further analysis.

To deduce the structural effects of the FBS missense variants on *SLC2A2*, variants that were functionally evaluated and expressed on the plasma membrane (nine mutations) were modeled and studied using molecular dynamics (MD) simulations.

CHARMM-GUI Membrane Builder [45] was used to construct and introduce the variants to the GLUT2 homology model in the outward-open conformation in lipid bilayer for MD simulations [46]. Protein and the lipid bilayer system were parameterized with AMBER formatted CHARMM36m force fields (Brooks et al., 2009; Huang et al., 2017).[47,48] Cholesterol (10%), phosphatidylcholine DOPC (55%), and phosphatidylethanolamine DOPE (35%) were used to construct the system in rectangular box with same x and y dimensions and at least 18 Å from protein to box edge. The final average area of the lipid bilayer was 7,600 $Å^2$ with 13,704 TIP3P water molecules[49] [49]and 0.15M NaCl concentration, with ~ 72,500-72,600 atoms for the various mutations. AMBER16 GPU-enabled PMEMD was used to perform the MD simulations.[50,51] The

systems were minimized with 500 steps and equilibrated with NVT and NPT at 310.15 K, while

hydrogens were constrained with SHAKE to enable 2-fs time step [52] 1 ns of equilibration

followed by 25 ns production run were performed for all systems. Trajectories were analyzed

with CPPTRAJ [53] and visualized with PyMOL (The PyMOL Molecular Graphics System,

Version 1.7 Schrödinger, LLC).

RESULTS

Mutations in GLUT2, which are cause for FBS, are present throughout the protein, and in general, alter an evolutionarily conserved amino acid residue

I collected information on over 70 patients with FBS including the patient's genotype,

demographic characteristics, and available clinical information (Supplemental Table 3.1). The

demographic information reveals that FBS has been reported in every major human ethnic group.

The first case of FBS was reported in Switzerland in 1949 [54]; the first Chinese patients diagnosed

with FBS was in 2011 [20]; and the first African-American patient diagnosed with FBS was in

2016.[55] Our updated FBS case list includes more than ten novel amino acid variants that were

published after the last published FBS review in 2002.[2] These novel variants plus other FBS

variants are scattered throughout the transporter, from the first transmembrane (TM) helix (e.g.

c.59_60delinsAA, p.Gly20Glu) to the start of the carboxyl terminal tail (e.g. c.1439C>G;

p.Thr480Arg) (Figure 3.1). PhyloP scores measure evolutionary conservation at individual

alignment sites expected under neutral drift [56]; positive scores indicate sites that are predicted to

be conserved while negative scores indicate sites that are predicted to be fast-evolving. With few

exceptions, most of the FBS *SLC2A2* variants occur in areas that are highly conserved amongst

one hundred vertebrate species (PhyloP scores ≥ 2) and highly conserved amongst GLUT1 to

GLUT5 transporters (data not shown).[56,57]

The SNP annotation scores of the five *SLC2A2* missense variants associated with metabolic traits

Five previously uncharacterized missense variants in *SLC2A2* that have significant associations with metabolic traits were found to have allele frequencies between 0.1-5% in certain ethnic populations. Amongst these variants, Polyphen-2 [58] predicted that one was probably damaging (c.301G>A, p.Val101Ile), two were possibly damaging (c.593C>A, p.Thr198Lys and c.1087G>T, p.Ala363Ser), and two were benign (c.1402C>G, p.Leu468Val, c.1556G>A, p.Gly519Glu). Among the three variants predicted to be probably or possibly damaging, two of them, p.Val101Ile and p.Ala363Ser, are conserved (PhyloP score > 2 and GERP score > 4). Interestingly, p.Val101Ile is predicted by various prediction scores to be damaging or deleterious (Polyphen-2, SIFT [59], CADD [60], GERP [61]) and predicted by REVEL [62] to be partially damaging (score 0.46 (range from 0 to 1.0)). In addition, valine in position 101 is conserved across human GLUT1 to GLUT5 (data not shown).[57]

Functional characterization of seventeen mutations associated with FBS reveals a mutant transporter with residual function

All of the FBS mutant transporters had significantly reduced uptake of substrate compared with wild-type GLUT2 (Figure 3.2, one-way ANOVA F = 467.7; p-value < 0.0001). The uptake rate of 2DG (2-deoxy-glucose) by most FBS mutant transporters was comparable to that of saline-injected oocytes. With the exception of one mutant (c.457_462delCTTATA, p.Leu153_Ile154del), all FBS mutants exhibited effectively no transport function, with uptake rates on average being approximately 0.5 pmol/0.5hr/oocyte (Figure 3.2). Remarkably, p.Leu153_Ile154del still exhibited residual uptake of 2DG (2.174 pmol/0.5hr/oocyte), which was significantly higher than the other FBS mutants (one-way ANOVA F = 5.591, *p-value = 0.0006) but still much lower than wild-type GLUT2 (student's t-test 37.16 ± 1.1 vs. 2.174 ± 0.23

pmol/0.5hr/oocyte; p-value < 0.0001). Nine of the seventeen evaluated FBS mutant transporters were expressed on the plasma membrane and had similar expression to wild-type GLUT2 (one-way ANOVA F-value = 0.6534; p-value = 0.74). The other eight mutants had partial or negligible plasma membrane expression compared to wild-type GLUT2 (Figure 3.2b, one-way ANOVA F-value = 23.05; p-value < 0.0001).

Figure 3.2 FBS mutant transporter uptake of 2DG and plasma membrane expression. The first graph (a) shows average transport rate of cRNA-injected *X. laevis* oocytes incubated with uptake buffer containing [^{3}H]-2DG and 10μM of unlabeled 2DG for 30 minutes. All FBS mutant transporters had significantly reduced uptake of substrate compared with wild-type GLUT2 (one-way ANOVA F = 467.7; p-value < 0.0001) except for p.Leu153_Ile154del, which had significantly higher 2DG uptake compared to the other FBS mutants (one-way ANOVA F = 5.591, *p-value = 0.0006). The second graph (b) shows luminescence emitted by cRNA-injected oocytes incubated sequentially with anti-GLUT2 (MAB #1414, R&D systems) and HRP-conjugated antibodies. Nine of the seventeen evaluated FBS mutant transporters were expressed on the plasma membrane at similar levels to wild-type GLUT2 (one-way ANOVA F-value = 0.6534; p-value = 0.74). The other eight mutants had partial or negligible plasma membrane expression compared with wild-type (one-way ANOVA F-value = 23.05; p-value < 0.0001). "PHL" in the figure represents phloretin, a canonical inhibitor of GLUT transporters. Data shown represent the mean ± SEM for 3 – 5 independent experiments.

Functional characterization of five missense variants in *SLC2A2*, associated with metabolic traits reveals mutations with partial function

We characterized five missense variants that were associated with metabolic traits. Three out of the five transporters had transport rates (one-way ANOVA F-value = 0.115, p-value = 0.95) statistically comparable to wild-type GLUT2 (Figure 3.3). p.Thr198Lys exhibited a partially reduced transport rate; it retained approximately 50% (23 pmol/0.5hr/oocyte) of transport activity compared to the wild-type (39.2 pmol0.5hr/oocyte, student's t-test, **p-value = 0.005). The rate of 2DG uptake by p.Ala363Ser was similar to that of saline-injected oocytes (0.42 pmol/0.5hr/oocyte).

Except for p.Ala363Ser, all of the above mutants were expressed on the oocyte plasma membrane (one-way ANOVA F-value = 0.3502, p-value = 0.88) in a manner comparable to wild-type GLUT2. The estimated K_m values of all mutant transporters associated with metabolic traits were not statistically different from that of wild-type GLUT2 (Figure 3.4, Table 1; one way ANOVA, p-value = 0.44). Similarly, the estimated V_{max} values of all mutant transporters were not statistically different from that of wild-type GLUT2 (one-way ANOVA, p-value = 0.48) with the exception of p.Thr198Lys, which had a V_{max} value much lower than the wild-type GLUT2 transporter (11±1.7 vs 53±2.0 nmol/0.5hr/oocyte, student's t-test ***p-value = 0.0044). Michaelis-Menten curves could not be generated for p.Ala363Ser (signal undetectable in oocytes).

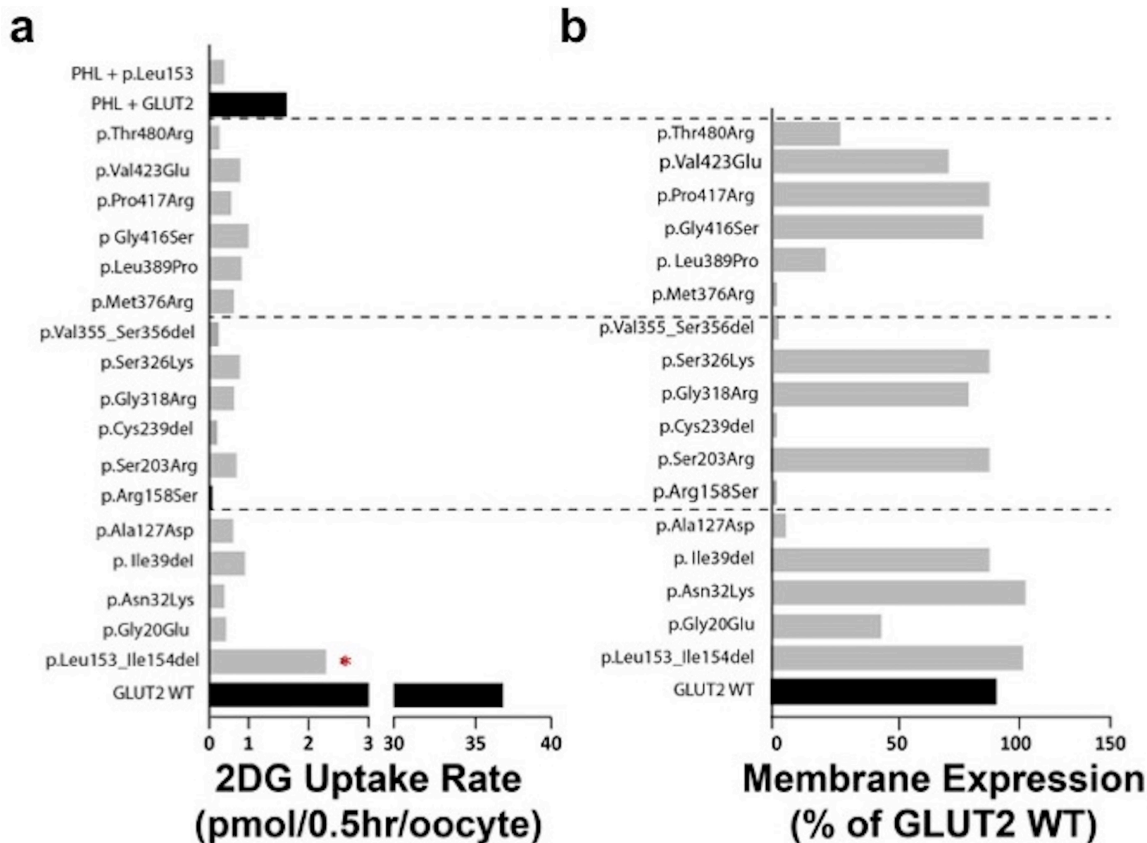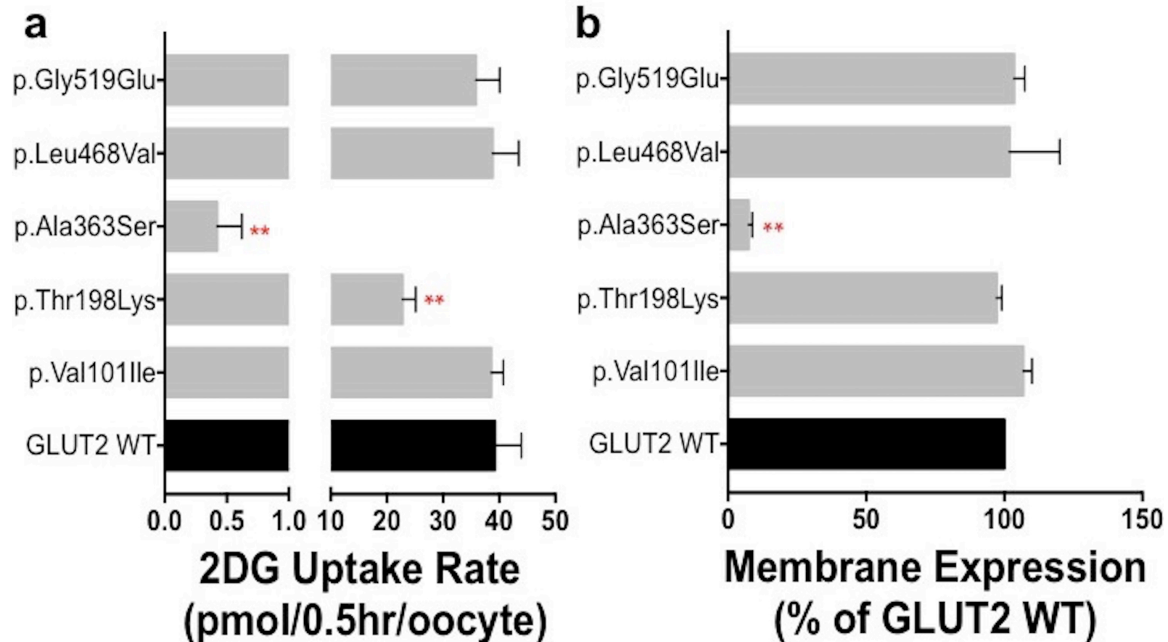**Figure 3.3** Less common mutant transporter uptake of 2DG and plasma membrane expression.

The first graph (a) shows average transport rate of cRNA-injected *X. laevis* oocytes incubated with uptake buffer containing [^{3}H]-2DG and 10μM of unlabeled 2DG for 30 minutes. The uptake rate of three out of the five mutant transporters (p.Val101Ile, p. Leu468Vale, p.Gly519Glu) was comparable to that of wild-type GLUT2 (one way ANOVA F= 0.1149; p-value =0.9509). The uptake rate of p.Thr198Lys (12 pmol/0.5hr/oocyte) was 42% lower than that of wild-type GLUT2 (40 pmol/0.5hr/oocyte, student t-test **p-value ≤ 0.0056). The uptake rate of p.Ala363Ser was negligible and similar to saline-injected oocytes. Figure 3b shows luminescence emitted by cRNA-injected oocytes incubated with anti-GLUT2 then HRP-conjugated antibodies. The second graph (b) reveals that all mutant transporters were expressed on the plasma membrane in similar levels to that of wild-type GLUT2 (one way ANOVA F value = 0.3502, p-value = 0.88) except for p.Ala363Ser, which was not expressed on the oocyte plasma membrane (students t-test 7.4% vs 100%, **p-value < 0.0001). Data shown represent the mean ± SEM for 3 – 5 independent experiments.
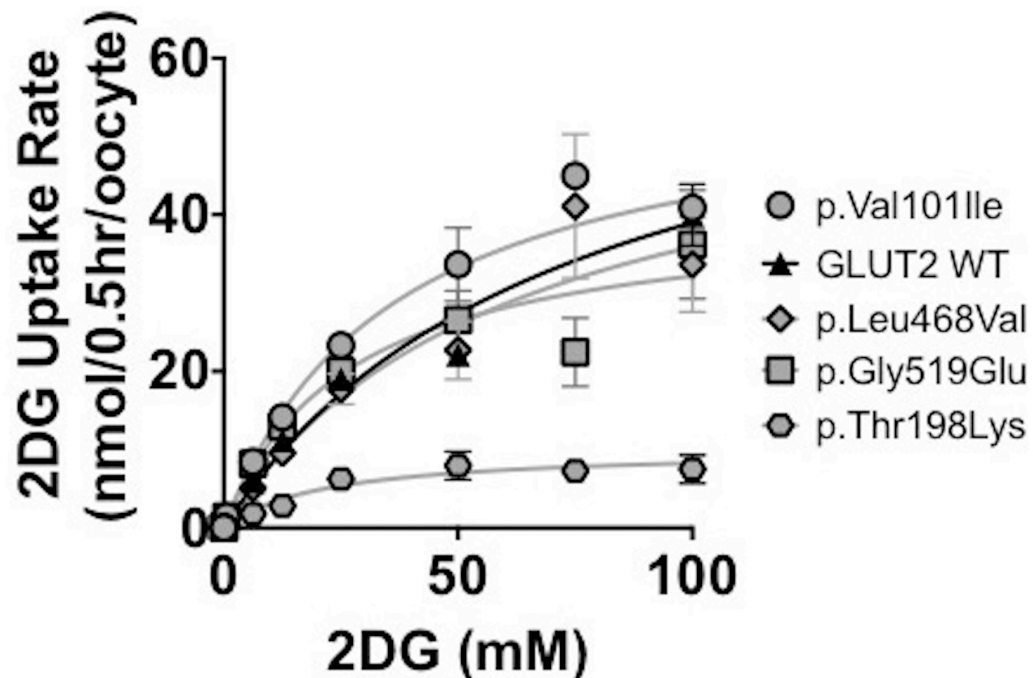
Figure 3.4 Kinetic characterization of less common variants using 2DG. Michaelis-Menten curves for each population-specific variant were generated using 2DG concentrations ranging from 0.001 to 100mM. K_m and V_{max} values of each mutant with the exception of p.Ala363Ser were calculated by fitting the data to a non-linear regression curve. The K_m and V_{max} values of each mutant (mean + SEM) are listed in Table 3.1.

Table 3.1 Estimated kinetic parameters (K_m, V_{max}) for 2DG transport by GLUT2 less common variants.

	K_m (mM) [mean (SE)]	V_{max} (nmol/0.5hr/oocyte) [mean (SE)]
GLUT2 WT	34 (3.3)	53 (2.0)
c.301G>A (p.Val101Ile)	35 (12)	56 (8.1)
c.593C>A (p.Thr198Lys)	26 (10)	11 (1.7)***
c.1087G>T (p.Ala363Ser)	n.d	n.d
c.1402C>G (p.Leu468Val)	56 (13)	62 (8.4)
c.1556G>A (p.Gly519Glu)	31 (13)	45 (7.8)

The K_m values for the four measured mutants are statistically similar to wild-type GLUT2 (one-way ANOVA p-value = 0.44). The V_{max} values of these mutants are also similar (one-way ANOVA p-value = 0.48) except for the significantly reduced V_{max} of p.Thr198Leu when compared with wild-type GLUT2 (11.7 vs 52.7, students t-test ***p-value = 0.0044). K_m and V_{max} shown represent the mean and standard error [mean (SE)] for 2 independent experiments.

Functional characterization of GLUT2 mutant transporters in mammalian cells reveal some differences with mutants expressed in oocytes

The maximum increase in 2DG uptake in Flp-In™ 293 cells expressing full-length wild-type GLUT2 and empty vector was approximately 50-80% (data not shown). To improve the signal-to-noise ratio, the uptake of radiolabeled fructose was tested. The fold difference in fructose uptake between pcDNA5-EV cells and pcDNA5-GLUT2 cells was approximately 2x (data not shown). The improvement in signal-to-noise ratio with fructose is due to its increased selectivity for GLUT2 compared to human GLUT1, GLUT3, and GLUT4 transporters that are present in immortalized mammalian cell lines.[63]

Therefore, [^{3}H]-fructose (fructose 100μM total concentration) was used as the radiolabeled substrate for the following Flp-In™ 293 radiolabeled substrate uptake experiments. As expected, the cells transfected with pcDNA5-GLUT2 FBS mutants had fructose uptake values similar to cells transfected with empty vector (one-way ANOVA, F=1.218, p-value = 0.2989). In contrast, fructose uptake was similar in cells transfected with mutants associated with metabolic traits (p.Val101Ile, p.Ala363Ser, p.Leu468Val, p.Gly519Glu) and cells transfected with full-length wild-type GLUT2 (one-way ANOVA, F=1.693, p-value = 0.1558), except for p.Thr198Lys. Fructose uptake rate in cells expressing p.Thr198Lys was slightly lower than in cells expressing wild-type GLUT2, though higher than in cells transfected with empty vector (student's t-test p-value <0.0001). The results in mammalian cells exhibited a similar trend with the results in the oocytes, except for p.Ala363Ser.

Unexpectedly, we found that p.Ala363Ser transporter, which was not expressed on the plasma membrane of *X. laevis* oocytes (Figure 3.3), was expressed well in Flp-In™ 293 cells and transported fructose at a similar rate as wild-type GLUT2. Interestingly, inhibition of the

proteasome in oocytes led to increased expression of p.Ala363Ser transporter on the plasma membrane, and the variant transporter had a similar activity as wild-type GLUT2 (data not shown). Discrepancies between heterologously expressed transporter activity in oocytes versus HEK293 cells have been previously observed. For example, Karniski and colleagues found that two out of the six variants in the sulfate transporter (DTDST) that were previously considered "partial activity" transporters in oocytes [64] demonstrated no activity in Flp-In™ 293 cells (Karniski et al., 2004).[65]

Temperature-sensitive mutants may also function differently in oocytes (generally studied between 25-30°C and mammalian cells which are usually studied at 37°C.[64,66,67] We determined that p.Ala363Ser is probably not temperature-sensitive because p.Ala363Ser transport function was similar to wild-type GLUT2 at 37°C (Flp-In™ 293 cells) and 18°C (oocytes plus proteasome inhibitor, data not shown). Our result showed that proteasome inhibition in Flp-In™ 293 cells does not change the transport function of p.Ala363Ser, which suggests that the proteasome machinery of *Xenopus laevis* oocytes differs from HEK293 cells. Normally, expression patterns of heterologously expressed proteins are similar in *Xenopus laevis* oocytes and mammalian cells [68], but clear examples of differences in the expression of heterologous proteins between cell lines have been reported.[65,66,67,69,70, 71,72,73] Therefore, extrapolation of these data to mammalian cells should be done with caution.

Some FBS *SLC2A2* alleles are represented in publicly available databases

Among the 17 FBS mutations characterized in this study, six were reported in various sequencing databases (c.115_117delATA, p.Ile39del; c.380C>A, p.Ala127Asp; p.Leu153_Ile154del; c.952G>A, p.Gly318Arg; p.Pro417Leu and c.1250C>G, p.Pro417Arg) like

the gnomAD and Geno2MP browsers, (Geno2MP:

http://geno2mp.gs.washington.edu/Geno2MP/#/).[75] Neither database reported a homozygote

carrier of the six variants. Most of these variants have one to three allele counts, except for,

p.Ile39del and p.Leu153_Ile154del which were more common (i.e. 8 and 22 allele counts,

respectively, in gnomAD; 3 alleles of p.Leu153_Ile154del reported in Geno2MP).

Structural analysis proposes molecular mechanisms of GLUT2 transporter dysfunction

SLC2A2 variants associated with FBS or metabolic traits that were expressed on the plasma

membrane were selected for structural analysis. The structure of GLUT2 is currently unknown,

however, structures of other human *SLC2* members as well as the closely related prokaryotic

homolog XylE, have been determined in atomic resolution, providing excellent templates to

model GLUT2.[76] We therefore generated homology models of GLUT2 in three different

conformations to analyze the variants (Figure 3.5a). While the majority of these FBS variants

occurred in the TM helices away from the substrate-binding site (> 5Å), we speculate that these

variants affected function through altered packing of the TM helices or protein dynamics. The

predicted effects of seven variants on the molecular attributes of GLUT2 that are needed for

transport are described below.

- c.96T>G, p.Asn32Lys and c.976_977delinsAA, p.Ser326Lys both lie in the extracellular

  channel of the transporter in TM1 and TM7, respectively. Their charged, bulky guanidinium

  side chain can sterically hinder the passage of the substrate through the channel (Figure

  3.5b).

- c.609T>A, p.Ser203Arg is positioned on TM5; the variant introduces large steric clashes and

  electrostatic repulsion that alter the packing of TM5 and TM1 and influence TM1 dynamics.

Furthermore, the basic, guanidinium side chain may reach into the substrate-binding site, directly affecting substrate binding.

- p.Gly318Arg, which is located next to substrate-binding site (Figure 3.5c), significantly influences the helix bending of TM7 while its large steric bulk affects the helix packing. These structural effects may alter the equilibrium of the conformation ensemble of the transporter and favour the inward-open conformation, resulting in the substrate being excluded from the substrate-binding site.

- c.1246G>A, p.Gly416Ser is a variant along the transporter's channel and can interact with the substrate. It is a variant on TM10 and in the critical (GPGPXP) helix kink formation, which can affect the helix packing and TM10 bending. In addition, p.Gly416Ser can directly interact with the binding site residues Glu412, and Gln193 on TM5, thus directly affecting binding site shape and biophysical properties.

- p.Pro417Leu, a commonly reported FBS variant, was functionally characterized previously and shown to be expressed on the plasma membrane.[11] According to our analysis, p.Pro417Leu directly and indirectly impacts several substrate-binding residues (Figure 3.5d). Specifically, p.Pro417Leu loses the backbone amide hydrogen bond to the *i+4* residue [77] and distorts the natural (GPGPXP) helix kink formation, which can lead to the indole sidechain of p.Trp420 to twist away and prevent interaction with the substrate. This variant introduces large steric bulk affecting the tight helix packing, resulting in the distortion of binding site residues (p.Glu412, p.Asn349, and p.Gln193).

- Another variant at position 417, p.Pro417Arg, also affects the kink of helix. Both p.Pro417Leu and p.Pro417Arg have the steric bulk that can affect the helix packing in the region (Figure 3.5d). But p.Pro417Arg establishes an extensive hydrogen bond network with

the surrounding residues and helices that potentially limits the expulsion movement of the other helices. p.Pro417Leu does not have this hydrogen bond network, and its steric bulk will expel the other packing helices.
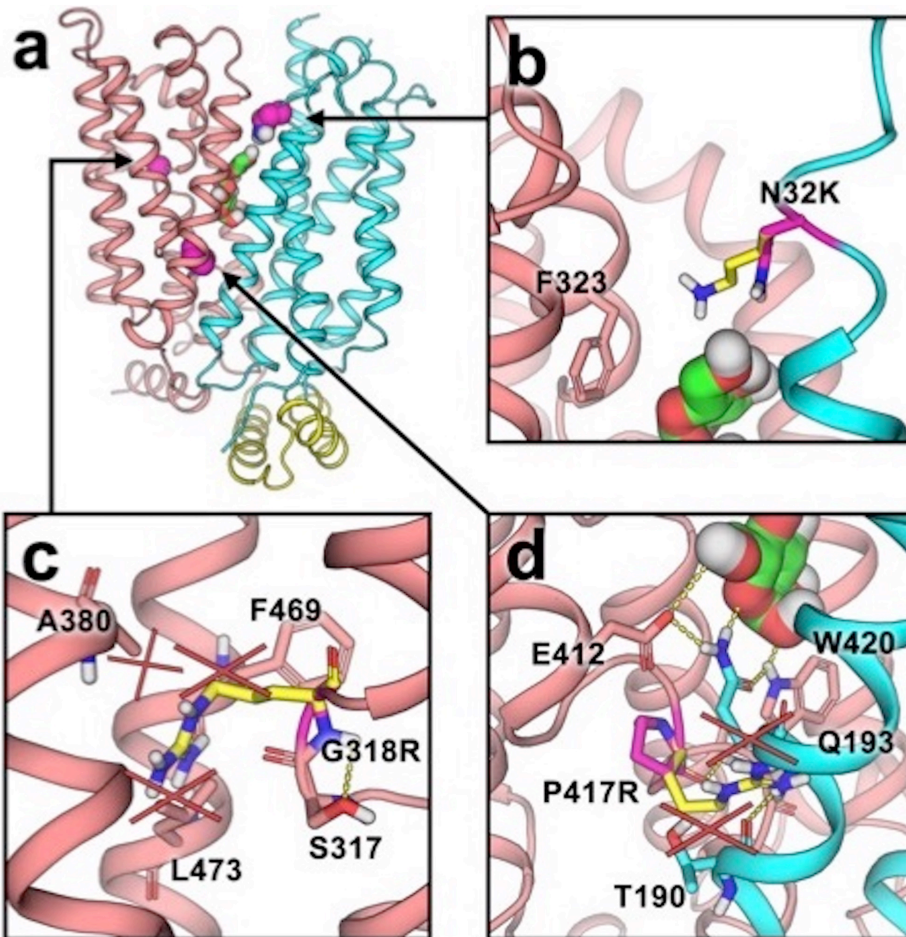
Figure 3.5 Structural analysis of GLUT2 variants associated with FBS.
GLUT2 homology model in an outward-open conformation (a). TM1-6, intracellular helix, and TM7-12 are represented as cyan, yellow, and salmon cartoon, respectively; the sugar maltose, derived from the template structure, is depicted as green sticks. Residues of interest and their mutations are depicted as magenta and yellow sticks, respectively.

Asn32 locates on TM2 and exposes to the extracellular channel (b). p.Asn32Lys (N32K) extends the reach of the residue sidechain into the channel by ~ 3.0 Å, potentially hindering the entry of substrate. Gly318 on TM7 is flanked by several bulky residues on TM9 and TM12, in addition to forms part of the SGXXG helix kink on TM7 (c). p.Gly318Arg (G318R) mutation introduces a bulky sidechain that clashes with bulky sidechains of TM9 and TM12 that disrupts the helix-helix packing interface, while affecting the formation of the native kink on TM7. Pro417 forms part of the TM10 helix kink and packs against TM5 and TM8, where its backbone carbonyl participates in the hydrogen-bond network of the substrate-binding site through interaction with the conserved Gln314 and Trp420. p.Pro417Arg (P417R) mutation, and similarly p.Pro417Leu, alters the formation of th TM10 helix kink (d). Furthermore, the introduction of a bulky sidechain clashes with Thr190 and Gln193, which disrupts the packing of TM5, TM8, TM10 as well as the integrity of the hydrogen-bond network of the substrate-binding site. Hydrogen bonds and steric clashes are depicted as yellow dotted lines and red crosses, respectively.

Commonly used pharmacochaperones and PTC-modulating agents did not improve membrane expression or transport activity of GLUT2 FBS variants

Several pharmacochaperone compounds were tested to determine if any improved the expression and/or activity of characterized FBS variants, specifically p.239delCys, p.Leu389Pro and p.Thr480Arg (Figure 3.6). None of the tested pharmacochaperones (4-PBA 2mM, Corr-4a 10μM, vorinostat 6mM, luminespib 100μM, URDA 500μM, 17-DMAG 500μM) provided a statistically significant increase in average transport activity in GLUT2 mutant oocytes compared to saline-injected oocytes (Figure 3.6a, student's t-test, p-value ≥ 0.1). Similarly, there was no statistical difference in average expression of mutant transporters on the plasma membrane (one-way ANOVA, p-value ≥ 0.21) compared to baseline (Figure 3.6b).

Additionally, two PTC-modulating agents (G418 100μM, Ataluren 50μM) were tested to evaluate their ability to increase the transport activity of four FBS nonsense variants (p.Arg301Ter, p.Arg365Ter, p.Trp395Ter, p.Trp420Ter; Figure 3.7). Exposure to G418 100μM or ataluren 50μM did not increase the average uptake of ribose in HEK cells transfected with any of the four GLUT2 nonsense plasmids (one-way ANOVA, p-value > 0.05; Figure 3.7).
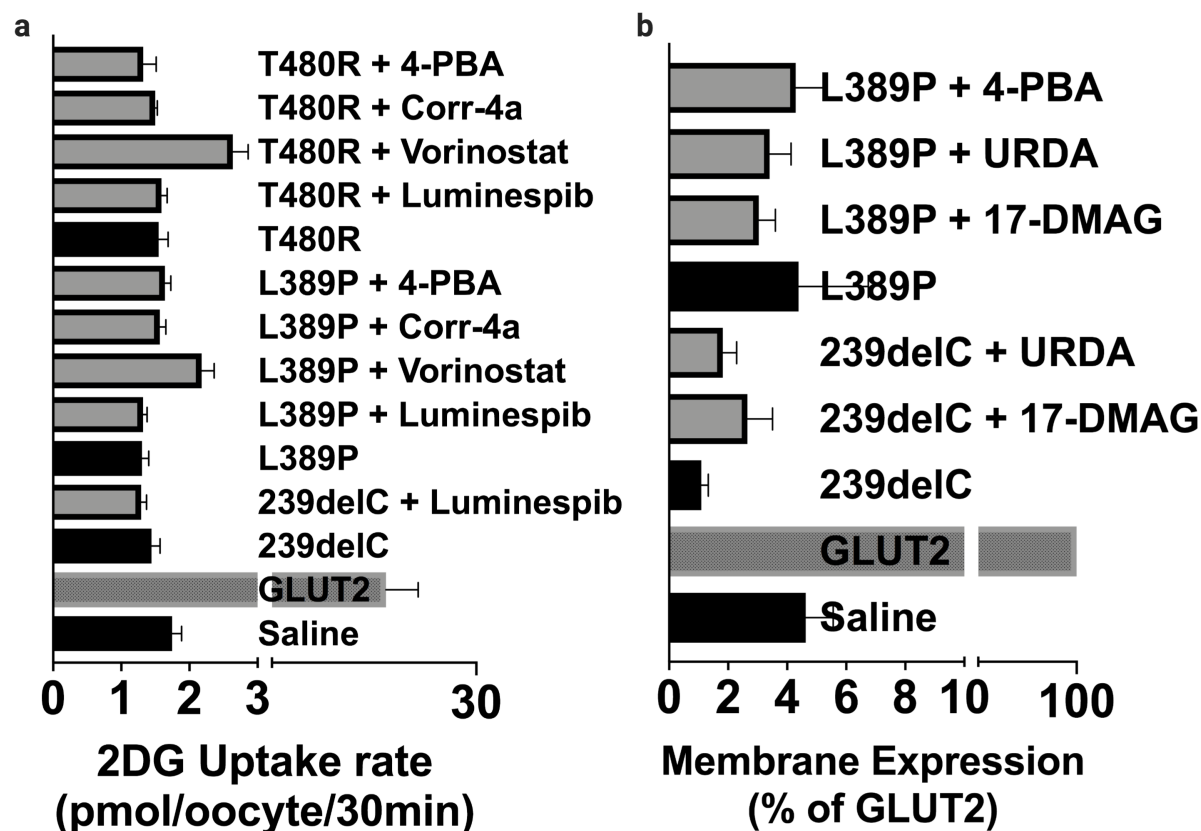
Figure 3.6 GLUT2 mutant oocytes exposed to pharmacochaperones. Radiolabeled uptake ([^{3}H]-2-deoxyglucose) (a) and plasma membrane expression (b) experiments were conducted with X.laevis oocytes injected with human GLUT2 wildtype or mutant cRNA (40ng) or saline. Oocytes were pre-exposed to MBS buffer containing specific concentration of each pharmacochaperone for 24 hours. Fourty-eight hours later experiments were conducted. Data shown represents the mean ± SEM for two independent experiments.
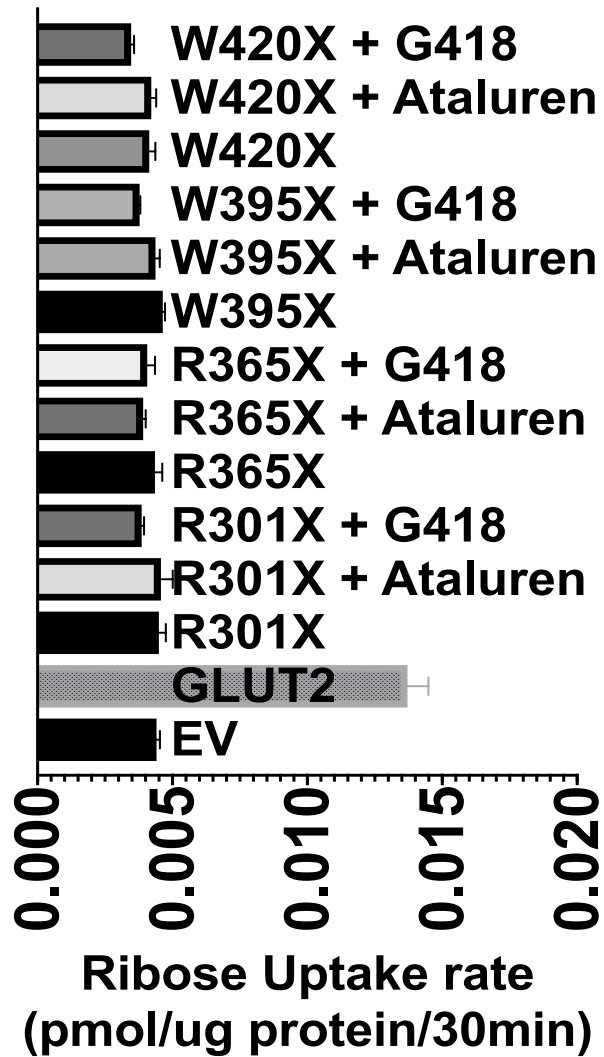
Figure 3.7 HEK cells transfected with GLUT2 nonsense plasmids exposed to PTC-modulating agents.

The graph shows the uptake of ribose in Flp-In™ 293 cells transfected with emptye vector, GLUT2 wildtype or four FBS nonsense variant plasmids with or without exposure to G418 (100μM) or Ataluren (50μM).

Three novel substrates of human GLUT2 were identified

A variety of radionuclide monosaccharides ([^{3}H]-ribose, [^{3}H]-xylose, [^{3}H]-1,5 AG, [^{3}H]-galactose, [^{3}H]-glucosamine) and disaccharides ([^{3}H]-sucrose) were evaluated to identify additional substrates of human GLUT2. As shown in Figure 3.8, xylose, 1,5 AG and ribose exhibited approximately a 3-fold, 5-fold and 30-fold higher uptake in oocytes expressing GLUT2 compared with saline-injected oocytes (students t-test, p-value < 0.005), respectively, suggesting that these compounds are substrates of the transporter.

The kinetic parameters (K_m, V_{max}) of each novel substrate for GLUT2 were evaluated (Figure 3.9). Notably, the V_{max} of ribose and 1,5 anhydroglucitol for GLUT2 was similar to or greater than that of glucose, (46 vs 110 vs 51 nmol/oocyte/30min, respectively), whereas the K_m of ribose and 1,5 anhydroglucitol was much larger than that of glucose (177 mM vs 148 mM vs 17 mM, respectively) suggesting that GLUT2 is a low affinity, high capacity ribose and 1,5 anhydroglucitol transporter. The kinetics of xylose were comparable to 2DG for GLUT2.
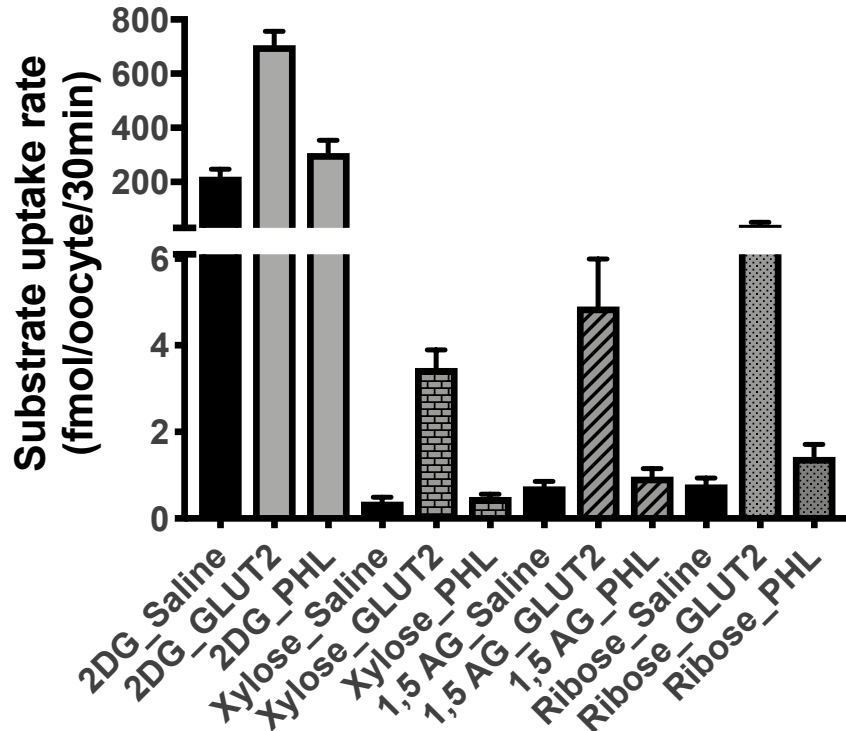
Figure 3.8 Discovery of novel substrates of GLUT2 transporter.
Radiolabeled substrate ([³H]-2-deoxyglucose, [³H]-xylose, [³H]-1,5 anhydroglucitol (1,5 AG), [³H]-ribose) experiments were conducted with X.laevis oocytes injected with human GLUT2 wild-type cRNA (40ng) or saline. Oocytes were exposed to buffer solution containing radiolabeled substrate with and without phloretin (PHL), a canonical GLUT transporter inhibitor. Duration of the uptake was 0.5 hour and rates were calculated into fentimoles of substrate. Data shown represents the mean ± SEM for two independent experiments.
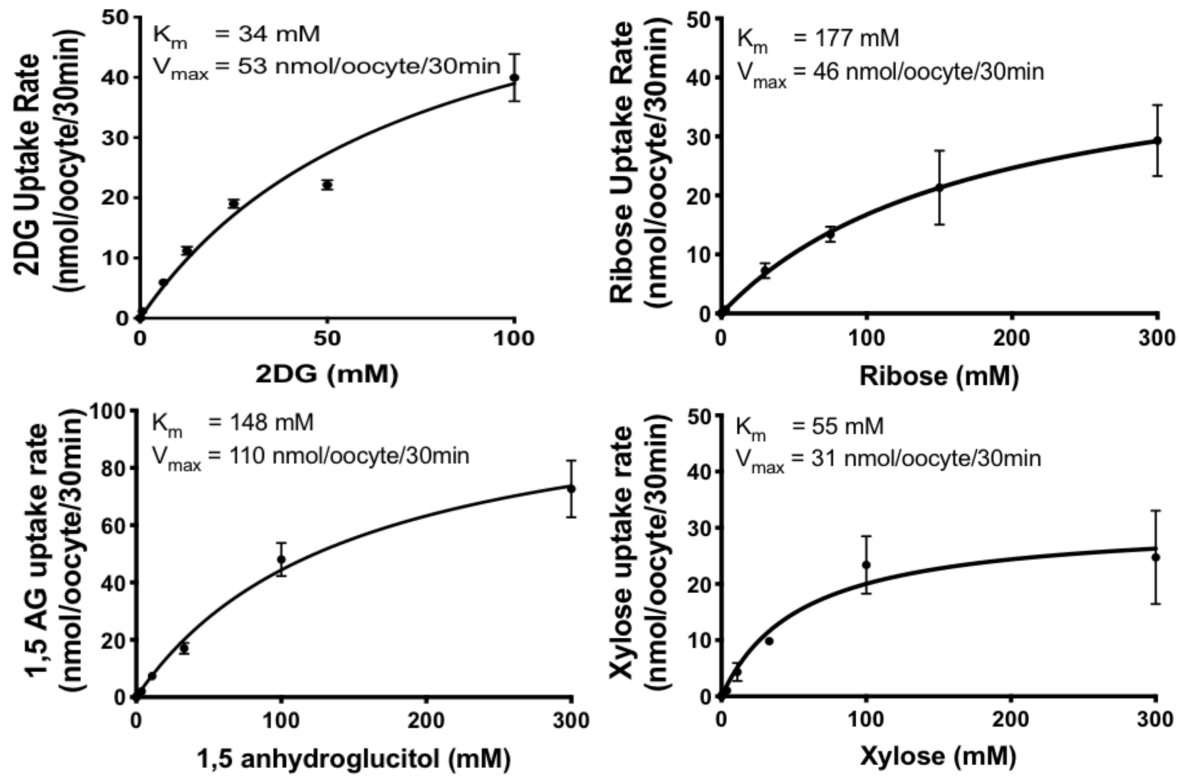
Figure 3.9 Kinetic evaluation of GLUT2 novel substrates.
Kinetic experiments were conducted with radiolabeled substrate ([³H]-2-deoxyglucose, [³H]-xylose, [³H]-1,5 anhydroglucitol (1,5 AG), [³H]-ribose) in combination with a series of unlabeled substrate concentrations (ranged from 0 - 300mM) using X.laevis oocytes injected with GLUT2 wild-type cRNA (40ng). Data was fitted to a Michaelis-Menten curve and kinetic parameters were estimated (K_m, V_{max}). Data shown represents the mean ± SEM for two independent experiments.

DISCUSSION

In this study, I performed a comprehensive functional characterization of seventeen *SLC2A2* missense and in-frame indel mutant transporters associated with a range of FBS phenotypes and five *SLC2A2* missense variants associated with various metabolic perturbations (e.g. T2D, elevated HbA1c concentration, increased fasting glucose levels). Previously, only five FBS mutant transporters had been characterized out of over 40 known variants and to our knowledge, only three *SLC2A2* missense variants associated with metabolic traits have been characterized (c.329C>T, p.Thr110Ile; c.203C>T, p.Pro68Leu and c.589G>A, p.Val197Ile.[7,11,13,16,17,55] Generally, in order for FBS to manifest in an individual, the person must be homozygous or compound heterozygous for deleterious *SLC2A2* variants. That is, FBS is, in most cases, an autosomal recessive disease. In this study, following characterization of variants associated with FBS, various pharmacochaperones and PTC-modulating agents were used in an attempt to recover some of the lost function of the mutant transporters.

The functional and structural evaluation of these variants led to four major findings. First, variants associated with FBS virtually abolished transport function; however, one variant, p.Leu153_Ile154del, found in two cases of mild FBS, exhibited residual transport function. Second, about half of the FBS mutant transporters were not expressed (or expressed at low levels) on the plasma membrane, whereas the other half were expressed on the membrane but exhibited no transport activity. Third, structural analysis provided a mechanistic rationale for the deleterious effects of GLUT2 genetic variants that were expressed on the plasma membrane. Fourth, two less common variants, p.Thr198Lys and p.Ala363Ser, which are associated with increased glycated hemoglobin levels and T2D, exhibited reduced activity in oocytes. Below we

70

discuss each of these findings in the context of the literature. This discussion is followed by discussions of our attempts to recover function of GLUT2 mutants with small molecule modulators of protein processing and our discovery of new substrates of GLUT2.

First, of the seventeen variants in GLUT2 that had not previously been characterized, all but one (p.Leu153_Ile154del) abolished transport function. To the best of our knowledge, p.Leu153_Ile154del is the first *SLC2A2* variant associated with FBS that results in a GLUT2 that retains some function, albeit minimal. The patients documented with a mild form of FBS were compound heterozygotes for the variants, p.Leu153_Ile154del and p.Pro417Arg. Our data show that p.Pro417Arg has no activity. Though speculative, the data suggest that the residual activity of p.Leu153_Ile154del transporter was sufficient to rescue the severe phenotype commonly associated with FBS, resulting in a milder phenotype. The data may suggest that the meager function of one GLUT2 allele can significantly mitigate the most severe signs and symptoms of FBS. This has been demonstrated in other rare genetic diseases in which variants that result in partially active enzymes, transporters or channels are associated with a mild to moderate form of the disease and variants that result in non-functional enzymes, transporters or channels lead to the most severe forms (e.g. DTDST variants and chrondrodysplasias).[65,69]

Second, the loss of function of about half of the mutant GLUT2 transporters was due to lack of expression on the plasma membrane. Various mechanisms for improper sorting of mutant transporters in GLUTs and other *SLC*s have been described. [78,79,80,70,81] Generally, these mechanisms involve the retention (in ER or Golgi apparatus) and eventual destruction of

aggregated or misfolded mutant protein or the misdirection of mutant protein to the incorrect membrane.[79,80,70,81,82]

GLUT1, GLUT3 and GLUT4 have known sorting sequences, DSQV [78], DRSGKDGVMEMN [83], and FQQI, TELEY, LL [84], respectively, that localize these GLUTs to their extracellular or intracellular membrane compartments. GLUT1 variants that occur near phosphorylation motifs that enhance membrane localization have been shown to reduce or prevent membrane expression of transporters, leading to GLUT1 deficiency syndromes.[85] MetaServer, a tool that predicts peptide binding sequences within proteins, predicted several sequences in GLUT2 that could bind to proteins that regulate membrane trafficking.[86,87,88,89] Of the 18 sequence-motifs in GLUT2 predicted to harbor binding sites for proteins that regulate trafficking, one is changed by an FBS variant characterized in our study (c.715_717delTGT, p.239delCys). Consistently, p.239delCys was not expressed on the plasma membrane. However, because these predicted motifs have not been validated *in vitro*, further work is needed to characterize the regulatory proteins and mechanisms by which these mutants alter plasma membrane expression of GLUT2.

Third, we observed that many variants associated with FBS (nine out of the seventeen FBS mutant transporters) result in expression levels on the plasma membrane that are comparable to levels of the wild-type GLUT2. Computational analysis of seven out of the nine expressed FBS variants revealed several potential mechanisms by which FBS variants may disrupt transporter activity: steric hindrance of substrate channel (e.g. p.Asn32Lys and p.Ser326Lys), distortion of substrate-binding residues (e.g. p.Ser203Arg, p.Gly416Ser), and instability of TM helices and the overall transporter structure (e.g. p.Gly20Glu). These findings expand on previous findings

of structural analysis of pathogenic *SLC2A1* variants that result in GLUT1 deficiency, which

suggest that the vast majority of pathological GLUT1 disrupt helix-helix interactions and lead to

structural instability of the transporter.[90]

For instance, variants located at or near the substrate channel pathway in GLUT1 have been

found to destabilize TM helix-helix interactions and helix packing (e.g.

NM_006516.2:c.197C>T, p.Ser66Phe; NM_006516.2:c.929C>T, p.Thr310Ile)[90] whereas our

analysis showed rare GLUT2 variants that reside on either side of the substrate entrance (e.g.

p.Asn32Lys and p.Ser326Lys; Figure 3.1, Figure 3.5) block the translocation pathway via their

bulky, polar side chains. Similarly, p.Gly318Arg, which is located next to GLUT2 substrate-

binding residues, affected TM7 helix bending and packing via it large steric bulk resulting in

conformational inflexibility. Only one pathogenic GLUT1 variant directly impacted sugar-

interacting residues (1 of 7)[90], whereas three of the nine FBS variants evaluated affected

substrate-binding residues in GLUT2. For GLUT1, a series of engineered variant in the Ile168

residue, located next to the hinge residue Gly167, which is adjacent to the sugar-binding site,

abolished GLUT1 function via increased rigidity and steric hindrance.[91] Future studies are

needed to confirm the mechanisms predicted by the structural model for each of the GLUT2

variants. Overall, these findings suggest that the majority of pathogenic *SLC2* variants disrupt

proper packing and dynamics of the TM helices or affect the shape and biophysical properties of

the substrate binding site, which in turn leads to ineffective transport.

Our final finding, which was not unexpected, was that most of the less common variants that had

not been associated with FBS showed similar function to the reference GLUT2 (Figure 3.3).

These variants, most of which retained normal GLUT2 transport activity, have been associated, but at weak p-values, with various metabolic traits (e.g. T2D, high body mass index). The minor allele of p.Val101Ile and p.Gly519Glu are on the same haplotype with the reference allele (T) of rs8192675, a known eQTL (expression quantitative trait locus) that associates with increased risk of T2D [92]. The minor allele of p.Leu468Val, which is associated with increased risk of T2D (OR=1.75, p=0.017) and fasting glucose (beta=0.58, p=0.003) is on the same haplotype with the reference G-allele of p.Thr110Ile, which itself has been associated with increased risk of T2D and fasting glucose. Structural analysis revealed that p.Thr198Lys (MAF < 0.5%), which displayed reduced substrate uptake in oocytes and Flp-In™ 293 cells (Figure 3.4, Table 3.1), may restrict helix movement and cause disturbance to the protein-lipid interface because of the charged lysine. This, along with a large reduction in glucose transport capacity (V_{max}), may explain the statistical associations of p.Thr198Lys with increased hemoglobin A1c levels (p=0.012).

Despite exposing three FBS missense variants and four FBS nonsense variants to a structural-varied pharmacochaperones and PTC-modulating agents, no increases in expression and/or activity were observed (Figure 3.6, Figure 3.7). There is in-vitro evidence of select pharmacochaperones (phenylbutyrate, ibogaine, noribogaine) increasing the expression and/or activity of missense mutant SLC transporters (OCTN2/*SLC22A5,* DAT/*SLC6A3,* SERT/*SLC6A4*).[93,94] But this area/topic is significantly understudied compared with the large amount of *in vitro* and *in vivo* efficacy data of pharmacochaperones in misfolded enzyme diseases. A lack of high-resolution structures for the vast majority of SLC transporters[95] (including GLUT2) make it difficult to simulate accurate models of transporter dynamics, which

ultimately contribute to a poor mechanistic understanding of transporter function and a lack of potential therapeutic interventions.[94]

Similarly, several PTC-modulating agents have demonstrated *in vitro* and *in vivo* efficacy in FAP and DMD, but *in vitro* increases in transporter expression by these agents have only been reported a few times (i.e. *ABCC6, CTNS/SLC66A4*).[96,97] High-throughput virtual screens using high resolution structures (crystal or cryo-elecronmicroscopy structures) of wildtype GLUT2 and partial/residual function GLUT2 mutants (i.e. p.Leu153_Ile154del) are needed to fast track discovery of pharmacochaperones and allosteric activators of GLUT2. Furthermore, establishing high-throughput *in vitro* screens for GLUT2 nonsense variants could expedite identification of novel compounds that promote PTC readthrough of FBS nonsense mutations.

Finally, in this study, we discovered three previously unreported substrates of human GLUT2: 1,5 AG, ribose and xylose. Previously, the mechanism of 1,5 anhydroglucitol (found in nearly all foods) absorption in the intestine and secretion in the kidneys was unknown, [34,98,36] but assumed to be carrier-mediated. 1,5 AG has an inverse relationship with glucose in the blood; the higher the 1,5 AG level, the lower the glucose level. Glycomark, a 1,5 AG clinical lab test, is approved by the Japanese Pharmaceuticals and Medical Devices Agency to measure and monitor short term glycemic control in diabetics.[98]

Ribose is sold as a dietary supplement to improve and accelerate muscle recovery after intense exercise. Based on the manufacturers' daily dose recommendations (15 grams/day)[99,100], the estimated intestinal concentration of ribose is 133 mM, close to the K_m for GLUT2.

Administration of ribose (intravenous or oral) has been associated with transient hypoglycemia.[101] Additional studies are needed to determine how GLUT2 influences ribose distribution and synthesis via the pentose phosphate pathway.

The clinical use of xylose is rare when compared to 1,5 AG or ribose. For example, intestinal absorption of xylose is used (rarely) as a clinical diagnostic test to determine if a patient's intestinal tract is absorbing nutrients and functioning properly.[102] Consumption of xylose in combination with sucrose reduced serum glucose and insulin levels compared with those who consumed sucrose-only beverages.[103] The K_m and V_{max} of GLUT2 for xylose, which is similar to that of glucose (Figure 3.9), suggests that competitive inhibition between xylose and glucose for GLUT2 may occur.

CONCLUSIONS

This comprehensive characterization of variants associated with FBS reveals that about half of the mutant transporters are expressed on the plasma membrane but do not function, and the other half are not significantly expressed on the plasma membrane. Further, we report for the first time that the in-frame deletion variant, p.Leu153_Ile154del, has residual glucose uptake. In contrast, the majority of the five *SLC2A2* missense variants associated with metabolic traits have normal glucose uptake and are expressed on the plasma membrane. The structural analysis corroborates the functional studies and provides mechanistic insights into how these low frequency variants inactivate or reduce GLUT2 transporter activity. The functional and structural information gained in this study suggests that different strategies should be used to produce targeted, safe and effective therapies for FBS and possibly other GLUT-related diseases.

REFERENCES

1.   Santer R, Schneppenheim R, Dombrowski A, Gotze H, Steinmann B, Schaub J. Mutations in GLUT2, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome. *Nat Genet*. 1997;17(3):324-326. doi:10.1038/ng1197-324

2.   Santer R, Groth S, Kinner M, et al. The mutation spectrum of the facilitative glucose transporter gene SLC2A2 (GLUT2) in patients with Fanconi-Bickel syndrome. *Hum Genet*. 2002;110(1):21-29. doi:10.1007/s00439-001-0638-6

3.   Mueckler M, Thorens B. The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med*. 2013;34(2-3):121-138. doi:10.1016/j.mam.2012.07.001

4.   Santer R, Schneppenheim R, Suter D, Schaub J, Steinmann B. Fanconi-Bickel syndrome - The original patient and his natural history, historical steps leading to the primary defect, and a review of the literature. *Eur J Pediatr*. 1998;157(10):783-797. doi:10.1007/s004310050937

5.   Manz F, Bickel H, Brodehl J, et al. Pediatric Nephrology. *Pediatr Nephrol*. 1987;1(October 1960):509-518.

6.   Berry GT, Baker L, Kaplan FS, Witzleben CL. Diabetes-like renal glomerular disease in Fanconi-Bickel syndrome. *Pediatr Nephrol*. 1995;9(3):287-291. doi:10.1007/BF02254185

7.   Pogoriler J, Neill AFO, Voss SD, Shamberger RC, Perez-atayde AR. Hepatocellular Carcinoma in Fanconi-Bickel Syndrome. 2017. doi:10.1177/1093526617693540

8.   Yoo HW, Shin YL, Seo EJ, Kim GH. Identification of a novel mutation in the GLUT2 gene in a patient with Fanconi-Bickel syndrome presenting with neonatal diabetes mellitus and galactosaemia. *Eur J Pediatr*. 2002;161(6):351-353. doi:10.1007/s00431-002-0931-y

9.   Karamizadeh Z, Saki F, Imanieh MH, Zahmatkeshan M, Fardaee M. A new mutation of

Fanconi-Bickel syndrome with liver failure and pseudotumour cerebri. *J Genet*. 2012;91(3):359-361. doi:10.1007/s12041-012-0198-7

10. Grünert SC, Schwab KO, Pohl M, Sass JO, Santer R. Fanconi-Bickel syndrome: GLUT2 mutations associated with a mild phenotype. *Mol Genet Metab*. 2012;105(3):433-437. doi:10.1016/j.ymgme.2011.11.200

11. Michau A, Guillemain G, Grosfeld A, et al. Mutations in SLC2A2 Gene Reveal hGLUT2 Function in Pancreatic β Cell Development. *J Biol Chem*. 2013;288(43):31080-31092. doi:10.1074/jbc.M113.469189

12. Mannstadt M, Magen D, Segawa H, et al. Fanconi-Bickel syndrome and autosomal recessive proximal tubulopathy with hypercalciuria (ARPTH) are allelic variants caused by GLUT2 mutations. *J Clin Endocrinol Metab*. 2012;97(10):1978-1986. doi:10.1210/jc.2012-1279

13. Amita M, Srivastava P, Mandal K. Fanconi-Bickel Syndrome : Another Novel Mutation in SLC2A2. 2017;84(March):236-237. doi:10.1007/s12098-016-2236-6

14. Bahíllo-Curieses MP, Garrote-Molpeceres R, Miñambres-Rodríguez M, del Real-Llorente MR, Tobar-Mideros C, Rellán-Rodríguez S. Glycosuria and hyperglycemia in the neonatal period as the first clinical sign of Fanconi-Bickel syndrome. *Pediatr Diabetes*. May 2017. doi:10.1111/pedi.12531

15. Odièvre MH, Lombès A, Dessemme P, et al. A secondary respiratory chain defect in a patient with Fanconi-Bickel syndrome. *J Inherit Metab Dis*. 2002;25(5):379-384. doi:10.1023/A:1020147716990

16. Wang W, Wei M, Song HM, et al. SLC2A2 gene analysis in three Chinese children with Fanconi-Bickel syndrome. *Chinese J Contemp Pediatr*. 2015;17(4):362-366.

doi:10.7499/j.issn.1008-8830.2015.04.014

17.   Abbasi F, Azizi F, Javaheri M, Mosallanejad A, Ebrahim-Habibi A, Ghafouri-Fard S. Segregation of a novel homozygous 6 nucleotide deletion in GLUT2 gene in a Fanconi-Bickel syndrome family. *Gene*. 2015;557(1):103-105. doi:10.1016/j.gene.2014.12.024

18.   Hadipour F, Sarkheil P, Noruzinia M, Hadipour Z, Baghdadi T, Shafeghati Y. Fanconi-Bickel syndrome versus osteogenesis imperfeeta: An Iranian case with a novel mutation in glucose transporter 2 gene, and review of literature. *Indian J Hum Genet*. 2013;19(1):84-86. doi:10.4103/0971-6866.112906

19.   Peduto A, Spada M, Alluto A, La Dolcetta M, Ponzone A, Santer R. A novel mutation in the GLUT2 gene in a patient with Fanconi-Bickel syndrome detected by neonatal screening for galactosaemia. *J Inherit Metab Dis*. 2004;27(2):279-280.

20.   Su Z, Du ML, Chen HS, Chen QL, Yu CS, Ma HM. Two cases of Fanconi-Bickel syndrome: First report from China with novel mutations of SLC2A2 gene. *J Pediatr Endocrinol Metab*. 2011;24(9-10):749-753. doi:10.1515/JPEM.2011.316

21.   Sakamoto O, Ogawa E, Ohura T, et al. Mutation analysis of the GLUT2 gene in patients with Fanconi-Bickel syndrome. *Pediatr Res*. 2000;48(5):586-589. doi:10.1203/00006450-200011000-00005

22.   Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes*. 2017;66(11):2888-2902. doi:10.2337/db16-1253

23.   Manning AK, Hivert M-F, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 2012;44(6):659-669. doi:10.1038/ng.2274

24.   Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting

glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2):105-116. doi:10.1038/ng.520

25.   Barker A, Sharp SJ, Timpson NJ, et al. Association of genetic Loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. *Diabetes*. 2011;60(6):1805-1812. doi:10.2337/db10-1575

26.   Wheeler E, Leong A, Liu C-T, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med*. 2017;14(9):e1002383. doi:10.1371/journal.pmed.1002383

27.   Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet*. 2016;48(9):1055-1059. doi:10.1038/ng.3632

28.   Type 2 Diabetes Knowledge Portal. http://www.type2diabetesgenetics.org/variantInfo/variantInfo/rs140138702#. Accessed March 15, 2017.

29.   McCafferty EH, Scott LJ. Migalastat: A Review in Fabry Disease. *Drugs*. 2019;79(5):543-554. doi:10.1007/s40265-019-01090-4

30.   Southern KW, Patel S, Sinha IP, Nevitt SJ. Correctors (specific therapies for class II CFTR mutations) for cystic fibrosis. *Cochrane database Syst Rev*. 2018;8:CD010966. doi:10.1002/14651858.CD010966.pub2

31.   Leturque A, Brot-Laroche E, Le Gall M. GLUT2 mutations, translocation, and receptor function in diet sugar managing. *Am J Physiol Endocrinol Metab*. 2009;296(5):E985-92. doi:10.1152/ajpendo.00004.2009

32.   Uldry M, Ibberson M, Hosokawa M, Thorens B. GLUT2 is a high affinity glucosamine transporter. *FEBS Lett*. 2002;524(1-3):199-203. doi:10.1016/s0014-5793(02)03058-2

33.   Mardones L, Ormazabal V, Romo X, et al. The glucose transporter-2 (GLUT2) is a low affinity dehydroascorbic acid transporter. *Biochem Biophys Res Commun*. 2011;410(1):7-12. doi:10.1016/j.bbrc.2011.05.070

34.   Yamanouchi T, Tachibana Y, Akanuma H, et al. Origin and disposal of 1,5-anhydroglucitol, a major polyol in the human body. *Am J Physiol*. 1992;263(2 Pt 1):E268-73. doi:10.1152/ajpendo.1992.263.2.E268

35.   Yamanouchi T, Ogata N, Yoshimura T, et al. Transport of 1,5-anhydro-D-glucitol into insulinoma cells by a glucose-sensitive transport system. *Biochim Biophys Acta*. 2000;1474(3):291-298. doi:10.1016/s0304-4165(00)00025-8

36.   Ying L, Ma X, Yin J, et al. The metabolism and transport of 1,5-anhydroglucitol in cells. *Acta Diabetol*. 2018;55(3):279-286. doi:10.1007/s00592-017-1093-8

37.   Clark PM, Flores G, Evdokimov NM, et al. Positron emission tomography probe demonstrates a striking concentration of ribose salvage in the liver. *Proc Natl Acad Sci U S A*. 2014;111(28):E2866-74. doi:10.1073/pnas.1410326111

38.   Consortium T 1000 GP. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

39.   Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815. doi:10.1006/jmbi.1993.1626

40.   Deng D, Xu C, Sun P, et al. Crystal structure of the human glucose transporter GLUT1. *Nature*. 2014;510(7503):121-125. doi:10.1038/nature13306

41.   Sun L, Zeng X, Yan C, et al. Crystal structure of a bacterial homologue of glucose

transporters GLUT1-4. *Nature*. 2012;490(7420):361-366. doi:10.1038/nature11524

42.   Deng D, Sun P, Yan C, et al. Molecular basis of ligand recognition and transport by glucose transporters. *Nature*. 2015;526(7573):391-396. doi:10.1038/nature14655

43.   Pei J, Grishin N V. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol Biol*. 2014;1079:263-271. doi:10.1007/978-1-62703-646-7_17

44.   Shen M-Y, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*. 2006;15(11):2507-2524. doi:10.1110/ps.062416606

45.   Wu EL, Cheng X, Jo S, et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J Comput Chem*. 2014;35(27):1997-2004. doi:10.1002/jcc.23702

46.   Lee J, Cheng X, Swails JM, et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput*. 2016;12(1):405-413. doi:10.1021/acs.jctc.5b00935

47.   Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*. 2017;14(1):71-73. doi:10.1038/nmeth.4067

48.   R. BB, L. BC, D. MA, et al. CHARMM: The biomolecular simulation program. *J Comput Chem*. 2009;30(10):1545-1614. doi:10.1002/jcc.21287

49.   Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935. doi:10.1063/1.445869

50. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput*. 2013;9(9):3878-3888. doi:10.1021/ct400314y

51. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. 2018;50(11):1505-1513. doi:10.1038/s41588-018-0241-6

52. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977;23(3):327-341. doi:https://doi.org/10.1016/0021-9991(77)90098-5

53. Roe DR, Cheatham TE 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput*. 2013;9(7):3084-3095. doi:10.1021/ct400341p

54. Fanconi G, Bickel H. Chronic aminoaciduria (amino acid diabetes or nephrotic-glucosuric dwarfism) in glycogen storage and cystine disease. *Helv Paediatr Acta*. 1949;4(5):359-396.

55. Gupta N, Nambam B, Weinstein DA, Shoemaker LR. Late Diagnosis of Fanconi-Bickel Syndrome. *J Inborn Errors Metab Screen*. 2016;4:232640981667943. doi:10.1177/2326409816679430

56. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res*. 2010;20(1):110-121. doi:10.1101/gr.097857.109

57. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*.

2009;25(9):1189-1191. http://dx.doi.org/10.1093/bioinformatics/btp033.

58.   Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249. doi:10.1038/nmeth0410-248

59.   Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812-3814. http://dx.doi.org/10.1093/nar/gkg509.

60.   Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310-315. doi:10.1038/ng.2892

61.   Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res*. 2005;15(7):901-913. doi:10.1101/gr.3577405

62.   Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*. 2016;99(4):877-885. doi:10.1016/j.ajhg.2016.08.016

63.   Seatter MJ, De la Rue SA, Porter LM, Gould GW. QLS motif in transmembrane helix VII of the glucose transporter family interacts with the C-1 position of D-glucose and is involved in substrate selection at the exofacial binding site. *Biochemistry*. 1998;37(5):1322-1326. doi:10.1021/bi972322u

64.   Baron D, Assaraf YG, Cohen N, Aronheim A. Lack of plasma membrane targeting of a G172D mutant thiamine transporter derived from Rogers syndrome family. *Mol Med*. 2002;8(8):462-474. doi:S1528365802804621 [pii]

65.   Karniski LP. Functional expression and cellular distribution of diastrophic dysplasia sulfate transporter (DTDST) gene mutations in HEK cells. *Hum Mol Genet*.

2004;13(19):2165-2171. doi:10.1093/hmg/ddh242

66.   Woodward OM, Tukaye DN, Cui J, et al. Gout-causing Q141K mutation in ABCG2 leads to instability of the nucleotide-binding domain and can be corrected with small molecules. *Proc Natl Acad Sci*. 2013;110(13):5223-5228. doi:10.1073/pnas.1214530110

67.   Machamer CE, Rose JK. Vesicular stomatitis virus G proteins with altered glycosylation sites display temperature-sensitive intracellular transport and are subject to aberrant intermolecular disulfide bonding. *J Biol Chem*. 1988;263(12):5955-5960.

68.   Sigel E. Use of Xenopus oocytes for the functional expression of plasma membrane proteins. *J Membr Biol*. 1990;117(3):201-221. doi:10.1007/BF01868451

69.   Karniski LP. Mutations in the diastrophic dysplasia sulfate transporter (DTDST) gene: correlation between sulfate transport activity and chondrodysplasia phenotype. *Hum Mol Genet*. 2001;10(14):1485-1490.

70.   Skach WR. Defects in processing and trafficking of the cystic fibrosis transmembrane conductance regulator. *Kidney Int*. 2000;57(3):825-831. doi:10.1046/j.1523-1755.2000.00921.x

71.   Denning GM, Anderson MP, Amara JF, Marshall J, Smith AE, Welsh MJ. Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature*. 1992;358(6389):761-764. doi:10.1038/358761a0

72.   Baroudi G, Pouliot V, Denjoy I, Guicheney P, Shrier A, Chahine M. Novel mechanism for Brugada syndrome: defective surface localization of an SCN5A mutant (R1432G). *Circ Res*. 2001;88(12):E78-83.

73.   Baroudi G, Carbonneau E, Pouliot V, Chahine M. SCN5A mutation (T1620M) causing Brugada syndrome exhibits different phenotypes when expressed in Xenopus oocytes and

mammalian cells. *FEBS Lett*. 2000;467(1):12-16.

74.   Geno2MP, NHGRI/NHLBI University of Washington-Center for Mendelian Genomics (UW-CMG), Seattle, WA. http://geno2mp.gs.washington.edu.

75.   Lek M, Karczewski KJ, Minikel E V, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291. doi:10.1038/nature19057

76.   Colas C, Ung PM-U, Schlessinger A. SLC Transporters: Structure, Function, and Drug Discovery. *Medchemcomm*. 2016;7(6):1069-1081. doi:10.1039/C6MD00005C

77.   von Heijne G. Proline kinks in transmembrane alpha-helices. *J Mol Biol*. 1991;218(3):499-503.

78.   Wieman HL, Horn SR, Jacobs SR, Altman BJ, Kornbluth S, Rathmell JC. An essential role for the Glut1 PDZ-binding motif in growth factor regulation of Glut1 degradation and trafficking. *Biochem J*. 2009;418(2):345-367. doi:10.1042/BJ20081422

79.   El-Kasaby A, Koban F, Sitte HH, Freissmuth M, Sucic S. A cytosolic relay of heat shock proteins HSP70-1A and HSP90beta monitors the folding trajectory of the serotonin transporter. *J Biol Chem*. 2014;289(42):28987-29000. doi:10.1074/jbc.M114.595090

80.   Koban F, El-Kasaby A, Hausler C, et al. A salt bridge linking the first intracellular loop with the C terminus facilitates the folding of the serotonin transporter. *J Biol Chem*. 2015;290(21):13263-13278. doi:10.1074/jbc.M115.641357

81.   Toye AM. Defective kidney anion-exchanger 1 (AE1, Band 3) trafficking in dominant distal renal tubular acidosis (dRTA). *Biochem Soc Symp*. 2005;(72):47-63.

82.   Asjad HMM, Kasture A, El-Kasaby A, et al. Pharmacochaperoning in a Drosophila model system rescues human dopamine transporter variants associated with infantile/juvenile parkinsonism. *J Biol Chem*. 2017;292(47):19250-19265. doi:10.1074/jbc.M117.797092

83. Inukai K, Shewan AM, Pascoe WS, Katayama S, James DE, Oka Y. Carboxy terminus of glucose transporter 3 contains an apical membrane targeting domain. *Mol Endocrinol*. 2004;18(2):339-349. doi:10.1210/me.2003-0089

84. Blot V, McGraw TE. Molecular mechanisms controlling GLUT4 intracellular retention. *Mol Biol Cell*. 2008;19(8):3477-3487. doi:10.1091/mbc.e08-03-0236

85. Lee EE, Ma J, Sacharidou A, et al. A Protein Kinase C Phosphorylation Motif in GLUT1 Affects Glucose Transport and is Mutated in GLUT1 Deficiency Syndrome. *Mol Cell*. 2015;58(5):845-853. doi:10.1016/j.molcel.2015.04.015

86. Kundu K, Backofen R. Cluster based prediction of PDZ-peptide interactions. *BMC Genomics*. 2014;15 Suppl 1:S5. doi:10.1186/1471-2164-15-S1-S5

87. Kundu K, Costa F, Huber M, Reth M, Backofen R. Semi-supervised prediction of SH2-peptide interactions from imbalanced high-throughput data. *PLoS One*. 2013;8(5):e62732. doi:10.1371/journal.pone.0062732

88. Kundu K, Mann M, Costa F, Backofen R. MoDPepInt: an interactive web server for prediction of modular domain-peptide interactions. *Bioinformatics*. 2014;30(18):2668-2669. doi:10.1093/bioinformatics/btu350

89. Kundu K, Costa F, Backofen R. A graph kernel approach for alignment-free domain-peptide interaction prediction with an application to human SH3 domains. *Bioinformatics*. 2013;29(13):i335-43. doi:10.1093/bioinformatics/btt220

90. Raja M, Kinne RKH. Pathogenic mutations causing glucose transport defects in GLUT1 transporter: The role of intermolecular forces in protein structure-function. *Biophys Chem*. 2015;200-201:9-17. doi:10.1016/j.bpc.2015.03.005

91. Ung PM-U, Song W, Cheng L, et al. Inhibitor Discovery for the Human GLUT1 from

Homology Modeling and Virtual Screening. *ACS Chem Biol*. 2016;11(7):1908-1916. doi:10.1021/acschembio.6b00304

92.   Machiela MJ, Chanock SJ. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. *Bioinformatics*. 2018;34(5):887-889. doi:10.1093/bioinformatics/btx561

93.   Amat di San Filippo C, Pasquali M, Longo N. Pharmacological rescue of carnitine transport in primary carnitine deficiency. *Hum Mutat*. 2006;27(6):513-523. doi:10.1002/humu.20314

94.   Freissmuth M, Stockner T, Sucic S. SLC6 Transporter Folding Diseases and Pharmacochaperoning. *Handb Exp Pharmacol*. 2018;245:249-270. doi:10.1007/164_2017_71

95.   Finkelstein JM. Milestone 24: Structures of membrane proteins. *Nat Publ Gr*. 2014;11(1):1. doi:10.1038/nature13372

96.   Zhou Y, Jiang Q, Takahagi S, Shao C, Uitto J. Premature Termination Codon Read-Through in the ABCC6 Gene : Potential Treatment for Pseudoxanthoma Elasticum. *J Invest Dermatol*. 2013;133(12):2672-2677. doi:10.1038/jid.2013.234

97.   Helip-Wooley A, Park MA, Lemons RM, Thoene JG. Expression of CTNS alleles: subcellular localization and aminoglycoside correction in vitro. *Mol Genet Metab*. 2002;75(2):128-133. doi:10.1006/mgme.2001.3272

98.   Dungan KM. 1,5-anhydroglucitol (GlycoMark) as a marker of short-term glycemic control and glycemic excursions. *Expert Rev Mol Diagn*. 2008;8(1):9-19. doi:10.1586/14737159.8.1.9

99.   Seifert JG, Brumet A, St Cyr JA. The influence of D-ribose ingestion and fitness level on

performance and recovery. *J Int Soc Sports Nutr*. 2017;14:47. doi:10.1186/s12970-017-0205-8

100.   MacCarter D, Vijay N, Washam M, Shecterle L, Sierminski H, St Cyr JA. D-ribose aids advanced ischemic heart failure patients. *Int J Cardiol*. 2009;137(1):79-80. doi:10.1016/j.ijcard.2008.05.025

101.   Ginsburg J, Boucher B, Beaconsfield P. Hormonal changes during ribose-induced hypoglycemia. *Diabetes*. 1970;19(1):23-27. doi:10.2337/diab.19.1.23

102.   Craig RM, Atkinson AJJ. D-xylose testing: a review. *Gastroenterology*. 1988;95(1):223-231. doi:10.1016/0016-5085(88)90318-6

103.   Jun YJ, Lee J, Hwang S, et al. Beneficial effect of xylose consumption on postprandial hyperglycemia in Korean: a randomized double-blind, crossover design. *Trials*. 2016;17(1):139. doi:10.1186/s13063-016-1261-0

Chapter 4: Pharmacologic inhibition of THTR-1 in hematological tissues as a potential mechanism of drug-induced megaloblastic anemia

ABSTRACT

Megaloblastic anemia (MA) is characterized by the presence of enlarged, immature red blood cells in the bone marrow and develops as a result of ineffective DNA synthesis in rapidly dividing hematopoietic stem cells. Exposure to specific drugs may cause drug-induced MA. An infrequent cause of MA is thiamine responsive megaloblastic anemia (TRMA), a genetic disorder caused by deleterious mutations in *SLC19A2* (encodes for thiamine transporter, THTR-1). THTR-1 inactivity leads to thiamine deficiency in hematopoietic stem cells resulting in MA. The major objective of our study was to determine whether drugs known to cause MA inhibit THTR-1 as a potential mechanism contributing to the anemia. In addition, we explored chemical properties of the THTR-1 inhibitors and used electronic health records (EHR) to validate our results. Drugs associated with MA were screened for THTR-1 inhibitor activity; IC_{50} curves were generated for drugs identified as inhibitors. Electronic health record information from patients on clinically relevant THTR-1 inhibitors was extracted and analyzed using R software.

Of the 63 drugs screened, ten inhibited THTR-1 transport $\geq$ 50% at 1 mM. IC_{50} curves, which were generated for seven drugs, indicated inhibition potencies ranging from 20 to 430 μM. Two of the seven drugs were predicted to have systemic plasma concentrations between 10 - 60% of

their IC_{50} values. LogP was statistically higher in the THTR-1 inhibitors compared to the non-inhibitors (p-value < 0.03). Two of the inhibitors, omeprazole and erythromycin, were predicted to inhibit THTR-1 at clinically relevant systemic concentrations and EHR data showed that use these drugs was significantly associated with reduced thiamine pyrophosphate (TPP) levels (p-value < 0.0016).

In conclusion, our study suggests a novel mechanism for drug-induced MA; specifically, inhibition of THTR-1, which may phenocopy TRMA and potentially contribute to the development of MA.

INTRODUCTION

Megaloblastic anemia (MA) is a term that describes a heterogeneous set of anemias that are characterized by the presence of enlarged, immature red blood cells, called megaloblasts, in the bone marrow. MA develops as a result of ineffective DNA synthesis in rapidly dividing hematopoietic stem cells.[1]

Major causes of MA include vitamin deficiencies (i.e. vitamin B_{12}, folic acid; Figure 4.1A) and exposure to certain drugs,[2] also known as drug-induced megaloblastic anemia (DIMA) (Figure 4.1). DIMA can be caused by a variety of drugs that inhibit DNA synthesis at different points of the pathway.[2] For example, drugs can cause MA by reducing cellular availability of vitamin B_{12} or folic acid through "interference with the absorption, plasma transport, or delivery of folate or vitamin B_{12}, competition for reducing enzymes, or end-product inhibition of cofactor-mediated reactions" (Figure 4.1A, #1, #2).[2] Folic acid and Vitamin B12 are required for the synthesis of thymidine, one of four nucleoside bases used for DNA synthesis. Furthermore, drugs can act as purine and pyrimidine antagonists or analogues, also causing reduction in DNA synthesis. For

example, certain chemotherapies (e.g., 5-fluorouracil), immune antagonists (e.g. leflunomide), and antiviral agents (Figure 4.1B, #4) may act as nucleoside antagonists or analogs, thus reducing DNA synthesis. Additional mechanisms of drug-induced MA include inhibition of key enzymes such as dihydrofolatereductase (Figure 4.1B, #3). Interestingly, some drugs cause MA through unknown mechanisms.[2]
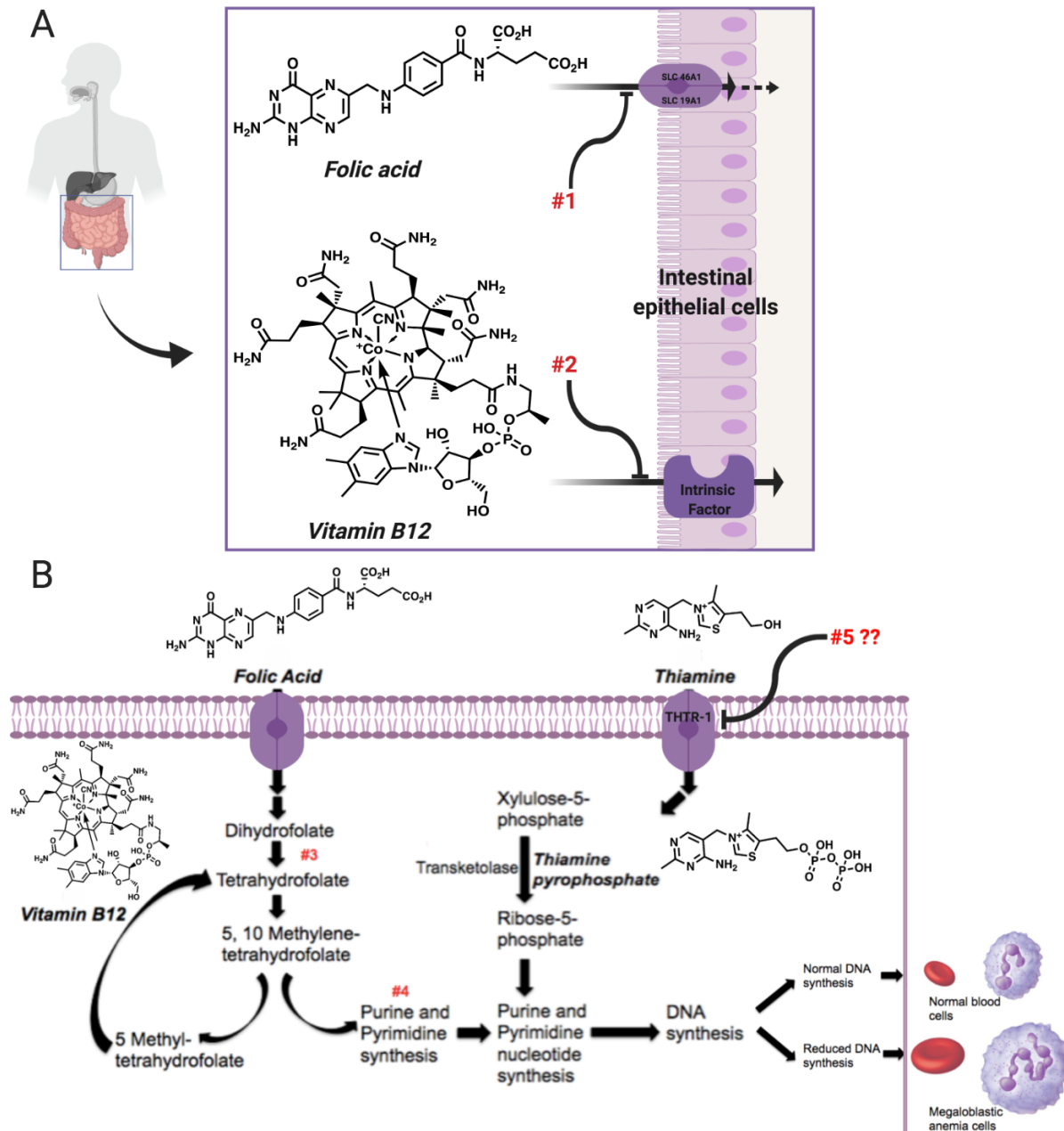
Figure 4.1 Mechanisms of drug induced megaloblastic anemia.

Megaloblastic anemia (MA) is a type of anemia that occurs as a result of **inhibition of DNA synthesis** during red blood cell production. Common causes of MA involve a disruption in folic acid or vitamin B12 absorption (#1, #2; Figure 4.1A). The etiology of drug induced MA can include drugs that reduce in folic acid absorption (#1), drugs that reduce vitamin B12 absorption (#2), drugs that inhibit dihydrofolatereductase (DHFR, #3) and drugs that inhibit purine/pyrimidine synthesis (#4) (Figure 4.1B). An occasional cause of MA includes thiamine transport deficiency (THTR-1; #5) (Figure 4.1B). MA is characterized by abnormally large, immature red blood cells.

93

In addition to common causes of MA such as vitamin deficiencies or use of particular prescription drugs, an infrequent cause of MA is an autosomal, recessive disorder known as thiamine-responsive megaloblastic anemia (TRMA). This syndrome is characterized by three hallmark signs: MA, non-autoimmune type 1 diabetes and sensorineural hearing loss.[3] TRMA is caused by deleterious mutations in the *SLC19A2*, which encodes for Thiamine Human Transporter 1 (THTR-1), which is expressed ubiquitously in many different tissues (e.g. skeletal muscle, cardiac muscle, liver, kidney, pancreas and heart.[4] THTR-1 is the main thiamine transporter in pancreatic islet tissue and hematopoietic progenitor cells.[5, 6] Interestingly, TRMA patients[7] and knockout mice (*Slc19a2 -/-*)[8] demonstrate normal intestinal thiamine uptake and normal plasma thiamine levels in contrast to humans and mice deficient in THTR-2 (encoded by *SLC19A3*), a second thiamine transporter expressed mainly in the intestine, liver and kidney. THTR-2 deficiency results in significantly reduced intestinal thiamine uptake and reduced plasma thiamine levels, but has not been associated with TRMA.[9]

The mechanism for TRMA, similar to MA from other causes, is related to inefficient DNA synthesis. In particular, the active metabolite of thiamine, thiamine pyrophosphate (TPP), serves as a crucial cofactor for transketolase, the primary enzyme used in the pentose phosphate pathway for *de novo* production of ribose [10]. Ribose is a necessary precursor for nucleic acid synthesis. Several studies have found that a lack of thiamine in fibroblasts from patients with TRMA or in fibroblasts from *Slc19a2 -/-* mice exhibit significant reductions in ribose levels and deoxyribose levels and accordingly, reduced nucleic acid synthesis.[10,11,12]

Based on the finding that TRMA and DIMA have similar mechanisms and that certain drugs associated with DIMA inhibit THTR-1 (i.e. trimethoprim),[2,13] we hypothesized that

a potential novel mechanism for DIMA may be through inhibition of THTR-1,

phenocopying TRMA. To test our hypothesis, we first explored *in vitro* reduction in

THTR-1 mediated thiamine uptake by genetic variants of THTR-1 associated with

TRMA and several prescription drugs previously identified as inhibitors of THTR-1.

Following these studies, we then screened drugs that are associated with DIMA to

determine if any were able to inhibit THTR-1, and examined the potency of inhibition

through IC_{50} experiments. Electronic medical record data were explored to evaluate the

clinical impact of prescription drugs that were predicted to inhibit THTR-1 on blood

levels of TPP.

METHODS

Chemicals and Reagents

 [3H]-Thiamine hydrochloride (catalog item ART 0710) was purchased from American

Radiolabeled Chemicals Incorporation (St. Louis, MO, USA). The specific activity of [3H]-

thiamine was 20 Ci/mmol. Non-radiolabeled compounds were purchased from Sigma-Aldrich,

Inc (USA), Selleck Chemicals (Houston, TX), VWR International, Inc (USA) and Thermo Fisher

Scientific (USA). Cell culture supplies were purchased from the Cell Culture Facility (UCSF,

California, USA) and Thermo Fisher Scientific (USA).

Radiolabeled thiamine inhibition assays in HEK 293 Cells

HEK 293 cells stably expressing THTR-1 were produced as previously described.[14] These cells

were plated in poly-D-lysine coated 24-well plates (200,000 cells/well); empty vector HEK 293

cells were also plated as a negative control. Once cells reached 90% confluency, THTR-1

expressing cells were transiently transfected with pcDNA5/FRT-SLC19A2 construct (500 ng per

well) using Lipofectamine® LTX with Plus™ reagent according to the manufacturer's

instructions (Life Technologies, Carlsbad, CA). Transfected cells were maintained in DMEM

containing 10% (v/v) FBS, 100 U/ml penicillin, 100 μg/ml streptomycin, and 5 μg/ml puromycin

(Thermo Fisher Scientific) at 37°C and 5% CO_2 for 48 hours.

Functional characterization of *SLC19A2* variants causal for TRMA

There are at least 30 *SLC19A2* mutations that are associated with TRMA[3,15], but only the activity

and/or expression of four TRMA variants has been characterized and published. We selected

three previously uncharacterized TRMA variants and characterized their activity using our *in*

vitro assay (Figure 4.2). Variants in the pcDNA5/FRT-SLC19A2 plasmid were generated using

site-directed mutagenesis (Genscript Inc); presence of each tested variant was confirmed using

Sanger sequencing (MCLAB Inc).

Selection of Drugs for Screening

Table 1 in the review article *Drug Induced Megaloblastic Anemia*,[2] contains a list of prescription

drugs and compounds associated with MA. Those compounds in addition to others belonging to

listed drug classes (e.g. proton-pump inhibitors; a total of 63 compounds) were screened in

triplicate at a 1 mM concentration to evaluate their potential to inhibit THTR-1 using the

protocol depicted in Figure 4.3.
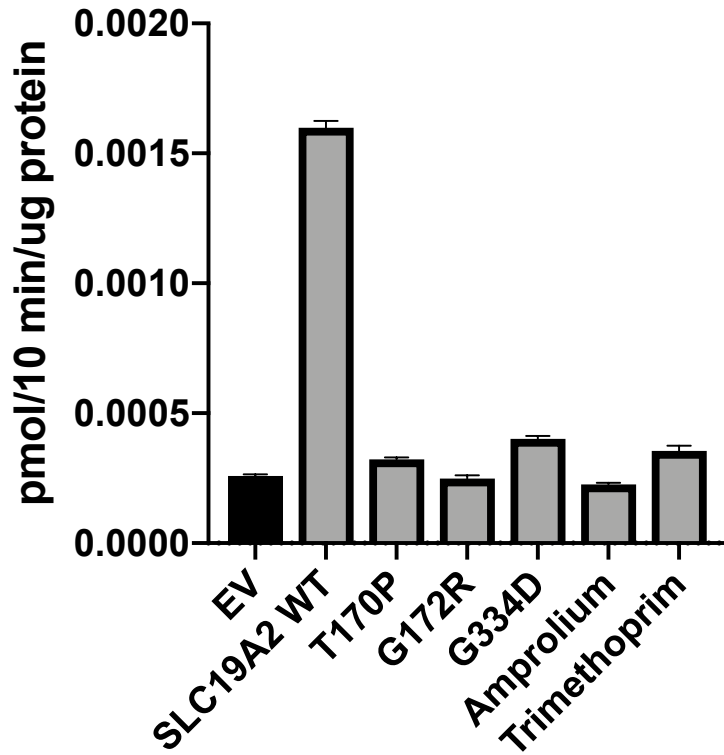
Figure 4.2 Functional characterization of *SLC19A2* TRMA variants.
Several *SLC19A2* variants known to be causal for TRMA and never previously tested for activity were functionally characterized for transport activity using radiolabeled thiamine. Thiamine transport activities of the G172R was not statistically different from empty vector activity (p-value = 0.5). At 200 μM, trimethoprim, a known inhibitor of *SLC19A2* and drug associated with DIMA, reduced transport activity of wild type *SLC19A2* to levels comparable to the activity of deleterious TRMA variants.
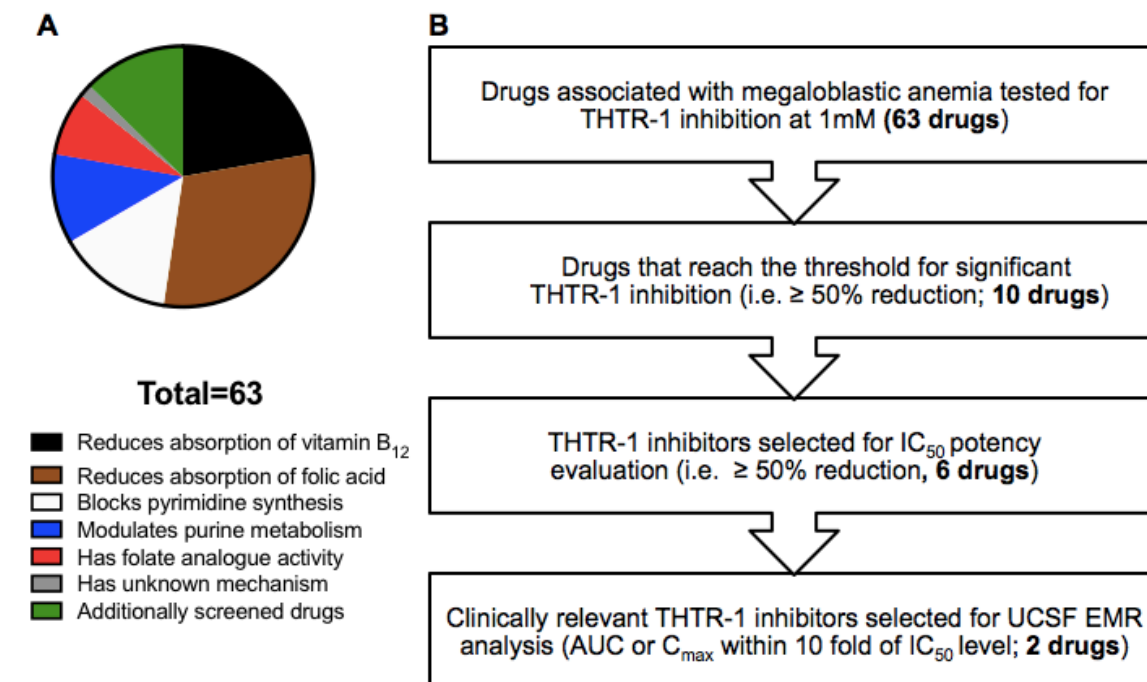
Figure 4.3 Common mechanisms of drug-induced megaloblastic anemia and THTR-1 inhibitor drug screen workflow.

Drugs associated with inducing megaloblastic anemia through various mechanisms (A) were obtained through commercial sources Solutions containing each of the drugs (1 mM) in buffer containing [^{3}H] thiamine were exposed to HEK 293 cells expressing THTR-1 for 10 minutes. Drugs that inhibited 50% or more of THTR-1 mediated [^{3}H] thiamine uptake were designated as inhibitors (10 total). Next, IC_{50} experiments were conducted on the THTR-1 inhibitors that did not have previously published IC_{50} information.   Last, we selected THTR-1 inhibitors for EHR analysis that were predicted to have systemic plasma levels or exposure within 10 fold of the estimated IC_{50} (B).

Methods Used for Screening THTR-1 for Inhibitors

Screening was performed using [^{3}H]-Thiamine (250 nM total of thiamine dissolved in Hank's balanced salt solution (HBSS) (Figure 4.4A); wells were exposed for ten minutes (within linear range) to radiolabeled thiamine (positive control), radiolabeled thiamine plus 1 mM of amprolium (negative control) and radiolabeled thiamine plus 1mM of screened drug (3 replicates per drug). Next, IC$_{50}$ assays were performed using a range of inhibitor concentrations (0.5 μM – 1 mM, three replicates per concentration, Figure 4.4B). After ten minutes, the reaction was terminated by washing cells with ice-cold HBSS three times; cells were lysed for one hour using 800 μl of lysis buffer (0.1 N NaOH and 0.1% SDS). The radioactivity (disintegrations per minute, DPM) of 450 μl of lysate was measured using a liquid scintillation counter machine and the DPM values were converted to picomoles of substrate. Total protein concentrations were determined using the bicinchoninic acid assay using 25 μl of lysate per well (ThermoFisher Scientific). Substrate uptake (in picomoles) was normalized by protein amount in each respective well.

Calculating and/or identifying systemic plasma levels of THTR-1 inhibitors

Using clinical pharmacokinetic data from human studies[16,17,18,19,20,21,22] and drug monograph databases (e.g. Micromedex, FDA) we obtained C$_{max}$, fraction unbound (f$_u$), maximum single dose and other parameters to determine if a single dose of inhibitor taken by patients will reach clinically relevant concentrations in the systemic circulation (i.e. if the R-value, the predicted ratio of the substrate's AUC in the presence and absence of an inhibitor, ≥ 1.1) (Table 4.1).[23]

Generation and analysis of physicochemical properties of THTR-1 drug screening library

Molecular structure files were obtained *via* ChemAxon Marvin Sketch (version 19.20.0). Based on the structures, 10 molecular descriptors (molecular weight, total polar surface area, logP, Van de Waals volume, formal charge, average molecular polarizability, hydrogen bond acceptor count, hydrogen bond donor count, number of rotatable bonds, and number of heavy atoms) were computed using ChemAxon Calculator (cxcalc; version 19.20.0). Statistical analysis and boxplots were generated using Graphpad Prism 8 (version 8.3.0). Boxplots display the following data distribution statistics: minimum, first quartile, median, third quartile, and maximum. Boxplot whiskers indicate the minimum and maximum values (Figure 4.5).

Electronic medical record analysis of potent THTR-1 inhibitors

As previously mentioned, thiamine levels in the blood are normal in TRMA patients and *Slc19a2($^{-/-}$)* knockout mice.[9] However, Xian and colleagues found that the concentration of TPP, the bioactive form of thiamine, was reduced in TRMA patients compared with normal patients.[24] Therefore, we used the UCSF electronic health record (EHR) database to evaluate TPP levels in the blood as a surrogate marker to track a potential effect of THTR-1 inhibition in hematopoietic cells and tissues by inhibitors identified in our screen and predicted to achieve plasma concentrations in the range of their THTR-1 IC_{50} values. The UCSF Research Data Browser was used to search for patients who had a TPP lab (measured in whole blood by the UCSF Health Clinical Laboratories) test value reported, which gave us a total of 1,730 patients and 2,486 lab values. TPP levels reported as an inequality were changed to a numerical value (i.e. < 1.6nM = 1.6nM).

Patients were divided into two groups depending on their medication use. Patients prescribed erythromycin, which was predicted to achieve clinically relevant concentrations for inhibition of

THTR-1, were grouped into the "on" drug group. Search terms included "Erythromycin", "Ery-tab", "Eryc", "Erythrocin", "Eryped", and "E.E.S." for erythromycin. Medication orders whose route of administration was topical, ophthalmic, and miscellaneous as well as orders without a medication order start date were excluded (Figure 4.6A). The remaining patients (i.e. individuals who were never prescribed erythromycin) were grouped into the "off" drug group. Only patients with one TPP level were included in the "off" group, which reduced our sample size to 1278 individuals for the erythromycin analysis (Figure 4.6A).

Patients in the "on" group were further filtered based on their laboratory collection date relative to their first medication order start date. For erythromycin, TPP levels measured within 7 to 60 days after their first medication order start date were included. A minimum of 7 days between medication start date and the TPP level was chosen to allow for an effect to be observed. For patients with more than one lab value, only the lab value closest to the first medication order start date was included. Lastly, patients were age- and sex-matched using the MatchIt package in R to be consistent in both the "on" and "off" groups which resulted in a final sample sizes of 4 patients "on" drug and 200 patients "off" drug for the erythromycin analysis (Table 4.2)

Despite omeprazole's R-value being < 1.1, EHR analysis of omeprazole, which had the lowest IC_{50} of all investigated proton pump inhibitors, was conducted. Search terms included "Omeprazole" and "Prilosec". Only patients with one TPP level were included in the omeprazole "off" group, which reduced the sample size to 893. These patients were age- and sex-matched using the MatchIt package in R, which resulted in 523 patients "on" drug and 523 patients "off" drug for the omeprazole analysis.

For patient population-specific analysis, patients were further sub-grouped based on a diagnosis of anemia. More specifically, anemia was defined as ICD10 Level 1 "Diseases of the blood and blood forming organs and certain disorders involving the immune mechanism (D50-D89)", ICD10 Level 2 "Nutritional anemias (D50-D53)", and ICD10 Level 3 "Folate deficiency", "Other nutritional anemia", and "Vitamin B_{12} anemia". There were 3,767 patients with one or more of these diagnoses in the EHR database. The same workflow as described above was used for the patient population-specific analysis, but only lab values taken before or on date of diagnosis were included. Again, patients were age- and sex-matched using the MatchIt package in R, which resulted in 7 patients in the omeprazole "on" drug group and 28 patients in the omeprazole "off" drug group (Table 4.2). Based on the aforementioned criteria, the number of patients with a diagnosis of nutritional anemia and "on" erythromycin was too small to conduct analyses. Welch's two sample t-test was performed to evaluate if there was a significant difference in lab values when comparing both groups; ggplot package in R (version 3.4.0) and GraphPad Prism 7 software was used to plot the data.

RESULTS

THTR-1 variants causal for TRMA show significantly reduced thiamine uptake, which is comparable to thiamine uptake by reference THTR-1 exposed to certain prescription drugs.

Three previously uncharacterized variants, associated with TRMA, significantly reduced thiamine uptake in the cells (Figure 4.2). In brief, HEK cells transfected with the TRMA variants had significantly reduced thiamine transport activity compared to cells transfected with wild-type THTR-1 (one-way ANOVA p-value < 0.0001). TRMA transporter activity of one of

the three mutant transporters (i.e. G172R) was not significantly different from that of empty

vector cells (Figure 4.2, p-value = 0.5) consistent with virtually no transport activity. We found

that known inhibitors of THTR-1 transport (trimethoprim, amprolium) at 200 μM significantly

reduced thiamine uptake by the reference THTR-1. Notably, we observed that the reduction in

thiamine uptake was not significantly different between the THTR-1 wild-type transporters

exposed to the drugs compared to thiamine uptake by the mutant THTR-1 transporters (p-value ≥

0.1). Though potency values need to be established, these results provide *in vitro* proof of

concept that drugs may potentially phenocopy TRMA. Based on these results we hypothesized

that certain drugs that are associated with megaloblastic anemia may alter the absorption and

distribution of thiamine in hematopoietic cells/tissues, by acting as inhibitors of THTR-1.

A number of drugs known to cause MA were identified as inhibitors of THTR-1 in screening studies.

Over 60 drugs known to induce DIMA through various mechanisms were screened to identify if

any were inhibitors of THTR-1 thiamine transport (Figure 4.3, Figure 4.4). Amprolium, a known

inhibitor of THTR-1 [14], was used as a negative control. Among the drugs screened, ten drugs

were found to be substantial inhibitors of THTR-1 (≥ 50% reduction in THTR-1 transport). Of

the ten inhibitors, six were novel (not previously known to inhibit the transporter) and included

erythromycin, mycophenolate mofetil, omeprazole, pantoprazole, lansoprazole and chloroquine

(Figure 4.4A); the other four compounds (fedratinib, amiloride, trimethoprim, amprolium) were

previously shown to inhibit THTR-1.[14,13,25] Leflunomide was found to be an intermediate

inhibitor of THTR-1 (> 40%, but < 50% reduction in transport). Erythromycin is a tetracycline

antibiotic used to treat a variety of infections. Mycophenolate mofetil is a drug used in solid-

organ transplant patients to prevent organ rejection. Omeprazole, pantoprazole and lansoprazole

are over-the-counter proton pump inhibitors commonly used to treat acid reflux disease. Chloroquine is an anti-malarial drug that is not commonly used in the U.S. Amiloride is a potassium-sparing diuretic often used in combination with other medications for blood pressure reduction and/or heart failure.
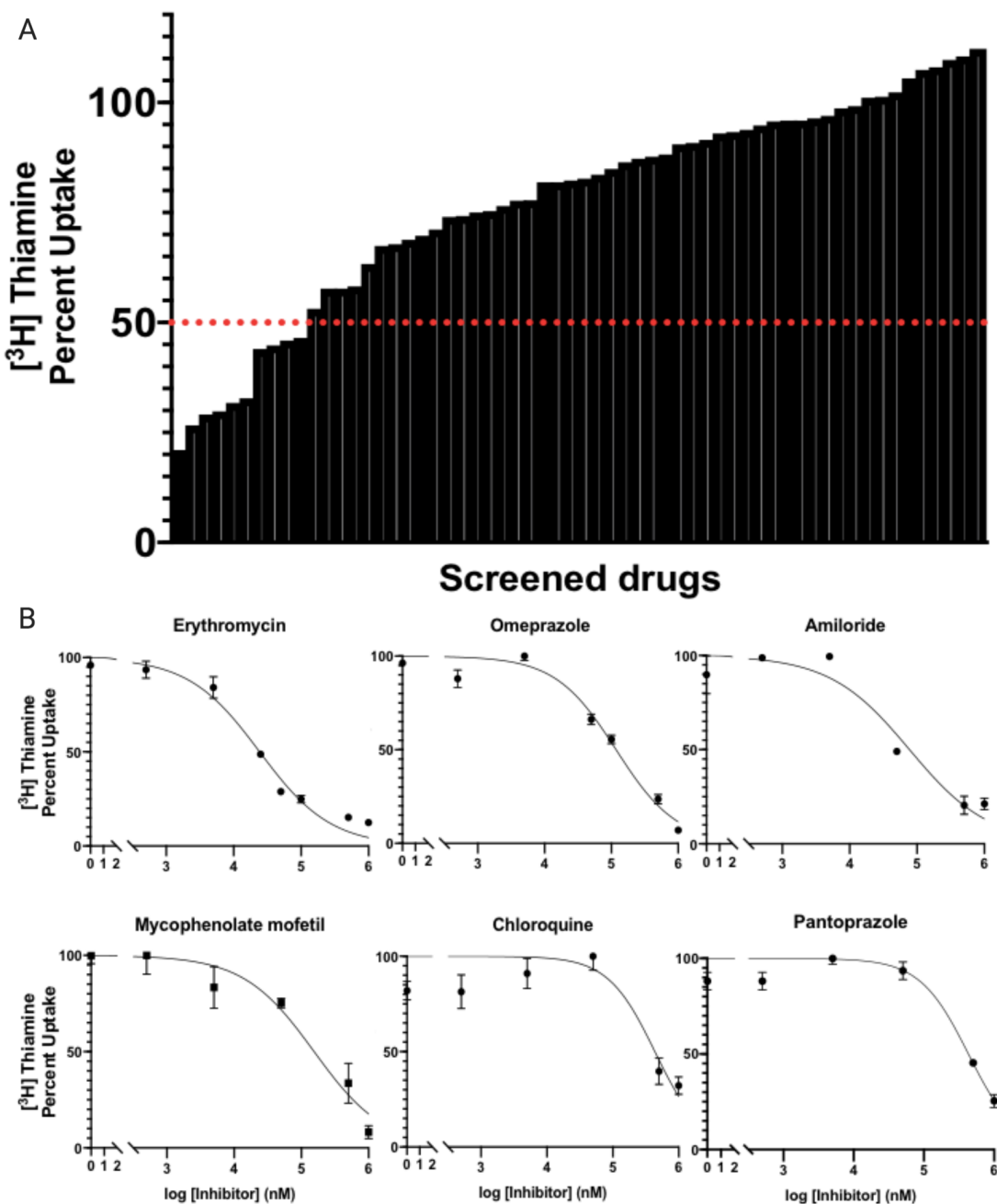
Figure 4.4 THTR-1 drug library screen and IC_{50} evaluation.
In the screen, THTR-1 inhibitors were defined as compounds that inhibit at least 50% (dotted line) of thiamine uptake (A). Out of the 63 drugs screened, ten reached that threshold. Experiments were conducted using inhibitor concentrations from 0 to 1 mM. IC_{50} curves for erythromycin, omeprazole, amiloride, mycophenolate mofetil, cholorquine and pantroprazole were generated by fitting data to a nonlinear model (B). Figures representative of 3 separate experiments except for pantoprazole (graph representative of one experiment).

Hydrophobicity associates with THTR-1 inhibition

Ten physicochemical properties of THTR-1 inhibitors and non-inhibitors were computed and compared: molecular weight, total polar surface area, logP, Van de Waals volume, formal charge, average molecular polarizability, hydrogen bond acceptor count, hydrogen bond donor count, number of rotatable bonds, and number of heavy atoms. The majority of molecular descriptors did not statistically differentiate THTR-1 inhibitors from non-inhibitors. The most statistically significant and informative physicochemical property was logP (Figure 4.5, student t-test p-value = 0.03) suggesting that THTR-1 inhibitors tend to be more hydrophobic than non-inhibitors.
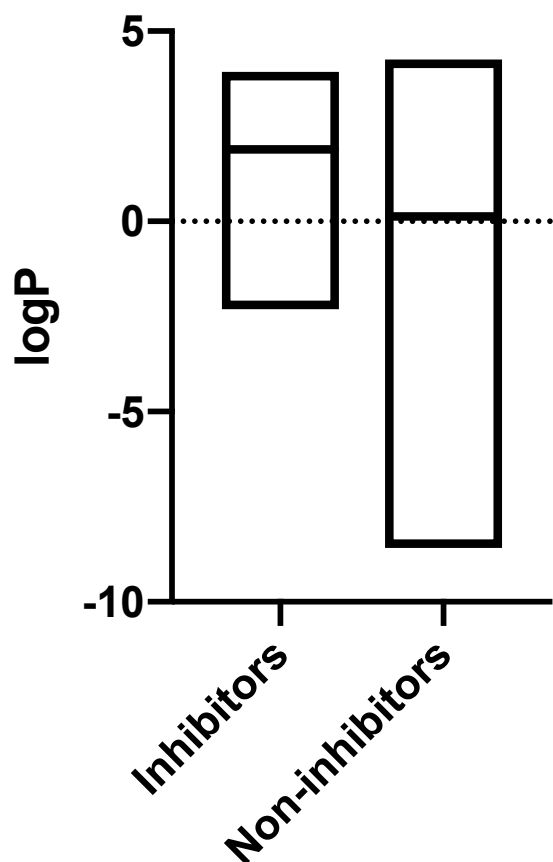
Figure 4.5 LogP associates with THTR-1 inhibition.
Ten molecular descriptors were generated via Chemaxon and DataWarrior. Out of the ten descriptors, increased logP (hydrophobicity) was found to be the only significant descriptor to differentiate inhibitors from non-inhibitors (p-value = 0.03). H bond donor count demonstrated a trend towards *SLC19A2* inhibitors having lower H-bond donors (p-value = 0.052).

Potency of inhibitors of THTR-1 ranged between 2.6 and above 400 µM

We evaluated the THTR-1 inhibition potency of the six novel inhibitors plus amiloride by conducting IC_{50} studies (Figure 4.4B). In contrast to fedratinib, amprolium and trimethoprim, previous IC_{50} studies on amiloride have not been published. Interestingly, amprolium was the most potent inhibitor of THTR-1 (Table 4.4). Of the novel inhibitors that were identified in our screen, erythromycin was the most potent inhibitor of THTR-1 with an estimated IC_{50} of 20 µM.

Table 4.1 THTR-1 inhibitor potency (IC$_{50}$).

Prescription drug	Mean IC$_{50}$ (µM)	SEM (µM)	SLC19A3 inhibitor?
Amiloride	69	17.7	Yes
Chloroquine	301	33.4	No
Erythromycin	21	8.24	No
Mycophenolate mofetil	145	9.93	No
Omeprazole	207	27.7	No
◆Pantoprazole	429	--	No
*Amprolium	2.6	0.93	Yes
*Fedratinib	7.1	1.3	Yes
*Trimethoprim	6.8	1.7	Yes

Experiments were conducted using inhibitor concentrations from 0 to 1 mM. IC$_{50}$ curves were generated by fitting data to a nonlinear model; average IC$_{50}$ estimates plus SEM were generated. * IC$_{50}$ ± SD values were extracted from previously published data (PMID: 27803021). Data representative of 3 separate experiments except for pantoprazole (◆ = graph representative of one experiment).

Erythromycin predicted to inhibit THTR-1 thiamine transport in the systemic circulation

To compare the THTR-1 IC_{50} to the inhibitor's systemic blood levels in humans, we collected and used the C_{max}, fraction unbound (f_u), and other parameters from clinical pharmacokinetic studies to calculate the R-value for all six THTR-1 inhibitors (Table 4.1). Drugs found to have an R-value ≥ 1.1 were considered potentially clinically relevant THTR-1 inhibitors.[23] After conducting R-value calculations, we found that erythromycin was the only THTR-1 inhibitor with an R-value ≥ 1.1. Essentially, a THTR-1 inhibitor with an R-value ≥ 1.1 is predicted to reach concentrations in the systemic circulation that reduce THTR-1 thiamine transport into hematopoietic progenitor cells, which are found in bone marrow.

Erythromycin, designated as a potentially clinically relevant THTR-1 inhibitor, reduces thiamine pyrophosphate levels in red blood cells

To determine the clinical relevance of inhibitors we identified in the screen, we mined the electronic health records at the University of California, San Francisco to identify drug-induced decreases in TPP laboratory values. Specifically, we compared TPP levels in individuals prescribed erythromycin or omeprazole versus levels in individuals never prescribed either of these medications. Based on the inclusion/exclusion criteria described in the methods, we were able to classify patients as "on" drug (i.e. on omeprazole or erythromycin) or "off" drug. Population demographics, such as age and sex, were matched and comparable between both groups (Figure 4.6A, Table 4.2). When examining data from all individuals in the EHR database, patients in the "on" group have significantly lower TPP levels compared to the "off" group for both omeprazole (student's t-test, p-value: 0.0004; Figure 4.6B) and erythromycin (student's t-test, p-value: 0.01). Figure 4.6 shows the results of our study and presents the flow of our analyses. When we performed a patient-specific analysis and looked at individuals

diagnosed with folate, vitamin B_{12}, and/or other nutritional anemia, although not significant, the trend still held where patients on omeprazole had lower TPP levels than patients not prescribed omeprazole (p-value: 0.35; n = 7 on drug, n = 28 off drug). There were no patients who met all the inclusion/exclusion criteria for the erythromycin patient-specific analysis.
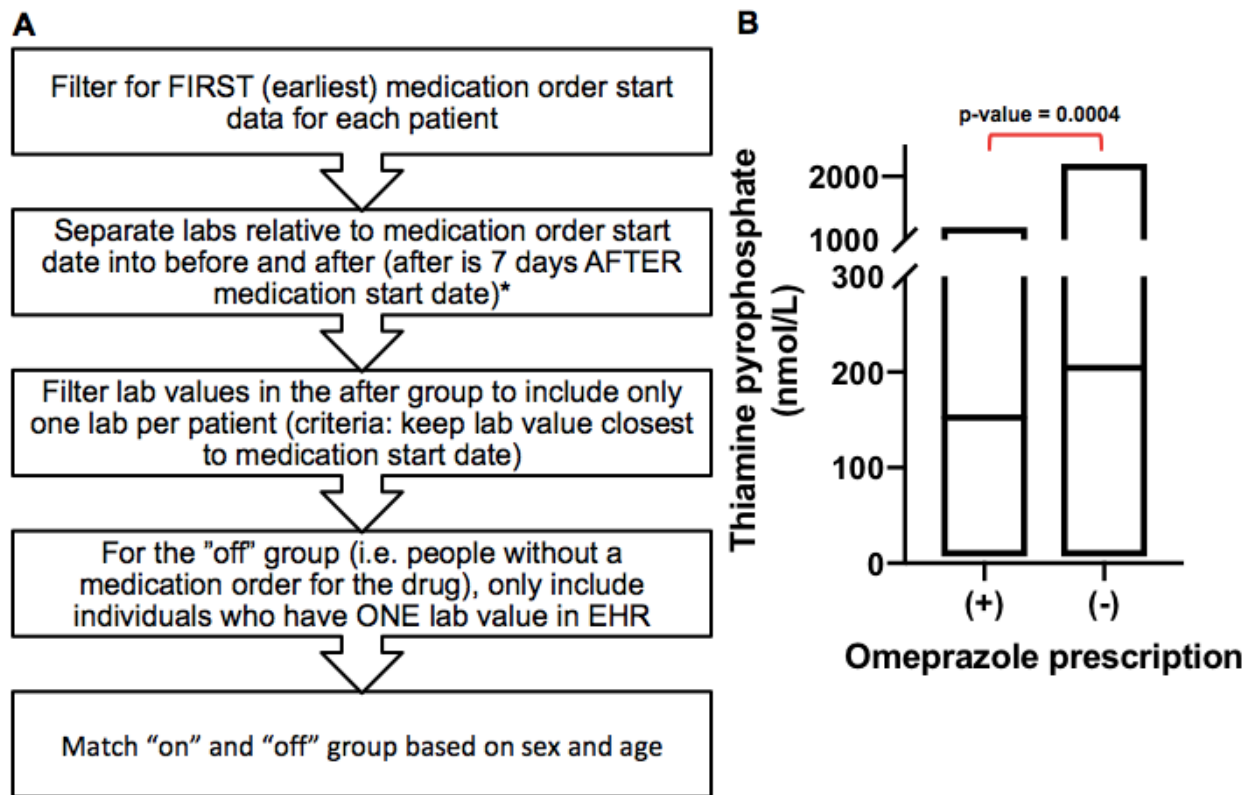
Figure 4.6 Average thiamine pyrophosphate levels in the general population of patients on or off omeprazole.

Patients with or without a diagnosis of megaloblastic anemia who had one thiamine pyrophosphate level were extracted from the UCSF EHR (A). The thiamine pyrophosphate levels of patients on an inhibitor (omeprazole or erythromycin) versus those not on an inhibitor were compared. The average TPP levels for patients on omeprazole was significantly lower (152 nM, N=510) than that of the TPP levels of patients not taking omeprazole (201 nM, N=510) (students t-test, p-value = 0.0004) (B). The average TPP levels for patients on erythromycin was lower (137 nM) than that of the TPP levels of patients not taking omeprazole (186 nM) (p-value = 0.0016)

Table 4.2 Average age and gender in drug-exposed versus unexposed groups.

Omeprazole analysis		
	On	**Off**
Gender		
% Female	71	71
% Male	29	29
Age		
Mean	48.8	49.3
Erythromycin analysis		
	On	**Off**
Gender		
% Female	50	54
% Male	50	46
Age		
Mean	32.3	33.8
Omeprazole - Megaloblastic Anemia analysis		
	On	**Off**
Gender		
% Female	57	61
% Male	43	39
Age		
Mean	55.1	51.8

Patients were age- and sex-matched using the MatchIt package in R to be consistent in both the "on" and "off" groups. The average age and gender percentage for each analysis was not statistically different between the "on" and "off" groups (Welch's two sample t-test, p-value > 0.5)

Discussion and Conclusions

Comprehensive THTR-1 inhibitor screens and kinetic evaluations of drugs associated with

DIMA were conducted in this study. Furthermore, EHR data of patients who were or were not

exposed to THTR-1 inhibitors predicted to reach plasma levels that may inhibit the transporter

were extracted and analyzed. The study led to five findings:

First our screen of drugs that inhibit THTR-1 indicates that a limited number of prescription

drugs may inhibit the transporter and suggests the possibility of THTR-1 drug-vitamin

interactions. Many prescription library inhibitor screens have been conducted for drug

transporters (i.e. OATPs, OCTs, MATEs), which modulate the disposition of many prescription

drugs. [26,27,28,29] The FDA recommends that any new investigational small molecule be

tested/screened *in vitro* for significant inhibition of intestinal, hepatic and renal drug transporters

such as organic anion transporting polypeptides (OATPs), organic cation transporter (OCTs) and

multi-antimicrobial extrusion protein (MATEs).[23] If the R-value for an OATP1B1/1B3 inhibitor

≥ 1.1, the inhibitor is predicted to affect the substrate disposition *in-vivo* and a clinical drug-drug

interaction trial may be warranted.[23] However, to date, the FDA has not developed

recommendations for testing/screening small molecules as inhibitors of critical vitamin

transporters to assess the liability of new drugs to cause drug-vitamin interactions. Though our

study represents the largest prescription drug library screen for inhibitors of THTR-1, the screen

was still limited. Notably, 60 drugs represent less than 10% of approved oral drugs suggesting

that larger screens may be warranted to identify prescription drug inhibitors of THTR-1 and

other nutrient transporters. Such screens should be guided by clinical information. For example,

if other detrimental drug-vitamin interactions (e.g. fedratinib) are exposed,[30, 31] additional

prescription drug library screens focused on critical vitamin transporters should be conducted to

identify other potential drug-vitamin interactions and inform clinicians and patients regarding risk for drug-vitamin interactions.

Our second finding is that inhibitors of THTR-1 have greater logP values in comparison to non-inhibitors (Figure 4.5). Greater logP values have been associated with inhibitors of other transporters (e.g., OCT1 and MATE1), [27] therefore a high logP value is likely a general property of many transporter inhibitors. As our study was limited to 60 drugs, future computational analysis together with high-throughput screens with larger libraries are needed to identify molecular/chemical signatures that are specific to THTR-1 inhibitors; these signatures can be used to pre-screen new molecular entities or other prescription drugs before use in populations vulnerable to MA and DIMA (e.g. elderly).

Our third finding is that several prescription drugs associated with MA inhibit THTR-1 with IC_{50} values of 200 μM or lower (i.e., trimethoprim, fedratinib, amiloride, mycophenolate mofetil, erythromycin, omeprazole). At least four (trimethoprim, mycophenolate mofetil, erythromycin, omeprazole) of the six drugs have been thought to cause MA through other validated mechanisms (e.g., absorptive inhibition of folic acid or vitamin B_{12}).[2] Though further studies are clearly needed, it is possible that inhibition of THTR-1 represents an additional mechanism for inducing MA for some of these drugs.

Fourth, our study demonstrates that erythromycin is a potent THTR-1 inhibitor that may inhibit the transporter clinically, that is, at concentrations associated with therapeutic doses of the drugs. In further studies, we found that erythromycin and omeprazole are associated with significantly

reduced blood levels of TPP, the bioactive form of thiamine, using EHR data. Most small molecule FDA approved drugs have systemic blood concentrations within the nanomolar to micromolar range or less.[32] We determined through examination of pharmacokinetic data that erythromycin's R-value for THTR-1 ≥1.1, suggesting that it may inhibit the transporter at therapeutic doses. Erythromycin is a macrolide antibiotic that was commonly used in the past to treat skin/soft tissue and blood infections but was associated with many drug-drug interactions and adverse effects, many of which were unexplained. Of note, is that hearing loss is a known adverse effect of erythromycin. Our data showing that erythromycin may inhibit THTR-1 suggest a potential mechanism for this adverse effect. That is, TRMA caused by genetic mutations in *SLC19A2* is generally characterized by hearing loss. Though speculative, inhibition of THTR-1 may be a mechanism of erythromycin induced hearing loss. However today, other macrolides with fewer drug-drug interactions and adverse effects are commonly used in place of erythromycin. Additionally, we found the R-value for omeprazole was < 1.1, which means it is not predicted to inhibit THTR-1 *in-vivo*. This result was based on a range of C_{max} levels found in patients who took a single dose of omeprazole between 40-90mg.[33] Omeprazole is a first-in-class proton pump inhibitor used to treat symptoms of gastrointestinal reflux disease (GERD) or gastrointestinal ulcers. Its estimated that over 70 million prescriptions for omeprazole were generated in 2016. [34] Importantly, the drug is chronically used with an average duration of treatment of 3 months with some patients on omeprazole for years.[35,36] Notably, a subset patients treated with omeprazole have hypersecretory disorders (e.g. Zollinger –Ellison syndrome) that require very high doses of omeprazole (i.e. 120mg three times a day for a total daily dose of 360mg). It is possible that the initial C_{max} and steady state C_{max} values in patients treated with high dose omeprazole could be much higher than those treated with normal doses, especially if

these patients are CYP2C19 poor metabolizers. Thus, patients chronically receiving

erythromycin or large doses of omeprazole with CYP2C19 inhibition may have sufficiently high

concentrations over extended periods of time to inhibit THTR-1 in hematopoietic cells and

potentially reduce the levels of TPP. Lower TPP levels in blood cells may result in reduced

DNA synthesis and lead to increased susceptibility to MA. Our EHR data analysis supports this

possibility (Figure 4.6).

Finally, our data show that erythromycin and omeprazole are associated with significantly lower

levels of TPP, the bioactive form of thiamine, in UCSF medical center patients. EHR data are

notoriously noisy; therefore, selected patients in the "on" and "off" medication groups were age

and sex-matched and the reduction in TPP levels kept its significance. Even in our subgroup

analysis (i.e. megaloblastic anemia) the trend of lower TPP levels in patients on omeprazole still

held. The reduction of TPP levels in patients exposed to erythromycin or omeprazole seems to

"phenocopy" the effects of genetic THTR-1 deficiency on TPP levels in TRMA patients.[24]

Additionally, we searched the EHR for patients who were exposed to either erythromycin or

omeprazole and had vitamin B_{12}, folic acid, RBC and TPP levels measured before their diagnosis

of MA. Out of thousands of patients diagnosed with MA, only two patients on omeprazole had

all three levels obtained before diagnosis. One of the patients was diagnosed with "Other

nutritional anemia" and had normal vitamin B_{12} and folic acid levels in RBCs, but low TPP

levels (data not shown). This result suggests that despite the clear ability of thiamine deficiency

to induce MA, TPP is not routinely ordered in patients suspected of having DIMA or MA;

therefore the incidence of DIMA caused by THTR-1 inhibitors may be significantly under

reported. Though further studies are clearly needed, our study highlights an underrated etiology

of megaloblastic anemia (thiamine deficiency) and suggests an additional mechanism of DIMA, THTR-1 inhibition.

REFERENCES

1. Green R. Megaloblastic Anemias Nutritional and Other Causes. *Med Clin NA*. 2017;101(2):297-317. doi:10.1016/j.mcna.2016.09.013

2. Hesdorffer CS, Longo DL. Drug-Induced Megaloblastic Anemia. *N Engl J Med*. 2015;373(17):1649-1658. doi:10.1056/NEJMra1508861

3. Raz T, Labay V, Baron D, et al. The spectrum of mutations, including four novel ones, in the Thiamine- Responsive Megaloblastic Anemia gene SLC19A2 of eight families. *Hum Mutat*. 2000;16(1):37-43. doi:10.1002/1098-1004(200007)16:1<37::AID-HUMU7>3.0.CO;2-9

4. Ortigoza-Escobar JD, Molero-Luis M, Arias A, et al. Treatment of genetic defects of thiamine transport and metabolism. *Expert Rev Neurother*. 2016;16(7):755-763. doi:10.1080/14737175.2016.1187562

5. Mee L, Nabokina SM, Sekar VT, Subramanian VS, Maedler K, Said HM. Pancreatic beta cells and islets take up thiamin by a regulated carrier-mediated process: studies using mice and human pancreatic preparations. *Am J Physiol Gastrointest Liver Physiol*. 2009;297(1):G197-206. doi:10.1152/ajpgi.00092.2009

6. Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419. doi:10.1126/science.1260419

7. Neufeld EJ, Fleming JC, Tartaglini E, Steinkamp MP. Thiamine-responsive megaloblastic anemia syndrome: a disorder of high-affinity thiamine transport. *Blood Cells Mol Dis*. 2001;27(1):135-138. doi:10.1006/bcmd.2000.0356

8. Oishi K, Hofmann S, Diaz GA, et al. Targeted disruption of Slc19a2, the gene encoding the high-affinity thiamin transporter Thtr-1, causes diabetes mellitus, sensorineural

deafness and megaloblastosis in mice. *Hum Mol Genet*. 2002;11(23):2951-2960. doi:10.1093/hmg/11.23.2951

9. Reidling JC, Lambrecht N, Kassir M, Said HM. Impaired intestinal vitamin B1 (thiamin) uptake in thiamin transporter-2-deficient mice. *Gastroenterology*. 2010;138(5):1802-1809. doi:10.1053/j.gastro.2009.10.042

10. Green R. Mystery of thiamine-responsive megaloblastic anemia unlocked. *Blood*. 2003;102(10):3464-a-3465. doi:10.1182/blood-2003-09-3025

11. Boros LG, Steinkamp MP, Fleming JC, Lee W-NP, Cascante M, Neufeld EJ. Defective RNA ribose synthesis in fibroblasts from patients with thiamine-responsive megaloblastic anemia (TRMA). *Blood*. 2003;102(10):3556-3561. doi:10.1182/blood-2003-05-1537

12. Sahai I, Montefusco MC, Fleming JC, et al. Role of Defective High-Affinity Thiamine Transporter slc19a2 in Marrow from a Mouse Model of Thiamine-Responsive Anemia Syndrome: Evidence for Defective Deoxyribose and Heme Synthesis. *Blood*. 2005;106(11):516. doi:10.1182/blood.V106.11.516.516

13. Giacomini MM, Hao J, Liang X, et al. Interaction of 2,4-Diaminopyrimidine-Containing Drugs Including Fedratinib and Trimethoprim with Thiamine Transporters. *Drug Metab Dispos*. 2017;45(1):76-85. doi:10.1124/dmd.116.073338

14. Liang X, Chien H-C, Yee SW, et al. Metformin Is a Substrate and Inhibitor of the Human Thiamine Transporter, THTR-2 (SLC19A3). *Mol Pharm*. 2015;12(12):4301-4310. doi:10.1021/acs.molpharmaceut.5b00501

15. Lagarde WH, Underwood LE, Moats-Staats BM, Calikoglu AS. Novel mutation in the SLC19A2 gene in an African-American female with thiamine-responsive megaloblastic anemia syndrome. *Am J Med Genet A*. 2004;125A(3):299-305. doi:10.1002/ajmg.a.20506

16. Landahl S, Andersson T, Larsson M, et al. Pharmacokinetic study of omeprazole in elderly healthy volunteers. *Clin Pharmacokinet*. 1992;23(6):469-476. doi:10.2165/00003088-199223060-00006

17. Nazir S, Iqbal Z, Ahmad L, Ahmad S. Variation in pharmacokinetics of omeprazole and its metabolites by gender and CYP2C19 genotype in Pakistani male and female subjects. :887-894.

18. Kanazawa S, Ohkubo T, Sugawara K. The effects of grapefruit juice on the pharmacokinetics of erythromycin. *Eur J Clin Pharmacol*. 2001;56(11):799-803. doi:10.1007/s002280000229

19. Limited TC. Amiloride Product monograph. 2014:1-18. https://pdf.hres.ca/dpd_pm/00025095.PDF.

20. Miller AK, Harrell E, Ye L, et al. Pharmacokinetic interactions and safety evaluations of coadministered tafenoquine and chloroquine in healthy subjects. *Br J Clin Pharmacol*. 2013;76(6):858-867. doi:10.1111/bcp.12160

21. Inc WP. Pantoprazole product monograph. 2012:1-32. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/020987s045lbl.pdf.

22. Inc RL. Mycophenolate mofetil product monograph. (7):1-43. https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/050722s021,050723s019,050758s019,050759s024lbl.pdf.

23. Food and Drug Administration. *In Vitro Metabolism- and Transporter- Mediated Drug-Drug Interaction Studies Guidance for Industry In Vitro Metabolism- and Transporter-Mediated Drug-Drug Interaction Studies Guidance for Industry.*; 2017. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/vitro-

metabolism-and-transporter-mediated-drug-drug-interaction-studies-guidance-industry.

24.   Xian X, Liao L, Shu W, et al. A Novel Mutation of SLC19A2 in a Chinese Zhuang Ethnic
      Family with Thiamine-Responsive Megaloblastic Anemia. *Cell Physiol Biochem*.
      2018;47(5):1989-1997. doi:10.1159/000491467

25.   Subramanian VS, Subramanya SB, Said HM. Relative contribution of THTR-1 and
      THTR-2 in thiamin uptake by pancreatic acinar cells: studies utilizing Slc19a2 and
      Slc19a3 knockout mouse models. *Am J Physiol Gastrointest Liver Physiol*.
      2012;302(5):G572-G578. doi:10.1152/ajpgi.00484.2011

26.   Khuri N, Zur AA, Wittwer MB, et al. Computational Discovery and Experimental
      Validation of Inhibitors of the Human Intestinal Transporter OATP2B1. *J Chem Inf
      Model*. 2017;57(6):1402-1413. doi:10.1021/acs.jcim.6b00720

27.   Wittwer MB, Zur AA, Khuri N, et al. Discovery of potent, selective multidrug and toxin
      extrusion transporter 1 (MATE1, SLC47A1) inhibitors through prescription drug profiling
      and computational modeling. *J Med Chem*. 2013;56(3):781-795. doi:10.1021/jm301302s

28.   Kido Y, Matsson P, Giacomini KM. Profiling of a prescription drug library for potential
      renal drug-drug interactions mediated by the organic cation transporter 2. *J Med Chem*.
      2011;54(13):4548-4558. doi:10.1021/jm2001629

29.   Chen EC, Khuri N, Liang X, et al. Discovery of Competitive and Noncompetitive Ligands
      of the Organic Cation Transporter 1 (OCT1; SLC22A1). *J Med Chem*. 2017;60(7):2685-
      2696. doi:10.1021/acs.jmedchem.6b01317

30.   Vora B, Green EAE, Khuri N, Ballgren F, Sirota M, Giacomini KM. Drug-nutrient
      interactions: discovering prescription drug inhibitors of the thiamine transporter ThTR-2
      (SLC19A3). *Am J Clin Nutr*. November 2019. doi:10.1093/ajcn/nqz255

31. Zhang Q, Zhang Y, Diamond S, et al. The Janus kinase 2 inhibitor fedratinib inhibits thiamine uptake: a putative mechanism for the onset of Wernicke's encephalopathy. *Drug Metab Dispos*. 2014;42(10):1656-1662. doi:10.1124/dmd.114.058883

32. Schulz M, Iwersen-Bergmann S, Andresen H, Schmoldt A. Therapeutic and toxic blood concentrations of nearly 1,000 drugs and other xenobiotics. *Crit Care*. 2012;16(4):R136-R136. doi:10.1186/cc11441

33. Andersson T, Cederberg C, Regardh CG, Skanberg I. Pharmacokinetics of various single intravenous and oral doses of omeprazole. *Eur J Clin Pharmacol*. 1990;39(2):195-197.

34. Mikulic M. Number of omeprazole prescriptions in the U.S. from 2004 to 2016. https://www.statista.com/statistics/780397/omeprazole-prescriptions-number-in-the-us/. Published 2018.

35. Bellou A, Aimone-Gastin I, De Korwin JD, et al. Cobalamin deficiency with megaloblastic anaemia in one patient under long-term omeprazole therapy. *J Intern Med*. 1996;240(3):161-164. doi:10.1046/j.1365-2796.1996.20846000.x

36. Analytics TH. Omeprazole. Drug Monograph [Micromedex]. Micromedex. www.micromedexsolutions.com. Published 2019. Accessed January 10, 2019.

Chapter 5: Conclusions and Perspective

This dissertation underscores the importance of mechanistic studies to understand the biological mechanisms underlying human disease, variation in drug response and drug-induced toxicities.[1,2] A large part of the research presented here centered on studies designed to discover rare genetic variants in *SLC2A2* that underlie inter-individual variation in metformin response (Chapter 2) and the characterization of rare variants in *SLC2A2* that have been associated with FBS (Chapter 3). Further, it examined the role of *SLC19A2* in drug-induced megaloblastic anemia. Below I summarize each of the chapters and provide some insights into limitations and future studies.

In Chapter 2, we used next-generation sequencing to identify genetic variants in *SLC2A2* that associate with extreme non-response. The study was based on the fact that common variants in *SLC2A2* associate with variation in response to the anti-diabetic drug metformin; therefore rare variants in the gene may have larger effects and associate with extreme non-response to the drug. One of the major issues in pharmacogenetic studies, in general, is power.[4] Pharmacogenetic GWAS usually contain smaller cohorts compared with disease GWAS due to problems in obtaining well-defined, drug response phenotypes and prescription information, which is often not available in electronic medical records. For pharmacogenomic sequencing association studies, power is not only affected by the lack of prescription data and a well-defined phenotype, but by the type and cost of sequencing that can be performed based on the available budget. Additionally, the total sample size of the cohort not only affects the power to find a statistical association but the power to detect rare or less common variants.

In Chapter 2, one of the greatest difficulties was choosing the type of next-generation sequencing method to select (i.e. whole genome, whole exome, target region) due to budget constraints. Whole genome or whole exome method would provide the data needed to conduct a truly

genome-wide evaluation of the impact of all genes and their rare/less common variants on metformin response. However, whole exome sequencing, which is cheaper than whole genome, on average cost a little under $1000 per sample;[5] with approximately one-thousand samples, the average cost of whole exome sequencing the entire METRO cohort would have been about $1,000,000 and above our budget. Therefore, we chose to move forward with our full sample size while reducing the size of the region to be sequenced. We sequenced extreme responders and non-responders in the METRO cohort using targeting sequencing of *SLC2A2*. Despite losing genome-wide breadth, the large sample size sequenced will increase the number of **rare** and **less common** variants we can detect in *SLC2A2* and therefore statistically analyze.

In the study conducted in Chapter 3, which focused on mutations in *SLC2A2* that are causal for FBS, we came to two major conclusions. First, p.Leu153_Ile154del mutant GLUT2 had residual transport activity, which may explain the mild phenotype of FBS exhibited in the two patients who are compound heterozygous for p.Leu153_Ile154del (which has residual transport activity, about 5% of wild-type activity) and p.Pro417Arg (which has no transport activity). The significant reduction in disease severity despite a small increase in GLUT2 transport hints at the therapeutic possibilities of "slight" or "residual" GLUT2 activation, and also to a "functional threshold" required to initiate disease. Deleterious variants in different SLC transporters cause a range of autosomal, recessive diseases (i.e. *SLC2A2*/Fanconi-Bickel syndrome, *SLC22A5*/Carnitine Transporter Deficiency (CTD) syndrome, *SLC19A2*/Thiamine-responsive megaloblastic anemia (TRMA)). All *SLC2A2* FBS variants functionally characterized thus far completely abolish GLUT2 transport except for one (p.Leu153_Ile154del).[3] But several *SLC22A5* missense variants causal for CTD still retain residual or partial transport activity,

despite the presence of clinically significant carnitine deficiency.[6] Therefore, the efficacy of treatments for these diseases will likely vary based whether the treatment can increase transport activity of the substrate(s) above the "functional threshold".

Our second conclusion was that 45% of GLUT2 FBS mutant transporters were not expressed or negligibly expressed on the plasma membrane. This conclusion demonstrates that significant portion of rare, deleterious variants in genes that encode for membrane transporters cause disease by preventing the trafficking to and/or expression of that membrane transporter on the plasma membrane. Other GLUT transporters (e.g. GLUT9) have sequence motifs that target their expression to specific parts of the plasma membrane.[7] Disruption of those motif sequences affects membrane trafficking. However, FBS variants not expressed on the plasma membrane are not clustered in one specific part of *SLC2A2* but are located throughout its sequence.[3] A few of those FBS variants exhibit some partial or "leaky" expression on the membrane, but most show no expression at all. Creation of a detailed map of GLUT2 folding and trafficking from nascent single-strand protein to folded, membrane-embedded transporter would help illuminate the mechanisms by which these variants prevent or hinder plasma membrane expression. Furthermore, these results suggest that either pharmacochaperones and/or allosteric activators (depending on the mechanism underlying the functional loss of the transporter) that stimulate the transport activity of FBS mutant transporters, even with a slight gain in function (as noted for the individuals harboring the p.Leu153_Ile154del of GLUT2) may provide therapeutic benefit.

Unfortunately, this dissertation also reveals that popular functional protein prediction algorithms (e.g. POLYPHEN, SIFT) have reduced accuracy when predicting the function of mutant

membrane transporters. Data from these protein prediction algorithms are often used, instead of functional characterization, to stratify or weigh variants in a statistical association analysis, especially for rare variant association analyses. Variants predicted to be destructive or dysfunctional to protein function are given higher weight in the association analyses. Similarly, after association analyses are conducted, protein prediction algorithms are used to triage associated variants and determine which variants alter protein function and are therefore likely to be a causal variant.[8] Out of twenty-one *SLC2A2* variants (15 FBS, six less common) functionally characterized, 5 were incorrectly predicted by POLYPHEN (24%). Three out of three *SLC19A2* variants functionally characterized were correctly predicted by POLYPHEN.

Finally, the latter portion of this dissertation evaluated the potential of drug-vitamin interactions to phenocopy signs and symptoms of corresponding genetic diseases. The study of drug-vitamin interactions is a relatively small field compared to drug-drug interaction studies. But the delayed FDA-approval of Janus kinase 2 inhibitor, fedratinib, due to the development of Wernicke's encephalopathy, highlighted the significant clinical toxicities that can occur from drug-vitamin interactions.[9] The study in Chapter 4 identified at least one prescription drug that may potentially inhibit THTR-1 transporters at physiological concentrations (erythromycin). Reduced blood levels of thiamine pyrophosphate were identified in patients prescribed either erythromycin and omeprazole, using real-world data from the electronic health record. Collectively, the studies suggested a novel mechanism of drug-induced megaloblastic anemia that resembles thiamine-responsive megaloblastic anemia (TRMA), a Mendelian disease caused by inactivation of THTR-1 transporters. Prospective, controlled in-vivo human trials are needed to confirm the

effect of erythromycin and omeprazole on thiamine pyrophosphate blood levels and risk of megaloblastic anemia.

In summary, this dissertation conducts an in-depth investigation of *SLC2A2* variation and its influence on inter-individual differences in metformin response and FBS disease severity. It also suggests the potential of prescription medications to cause adverse reactions through chronic THTR-1 (thiamine transporter) inhibition. Notably, we identified the first FBS-associated mutation that still retains residual transport activity (p.Leu153_Ile154del) in patients that exhibit a mild form of FBS.[3] And we identified two novel inhibitors of THTR-1 (omeprazole, erythromycin) that associate with reduced levels of thiamine pyrophosphate. These findings expand our knowledge regarding the genotype-to-phenotype relationship of FBS and highlight the importance of functional characterization and annotation of membrane transporters. While additional replication cohorts of metformin response would help to increase power and confirm rare-variant associations, in-depth mechanistic studies investigating the relationship between *SLC2A2* and metformin response will offer further biological insights into the mechanisms of metformin action and inform its expanded use in diverse patient populations.

REFERENCES

1. Zhou K, Yee SW, MetGen coauthors, Giacomini K PE. *Variation in the Glucose Transporter Gene GLUT2 (SLC2A2) Is Associated with Glycaeic Response to Metformin: A MetGen Study*.; 2016.

2. Grünert SC, Schwab KO, Pohl M, Sass JO, Santer R. Fanconi-Bickel syndrome: GLUT2 mutations associated with a mild phenotype. *Mol Genet Metab*. 2012;105(3):433-437. doi:10.1016/j.ymgme.2011.11.200

3. Enogieru OJ, Ung PMU, Yee SW, Schlessinger A, Giacomini KM. Functional and structural analysis of rare SLC2A2 variants associated with Fanconi-Bickel syndrome and metabolic traits. *Hum Mutat*. 2019;40(7):983-995. doi:10.1002/humu.23758

4. Holmes M V, Shah T, Vickery C, Smeeth L, Hingorani AD, Casas JP. Fulfilling the promise of personalized medicine? Systematic review and field synopsis of pharmacogenetic studies. *PLoS One*. 2009;4(12):e7960. doi:10.1371/journal.pone.0007960

5. PHG Foundation. *The Budget Impact and Cost-Effectiveness of Introducing Whole-Exome Sequencing-Based Virtual Gene Panel Tests into Routine Clinical Genetics*. Cambridge, UK; 2017.

6. Lahjouji K, Mitchell GA, Qureshi IA. Carnitine transport by organic cation transporters and systemic carnitine deficiency. *Mol Genet Metab*. 2001;73(4):287-297. doi:10.1006/mgme.2001.3207

7. Bibee KP, Augustin R, Gazit V, Moley KH. The apical sorting signal for human GLUT9b resides in the N-terminus. *Mol Cell Biochem*. 2013;376(1-2):163-173.

doi:10.1007/s11010-013-1564-3

8.   Ramsey LB, Bruun GH, Yang W, et al. Rare versus common variants in

pharmacogenetics: SLCO1B1 variation and methotrexate disposition. *Genome Res*.

2012;22(1):1-8. doi:10.1101/gr.129668.111

9.   Giacomini MM, Hao J, Liang X, et al. Interaction of 2,4-Diaminopyrimidine-Containing

Drugs Including Fedratinib and Trimethoprim with Thiamine Transporters. *Drug Metab

Dispos*. 2017;45(1):76-85. doi:10.1124/dmd.116.073338

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Osatohanmwen Enogieru

7B53811AB30740E...   Author Signature

3/17/2020

Date