2022

Mode of substrate binding and specificity for ketohexokinase across isozymes implies an induced-fit mechanism

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Thesis

MODE OF SUBSTRATE BINDING AND SPECIFICITY FOR KETOHEXOKINASE ACROSS ISOZYMES IMPLIES AN INDUCED-FIT MECHANISM

by

SO YOUNG BAE

B.A., Boston University, 2018

Submitted in partial fulfillment of the

requirements for the degree of

Master of Arts

2022

Approved by

First Reader _____

Dean R. Tolan, Ph.D.
Professor of Biology

Second Reader _____

Karen N. Allen, Ph.D.
Professor and Chair of Chemistry

Third Reader _____

Blair Szymczyna, Ph.D.
Lecturer in Chemistry

Fourth Reader _____

Jeffrey W. Bacon, Ph.D.
Specialist in Chemistry

ACKNOWLEDGMENTS

Firstly, I am thankful for my family for years of financial and moral support that allowed me to finish a degree program. I am extremely grateful for the boundless help and provisioning of Dr. Dean R. Tolan. I thank him for steering me to the foundational and historical knowledge in biochemistry and molecular biology and clarifying my view on science with the skill of synthesis. I am grateful to have been introduced to the discipline of structural biology under his guidance. I also thank Dr. Karen N. Allen for teaching me crystallography and bolstering my developing insights in the realm of structural biology. I wish to thank the other members of my committee Drs. Jeffrey W. Bacon and Blair Szymczyna for reading this thesis. I am grateful for Ashley Marie-Alice Rebelo and Dr. William Clarke Gasper who have helped me to learn the more technical aspects of crystallography and data processing, and in particular for the previous work of Clarke, and his patience in training me that I was able to ask an independent question with the right skill set. In addition, I am thankful for Dr. Jose Luis Medrano Jr. and Dr. Dean R. Tolan who had provided ample feedback on the literature review I performed on the paper describing the ribokinase structure, which lead me to the hypothesis that is the subject of this thesis. I am extremely grateful for all the Tolan lab members, past and current, for their fervor and inquisitive positivity that it allowed me to grow in doing science. I am also grateful for the members of the Allen lab for all their support and welcoming me into the environment in which I could grow. Lastly, I am also extremely grateful for the exposure

and participation as a full-fledged team member of the Centennial Therapeutics Inc. effort in the pursuit of drug discovery.

MODE OF SUBSTRATE BINDING AND SPECIFICITY FOR KETOHEXOKINASE ACROSS ISOZYMES IMPLIES AN INDUCED-FIT MECHANISM

SO YOUNG BAE

ABSTRACT

Ketohexokinase (KHK), in an adenosine triphosphate (ATP) dependent reaction, catalyzes the first reaction in fructose metabolism, which converts the furanose form of D-fructose into fructose-1-phosphate. This enzyme has become a target for pharmacological development against fatty liver and metabolic syndrome. KHK exists in two isoforms, A and C, which differs by alternative splicing of exon 3 which encodes 45 out of 298 amino acids. Normally KHK exists as a homodimer and is comprised of an alpha/beta domain interlocking with a β-clasp domain. For KHK-C, there appear to be at least two conformations of the β-clasp domain. Previous work on KHK-A reveals it does not adopt the same conformations. A structure of the mouse KHK-A in its unliganded form is solved and shows that these two conformations also exist for KHK-A. Furthermore, this property is conserved across species. While crystals of human KHK-A in its unliganded form were grown, a structure was not achieved. However, unpublished structures of human KHK-A in its unliganded form also shows different conformations in β-clasp domain when in juxtaposition with the same enzyme complex with

ligands. Defining the role of conformational changes in KHK-A is important, because this isozyme has been reported to have a role in cancer metastasis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMATICS

LIST OF ABBREVIATIONS

AldoB ..... aldolase B

AMP-PNP ..... Adenylyl-imidodiphosphate

AMPD ..... AMP deaminase-2

Unliganded ..... structure of enzyme without any substrates bound

AR ..... aldehyde reductase

BSA ..... Bovine Serum Albumin

BU-CIC ..... Boston University Chemical Instrument Center

CD ..... Circular Dichroidism

CCD ..... Charge Coupled Device

dH_2O ..... deionized water

DNAse ..... deoxyribonuclease

DHAP ..... dihydroxy acetone phosphate

DTT ..... dithiothreitol

DYT ..... Double yeast tryptone

EtOH ..... Ethanol

Fru-1-P ..... Charge Coupled Device

GK ..... glucokinase

HEPES ..... (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HFI ..... Hereditary Fructose Intolerance

IPTG ..... Isopropyl β-D-1-thiogalactopyranoside

k_{cat} ..... turnover number

k_D ..... kilodaltons

KHK ..... Ketohexokinase

K_m ..... Michaelis constant

LDH ..... Lactate Dehydrogenase

MOPS ..... (3-(N-morpholino)propanesulfonic acid)

MW ..... Molecular Weight

MWCO ..... Molecular Weight Cut Off

NAFLD ..... Non-alcoholic fatty liver disease

NMR ..... Nuclear Magnetic Resonance

OD ..... optical density

PAGE ..... polyacrylamide gel electrophoresis

PDB ..... Protein Database

PEP ..... Phosphoenol-pyruvate

pfkB ..... phosphofructo kinase B

PK ..... Pyruvate Kinase

RNAse ..... ribonuclease

SPR ..... Surface Plasmon Resonance

SSRL ..... Stanford Synchrotron Radiation Lightsource

T2D ..... Type-2 diabetes

TAE ..... Tris-acetate-EDTA

TEA ..... Triethylamine

TEMED..... Tetramethylethylenediamine

Ternary..... structure of enzyme with both ATP, or an analog, and fructose bound

TK ..... triosekinase

WT .....wildtype, i.e. native

YT .....yeast tryptone

CHAPTER 1: INTRODUCTION

Roles for KHK in Sugar Metabolism

Human consumption of dietary sugars has experienced a spike in recent years, and is now a hallmark of what is called the western diet. In the past decades, added sugar accounts for nearly 15% of the US diet (George *et al.*, 2005). A variety of health problems have been blamed on this western diet, including type-2 diabetes (T2D), obesity, hypertension, heart disease, and non-alcoholic fatty liver disease (NAFLD) (Abdelmalek *et al.*, 2012; Ouyang *et al.*, 2008; Sheiham *et al.*, 2015). This places a significant financial burden on the US healthcare system with costs projected to exceed $100 billion by 2030 (Tremmel *et al.*, 2017). Sugar consumption is on the rise in other countries including most of large parts of India and Asia. While there are recommendations to just change the diet, these efforts have largely been unsuccessful and it lends credence to theories of sugar addiction (Johnson *et al.*, 2011).

As the first step of fructose metabolism, ketohexokinase (KHK) catalyzes, with ATP, the conversion of the β-anomer of D-fructose into fructose 1-phosphate (Fru1P) using ATP. The enzyme initiates the intracellular catabolism of sugar molecules and regulates glucose metabolism in the liver (Jensen *et al.*, 2018). Parallel to the fructose metabolism pathway, the depletion of the ATP levels transiently occurs due to KHK catalysis and its subsequent activation of AMP deaminase, resulting in the build-up of the uric acid

(Johnson *et al.*, 2013; Johnson *et al.*, 2009; Stirpe *et al.*, 1970). It is thought that these increases in uric acid lead to potential mitochondrial dysfunction (Johnson *et al.*, 2013; Lanaspa *et al.*, 2012; Sanchez-Lozada *et al.*, 2012). Figure 1 shows these pathways with KHK (denoted as FK) utilizing ATP and the parallel degradation of ADP to uric acid.

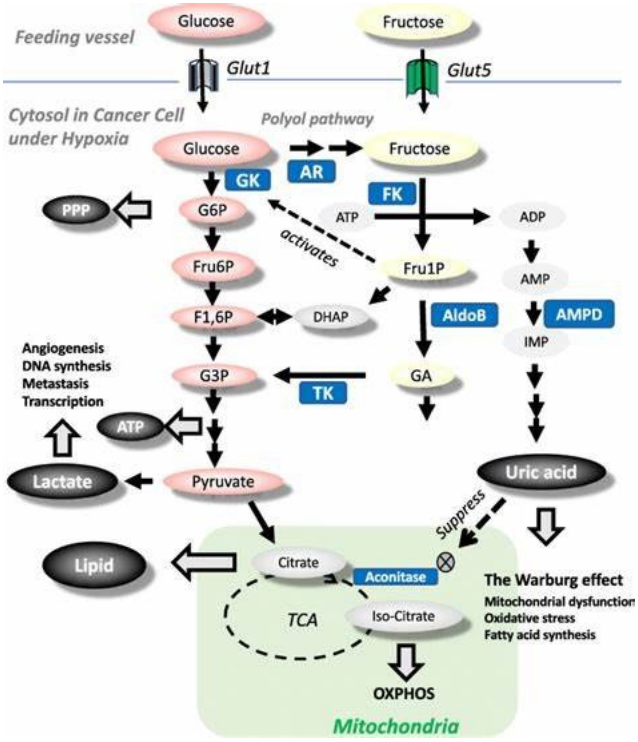


Figure 1. Fructose metabolism in the liver (from (Nakagawa *et al.*, 2020)). Fructose and intermediates in catabolism are in yellow ovals, glucose and intermediates are in pink ovals, adenine-nucleotides are in light gray ovals, with end products in dark gray ovals. Enzymes in blue boxes; KHK (FK), aldehyde reductase (AR), glucokinase (GK), aldolase B (AldoB), triosekinase (TK), AMP deaminase-2 (AMPD).

KHK is a critical entry point for fructose in these pathways, which were thought to operate mostly in the liver and kidney. This was consistent with the expression of a liver and

kidney specific form of the enzyme, KHK-C (Bonthron *et al.*, 1994; Diggle *et al.*, 2009; Hayward *et al.*, 1998).

KHK Isozymes

Soon after cloning the KHK gene, it became clear that exon 3 is alternatively spliced and results in the formation of two different enzymes, KHK-A and KHK-C. The isozymes differ by only 45 amino acids (135 nucleotides in exon 3) as a result of alternative mRNA splicing (see Appendix)(Hayward *et al.*, 1998). Using a probe for each isozyme (Hayward *et al.*, 1998), it became apparent that one isozyme, KHK-C was expressed in a tissue specific way, appearing in liver, kidney, and small intestine, while KHK-A was expressed at low levels in most other tissues (Diggle *et al.*, 2009). Originally it was thought that KHK-A might not contribute to fructose metabolism and may be a pseudogene due to its relatively low activity in these tissues (Asipu *et al.*, 2003; Funari *et al.*, 2005; Weiser *et al.*, 1975). This low activity was eventually determined as a K_m effect, where KHK-C has a 10-fold lower K_m value than KHK-A (Asipu *et al.*, 2003). The kinetic parameters of the isozymes from human are shown in Table 1. The affinity and capacity of KHK-C for

Table 1. Kinetic Parameters of KHK across Isozymes and Species[1]

	hKHK-A WT		hKHK-C WT		mKHK-C WT	
	fructose	ATP	fructose	ATP	fructose	ATP
KM (mM)	7	0.36	0.8	0.15	1.7	0.27
kcat (1/sec)	6.9		7.6		7.1	

[1]Data from (Asipu *et al.*, 2003; Gasper, 2020)

fructose is much greater than KHK-A due to a much lower K_m value. Others have reported even larger differences between these values (0.73 mM for KHK-C vs 28.6 mM for KHK-A (Futatsugi *et al.*, 2020; Huard *et al.*, 2017)). The low level of expression and a K_m value higher than the physiological concentrations of fructose (Asipu *et al.*, 2003), generated the thought that KHK-A was of little physiological consequence. However, the production of two knockout strains for KHK, one with both isozymes missing and one with only KHK-A missing (Diggle *et al.*, 2009), showed the valuable role for KHK-A in mice. Mice fed a high fructose diet for several months showed signs of metabolic syndrome, whereas the complete KHK knockout mouse (missing both KHK-A and KHK-C) was resistant to these pathologies (Ishimoto *et al.*, 2012). The interesting effect was that the KHK-A-only knockout mice showed an even greater pathology. This result revealed that KHK-A plays a buffering role for fructose on a chronically high fructose diet. The KHK-A-only knockout mice forced all the ingested fructose to the liver accentuating the pathologies originating there.

Recent literature reports that repressing KHK-C activity may serve as a therapeutic target to ameliorate Hereditary Fructose Intolerance (HFI) disease or fructose induced metabolic syndromes (Lanaspa *et al.*, 2018). When the KHK-C isozyme in the liver is inhibited, less Fru1P is phosphorylated and cleaved by the downstream enzyme aldolase B into dihydroxyacetone phosphate (DHAP) and glyceraldehyde. This alleviates the significant adenosine triphosphate (ATP) depletion and consequential uric acid build-up (see Figure 1) aiding patients with inborn fructose intolerance disease. Once less phosphate is sequestered in Fru1P, glycogen levels increase due to the now-available substrate for

glycogen phosphorylase b and the alleviation of the competitive inhibition of phosphoglucomutase by Fru1P (Cain *et al.*, 1971; Thurston *et al.*, 1974).

Realizing that KHK-C might be a valuable drug target, major pharmaceutical companies and teams of scientists had been pursuing a potent inhibitor molecule for KHK. If inhibited sufficiently, uric acid is not accumulated, and mitochondria would not resort to fat synthesis (Gibbs *et al.*, 2010; Gutierrez *et al.*, 2021; Huard *et al.*, 2017; Lanaspa *et al.*, 2018; Sanchez-Lozada *et al.*, 2010) (see Figure 1). For these reasons, this enzyme has become a target for pharmacological development against fatty liver and metabolic syndrome.

Structural Relationships of KHK Isozymes

As mentioned above, KHK exists in two isoforms, A and C, which differ by alternative splicing of exon 3 which encodes 45 out of 298 amino acids. Normally KHK exists as a homodimer and is comprised of α and β domain interlocking at the β-clasp domain (Figure 2) (Trinh *et al.*, 2009).

There is a total of 26 structures now deposited in the Research Collaboratory for Structural Bioinformatic PDB (rcsb.org) (Berman *et al.*, 2000). These are listed in Table 2. Most of these structures are from human in complex with inhibitors and determined by groups interested in drug development. However, some of the first structural models were the human KHK-A and C isozymes with either substrate memetics bound or are so-called "apo" structures without any ligands bound. The structure of wild-type unliganded hKHK-A (2HQQ), wild-type unliganded hKHK-C (3B3L), and wild-type hKHK-A co-complexed

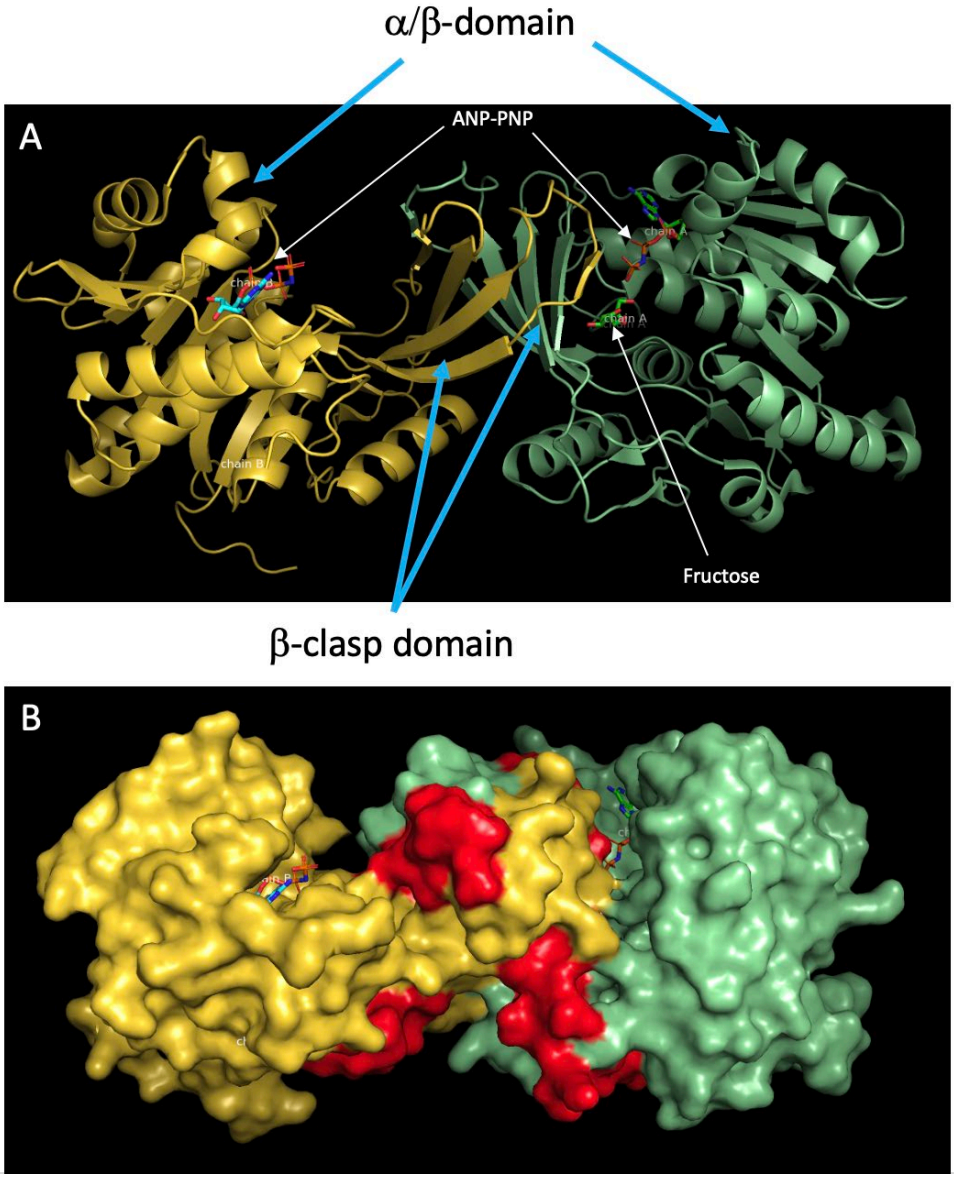


Figure 2. Overall Structure of KHK. Panel A: The dimeric human KHK-C is depicted in ribbon structure with its interlocking subunits, which is complexed with AMP-PNP and fructose (indicated with white arrows) at the active site shown as sticks. Chain A and B from PDBID:3NBV (Gibbs *et al.*, 2010) are shown in green and yellow, respectively. The domains are identified with light-blue arrows. Panel B: the same depiction as in panel A, but depicted as space-filling model, except the region shown in red, which depicts the region of the protein encoded by alternative mRNA splicing of exon 3.

Table 2. Summary of All the KHK Structures Deposited in the PDB.

No.	PDB Deposit / Publication Year	PDB ID	Reference	Species	Isozyme	Ligand	Resolution	Copies in ASU	Rfree (%)	Space Group	Well Condition
1	2006/ -	2HLZ	Rabeh	human	A	Apo	1.85 Å	4	24.9	P 1	17% PEG8000, 5mM Cobalt Chloride, 200mM ammonium sulfate, 100mM NaCitrate, pH 4.2, vapor diffusion, sitting drop, temperature 291K
2	2007/2009	2HQQ	Trinh	human	A	Apo	1.86 Å	1	23.4	I 2 2 2	0.7M ammonium sulphate, 0.5M lithium sulphate, 0.1M sodium citrate, pH 5.7, VAPOR DIFFUSION, HANGING DROP, temperature 291K
3	2007/2009	2HW1	Trinh	human	A	AMP-PNP + Fru	2.10 Å	1	22.8	I 2 2 2	0.7M ammonium sulphate, 0.5M lithium sulphate, 0.1M sodium citrate, pH 5.7, VAPOR DIFFUSION, HANGING DROP, temperature 291K
4	2008/2009	3B3L	Trinh	human	C	Apo	2.90 Å	4	28.1	P 21 21 21	20% PEG 3000, 0.1M sodium acetate, pH 4.2, VAPOR DIFFUSION, HANGING DROP, temperature 291K
5	2010/2010	3NBV	Gibbs	human	C	AMP-PNP + Fru	2.30 Å	2	23.6	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
6	2020/ -	6P2D	Gasper	mouse	C	ADP + Fru	1.79 Å	1	22	C 2 2 21	1.2 M Ammonium Citrate Tribasic, 20% glycerol, 1.3 M KNO3, 100 mM MgCl2, 220 mM fructose, and 52 mM ADP
7	2010/2010	3NBW	Gibbs	human	C	pyrazole	2.34 Å	2	27.4	P 21 21 21	17% PEg 8k, 0.1M Na-Citrate, 0.2M Ammonium Sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
8	2010/2010	3NC9	Gibbs	human	C	Indazole-6	2.40 Å	2	28.5	P 21 21 21	17% PEG8k, 0.1M Na-Citrate, 0.2M ammonium sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
9	2010/2010	3NC2	Gibbs	human	C	pyrimdobenzene	2.50 Å	2	27.1	P 21 21 21	17% PEg 8k, 0.1M Na-Citrate, 0.2M Ammonium Sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
10	2010/2010	3NCA	Gibbs	human	C	thieno pyridinol	2.60 Å	2	23.9	P 21 21 21	17% PEG 8K, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
11	2011/2011	3RO4	Abad	human	C	Indazole-9	2.60 Å	2	29.7	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
12	2012/2011	3Q92	Abad	human	C	pyrimidinopyrimidine-8	2.80 Å	2	26.3	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium Sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
13	2012/2011	3QA2	Abad	human	C	pyrimidinopyrimidine-8	2.52 Å	2	23.7	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium Sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
14	2012/2011	3QAI	Abad	human	C	pyrimidinopyrimidine-8	2.70 Å	2	29.8	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium Sulfate, pH 4.5, VAPOR DIFFUSION, SITTING DROP, temperature 295K
15	2017/2017	5WBQ	Pandit	human	C	hydroxy Me pyrrolidine + pyrrolopyrimidine	2.40 Å	2	24.8	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5
16	2017/2017	5WBM	Pandit	human	C	piperidin methanol + pyrrolopyrimidine	2.16 Å	2	22.7	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5
17	2017/2017	5WBP	Pandit	human	C	Quinoxaline	2.74 Å	2	22.9	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5
18	2017/2017	5WBR	Pandit	human	C	2 piperidine + pyridiine	2.58 Å	2	24.1	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5
19	2017/2017	5WBZ	Pandit	human	C	2 pyrrolidines + pyridine	2.40 Å	2	21.4	P 21 21 21	17% PEG 8k, 0.1M Na-Citrate, 0.2M Ammonium sulfate, pH 4.5
20	2020/2020	6W0N	Jasti	human	C	2 pyrrolidines + pyridine	2.41 Å	2	20.4	P 21 21 21	17% PEG 8000, 100 mM Na citrate, pH 4.5, and 0.2 M ammonium sulfate
21	2020/2020	6W0W	Jasti	human	C	pyrrolidinopyridine + Me-hydroxyl-Az	2.80 Å	2	21.6	P 21 21 21	17-20% PEG 8000, 100 mM sodium citrate, pH 4.4-4.7, and 0.2 M ammonium sulfate
22	2020/2020	6W0X	Jasti	human	C	azabicyclohexanyl trifluromethyl carbonitrile pyrimidine	2.38 Å	2	20.6	P 21 21 21	17-20% PEG8000, 100 mM sodium citrate, pH 4.4-4.7, 0.2 M ammonium sulfate
23	2020/2020	6W0Y	Jasti	human	C	ethanoic acid + W0Y	20.5 Å	2	20.5	P 21 21 21	17-20% PEG8000, 100 mM sodium citrate, pH 4.4-4.7, 0.2 M ammonium sulfate
24	2020/2020	6W0Z	Jasti	human	C	azabicyclohexanyl trifluromethyl MeAz pyridimidine	2.30 Å	2	19.5	P 21 21 21	17-20% PEG8000, 100 mM sodium citrate, pH 4.4-4.7, 0.2 M ammonium sulfate
25	2021/ -	6UL7	Gasper	human	C	Osthole + Fru	2.30 Å	1	27.8	P 43 21 2	0.1 M Bis-Tris, pH 5.5, 2 M ammonium sulfate, 1.3 M potassium nitrate, 100 mM magnesium chloride, 220 mM fructose, 3.8 mM osthole

with AMP-PNP (an ATP mimetic), and fructose (2HW1), which is called a "ternary" complex, were determined in 2007 (Trinh *et al.*, 2009). Gibbs *et al*. (2010) reported a ternary complex of hKHK-C with a proposed inhibitor. The highly conserved structure and fold is depicted in Figure 2.

The KHK isozymes are in the protein family called the pfkB family of carbohydrate kinases (Pfam, PF00294) (Sigrell *et al.*, 1998; Titgemeyer *et al.*, 1994). Phospho fructokinase B family of carbohydrate kinases phosphorylate the hydroxymethyl groups of different sugar moieties, like two other families of sugar kinases, hexokinase and galactokinase. (Park & Gupta 2008). The enzymes in this family includes human homologs like ribose kinase and adenosine kinase, which along with KHK are the only members expressed in humans, and includes a pseudouridine kinase (Kim *et al.*, 2021). The family is often referred to as the ribokinase family as well (Park *et al.*, 2008) and is distinct from the more well-known glycolytic kinases. All members transfer a phosphate to an alcohol from ATP. The common structural features include a central eight-stranded alpha/beta sandwich with an eight-stranded sheet flanked by five and three conserved helices. There is also often a beta-clasp domain.

Roles for KHK-A

While KHK-C is clearly responsible for the catabolism of fructose, the role of KHK-A is becoming more intriguing. First, different pathologies cause a switch in expression from A to C via alternative splicing (Mirtschink *et al.*, 2015). Second, it has been shown to not only be a *substrate* for a protein kinase, but to act as a protein kinase in different conditions

(Xu *et al.*, 2019). This protein kinase activity has been proposed to explain the long-known association of fructose and breast-cancer metastasis (Kim *et al.*, 2020; Zamora-Leon *et al.*, 1996). Although more structures of the liver KHK-C isozyme have been determined, likely due to its identification as a pharmacological target, the structure and functions of KHK-A could be equally revealing and important. One structure/function relationship is for changes in conformation of the β-clasp domain of KHK. It appears that this domain from one subunit moves a critical Asp-27 into the active site of the other subunit of the dimer, a change that is correlated with fructose binding to the active site (Gasper, 2020). A similar conformational change in this domain has been studied extensively for ribokinase using X-ray crystallography (Sigrell *et al.*, 1997, 1998; Sigrell *et al.*, 1999), and led the authors to propose an induced-fit model for substrate binding. There are much fewer reports of KHK-A structures, and among this previous work are reports that KHK-A does not undergo the same conformational changes as KHK-C (Trinh *et al.*, 2009).

Aims of MA project

Previous work had shown that there is no conformational change observed for KHK-A when solved with and without (unliganded) substrates bound (Trinh *et al.*, 2009). Here it is shown that alternate conformations do indeed exist for KHK-A. Furthermore, this appears that these changes are conserved across species. Defining the role of conformational changes in KHK-A is important, because of the various roles reported for KHK-A in cellular pathology such as in cancer metastasis (Kim *et al.*, 2020).

CHAPTER 2: MATERIALS AND METHOD

Construction of mouse KHK-A expression plasmid

A plasmid for expression of mouse KHK-A was commercially synthesized at Genewiz, Inc. (Order Number: 30-235096037). The codons, 906 base-pairs in length, were optimized for expression in *Escherichia coli,* and inserted in the *Nde* I and *Hin*d III sites of pET28a. The reference sequences of the mRNA for mouse KHK-A and the alignment of the codon-optimized sequence with the reference sequence are shown in the Appendix. There were 223 base changes. Plasmid DNA was attained and dissolved in TE (10 mM Tris-HCl, pH 8, 1 mM EDTA) to ~10 ng/μL. DNA was used for transformation into the BL21(DE3) strain for expression and the DH5α strain for production of plasmid DNA for verification by gel electrophoresis and DNA sequence determination. The human KHK-A expression plasmid, also constructed in pET28a, was a gift from C. Gasper [cite thesis].The alignment of human and mouse KHK-A proteins (89% identical), as well as the KHK-C isozymes from both species, is shown in the Appendix.

Bacterial Transformation

Commercially purchased competent *E. coli* cells, BL21(DE3) or DH5α, were used to insert expression plasmids by transformation as recommended by the manufacturer (New England Biolabs). Briefly, cells (50 μL aliquots per reaction) were incubated with 1-10 ng or plasmid DNA on ice for 40 min, then quickly heat shocked for 45 s at 42 °C.

Reactions were diluted with 2 mL of DYT media (1.6% (w/v) bactotryptone, 1% (w/v) yeast extract, and 0.5% (w/v) NaCl), and allowed to recover by incubation at 37 °C with shaking at 200 rpm for 30 min. Various volumes of each reaction were spread on agar petri dishes containing YT-agar media (0.8% (w/v) bactotryptone, 0.5% (w/v) yeast extract, and 0.5% (w/v) NaCl, plus 0.5% (w/v) bactoagar), and containing 50 μg/mL of kanamycin. Plates were dried beforehand by placing lids of plates slightly opened for 30 min under a Bunsen burner. Transformations include control plates of reactions without any DNA and using 1 ng of pET28a supercoiled DNA. Plates were placed in a 37 °C incubator for 16-20 h.

Plasmid Preparation

Individual colonies of expression plasmids transformed into *E. coli* strain DH5α were picked from DYT plates and used to inoculate 2 L of DYT containing 50 μg/mL of kanamycin. The cultures were incubated at 37 °C with shaking at >200 rpm for 16-20 h. Cells were pelleted by centrifugation at 11,000*g* for 20 s. Plasmids were purified from cell pellets using a Promega Wizard SV Plus Miniprep Kit as instructed by the manufacturer. The procedure resulted in recovery of 1 μg DNA per 2 mL culture. Purity and concentration of DNA was determined by agarose gel electrophoresis.

Preparation of Permanent Stabs

Individual colonies from DYT plates were picked and used to inoculate 5 mL of DYT containing 50 μg/mL of kanamycin. The cultures were incubated at 37 °C with

shaking at >200 rpm until mid-log phase growth was observed (OD_{600} 0.5-1.5). An aliquot of the culture 350 µL was mixed with 35 µL DMSO in a vial or tube and frozen immediately at –80 °C.

DNA sequence determination

Verification of pET28a plasmids containing cDNA of mouse KHK-A was performed by commercial DNA-sequence determination. Standard oligonucleotide primers for the T7 promoter and T7 termination sequences in pET28a were provided by Eurofins. A minimum concentration of 50 ng/µL of purified plasmid in a 10 µL of purified plasmid at was sent to Eurofins for sequencing.

Agarose gel electrophoresis of DNA

DNA samples were mixed with loading buffer (15% (w/v) glycerol, 1% (w/v) SDS, 0.04% (w/v) xylene cyanol FF, 0.04% (w/v) bromophenol blue) at a 5:1 ratio, and loaded into wells of 1-1.5% (w/v) agarose in TAE (40 mM Tris, pH 8.1, 20 mM acetate, 1 mM EDTA). Gels were electrophoresed at 80-110 mV for an appropriate amount of time, and then stained by immersion in 0.5 µg/mL of ethidium bromide for 10-15 min. Gels were de-stained for 10 min in dH_2O, and visualized using a Gel Doc imaging system (Bio-Rad). DNA size and concentration were measured by comparing band running distance and intensity, respectively, to an aliquot of known DNA standard run on the same gel (Hi-Lo DNA Marker, Minnesota Molecular Inc.).

Expression verification

Individual colonies of expression plasmids transformed into *E. coli* strain BL21(DE3) were picked from DYT plates and used to inoculate 10 mL of DYT containing 50 µg/mL of kanamycin. The cultures were incubated at 37 °C with shaking at >200 rpm until mid-log phase growth was observed (OD_{600} 0.5-1.5). An aliquot (1 mL) was removed. The expression was induced by addition of IPTG to a final concentration of 1 mM. The cultures were incubated at 37 °C with shaking at >200 rpm for 4-24 h before another aliquot (1 mL) was removed. Cells in aliquots were pelleted by centrifugation at 11,000*g* for 20 s and its wet pellet weight was determined. The typical wet weight of these 1 mL aliquots was 2.2 mg. Cells were lysed in 1 mL of 6XSDS sample buffer (250 mM Tris, pH 6.8, 10% (w/v) SDS, 0.5% (w/v) bromophenol blue, 50% (v/v) glycerol, and 715 mM β-mercaptoethanol), and 0.5 µL (~ 1 ug) was used to load on a gel. Expression was determined by SDS- polyacrylamide gel electrophoresis (PAGE).

SDS- polyacrylamide gel electrophoresis

Protein samples were separated by electrophoresis using a discontinuous gel sy stem, consisting of a stacking gel (top layer) and a resolving gel (bottom layer) [(Laemmli 1970)]. The stacking gel consisted of 5% (w/v) acrylamide/bis-acrylamide (29:1 ratio), 125 mM Tris, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate and 0.04% (v/v) TEMED. The resolving gel consisted of 12% (w/v) acrylamide: bis-acrylamide (29:1), 375 mM Tris, pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, and 0.06% (v/v) TEMED.

Protein samples were prepared for electrophoresis by combining a protein aliquot with 6xSDS sample buffer at a 5:1 ratio for a total volume of 20 μL, heated at 95 °C for 5 min, then loaded onto gel wells. Proteins of known size (Broad Range Molecular Weight Marker (Bio-Rad, Hercules CA) containing myosin (200 kD), β-galactosidase (116 kD), phosphorylase b (97.4 kD), serum albumin (66.2 kD), ovalbumin (45 kD), carbonic anhydrase (31 kD), trypsin inhibitor (21.5 kD), lysozyme (14.4 kD), and aprontinin (6.5kD), were included as a "Marker" set according to manufacturer's instructions. Gels were electrophoresed for 35 min at 75 mV to move samples through the stacking gel, and then at 200 mV until the dye ran off the resolving gel.. Proteins were visualized by staining with AquaStain (BullDog Bio) according to the manufacturer's instructions.

Expression and Purification of Recombinant Ketohexokinase-A

Cultures (1-2 L) in DYT media containing the resistant antibiotic, kanamycin, (50 ug/mL) were inoculated with 3-5 single colonies of pET28a-based expression plasmids for either human or mouse KHK-A from a fresh YT plate containing the same antibiotic. The culture was incubated at 37 °C with shaking. The cell growth was monitored by media turbidity until the OD at 600 nm reached 0.8-0.85, whereupon IPTG was added to 1 mM. Induction continued at 30 °C with a shaking for 12-14 hr. Prior to induction and after induction, aliquots (~1 mL) were collected in a pre-weighed centrifuge tubes and the cells collected by centrifugation at 13400 rpm for 30 s. The net cell wet weight was calculated from the weight difference. After induction, cultures were centrifuged at 4000*g* at 4 °C for 15 min, and the cell pellet was gently resuspended without creating foam at 4 °C in a

volume that was 2.5% of culture volume using a buffer containing 250 mM MOPS, 10 mM EDTA, 2 mM DTT, 1% (v/v) glycerol, 0.1 mg/mL DNAse, 0.1 mg/mL RNAse, and 10 uL/mL Protease-Inhibitor Cocktail (Sigma). The cells were lysed using a French-pressure cell under 20,000 – 25,000 psi. The cell lysate was centrifuged at 37,000*g* for 30 minutes. As an alternative, the cell lysate could be ultracentrifuged at 50,000 rpm for 30 minutes to remove the ribosomes and cell debris to enhance the result in crystallography. Supernatant fraction was loaded onto a 5 mL pre-packed His Trap^{TM} FF nickel column pre-equilibrated with binding buffer (10 mM HEPES, pH 7.5, and 20 mM imidazole) at 0.1 mL/min by creating a loop at 4 °C. The column was washed with 40-80 column volumes of binding buffer at ≤0.25 mL/min. The His-tagged protein was eluted with an elution buffer (10 mM HEPES, pH 7.5 plus 500 mM imidazole) at the flow rate of 1 mL/min by collecting 1.5 mL fractions. Protein concentration was determined for each fraction [Bradford 1976] and peak fractions were combined after its activity was confirmed (see *Protein Concentration Determination and Enzymatic Enzyme Assays).*

The most concentrated fractions were pooled and dialyzed (1 cm diameter 10,000 MW cut-off) at 4 °C against >100x volume of dialysis buffer (10 mM HEPES, pH 7.5, 0.1 mM DTT) for four hours or longer and repeated. Dialysate was centrifuged for 10 min at 12,000*g* to remove any of the white precipitate. The activity of the supernatant fraction after dialysis was compared to the activity before concentration. The dialysate was concentrated to ≥10 mg/mL by ultrafiltration using an Amicon ultrafiltration apparatus using a 10 kDA MWCO membrane <75 psi.

Protein Concentration Determination

Protein concentrations were determined by a Coomassie dye-binding assay (Bradford, 1976). The assays were preformed in a 1 mL plastic cuvette or a 96 well plate (final volume 250 µL), by combining 20 µL or 50 µL protein sample of appropriate dilution with either 980 µL or 950 µL, respectively, of Bradford reagent (0.01% Coomassie Brilliant Blue G-250, 4.7% EtOH, 8.5% phosphoric acid). Samples were incubated for 2-5 min at room temperature, and absorbance read at 595 nm. Protein concentration was calculated from an equation derived from a standard curve (2-15 µg) with known concentrations of bovine serum albumin (BSA).

Protein concentration of purified KHK was determined using its absorbance at 280 nm. Beer's law was used with an absorptivity coefficient of $^{280}E_{0.1\%}$ of 0.84 (Scheme I). The Absorptivity Coefficient was determined from converting from the predicted Molar Extinction Coefficient as follows:. $\frac{26970 \frac{mL}{mmol*cm}}{32000 \frac{mg}{mmol}}$ ($E_{0\ 1\%}$ is equivalent to 0.1 g/100 mL, or 1 mg/mL).

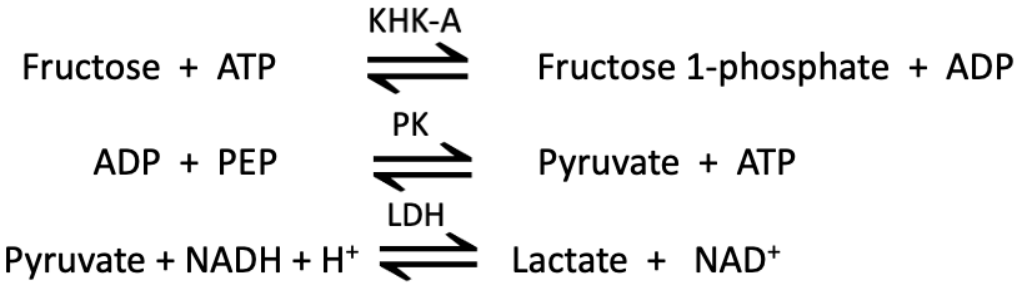
$$A^{280} \div \left(\text{Absorptivity Coefficient } \left(\frac{mL}{mg * cm}\right) \times \text{Path length } (cm)\right) = [\text{Protein}] \frac{mg}{mL}$$

Schematics I. Protein Concentraton Calculation from Absorbtivity Coefficient.

Enzymatic Activity Assays

Activity of KHK was measured by using a coupled enzyme assay based on the oxidation of NADH [Scheme II]. Assays of 1 mL samples contained 33 mM TEA (pH 7.4),

100 mM KCl, 20 mM $MgCl_2$, 0.3 mM NADH, 1.33 mM phosphoenolpyruvate (PEP), 4 mM ATP, 70 mM D-Fructose, and ~1.0 U/mL of the coupling enzymes (pyruvate kinase (PK) and lactate dehydrogenase (LDH)). Dilutions of KHK enzyme preparations were chosen that gave a rate of ΔA_{340}/min of 0.005 to 0.050, and absorbance was measured for 5–30 min. The specific activity of each mouse and human KHK-A was confirmed by activity assay designed to contain excess concentrations of substrates $>10 \cdot K_m$ [Asipu 2007]. During each step of the purification, column elution, post-dialysis, post-ultrafiltration by Amicon device or conical centricons, with 10,000 MWCO filter, the tests confirmed samples retained KHK activity. The ΔA_{340}/min rates were converted to U/mg using Beer's law and the molar extinction coefficient of NADH (6200 $M^{-1}cm^{-1}$), then multiplying by the volume (mL) of the assay and dividing by the amount of enzyme added (mg) to get μmol/min•mg, or U/mg (see Scheme III).



Schematics 2. Schematics KHK Coupled Enzyme Assay

$$(\Delta A\ min^{-1})/(1\ cm \bullet 6.2\ (mM \bullet cm)^{-1}) = (mM\ min^{-1} \bullet assay\ volume\ (mL))/enzyme\ (mg)^{*} = \mu mol/min \bullet mg = U/mg$$

*enzyme volume added (mL)•[enzyme](mg/mL)

Schematics III. Specific Activity Calculation.

Crystallography

Given that the structure of mouse KHK-A has never been determined before, mouse KHK-A crystals were grown by the hanging-drop vapor diffusion method using PCT Pre-crystallization Screen which included 4 different crystallization solutions (Hampton Research). A matrix of 4 different conditions at 6 different protein concentrations (8, 10, 12, 14, 16, 18 mg/mL) were set-up to see if any crystallization of the protein. Then, high-throughput Crystal Screen (Hampton Research, HR2-130), Index Screen (Hampton Research, HR2-134), and PEG/Ion Screen (Hampton Research, HR2-139) for further screen of crystallization conditions by sitting-drop vapor diffusion methods using a 96-well Intelli-Plate 96-3 LVR (Hampton Research). For all conditions a 0.7:1.2 ratio of protein to well solution was used to compose the total volume of 2 µL of sitting-drop equilibrated against 25 µL well reservoir. Crystals were grown at 17 °C.

For human KHK-A unliganded crystal conditions were screened as described above for the mouse KHK-A. Hits from the PCT Pre-crystallization Screen were optimized by varying the Ammonium sulfate concentration (0.4 – 1.4 M in 0.2 increments).

X-ray Diffraction and Data Collection

X-ray diffraction data were collected at the BU-CIC using the Bruker AXS X8 Proteum-R diffractometer with a Microstar Cu rotating anode X-ray source and a PLATINUM135 charge-coupled device (CCD) area detector. The crystallographic data were collected under constant flow N2 cryo-stream at 100 K, while crystals were cryoprotected by 25% Glycerol and 75% mother liquor solution after a flash-freezing test. PROTEUM3 was used to determine the unit cell and strategy for data collection.

Data Processing: Integration, Scaling, Phasing, and Refinement for Structural Solution

The data were processed in PROTEUM3 using SAINT for integration, and SADABS for scaling. Then aimless and pointless were used to convert the reflection files into a single merged .mtz file. Either the .mtz or the .sca file were used as a proxy to determine the quality of the data using Xtriage. Then, the phases were solved by using PHENIX.phaser molecular replacement. A monomer model of native mouse KHK-C complexed with fructose and ADP (PDB:6P2D; [Gasper et al. 2020]) was used as a reference molecule for phasing. Confirming a good quality of single MR solution is assessed by a value in the thousands of Top LLG score and tens of TFZ score. The first round of refinement was done using PHENIX. Refine. This round one confirmed that the unbiased electron density map fits the 298 backbone residues of the protein, then waters were added. Then subsequent rounds of refinement were done to assure a decrease in the R *free* and R *work* in the refinement statistics. PHENIX. Molprobity were run with the final .pdb file as a validation.

CHAPTER 3: RESULTS

The expression plasmids for mouse KHK-A were synthesized by Genewiz. The human KHK-A expression plasmid was subcloned from a mammalian vector, pPAL7, into pET28a previously (Gasper, 2020). These were transformed into two different *E. coli* strains. DH5α cells were used to prepare plasmid DNA to verify the DNA sequence. The optimized sequence of the plasmid DNA encoded the reference sequence for the protein (see Appendix). BL21(DE3) cells were used for preparation of permanent stabs and testing of expression. Each KHK-A gene in pET28a(+) was induced with 1 mM IPTG for 12 hours and total cellular protein was separated by SDS-PAGE. The expression of mouse KHK-A from the pET28a(+) plasmid was clearly inducible (Figure 3, lanes 2-3).

Sufficient quantities of KHK-A for crystallography was prepared from 2 L cultures followed by cell lysis and isolation by Ni-chelate chromatography as described in methods. Aliquots obtained during the purification were analyzed by SDS-PAGE along with the final product of purification (Figure 3, (lanes 4-7)). The pooled fractions from the column eluate for both mouse and human KHK-A, as well as the final concentrated preparation used crystallography, for were tested for activity and protein concentration (Table 3). The final yields were >60 mg. The increase in both activity and protein after dialysis was likely due to the high imidazole concentration [C Gasper, personal communication].

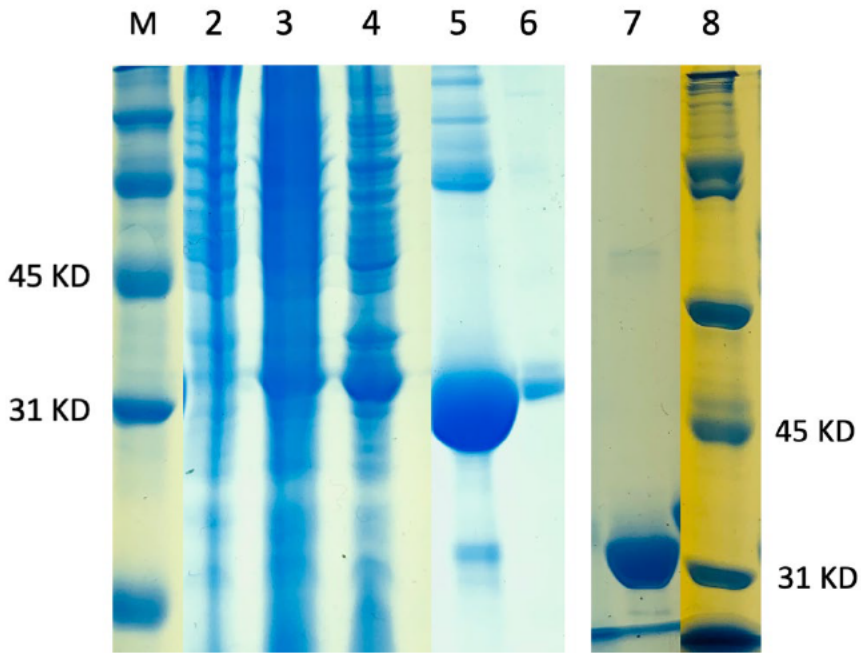


Figure 3. SDS-PAGE analysis of Induction and purification of mouse and human KHK-A. Samples were made from 1 mL aliquots of the culture before induction Lane 2), and after induction (Lane 3). The crude extract (supernatant fraction after lysis) (Lane 4) and purified preparations of mouse KHK-A (lanes 5 (26 µg) and 6 (3.5 µg)) and human KHK-A (lane 7). Lane 1 and 8 are the marker used for the gel for lanes 2-6 and 7, respectively.

Table 3. Purified and Concentrated Protein Preps

	[Protein] (mg/mL)	[Activity] (U/mL)	Volume (mL)	Total Units (U)	Total Protein (mg)	Specific Activity (U/mg)
Mouse KHK-A						
Column eluate	7.5	1.7	9	14	63	0.22
Post dialysis & Ultrafiltration	26	8	2.8	23	73	0.31
Human KHK-A						
Column eluate	7	4.4	8.5	38	60	0.63
Post dialysis & Ultrafiltration	26	52	2.5	130	65	2

Attempts at crystallography were performed using a number of conditions. Furthermore, Crystal screen, and an Index screen were attempted as described in methods. Each screen tested 96 different conditions. Both screens were unsuccessful; no crystals were found for either mouse and human KHK-A. However, a PEG/Ion screen yielded 11 hits for mouse KHK-A (Figure 4). The crystallization condition yielded >50 different crystals with dimensions of multiples of 0.1 mm. For example, the large crystal in Figure 4B was 2 mm x 1.5 mm, the crystal in Figure 4G was 0.4 mm in height, and the crystals in Figure 4F were 0.3 mm in width for middle poly-crystal, 0.3 mm for the right-side auxiliary crystal, and 0.4 mm for the additional crystal on the right left corner of the well. All these crystals formed within 21-27 days (crystals in Figure 4J took 27 days). The crystal in Figure 4A was tightly bound to the concaved plastic bottom of the well plate and could not be recovered.

Several crystals were looped for diffraction. When screened at BU-CIC, two cuboid crystals in Figure 4J didn't diffract well and had insufficient intensity peak counts. The clear crystal in Figure 4B and 4C broke while looping for inclusion in a trip to SSRL. Both crystals in Figure 4D were mounted and screened on the same day at BU-CIC, but the crystal on the left, which was originally 0.3 mm in diameter, was deformed as part of it was exposed to the oil-droplet and was undergoing a phase change. A small breakage also was observed on the same crystal due to vibrational force from moving the plates. Most importantly, a satisfactory diffraction pattern for apo mouse KHK-A was obtained from the octahedron crystal shown in Figure 4H that is closest to the surface or skin of the

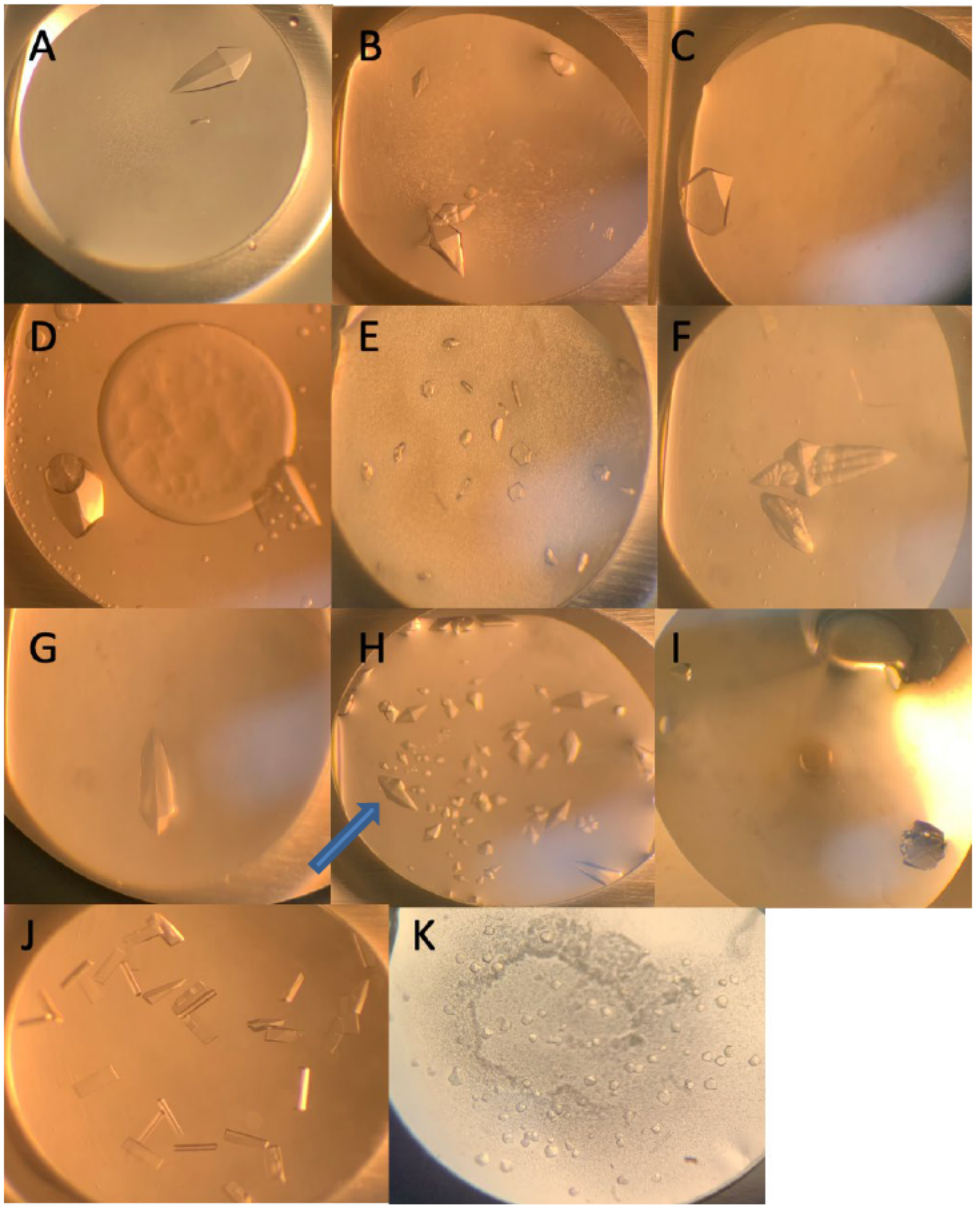


Figure 4. Crystals of Unliganded Mouse KHK-A. Mouse KHK-A crystallized in 4 % v/v Tacsimate and 12% w/v Polyethylene glycol 3350, pH 7 (Panel A - pyramid, flat bottom); 0.1 M Succinic acid and 12% w/v Polyethylene glycol 3350, pH 7 (Panel B - octahedron); 0.2 M Sodium formate and 0.2 M Sodium formate, pH 7.2 (Panel C -faceted cube); 0.2 M Sodium citrate tribasic dihydrate and 20% w/v Polyethylene glycol 3350, pH 8.3 (Panel D – faceted cube, rhombohedron); 8 % v/v Tacsimate and 20% w/v Polyethylene glycol 3350, pH 4 (Panel E – six-sided plate); 0.2 M Ammonium citrate dibasic and 20% w/v Polyethylene glycol 3350, pH 5.1 (Panel F – poly-crystalline); 0.2 M Sodium tartrate dibasic dihydrate and 20% w/v Polyethylene glycol 3350, pH 7.3 (Panel G -dagger); 0.2 M Sodium tartrate dibasic dihydrate and 20% w/v Polyethylene glycol 3350, pH 6.2 (Panel H - octahedron); 0.8 M Ammonium Tartrate Dibasic, pH 7.0 (Panel I – pyramids; poly-cube); 0.2 M Magnesium acetate tetrahydrate and 20% w/v Polyethylene glycol 3350, pH 7.9 (Panel J- cuboid); and 2 % v/v Tacsimate pH 5.0, 0.1 M sodium citrate tribasic dihydrate pH 5.6 (Panel K- octagon). All were 2 µL sitting drops with 25 µL well solutions set up as described in materials and methods. The crystals were observed using Olympus SZX12 stereo microscope under 37.5 magnification. Pictures were taken using iPhone XR camera.

drop (see blue arrow in Figure 4H). Since a crystal in Figure 4H diffracted well data were collected, no further crystals were screened.

Using PROTEUM3, PHENIX Xtriage, Phaser MR, and Refinement, a final structure was solved to a resolution of 1.85 Å. Mouse KHK-A crystallized in space group C 2221, with unit-cell parameters a= 42.2 Å, b=78.2 Å, c=149.8 Å, α= 90 °, β= 90 °, γ= 90 °. See data-collection statistics in Table 4. The final apo mKHK-A structure model was solved as a monomer in the asymmetric unit (Figure 5). Figure 6 shows the electron density map created by the diffraction data caging the mouse KHK-A-specific residues, His112-Ala113-Tyr114. This confirms that the expressed, purified and crystallized protein was indeed mouse KHK-A.

As mentioned in Chapter 1, the analysis of this new structure involved the question of whether KHK-A, like KHK-C and other members of the pfkB family (Matthews et al., 1998; Sigrell et al., 1998; Park & Gupta 2008), can undergo conformational changes in the β-clasp domain. Figure 6 demonstrates these conformational changes and/or lack thereof. KHK-C in its unbound (unliganded) and bound (ternary) forms show a nearly 30° change in the angle of the β-clasp domain relative to the α/β domain (Figure 7A). Whereas, as shown by Trinh et al. (2009), the KHK-A in its unbound (unliganded) and bound (ternary) form are superimposable (Figure 7B).

When the mouse unliganded KHK-A structure was aligned similarly to the unliganded and ternary KHK-A structures as in Figure 7B, there was an observable change in the conformation of the β-clasp domain (Figure 8). The alignment with the ternary human KHK-A shows over 10° change in conformation. A similar alignment with the

Table 4. Crystallographic Data Collection and Refinement Statistics for Mouse KHK-A Unliganded Structure.

	Mouse KHK-A Unliganded
Source	Microstar Cu rotating anode
Detector	PLATINUM135 CCD
Wavelength (Å)	1.54
Resolution range (Å)	21.94 – 1.85 (1.92 – 1.85)*
Space group	C 222_1
Cell dimensions and Refinement	
a, b, c (Å)	42.18, 78.22, 149.84
a, b, g (°)	90, 90, 90
Total reflections	43208 (4256)
Unique reflections	21608 (2130)
Multiplicity	2.0 (2.0)
Completeness (%)	99.68 (99.95)
Mean I/sigma (I)	30.19 (8.76)
Wilson B-factor	13.72
R-merge†	0.016 (0.076)
R-meas	0.022 (0.107)
CC1/2	1.00 (0.995)
Reflections used in refinement	21570 (2535)*
Reflections used for R*free*	1117 (124)*
R*work*	0.188 (0.242)*
R*free*	0.230 (0.303)*
Number of atoms	
Non-hydrogen	2606
Protein	2301
Water	305
Protein residues	298
B factors (Å^2)	
Average	16.29
Protein	15.35
Water	23.24

*Statistics for the highest resolution shell are shown in parenthesis. †$R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

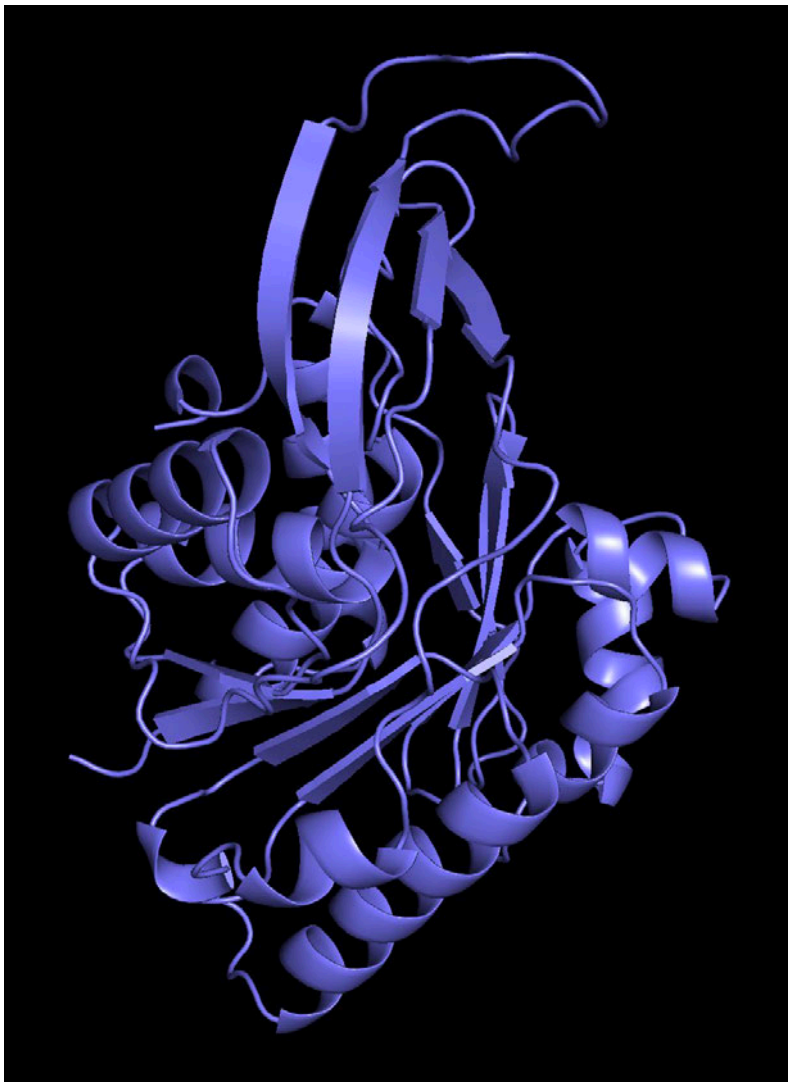


Figure 5. The Structure of Mouse Unliganded KHK-A. The 1.85 Å-resolution alpha-carbon backbone is depicted as a ribbon structure with the β-clasp domain oriented at the top. It is depicted as a monomer, which was the asymmetric unit.

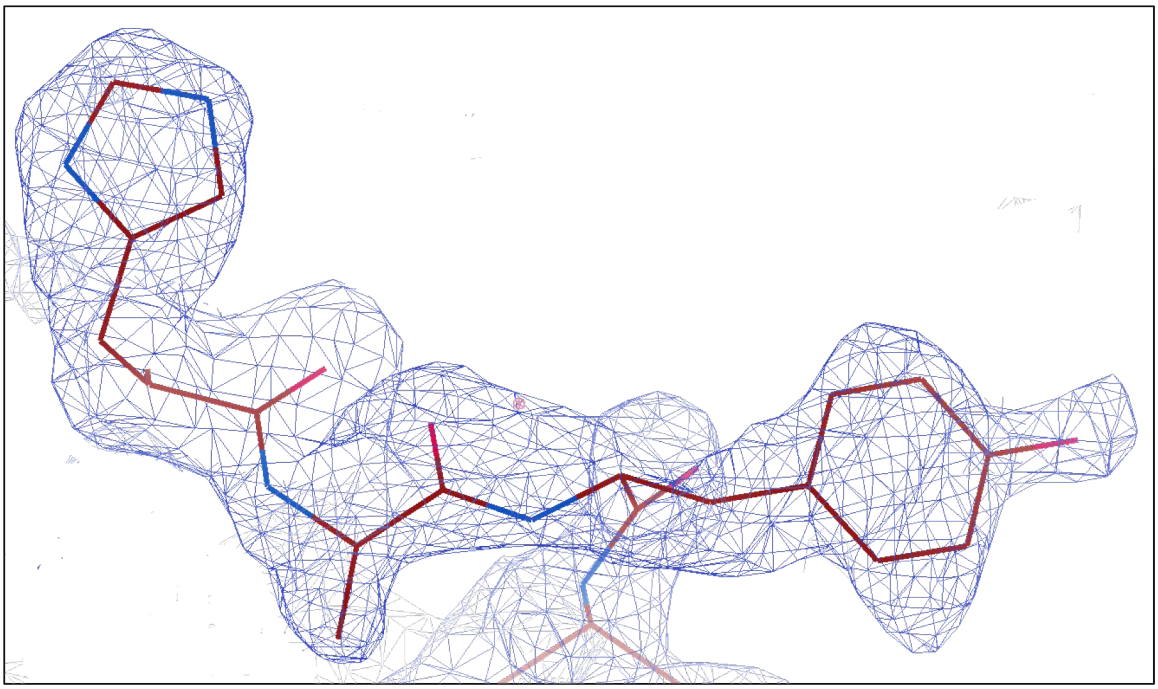


Figure 6. Amino-acid Residues 112-114 of Mouse KHK-A. The $2F_o$-F_c electron-density map (blue cages) contoured at 0.7 σ for mouse KHK-A residues His112-Ala113-Tyr114, which are specific for mouse KHK-A among KHK enzymes (see alignment in Appendix).

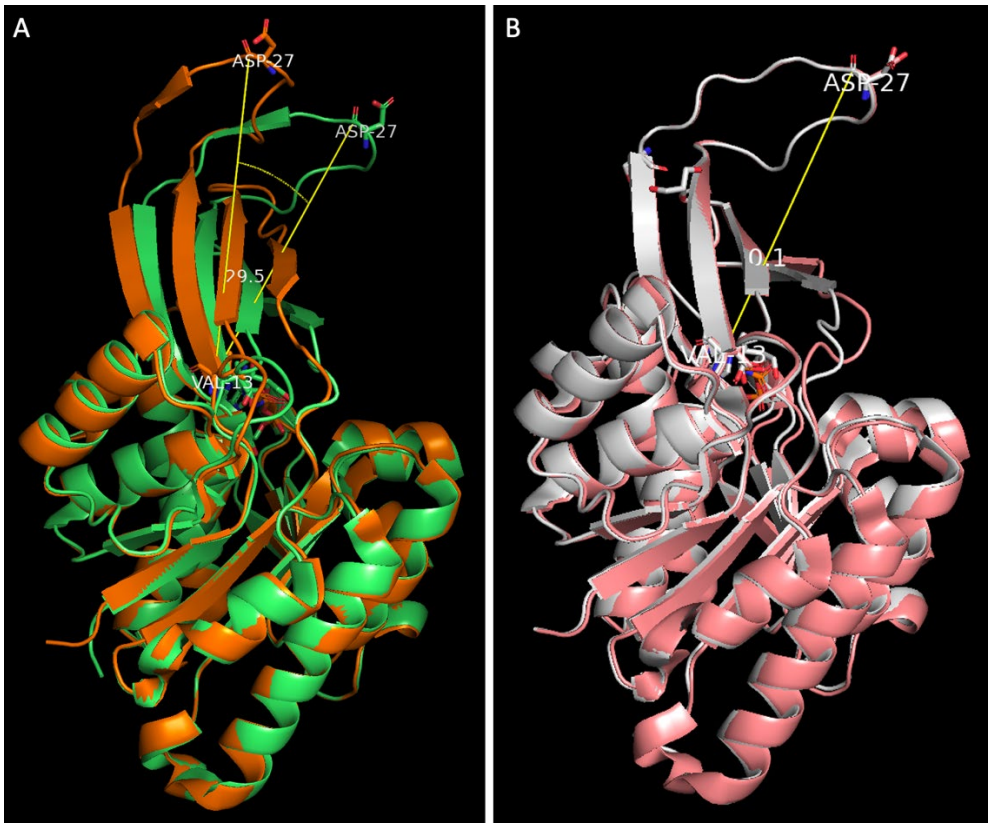


Figure 7. Overlays of Unliganded and Ternary Human KHK-C and KHK-A. **Panel A**: Chain A of unliganded KHK-C (green, PDBID:3B3L) and chain A of ternary KHK-C (orange, PDBID:3NBV) were aligned in PyMol and oriented as in Fig. 5A. The degree of change in the β–clasp domain was quantified by measure of the angle between the carboxyl carbon of Asp27 (identified in white text) of each β-clasp in relation to a fulcrum point, the Cα of Val13, located at the N-terminus of β2 strand. The number is the angle value between the lines (yellow). **Panel B**: unliganded KHK-A (pink, PDBID:2HQQ) and ternary KHK-A (light gray, PDBID:2HW1) were aligned as described in panel A.

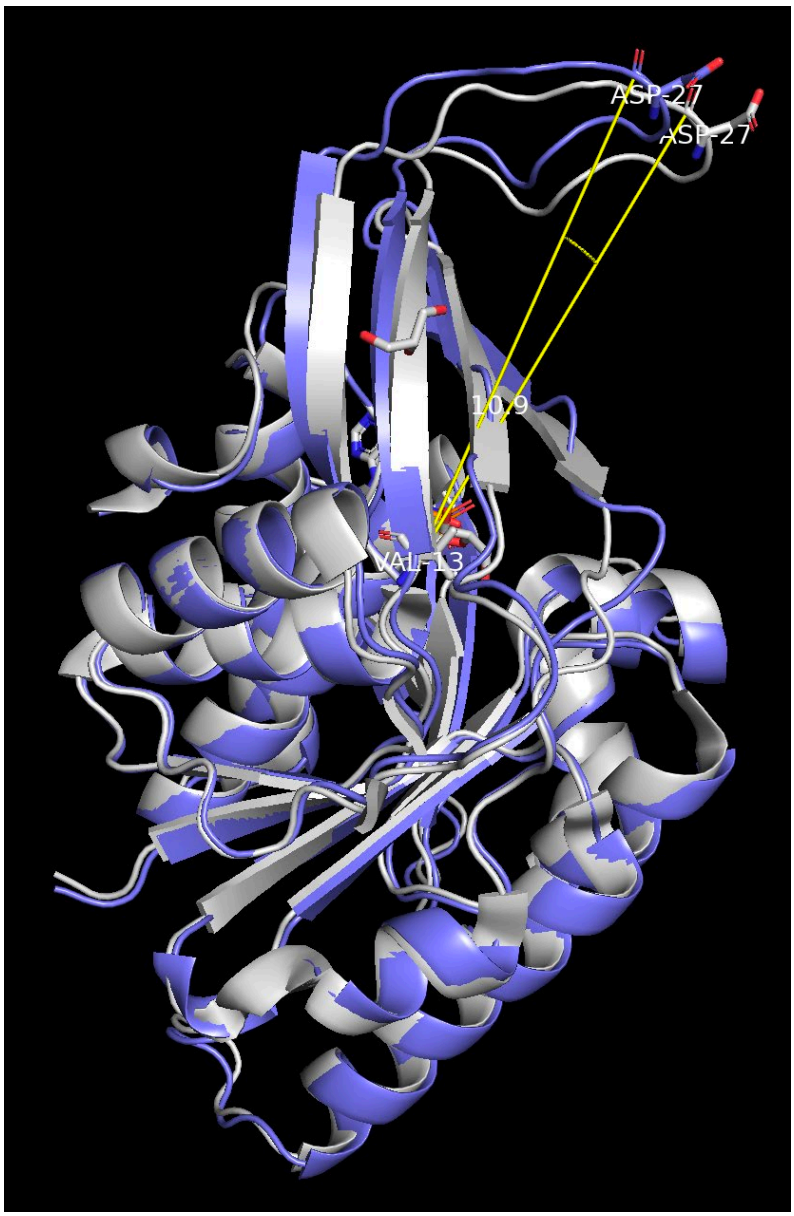


Figure 8. Overlay of Unliganded Mouse KHK-A and Ternary Human KHK-A. The ribbon structures of mouse KHK-A (purple) (see Fig. 5) and ternary human KHK-A (light gray, PDBID:2HW1) were aligned and distinguished as described in Figure 7. When substrates are bound, the β-clasp rotates to a closed conformation by 10.9°(Panel A). The distance between the two carbonyl carbon of Asp27 is 5.3 Å. Demonstrated is that mouse KHK-A has the ability to make a conformational change unlike what was claimed by Trinh *et al*. (2009). However, since this was a comparison between the mouse and human KHK-A, attempts were made to determine the human unliganded structure (Figure 9).

unliganded human KHK-A was the same (data not shown). It appears that the unbound form of KHK-A from mouse has a different structure, or ability to change structure, than the human unbound form. To distinguish whether this is due to a species difference or a different observation for KHK-A from what was published (Trinh et al. 2009), the crystals of human unliganded KHK-A were obtained but a structure was not achieved, since it has been solved previously. An unpublished deposited structure was used in the analysis.

Purified human KHK-A was used for crystallography using the same screens as was performed for the mouse structure. A PCT Pre-crystallization screen revealed a hit of human KHK-A crystals using condition B1 (0.1 M Tris-HCl, pH 8.5, 1.0 M ammonium sulfate) and was optimized with a 4 x 6 matrix as described in methods, but yielded no success. A total of 292 different conditions were screened, and three were successful (Figure 9).

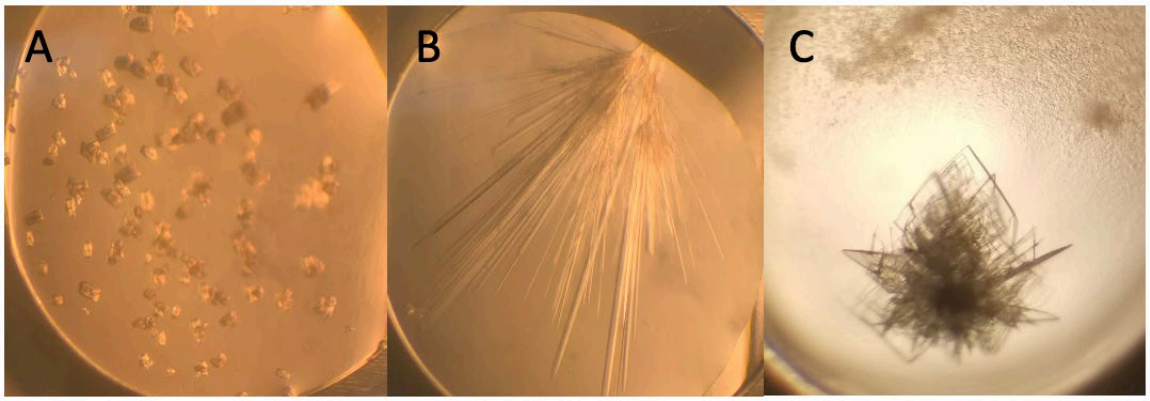


Figure 9. Crystals of Unliganded Human KHK-A. Purified human KHK-A was crystallized as described in Chapter 2 by sitting-drop vapor diffusion using either 0.15 M Cesium chloride and 15% w/v Polyethylene glycol 3350 (Panel A- fine cluster), or 0.02 M Calcium chloride dihydrate, 0.02 M Cadmium chloride hydrate, 0.02 M Cobalt (II) chloride hexahydrate and 20% w/v Polyethylene glycol 3350 (Panel B- needles), or 1.1 M Ammonium tartrate dibasic (Panel C- poly-crystaline, layered platelets) as well solutions. The crystals were observed using Olympus SZX12 stereo microscope under 37.5 magnification. Pictures were taken using iPhone XR camera.

Interestingly, the human enzyme did not crystalize under any of the conditions under which the mouse KHK-A crystalized (see Figure 3). Instead three different conditions were successful. The crystallization conditions yielded micro (<0.1 mm) crystals, needles, and plates. None of these were used for diffraction because there was an independently determined structure of human unliganded KHK-A already deposited in the PDB in 2006 (PDBID:2HLZ; W. M. Rabeh, W. Tempel, L. Nedyalkova, R. Landry, C. H. Arrowsmith, A. M. Edwards, M. Sundstrom, J. Weigelt, A. Bochkarev & H. Park, unpublished work). This structure has four chains in the asymmetric unit with three different conformations of the β-clasp domain.

All four chains from the unliganded human KHK-A structure (PDBID:2HLZ) were aligned with the mouse unliganded KHK-A structure (Table 5). For each alignment, the angle between the α/β domain and the β-clasp domain was measured as well as the resulting distance between the Asp27 residues at the tip of the β-clasp domain. Chains B and C had similar conformations with a 7° angle, whereas chain D had a 10° angle, similar to the difference between the ternary human KHK-A (PDBID: 2HW1) and the unliganded mouse KHK-A shown in Figure 7. As shown in Figure 10, one of these human chains (chain A) adopts a very similar conformation to the mouse KHK-A. These two structures have only a ~5° angle change with a distance at the slightly more than level as the resolution of the structure (1.8 Å, see Table 2). Although the degree of the apparent conformational change in the β-clasp domain upon substrate binding for KHK-A (see Figure 7) was not as large as that of KHK-C (see Figure 6A), it's clear that there is a conformational change and that it is not species specific.

Table 5. Different Conformations of the β-clasp Domain for Unliganded Human KHK-A[1] (Chains A-D).

	Distance[2]	Angle[3]
Chain A	2.8 Å	5.5 degrees
Chain B	3.3 Å	7 degrees
Chain C	3.9 Å	7 degrees
Chain D	4.8 Å	10 degrees

[1]human KHK-A (PDBID:2HLZ) has four chains in the asymmetric unit. Each chain was aligned with mouse KHK-A and distinguished as described in Figure 7.
[2]The distance between the two carbonyl carbons of Asp27
[3]The angles between lines from the Cα of Val13 and the carbonyl carbon of Asp27

Figure 10. Unliganded Mouse KHK-A and Unliganded Human KHK-A have Similar Conformations. The ribbon structures of unliganded mouse KHK-A (purple) (see Fig. 5) and chain A of unliganded human KHK-A (green, PDBID:2HLZ) were aligned and distinguished as described in Figure 8.

CHAPTER 4: **DISCUSSION**

Previously, a case by Trinh *et al*. (2009) had been made that there is no significant change in overall conformation when comparing unliganded KHK-A and its ternary complex, which they supported by no evidence of (β-clasp) domain movement using *DynDom* (Trinh *et al.*, 2009). Furthermore, they supported this conclusion in the discussion, stating, "KHK-A does not undergo a conformational change upon substrate binding. Indeed, …. aminoimidazole riboside kinase with and without substrate have been shown not to undergo any conformational changes …. (Zhang *et al.*, 2004)" (Trinh *et al.*, 2009). Proposed here is to reassess Trinh et al.'s discounting of the "small rotation between domains" that corresponds to hinge motion, exposing the chemistry site, and to provide a new KHK-A model from a different species other than human. Here, the mouse KHK-A structural models without substrates bound, as well as a synthesized analysis of the some previously unpublished structures (Gibbs *et al.*, 2010) show that there is a possibility of a conformational change in KHK-A.

Correlation of fructose metabolism with cancer, first described for breast cancer (Zamora-Leon *et al.*, 1996), and now for colorectal, esophageal, hepatic, and small-cell lung cancer have been reported (Deng *et al.*, 2022; Kim *et al.*, 2020; Li *et al.*, 2016; Shen *et al.*, 2022; Xu *et al.*, 2019; Yang *et al.*, 2021; Yau *et al.*, 2017). The recent work has shown that this pathology is associated specifically with KHK-A. Other pathologies have

involved a switch from KHK-A to KHK-C, as was shown in cardiac hypertrophy (Mirtschink *et al.*, 2015). The KHK-A isozyme appears to be of medical importance.

One study by Kim *et al*. (2020) developed a mechanism behind this pathology. KHK-A is not expressed in the liver but is ubiquitously upregulated in human cancer cells, for example breast cancer. The increase in metastasis was shown using MDA-MB-231 cells. The cells were morphologically altered by F-actin rearrangement and an epithelial–mesenchymal transition, which changed cell adhesion by suppressing the expression of E cadherin, when exposed to fructose. Silencing the overexpression of KHK-A, attenuated the fructose-induced invasion. Furthermore, the mechanism proposed involves cytoplasmic KHK-A entering the nucleus following fructose stimulation, which was mediated by LRRC59 (Leucine-rich repeat-containing protein 59) and KPNB1 (Karyopherin Subunit Beta 1). In the nucleus, KHK-A phosphorylates YWHAH (a form of a 14-3-3 ETA protein) at Ser25 and YWHAH recruits SLUG to the E-cadherin promoter, triggering cell migration. This protein kinase activity has also been reported to phosphorylate and up-regulate PRPS1 in oesophageal squamous cell carcinoma (Kim *et al.*, 2020; Yang *et al.*, 2021). Others have shown the role of oxidative stress in the activation of KHK-A (versus KHK-C) expression (Xu *et al.*, 2019) wherein KHK-A is itself phosphorylated at the KHK-A-specific residue Ser80 (see Appendix). It has been reported that AMPK (5'-adenosine monophosphate-activated protein kinase) phosphorylates KHK-A at Ser80 (Xu *et al.*, 2019). It seems clear the KHK-A has some specific functions that are distinct from KHK-C. Both isozymes represent clear targets for medicinal chemistry with disease pathologies, and in particular the role, if any, of substrate-

induced conformational changes in these activities highlights the importance of such studies.

As mentioned in the introduction, ribokinase is in the same pfKB family (Bork *et al.*, 1993; Hansen *et al.*, 2007) of carbohydrate kinases. The conserved residues across all enzymes in the family, are two motifs, 1) a GG near the N-terminus and a DTXGAGD near the C-terminus (shown highlighted in green) [see Appendix]. Most importantly, it is a structurally homologous to KHK, while it only has 19% sequence similarity. Despite the low sequence homology, Adenosine Kinase (AK) and Ribokinase (RK) share structural homology by large. For both AK and RK, the hinge motions by 30 ° and 17°, subsequently, were detected upon each substrate sugar binding. The subsequent binding of the nucleotide induced a comparatively a less drastic but equally essential change were observed done the unliganded versus the substrate-bound forms (Matthews et al., 1998; Schumacher et al., 2000; Sigrell et al., 1998). Ribokinase has been studied crystallographically with the implications of conformational changes by induced-fit changes upon binding its substrate. According to Sigrell *et al*. (1999) the sugar substrate, ribose, initially binds to an open active site and displaces water to form the more closed form of the β-clasp domain from the opposite subunit onto the active site (for the monomer it appears more open). This small conformational change affects the close-by nucleotide binding site, potentially increasing the affinity for the second-substrate binding. At that point, after closure of the active-site lid, the ribose is close enough for a nucleophilic attack on the γ-phosphate of ATP.

Sigrell *et al*. (1999) proposed that ribokinase operates in sequential ordered mode by induced-fit binding during their catalysis. The similar mode of binding and induction

of the nucleotide binding for Toxoplasma gondii adenosine kinase is shown by Schumacher et al., in 2000. On the other hand, while KHK has a good deal of sugar specificity (Adelman *et al.*, 1967; Bais *et al.*, 1985; Raushel *et al.*, 1973), it reportedly used a sequential random kinetic mechanism (Raushel *et al.*, 1973), as depicted in Figure 11. Although, this report did not show the data, KHK may also exhibit a similar ordered substrate-binding mode. Subsequent studies, testing substrate induced-fit binding (mode of conformational change) may be conducted to see the differences of snap shots of the unliganded (open) form, the single substrate (open or closed), and the ternary (open or closed). These studies could be supported by binding studies (SPR, equilibrium dialysis), spectroscopic (NMR and CD) analysis, and activity kinetics studies.

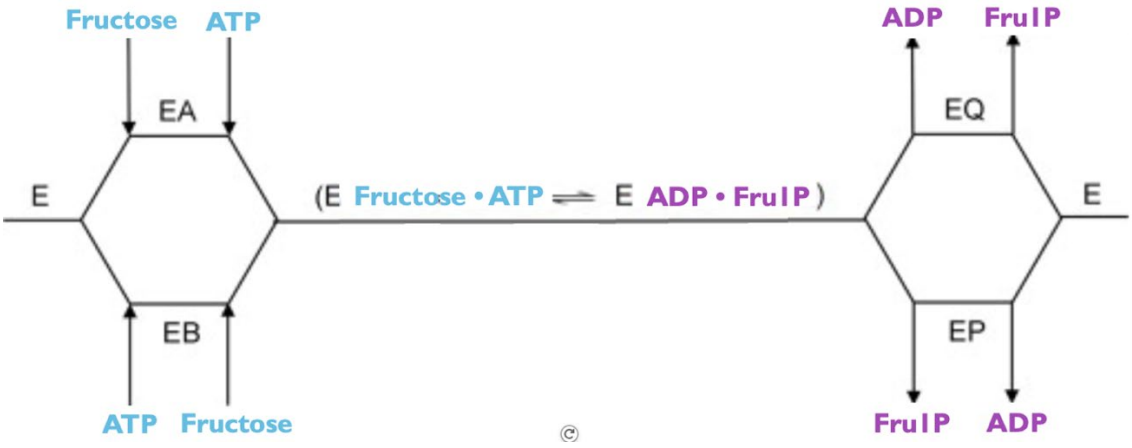


Figure 11. Diagram of KHK Kinetic Mechanism. The Sequential-random kinetic mechanism of KHK (modified from (Gasper, 2020)).

APPENDIX

Figure 1S. Genewiz Expression Plasmid Sequences

Mouse KHK-A DNA Reference Sequence: **NM_001349066**

GAAGAGAAGCAGATCCTGTGCGTGGGGCTGGTGGTGCTGGACATCATCAATGTGGTGGACAAATACCCAGAGGAAGACACGGATCGCAGGTGCCTGTCCCAGAGATGGCA
GCGTGGAGGCAACGCATCCAACTCCTGCACTGTCCTTTCCTTGCTTGGAGCCCGCTGTGCCTTCATGGGCTCTTTGGCCCCTGGCCACGTTGCCGACTTTGTCCTGGATG
ACCTCCGCCAACATTCTGTGGACTTACGATATGTGGTCCTTCAGACCGAGGGCTCCATCCCCACTTCTACAGTCATCATCAACGAGGCCAGCGGCAGCCGCACCATTCTG
CACGCCTACAGGAACCTGCCAGATGTGTCTGCTAAGGACTTTGAGAAGGTCGATCTGACCCGGTTCAAGTGGATCCACATTGAGGGCCGGAATGCATCGGAACAGGTGAA
GATGCTGCAGCGGATAGAGGAGCACAATGCCAAGCAGCCTCTGCCACAGAAGGTCCGGGTGTCGGTGGAGATAGAGAAGCCCCGTGAGGAGCTCTTCCAGTTGTTTAGCT
ATGGTGAGGTGGTGTTTGTCAGCAAAGATGTGGCCAAGCACCTGGGGTTCCAGTCAGCAGTGGAGGCCCTGAGGGGCTTGTACAGTCGAGTGAAGAAAGGGGCTACGCTT
GTCTGTGCCTGGGCTGAGGAGGGTGCCGATGCCCTGGGCCCCGATGGTCAGCTGCTCCACTCAGATGCCTTCCCACCGCCCCGAGTAGTAGACACTCTTGGGGCTGGAGA
CACCTTCAATGCCTCTGTCATCTTCAGCCTCTCGAAGGGAAACAGCATGCAAGAGGCCCTGAGATTCGGGTGCCAGGTGGCTGGCAAGAAGTGTGGCTTGCAGGGGTTTG
ATGGCATTGTG

The Optimized (for *Escherichia coli*) sequence of mKHK-A (black) aligned with the NM_001349066 (blue). A 200 T 193 C 230 G 283 | GC%: 56.62% | Length: 906; (*) indicate changes, restriction sites indicated by underline, start codon indicated in green, and stop codon in red.

```
CATATGGAGGAAAAACAGATTCTGTGCGTTGGTCTGGTGGTGCTGGACATCATCAACGTGGTGGACAAGTACCCGGAAGAGGATACCGACCGCCGCT
NdeI  GAAGAGAAGCAGATCCTGTGCGTGGGGCTGGTGGTGCTGGACATCATCAATGTGGTGGACAAATACCCAGAGGAAGACACGGATCGCAGGT
        *  *  *     *        *   *                        *           *     *   *  *  *  *  *   ***

GTCTGAGTCAGCGTTGGCAACGTGGCGGCAATGCCAGTAATAGTTGCACCGTGCTGAGTCTGCTGGGTGCCCGTTGTGCCTTCATGGGCAGTCTCGC
GCCTGTCCCAGAGATGGCAGCGTGGAGGCAACGCATCCAACTCCTGCACTGTCCTTTCCTTGCTTGGAGCCCGCTGTGCCTTCATGGGCTCTTTGGC
 *   ***  ** *      *     *     *  ****  ****     *   *  *****    *  *     *                ** * *

GCCGGGCCATGTTGCCGATTTCGTTCTGGACGATCTGCGCCAGCATAGCGTTGATCTGCGCTACGTTGTGCTGCAGACGGAAGGCAGCATTCCGACC
CCCTGGCCACGTTGCCGACTTTGTCCTGGATGACCTCCGCCAACATTCTGTGGACTTACGATATGTGGTCCTTCAGACCGAGGGCTCCATCCCCACT
*  *     *        *   *  *     *  *  *     *   ***   * **  *  *  *  *  *  *     *   *  **    *     *

AGCACCGTGATTATCAACGAAGCCAGTGGCAGCCGTACCATTCTGCACGCCTATCGCAATCTGCCAGACGTGAGCGCCAAAGATTTCGAGAAGGTGG
TCTACAGTCATCATCAACGAGGCCAGCGGCAGCCGCACCATTCTGCACGCCTACAGGAACCTGCCAGATGTGTCTGCTAAGGACTTTGAGAAGGTCG
***  *  *  *         *     *        *                 **  *  *        *   ***  *   *  *  *        *

ATCTGACGCGCTTTAAGTGGATCCACATCGAGGGCCGTAACGCCAGCGAACAAGTTAAAATGCTGCAGCGCATCGAAGAACACAACGCGAAACAGCC
ATCTGACCCGGTTCAAGTGGATCCACATTGAGGGCCGGAATGCATCGGAACAGGTGAAGATGCTGCAGCGGATAGAGGAGCACAATGCCAAGCAGCC
       *  *  *               *        *  *  ****      *  *  *           *  *  *  *      *  *  *

ACTCCCGCAGAAAGTGCGTGTGAGCGTTGAAATCGAGAAGCCGCGCGAAGAACTGTTCCAGCTGTTCAGCTACGGCGAGGTGGTGTTCGTGAGTAAG
TCTGCCACAGAAGGTCCGGGTGTCGGTGGAGATAGAGAAGCCCCGTGAGGAGCTCTTCCAGTTGTTTAGCTATGGTGAGGTGGTGTTTGTCAGCAAA
*  *  *     *  *  *    ***  *  *  *           *  *   *  *      *    *     *  *            *  *  *  *

GATGTGGCGAAACACCTCGGCTTCCAAAGCGCCGTTGAAGCGCTGCGCGGTCTGTATAGCCGCGTTAAGAAAGGTGCCACGCTGGTTTGTGCGTGGG
GATGTGGCCAAGCACCTGGGGTTCCAGTCAGCAGTGGAGGCCCTGAGGGGCTTGTACAGTCGAGTGAAGAAAGGGGCTACGCTTGTCTGTGCCTGGG
           *         *     ****  *  *  *  *  ** *   **    *  *  *  *        *  *      *  *     *

CCGAAGAAGGTGCGGATGCGCTGGGTCCAGATGGCCAGCTGCTGCATAGCGATGCCTTTCCGCCGCCACGCGTTGTTGATACCCTCGGCGCGGGCGA
CTGAGGAGGGTGCCGATGCCCTGGGCCCCGATGGTCAGCTGCTCCACTCAGATGCCTTCCCACCGCCCCGAGTAGTAGACACTCTTGGGGCTGGAGA
 *  *  *     *      *     *  *     *        *  ****         *  *     *  *  *  *  *   *  *  *  *  *

TACGTTCAACGCCAGCGTGATCTTTAGTCTGAGCAAGGGCAACAGCATGCAAGAAGCGCTGCGTTTTGGTTGTCAAGTTGCCGGCAAGAAGTGCGGT
CACCTTCAATGCCTCTGTCATCTTCAGCCTCTCGAAGGGAAACAGCATGCAAGAGGCCCTGAGATTCGGGTGCCAGGTGGCTGGCAAGAAGTGTGGC
*  *     *   ***  *      *  *  ****     *               *  *   **   *  *  *  *  *  *            *  *

CTGCAAGGCTTTGATGGCATCGTTTAAAAGCTT
TTGCAGGGGTTTGATGGCATTGTGSTOP-HindIII
*    *  *            *  *
```

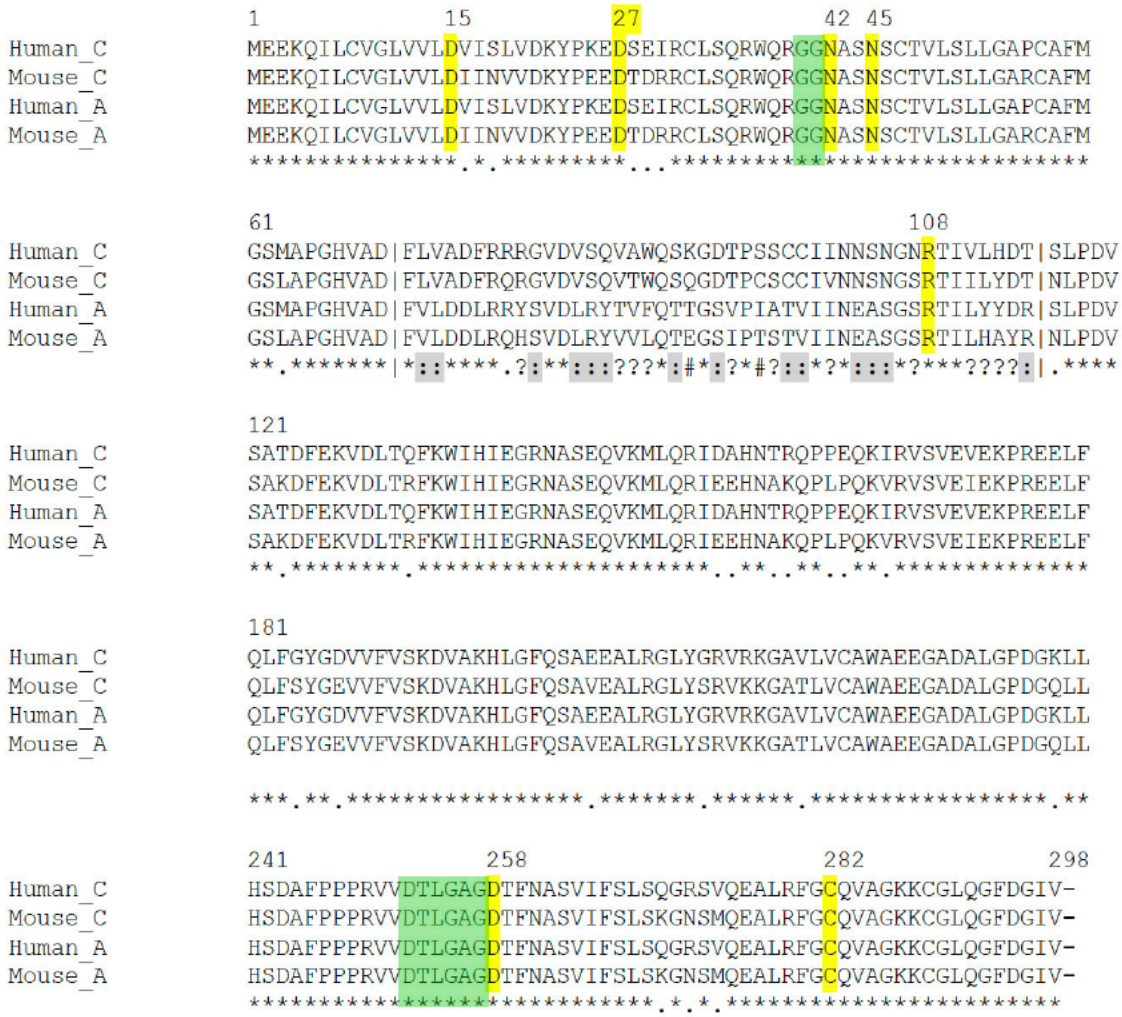
Mouse KHK-A Protein Sequence: **NP_001335959**

MEEKQILCVGLVVLDIINVVDKYPEEDTDRRCLSQRWQRGGNASNSCTVLSLLGARCAFMGSLAPGHVADFVLDDLRQHSVDLRYVVLQTEGSIPT
STVIINEASGSRTILHAYRNLPDVSAKDFEKVDLTRFKWIHIEGRNASEQVKMLQRIEEHNAKQPLPQKVRVSVEIEKPREELFQLFSYGEVVFVS
KDVAKHLGFQSAVEALRGLYSRVKKGATLVCAWAEEGADALGPDGQLLHSDAFPPPRVVDTLGAGDTFNASVIFSLSKGNSMQEALRFGCQVAGKK
CGLQGFDGIV

Genewiz KHK-A Protein Sequence: Translated from the blue sequence above (exon 3-encoded region underlined)

MEEKQILCVGLVVLDIINVVDKYPEEDTDRRCLSQRWQRGGNASNSCTVLSLLGARCAFMGSLAPGHVADFVLDDLRQHSVDLRYVVLQTEGSIPT
STVIINEASGSRTILHAYRNLPDVSAKDFEKVDLTRFKWIHIEGRNASEQVKMLQRIEEHNAKQPLPQKVRVSVEIEKPREELFQLFSYGEVVFVS
KDVAKHLGFQSAVEALRGLYSRVKKGATLVCAWAEEGADALGPDGQLLHSDAFPPPRVVDTLGAGDTFNASVIFSLSKGNSMQEALRFGCQVAGKK
CGLQGFDGIV

Figure 2S. Protein Sequence Alignment between mouse and human KHK-A and C.



Key:

Yellow highlight = active site residues
Green highlight = pfkB conserved motifs
|| = exon 3, alternative spicing region
***= same**
. = same between species (=different across species)
: = same between isozymes (=different across isozymes)
 # = all four are different
? = different, not included above

REFERENCES

Abdelmalek, M.F., Lazo, M., Horska, A., Bonekamp, S., Lipkin, E.W., Balasubramanyam, A., Bantle, J.P., Johnson, R.J., Diehl, A.M., Clark, J.M., and Fatty Liver Subgroup of Look, A.R.G. (2012) Higher Dietary Fructose Is Associated with Impaired Hepatic Adenosine Triphosphate Homeostasis in Obese Individuals with Type 2 Diabetes. *Hepatology* **56**, 952-960. 10.1002/hep.25741.

Adelman, R.C., Ballard, F.J., and Weinhouse, S. (1967) Purification and Properties of Rat Liver Fructokinase. *Journal of Biological Chemistry* **242**, 3360-3365.

Asipu, A., Hayward, B.E., O'Reilly, J., and Bonthron, D.T. (2003) Properties of Normal and Mutant Recombinant Human Ketohexokinases and Implications for the Pathogenesis of Essential Fructosuria. *Diabetes* **52**, 2426-2432.

Bais, R., James, H.M., Rofe, A.M., and Conyers, R.A. (1985) The Purification and Properties of Human Liver Ketohexokinase. A Role for Ketohexokinase and Fructose-Bisphosphate Aldolase in the Metabolic Production of Oxalate from Xylitol. *Biochemical Journal* **230**, 53-60.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242. 10.1093/nar/28.1.235.

Bonthron, D.T., Brady, N., Donaldson, I.A., and Steinmann, B. (1994) Molecular Basis of Essential Fructosuria: Molecular Cloning and Mutational Analysis of Human Ketohexokinase (Fructokinase). *Human Molecular Genetics* **3**, 1627-1631.

Bork, P., Sander, C., and Valencia, A. (1993) Convergent Evolution of Similar Enzymatic Function on Different Protein Folds: The Hexokinase, Ribokinase, and Galactokinase Families of Sugar Kinases. *Protein science : a publication of the Protein Society* **2**, 31-40. 10.1002/pro.5560020104.

Bradford, M.M. (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry* **72**, 248-254.

Cain, A.R., and Ryman, B.E. (1971) High Liver Glycogen in Hereditary Fructose Intolerance. *Gut* **12**, 929-932. 10.1136/gut.12.11.929.

Deng, Q., Wu, M., and Deng, J. (2022) Usp36 Promotes Tumor Growth of Non-Small Cell Lung Cancer Via Increasing Khk-a Expression by Regulating C-Myc-Hnrnph1/H2 Axis. *Human Cell* **35**, 694-704. 10.1007/s13577-022-00677-6.

Diggle, C.P., Shires, M., Leitch, D., Brooke, D., Carr, I.M., Markham, A.F., Hayward, B.E., Asipu, A., and Bonthron, D.T. (2009) Ketohexokinase: Expression and Localization of the Principal Fructose-Metabolizing Enzyme. *Journal of Histochemistry and Cytochemistry* **57**, 763-774.

Funari, V.A., Herrera, V.L.M., Freeman, D., and Tolan, D.R. (2005) Genes Required for Fructose Metabolism Are Expressed in Purkinje Cells in the Cerebellum. *Molecular Brain Research* **142**, 115-122.

Futatsugi, K., Smith, A.C., Tu, M., Raymer, B., Ahn, K., Coffey, S.B., Dowling, M.S., Fernando, D.P., Gutierrez, J.A., Huard, K., Jasti, J., Kalgutkar, A.S., Knafels, J.D., Pandit, J., Parris, K.D., Perez, S., Pfefferkorn, J.A., Price, D.A., Ryder, T., Shavnya, A., Stock, I.A., Tsai, A.S., Tesz, G.J., Thuma, B.A., Weng, Y., Wisniewska, H.M., Xing, G., Zhou, J., and Magee, T.V. (2020) Discovery of Pf-06835919: A Potent Inhibitor of Ketohexokinase (Khk) for the Treatment of Metabolic Disorders Driven by the Over-consumption of Fructose. *Journal of Medicinal Chemistry* **63**, 13546–13560. 10.1021/acs.jmedchem.0c00944.

Gasper, W.C. (2020). Characterization of Ketohexokinase as a Therapeutic Target for Hereditary Fructose Intolerance and Metabolic Syndrome. Ph.D. dissertaion (Boston University).

George, R.A., Spriggs, R.V., Bartlett, G.J., Gutteridge, A., MacArthur, M.W., Porter, C.T., Al-Lazikani, B., Thornton, J.M., and Swindells, M.B. (2005) Effective Function Annotation through Catalytic Residue Conservation. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12299-12304. 10.1073/pnas.0504833102.

Gibbs, A.C., Abad, M.C., Zhang, X., Tounge, B.A., Lewandowski, F.A., Struble, G.T., Sun, W., Sui, Z., and Kuo, L.C. (2010) Electron Density Guided Fragment-Based Lead Discovery of Ketohexokinase Inhibitors. *Journal of Medicinal Chemistry* **53**, 7979-7991. 10.1021/jm100677s.

Gutierrez, J.A., Liu, W., Perez, S., Xing, G., Sonnenberg, G., Kou, K., Blatnik, M., Allen, R., Weng, Y., Vera, N.B., Chidsey, K., Bergman, A., Somayaji, V., Crowley, C., Clasquin, M.F., Nigam, A., Fulham, M.A., Erion, D.M., Ross, T.T., Esler, W.P., Magee, T.V., Pfefferkorn, J.A., Bence, K.K., Birnbaum, M.J., and Tesz, G.J. (2021) Pharmacologic Inhibition of

Ketohexokinase Prevents Fructose-Induced Metabolic Dysfunction. *Molecular Metabolism* **48**, 101196. 10.1016/j.molmet.2021.101196.

Hansen, T., Arnfors, L., Ladenstein, R., and Schönheit, P. (2007) The Phosphofructokinase-B (Mj0406) from Methanocaldococcus Jannaschii Represents a Nucleoside Kinase with a Broad Substrate Specificity. *Extremophiles* **11**, 105-114. 10.1007/s00792-006-0018-1.

Hayward, B.E., and Bonthron, D.T. (1998) Structure and Alternative Splicing of the Ketohexokinase Gene. *European Journal of Biochemistry* **257**, 85-91.

Huard, K., Ahn, K., Amor, P., Beebe, D.A., Borzilleri, K.A., Chrunyk, B.A., Coffey, S.B., Cong, Y., Conn, E.L., Culp, J.S., Dowling, M.S., Gorgoglione, M.F., Gutierrez, J.A., Knafels, J.D., Lachapelle, E.A., Pandit, J., Parris, K.D., Perez, S., Pfefferkorn, J.A., Price, D.A., Raymer, B., Ross, T.T., Shavnya, A., Smith, A.C., Subashi, T.A., Tesz, G.J., Thuma, B.A., Tu, M., Weaver, J.D., Weng, Y., Withka, J.M., Xing, G., and Magee, T.V. (2017) Discovery of Fragment-Derived Small Molecules for in Vivo Inhibition of Ketohexokinase (Khk). *Journal of Medicinal Chemistry* **60**, 7835-7849. 10.1021/acs.jmedchem.7b00947.

Ishimoto, T., Lanaspa, M.A., Le, M.T., Garcia, G.E., Diggle, C.P., Maclean, P.S., Jackman, M.R., Asipu, A., Roncal-Jimenez, C.A., Kosugi, T., Rivard, C.J., Maruyama, S., Rodriguez-Iturbe, B., Sanchez-Lozada, L.G., Bonthron, D.T., Sautin, Y.Y., and Johnson, R.J. (2012) Opposing Effects of Fructokinase C and a Isoforms on Fructose-Induced Metabolic Syndrome in Mice. *Proceedings of the National Academy of Sciences of the U.S.A.* **109**, 4320-4325. 10.1073/pnas.1119908109.

Jensen, T., Abdelmalek, M.F., Sullivan, S., Nadeau, K.J., Green, M., Roncal, C., Nakagawa, T., Kuwabara, M., Sato, Y., Kang, D.H., Tolan, D.R., Sanchez-Lozada, L.G., Rosen, H.R., Lanaspa, M.A., Diehl, A.M., and Johnson, R.J. (2018) Fructose and Sugar: A Major Mediator of Non-Alcoholic Fatty Liver Disease. *Journal of Hepatology*. 10.1016/j.jhep.2018.01.019.

Johnson, R.J., Gold, M.S., Johnson, D.R., Ishimoto, T., Lanaspa, M.A., Zahniser, N.R., and Avena, N.M. (2011) Attention-Deficit/Hyperactivity Disorder: Is It Time to Reappraise the Role of Sugar Consumption? *Postgraduate Medicine* **123**, 39-49. 10.3810/pgm.2011.09.2458.

Johnson, R.J., Nakagawa, T., Sanchez-Lozada, L.G., Shafiu, M., Sundaram, S., Le, M., Ishimoto, T., Sautin, Y.Y., and Lanaspa, M.A. (2013) Sugar, Uric Acid,

and the Etiology of Diabetes and Obesity. *Diabetes* **62**, 3307-3315. 10.2337/db12-1814.

Johnson, R.J., Perez-Pozo, S.E., Sautin, Y.Y., Manitius, J., Sanchez-Lozada, L.G., Feig, D.I., Shafiu, M., Segal, M., Glassock, R.J., Shimada, M., Roncal, C., and Nakagawa, T. (2009) Hypothesis: Could Excessive Fructose Intake and Uric Acid Cause Type 2 Diabetes? *Endocrine Reviews* **30**, 96-116. 10.1210/er.2008-0033.

Kim, J., Kang, J., Kang, Y.L., Woo, J., Kim, Y., Huh, J., and Park, J.W. (2020) Ketohexokinase-a Acts as a Nuclear Protein Kinase That Mediates Fructose-Induced Metastasis in Breast Cancer. *Nature Communications* **11**, 5436. 10.1038/s41467-020-19263-1.

Kim, S.H., Witte, C.P., and Rhee, S. (2021) Structural Basis for the Substrate Specificity and Catalytic Features of Pseudouridine Kinase from Arabidopsis Thaliana. *Nucleic Acids Research* **49**, 491-503. 10.1093/nar/gkaa1144.

Lanaspa, M.A., Andres-Hernando, A., Orlicky, D.J., Cicerchi, C., Jang, C., Li, N., Milagres, T., Kuwabara, M., Wempe, M.F., Rabinowitz, J.D., Johnson, R.J., and Tolan, D.R. (2018) Ketohexokinase C Blockade Ameliorates Fructose-Induced Metabolic Dysfunction in Fructose-Sensitive Mice. *Journal of Clinical Investigation* 10.1172/JCI94427.

Lanaspa, M.A., Sanchez-Lozada, L.G., Choi, Y.J., Cicerchi, C., Kanbay, M., Roncal-Jimenez, C.A., Ishimoto, T., Li, N., Marek, G., Duranay, M., Schreiner, G., Rodriguez-Iturbe, B., Nakagawa, T., Kang, D.H., Sautin, Y.Y., and Johnson, R.J. (2012) Uric Acid Induces Hepatic Steatosis by Generation of Mitochondrial Oxidative Stress: Potential Role in Fructose-Dependent and -Independent Fatty Liver. *The Journal of biological chemistry* **287**, 40732-40744. 10.1074/jbc.M112.399899.

Li, X., Qian, X., Peng, L.X., Jiang, Y., Hawke, D.H., Zheng, Y., Xia, Y., Lee, J.H., Cote, G., Wang, H., Wang, L., Qian, C.N., and Lu, Z. (2016) A Splicing Switch from Ketohexokinase-C to Ketohexokinase-a Drives Hepatocellular Carcinoma formation. *Nature Cell Biology* **18**, 561-71. 10.1038/ncb3338.

Mathews, II., Erion, M.D., Ealick, S.E. (1998) Structure of Human Adenosine Kinase at 1.5 Å Resolution. *Biochemistry* 37(45): 15607-20. 10.1021/bi9815445.

Mirtschink, P., Krishnan, J., Grimm, F., Sarre, A., Hörl, M., Kayikci, M., Fankhauser, N., Christinat, Y., Cortijo, C., Feehan, O., Vukolic, A., Sossalla, S., Stehr, S.N., Ule, J., Zamboni, N., Pedrazzini, T., and Krek, W. (2015) Hif-

Driven Sf3b1 Induces Khk-C to Enforce Fructolysis and Heart Disease. *Nature* **522**, 444-449. 10.1038/nature14508.

Nakagawa, T., Lanaspa, M.A., Millan, I.S., Fini, M., Rivard, C.J., Sanchez-Lozada, L.G., Andres-Hernando, A., Tolan, D.R., and Johnson, R.J. (2020) Fructose Contributes to the Warburg Effect for Cancer Growth. *Cancer & Metabolism* **8**, 16. 10.1186/s40170-020-00222-9.

Ouyang, X., Cirillo, P., Sautin, Y., McCall, S., Bruchette, J.L., Diehl, A.M., Johnson, R.J., and Abdelmalek, M.F. (2008) Fructose Consumption as a Risk Factor for Non-Alcoholic Fatty Liver Disease. *Journal of Hepatology* **48**, 993-999.

Park, J., and Gupta, R.S. (2008) Adenosine Kinase and Ribokinase--the Rk Family of Proteins. *Cellular and Molecular Life Sciences* **65**, 2875-2896. 10.1007/s00018-008-8123-1.

Raushel, F.M., and Cleland, W.W. (1973) The Substrate and Anomeric Specificity of Fructokinase. *The Journal of biological chemistry* **248**, 8174-8177.

Sanchez-Lozada, L.G., Lanaspa, M.A., Cristobal-Garcia, M., Garcia-Arroyo, F., Soto, V., Cruz-Robles, D., Nakagawa, T., Yu, M.A., Kang, D.H., and Johnson, R.J. (2012) Uric Acid-Induced Endothelial Dysfunction Is Associated with Mitochondrial Alterations and Decreased Intracellular Atp Concentrations. *Nephron. Experimental Nephrology* **121**, e71-78. 10.1159/000345509.

Sanchez-Lozada, L.G., Mu, W., Roncal, C., Sautin, Y.Y., Abdelmalek, M., Reungjui, S., Le, M., Nakagawa, T., Lan, H.Y., Yu, X., and Johnson, R.J. (2010) Comparison of Free Fructose and Glucose to Sucrose in the Ability to Cause Fatty Liver. *Eur J Nutr* **49**, 1-9. 10.1007/s00394-009-0042-x.

Sheiham, A., and James, W.P. (2015) Diet and Dental Caries: The Pivotal Role of Free Sugars Reemphasized. *Journal of Dental Research* **94**, 1341-1347. 10.1177/0022034515590377.

Shen, Z., Li, Z., Liu, Y., Li, Y., Feng, X., Zhan, Y., Lin, M., Fang, C., Fang, Y., and Deng, H. (2022) Glut5-Khk Axis-Mediated Fructose Metabolism Drives Proliferation and Chemotherapy Resistance of Colorectal Cancer. *Cancer Letters* **534**, 215617. 10.1016/j.canlet.2022.215617.

Schumacher , M.A., Scott D.M., Mathews I.I., Ealick S.E., Roos D.S., Ullman B., Brennan R.G. (2000) Crystal Structures of Toxoplasma gondii Adenosine Kinase Reveal a Novel Catalytic Mechanism and Prodrug Binding. *Journal of Molecular Biology* **296**(2): 549-67. 10.1006/jmbi. 1999.3474.

Sigrell, J.A., Cameron, A.D., Jones, T.A., and Mowbray, S.L. (1997) Purification, Characterization, and Crystallization of Escherichia Coli Ribokinase. *Protein science : a publication of the Protein Society* **6**, 2474-2476. 10.1002/pro.5560061124.

Sigrell, J.A., Cameron, A.D., Jones, T.A., and Mowbray, S.L. (1998) Structure of Escherichia Coli Ribokinase in Complex with Ribose and Dinucleotide Determined to 1.8 a Resolution: Insights into a New Family of Kinase Structures. *Structure* **6**, 183-193. 10.1016/s0969-2126(98)00020-3.

Sigrell, J.A., Cameron, A.D., and Mowbray, S.L. (1999) Induced Fit on Sugar Binding Activates Ribokinase. *Journal of Molecular Biology* **290**, 1009-1018. 10.1006/jmbi.1999.2938.

Stirpe, F., Della Corte, E., Bonetti, E., Abbondanza, A., Abbati, A., and De Stefano, F. (1970) Fructose-Induced Hyperuricaemia. *Lancet* **2**, 1310-1311.

Thurston, J.H., Jones, E.M., and Hauhart, R.E. (1974) Decrease and Inhibition of Liver Glycogen Phosphorylase after Fructose. An Experimental Model for the Study of Hereditary Fructose Intolerance. *Diabetes* **23**, 597-604. 10.2337/diab.23.7.597.

Titgemeyer, F., Reizer, J., Reizer, A., and Saier, M.H., Jr. (1994) Evolutionary Relationships between Sugar Kinases and Transcriptional Repressors in Bacteria. *Microbiology (Reading)* **140 (Pt 9)**, 2349-2354. 10.1099/13500872-140-9-2349.

Tremmel, M., Gerdtham, U.G., Nilsson, P.M., and Saha, S. (2017) Economic Burden of Obesity: A Systematic Literature Review. *International Journal of Environmental Research and Public Health* **14**. 10.3390/ijerph14040435.

Trinh, C.H., Asipu, A., Bonthron, D.T., and Phillips, S.E. (2009) Structures of Alternatively Spliced Isoforms of Human Ketohexokinase. *Acta Crystallographica. Section D, Biological Crystallography* **65**, 201-211.

Weiser, M.M., and Quill, H. (1975) Estimation of Fructokinase (Ketohexokinase) in Crude Tissue Preparations. *Methods in Enzymology* **41**, 61-63.

Xu, D., Li, X., Shao, F., Lv, G., Lv, H., Lee, J.H., Qian, X., Wang, Z., Xia, Y., Du, L., Zheng, Y., Wang, H., Lyu, J., and Lu, Z. (2019) The Protein Kinase Activity of Fructokinase a Specifies the Antioxidant Responses of Tumor Cells by Phosphorylating P62. *Science Advances* **5**, eaav4570. 10.1126/sciadv.aav4570.

Yang, J., Yang, S., Wang, Q., Pang, J., Wang, Y., Wang, H., and Fu, X. (2021) Khk-a Promotes the Proliferation of Oesophageal Squamous Cell Carcinoma through the up-Regulation of Prps1. *Arab Journal of Gastroenterology* **22**, 40-46. 10.1016/j.ajg.2020.08.007.

Yau, E.H., Kummetha, I.R., Lichinchi, G., Tang, R., Zhang, Y., and Rana, T.M. (2017) Genome-Wide Crispr Screen for Essential Cell Growth Mediators in Mutant Kras Colorectal Cancers. *Cancer Research* **77**, 6330-6339. 10.1158/0008-5472.Can-17-2043.

Zamora-Leon, S.P., Golde, D.W., Concha, II, Rivas, C.I., Delgado-Lopez, F., Baselga, J., Nualart, F., and Vera, J.C. (1996) Expression of the Fructose Transporter Glut5 in Human Breast Cancer. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 1847-1852. 10.1073/pnas.93.5.1847.

Zhang, Y., Dougherty, M., Downs, D.M., and Ealick, S.E. (2004) Crystal Structure of an Aminoimidazole Riboside Kinase from Salmonella Enterica: Implications for the Evolution of the Ribokinase Superfamily. *Structure* **12**, 1809-1821. 10.1016/j.str.2004.07.020.

CURRICULUM VITAE