ABSTRACT

Title of Document: TRANSCRIPTIONAL REGULATION OF N-ACETYLGLUTAMATE SYNTHASE AND ITS CLINICAL RELEVANCE

Sandra Kirsch Heibel, Ph.D, 2012

Directed By: Professor Mendel Tuchman, Department of Pediatrics, Children's National Medical Center, George Washington University
Professor David Mosser, Department of Cell Biology and Molecular Genetics

The urea cycle converts ammonia, the toxic byproduct of protein metabolism, to non-toxic urea utilizing six enzymes in mammalian liver hepatocytes. In the mitochondria, the enzyme N-acetylglutamate synthase (NAGS), produces the essential allosteric activator, N-acetylglutamate (NAG), for the first enzyme of the cycle, carbamylphosphate synthetase 1 (CPS1). Ammonia sequestered by CPS1 is condensed with bicarbonate and phosphate from ATP to produce to carbamylphosphate, which is then condensed with ornithine to produce citrulline by ornithine transcarbamylase (OTC). Citrulline is then transported to the cytoplasm where it is further converted by the distal urea cycle enzymes argininosuccinate synthetase, argininosuccinate lyase, and arginase 1 to the end product, urea.

Prior to sequestration of ammonia, signaling pathways sense dietary protein load. Analysis was conducted to determine the effect of protein composition on metabolic signaling and nutrient sensing pathways AMPK, mTOR, and eIF2, promote

transcriptional activation of urea cycle genes and translation of urea cycle proteins. Future studies will determine the importance of these pathways on the urea cycle and whether post-translational control mechanisms, such as phosphorylation, also respond to protein in the diet.

Since NAGS plays an important role in controlling the rate of urea production by activating the rate limiting CPS1, its regulatory mechanisms control ureagenesis flux. This project elucidated the regulatory domains of *NAGS* including a promoter which contains multiple tissue- and species- specific transcription initiation sites and binding sites for transcription factors CREB and Sp1. It also found and characterized an enhancer 3kb upstream of the translational start site which confers liver specificity and has binding sites for NF-Y and HNF-1. These transcription factors are regulated by glucocorticoid and glucagon hormone signaling pathways and are also regulated in a tissue selective manner.

A deficiency of NAGS in humans and mice leads to high ammonia, low urea, and high glutamine levels in the plasma, and can be overcome by treatment with N-carbamylglutamate (NCG). This project identified a patient with NAGS deficiency and hyperammonemia caused by a deleterious mutation in the HNF-1 binding site within the enhancer of the NAGS gene, which was identified and characterized by this research. This mutation caused decreased transcription of NAGS due to reduced HNF-1 binding. Subsequently, this work found disease-causing mutations in the promoter region of OTC which have been shown to interfere with HNF-4 binding.

The knowledge garnered by this project significantly increases our understanding of the regulation of urea cycle RNA and protein expression and their role in disease

pathophysiology. Understanding of these mechanisms will lead to improved diagnoses and continued development of effective treatments for people with urea cycle disorders and hyperammonemia.

TRANSCRIPTIONAL REGULATION OF N-ACETYLGLUTAMATE SYNTHASE AND ITS CLINICAL RELEVANCE

By

Sandra Kirsch Heibel

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor David Mosser, Chair
Professor Mendel Tuchman, Co-Chair
Associate Professor Ljubica Caldovic
Assistant Professor Kenneth Frauwirth
Professor Tom Porter, Dean's Representative

Dedication

To my husband TJ and to our future.

Acknowledgements

First, I would like to thank my advisors Mendel Tuchman and Ljubica Caldovic. Mendel, as your first graduate student, I am proud to say that you have excelled as a mentor. Your guidance has shaped my scientific thinking and led me to develop a passion for translational medicine that I intend to carry into my future career. Ljuba, you have been a wonderful mentor and friend. Through your dedication, optimism, and patience, you have not only taught me to think critically, but have given me the confidence to pursue my research interests while keeping the day to day failings of experiments and negative reviewers comments in perspective. Your future graduate students and technicians will be lucky to have you as their mentor.

I thank my committee members David Mosser, Leslie Pick, and Kenneth Frauwirth for all your suggestions, encouragement, and support. Your councel ensured the applicability of my research to a large audience while your enthusiasm for my progress has given me the confidence to leave the safety of graduate school for a “real job”. I would like to thank Tom Porter for serving as my Dean’s Rep and for always taking the time to chat with me about the Animal Science department, our research interests, and life in general.

I thank my lab mates for their incredible support over the years. Hiroki, your encyclopedic knowledge of just about everything has made you a wonderful resource. Abbe, you are amazing at everything from trouble shooting an experiment to planning a happy hour. I know you will be a great success in your future endeavors. Net, Emilee, David, Juan, Himani, Elena, Amy, Dina, Alex, Marc and all my friends in Gen. Med. thank you for sharing your knowledge, sympathy, and laughter with me over the years.

I have benefitted from the support and friendship of so many people at the University of Maryland. I especially want to thank my friends Amanda, Jen, Lisa, and Trinity. When I think of these last few years at UMD, you are part of the best memories. I am glad our shared miseries and joys from graduate school have brought us together, and I look forward to sharing all the highs and lows of our future careers as well.

My family deserves an extra special thank you, you are the rocks in my life and I could not have done any of this without you. Mom and Dad, thank you so much for your love, for always encouraging me to work hard and, in retrospect, for telling me I could not move home after college and asking me when I planned to graduate. Your emotional (and time to time, financial) support means so much to me, and I look forward to continuing to make you proud. Stephanie, your dedication to your work and family has always been an inspiration to me. Thank you for all of your advice over the years, and for reminding me that the big picture and shoe shopping are key to being a well-rounded person. Jeff, thank you for genuinely taking the time to ask about my research, I look forward to challenging you with completely new research topics in the future. Wesley, you are a wonderful little person and I look forward to seeing you turn into a fellow biology lover. Allison, from PSU to UMD you have been like a sister to me. I am so lucky to have you in my life for your friendship, your council, and your delicious delicious baked goods. Lily, thank you for convincing me that you belong in my lap instead of the computer. You and graduate school came into my life together and I am happy you and I will continue the journey together.

Finally, the biggest thank you goes to my husband TJ for every big and little thing you do. We've been through so much together, and I can't wait to see what's next!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAP- antarctic alkaline phosphatase

ACC- acetyl CoA carboxylase

AICAR- 5-aminoimidazole-4-carboxamide-1-beta-4-ribofuranoside

AMPK- adenosine 5-monophosphate-activated protein kinase

AP-2- activating protein-2

ASS- argininosuccinate synthase

ASL- argininosuccinate lyase

ARG1- arginase 1

BCAA- branched chain amino acid

BUN- blood urea nitrogen

C/EBP- CAAAT enhancer binding protein

CLOVER- cis-element overrepresentation

CPS1- carbamylphosphate synthetase 1

CREB- cyclic AMP response element binding

CTF- CCAAT-binding transcription factor

dbSNP- database of single nucleotide polymorphisms

eIF2/GCN2- eukaryotic translation initiation factor-2/ general control nondepressible 2 kinase

FASN- fatty acid synthase

FBS- fetal bovine serum

GenMAPP- Gene Mapp Annotator and Pathway Profiler

GR- glucocorticoid receptor

GSK-3β- glycogen synthase kinase 3-beta

GYS- glycogen synthase

HCE- Hierarchical Clustering Explorer

HepG2- human hepatoma cell line G2

HGDP- Human Genome Diversity Panel

HMGCR- hydroxymethylglutaryl-CoA reductase

HNF- hepatic nuclear factor

IPA- Ingenuity Pathway Analysis

MAS- Microarray Suite

MIM- Mendelian inheritance in Man

mTOR- mammalian target of rapamycin

NAG- N-acetylglutamate

NAGS- N-acetylglutamate synthase

NAGSD- N-acetylglutamate synthase deficiency

NCG – N-carbamyl-L-glutamate

NF-1- nuclear factor 1

NF-Y- nuclear factor Y

OTC- ornithine transcarbamylase

OTCD- ornithine transcarbamylase deficiency

P70S6K- kDa ribosomal protein S6 kinase

PABP- poly(A) binding protein

PCR- polymerase chain reaction

PDK-1- 3-phosphoinositide dependent protein kinase-1

PGC-1α- peroxisome proliferator-activated receptor gamma co-activator

PI3K- phosphatidyinositol-3-kinase

PICC – peripheral inserted central catheter

PKA- protein kinase A

PP2A- protein phosphatase 2A

PP2C- protein phosphatase 2C

RPS6- ribosomal protein S6

SMAD2/3- mothers against decapentaplegic homolog 2 or 3

Sp1- specificity protein 1

TBP- TATA binding protein

TSC1 & 2- tuberous sclerosis 1 & 2

UCD – urea cycle disorders

UCDC – Urea Cycle Disorders Consortium

VDR- vitamin D receptor

ZMP- 5-amino-4-imidazolecarboxamide ribotide

Chapter 1: Introduction

Urea Cycle Function and Regulation

The urea cycle, discovered by Hans Krebs and Kurt Henseleit in 1932 [1], is essential for the protection of the central nervous system from the toxic effects of ammonia. Accumulate ammonia in the blood can cause lethargy, coma and death and can lead to permanent brain damage and intellectual and developmental disabilities. As the by-product of protein catabolism, ammonia is converted to non-toxic urea in mammalian liver hepatocytes through six enzymes and two mitochondrial membrane transporters. One of the enzymes, N-acetylglutamate synthase (NAGS), is found in the mitochondria of liver hepatocytes and small intestine epithelial cells. NAGS catalyzes the formation of N-acetylglutamate (NAG) from glutamate and acetyl CoA, producing the obligatory cofactor (NAG) for the first enzyme of the urea cycle, carbamylphosphate synthase 1 (CPS1). Following the binding of NAG, mitochondrial CPS1 sequesters ammonia, the first of two waste nitrogen molecules destined to make urea. Ammonia, bicarbonate and phosphate from ATP are condensed by CPS1 into carbamyl phosphate. The next enzyme in the cycle, ornithine transcarbamylase (OTC), forms citrulline from carbamyl phosphate and ornithine in the mitochondria. Aspartate, the source of the second waste nitrogen, and citrulline are then transported out of the mitochondria to the cytosol by the ornithine transporter (ORNT) and citrin, respectively. In the cytoplasm, the next enzyme, argininosuccinate synthase (ASS), condenses citrulline and aspartate to produce argininosuccinate. Argininosuccinate lyase then converts argininosuccinate to arginine and fumarate, arginine is then broken down by Arginase 1 (ARG1) to urea and ornithine.

Urea is excreted from the body through the kidneys and ornithine is transported back into the mitochondria by the ornithine transporter to be reused by the urea cycle (Figure 1).

In 1939, one of the earliest studies on metabolic adaptations to diet examined urea cycle enzyme activity and showed that arginase activity changed in response to diet [2]. Since then, studies have shown that urea production is regulated by dietary protein intake such that it increases following consumption of a high protein diet, decreases following consumption of a low protein diet, and increases during acute starvation due to increased muscle catabolism and production of endogenous nitrogen waste. This mechanism regulates the activity of CPS1, OTC, ASS, ASL and ARG1 [3,4,5] by controlling the rate of enzyme synthesis [6,7]. Early studies did not examine the regulation of NAGS as it is of very low abundance and is difficult to assay [8]. Later studies, in 1977, examined the regulation of NAGS activity and found that its activity also increases in response to diet [9].

Precursrs and metabolites of the urea cycle have been shown to regulate activity of the individual enzymes. In addition to dietary protein, NAGS activity is increased by arginine, ammonia, and glutamate [10,11,12]. Activity of CPS1 is completely dependent on the availability of NAG and is stimulated by ammonia and ornithine [13,14,15,16]. OTC activity is dependent on the abundance of its substrate, carbamyl phosphate, and is inhibited by an excess of citrulline [13,14,17]. ASS activity is dependent on the abundance of citrulline and is stimulated by ammonia as well as aspartate [7,14,18]. Finally, the abundance of arginine mediates repression of ASS and ASL and limits the activity of ARG1 [2,6,14].

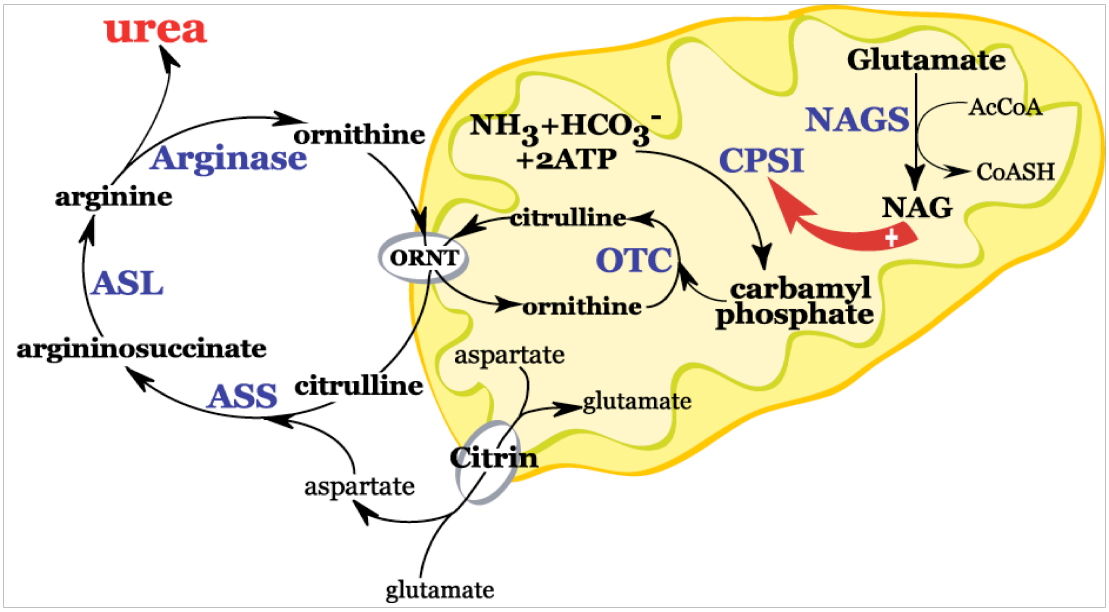

Figure 1. The urea cycle.

The urea cycle system converts waste nitrogen from ammonia and aspartate to urea using six enzymes (blue) and two transporters (in circles) in the mammalian liver.

Short-term regulation of urea production is most likely based on both substrate availability and post-translational processing of urea cycle enzymes considering their capacity for ureagenesis exceeds the physiological rate of urea production [19]. CPS1 is the rate-limiting enzyme for short-term ureagenesis and its activity is directly related to the production of its allosteric activator, NAG [12,20]. Early studies of ureagenesis failed to show a short-term increase in NAGS activity in response to increased amino acid load and suggested, instead, that the increased production of NAG was due to an increase in hepatic glutamate [12,21]. While glutamate substrate load may be a major regulator of NAG production, it seems unlikely that regulation of the complex interactions required to maintain nitrogen homeostasis are dependent only on substrate concentration [22], but in addition may involve a combination of amino acid sensing to stimulate gene transcription [23] and post-translation processing of urea cycle proteins [24,25,26,27,28,29].

Amino Acid Sensing

Hepatic metabolism of fats, carbohydrates, and proteins is kept in balance by complex interactions between the regulatory networks that sense nutrients and the pathways that utilize them. These nutrient sensing pathways include: 1. the adenosine 5'-monophosphate-activated protein kinase (AMPK) pathway, which acts as an intracellular energy sensor [30,31], 2. the mammalian target of rapamycin (mTOR) pathway, which stimulates translation [32], and 3. the eukaryotic initiation factor 2 (eIF2)/general control nondepressible kinase 2 (GCN2) pathway, which inhibits translation upon sensing of amino acid deprivation [32,33,34]. The mechanisms that allow hepatocytes to adapt to changes in dietary carbohydrate and protein composition are in part signaled by amino acids [23,35]. As such, in cultured primary hepatocytes, phosphorylation of AMPK,

GCN2, and the downstream target of GCN2, eukaryotic initiation factor 2α (eIF2α), decreased [23] while phosphorylation of mTOR and the downstream targets of the mTOR kinase, eukaryotic initiation factor 4E-binding protein-1 (eIF4E-BP1) and p70 S6 kinase (p70S6K) increased in response to a high protein diet and particularly leucine [23,36]. Targeted manipulation of these pathways by rapamycin, inhibitor of mTOR [35,37], and 5-aminoimidazole-4-carboxamide riboside (AICAR), activator of AMPK [38,39,40], can be used to discern the role these pathways play in expression of urea cycle genes and enzymes.

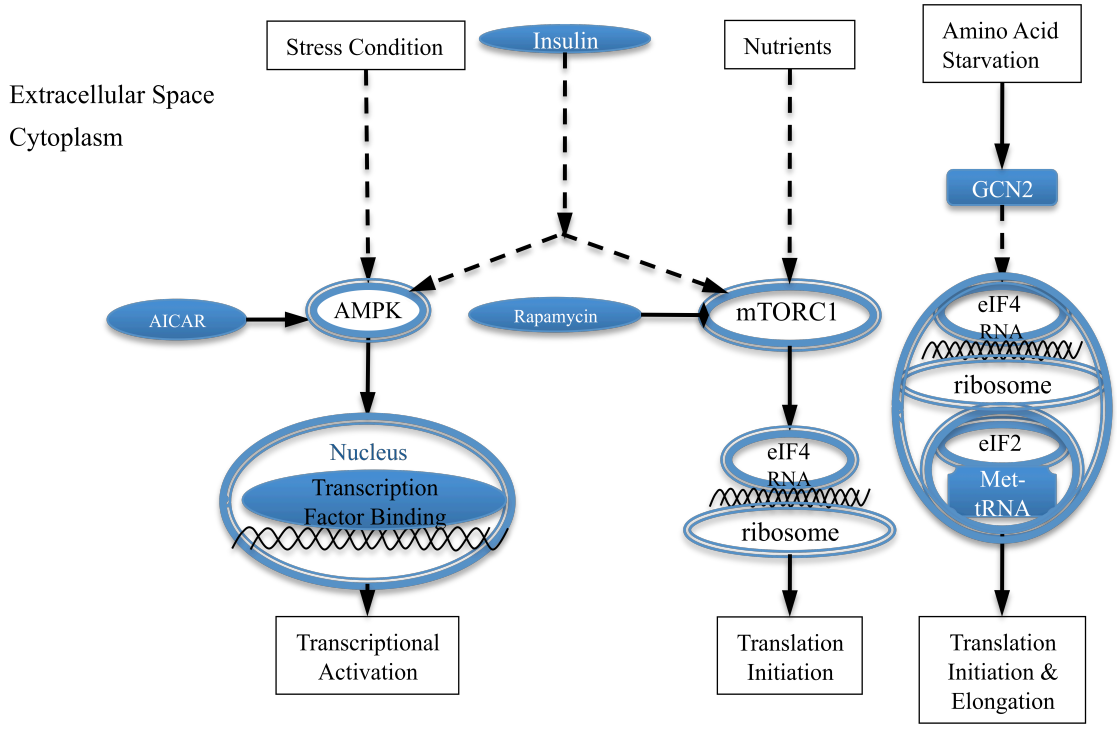

Figure 2. Signaling pathways involved in nutrient sensing.

The AMPK, mTOR, and eIF2 pathways regulate transcription and translation in response to signals from the diet.

Transcriptional Regulation of Urea Cycle Genes

Regulation of the urea cycle genes is mediated, at least in part, by hormones and dietary amino acid concentration. In general, glucocorticoids increase enzyme activity, and glucagon, or analogs of its second messenger, cAMP, increase mRNA levels in response to fasting [41]. The mechanisms for transcriptional regulation are not fully understood, but previous studies of the promoter and enhancer regions of C*PS1, OTC, ASS, ASL,* and *ARG1* have identified sites where hormone responsive transcription factors may bind and differentially regulate these genes. Table 1 lists the transcription factors shown to bind to these sites and regulate each gene's expression. The DNA sequence of *NAGS* was identified in 2002 by Caldovic, *et al.* [42,43] and our elucidation of the transcriptional regulation of *NAGS* has recently been published [44] and is presented in chapter 2. In summary, the promoter of *NAGS* has binding sites for cyclic AMP responsive element binding protein (CREB) and multiple specificity protein 1 (Sp1) sites, whereas hepatic nuclear factor 1 (HNF-1) and nuclear factor Y (NF-Y) bind to the upstream enhancer. A naturally occurring mutation of the HNF-1 binding site in the *NAGS* enhancer of a patient with NAGS deficiency resulted in the inhibition of transcription factor binding and decreased transcription of *NAGS* [44]. Transcription of *CPS1* is activated by TATA-binding protein (TBP). The proximal and distal enhancers of the *CPS1* gene contain binding sites for glucocorticoid response units including CCAAT-enhancer binding protein (C/EBP), activator protein-1 (AP-1), glucocorticoid receptor (GR), CREB, and tissue selective factor binding sites for hepatic nuclear factor 3 (HNF-3) [45,46,47]. The *OTC* gene is expressed in a tissue specific manner such that activation of part of the promoter induces intestinal expression [48,49] while activation of

the enhancer induces liver expression [50,51,52]. The promoter contains binding sites for the liver selective hepatic nuclear factor 4 (HNF-4) [53,54] and repressor chicken ovalbumin upstream promoter-transcription factor (COUP-TF) [51,55]. The enhancer requires two transcription factors, HNF-4 and the glucocorticoid response factor C/EBP, for regulation of TBP activated transcription [47]. Study of the regulation of *ASS*, has identified multiple Sp1 binding sites in the promoter, which have been shown to cooperatively activate transcription; multiple Sp1 binding sites have been identified on the promoters of *ASL* and *ARG1* as well [56,57]. *ASS* is also regulated through cAMP-dependent phosphorylation of CREB [18,58] and AP-2 may play a role in the arginine-mediated repression of this gene [59]. Within the promoter of *ASL*, Sp1 activates transcription, but it is unknown whether the binding of nuclear factor Y (NF-Y) acts to activate or repress transcription [60,61,62]. Sp1 and nuclear factor 1 (NF-1)/CCAAT-binding transcription factor (CTF), both of which recognize the same region of DNA and bind in a mutually exclusive manner, can activate *ARG1* transcription [63]. An enhancer of *ARG1* resides approximately 11kb downstream of the translational start site, within intron 7, where two C/EBP factors bind along with two unidentified proteins to confer glucocorticoid responsiveness [63,64].

Table 1. Transcription factors that regulate urea cycle genes.

Gene	Transcription Activators	Transcription Factors	
		Proximal Enhancer	Distal enhancer
NAGS	Sp1	CREB, FXR	HNF-1, NF-Y
CPS1	TBP	C/EBP	CREB, C/EBP, GR, HNF-3, AP-1, P1-3
OTC	TBP	HNF-4, COUP-TF	C/EBP, HNF-4/COUP-TF
ASS	Sp1	AP-2	CREB
ASL	Sp1	NF-Y	
ARG1	Sp1, NF-1/CTF	C/EBP	C/EBP, P4, P5

As mentioned in the previous section, the nutrient sensors' (mTOR, AMPK, and eIF2) phosphorylation state changes in response to dietary amino acid intake. One of the downstream targets of mTOR and AMPK, peroxisome proliferator-activated receptor gamma co-activator (PGC-1α) [65,66,67], regulates the expression of HNF-4 and GR [68,69], which in turn directly regulate the expression of *OTC* and *CPS1* [55,70,71,72,73]. Additional transcription factors, including C/EBP and CREB, require phosphorylation by the AMPK pathway members [74,75] to activate transcription of *NAGS*, *CPS1*, *OTC*, *ASS*, and *ARG1* [47,63,73,76,77]. While HNF-1, regulator of *NAGS*, and HNF-3, regulator of *CPS1*, are not direct downstream targets of these pathways, HNF-1, -3, -4, and -6 regulate each other via complex feedback mechanisms. While HNF-6 does not bind to the regulatory regions of urea cycle genes, it binds to the promoter of HNF-4 [53,78] and simple feedback loops exist between HNF-1 and HNF-4 [79,80,81] and between HNF-3 and HNF-6 [82,83]. In addition, many genes involved in hepatocyte metabolism exhibit multi-input regulation by a combination of two or more of the HNFs and synergistic activation through their interactions [53,81].

Hormone Signaling

Many of the transcription factors described in the previous section are in part regulated by hormones. In fact, studies by Morris *et. al*, in the 1980s and 1990s, showed that transcription of *CPS1, OTC, ASS, ASL* and *ARG1* was coordinately regulated in response to cAMP, glucocorticoids, and insulin [19,47,84]. Glucocorticoids increase enzyme activity of the urea cycle enzymes, and while they have little effect on the gene expression of *OTC*, transcription rates of *CPS1* and *ASL* are stimulated by glucocorticoids in hepatocytes [41,85]. Recently, a transcriptional profiling study was

conducted in rats to examine the stimulatory effect of the glucocorticoid analog methylprednisolone on the urea cycle genes [86]. The authors concluded that glucocorticoid regulation of mRNA expression differs between urea cycle enzymes. *CPS1, OTC,* and *ARG1* exhibit an acute response while *ASS* and *ASL* have a delayed response to glucocorticoids [86]. This difference in response may be explained by the glucocorticoid ability to stimulate or inhibit gene expression based on the regulatory proteins and transcription factors that are already present on the DNA control regions [86].

Transcription of *CPS1*, *ASS*, and *ASL* is rapidly stimulated by glucagon and cAMP analogs in rat liver [85] while transcription of *OTC* fails to be stimulated on a short-term time scale but is induced following 48 hour incubation with glucagon [41,84,87]. While the glucocorticoid response mechanism is still unknown and most likely mediated by proteins that are synthesized in response to the hormone [41], cAMP stimulation of mRNA expression does not require ongoing protein synthesis in rat hepatocytes [41]. Following short-term fasting or feeding of a low protein diet *in vivo*, an increase in glucagon activates the cAMP pathway and stimulates CREB phosphorylation by protein kinase A (PKA1) [88]. This results in diffusion of phosphorylated CREB into the nucleus and increased transcriptional activation of genes with CREB binding sites [89,90]. This mechanism may in part regulate cAMP stimulation of the urea cycle enzymes with known CREB binding sites, namely *NAGS*, *CPS1* and *ASS* [46,76]. A combination of glucocorticoids and cAMP analogs produced large increases in transcription rates for all of the urea cycle enzymes except OTC in rats [85].

Expression of the urea cycle genes is differentially regulated in response to treatment with insulin and amino acids [21]. No significant difference in urea production was noted when hepatocytes were treated with high or low insulin alone when no additional source of nitrogen was presented to the cells. However, treatment with low levels of insulin in the presence of amino acids optimized hepatocellular function and increased urea synthesis due to increased expression of urea cycle enzymes while high levels of insulin and amino acids decreases enzyme expression [21]. In contrast to the noted induction of urea cycle enzyme expression by glucagon and cAMP, insulin has been shown to play an inhibitory role on expression of urea cycle genes in hepatocytes. [21,85]. Thus, urea cycle genes are differentially regulated in response to hormone signaling and each of the urea cycle enzymes may be subject to a combination of glucocorticoid, cAMP/glucagon, and insulin signaling.

N-Acetylglutamate Synthase Deficiency

A deficiency of any of the cycle enzymes can lead to a urea cycle disorder (UCD) manifested clinically as hyperammonemia. NAGS deficiency (NAGSD) (OMIM 237310) can result in the reduction or absence of NAG causing reduced or absent CPS1 activity. The resulting hyperammonaemia can cause vomiting, seizures, mental retardation, coma, and possibly death [91,92,93]. The phenotype of NAGSD is indistinguishable from CPS1 deficiency (CPS1D) based on the clinical symptoms described above, and biochemical features include elevated plasma ammonia and glutamine and reduced urea and citrulline levels [94]. Since the enzymatic diagnosis of NAGSD requires an open liver biopsy and the assay can be inaccurate, this disorder is best diagnosed by sequencing of the coding and non-coding regions of the gene [95]. The first suspected case of NAGSD was

described in 1981 [96]. Several patients had been reported subsequently, but their diagnosis could not be confirmed until the NAGS DNA sequence was identified in our laboratory. To date, twenty-two disease causing mutations in the *NAGS* coding sequence or intron/exon boundries [94] and one disease causing mutation in the enhancer region of *NAGS* [95] have been reported. The severity of this autosomal recessive disease seems to depend on the genotype. Patients with nonsense or frameshift mutations that lead to absence of the NAGS protein present in the neonatal period [97,98,99] while patients with hypomorphic alleles such as missense mutations that allow for some residual NAGS enzyme activity typically present after the neonatal period [95,100,101,102,103,104,105] and sometimes during adulthood (late onset) [93,106,107].

During the first clinical presentation of any UCD, the initial focus of treatment is to lower plasma ammonia [108] by using scavenging agents such as sodium benzoate, sodium phenylacetate, or sodium phenylbutyrate [109], antibiotics to reduce nitrogen-producing gut bacterial load, and dialysis to acutely remove excess ammonia [110,111]. N-carbamylglutamate synthase (NCG), a non-hydrolyzed glutamic acid derivative [112] that can activate CPS1 [96] in place of NAGS, is the only specific treatment for any UCD and it effectively restores ureagenesis and reverses the hyperammonemia in NAGSD. Administration of NCG during acute hyperammonemia crisis has rescued patients [113,114] and long-term treatment with NCG completely corrects the urea cycle block such that patients may consume liberalized dietary protein [114,115,116] without additional ammonia scavenging agents [98,102,113,114,117,118]. Currently, NCG is the only drug, for treatment of a UCD, in which the disease-causing enzyme can be by-passed and the remaining urea cycle enzymes can carry on their normal function.

Ornithine Transcarbamylase Deficiency

In contrast to the low prevalence of NAGSD, OTC deficiency (OTCD) (OMIM 311250) is the most commonly occurring UCD with an estimated frequency of 1 in 14,000 [91,119] due to location of the OTC gene on the X chromosome. When OTC is defective or absent, citrulline formation from carbamylphosphate and ornithine, is disrupted and results in increased concentrations of plasma ammonia, glutamine, alanine, urinary orotic acid and orotidine and decreased urea and citrulline. Conclusive diagnosis of OTCD consists of mutation analysis and/or liver biopsy to measure enzyme activity [120]. Because OTCD is x-linked, most of the patients are hemizygous males; when OTC activity is completely absent they present in the neonatal period and exhibit the most severe symptoms of vomiting, hyperammonemia and coma within the first few days of life [121,122,123]. Late-onset male patients and female heterozygotes, however, vary in disease severity and presentation of symptoms may occur from infancy to adulthood, or the disease may go unrecognized entirely [120,124,125,126]. The disease most often manifests in previously asymptomatic patients due to infection, injury, childbirth and other physiological stresses [127]. In general, patients with neonatal onset of disease exhibit mutations that abolish enzyme activity while late-onset patients have partial activity caused by mutations that destabilize OTC or affect the substrate binding sites [128].

The long-term treatment of OTCD includes a low protein diet with sodium benzoate or sodium phenylbutyrate as ammonia scavenging agents. If the severity of the disease is such that life-threatening symptoms are expected to occur despite treatment [38,39,40], liver transplantation is currently the only option for a cure. As a prevention

for hyperammonemic episodes due to complete OTCD, liver transplantation provides the best opportunity for these patients to develop normally [129]. Due to limitations associated with liver transplantation including availability of donor livers, dangers of the procedure, and the necessary immunosuppressive drugs, gene therapy is currently being evaluated as an alternative treatment. Therapy with adeno-associated virus 8 (AAV8) has led to successful transgene expression in mice, dogs, and non-human primates [130,131,132,133] and has shown to restore activity of the human OTC enzyme in mice [134].

In summary, the combined body of knowledge on urea cycle regulation at the mRNA and protein level spans from the dietary protein that triggers increased gene expression and phosphorylation of nutrient sensing pathway members to downstream targets of these pathways that act as the transcriptional regulators of the urea cycle genes. The contributions of the body of research contained within this dissertation substantially increases knowledge of the pathways involved in amino acid signaling and the role nutrient sensing plays in transcriptional regulation of the urea cycle genes, describes transcriptional regulation of *NAGS*, enables mutation identification in gene transcription regulatory regions and subsequent diagnosis of OTCD and NAGSD, and has led to treatment with markedly improved long-term outlook of NAGSD.

Chapter 2: Transcriptional Regulation of N-Acetylglutamate Synthase

Publication

Sandra Kirsch Heibel, Giselle Yvette Lopez, Maria Panglao, Sonal Sodha, Leonardo Mariño-Ramírez, Mendel Tuchman, Ljubica Caldovic. (2012) Transcriptional Regulation of N-acetylglutamate Synthase. *PloS One.* 7(2):e29527:

http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0029527

Introduction

Ammonia, the toxic product of protein catabolism, is converted to urea by the urea cycle in the liver of mammals. Incorporation of two nitrogen atoms into urea is catalyzed by six enzymes: three of them mitochondrial, N-acetylglutamate synthase (NAGS; EC 2.3.1.1), carbamylphosphate synthetase 1 (CPS1; EC 6.4.3.16) and ornithine transcarbamylase (OTC; EC 2.1.3.3), and the other three cytosolic, argininosuccinate synthetase (ASS; EC 6.3.4.5), argininosuccinate lyase (ASL; EC 4.3.2.1) and arginase 1 (Arg1; EC 3.5.3.1).

NAGS catalyzes the formation of N-acetylglutamate (NAG), an essential allosteric activator of CPS1, in the mitochondrial matrix of hepatocytes and small intestine epithelial cells [42,43]. Within hepatocytes, NAGS activity and NAG abundance are regulated by L-arginine, ammonia, and dietary protein intake [4,8,10] and therefore, the NAGS/NAG system may play a critical role in the regulation of ureagenesis in

response to these factors [11]. While studies in the 1980s and 1990s identified the *cis*-acting motifs regulating transcription of the urea cycle enzymes CPS1 [77,135,136,137], OTC [50,51,55,73], ASS [57,76,138], ASL [61,62,139], and ARG1 [63,140], the mammalian *NAGS* gene was not identified until 2002 [43] and we can now report for the first time on its transcriptional regulation.

Many studies have identified regulatory links between the urea cycle genes and glucocorticoids and glucagon [19,47,84], however the mechanism of regulation differs for each gene [3,41,84,141,142]. Transcription of *CPS1* is activated by TATA-binding protein (TBP) while its proximal and distal enhancers contain binding sites for glucocorticoids and cAMP responsive factors including CCAAT-enhancer binding protein (C/EBP), activator protein-1 (AP-1), glucocorticoid receptor (GR) and cAMP response element binding (CREB). Sites for binding tissue selective factors including hepatic nuclear factor 3 (HNF-3) are also present [45,46,47]. Tissue specific expression of the *OTC* gene is induced in the intestine and liver by HNF-4, which binds in the promoter [51,73,143] while binding of both HNF-4 and C/EBP to the enhancer, induces high expression levels in the liver [47,50,51,52,73]. *ASS* transcription is regulated by cooperative binding of multiple copies of specificity protein 1 (Sp1) [57]. *ASL* is regulated through Sp1 and the positive regulator, nuclear factor Y (NF-Y), which binds within the promoter of ASL to activate its transcription [60,61,62,139]. Sp1 and nuclear factor 1 (NF-1)/CCAAT-binding transcription factor (CTF) activate *ARG1* transcription while two C/EBP factors and two unidentified proteins bind within an enhancer in intron 7 to confer glucocorticoid responsiveness [63].

Abundance of urea cycle enzymes correlates with dietary protein intake [3,4]. Transcription of urea cycle genes is in part regulated by the glucocorticoid and glucagon signaling pathways [142,144]. Therefore, we postulate that there exists a nitrogen sensing mechanism that is both responsive to amino acid(s) and hormone stimulation and that an understanding of the transcriptional regulation of *NAGS* could contribute to the understanding of such mechanism.

In this study, we identified two regulatory regions upstream of the NAGS translation start site that contain highly conserved protein-binding DNA motifs. We subsequently confirmed that these regions function as promoter and enhancer and that the enhancer is most effective in liver cells. Avidin-agarose protein-DNA pull-down assays have been used to confirm binding of Sp1 and CREB within the *NAGS* promoter and Hepatic Nuclear Factor 1 (HNF-1) and NF-Y within the enhancer regions. Chromatin Immunoprecipitation (ChIP) and quantitative real-time PCR have been used to independently verify that Sp1 and CREB bind to the promoter region and HNF-1 and NF-Y bind to the enhancer region. We also used 5'RACE analysis to identify multiple transcription start sites for *NAGS* that may be species and tissue specific. These findings provide new information on the regulation of the *NAGS* gene, and suggest possible mechanisms for coordinated regulation of the genes involved in ureagenesis.

Materials and Methods

Bioinformatic Analysis of the Upstream Regulatory Regions

Multiple and Pair-wise Alignment Analysis

Identification of highly conserved regions was conducted by gathering 15 kilobases of genomic sequence 5' of the *NAGS* translational start site and sequence of intron one in 7 mammalian species including: human (NM_153006.2), chimpanzee (XM_001152480.1), dog (XM_548066.2), cow (XM_618194.4), horse (XM_001917302.1), mouse (NM_145829.1) and rat (NM_001107053.1). The highly conserved regulatory regions of *CPS1* were identified by gathering 15 kilobases of genomic sequences 5' of the translational start site from human (NM_001875), chimpanzee (XM_001146604), dog (XM_856862), mouse (NM_001080809), and rat (NM_017072). Genomic sequences were subject to pair-wise blast comparison using megaBLAST 2.2.21+ [145]. Parameters included expect threshold of 10, match and mismatch scores of 1 and -2, respectively, gap existence and extension scores of 5 and 2 respectively, and maximum expected value E=0.001. Regions of high conservation were identified as sequences with more than 80% identity that were at least 100bp and are present in four or more species.

Cis-eLement OVERrepresentation Analysis

The *Cis*-element OVERrepresentation (CLOVER) [146] program was used to predict the over-represented motifs within the highly conserved regulatory regions of *NAGS* and *CPS1*. CLOVER analysis of these conserved regions identified known protein binding DNA motifs in the TRANSFAC Pro database by calculating over-representation of these sequences compared to a background of ppr_build_33.fa generated from NCBI build 33 [147]. Matrices recognized by multiple transcription factors in the same family are represented by one family member unless otherwise noted. Genomic sequences of the highly conserved regions were aligned using CLUSTALW version 2.0.10 [148].

Plasmid Constructs

The promoter and enhancer of *NAGS* were amplified from human genomic DNA with primer pairs hPromXH and hEnhXH or hPromHXrev and hEnhHXrev (Table 2), respectively, to introduce *XhoI* and *HindIII* restriction enzyme sites and allow subcloning in forward and reverse orientation. Platinum Taq PCRx DNA Polymerase (Invitrogen) was used for amplification with the following conditions: initial denaturation at 95˚C for 2 min, followed by 35 cycles of denaturation at 95˚C for 30 sec, annealing at 57˚C for 30 sec and extension at 68˚C for 1 min, and final extension at 68˚C for 6 min. Promoter and enhancer PCR products were ligated with TOPO-TA sequencing vector (Invitrogen) according to manufacturer's instructions and referred to as TOPOProm, TOPOEnh, TOPOPromRev, and TOPOEnhRev respectively. Mouse *Nags* (m*Nags*) promoter and enhancer were inserted into TOPO-TA vector following the same methods. Correct DNA sequences were confirmed using sequencing primers specified by Invitrogen.

TOPOProm, TOPOEnh, TOPOPromRev, TOPOEnhRev, pGL4.10 (Promega) basic vector containing firefly (*Photinus pyralis*) luciferase *luc2*, and pGL4.23 (Promega) vector containing a minimal TATA promoter with *luc2* were cut with *XhoI* (New England Biolabs) and *HindIII* (New England Biolabs). The vectors were treated with Antarctic Alkaline Phosphatase (AAP) (New England Biolabs) according to manufacturer's instructions, and the *NAGS* regions were ligated with the vectors to form the plasmids in Table 3. TOPOEnh was also amplified with primer pair hEnhBS (Table 2), to introduce *BamHI* and *SalI* restriction enzyme sites at the 5' and 3' ends of the enhancer, respectively. The amplified enhancer product and 4.10Prom were cut with *BamHI* (New England Biolabs) and *SalI* (New England Biolabs), the vector was treated

with AAP, and the enhancer was ligated with the vector (Table 3). Plasmids containing mouse *NAGS* promoter and enhancer were generated using the same methods with the primer pairs listed in Table 2 and plasmids in Table 3. Correct sequences were confirmed using primers specified by Promega.

Point mutations in the binding sites for transcription factors Sp1, HNF-1 and NF-Y were selected based on functional analysis of Sp1 [149,150,151], HNF-1 [152,153], and NF-Y [154,155] binding in other genes. Mutations were engineered by Integrated DNA Technologies and provided in pIDTSMART-KAN vectors (Integrated DNA Technology) (Table 4). Plasmids with mutant Sp1, HNF-1, and NFY were cut with *XhoI* and *HindIII*. Reporter plasmids pGL4.10, and pGL4.23 were cut with *XhoI* and *HindIII* and treated with AAP. Mutated inserts were ligated with vectors to form the plasmids 4.10Sp1m, 4.23HNF-1m, and 4.23NFYm (Table 3). Correct sequences were confirmed using primers specified by Promega.

Point mutations in the CREB binding site, c.-7T>C and c.-5T>A (Table 4), were selected based on functional analysis of CREB binding [156,157] in other genes and were engineered into the *NAGS* gene using QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to manufacturer's instructions. Primers hCREBm Fw and Rv (Table 2) amplified 50ng of template plasmid 4.10Prom to create 4.10CREBm. The correct sequence was confirmed using primers specified by Promega.

The expression vectors encoding Sp1 or HNF-1 cDNA were under control of the cytomegalovirus promoter (Origene).

Table 2. Primer sequences used for cloning of promoter and enhancer regions of NAGS for reporter assays.

Promoter Name	Promoter Sequence
hPromXH Fw	5'-CTCGAGTGGAGGCTGCAGTGAGCTATGATT-3'
hPromXH Rv	5'-AAGCTTTCGCCATGACGACAACCAACTCTT-3'
hEnhXH Fw	5'-CTCGAGAGGACCCTTCTGGGTGGAAGTTAT-3'
hEnhXH Rv	5'-AAGCTTTTCCTAGGGATCCACCCAATTCAGTC-3'
hPromHXrev Fw	5'-AAGCTTTGGAGGCTGCAGTGAGCTATGATT-3'
hPromHXrev Rv	5'-CTCGAGTCGCCATGACGACAACCAACTCTT-3'
hEnhHXrev Fw	5'-AAGCTTAGGACCCTTCTGGGTGGAAGTTAT-3'
hEnhHXrev Rv	5'-CTCGAGTTCCTAGGGATCCACCCAATTCAGTC-3'
hEnhBS Fw	5'-GGATCCAGGACCCTTCTGGGTGGAAGTTAT-3'
hEnhBS Rv	5'-GTCGACTTCCTAGGGATCCACCCAATTCAGTC-3'
hCREBm Fw	5'-GGGGGCAAGAGTTGGTCGACGTCATGGCGAAGC-3'
hCREBm Rv	5'-GCTTCGCCATGACGTCGACCAACTCTTGCCCCC-3'
mPromXH Fw	5'-CTCGAGGCTTAGGCTGGCCTTGAATTGCTT-3'
mPromXH Rv	5'-AAGCTTCCATGACGACAACCAAACCCACT-3'
mEnhXH Fw	5'-CTCGAGGGGAATGGCACTGAGACTGTGT-3'
mEnhXH Rv	5'-AAGCTTGCTCTCCCTTCCAAACATCTCTTCCT-3'
mEnhBS Fw	5'-GGATCCGGGAATGGCACTGAGACTGTGT-3'
mEnhBS Rv	5'-GTCGACGCTCTCCCTTCCAAACATCTCTTCCT-3'

Primers were used to amplify human or mouse DNA by PCR for insertion of the promoter and enhancer regions into sequencing and reporter assay vectors.

Table 3. Plasmids generated for luciferase reporter assays.

Name	Vector	Insert
4.10Prom	pGL4.10	hNAGS promoter
4.10Enh	pGL4.10	hNAGS enhancer
4.23Enh	pGL4.23	hNAGS enhancer
4.10PromEnh	4.10Prom	hNAGS enhancer
4.10PromRev	pGL4.10	hNAGS promoter reverse
4.23EnhRev	pGL4.23	hNAGS enhancer reverse
m4.10Prom	pGL4.10	mNAGS promoter
m4.10Enh	pGL4.10	mNAGS enhancer
m4.23Enh	pGL4.23	mNAGS enhancer
m4.10PromEnh	4.10Prom	mNAGS enhancer
4.10Sp1m	pGL4.10	hNAGS promoter with Sp1 mutations
4.10CREBm	pGL4.10	hNAGS promoter with CREB mutations
4.23HNF-1m	pGL4.23	hNAGS enhancer with HNF-1 mutations
4.23NF-Ym	pGL4.23	hNAGS enhancer with NF-Y mutations

Plasmids generated by ligation of the human or mouse promoter or enhancer with pGL4 vectors for use with luciferase reporter assays.

Table 4. Mutations of Transcription Factor Binding Sites in the Promoter and Enhancer of human *NAGS*

Factor	Wild-type	Mutant
Sp1	5'-CCGCCCCCGCC-3'	5'-AAGAACAAGAA-3'
	5'-GGGGCGGGGG-3'	5'-GGTTCTTTGG-3'
	5'-CCCCGCCCCC-3'	5'-CCAAGAAACC-3'
	5'-CCCCGCCCCG-3'	5'-CCAAGAAACG-3'
CREB	5'-GGTTGTCGTCATGG-3'	5'-GGTCGACGTCATGG-3'
HNF-1	5'-TGGAGTTAATCATCTACTCTG-3'	5'-TGGAGTAAGTCTGCAACCAGG-3'
NF-Y	5'-GGCCCCATTGGCTGCCT-3'	5'-GGCCCCTCCAGCTG-3'

Transcription factor binding sites mutations in the human *NAGS* promoter, Sp1 and CREB, and enhancer, HNF-1 and NF-Y.

Tissue Culture

Cell Culture and Transfections

Human hepatoma cells (HepG2) (donated by Dr. Marshall Summar, Children's National Medical Center, Washington, DC) were cultured in complete media containing RPMI 1640 medium (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (ATCC) and 5% Penicillin/Streptomycin (Invitrogen) under 5% CO_2 at 37°C. Human alveolar basal epithelial cells (A549) (donated by Dr. Mary Rose, Children's National Medical Center, Washington, DC) were cultured in complete media containing Ham's F-12 medium (Invitrogen) supplemented with 10% FBS and 5% Penicillin/Streptomycin. Human colorectal adenocarcinoma cells (Caco-2) (ATCC) were cultured in Eagle's Minimum Essential Medium (Invitrogen) supplemented with 20% FBS. Cells were plated at a density of 5 x 10^5 cells/well on 24-well culture plates 24 hours prior to transfection. The cells (90-95% confluent for HepG2 and A549, 80-85% confluent for Caco-2) were then transfected using Lipofectamine 2000 reagent (Invitrogen) and cultured in transfection media containing medium and serum only. A total of 0.25ug of DNA was transfected with 0.225ug of vector expressing *luc2* and 0.025ug of pGL4.74 vector containing *Renilla reniformis* luciferase (*hRluc)* as an internal control (Promega). For co-transfections 0.225ug of *luc2* vector was combined with either 0.25ug of expression vector or empty vector pUC19 (Invitrogen), and 0.025ug of *hRluc* control vector.

Reporter Assays

24 hours following transfection, cells were assayed for both firefly and *Renilla* luciferase activity using Dual-Luciferase Reporter Assay System (Promega) and Berthold

Centro 960 luminometer (Berthold) according to the manufacturer's protocol. All reporter assay measurements were corrected for transfection efficiency by normalizing the firefly luciferase signal to the *Renilla* luciferase values. Expression level of each construct was determined relative to luciferase expression under control of the *NAGS* promoter in each cell line. All results are an average of three independent experiments that were each carried out in triplicate. Values were expressed as mean ± SEM and analyzed using Student's *t*-test.

5'Rapid Amplification of cDNA Ends (RACE)

5' RACE (Version 2.0; Invitrogen) was performed using RNA isolated from donated mouse livers by Trizol reagent (Invitrogen). RNA from mouse small intestine (Origene), human duodenum (Ambion), or human liver (Ambion) was commercially available. Products were synthesized with human or mouse *NAGS* specific primers complementary to sequence within Exon 1 (Table 5). All reactions began with 5ug of total RNA and the RACE procedure was conducted according to manufacturer's instructions. Second strand synthesis was conducted using Ex Taq Polymerase (TaKaRa Bio Inc.) PCR products were subcloned into pCR 2.1-TOPO vector (Invitrogen) and RACE products were sequenced with primers specified by the manufacturer.

Table 5. Primer sequences used to determine transcription start sites of *NAGS* with 5' RACE.

Primer Name	Primer Sequence
hNAGS-GSP1	5'- GACGGCGAAGGGCTTGT-3'
hNAGS-GSP2	5'- GTCTGGAACTGCGTGAGCC-3'
hNAGS-GSP2B	5'- CGACGCACTGGTTCAGGA-3'
mNAGS-GSP1	5'- CATGACGGCGAAGGGCTTGT-3'
mNAGS-GSP2	5'- GGTAGCAGGTCTGGAATTGCGTG-3'
mNAGS-GSP2B	5'- CGATGAGCCAGTGGCGTGCTTC-3'

Primers were designed according to manufacturer's instructions and used to determine transcription start sites of human and mouse *NAGS* in liver and small intestine RNA using 5' RACE.

Avidin-Agarose DNA Protein Pull-Down Assay

Biotinylated DNA Probes

Probes for Avidin-Agarose DNA-Protein Pull-Down Assays were generated by PCR amplification of genomic DNA isolated from donated mouse tails using Pure Gene DNA Purification Kit (Gentra). Probes were generated using biotinylated or non-biotinylated forward primers and non-biotinylated reverse primers with Platinum Taq PCRx DNA Polymerase (Invitrogen) and amplification conditions: initial denaturation at 95˚C for 2 min, followed by 35 cycles of denaturation at 95˚C for 30 sec, annealing at 60˚C for 30 sec and extension at 68˚C for 1 min, and final extension at 68˚C for 6 min. The m*Nags* promoter regions A and B (Figure 3) were amplified with primer pair mNAGS-Prom Region A, from +97 to -259, relative to the translation initiation codon and with mNAGS-Prom Region B, from -302 to -776, respectively (Table 6). A region of m*Nags*, that is not highly conserved in mammals, -1056 to -1320, was amplified using primer pair mNAGS-Prom-NC to serve as a negative control for the promoter regulatory region. The enhancer region of mNAGS, spanning from -2834 to -3167, was amplified using forward primer pair mNAGS-enh. The non-conserved probe, for use as a negative control for the enhancer region, was the amplification product of primer pair mNAGS-Enh-NC spanning -5569 to -5997 upstream of m*Nags*. For additional negative controls, non-biotinylated probes were generated using each primer pair.

Table 6. Primer sequences used to generate DNA probes of the specified regions of m*Nags*.

Primer Name	Primer Sequence
mNAGS-Prom Region A Fw	5'-AATCTGACCTTCGTGACCCTCACT-3'
mNAGS-Prom Region A Rv	5'- CACTCAGTCTCCGTGAGCCC-3'
mNAGS-Prom Region B Fw	5'-ACATTCCCAAATGTGGCCATCACC-3'
mNAGS-Prom Region B Rv	5'-AAAGCGCTCAGATGTCCTAGGTGT-3'
mNAGS-Prom-NC Fw	5'-AGCTGTACAGATGGTTGTGAGCCT-3'
mNAGS-Prom-NC Rv	5'-TCAGCGGGTAAGAGTACTGACTGCT-3'
mNAGS-Enh Fw	5'-GCTTTGTTGGAAGGTCAAGTCG-3'
mNAGS-Enh Rv	5'-GTGCCCTTCATCTTTGTCCCA-3'
mNAGS-Enh-NC Fw	5'-TAGGGCGTGTTCAAGACAGGGTTT-3'
mNAGS-Enh-NC Rv	5'- AGCTTGGTGGTGAGTGCCTTTAGT-3'

Primers were used to generate DNA probes, by PCR, of the promoter, enhancer, or non-specific specified regions of m*Nags*.

Preparation of Nuclear Extracts

Nuclear extract was isolated from donated adult mouse livers of C57BL/6 mice using Nuclear Extraction Kit (Origene) according to manufacturer's instructions. The protein concentration of the nuclear extract was determined using bovine serum albumin as the protein standard with Bradford Assay dye concentrate reagents (Bio-Rad). On average, 10mg of nuclear protein was obtained from mouse liver.

Binding Protocol and Immunoblot

For the avidin-agarose protein-DNA pull-down assay [158] 1 mg of nuclear extract in PBS buffer containing inhibitors (PBSI; 1x PBS with 0.5mM PMSF, 25mM β-glycerophosphate, mM NaF), 15 ug of DNA probe, and avidin-agarose beads (Sigma) were combined and incubated for 16 hrs on a rotating shaker at 4°. The probe and bead concentrations were in excess to ensure complete pull-down of DNA–protein complexes. Following incubation, the supernatant was reserved while the beads were washed 3 times with cold PBSI and then resuspended and boiled in Laemmli protein denaturing buffer (Bio-Rad) with 0.2 M DTT. The supernatant was also mixed with protein denaturing buffer and DTT and then boiled prior to application on a 10% SDS–polyacrylamine gel. The proteins were separated by electrophoresis, transferred to a nitrocellulose membrane, and then identified by immunoblotting using primary antibodies at 1:2000 dilution of antibody to Sp1 (Santa Cruz Biotech; Millipore), 1:1000 dilution of CREB-1α/β (Santa Cruz Biotech), and 1:3000 dilution of C/EBPα/β (Santa Cruz Biotech) for the promoter region and 1:500 dilution of HNF-1α/β (Santa Cruz Biotech), 1:1000 dilution of NF-Yα (Santa Cruz Biotech), 1:1000 dilution of AP-2 (Santa Cruz Biotech), and 1:2000 dilution of SMAD2/3 (Santa Cruz Biotech) for the -3kb conserved region. The membrane was

than incubated with 1:20,000 dilution of donkey anti-rabbit secondary antibody conjugated to horseradish peroxidase (Pierce) and bands were visualized using SuperSignal West Pico Kit (Pierce) according to manufacturer's instructions.

Chromatin Immunoprecipitation

Tissue Preparation and DNA Immuniprecipitation

Donated livers from adult C57BL/6 mice were minced and chromatin was precipitated using SimpleChIP Enzyme Chromatin Kit (Origene) with the variation for whole tissue. Briefly, fresh tissue was minced and washed with PBS including Protease Inhibitor Complete tablets (Roche). Proteins and DNA were cross-linked with 1.5% formaldehyde, and tissue was disaggregated with dounce homogenizer. Chromatin was sheared to an approximate size of 100-1000bp by micrococcal nuclease digestion followed by sonication. Immunoprecipitation was conducted using antibodies to transcription factors Sp1 (Millipore), CREB (Santa Cruz Biotech), C/EBP (Santa Cruz Biotech), HNF-1 (Santa Cruz Biotech), NF-Y (Santa Cruz Biotech), SMAD2 (Santa Cruz Biotech) and AP-2 (Santa Cruz Biotech) and control antibodies to histone H3 and non-specific rabbit IgG (Cell Signaling Technologies). Chromatin was eluted from protein G agarose beads, cross-linking was reversed, and DNA was purified according to manufacturer's instructions.

Real-Time PCR Quantification

ChIP enriched DNA samples included 2% input control and dilutions for a standard curve, positive control immunoprecipitate from anti-histone H3 antibody sample, negative control immunoprecipitation from anti-rabbit IgG antibody, no antibody

control, water control, and test antibodies. Enriched DNA was subject to quantitative real-time PCR using iTaq SYBR Green Supermix with ROX (Bio-Rad) and gene specific primers (Table 7) including negative locus primers to Chemokine ligand 2 (MIP-2) on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Amplification conditions included initial denaturation at 95˚C for 2 min, followed by 50 cycles of denaturation at 95˚C for 30 sec, annealing at 60˚C for 30 sec and extension at 72˚C for 30 sec, with dissociation steps of 95° for 15 sec followed by 50° for 15 sec and finally 95° for 15 sec. Samples were amplified and analyzed using 7900HT Sequence Detection System Software (Applied Biosystems). Values were expressed as mean ± SEM and analyzed using Student's *t*-test.

Table 7. Primer sequences for quantitative real-time PCR analysis of chromatin immunoprecipitation samples.

Primer Name	Sequence
mNAGS promoter FW	CATACACAAGGGGCGGAGA
mNAGS promoter RV	GGGTTCTTAACTTGCCGCTGA
mNAGS enhancer FW	GGCCTTCCGTAAGTAGGAAGCA
mNAGS enhancer RV	CCCACCTAGAGGGCTGTGT
MIP-2 FW	GAAGGGCAGGGCAGTAGAAT
MIP-2 RV	ATGCACGATGTCTGGAAAAG
mCPS1-CEBP Fw	GGAACATCTCTGGACATCA
mCPS1-CEBP Rv2	CCAATTTGTTTGTAACCAGTGTGAA

Primers were used for quantitative real-time PCR analysis of chromatin immunoprecipitation samples.

Results

Selected regions of non-coding DNA upstream of NAGS are highly conserved

A 15 kilobase region of genomic DNA 5' of the translational start site of *NAGS* and sequence of the first intron from human, chimpanzee, dog, horse, cow, mouse and rat were aligned and compared using pair-wise BLASTn. Comparisons showed three highly conserved regions upstream of human *NAGS* at -57 to -284, -498 to -576, and -2978 to -3344 relative to the start ATG, and no significant conservation within the intron or between -5 and -15kb upstream (Figure 3). The region within -1kb of the translational start site was designated as the putative promoter while the region 3kb upstream was designated a putative regulatory element. Figure 3 also shows an alignment of mammalian *NAGS* genes using UCSC Genome Browser software phastCons (green) and phyloP (blue), which identified three non-coding regions of conservation located 3kb upstream, immediately upstream, and within the first intron of *NAGS*, respectively (Figure 3). The phastCons, phyloP and our analyses of conservation within the NAGS gene differed due to different algorithms that were used to identify regions of conservation [145,159,160].

To validate our strategy for identification of conserved regions, the same analyses were conducted for *CPS1*, a gene in which a proximal promoter and an enhancer element located 6.3kb upstream of rat *Cps1*, have been characterized [70,161,162]. 15 kb of *CPS1* genomic DNA sequence 5' of the translational start site was collected from human, chimpanzee, dog, mouse and rat and compared using pair-wise BLASTn. Five regions of high conservation were identified including the previously reported proximal promoter

located immediately upstream of the translation initiation codon and the enhancer at -7392 to -7966 relative to ATG of the human *CPS1* gene (Figure 4). In addition, three previously unknown regions, termed A, B and C, were also identified at -5, -10.5 and -12 kb relative to *CPS1* translation initiation codon (Figure 4). PhastCons and phyloP alignment of mammalian genomic DNA identified the same 5 conserved regions (Figure 4).

Highly conserved, non-coding regions of NAGS function as promoter and enhancer elements for gene transcription

Reporter assays were used to examine the functionality of each of the following: wild type *NAGS* promoter (4.10Prom), control reversed promoter (4.10PromRev), enhancer alone (4.10Enh), promoter and enhancer (4.10PromEnh), and enhancer in both orientations with the heterologous TATA-box promoter (4.23Enh and 4.23EnhRev) by measuring the expression of a luciferase reporter gene in cultured HepG2 cells (Figure 5A). Vectors pGL4.13, pGL4.23, and pGL4.10 containing firefly luciferase *luc2,* with an SV40 promoter, a minimal TATA-promoter, or without a promoter respectively, were used as positive, baseline reference, and negative assay controls. Vector pGL4.74, containing *Renilla* luciferase *hRluc*, was co-transfected with each plasmid to control for transfection efficiency.

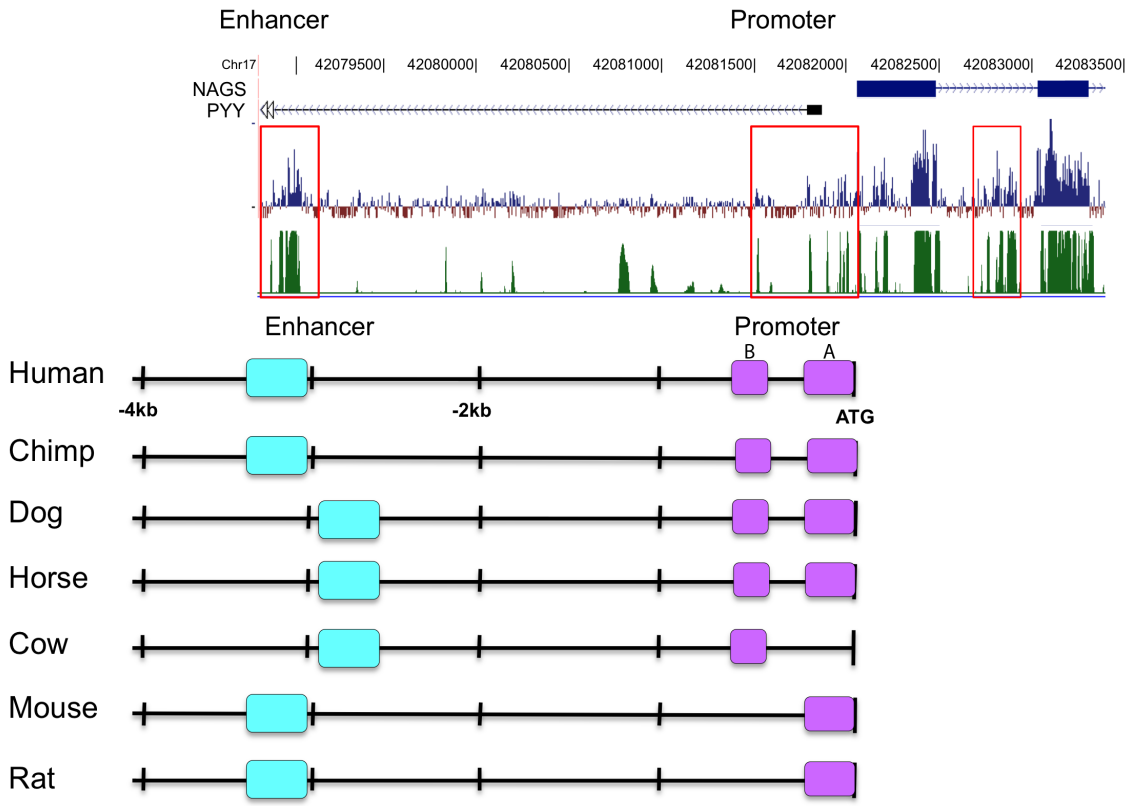

Figure 3. Upstream regions of the mammalian *NAGS* genes are highly conserved.

Conservation of mammalian *NAGS* DNA by phastCons (green) and phyloP (blue) algorithms is shown with the highly-conserved regions indicated in red boxes. Pair-wise BLASTn analysis of mammalian non-coding regions of *NAGS* identified highly conserved sequences upstream of the translational start site termed the promoter (purple) and enhancer (cyan).

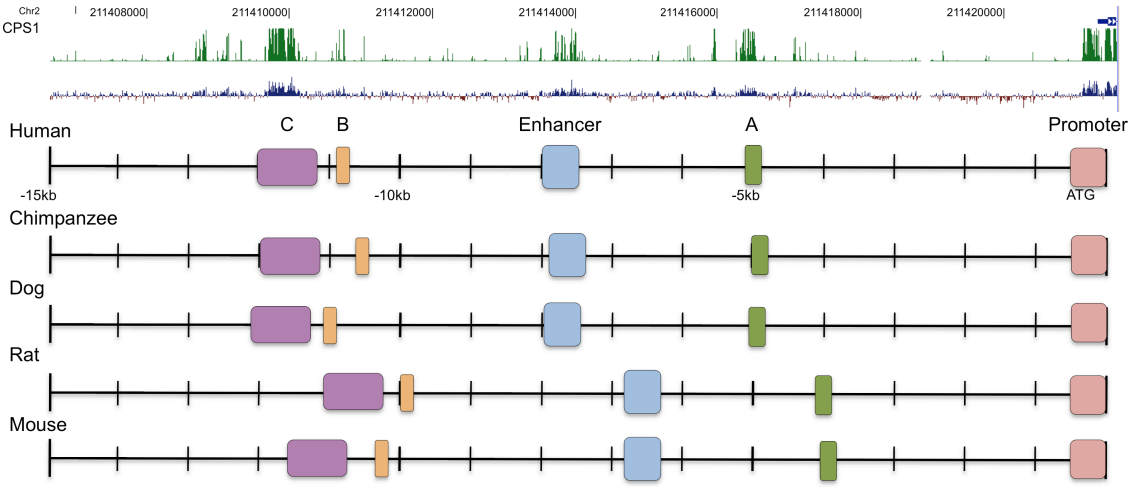

Figure 4. Regions Upstream of mammalian *CPS1* genes are highly conserved.

Three new highly conserved regions were identified within 15kb 5' of the *CPS1* translational start site. Conservation algorithms phastCons (green) and phyloP (blue) from the UCSC genome browser indicate regions that are highly conserved across all mammals (A). Pair-wise blast analysis of human, chimpanzee, dog, mouse, and rat 5' non-coding region of *CPS1* were used to identify two known and three previously unknown regions of high conservation, referred to enhancer/repressor regions A, B, and C. Highly conserved regions within the *CPS1* 5' non-coding sequence include the proximal promoter, region A, the enhancer, region B, and region C.

The human *NAGS* promoter alone (plasmid 4.10Prom), stimulated transcription of the luciferase gene while the upstream regulatory region (plasmid 4.10Enh) alone, did not (Figure 5A). When the *NAGS* promoter and upstream regulatory region were both present (4.10PromEnh plasmid), transcription increased by 50% compared to the promoter alone confirming that the upstream conserved region can function as an enhancer of transcription. When the *NAGS* enhancer was paired with a heterologous promoter containing a TATA-box, in the 4.23Enh construct, it activated transcription of luciferase but the activity was only one third of that observed for 4.10PromEnh plasmid. Neither backbone vector 4.10 nor 4.23 containing the minimal TATA-box stimulated expression of the luciferase gene. As expected, positive control vector 4.13, containing a strong promoter, activated transcription in this cell culture system (Figure 5A). The promoter in the reverse orientation (4.10PromRev) did not activate luciferase expression indicating that the *NAGS* promoter acts in a direction dependent manner (Figure 5B). The ability of the *NAGS* enhancer (4.23EnhRev) to stimulate transcription with the heterologous promoter was orientation independent (Figure 5C). Similar results were obtained for reporter assays using mouse promoter and enhancer (Figure 6).

Transcription of NAGS initiates at multiple sites

Following discovery of the *NAGS* promoter, the transcriptional start sites (TSS) in human and mouse liver and small intestine were identified using 5' RACE (Figure 7A and B). Cloned and sequenced amplification products from 5'RACE were aligned along the 5' non-coding region of *NAGS* with TSS identified in the Database of Transcriptional Start Sites (DBTSS) and expressed sequence tags (ESTs) from Genbank. Results suggest that *NAGS* has multiple TSS and that some may be species and tissue-specific. Combined

5'RACE, DBTSS, and Genbank results indicate that within human liver, the most frequently occurring TSS was at -42bp upstream of the ATG codon, while in human small intestine it was at -146bp (Figure 7A). Within mouse tissues, no dominant TSS was evident, but transcription of the *NAGS* gene initiated most often from -20bp and -108bp in liver and -20bp and -95bp in small intestine (Figure 7B). Figure 7 also shows several other rare TSS that were identified.

Transcription factors bind conserved motifs within the promoter and enhancer of NAGS

When promoters and enhancers from six mammalian NAGS genes were aligned, there were multiple regions of base pair conservation (Figure 8). Cis-eLement OVER-representation (CLOVER) software analysis was employed to identify transcription factor binding motifs in regulatory regions of human, chimpanzee, horse, cow, dog, mouse, and rat *NAGS*. Analyses of the region +9 to -996 bp (relative to the translational start codon, promoter, Table 8) and –2866 to -3620 bp (enhancer, Table 9) predicted several transcription factor binding motifs that are expressed in the liver, but no TATA-box for transcription initiation. Sp1 binding sites, within the promoter, and the HNF-1 binding motif, within the enhancer, received the highest over-representation scores, but additional motifs with lower scores were also over-represented.

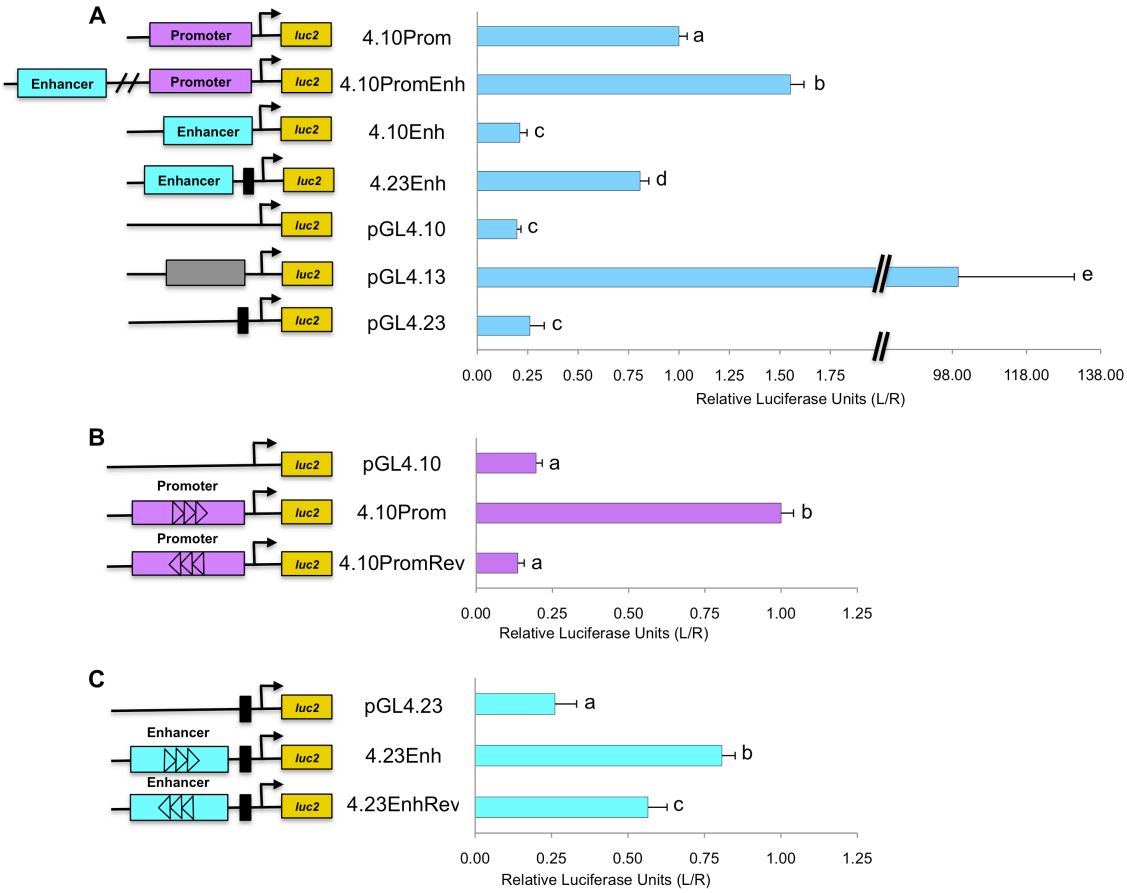

Figure 5. Highly conserved regulatory regions, upstream of the *NAGS* gene, function as promoter and enhancer elements.

In liver derived cells the *NAGS* promoter (4.10Prom), promoter + enhancer (4.10PromEnh), enhancer with TATA promoter (4.23Enh), and positive control promoter vector (pGL4.13) significantly simulate transcription while the enhancer (4.10Enh), basic TATA protein alone (pGL4.23), and vector (pGL4.10) do not stimulate transcription above baseline (A). Reverse insertion of the promoter (4.10PromRev) did not stimulate transcription compared to 4.10Prom and pGL4.10 vector (B), but reverse enhancer (4.23EnhRev) significantly stimulated transcription compared to 4.23Enh and pGL4.23 vector (C). Calculated results are an average of three independent experiments that were each carried out in triplicate, normalized to *Rluc* expression, and expressed relative to the promoter for each experiment with error reported as ±SEM. Different lowercase letters indicate statistically significant differences ($p<0.05$).

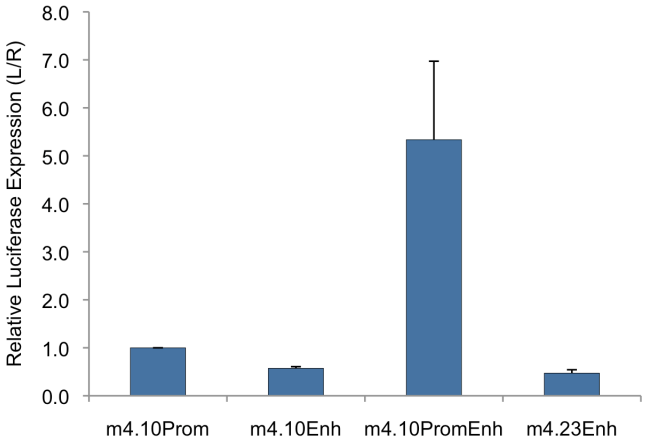

Figure 6. Highly conserved regulatory regions, upstream of the mouse *Nags* gene, function as promoter and enhancer elements.

Mouse promoter (m4.10Prom), promoter + enhancer (m4.10PromEnh), and enhancer with TATA promoter (m4.23Enh) stimulated transcription in HepG2 cells while enhancer lacking a promoter (m4.10Enh) did not. Calculated results are an average of three independent experiments that were carried out in triplicate, normalized to *Rluc* expression, and expressed relative to the promoter for each experiment with error reported as ±SEM.

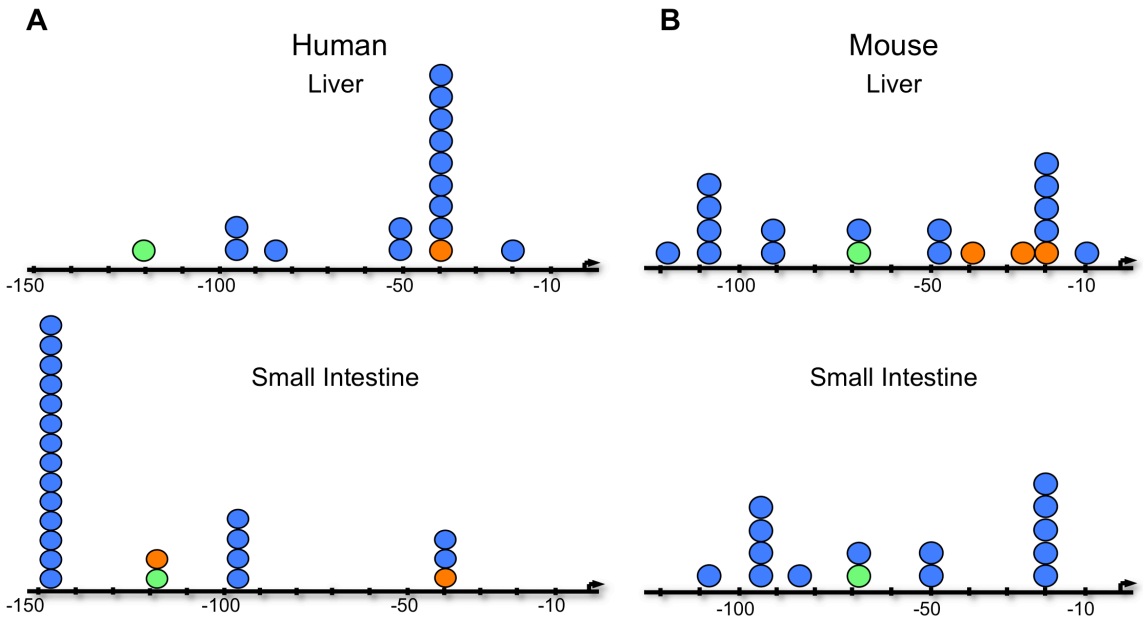

Figure 7. Transcription start sites (TSS) are species and tissue specific.

TSS identified in the promoter of *NAGS* by 5'RACE analysis (blue circles), the Database of Transcriptional Start Sites (DBTSS) (green circles) and 5' termination sites of Expressed Sequence Tags (ESTs) from Genbank (orange circles) were aligned on the DNA sequence 5' of the human (A) and mouse (B) *NAGS* coding sequence. The arrow indicates the translation start site.

Figure 4

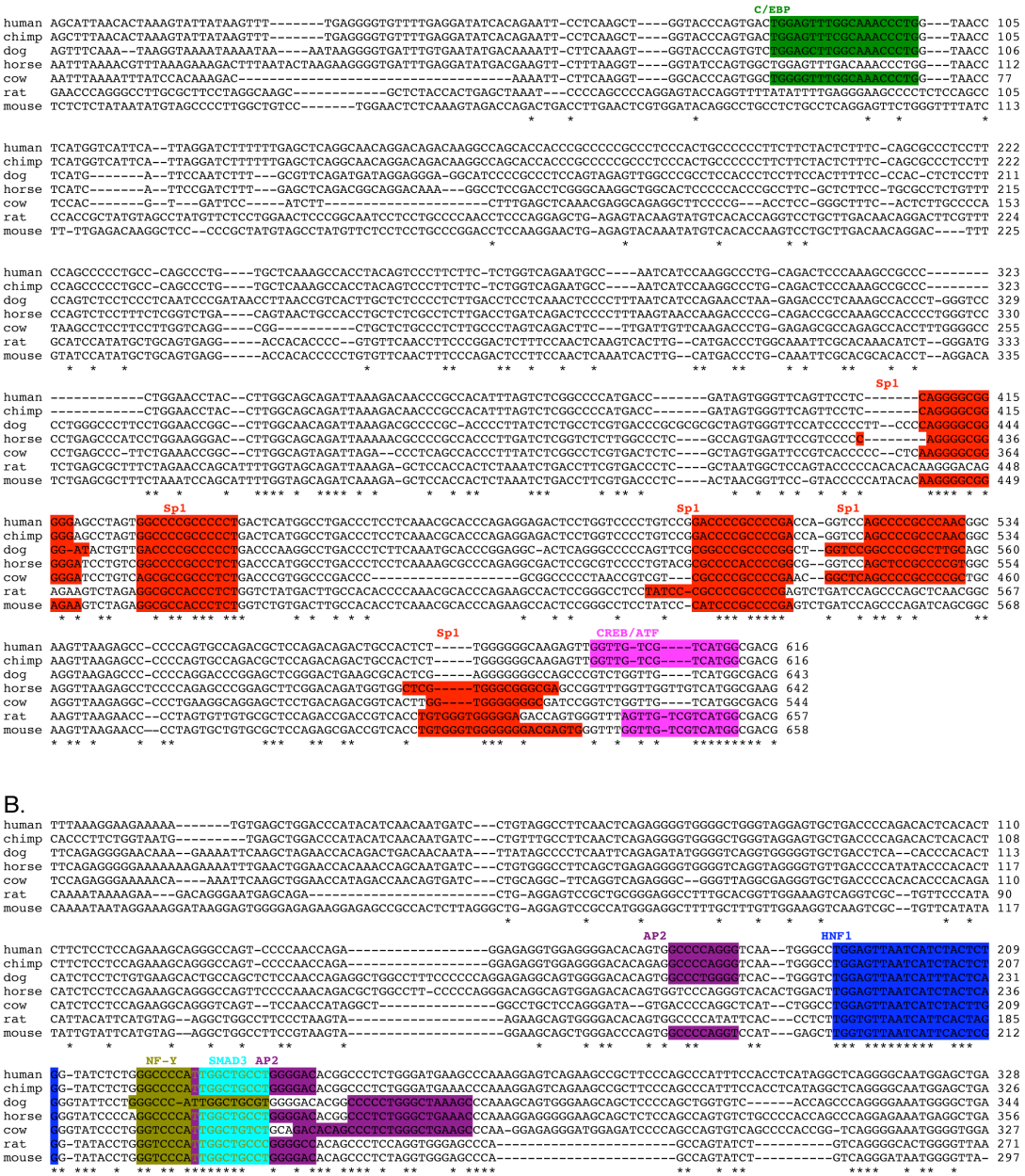

Figure 8. Sequence alignment of *NAGS* promoters and enhancers from seven mammalian species indicate conserved motifs.

DNA sequence of the promoter (A) and enhancer (B) regions were aligned using CLUSTALW alignment software. CLOVER analysis was used to identify transcription factor binding motifs. Binding sites for C/EBP (green), Sp1 (red), CREB/ATF (pink), AP-2 (purple), HNF-1 (blue), NF-Y (olive), and SMAD 3 (cyan) were highly conserved.

Table 8. CLOVER analysis of the promoter region with sequence information for human and mouse *NAGS*.

Sequence file: NAGS_promoter.txt (7 sequences, 7031 bp, 52.8% C+G)
Motif file: transfac_pro_n (588 motifs)

Motif	Raw score	P-value from randomizing
M00933|Sp1	20.3	0
M00196|Sp1	19.8	0
M00932|Sp1	19.7	0
M00931|Sp1	19.6	0
M00255|GC box	14.7	0.0001
M00257|RREB-1	11.3	0.006
M00008|Sp1	8.38	0
M00721|CACCC-binding factor	8.05	0.001
M00691|ATF1	5.05	0
M00916|CREB	1.42	0.006
M00017|ATF	1.03	0.003

Motif	Location	Strand	Sequence
>humanNAGS_promoter			
M00932|Sp1	22 - 34	+	atagggtgggact
M00932|Sp1	138 - 150	+	agtgggaggactg
M00008|Sp1	233 - 242	+	tgggcatggt
M00257|RREB-1	244 - 257	-	gtgtgcatttgtgg
M00255|GC box	302 - 315	+	gggaggtggaggct
M00932|Sp1	302 - 314	+	gggaggtggaggc
M00257|RREB-1	474 - 487	-	aggggtgttttgag
M00916|CREB	542 - 555	+	ggtaacctcatggt
M00255|GC box	607 - 620	-	accacccgcccccg
M00932|Sp1	608 - 620	-	ccacccgcccccg
M00255|GC box	613 - 626	-	cgcccccgccctcc
M00932|Sp1	614 - 626	-	gcccccgccctcc
M00932|Sp1	618 - 630	-	ccgccctcccact
M00935|NF-AT	646 - 655	-	ctctttccag
M00932|Sp1	850 - 862	+	caggggcggggga
M00255|GC box	850 - 863	+	caggggcgggggag
M00255|GC box	869 - 882	-	tggccccgcccccct
M00932|Sp1	870 - 882	-	ggccccgcccccct
M00255|GC box	940 - 953	-	ggaccccgccccga
M00932|Sp1	941 - 953	-	gaccccgccccga
M00255|GC box	961 - 974	-	cagccccgcccaac
M00196|Sp1	962 - 974	-	agccccgcccaac
M00017|ATF	1041 - 1054	-	ggttgtcgtcatgg
M00916|CREB	1042 - 1055	+	gttgtcgtcatggc

>mouseNAGS_promoter

M00691|ATF1	53 - 63	-	tgagttcaagg
M00932|Sp1	247 - 259	-	atcaccgcccccc
M00932|Sp1	252 - 264	-	cgccccccccccc
M00257|RREB-1	274 - 287	-	gttttgttttgtgt
M01082|BRCA1:USF2	566 - 573	-	caacagga
M00255|GC box	598 - 611	-	ggaccacacccct
M00932|Sp1	599 - 611	-	gaccacacccct
M00721|CACCC-binding factor	785 - 800	-	ccatacacaaggggcg
M00932|Sp1	793 - 805	+	aaggggcggagaa
M00932|Sp1	813 - 825	-	ggcgccaccctct
M00257|RREB-1	844 - 857	+	cctcaaacgcaccc
M00255|GC box	882 - 895	-	ccatcccgccccga
M00932|Sp1	883 - 895	-	catcccgccccga
M00721|CACCC-binding factor	959 - 974	+	cgtcacctgtgggtgg
M00257|RREB-1	965 - 978	-	ctgtgggtgggggg
M00255|GC box	966 - 979	+	tgtgggtgggggggg
M00196|Sp1	966 - 978	+	tgtgggtgggggg
M00932|Sp1	970 - 982	+	ggtgggggggacg
M00932|Sp1	974 - 986	+	gggggggacgagtg
M00257|RREB-1	982 - 995	-	gagtgggtttggtt
M00257|RREB-1	986 - 999	-	gggtttggttgtcg
M00017|ATF	992 - 1005	-	ggttgtcgtcatgg
M00916|CREB	993 - 1006	+	gttgtcgtcatggc

Table 9. CLOVER analysis of the enhancer region with sequence information for human and mouse *NAGS*.

Sequence file: NAGS_enhancer.txt (7 sequences, 5280 bp, 55.2% C+G)
Motif file: transfac_pro_n (588 motifs)

Motif	Raw score	P-value from randomizing
M00206|HNF1	22.4	0
M01011|HNF1	18.9	0
M00687|alpha-CP1	17.9	0
M00287|NF-Y	10.5	0
M00444|VDR	10.1	0.007
M00056|myogenin / NF-1	8.02	0
M00701|SMAD3	6.29	0.001
M00185|NF-Y	5.21	0.0001
M00800|AP-2	4.91	0.01
M00188|AP-1	2.38	0.004
M00514|ATF4	2.36	0.006

Motif	Location	Strand	Sequence
>humanNAGS_enhancer			
M01011|HNF1	19 - 39	+	tactgtcaaatgtcaacccca
M00444|VDR	378 - 392	-	tgaccccagacactc
M01011|HNF1	477 - 497	-	tggagttaatcatctactctg
M00188|AP-1	480 - 490	-	agttaatcatc
M00287|NF-Y	507 - 522	-	ggccccattggctgcc
M00687|alpha-CP1	510 - 520	-	cccattggctg
M00185|NF-Y	510 - 520	-	cccattggctg
M00800|AP-2	514 - 529	-	ttggctgcctggggac
M00701|SMAD3	515 - 523	+	tggctgcct
>mouseNAGS_enhancer			
M00188|AP-1	10 - 20	+	gatggctcagt
M00444|VDR	350 - 364	-	ggacccagtggcccc
M00185|NF-Y	353 - 363	-	cccagtggccc
M01011|HNF1	378 - 398	-	tggtgttaatcattcactcgg
M00188|AP-1	381 - 391	-	tgttaatcatt
M00206|HNF1	381 - 397	+	tgttaatcattcactcg
M00287|NF-Y	408 - 423	-	ggtcccattggctgcc
M00185|NF-Y	411 - 421	-	cccattggctg
M00687|alpha-CP1	411 - 421	-	cccattggctg
M00800|AP-2	415 - 430	-	ttggctgcctggggac
M00701|SMAD3	416 - 424	+	tggctgcct
M00056|myogenin / NF-1	432 - 460	+	cagccctctaggtgggagcccagccagta
M00701|SMAD3	459 - 467	+	tatctgtca

Next, over-represented motifs were mapped on the CLUSTALW alignments (Figure 8A and B) and motifs with high conservation, having been identified in at least four out of the seven mammalian species, were examined further. Throughout the promoter, five binding sites for Sp1 were highly conserved, two of which were conserved in all examined species. A binding site recognized by CREB and Activating Transcription Factor-1 (ATF-1) was conserved in four species and overlapped with the translation start codon; a C/EBP binding site was identified farther upstream in region B of the promoter (Figures 8A & 9A). Within the enhancer, a binding site for HNF-1 was conserved in all species. Overlapping binding sites for NF-Y, AP-2 and Mothers Against Decapentaplegic Homolog 3 (SMAD3) were also conserved in all species, while an additional AP-2 binding site, located 5' of the HNF-1 site, was conserved in four out of seven species (Figure 8B & 9B).

To validate computational strategy for identification of transcription factor binding sites, the enhancers of human, chimpanzee, dog, mouse, and rat *CPS1* were analyzed using CLOVER, and the experimentally identified binding motifs for C/EBP, CREB, GR, AP-1 and HNF-3 [70,161,162] were detected along with additional unreported motifs for HNF-4, AR, C/EBP and HNF-3 (Table 10). The detection of experimentally confirmed binding motifs in *CPS1* has made the use of CLOVER for bioinformatic analysis of *NAGS* credible.

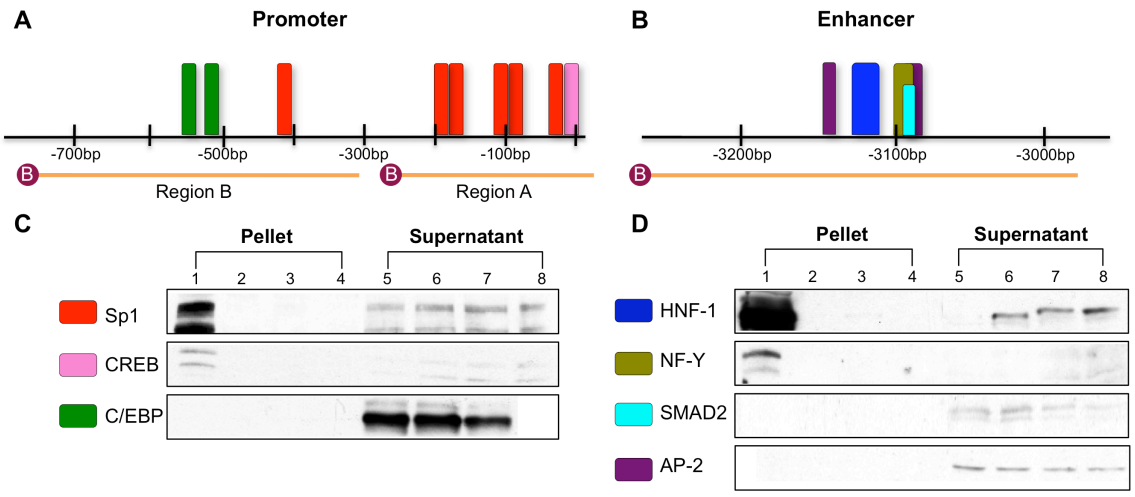

Figure 9. DNA-protein avidin-agarose pull-down assay results confirm transcription factor binding.

Two probes for the promoter (A) and one probe for the enhancer (B) encompass the highly conserved transcription factor binding motifs of *NAGS*. The motif colors reflect the colors used in figures 8A and B. Assays followed by immunoblot confirmed binding of Sp1 and CREB, but not C/EBP within the promoter (C) and HNF-1 and NF-Y, but not SMAD3 or AP-2 within the enhancer regions (D). Lanes 1-4 represent precipitated proteins from mouse liver nuclear extract bound to biotinylated probes of the regions of interest (Lane 1), non-biotinylated probes of the regions of interest (Lane 2), biotinylated probes of non-specific regions (Lane 3), and no probe (Lane 4). Lanes 5-8 represent supernatant fluid from overnight incubation of biotinylated probes of the region of interest (Lane 5), non-biotinylated probes of the region of interest (Lane 6), biotinylated probes of the non-specific regions (Lane 7), or no probe (Lane 8). Immunoblots are representative of at least three replicate experiments.

Table 10. CLOVER analysis of the enhancer region with sequence information for human and mouse *CPS1*.

Sequence file: CPS1_enhancer_element.txt (5 sequences, 2875 bp, 40.4% C+G)
Motif file: transfac_pro_n (588 motifs)

Motif	Raw score	P-value from randomizing
M00447|AR	10.8	0
M00724|HNF3alpha	10.1	0
M00912|C/EBP	6.34	0
M00117|C/EBPbeta	6.07	0.001
M00109|C/EBPbeta	5.83	0.002
M00123|c-Myc:Max	4.52	0
M00921|GR	4.09	0.002
M00967|HNF4, COUP	3.64	0.004
M00116|C/EBPalpha	3.44	0.006
M00707|TFIIA	3.26	0.003
M00981|CREB, ATF	2.52	0
M00801|CREB	1.01	0

Motif	Location	Strand	Sequence	Score
>humanCPS1_enhancer				
M00921|GR	256 - 263	+	cctgttct	6.99
M00981|CREB, ATF	264 - 272	-	ctacgtcat	8.05
M00109|C/EBPbeta	325 - 338	-	atgttgcaccacat	7.08
M00123|c-Myc:Max	331 - 342	+	caccacatgctt	7.87
M00707|TFIIA	349 - 360	-	gatcctcaaata	6.31
M00967|HNF4, COUP	375 - 383	+	agggtccag	7.47
M00921|GR	385 - 392	+	agtgtcct	6.49
M00109|C/EBPbeta	454 - 467	-	atcttgcaaaatca	7.41
M00724|HNF3alpha	470 - 480	+	tgtttactctt	8.11
M00724|HNF3alpha	489 - 499	-	ttaagaaaaca	6.11
M00447|AR	508 - 522	+	agagttgtttgttct	6.57
M00724|HNF3alpha	513 - 523	+	tgtttgttctg	9.29
M00921|GR	515 - 522	+	tttgttct	6.65
M00117|C/EBPbeta	560 - 573	-	tcgttgtgcaaaga	6.17
>mouseCPS1_enhancer				
M00109|C/EBPbeta	117 - 130	+	agtttgtgaaagca	6.03
M00981|CREB, ATF	284 - 292	-	caacgtcat	6.98
M00117|C/EBPbeta	311 - 324	+	ctcttgcttcacta	6.35
M00967|HNF4, COUP	337 - 345	+	caagtccat	6.3
M00109|C/EBPbeta	342 - 355	-	ccattacacaacat	8.34
M00123|c-Myc:Max	372 - 383	+	catcacacgtgt	6.82
M00967|HNF4, COUP	394 - 402	+	agggtccag	7.72

M00921|GR	404 - 411	+	agtgtcct	6.49
M00001|MyoD	428 - 439	-	taccacctctct	6.18
M00189|AP-2	440 - 451	+	ctcctcatggcg	6.06
M00912|C/EBP	470 - 481	+	cttgcaaaatca	7.28
M00724|HNF3alpha	486 - 496	+	tgtttactctt	7.9
M00447|AR	522 - 536	+	agagcagtttgttct	11
M00724|HNF3alpha	527 - 537	+	agtttgttctg	6.71
M00921|GR	529 - 536	+	tttgttct	6.54
M00921|GR	534 - 541	+	tctgttca	6.31
M00707|TFIIA	548 - 559	+	tataagaggggg	8.67
M00912|C/EBP	562 - 573	-	ggttttggcaat	6.11

A DNA-protein pull-down assay was devised to test the bioinformatic prediction of specific binding sites. Two biotin-labeled DNA probes for the promoter (Figure 9A) encompassed regions A and B (Lane 1 in Figure 9C) and one probe (Figure 9B) encompassed the enhancer (Lane 1 in Figure 9D). A biotinylated probe to a region upstream of the *NAGS* gene, lacking any highly conserved motifs (Lane 3 in Figures 9C and D), and non-biotinylated probes to region A or B (Lane 2 in Figures 9C and D) were used as negative controls. The supernatant fluid from each pull-down was included as a positive control for the presence of the transcription factor protein (Lanes 5-8). Intensities of bands corresponding to each transcription factor in supernatant fluids were also used as indicators of pull-down efficiency.

Factors Sp1 and CREB bound to the probe of promoter region A (Lane 1 in Figure 9C). Sp1 also bound to the probe of promoter region B (data not shown) while C/EBP did not bind to this probe (Lane 1 in Figure 9C). Within the enhancer region, transcription factors HNF-1 and NF-Y bound to the probe, however SMAD2/3 and AP2 did not (Lane 1 in Figure 9D). Binding of Sp1, CREB, C/EBP, HNF-1, NF-Y, SMAD2/3, and AP-2 was not detected in the negative controls (Lanes 2–4 in Figures 9C and D) while each transcription factor was detected in the positive controls of liver nuclear extract supernatants (Lanes 5-8 in Figures 9C and D). Each immunoblot result is representative of at least three replicate experiments.

Binding of transcription factors to the predicted motifs was also confirmed using Chromatin Immunoprecipitation (ChIP) followed by Real-Time PCR. Measurement compared the enrichment of target DNA regions to the negative control locus MIP-2. ChIP with Sp1 and CREB antibodies significantly enriched the *NAGS* promoter DNA

compared to MIP-2 ($p<0.005$ and $p<0.05$, respectively; Figure 10A). ChIP with C/EBP antibody did not enrich the *NAGS* promoter DNA compared to the negative locus ($p>0.05$; Figure 10A). The *NAGS* enhancer was enriched in chromatin immunoprecipitated with antibodies against HNF-1 and NF-Y ($p<0.005$ and $p<0.05$, respectively; Figure 10B), but not with antibodies against AP-2 and SMAD2/3 ($p>0.05$; Figure 10B). Thus, pull-down and ChIP assays confirmed that Sp1 and CREB bind along the *NAGS* promoter and HNF-1 and NF-Y bind along the enhancer.

Transcription factors and binding motifs are functionally important

Reporter assays in liver hepatoma cells with mutated transcription factor binding motifs demonstrate the functional importance of each site. Following these sequence substitutions (Table 4), transcription factor binding motifs were no longer detected by CLOVER. Within the promoter, point mutation in the Sp1 binding sites decreased the expression of reporter gene by 75% ($p<0.005$) and point mutations in the CREB binding site resulted in a 40% decrease ($p<0.005$; Figure 11A). Point mutations in the HNF-1 or NF-Y binding sites, in the enhancer, decreased expression of luciferase reporter by 50% ($p<0.005$ for both; Figure 11B).

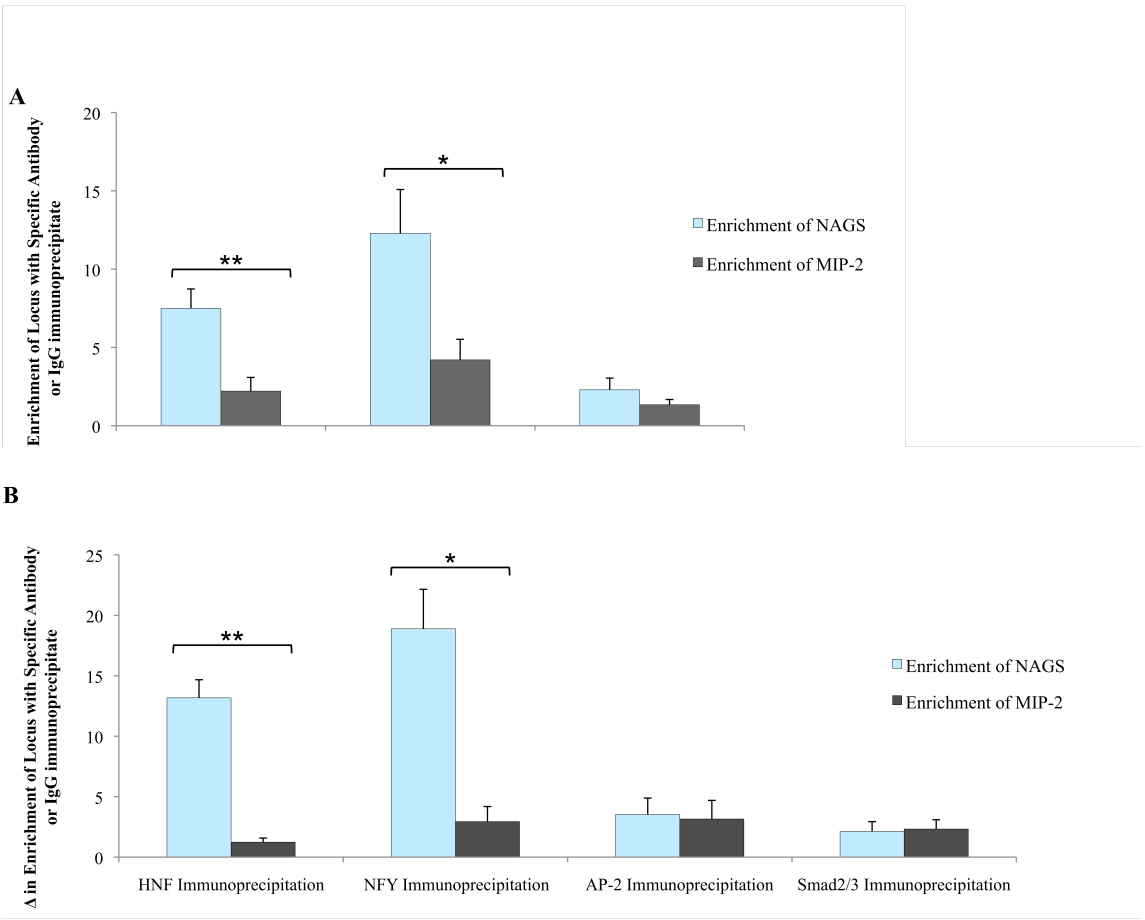

Figure 10. Chromatin Immunoprecipitation (ChIP) results confirm transcription factor binding.

ChIP with transcription factor antibodies was compared to negative control IgG antibody. Real-Time PCR using promoter or enhancer specific primers was compared to primers for the negative locus MIP-2. The results confirmed that Sp1 and CREB but not C/EBP bind within the promoter (A) and HNF-1 and NF-Y but not AP-2 or SMAD2/3 bind within the enhancer region (B) of *NAGS*. Calculated error was from three replicate experiments and reported as ± SEM. One asterisk (*) indicates $p<0.05$ and two asterisks (**) indicate $p<0.005$.

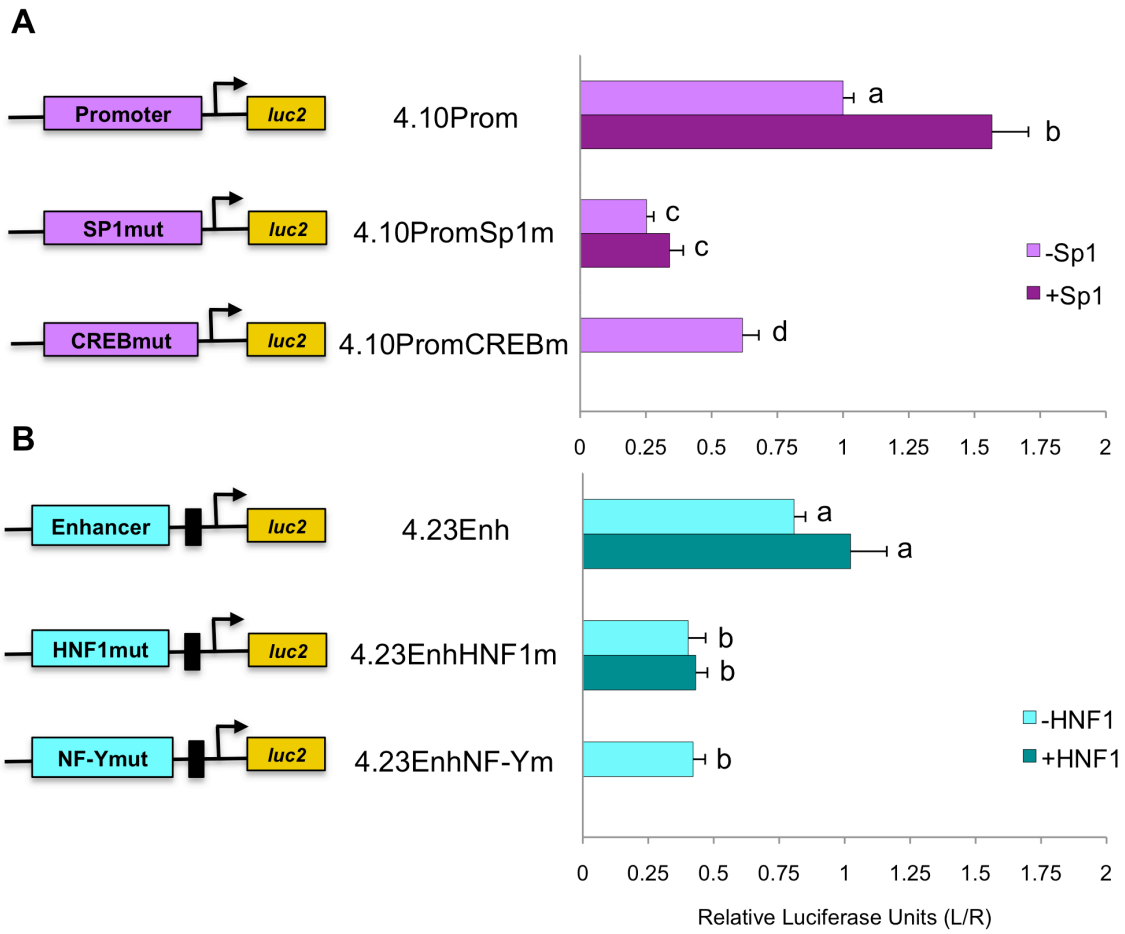

Figure 11. Transcription factors Sp1, CREB, HNF-1, and NF-Y are functionally important for stimulating expression of reporter gene transcription.

Mutagenesis of the putative transcription factor binding sites significantly decreases transcription by the promoter (A) and the enhancer with TATA promoter (B) in liver derived cells when compared to non-mutated sites. Addition of Sp1 with the promoter (A) and HNF-1 with the enhancer (B) increases transcription driven by non-mutated constructs. Calculated results are an average of three independent experiments that were each carried out in triplicate, normalized to *Rluc* expression, and expressed relative to the promoter for each experiment with error reported as ±SEM. Different lowercase letters indicate statistically significant differences ($p<0.05$).

While these results confirm that each motif is important for transcription, the functional importance of Sp1 and HNF-1 proteins is demonstrated by co-expression of the proteins with reporter assay constructs. Co-transfection of Sp1 expression plasmid with the *NAGS* promoter (4.10Prom) increases expression of luciferase more than 50% ($P<0.005$; Figure 11A) while co-transfection of HNF-1 expression construct with the enhancer and minimal TATA promoter (4.23Enh), increases expression of the reporter gene by 25% ($p>0.05$; Figure 11B) suggesting that endogenous Sp1 and, less so, HNF-1 do not saturate their binding motifs on the transfected reporter plasmids.

Reporter assays to compare the effect of the enhancer in liver, intestine and lung cells, included data that were normalized to the reporter expression driven by the promoter. While the *NAGS* enhancer (4.10PromEnh) increased expression of the reporter gene by 50% in liver derived cells (Figure 5A), expression of the luciferase gene did not increase in the intestine or lung derived cells (Figure 12) and therefore the enhancer was determined to be tissue specific. When HNF-1 expression plasmid and 4.10PromEnh were co-transfected into intestine and lung derived cells, transcription was stimulated to levels that were not significantly different from 4.10PromEnh in liver cells ($p>0.05$) (Figure 12). Because intestine and lung derived cells lack HNF-1 (data not shown), this demonstrated the importance of HNF-1 and NAGS enhancer for the tissue specificity of NAGS expression.

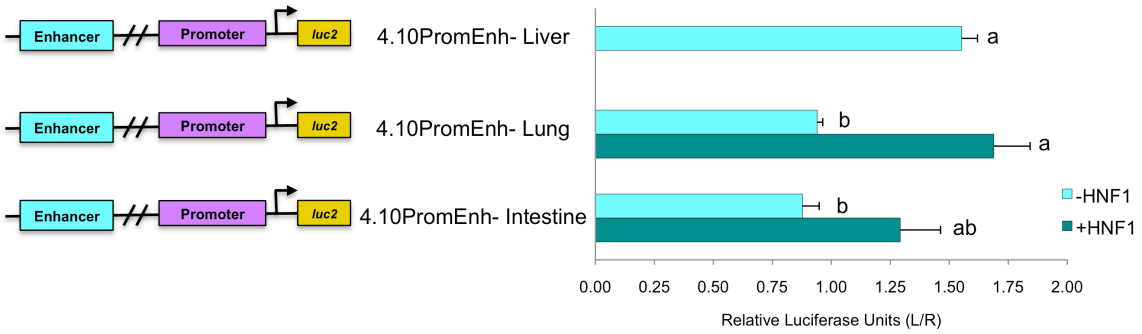

Figure 12. The *NAGS* enhancer shows tissue specificity.

The enhancer with *NAGS* promoter (4.10PromEnh) increases transcription relative to the promoter in liver derived cells but not in intestine or lung derived cells (cyan bars) without the addition of HNF-1 protein (teal bars). Calculated results are an average of three independent experiments that were carried out in triplicate, normalized to *Rluc* expression, and expressed relative to the promoter for each experiment with error reported as ±SEM. Lowercase letters indicate statistically significant differences.

Discussion

In this study we used bioinformatic analyses to predict regulatory regions based on the hypothesis that non-coding DNA sequences that are highly conserved between species are important for gene regulation. Multiple pair-wise BLASTn alignments and sequence alignment from the UCSC genome browser were used to identify two conserved regions within *NAGS*, which were determined to be a promoter and an enhancer. The efficacy of this method was confirmed by successful identification of the experimentally identified promoter and -6.3kb enhancer [46,47], along with three additional highly conserved regions, in the non-coding region upstream of *CPS1*. It should be noted that the high stringency of our BLASTn analysis (80% identity and at least 100bp of aligned sequence in four or more species) was selected to identify conserved regions that could support multiple binding sites where complexes of transcription factors may form [47,163]. This may have caused us to overlook species specific or isolated binding motifs, such as the recently identified FXR binding site [164].

The reporter assay results confirm that the two highly conserved regions within 1kb and 3kb upstream of the translational start site function as promoter and enhancer, respectively. The promoter activates expression of the luciferase reporter gene and we therefore infer that it will activate transcription of *NAGS in vivo*. Similarly, the enhancer in either orientation increases expression of luciferase by approximately 50% relative to the promoter alone, suggesting that it increases *NAGS* transcription as well. The relatively small but significant effect of the enhancer could be due to spacing differences between the genomic *NAGS* promoter and enhancer and their spacing in the reporter constructs. Alternatively, while HepG2 cells express transcription factors that we identified using

bioinformatic tools, the NAGS enhancer may bind additional factors, absent in HepG2 cells, and have larger effect in vivo than in cultured cells. Another explanation for the relatively small effect of the *NAGS* enhancer is the possible presence of a proximal enhancer within the region we termed the promoter. Additional experiments are necessary to distinguish between these two possibilities.

Our analysis of the *NAGS* transcriptional start sites identified multiple TSS that may be species and tissue specific. While the function of each TSS is unknown, these results are consistent with transcription initiation by Sp1 [57,165,166], and future experiments may find that they are involved in transcriptional control for tissue specific expression, developmental-stage specific expression, quantitatively different levels of mRNA expression, or may even determine the transcript stability [167].

After we confirmed that the promoter and enhancer initiate and increase transcription, we looked for transcription factors that bind and regulate *NAGS* in these regions. By filtering for the highly over-represented and spatially conserved binding sites, relative to the translational start codon, we identified Sp1, CREB, and C/EBP in the promoter and HNF-1 AP-2, NF-Y, and SMAD-3 in the enhancer as transcription factors that could bind to the *NAGS* upstream region. This filtering method was confirmed by analysis of the -6.3kb enhancer of *CPS1* in which binding sites for the previously published C/EBP, CREB, GR, and HNF-3 were identified.

The protein-DNA pull down assays, designed to test which transcription factors among a pool of nuclear proteins bind to amplified sequence of conserved upstream DNA, confirmed that Sp1, CREB, HNF-1 and NF-Y bind to *NAGS* promoter and enhancer, while we could not detect binding of C/EBP, AP-2 and SMAD3 (Figure 9). We

initially used 60bp probes encompassing a specific binding motif for the protein–DNA pull down assays. However, probes encompassing the entire region were better able to bind transcription factors (data not shown), suggesting that binding is facilitated by interactions with DNA sequences outside predicted binding sites and possibly other transcription factors and co-activators. ChIP analysis was used to confirm binding of the predicted transcription factors to the DNA regions of interest under physiological conditions. ChIP and DNA-pull down assays confirmed that Sp1 and CREB bind to the promoter and HNF-1 and NF-Y bind to the enhancer of *NAGS*, while reporter assays demonstrated the functional importance of each binding motif by a decrease in transcription following mutagenesis of the binding sites (Figure 11).

Furthermore, we have demonstrated that Sp1 and HNF-1 are important for stimulation of transcription of *NAGS* and that HNF-1 determines tissue specificity of *NAGS* expression. In the liver derived cell line, co-transfection of either Sp1 or HNF-1 expression plasmids with reporter constructs containing the *NAGS* promoter and enhancer led to increased expression of the reporter gene (Figure 11) suggesting that these two transcription factors regulate expression of *NAGS* in the liver. In the lung and intestine derived cell lines, expression of HNF-1 was sufficient to activate expression of reporter gene in constructs containing *NAGS* enhancer and promoter (Figure 12). This suggests that HNF-1 binding to the *NAGS* enhancer determines tissue specificity of NAGS expression. Testing the effect of CREB protein was hindered by its capacity to homo- and heterodimerize with multiple partners [168,169]. The effect of NF-Y was not tested because this transcription factor is a complex heterotrimer [170] and its co-expression with reporter plasmids would require stable expression of NF-Y subunit proteins by *in*

vitro cell culture before reporter plasmids can be co-transfected and assayed for their effect on transcription.

From the data provided herein, we can speculate on the potential role these factors play in regulating *NAGS* transcription. First, in the absence of a canonical TATA-box, transcription initiated by Sp1 often results in multiple transcriptional start sites [171,172]. Sp1 is a strong activator of transcription [57,173,174,175,176] and when multiple Sp1 sites are present, as in *NAGS*, multiple Sp1 proteins can form complexes with each other and synergistically activate transcription [57,174]. Because transcription is significantly increased by co-expression with Sp1 protein and decreased following mutation of the Sp1 binding sites, Sp1 may prove to be the activator of *NAGS* transcription, similar to its role for *ASS*, *ASL* and *ARG1* [47,76].

Second, studies have shown that glucagon and second messenger cAMP trigger a cascade that phosphorylates CREB and allows for DNA binding and activation of transcription [177,178]. In *CPS1* and *ASS*, CREB stimulates transcription upon glucagon signaling [46,76]. Decrease in transcription following CREB binding site mutation and the close proximity of Sp1 and CREB binding sites among the TSS suggests that the transcription initiation machinery may be recruited by these factors, and future research should examine this postulate.

Our experiments and other studies [95] confirm the role of HNF-1 in *NAGS* expression. HNF-1 is essential for stimulation of NAGS expression by its enhancer. This factor is in part regulated by HNF-3, HNF-4, and C/EBP, each of which are known to regulate other urea cycle genes [79,179,180]. Future research could focus on the mechanism of control between these factors, HNF-1, and *NAGS*. Our study has also

shown that NF-Y is an activator of *NAGS* expression, and future studies will focus on the exact mechanism of its function in this context.

The human NAGS gene on the forward strand of chromosome 17 partially overlaps with the peptide YY (*PYY*) gene, which is on the reverse strand. This overlap was identified with a *PYY* cDNA isolated from a brain astrocytoma cDNA library that has an 80 nucleotide long exon located between regions A and B of the *NAGS* promoter [181,182] (Figure 3). Other full-length *PYY* transcripts initiate about 500 bp upstream of the *PYY* coding region, which is located 51 kb upstream of the *NAGS* translation initiation codon. Recent analysis of human transcripts revealed that many protein coding loci are associated with at least one transcript that initiates from a distal site [183], but the significance or function of these transcripts remains to be elucidated. Partial overlap between human *NAGS* and *PYY* genes raises the interesting possibility that these two genes share *cis*-acting regulatory elements and might be co-regulated [182,184]. The mechanism of co-regulation of human *NAGS* and *PYY* is likely to be complex because of their differing tissue expression patterns [42,185,186,187] including differing cell types within the intestine. *PYY* is expressed in the intestinal neuroendocrine cells [188,189] while enterocytes in the intestinal mucosa express *NAGS* [190,191], together with *OTC* and *CPS1* [51,192]. Inspection of the transcription factor binding track of the UCSC genome browser revealed two binding sites for the CTCF transcription repressor between *NAGS* and *PYY* genes; they are located approximately 9.5 and 21 kb upstream of the *NAGS* coding region. The CTCF binding sites could act as chromatin insulators [193,194,195] and either block regulation of *PYY* by the *NAGS* enhancer or enable cell type specific regulation of each gene by the *NAGS* enhancer and promoter. Our results

show that the *NAGS* promoter in the reverse orientation does not activate transcription of the reporter gene in liver derived cells (Figure 5), but this does not preclude transcription activation in other cell types, not tested in this study. It is possible that the *NAGS* promoter, enhancer, or other *NAGS* region, regulates expression of *PYY* [187], and reporter assays in tissues and cultured cells which express *PYY* would test this hypothesis.

While regulation of *NAGS* by Sp1, CREB, HNF-1, NF-Y, and factors that regulate them, requires additional study, identification of regions that regulate human *NAGS* and *OTC* have enabled diagnosis of patients with clinical symptoms of urea cycle disorders, but lacking disease causing mutations in the coding regions of the genes [95,196]. Recently, we identified a patient with a mutation in the enhancer of *NAGS* and confirmed the diagnosis of NAGS deficiency by showing that the mutation significantly decreases transcription of *NAGS* [95]. This example suggests that identification of regulatory regions within genes will lead to more and better diagnoses of urea cycle disorders and other genetic diseases and to accurate genetic counseling.

In conclusion, this study identified a promoter and a tissue specific enhancer of *NAGS* and functionally relevant transcription factor binding motifs within these regions. The results show that Sp1 and CREB bind to the *NAGS* promoter, suggesting that glucagon and cAMP signaling may regulate the expression of *NAGS*. Within the enhancer, HNF-1 may be an important factor in the coordinated regulation of this urea cycle gene transcription through its interaction with HNF-3, HNF-4 and C/EBP while the role of NF-Y is less clear considering that NF-Y may function as an activator or repressor. While additional studies will be needed to further define the roles of these factors, these results contain the first thorough analysis of NAGS and suggest networks of

control between signaling cascades, *NAGS* and the coordinated regulation of the other urea cycle genes.

Acknowledgements

This work was supported by Public Health Service Grant 5R01DK064913 from the National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health. Dr. Mariño-Ramírez was supported by the Intramural Program of the National Library of Medicine, National Institutes of Health. We would like to thank Dr. Marshall Summar for providing HepG2 cells and Dr. Mary Rose for providing A549 cells.

Chapter 3: N-Carbamylglutamate Enhancement of Ureagenesis Leads to Discovery of a Novel Deleterious Mutation in a Newly Defined Enhancer of the *NAGS* Gene and to Effective Therapy

Publication

Sandra K. Heibel[‡], Nicholas Ah Mew[‡], Ljubica Caldovic, Yevgeny Daikhin, Marc Yudkoff, Mendel Tuchman. (2011) N-carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the *NAGS* gene and to effective therapy. *Human Mutation* Vol:32 Iss:10

[‡]These authors contributed equally.

http://onlinelibrary.wiley.com/doi/10.1002/humu.21553/abstract

Introduction

The urea cycle requires the coordinated function of 6 enzymes and 2 mitochondrial membrane transporters to convert ammonia into urea [197]. An interruption in any step of the urea cycle can lead to hyperammonemia, and may result in brain damage, coma and death [197].

The initial step of the urea cycle is catalyzed by carbamyl phosphate synthetase I (CPS1) (EC 6.3.4.16), generating carbamyl phosphate from ammonia, bicarbonate and ATP. However, CPS1 is only active in the presence of its allosteric activator, N-

acetylglutamate (NAG), biosynthesized by N-acetylglutamate synthase (NAGS; EC 2.3.1.1) from glutamate and acetyl-coenzyme A.

Biochemically, a deficiency of NAG, as observed in inherited NAGS deficiency [198] (MIM #237310), is indistinguishable from CPS1 deficiency (MIM #237300), as both conditions result in hyperammonemia due to reduced flux through the CPS1 enzyme. Initial diagnoses of NAGS deficiency were based on measurements of hepatic NAGS activity [198,199], though these assays may not be entirely reliable [101,200,201]. The cloning of the human *NAGS* gene (MIM #608300) in 2002 [42] has allowed molecular testing to become the primary method of diagnosis. Mutations in the coding region of the *NAGS* gene have been identified in all reported cases of NAGS deficiency since 2002 [97,99,105,106,114,116,117,202].

Patients with NAGS deficiency respond remarkably well when administered N-carbamylglutamate (NCG), a stable structural analog of NAG, which has been shown to activate CPS1 *in vitro* [203], normalize blood ammonia [116,204], and restore ureagenesis [116,204]. In fact, NAGS deficiency is the only urea cycle disorder (UCD) in which urea cycle function may be restored through a pharmacological agent. In other urea cycle disorders, medical treatment is suboptimal, and consists of dietary protein restriction and ammonia scavenging agents [109].

The regulatory regions of the *NAGS* gene were recently identified [44] *NAGS* transcription is regulated by a promoter, which contains multiple Sp1 (Specificity protein 1) and one CREB (cAMP response element binding) binding sites, and an enhancer 3kb upstream of the coding sequence to which HNF-1 (Hepatic Nuclear Factor 1) and NF-Y (Nuclear factor Y) bind.

Mutations within the regulatory regions of genes currently represent a very small proportion of known disease causing mutations [205], and are likely underdiagnosed. Indeed, there has been only one report of a disease-causing mutation in a regulatory region of a urea cycle gene [196].

In this report, we describe the first *NAGS* enhancer mutation, identified in a patient with an unspecified proximal urea cycle disorder, who despite the absence of any identifiable *NAGS* mutations by conventional mutation analysis [99,206], demonstrated a marked response to a short-term trial of NCG. We demonstrate that the mutation ascertained in our patient is indeed deleterious, and results in decreased transcription factor binding and expression of the *NAGS* gene. We also document the patient's favorable subsequent response to chronic treatment with NCG, leading to apparent normalization of her health.

Materials and Methods

Case Report

The patient was a 17-year-old girl of Eastern European origin, with an unspecified proximal urea cycle disorder. At 10 years of age, she presented to the Emergency Department with a 3-day history of nausea, ataxia, and somnolence. Initial plasma ammonia was 203 μM, anion gap 13, glutamine 868 μM, citrulline 12 μM, BUN 3.6 mM. She was tentatively diagnosed with partial OTC deficiency, however, a subsequent allopurinol challenge revealed normal orotic acid excretion in response to the drug, making this diagnosis improbable. In addition, sequencing of exons and intron/exon boundaries for *NAGS*, *CPS1* (MIM #608307) and *OTC* (MIM #300461) did not reveal

any pathogenic mutation. In a further attempt to reach a diagnosis, a liver biopsy was performed, but enzymatic analysis of hepatocyte CPS1 and OTC demonstrated normal activity. Analysis of hepatic NAGS enzyme activity was not performed. Because of the succession of non-diagnostic investigations, we considered enrolling her into a short-term clinical trial of N-carbamylglutamate (NCG) [116,204]. We deemed that a diagnosis of partial NAGS deficiency was still plausible, and that a short-term treatment with NCG might not only be diagnostic, but would also correct the metabolic defect in this condition, as observed in previous studies [116,204,205]. In addition, a short-term trial was not likely to be harmful in the other proximal urea cycle disorders. Indeed, a partial CPS1 deficiency might in fact respond to supraphysiological levels of a CPS1 activator [207,208]. We discussed an NCG trial with the patient and her parents, and they consented to the study.

N-Carbamylglutamate study

The study design (Figure 13) has been used in our prior NCG studies [116,209]. As shown, the patient underwent an identical study procedure before and immediately following treatment for 3 days with oral NCG (Carbaglu, Orphan Europe, Paris, 2.2g/m^2/d in four divided doses). The study was carried out by Dr. Nicholas Ah Mew and was approved by the Children's National Medical Center Institutional Review Board with informed consent was obtained before enrollment.

On day 1 of the clinical study, an indwelling venous catheter was used for blood drawing. On day 4, however, peripheral intravenous access could not be obtained, and a peripherally inserted central catheter (PICC) was introduced. The PICC remained in place for the remainder of the study, and was removed upon obtaining the last blood sample.

Following an overnight fast, a baseline sample (5 mL) of heparinized blood was obtained (time 0). Then, the subject ingested 0.33 mmol/kg, (27.5 mg/kg) of [1-^{13}C] sodium acetate (98 atom % excess) dissolved in 60 mL of water. Blood samples (4 mL) were subsequently obtained at 15, 30, 45, 60, 90, 120, 180, and 240 minutes, and drawn into a pre-cooled, heparinized tube. The sample was immediately centrifuged at 4°C to separate the plasma, which was directly transferred to cryogenic tubes, placed in dry ice, and kept frozen at -70°C until analysis, performed within 2 days after collection, as previously described [116,210]. Each sample was analyzed for plasma ammonia and urea (RXL Dade Behring, Siemens Healthcare Diagnostics, Deerfield, IL), quantitative amino acids (Biochrom, Cambridge, UK), and [^{13}C]urea [116,210].

Mutation Detection

The patient's *NAGS* gene was amplified using primers listed in Table 11. Polymerase Chain Reaction (PCR) conditions were: 3m initial denaturation at 98°C, followed by 30 cycles of 30s denaturation at 98°C, 30s of annealing at 65°C, 50s extension at 72°C, and final extension of 5m at 72°C. The entire *NAGS* gene, from 3.5kb upstream of the start codon to 250bp downstream of the termination codon, including introns, was amplified and sequenced. Two small regions from c.-1414 to -1346 (68bp) and c.2301 to 2575 (274bp), could not be sequenced likely due to the GC-rich nature of these segments. These two segments fall within a non-conserved 5'-untranscribed location, and within intron 4, respectively.

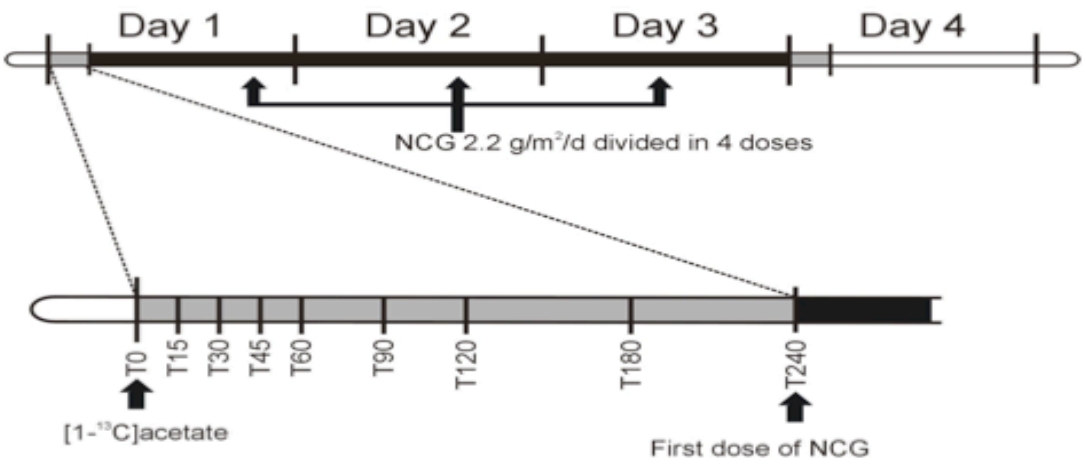

Figure 13. Experimental time line.

Identical studies were performed before and immediately following a 3-day course of NCG at a dose of 2.2 g/m^2/d. Plasma samples for ammonia, urea, amino acids, and [^{13}C]urea were obtained at 0, 15, 30, 45, 60, 90, 120, 180, and 240 minutes prior (time 0) and following the ingestion of isotopic acetate.

The patient's parents' *NAGS* enhancers were also sequenced using forward primer hNAGS-Enh 5'-CTGAACCTGGGCACAATGACAACA-3', and reverse primer hNAGS-Enh 5'-TATGAGGTGGAAATGGGCTGGGAA-3'. The genomic DNA from a healthy individual was used as a control. The NM_153006.1 (www.ncbi.nlm.nih.gov/nuccore) sequence of the human *NAGS* gene was used as reference for sequence comparisons.

SNP Genotyping

Genotyping was completed using TaqMan allele discrimination assays from Applied Biosystems Inc. according to the manufacturer's instructions (See Table 12 for the PCR primers and allele specific probes). The end point fluorescent readings were performed on an ABI 7900HT Sequence Detection System (SDS V 2.3 software; Applied Biosystems). For quality control, negative controls and duplicate samples (100% agreement) were included in the analysis.

Consensus sequence of HNF-1 Binding Motif

A consensus binding site for transcription factor HNF-1 was determined based on similarity matrices M01011, M00790, M00206, and M00132 in the TRANSFAC database (Figure 14A) [211,212,213,214,215]. At sites with ambiguous nucleotide identity, selection of nucleotide deferred to human *NAGS*. The consensus binding site for HNF-1 at c.-3079 to -3059 was determined to be 5'-AGTTAATAATTAACTAT-3'.

Table 11. Primers used to amplify the human *NAGS* gene for mutation screening.

Name	Sequence
Tile1-1f	5- GTG ACT GTG GCT GGT GTG GGT CTC C -3
Tile1-1r	5- GCC AGC TCT TCA TCA CAG CAG GCC ACC -3
Tile 2-1f	5- GAT ACC AGT CCT GCC CTC AGG GAG CC -3
Tile2-1r	5- CAG GGT CAC TCA GCT CAC AAG CCG TGG -3
Tile3-1f	5- GGGTCT GGC CAT GCC ACT GAC GAG C -3
Tile3-1r	5- GTG AGG AGG CGG GTG GGT CAT TTG AGG -3
Tile4-2f	5- AGG TTA AAG AGG GCC GGG CGT GGT GG -3
Tile4-2r	5- GGG TGG TGC TGG CCT TGT CTG TCC TG -3
Tile5-2f	5- GTC TGG GCA ACA TGG TGA ACA CCC GTC -3
Tile5-2r	5- CAG AAC CAC AGC CAT CAG CGC CGT C -3
Tile6-2f	5- CGA CGG CGC TGA TGG CTG TGG TTC TG -3
Tile6-2r	5- CCG TCT CGG TGA CCC TAC CCT TCC AG -3
Tile7-2f	5- CCA TCA CTC CGC GGA CAA GCC CTT CG -3
Tile7-2r	5- CGG GCA GAG TAG CAC GCT CTG AGG AC -3
Tile8-2f	5- TTT GGC GGC GGG TCT GTG CTA CG -3
Tile8-2r	5- TCC TGG ACA GGA GGA GAA GCC GC -3
Tile9-1f	5- CCC ACC ACT CCT CGG CCG TCA TCA CC -3
Tile9-1r	5- CCT TGG TCC GAA GCT GCC CGA CTG CCC -3
Tile10-1f	5- GGG ACG ACT ACC TGG CCT CGC TG -3
Tile10-1r	5- GAG GAC CTA CTG CAT GCT GGG CAC C -3
Tile11-1f	5- CCC ACG CAG CCC ACC TCT CAC AG -3
Tile11-1r	5- CCT GAG TAG AGC CCT GGA AGG CAC CC -3

f-forward and r-reverse

Table 12. TaqMan primers and probes used for SNP genotyping.

Mutation	Forward Primer	Reverse Primer	Common allele probe (5' VIC)	Rare allele probe (5' FAM)
-3064C>A	5’GGTCAATGGGC CTGGAGTTAAT-3’	5’TCCCCAGGC AGCCAATG-3’	CAGAGATACC AGAGTAGATG	CAGAGATACC AGATTAGATG

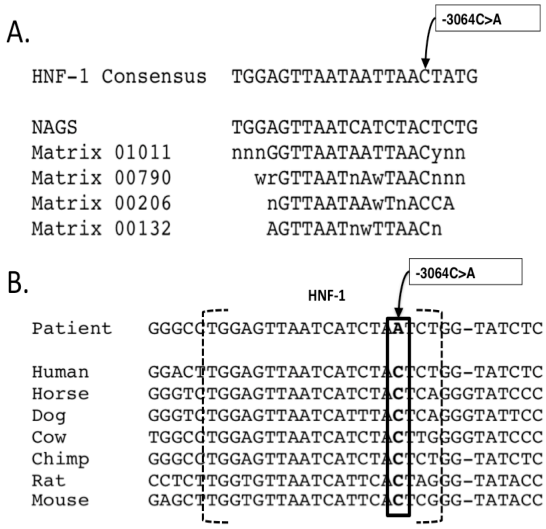

Figure 14. Sequence alignments.

(A) The human *NAGS* HNF-1 binding site and matrices for HNF-1 binding sites from the TRANSFAC database and indication of the c.-3064C>A mutation. (B) Reference mammalian DNA sequence of the HNF-1 binding site was aligned with the patient's HNF-1 binding site. The c.-3064C>A mutation (box) is indicated. Standard nucleotide codes are used.

Site-Directed Mutagenesis

The c.-3064C>A patient mutation and c.-3069C>A, c.-3066C>T, c.-3065T>A, c.-3061C>A consensus HNF-1 binding site mutations were engineered into the *NAGS* gene using QuickChange Lightening Site-Directed Mutagenesis Kit (Agilent) according to manufacturer's instructions. Primers NAGS c.–3064C>A Fw and Rv (Table 13) amplified 100ng of template plasmid 4.10Enh [44] to create 4.10EnhPat. Amplification of 100ng of 4.10WT using primers NAGS HNF-1con Fw and Rv (Table 13) created 4.10EnhCon plasmid. The presence of each mutation was confirmed by sequencing. Cloning of wild-type and mutated NAGS enhancers into plasmids for analysis of transcription is described in detail in the previous chapter.

Tissue Culture and Reporter Assays

Cell culture and transfections

Human hepatoma cells (HepG2) (donated by Dr. Marshall Summar, Children's National Medical Center, Washington, DC) were cultured in complete media containing RPMI 1640 supplemented with 10% FBS (fetal bovine serum) and 5% Penicillin/Streptomycin under 5% CO_2 at 37°C. Cells were plated at a density of 5 x 10^5 cells/well on 24-well culture plates 24 hours prior to transfection. The cells (90% confluent) were then transfected using Lipofectamine 2000 (Invitrogen) and cultured in transfection media containing RPMI 1640 supplemented with 10% FBS. Cells were transfected with reporter plasmids 4.10WT, 4.10Pat, 4.10Con, 4.23WT, 4.23Pat, or 4.23Con, described in the electronic supplement, or control plasmids (Table 14). A total

of 0.25 μg of DNA was co-transfected with 0.225 μg of *luc2* containing expression vector and 0.025μg of transfection control vector pGL4.74.

Reporter Assays

24 hours following transfection, cells were assayed for both firefly and *Renilla* luciferase activity using Dual-Luciferase Reporter Assay System (Promega) and Berthold Centro 960 luminometer (Berthold) according to the manufacturer's protocol. All reporter assay values were corrected for transfection efficiency by normalizing the firefly luciferase signal to the *Renilla* luciferase values. Mean values are expressed relative to wild-type and calculated error was from three trials and reported as ± standard error of the means (SEM). To evaluate the effects of enhancer and promoter from repeated measurements over three series of experiments, mixed-effects linear regression (Stata 10) was employed.

Avidin-Agarose DNA-Protein Pull-Down Assay

Biotinylated DNA probes

Biotinylated and non-biotinylated probes for the enhancer region of human *NAGS* were amplified from genomic DNA isolated from donated mouse tail using Pure Gene DNA Purification Kit (Gentra) or 4.10EnhPat and 4.10EnhCon plasmids with primers in Table 15 and conditions as described previously [44]. Resulting probes were referred to as pWT, pPat, or pCon respectively.

Table 13. Primers used to generate c.-3064C>A and consensus HNF-1 binding site mutants via site-directed mutagenesis.

Primer Name	Primer Sequence
NAGS-3064C>A FW	5'-GGCCTGGAGTTAATCATCTAATCTGGTATCTCTGG-3'
NAGS-3064C>A RV	5'-CCAGAGATACCAGATTAGATGATTAACTCCAGGCC-3'
NAGS HNF-1con Fw	5'-CCAGGGTCAATGGGCCTGGAGTTAATAATTAACTATG GTATCTCTGGGCC-3'
NAGS HNF-1con Rv	5'-GGCCCAGAGATACCATAGTTAATTATTAACTCCAGGCC CATTGACCCTGG-3'

Table 14. Plasmids for reporter assays with promoter, enhancer and reporter genes as indicated.

Plasmid	Enhancer	Promoter	Reporter Gene
pGL4.10WT	NAGS	NAGS	*luc2*
pGL4.10Pat	c.-3064C>A NAGS	NAGS	*luc2*
pGL4.10Con	Consensus NAGS	NAGS	*luc2*
pGL4.23WT	NAGS	TATA-box	*luc2*
pGL4.23Pat	c.-3064C>A NAGS	TATA-box	*luc2*
pGL4.23Con	Consensus NAGS	TATA-box	*luc2*
pGL4.10	-	-	*luc2*
pGL4.13	-	SV40	*luc2*
pGL4.23	-	TATA-box	*luc2*
pGL4.74	-	HSV-TK	*hRluc*

Reporter genes were from Photinus pyralis (luc2) and Renilla reniformis (hRluc).

Preparation of nuclear extracts

Nuclear extract was isolated from donated adult mouse livers of C57BL/6 mice using Nuclear Extraction Kit (Origene) according to manufacturer's instructions. The protein concentration of the nuclear extract was determined using Bradford Assay dye concentrate reagents (Bio-Rad, Hercules, CA). On average 10mg of nuclear protein was obtained from one mouse liver.

Binding Protocol

The binding assay was carried out as described in the previous chapter. Briefly, nuclear extract was incubated with 15 μg of the DNA probe for 16 hrs at 4°. Following incubation, the beads were washed and resuspended in denaturing buffer. The proteins were resolved on SDS-PAGE, transferred to a nitrocellulose membrane, and identified by immunoblotting using primary antibodies against HNF-1α/β at 1:500 dilutation and NF-Yα at 1:1000 dilution (Santa Cruz Biotech) and HRP conjugated secondary donkey anti-rabbit antibody at 1:20,000 dilutation (Pierce) The bands were visualized using SuperSignal West Pico Kit (Pierce) according to manufacturer's instructions. Band intensity was measured using GS-800 Calibrated Densitometer (Bio-Rad and Quantity One 1-D Analysis Software (Bio-Rad). Values were expressed as mean ± SEM and analyzed using Student's *t*-test.

Results

N-carbamylglutamate trial

The stable isotope method for measuring *in vivo* ureagenesis has been previously applied to assess ureagenesis in normal unaffected individuals [116], as well as to

evaluate response to a 3-day trial with NCG in patients with NAGS deficiency and propionic acidemia [116,209]. Figure 15 summarizes the results of the tracer study before and at the end of the 3-day treatment with NCG. In the described patient, the pre-treatment production of [^{13}C]urea attained a peak of only 0.017 atom % excess (Fig 15A). Therapy with NCG for 3 days augmented this parameter over 7-fold to 0.12 atom % excess. Correspondingly, peak blood [^{13}C]urea concentration before administration of NCG was 0.074μM, and 2.0μM, following treatment, a remarkable 27-fold increase above baseline (Fig 15B). In comparison, the peak concentration of this parameter in 38 untreated adult control studies was 4.01±0.59 μM (mean±SEM) [116].

This marked increase in nitrogen disposal was corroborated by changes in other biomarkers, as illustrated in Table 15. Ammonia at baseline was normal at 21 μM, and remained essentially unchanged at 18 μM following therapy. However, blood Urea Nitrogen (BUN) increased over three-fold (Table 15) and mean glutamine and alanine levels both decreased.

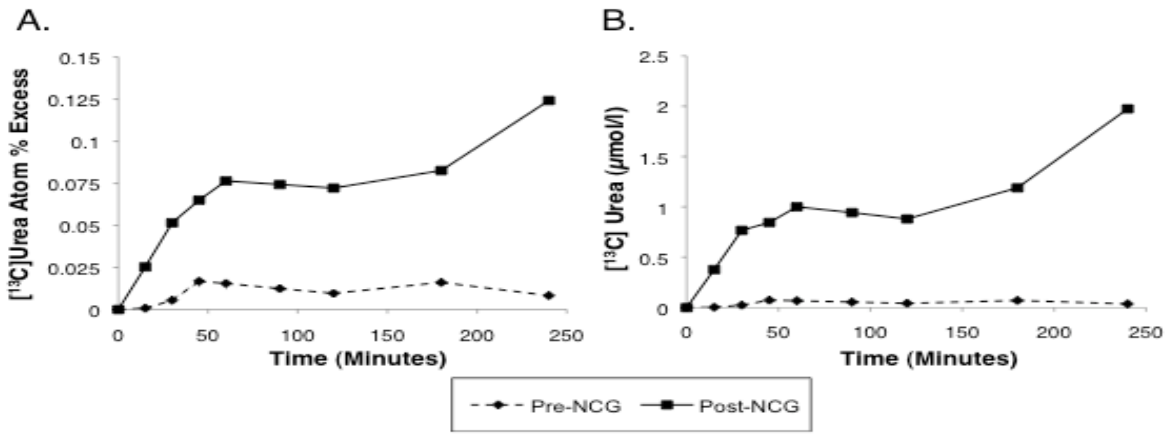

Figure 15. Isotopic enrichment of [^{13}C]urea.

Increase over time in isotopic enrichment of [^{13}C]urea (A) and plasma concentration of [^{13}C]urea (B) in the patient, who was administered 27.5 mg/kg of [1-^{13}C]sodium acetate at time 0 of each study day.

Table 15. Plasma ammonia, urea, glutamine and alanine (mean ± SD) in the patient before and after 3 days treatment with NCG.

Biomarker	Pre-NCG	Post-NCG
Ammonia (μmol/l)	21 ± 6	18 ± 3
Urea (mmol/l)	0.4 ± 0.1	1.4 ± 0.1
Glutamine (μmol/l)	553 ± 61	309 ± 29
Alanine (μmol/l)	616 ± 76	363 ± 31

Mutation detection and SNP Genotyping

Sequencing of the patient's *NAGS* gene, from 3.5kb upstream of the translation start codon to 250bp downsteam of the termination codon, revealed 16 homozygous and 1 heterozygous sequence alteration compared to the reference *NAGS* sequence from the NCBI database. All but one were identified as commonly-occurring polymorphisms in dbSNP (build 132: http://www.ncbi.nlm.nih.gov/SNP/) [216], or were also present in the wild-type control sequence. The remaining change, a homozygous C to A transversion 3064bp upstream of the *NAGS* translation start codon (c.-3064C>A) (Figure 16), was determined to be in the recently identified hepatic nuclear factor 1 (HNF-1) binding site of a conserved *NAGS* regulatory region (Figure 17) [44]. The patient's parents were both heterozygous for this change.

While the mutated cytosine nucleotide is completely conserved among 7 mammalian species (Figure 14), single nucleotide polymorphism (SNP) genotyping was conducted to determine whether the c.-3064C>A change could be an uncommon change found in the general population. The mutation was not detected in 1086 control alleles (Table 16), with representation of the patient's ethnic origin. The diverse cohort of subjects obtained through the Human Genome Diversity Panel (HGDP) [217] included 36 individuals from Russia, 150 subjects from Asia including Japan, China Cambodia, and Papua New Guinea, 56 subjects from South America including native Mexicans and Columbians, and 36 subjects from Western Africa including Nigeria and Senegal, and 265 subjects of Northern European Caucasian decent, obtained from the FAMUSS population [218].

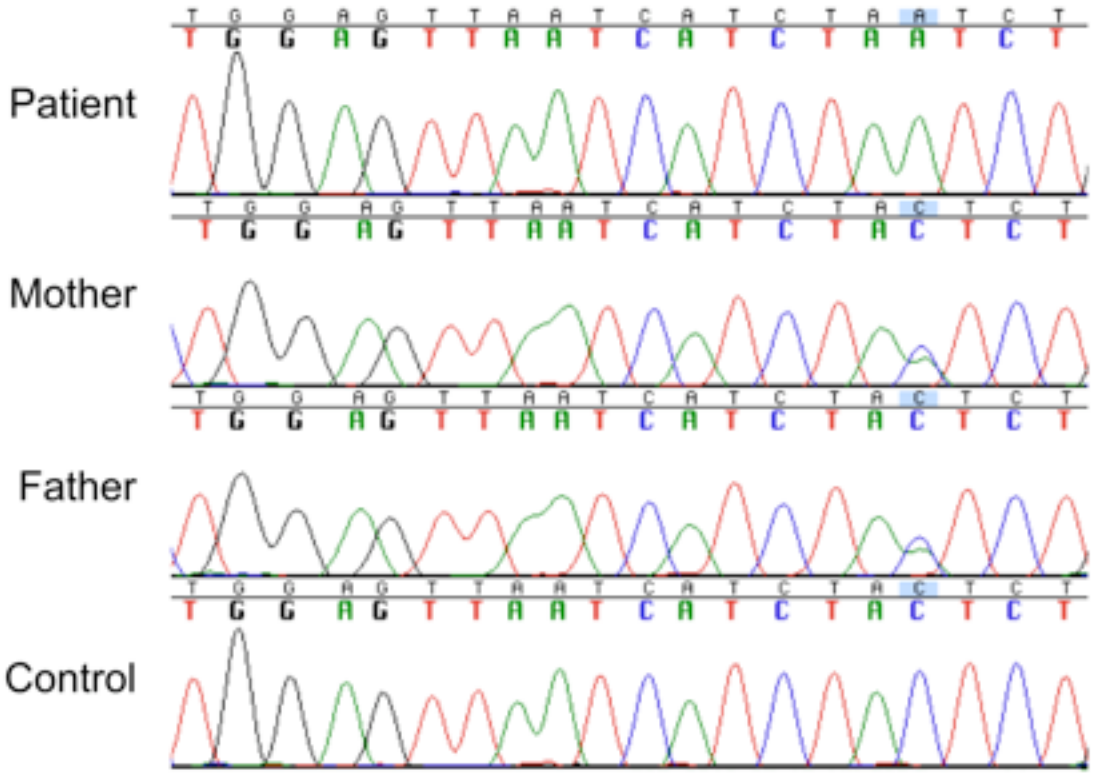

Figure 16. DNA sequence of the HNF-1 binding site within the enhancer of the *NAGS* gene.

The patient is homozygous for a single base substitution at position -3064. The patient's parents are heterozygous for this allele at the same position. Highlighted base indicates the location of the mutation.

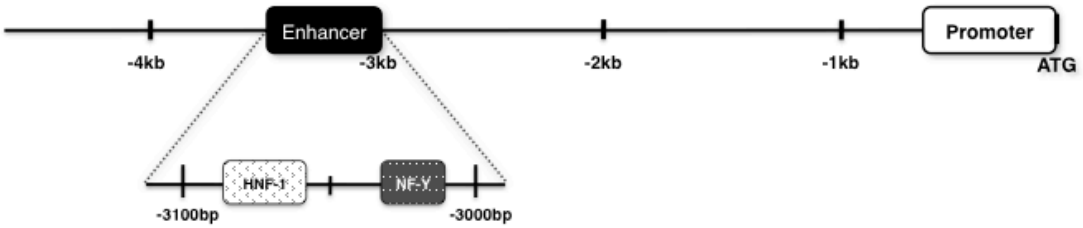

Figure 17. Transcription factor binding sites in the *NAGS* enhancer.

Transcription factors HNF-1 and NF-Y bind within the human *NAGS* enhancer.

Table 16. Allele frequencies in populations analyzed by SNP genotyping.

		Genotype	
		CC	CA/AA
Number of Screened Individuals		*N=543*	*N*=543
Ethnicity	Russian	36	0
	Northern European	265	0
	South American*	56	0
	Asian[+]	150	0
	Western African[†]	36	0

[*]Subjects were native Mexican and native Columbian

[+]Subjects were Japanese, Cambodian, Papau New Guinean, and Chinese

[†]Subjects were Senegalese and Nigerian

Reporter Assays

The functional effect of the mutation on *NAGS* transcription was examined using a luciferase reporter assay, driven by the *NAGS* enhancer with either the wild-type, mutated, or consensus HNF-1 binding sites. Reporter plasmids included the *NAGS* wild-type promoter with the wild-type enhancer, 4.10WT, the enhancer containing the patient's mutation -3064C>A, 4.10Pat, or the enhancer containing the consensus HNF-1 binding site, 4.10Con (Figure 18A). Additional reporter plasmids included a minimum TATA–promoter and the wild-type enhancer, 4.23WT, the enhancer containing the patient's mutation, 4.23Pat, or the consensus enhancer, 4.23Con (Figure 18B) to isolate the role of the enhancers. HepG2 cells were transfected with the reporter plasmids or pGL4.10, pGL4.13, and pGL4.23 as negative, positive, and reference controls respectively. Cells were co-transfected with vector pGL4.74, containing *Renilla* luciferase *hRluc*, as a transfection efficiency control.

Transcription of luciferase was decreased by 25% in cells transfected with plasmids harboring the -3064C>A mutation, compared to the wild-type enhancer, and either the *NAGS* or TATA-containing promoter (4.10Pat $p<0.001$, 4.23Pat $p<0.001$) (Figure 18A & B). Transcription of luciferase increased in cells transfected with plasmids containing the consensus enhancer, compared to wild-type, and either promoter (4.10Con, 35% greater, $p<0.001$, 4.23Con, 75% greater, $p=0.006$) (Figure 18A & B). No luciferase expression was detected for each enhancer alone, in the absence of a promoter (data not shown). Control vectors pGL4.10 and pGL4.23 did not stimulate luciferase expression while positive control vector pGL4.13 activated transcription in this cell culture system as previously shown [44]. Measured luciferase activity was

normalized to *hRluc* expression to adjust for transfection efficiencies. Statistical analysis was performed for normalized values while graphical representation is displayed as normalized values relative to the selected promoter with the wild-type enhancer. The reporter assay results compared to wild-type, showed an increase in transcription of luciferase when the consensus HNF-1 binding site was present and a decrease when the wild-type HNF-1 binding site contained the c.-3064C>A mutation.

Pull-Down Assays

To examine the direct effect of the mutation, a DNA-protein pull-down assay was devised in which biotin-labeled DNA probes containing the wild-type (pWT), c.-3064C>A mutation (pPat), and consensus (pCon) HNF-1 binding site were incubated with nuclear proteins, precipitated using avidin-agarose beads, and analyzed by immunoblot.

The biotinylated enhancer probes pWT, pPat, and pCon contained 333bp of the *NAGS* sequence from c.-2834 to -3167 (Figure 19B, lanes 1, 3, and 5). Non-biotinylated enhancer probes were used as negative controls (Figure 19B, lanes 2, 4, and 6). Normalized results showed a 30% increase in band intensity for pCon (p=0.015) and a 40% decrease in band intensity for pPat (p=0.001) compared to pWT (Figure 19A). Negative control DNA probes did not bind HNF-1 (Figure 19B, lanes 2, 4, and 6). These results indicate that HNF-1 binding is reduced in the presence of the c.-3064C>A mutation and increased in the presence of the consensus site sequence.

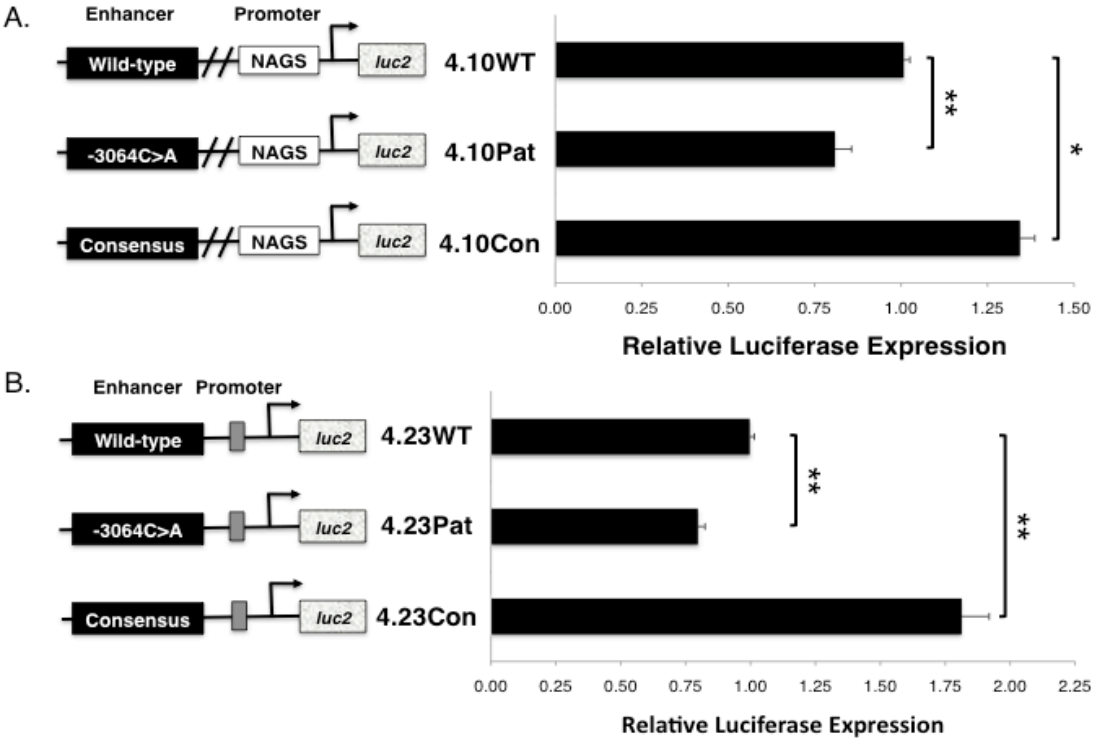

Figure 18. Effects of the c.-3064C>A mutation and consensus HNF-1 binding site on gene expression.

Luciferase expression is significantly different between wild-type, -3064C>A, and consensus enhancers when expressed in conjunction with the human wild-type *NAGS* promoter (A) or the minimal TATA-box promoter (B). All results are an average of three independent experiments that were each carried out in triplicate. Calculated error was from three trials and reported as ±SEM. One asterisk (*) indicates $p<0.05$ and two asterisks (**) indicate $p<0.005$.

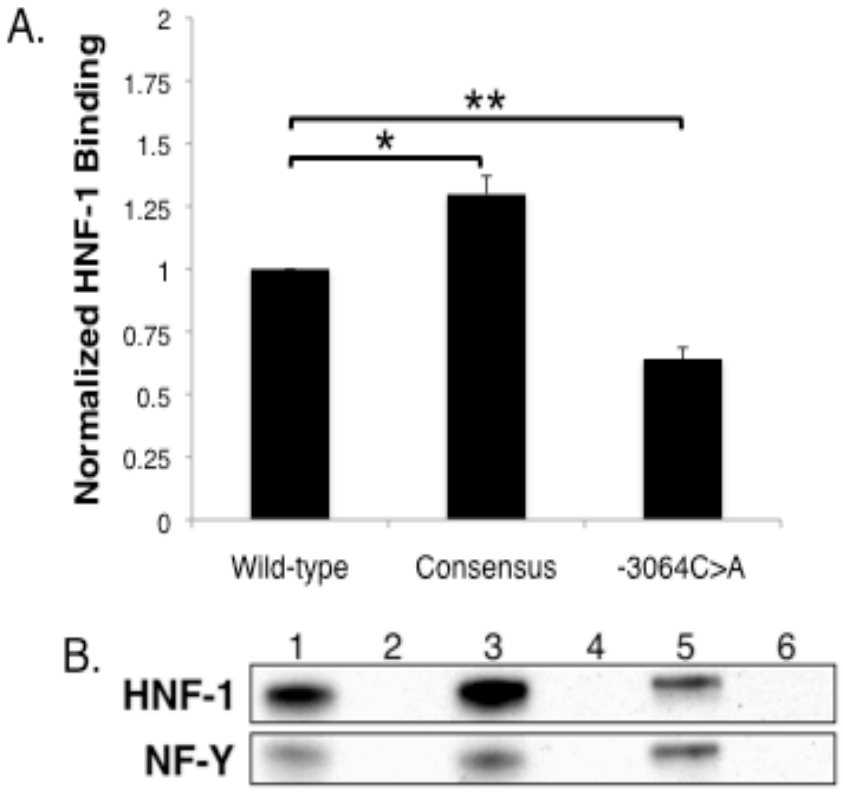

Figure 19. Effects of the -3064C>A mutation and consensus HNF-1 binding site on pull-down of HNF-1 protein.

Amounts of HFN-1 protein pulled down were significantly different between enhancers containing the wild-type, c.-3064C>A mutated, or consensus HNF-1 binding site (A). Immunoblots (B) show HNF-1 and NF-Y protein following overnight incubation with whole nuclear extract. Values are normalized means from three replicate experiments. Immunoblots are representative of the results obtained by three replicate experiments. Calculated error was from three trials and reported as ±SEM. One asterisk (*) indicates $p<0.05$ and two asterisks (**) indicate $p<0.005$.

Chronic Therapy with N-carbamylglutamate

Given the substantial experimental evidence that the patient's mutation was deleterious, as well as her marked response to a short-term trial of NCG, she was prescribed daily NCG treatment.

On her clinic visit just prior to NCG administration, plasma ammonia level was 118μM, glutamine 601μM, alanine 579 μM, blood urea nitrogen 0.7 mM, and arginine 25 μM. Also, essential amino acids, including branched chain amino acids, were low, presumably as a result of protein restriction and alternate pathway therapy. At the time her dietary protein consumption was limited to 0.5g/kg/d (~30-35g/d), and she was on sodium phenylbutyrate at 18g/day.

We prescribed a continuous daily regimen of oral NCG, 2.2 $g/m^2/d$, identical to her dose during the short-term trial. Phenylbutyrate intake was tapered off over a month, and the patient has been able to entirely liberalize her dietary protein. During the three months of therapy to date, she has been clinically well. Her most recent plasma ammonia level was 4 μM, glutamine 379 μM, alanine 374 μM, and arginine 105 μM. Blood urea nitrogen was 4.6 mM, the highest ever recorded in her.

Discussion

This is the first report of a disease-causing mutation in the regulatory region of *NAGS* and the second in any urea-cycle gene [196]. NAGS deficiency is the rarest of the urea cycle disorders [93]. In the 2008 report from the Urea Cycle Disease Consortium Longitudinal Study, there were no enrolled patients with NAGS deficiency [93], yet there were 11 patients without a current specific diagnosis that had symptomatic hyperammonemic episodes and biochemical measures highly suggestive of a urea cycle

disorder (UCD). It is likely that these patients have proximal UCDs since distal UCDs and transporter defects typically present with additional amino acid abnormalities. These undiagnosed patients may in fact have mutations in the regulatory region of *NAGS* or *CPS1* genes.

Although it is well documented that mutations in non-coding regions of a gene can result in disease [219,220,221], currently available clinical molecular diagnostics frequently do not screen for mutations in regulatory regions, and therefore such patients are likely to remain undiagnosed. In a recent report from the Human Genome Mutation Database (HGMD, www.hgmd.cf.ac.uk) [205], only 1.6% of all reported mutations were in regulatory regions of genes. As a preliminary survey, we searched the public version of HGMD for apparently regulatory mutations in the 8 genes of the urea cycle as well as the genes of the American College of Medical Genetics newborn screening core panel, not including congenital deafness, cystic fibrosis or the hemoglobinopathies. As of March 10, 2011, this search only identified a single (aforementioned) disease-causing regulatory mutation in the *OTC* promoter [196], out of 3316 reported mutations in genes associated with inherited metabolic disease.

In the patient described herein, with a high clinical suspicion of NAGS deficiency, a successful diagnostic strategy involved sequencing nearly the entire *NAGS* gene, including the upstream regulatory regions. Continued advances in molecular genetics have allowed for more rapid and accurate sequencing of vast regions of DNA. These technologies will be applied to sequence not just exons, but entire genes, in an attempt to identify deleterious mutations in other patients with a highly-likely clinical diagnosis. However, this strategy will certainly identify numerous single base

substitutions, and subsequent SNP genotyping, as well as functional studies, will be required to differentiate between common polymorphisms and disease causing mutations.

In our SNP genotyping studies, our patient's mutation was not identified in 1086 control alleles, from individuals of varied ethnic backgrounds. Also, the patient's parents were both heterozygous for the c.-3064C>A substitution, consistent with the autosomal recessive inheritance. These results were highly suggestive that the patient's mutation was indeed deleterious.

Functional studies ultimately determined the deleterious effect of the HNF-1 binding-site mutation. While the role of HNF-1 in *NAGS* regulation is not yet entirely understood, these studies demonstrate that it plays an important role in the regulation of *NAGS* expression. The engineering of a consensus HNF-1 sequence resulted in increased HNF-1 binding (Figure 18). It is this enhanced affinity to HNF-1 that likely resulted in increased transcription of our reporter gene (Figure 18). Conversely, the c.-3064C>A mutation results in decreased binding of HNF-1 (Figure 19), and thus significantly reduces, but does not abolish transcription (Figure 18).

This patient, like others with regulatory mutations in genes coding for enzymes [196,222], presented at an age well after the neonatal period. This likely reflects the nature of promoter or enhancer mutations, which might decrease but not eliminate expression of the genes they regulate. The residual gene expression and enzymatic activity retained in these patients results in a less severe clinical presentation.

Treatment with NCG is the therapy of choice in patients with NAGS deficiency, as it corrects the metabolic defect in this condition. In affected patients, NCG has been

shown to normalize blood ammonia [97,98,113,116,223] and restore ureagenesis [116]. This is likely the case in our patient, who has been symptom-free while on NCG, despite discontinuing ammonia scavengers and liberalizing protein in her diet. NCG may also have a salutary effect in other conditions where hepatic NAG is reduced [224], as well as in some patients with partial CPS1 deficiency [207,208].

In the short-term N-carbamylglutamate trial, we used a stable isotope method to assess *in vivo* changes to patients' capacity for ureagenesis in response to NCG. Stable isotopes are safe, and in this case, the analysis of isotopic enrichment in blood urea, by ion-ratio mass spectrometry, is highly sensitive, robust and reproducible [225,226]. This method reliably documented the augmentation in ureagenesis following NCG administration in previous 8 patients studied by our group, including seven patients with propionic acidemia and one with NAGS deficiency [116,209]. In the previously-described patient with NAGS deficiency, there was a greater than 4-fold increase in ^{13}C label incorporation into urea following NCG administration – a marked augmentation that was also observed in the patient described herein. In fact, in this patient, this restoration of ureagenesis provided an important diagnostic clue, and triggered an extensive mutation search in the *NAGS* gene, culminating in the identification of the deleterious mutation and confirmation of the regulatory function of the site affected by it. This marked increase in isotope-enriched urea following a trial of NCG is likely pathognomonic of NAGS deficiency, and may be a useful diagnostic tool in future patients where this diagnosis is suspected, but in whom no deleterious mutations are identified. Furthermore, subsequent long-term treatment with NCG appears to have corrected the metabolic defect in our patient, suggesting that the increase in ureagenesis

observed during a short-term trial with NCG is predictive of the long-term response to this medication.

Acknowledgements

This work was supported by public health service grants R01HD058567, R01DK064913, P30HD026979, M01RR020359, from the National Institutes of Health and by grants from the O'Malley Family Foundation.

We thank Dr. Uta Lichter for her expert patient care, Dr. Robert McCarter for his expertise in statistics, Dr. Marshall Summar for providing HepG2 cells, and the Clinical Research Center staff of Children's National Medical Center for help in performing the study.

Chapter 4: Disease-causing Mutations in the Promoter of the Ornithine Transcarbamylase Gene in Patients Without Molecular Diagnosis

Publication

Disease-causing Mutations in the Promoter of the Ornithine Transcarbamylase Gene in Patients Without Molecular Diagnosis. Sandra K. Heibel, Abigail L. LaBella, Timothy P. Feeney, Maria Panglao, Sonal Sodha, Giselle Y. Lopez, Mendel Tuchman, Leonardo Marino-Ramirez, Hiroki Morizono, Ljubica Caldovic. Annals of Human Genetics [submitted]

Introduction

Within the liver, the urea cycle converts ammonia, a neurotoxic product of protein catabolism, into non-toxic urea, which is easily excreted in the urine [227]. The second enzyme in the cycle, ornithine transcarbamylase (OTC; EC 2.1.3.3), is a mitochondrial enzyme that catalyzes the formation of citrulline from ornithine and carbamylphosphate. The human *OTC* gene is 73 kb long and has ten exons, which contain a 1062 bp long coding sequence [228]. The human OTC protein is 354 amino acids long with a 32 amino acid mitochondrial targeting peptide at its N-terminus [229,230]. The human OTC protein is a homotrimer with three active sites, each shared between two adjacent monomers [231].

OTC deficiency (OTCD) accounts for one-half of all urea cycle disorders [121] and is caused by mutations in the *OTC* gene that lead to reduced or absent functional

OTC enzyme. Because the human *OTC* gene is located on the X-chromosome, OTCD affects primarily hemizygous males and while the true prevalence of OTCD is unknown, it has been estimated to afflict approximately 1 in 14000 [91,119]. Severity and onset of disease in males varies due to mutation as mutations that severely impair OTC function manifest OTCD within the first five days of life [122] while mutations that reduce OTC activity or abundance may present clinically at any time after the first week of life [123,232]. Female carriers of these mutations may exhibit symptoms of OTCD later in life due to skewed lyonization [122,123]. The toxic effects of ammonia yield clinical symptoms of OTCD that include lethargy, vomiting, behavioral and neurological abnormalities, and, in severe cases, coma and death [233]. Biochemical abnormalities include high plasma ammonia, elevated plasma glutamine, low or absent plasma citrulline, and increased excretion of orotic acid and orotidine in the urine [227].

OTC is expressed in the liver and intestine of humans and other mammals [227]. Transcription of the human *OTC* gene appears to initiate at multiple transcription start sites (TSS) [196], while transcription of the mouse and rat *Otc* genes initiate at 136 bp and 98 bp upstream of the translation initiation codon, respectively [234,235]. Four regions within the rat *Otc* promoter, A through D, bind transcription factors that regulate expression of the *Otc* gene [51]. Region A seems to negatively regulate transcription [51], but transcription factors that bind to this region have not been identified. Regions B and C bind transcriptional activator hepatocyte nuclear factor 4 (HNF4), and repressor chicken ovalbumin upstream promoter – transcription factor (COUP-TF) [51,55]. Region D is located downstream of the transcription start site and its role in expression of the *Otc* gene remains to be elucidated [51].

Studies in transgenic mice determined that the rat *Otc* promoter was sufficient to direct expression of transgenes in the liver and intestine of transgenic mice, with higher expression in the intestine [49,143]. Expression studies in cultured cells and transgenic animals revealed that an enhancer, located approximately 11 kb upstream of the first exon of the rat *Otc* gene, is essential for a high level of expression of the *Otc* gene in the liver [51]. *In vitro* binding studies revealed four sites, designated I-IV, that are important for function of the -11 kb enhancer [51]. Transcription factors CCAAT enhancer binding protein (C/EBP), which binds to sites I and II, and HNF4, which binds to sites I and IV, activate expression of the *Otc* gene [72,73,236].

While most cases of OTCD are caused by mutations in the exons or at the intron/exon boundaries of the *OTC* gene [123], some patients exhibit large deletions or insertions which interfere with the open reading frame or processing of the *OTC* mRNA and lead to absence of functional OTC [232,237,238,239,240,241,242,243]. Some mutations in the introns of *OTC* have created cryptic splice sites and resulted in defective mRNA [241,244]. Many molecular causes of OTCD have been accounted for in these regions, and to date a single mutation in the *OTC* promoter in one family resulted in late onset OTCD [196]. However, it is important to note that a significant proportion of patients with clinical OTCD still lack the molecular cause of the disease [123].

The first goal of this study was to determine whether OTCD patients lacking an identified molecular cause of the disease have a deleterious mutation in the 5' flanking region of the *OTC* gene. To that end, we identified a group of patients with clinically diagnosed OTCD, but no disease-causing mutations in the coding region or intron/exon boundaries and then identified two new mutations in the HNF4 binding site in the

promoter of the *OTC* gene that reduced transcription factor binding in three patients. The second goal was to evaluate the feasibility of identifying disease-causing mutations outside of the coding regions and splice junctions of genes without extensive functional tests. Through bioinformatic analyses and SNP genotyping, we identified disease-causing mutations in the regulatory regions of *OTC*. These methods will lead to improved diagnosis of OTCD, allow more accurate genetic counseling and can be extended to screening regulatory regions of other disease-causing genes.

Materials and Methods

Study Subjects

Subjects selected by Dr. Ljubica Caldovic and Timothy Feeney included three patients with reduced or absent OTC enzyme activity and/or urinary orotate above 30 μmoles/mmoles of urinary creatinine, but no disease-causing mutations in the exons and intron/exon boundaries of the *OTC* gene. Patient 1 was a male diagnosed with late onset OTCD based on OTC activity of 6.5 μmoles min^{-1} g^{-1}, which was approximately one-tenth of the normal value (56.7 μmoles min^{-1} per gram of liver tissue). Patient 2 was a male diagnosed with late onset OTCD based on OTC activity of 10.3 μmoles min^{-1} g^{-1}, and urinary orotate concentration of 77 μmoles/mmoles creatinine. Patient 3, a male, was included in the study based on high concentration of urinary orotate (261 μmoles/mmoles creatinine).

Identification of Conserved Sequences in the Upstream Regulatory Regions of Mammalian OTC genes.

15kb of conserved sequence upstream of the translational start codon of human, chimpanzee, rat, mouse, dog, pig, cow and horse *OTC* (Table 17) were identified using pairwise BLAST2N [245] to identify upstream regulatory regions. Regions of at least 100 bp that exhibited more than 70% identity were considered to be highly conserved and were used for predicting functional roles such as transcription factor binding.

Identification of Transcription Factor Binding Sites

Transcription factor binding sites within the promoter, liver specific enhancer (LSE) and upstream regulatory element (URE) from each organism were identified using Cis-element over-representation (CLOVER) software [146] in conjunction with the TRANSFAC Pro Database [147,246]. DNA sequences of the promoter, LSE and URE were aligned using CLUSTAL W [148] and predicted transcription factor binding motifs were identified within each region of the alignment. Transcription factor binding sites were identified based on raw score ranking and conservation of their position in multiple sequence alignments. Binding sites that were present in six out of eight species were considered as candidates for binding of transcription factors.

Table 17. Sequences that were used for identification of conserved regions upstream of the OTC coding sequence.

Species	Genome Build	Sequence	Range
Human	GRCh37.p5	**NC_00023.10**	38196950…38211950
Chimpanzee	Pan_troglodytes-2.1.4	**NC_006491**	38679212…38694212
Mouse	Build 37.2	**NC_000086.6**	9514863…9829563
Rat	RGSC v3.4	**NC_005120.2**	24594496…24609496
Dog	Build 2.1	**NC_006621.2**	33065945…33080945
Pig	Sscrofa10	**NC_010461.3**	36498763…36513763
Cow	5.2	**NC_007331.3**	65364456…65379456
Horse	EquCab2.0	**NC_009175.2**	30935959…30950959

Mutation Analysis of the Conserved Regulatory Regions in the Human OTC Gene

PCR amplification of the promoter, LSE, and URE was conducted by Timothy Feeny on the patient's genomic DNA. Reagent concentration for PCR amplification were as follows: 1.25 U of AmpliTaq DNA polymerase (Applied Biosystems Inc., Carlsbad, CA, USA), 0.2 mM of each dNTP, 0.2 μM of each primer, 1X GeneAmp PCR buffer with 1.5 mM MgCl2 (Applied Biosystems Inc.), and 100 ng of genomic DNA in a final volume of 50 μl. The same conditions were used for amplification of all three regions: 4 min. initial denaturation at 94˚C, followed by 28 cycles of 30 sec. denaturation at 94˚C, 30 sec. annealing at 62˚C, and 1 min. extension at 72˚C, and a 5 min. final extension at 72˚C. Amplification products were resolved using 1% agarose gel electrophoresis to confirm successful PCR. Unincorporated primers were removed from amplification products using ExoSAP-IT (USB, Molecular Biology Reagents and Biochemicals, Cleveland, OH, USA) according to manufacturer's instructions. In addition to the primers that were used for amplification of promoter and LSE, primers that anneal in the middle of promoter and LSE amplicons were used for sequencing of these two regulatory regions (Table 18). Promoter, LSE and URE were sequenced using Big Dye Terminator kit v3.1 (Applied Biosystems Inc.) according to manufacturer's instructions.

SNP Genotyping

Genotyping was carried out using TaqMan allele discrimination assays from Applied Biosystems Inc. according to manufacturer's instructions. Primers and allele specific probes are listed in Table 19. The end point fluorescence readings were performed on an ABI 7900HT sequence detection system (SDS v 2.3 software package,

Applied Biosystems Inc.). Samples without template added were used as negative controls and 100% agreement between duplicate samples were used for quality control.

DNA-Protein Pull-Down Assay

The DNA-protein pull-down assay was carried out by Abigail LaBella using the method described previously [95] with some modifications. Biotinylated and non-biotinylated double stranded DNA with wild-type and mutated HNF4 binding sites were purchased from Integrated DNA Technologies (Coralville, IA, USA). Resulting probes were referred to as pWT, pSNP1 and pSNP2, respectively; their sequences and the sequence of non-specific control DNA are listed in Table 20.

Nuclear extract was prepared from adult mouse livers of C57Bl/6 mice using Nuclear Extraction Kit (Origene, Rockville, MD, USA) according to manufacturer's instructions. The protein concentration in the nuclear extract was determined using Bradford Assay Dye Concentrate (Bio-Rad, Hercules, CA, USA). Nuclear extract was diluted to final protein concentration of 2 mg/ml using PBS containing protease and phosphatase inhibitors (Roche, Indianapolis, IN, USA). The nuclear extract was pre-cleared with streptavidin coupled Dynabeads (Invitrogen, Carlsbad, CA, USA). Pre-cleared extract was combined with 20 µg of DNA probe and 100 µl of streptavidin beads and incubated for 16 hrs. on a rotating shaker at 4˚C. Following incubation, the beads were washed, resuspended in 2x Laemmli sample buffer containing 200 mM DTT, and boiled to dissociate bound proteins. Eluted proteins were resolved using SDS-PAGE, transferred onto nitrocellulose, and probed with the goat anti-HNF-4α primary antibody (Abcam, Cambridge, MA, USA) at 1:1000 dilution followed by HRP-conjugated donkey secondary antibody at 1:20,000 dilution (Pierce, Rockford, IL, USA). The bands were

visualized using Super Signal West Pico Kit (Pierce) according to the manufacturer's instructions. Band intensity (optical density per mm^2) was determined using a GS-800 Calibrated Densitometer (Bio-Rad).

Table 18. Sequences of primers used to amplify and sequence upstream regulatory regions of the OTC gene.

Amplified Region	**Amplicon Size**	**Primer Sequence**	
Promoter	816 bp	Right	5'-CACCACCACGTCTAGCTAATTTTTGTATTT-3'
		Left	5'-GTTTAACAGGATCCTCAGATTAAACAGCAT-3'
LSE	998 bp	Right	5'-CACCAAAAAGTAGAAATGGGCAGAGAAG-3'
		Left	5'-TTAACAGATGGTAATGGTAATGGTAGTTTGTC-3'
URE	544 bp	Right	5'-ACCACCACGTCTAGCTAATTTTTGTATTTT-3'
		Left	5'-ATTTCTAAAAGCTGCATTGTTTAACAGGAT-3'
Sequenced Region		**Primer Sequence**	
Promoter		Right	5'-TCCTCAATGAGTACTTGTCAATTGATTTTG-3'
		Left	5'-TTTTGGAAGAAATCCTGACTCTTTTATTGC-3'
LSE		Right	5'-GTTGTGCATTTCCTAATCTTGTCTTCTGTG-3'
		Left	5'-CATTATTCTGTTAATCATTTGCCCTGGAAG-3'

Table 19. Sequence of primers and allele specific probes used for determining frequency of SNP1 and SNP2 in general population.

Primer Name	Primer Sequence
SNP1 F	5'-AGGAAGCTGAAGGGTGATATTACCT-3'
SNP1 R	5'-GGCCACCTAAAAACTAAGAAATGTGT-3'
SNP1 VIC	5'-TGCAGTGAGGGAGCAA-3'
SNP1 FAM	5'-TGCAGTGATGGAGCAA-3'
SNP2 F	5'-TGCCCTGCAGTATCTCTAACCA-3'
SNP2 R	5'-GGCCACCTAAAAACTAAGAAATGTGT-3'
SNP2 VIC	5'-AGGGAGCAAAGGTAATAT-3'
SNP2 FAM	5'-AGGGAGCAAAAGTAATAT-3'

Table 20. DNA probes used to determine binding of HNF4 to the wild-type, mutated and non-specific control binding sites.

Name	Sequence
pWT	5'-AACCAGGGGACTTTGATAAGGAAGCTGAAGGGTGATATTACCTTTGCTCCCTCACTGCAA-3'
pSNP1	5'-AACCAGGGGACTTTGATAAGGAAGCTGAAGGGTGATATTACCTTTGCTCCATCACTGCAA-3'
pSNP2	5'-AACCAGGGGACTTTGATAAGGAAGCTGAAGGGTGATATTACTTTTGCTCCCTCACTGCAA-3'
Control	5'-TTTTAACATGCCTTACTAGCTCAGTGGTTAAGCACTTATAAGCCTCCCTG-3'

Results

Regulatory Elements and Transcription Factor Binding Sites in Mammalian *OTC*

The urea cycle is essential for disposal of toxic ammonia in mammals, therefore we reasoned that *cis*-acting elements regulating expression of the *OTC* gene will be conserved. Using bioinformatic approaches to establish regions of conservation within the 5'-flanking sequence, we identified possible binding sites for transcription factors that regulate expression of mammalian *OTC*. Pair-wise comparisons identified three conserved regions upstream of the first exon of *OTC* (Figure 20). The first region is 571 bp long and located immediately upstream of the initiating ATG codon. This area encompasses the *OTC* promoter and the 5'-untranslated regions of the *OTC* mRNA. The second region is approximately 900 bp long and corresponds to the -11 kb enhancer in the rat *Otc* gene [51,73]. Because this enhancer is essential for high expression of *OTC* in the liver [73], we termed this *cis*-acting element the liver specific enhancer (LSE; Figure 20). The LSE is located between 5 and 11 kb upstream of the *OTC* promoter in the eight examined mammals and it is located approximately 9 kb upstream of the human *OTC* promoter (Figure 20). The third region, located approximately 2 kb upstream of the LSE, has not been identified before and we termed this sequence the upstream regulatory element (URE) (Figure 20). Figure 20 also shows the conservation track of the 15 kb genomic sequence with additional conserved regions between the promoter and the LSE, and upstream of the URE. This difference is likely due to different computational methods and algorithms used to identify conservation of DNA sequences [160,245,247].

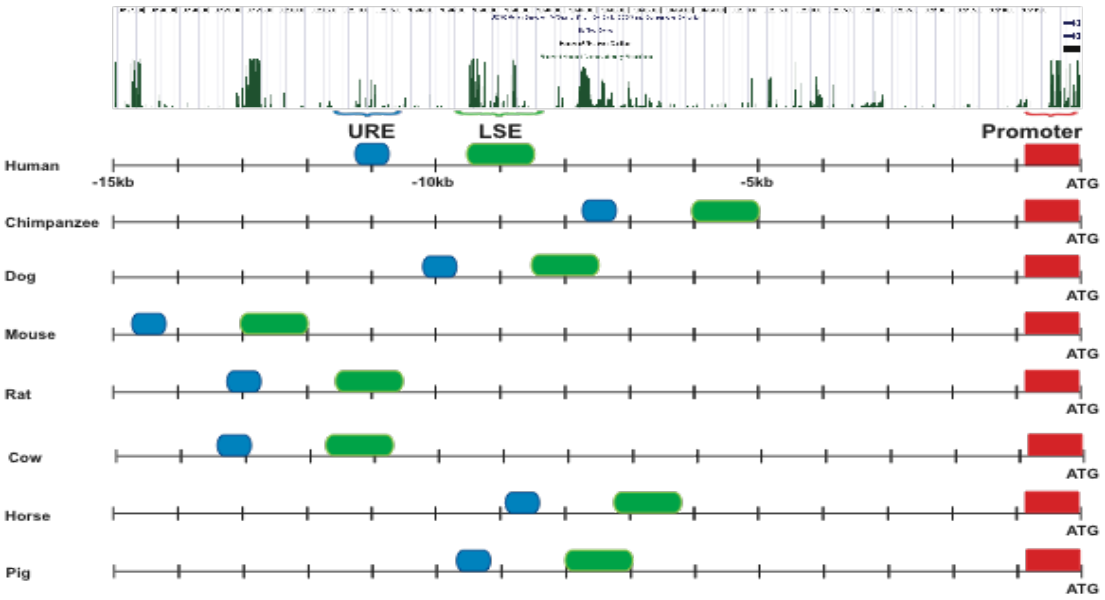

Figure 20. Conserved elements in mammalian *OTC* genes.

Pair-wise BLAST analysis was used to identify three regions of conservation that were present in 5'-flanking sequences of all eight mammalian *OTC* genes. All three regions can be identified in the conservation track that displays sequence conservation in placental mammals. Red – promoter; green – liver specific enhancer (LSE); blue – upstream regulatory element (URE).

Promoter, LSE and URE from all eight mammals were then analyzed using CLOVER software to identify transcription factor binding sites that may be responsible for regulation of the OTC gene. The ability of CLOVER to identify known transcription factor binding sites [51,55,73] was used as positive control. Figure 20 shows LOGO alignments of the OTC regulatory regions and the predicted transcription factor binding sites (magenta) together with the location of cis-acting regulatory elements identified in the rat Otc gene (yellow boxes); [51,55,73]. CLOVER correctly predicted binding sites for HNF4 and COUP-TF [55] in regions B and C in the OTC promoters of the mammalian species (Figure 21A). Binding sites for HNF4 and COUP-TF were predicted in region D, while hepatic nuclear factors 3 and 6 (HNF3 and HNF6) were predicted to bind to region A (Figure 21A). Although regions A and D have been identified as cis-acting regulators of OTC expression [51], transcription factors that bind to these two regions have not been identified previously. Finally, CLOVER identified conserved binding sites for transcription factors Oct-1 and HNF3 upstream of region A (Figure 21A).

The region of conservation located approximately 9 kb upstream of the human OTC translation initiation codon is approximately 600 bp longer than the experimentally identified LSE [51] (Figure 21B). Binding of HNF4 was correctly predicted at sites I and IV in the rat [51,73] and the other seven mammals included in this analysis (Figure 21B). Two binding sites for C/EBP and one for HNF3 were predicted to reside downstream of site IV in all eight species (Figure 21B). However, binding sites for transcription factor C/EBP, which binds to sites II and III of the rat LSE [51,73], were not identified. Instead, our analysis identified conserved HNF3 and GATA-2 binding sequences in site II and

adjacent to site III, respectively (Figure 21B). Within the URE, CLOVER predicted binding sites for immunoglobulin transcription factor 4 (TCF-4) and homeobox transcription factor Cdx-2 (Figure 21C).

Mutation Analysis of the OTC Gene Regulatory Regions in Patients with OTCD

The sequence of the *OTC* promoter, LSE and URE were screened for mutations in the three subjects by Timothy Feeny. Sequencing results were compared to a wild-type sequence from an individual without OTCD and to the reference sequence of the human genome. Two sequence changes, found in the promoter, were designated SNP1 and SNP2 (Figure 22). SNP1 and SNP2 are located in region C, where HNF4 binds and activates transcription of the rat *Otc* gene [51,55]. Patients 1 and 2 had SNP1, which is a C to A transversion, located 95 bp upstream of the translation initiation codon (Figure 22). SNP1 is immediately adjacent to the HNF4 binding site, and therefore it could disrupt binding of this transcription factor and activation of OTC transcription. Patient 3 had SNP2, a C to T transition located 105 bp upstream of the translation initiation codon. This sequence change affects a base pair that is highly conserved in HNF4 binding sites [248] and, therefore, is expected to decrease binding of HNF4 to its mutated binding site.

A.

B.

C.

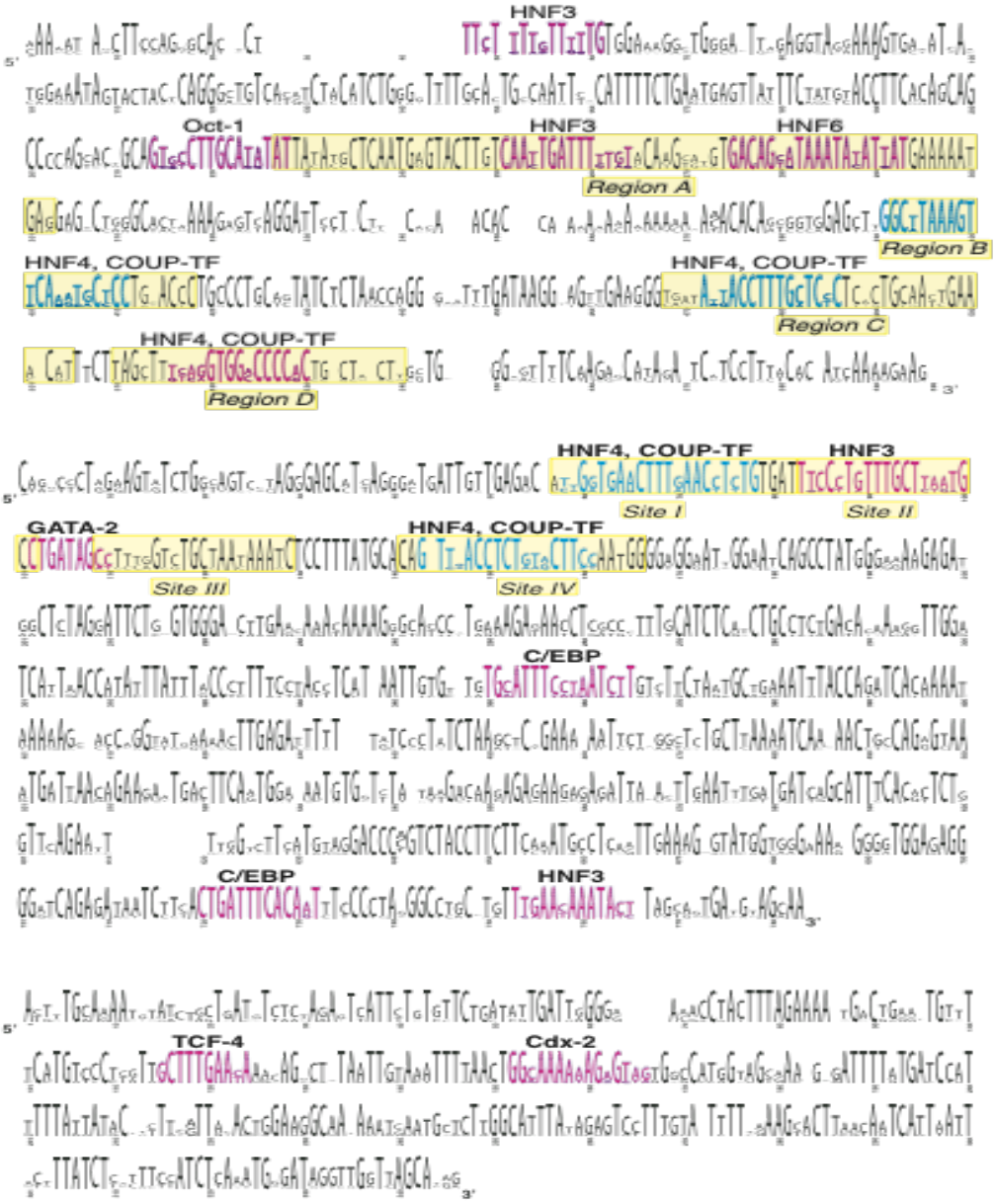

Figure 21. LOGO representations of sequence alignment of conserved elements from 5'-flanking regions of *OTC* genes from eight mammalian species.

OTC promoters (A), liver-specific enhancers (B) and upstream regulatory regions (C) were aligned using Clustal W and analyzed with CLOVER to identify transcription factor binding sites. Experimentally verified transcription factor binding sites are shown in cyan. Predicted binding sites are shown in magenta. Yellow boxes indicate regions that bind transcription factors within rat *Otc* promoter and enhancer. Up-arrows indicate positions of SNPs in the regulatory regions of the human *OTC* gene; a previously identified mutation in *OTC* promoter (clinically important SNP) is shown in red and italic typeface.

We queried build 135 of Single Nucleotide Polymorphism Database (dbSNP) to determine if the sequence changes may have been identified previously. SNP1 and SNP2 were not among the 10 polymorphisms identified in the *OTC* promoter (Figure 22). To further exclude the possibility that the sequence changes may be present in the general population, we conducted SNP genotyping in 362 healthy individuals at these two positions. All tested individuals had wild-type alleles at DNA positions corresponding to SNP1 and 2 (Table 21). Tested populations included a diverse cohort of subjects from the human genome diversity panel [217] and 256 subjects of Northern European Caucasian descent from the FAMUSS study [218]. Assuming equal distribution of genders in both test populations, the three sequence changes were absent in 540 alleles, corresponding to less than 0.18% frequency of the patients' A and T alleles for SNP1 and 2, respectively. This suggests that SNP1 and SNP2 may be disease-causing in patients who carry them.

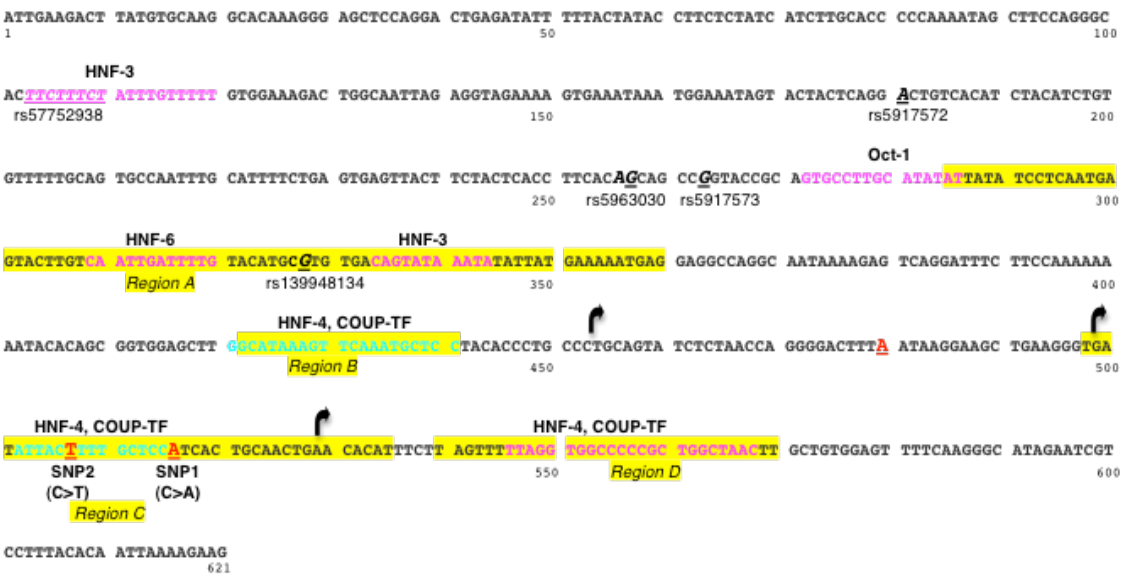

Figure 22. Mutations, polymorphisms, and transcription factor binding sites in the human *OTC* promoter.

SNP1 and SNP2, identified in this study, are shown in red and underlined typeface. Common polymorphisms are shown in bold italic and underlined typeface. Rare polymorphisms, identified through 1000 genomes project, are shown in boxed typeface. A previously identified mutation in *OTC* promoter is shown in bold italics typeface. Regions A, B, C and D (yellow) correspond to binding sites of transcription factors within the rat *Otc* promoter. Predicted transcription factor binding sites are shown in magenta. Binding sites for transcription factors that bind to regions B and C in the rat *Otc* promoter are shown in cyan. Angled arrows indicate transcription starts sites.

Table 21. Allele frequencies in populations analyzed by SNP genotyping.

		SNP1		SNP2	
		A (patient)	**C** (wild-type)	**T** (patient)	**C** (wild-type)
Number of screened individuals.		N=362		N=361	
Ethnicity	Caucasian	0	188	0	187
	Hispanic	0	13	0	13
	African American	0	131	0	137
	Unknown	0	30	0	24

SNP1 NC_000023.10:g.38211844C>A, NG_008471.1:g.5109C>A

SNP2 NC_000023.10:g.38211835C>T, NG_008471.1:g.5100C>T

The locations of SNP1 and SNP2 suggest that the mutations could affect binding of HNF4 and COUP-TF to their binding site and decrease transcription of the OTC gene. DNA pull-down assays, conducted by Abigail LaBella, were used to determine if SNP1 and SNP2 affect binding of HNF4 to region C of the OTC promoter. Binding of COUP-TF to region C was not detectable using DNA pull-down assay. Additionally, functional studies in cultured cells have shown that region C may not be important for repression of OTC transcription by COUP-TF [55]. To isolate the effects of SNP1 and SNP2 on region C, wild-type (pWT) and mutated (pSNP1 and pSNP2) probes were used in DNA pull-down experiments. DNA sequence of similar length located outside OTC promoter and non-biotinylated DNA were used as negative controls. The migration of two bands (approximately 44 and 52 kDa), indicative of HNF4 binding to pWT, pSNP1 and pSNP2 (lanes 1, 5 and 7 in Figure 23), were in good agreement with the predicted sizes of two HNF4 isoforms, translated from alternatively spliced transcripts [248]. The intensities of bands which indicate HNF4 binding to either pSNP1 or pSNP2 were compared to the intensity of bands which indicate HNF4 binding to pWT. The bands in lane 5 were approximately 37±7% (mean±SEM; n=3) less intense than bands in lane 1 suggesting that pSNP1 binds about 40% less HNF4 than pWT (Figure 23). The bands in lane 7 were approximately 67±13% (mean±SEM; n=3) less intense than bands in lane 1, similarly suggesting that pSNP2 binds less HNF4 than pWT (Figure 23). These data, combined with absence of SNP1 and SNP2 in test populations, suggest that these two mutations are likely to cause OTCD by reducing binding of HNF4 to the OTC promoter, leading to reduced transcription of the OTC gene and abundance of its product.

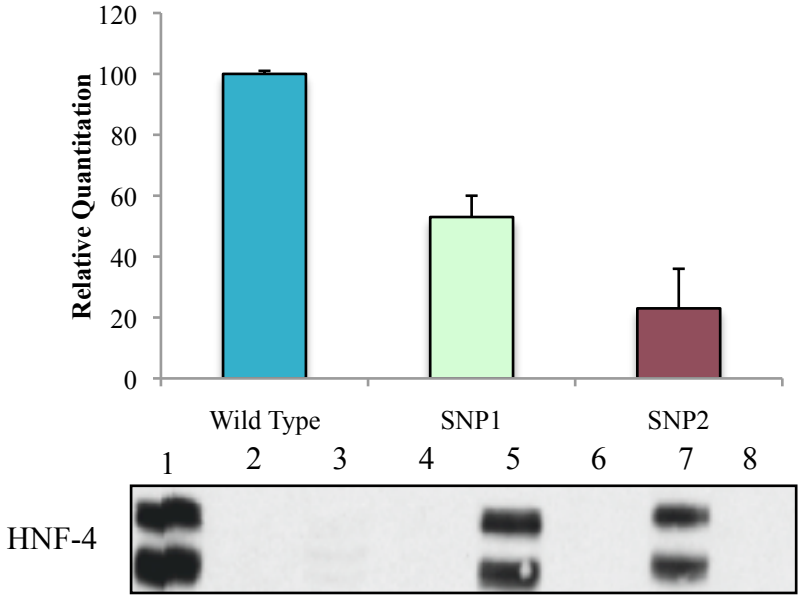

Figure 23. Reduced binding of HNF4 to sites with SNP1 and SNP2 mutations.

DNA pull-down assays were used to compare binding of HNF4 to its wild type and mutated binding sites. Lanes: 1 – biotinylated DNA with wild-type binding site; 2 – unlabeled DNA with wild-type binding site; 3 – non-specific biotinylated DNA; 4 – pull-down without DNA; 5 – biotinylated binding site containing SNP1; 6 – non-biotinylated binding site containing SNP1; 7 – biotinylated binding site containing SNP2; 8 – non-biotinylated binding site containing SNP2. Results are representative of three independent experiments.

Discussion

Herein we report two new mutations in the promoter of the *OTC* gene and show that including the *OTC* regulatory regions in mutation analysis of the *OTC* gene will likely increase the number of patients with definitive molecular diagnosis of their disease. Our bioinformatic analysis of the 5'-flanking regions of eight mammalian *OTC* genes identified the promoter and LSE of the human *OTC* gene based on similarity to rat *Otc* [196]. A new conserved element located approximately 2 kb upstream of the LSE was also identified. CLOVER analysis extended our findings by identifying new potential transcription factor binding sites that extend the previously identified conserved regions [51,55,73].

We used CLOVER software to determine that binding sites for HNF4, COUP-TF and C/EBP in the promoter and LSE of rat *Otc* gene are conserved in other mammals. The CLOVER software also predicted binding of HNF3 and HNF6 to region A, and binding of HNF4 and COUP-TF to region D of the *OTC* promoter. Reporter gene assays in cultured cells have identified regions A and D as negative and positive *cis*-regulatory elements of the rat *Otc* gene [51], yet transcription factors that bind and mediate the effects of these two regions have not been identified until now. Our analysis also identified conserved binding sites for HNF3 and Oct-1 upstream of region A in the *OTC* promoter and HNF3 and C/EBP in the conserved region that extends downstream of the four known regulatory sites in the *OTC* enhancer (Figure 21). Regulation of *OTC* transcription by either HNF3 or GATA-2 has not been shown before. We queried results of ENCODE project [183,249] and found that binding of HNF3 and C/EBP transcription factors, and p300 transcription co-activator to *OTC* LSE was detected in HepG2 cells

using chromatin immunoprecipitation followed by high throughput sequencing. This increased our confidence in the predicted binding of HNF3 to LSE and suggests a role for this transcription factor in regulating *OTC*.

Mutations in transcription factor binding sites cause disease, including urea cycle disorders [95,196]. Up until now, a combination of DNA sequencing and studies of decreased transcription factor binding and/or gene function due to a mutation have been required to gain acceptance for this kind of causation. In this study, two sequence changes were found in the promoter region of three patients with OTCD, and these changes were absent in 540 alleles from healthy individuals of diverse ethnic background. SNP1 is located immediately adjacent to HNF4 binding site while SNP2 affects a base pair that is conserved in most HNF4 binding sites [248]. Both sequence changes decreased HNF4 binding to the mutated binding sites *in vitro* (Figure 23). Interestingly, although SNP1 and SNP2 decrease HNF4 binding to different extents *in vitro*, their effects on OTC activity in the patients are similar – reduction of OTC activity to about 10% of control. We can conclude that reduction in OTC activity is due to decreased transcription of the *OTC* gene, resulting in reduced abundance of *OTC* mRNA and protein.

Mutations that are not in coding regions or splice junctions, such as these, can disrupt any stage in the process of gene expression, from epigenetic regulation, transcription, and RNA editing to mRNA stability and translation. Although next generation sequencing may become a frequently used diagnostic tool in the near future, a method to identify a disease causing sequence change in a timely and cost effective manner is lacking. Even with detailed understanding of the roles that *cis*-acting elements

play in gene expression, *in vitro* studies such as reporter and DNA binding assays may not accurately correlate with the severity of disease phenotype in patients. This is well illustrated by our experiments with binding of HNF4 to its mutated binding sites. Furthermore, functional consequences of sequence changes in introns are even harder to test because that would require transformation of cultured cells with artificial chromosomes, transgenic animals or even *in vivo* genome editing. Inspection of conservation tracks for the human OTC gene in the UCSC Genome browser revealed several putative regulatory regions in introns 1, 2, 3, 4 and 6 that are conserved in placental mammals. It is possible that mutations in these regions could cause OTCD but their role in expression of OTC must first be established.

Given the cost and time needed for functional tests of mutated sequences, we propose a practical set of steps for diagnosing disease-causing mutations outside of coding regions and splice junctions of genes: 1) Combine known information about regulation of gene expression with bioinformatic analysis of its sequence including conservation of sequences across species, prediction of transcription factor binding sites, and sequence variation in the general population; 2) Identify sequence changes in patients' DNA and determine whether they reside in functionally important regions based on analysis in step 1; 3) If possible, determine whether patients parents are carriers of detected sequence change, and/or whether the mutation segregates with the disease in patient's relatives; 4) Determine frequency of each allele in the population; and 5) Use functional tests, if feasible, to show that detected sequence change indeed affects gene expression. This framework can be extended to other conserved regions of the *OTC* gene and to other genes.

Acknowledgments

This work was supported by public health service grants R01DK047870 and R01DK064913 from the National Institute of Diabetes Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services.

Chapter 5: Nitrogen load sensing pathways

Introduction

The nutrient sensing pathways mTOR, eIF2/GCN2 and AMPK have been shown to regulate carbohydrate and fat metabolism [30,31,33,34,250]. Recent studies suggest that these pathways also control amino acid metabolism in the liver through changes in phosphorylation [23,251,252,253] and may affect changes in the urea cycle enzymes through regulation of their *trans*-acting factors [68,69,74,75,254,255,256]. While studies have shown that urea cycle protein expression changes in response to diet [4,5,84], the signaling networks that sense and coordinate this response have not been identified. Although numerous studies have been conducted to test the effects of nitrogen load on individual genes, and recent studies have examined the effects of a carbohydrate only diet or fasting on the expression of genes, proteins, and signaling pathways [257], no global study, tracking the coordinated changes in gene and protein expression, has been conducted using controlled dietary protein.

Several lines of evidence suggest that AMPK, eIF2 and mTOR may be involved in the regulation of ureagenesis. As key nutrient sensors in many cell types [30,31,33,34,250], their activity in cultured cells depends on the concentration of amino acids, and specifically leucine in the culture medium [23,258,259,260,261]. Therefore, it is possible that the branched chain amino acids (BCAA) directly regulate the expression of the urea cycle enzymes and urea production mTOR, AMPK and eIF2 signaling. Rats fed a high (42%) protein diet exhibited increased phosphorylation of mTOR, while phosphorylation of AMPK and eIF2 decreased compared to animals fed a control (12%

protein) diet [23]. Similarly, urea cycle enzyme mass and activity have been shown to increase upon administration of a high protein diet [3,262]. Along with these data, additional evidence to suggest a role for mTOR, AMPK, and eIF2 in regulation of the urea cycle genes includes the downstream targets of these pathways. Peroxisome proliferator-activated receptor gamma co-activator (PGC-1α) is activated by the AMPK pathway [65,66,67] and then directly regulates the expression of HNF4α (hepatic nuclear factor 4 alpha) and the glucocorticoid receptor (GR) [68,69]: transcription factors known to control expression of CPS1 and OTC [55,70,71,72,73]. Furthermore, the transcription factors CCAAT/element binding protein (C/EBP) and cAMP response element binding (CREB) are phosphorylated by AMPK [74,75] and regulate NAGS, CPS1, OTC, ASS, and ARG1 [47,63,73,76,77].

Materials and Methods

Long Term Feeding Study

Syngeneic adult male C57BL/6 mice (Jackson Labs) in three groups of 8 were fed an isocaloric chow (Harlan Laboratories) with protein content of 10% (low), 20% (recommended- control), or 60% (high) *ad libitum* for 11 days to elicit long-term adaptive response to different nitrogen loads. Mice were weighed and their food and water consumption was measured each day. On the 12th day, mice were sacrificed and the livers were excised and immediately frozen in liquid nitrogen.

Acute Feeding Study

Syngeneic adult male C57BL/6 mice were fed isocaloric control (20% protein, CP) or high (60% protein, HP) protein diets on a fast/feast cycle consisting of 8 hours

fasting, with unlimited access to water, followed by 16 hours of *ad libitum* feeding for the purpose of measurement of RNA and protein levels in response to eating a meal. Mice were allowed 11 days to acclimate to this diet and on the 12th day of the study mice were sacrificed immediately before scheduled food reintroduction (T0), and at 30 min (T30), 60 min (T60) and 120 min (T120) after reintroduction of food. Eight mice were included at each time point and their livers were collected and immediately frozen in liquid nitrogen.

RNA sample preparation and Microarray Processing

Hierarchical Clustering Explorer (HCE) 3.5 power analysis tool [263], developed at Children's National Research Center for Genetic Medicine, was used to determine that 3 replicates would be sufficient to allow for detection of significant differences in mRNA levels for 60-80% of probe sets present on each microarray.

RNA was isolated from frozen liver using TRIzol reagent (Invitrogen) and its quality was tested by analyzing RNA concentration through 260/280nm absorbance and visualization of 28s and 18s ribosomal bands. RNA was converted to double-stranded cDNA using One-Cycle cDNA Synthesis Kit (Affymetrix) according to manufacturer's instructions, and the cDNA was amplified to ensure greater than 4-fold amplification. The cDNA was then transcribed and biotin tagged with the IVT labeling kit (Affymetrix), according to manufacturer's instructions. The cRNA was purified with the RNAeasy Kit (QIAGEN), and fragmented to approximately 200 bp. To ensure proper fragmentation, the IVT fragmented and unfragmented samples were visualized on an agarose gel. Following completion of all quality control checks, cRNA was hybridized to GeneChip Mouse Genome 430 2.0 (Affymetrix) and the microarray chips were washed and stained

on an Affymetrix Fluidics Station 400 within the Center for Genetic Medicine's Expression Profiling core facility, according to manufacturer's instructions. Fluorescent images were scanned using a Hewlett-Packard G2500A Gene Array Scanner [264].

Analysis of Transcriptional Profiling Results

Affymetrix image data was analyzed using Probe Microarray Suite (MAS) version 5.0 and dCHIP [265] probe set algorithms to highlight genes both algorithms identified as having significant expression level change from the control and to eliminate false positives. Genes identified by each algorithm were imported into Partek software package (Partek Incorporated), normalized by Log2, and analyzed by 2-way ANOVA with p-value of $P<0.05$. Comparisons for the long-term adaptation study included control vs. high, control vs. low, and high vs. low diets. Comparisons for the acute adaptation study included control T0 vs high T0, control T30 vs high T30, control T60 vs high T60, control T120 vs high 120, control T0 vs control T30, control T0 vs T60, control T0 vs control T120, high T0 vs high T30, high T0 vs high T60, and high T0 vs high T120.

ANOVA results were imported into Ingenuity Pathway Analysis (IPA) software (Ingenuity) using Affymetrix probe set IDs as identifiers and absolute fold-change for observation values. The long-term adaptation data had expression value cutoffs of -1.7 and 1.7. The acute adaptation data had expression value cutoffs of -1.5 to 1.5.

Protein Sample Preparation

As with transcriptional profiling analysis, HCE 3.5 power analysis tool [263] was used to determine that 3 replicas would be sufficient to allow for detection of significant differences in liver protein abundance from each of the diets.

Total liver protein was extracted from frozen and ground mouse tissue by homogenization of tissue in 50 mM Tris-acetate buffer, pH 7.5 containing 250 mM sucrose, 1mM EDTA, 1% Nonident P40 (NP40), and Complete Mini Protease Inhibitor Cocktail (Roche). Proteins were purified and desalted using Micro Bio-Spin Chromatography Columns (Bio-Rad) in Tris buffer. Proteins were then resolved by electrophoresis on SDS-PAGE gel, stained, and 1 mm wide sections were excised from the gel. Proteins in the gel were digested using trypsin, and the resulting peptides were extracted from the gel using organic solvents.

Mass Spectrometry and Analysis of Proteomic Profiling Results

Concentrated peptides from each band were injected via an autosampler (6 uL) and loaded onto a Symmetry C18 trap column (5 μm, 300 μm i.d. x 23 mm, Waters) for 10 min at a flow rate of 10 μL/min, 100% A. The sample was subsequently separated by a C18 reverse-phase column (3.5 μm, 75 μm x 15 cm, LC Packings) at a flow rate of 250 nL/min using an Eksigent nano-hplc system (Dublin, CA). The mobile phases consisted of water with 0.1% formic acid (A) and 90% acetonitrile (B). A 65 minute linear gradient from 5% to 60% B was employed. Eluted peptides were introduced into the mass spectrometer via a 10 μm silica tip (New Objective Inc., Ringoes, NJ) adapted to a nano-electrospray source (ThermoFisher Scientific). The spray voltage was set at 1.3 kV and the heated capillary at 200 °C. The LTQ-Orbitrap-XL (ThermoFisher Scientific) was operated in data-dependent mode with dynamic exclusion in which one cycle of experiments consisted of a full-MS in the Orbitrap (300-2000 m/z) survey scan and five subsequent MS/MS scans in the LTQ of the most intense peaks using collision-induced

dissociation with the collision gas (helium) and normalized collision energy value set at 35%.

Each file was searched for protein identification using the Sequest algorithm in the Bioworks Browser software (ThermoFisher Scientific) against the Uniprot database (downloaded June 2009) indexed mouse species and for fully tryptic peptides, 2 missed cleavages and no modifications. DTA generation parameters were Peptide Tolerance of 50 ppm and Fragment Ion Tolerance of 1 Da.

Results

Physiological effects of different feeding regimens on mice

Mice fed low (10%), control (20%), or high (60%) protein diet did not differ in body weight following 12 days of *ad libitum* (Figure 24A) or time-restricted (Figure 24B) feeding. Spontaneous food intake was lower in mice fed the high protein diet (Figure 24C and D), most likely due to the satiety signaling to the digestive system following consumption of increased dietary protein [266,267]. While urea excretion was not measured in this study, the high protein diet most likely results in increased urea production as documented previously [268]. Water consumption was increased in mice fed the high protein diet (24E and F), presumably to allow for increased urea excretion through urination. No significant differences in weight, food intake, or water intake were noted in mice on the *ad libitum* control and low protein diets (Figure 24A, C and E).

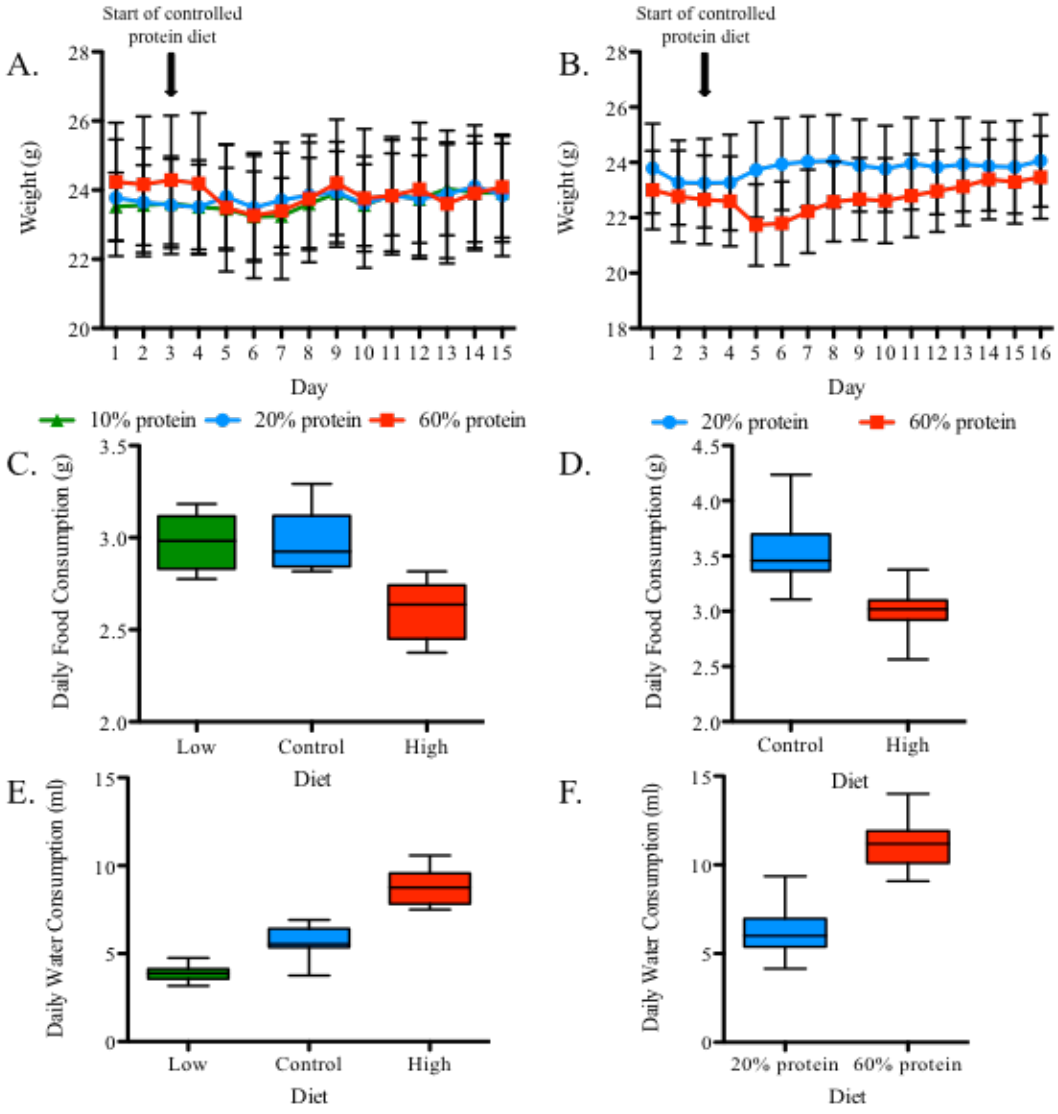

Figure 24. Weight, food consumption, and water consumption of mice in 12 day feeding studies.

Mice fed an *ad libitum* diet of low (10%), control (20%) or high (60%) protein (n=8/diet) (A) and mice fed on a time-restricted diet of control (20%) or high (60%) protein (n=24/diet) (B) showed no difference in weight during 12 days of feeding. *Ad libitum* fed mice on the high protein diet consumed less food (C) and more water (E) than mice on the low or control protein diets. High protein fed mice on the time-restricted diet consumed less food (D) and more water (F) than mice on the control protein diet.

Changes in mRNA expression in mice fed *ad libitum*

mRNA was extracted from the livers of mice fed *ad libitum* high, control, or low protein diet for a 12 day period. Microarrays consisting of 45101 probe sets corresponding to 21554 unique known genes, were used to determine differences in liver mRNA expression. Diet significantly changed ($p<0.05$) a total of 13771 genes, with the greatest difference between the high and low protein diets, 13141 genes. Fewer genes significantly differed between the control diet and the low diet, 3926 genes, or high diet, 5672 genes.

Hierarchical clustering analysis was carried out on the significant genes in each diet group (Figure 26). The replicates of each treatment (n=4) were clustered together, indicating the effectiveness of the treatments and robustness of the filtering criteria.

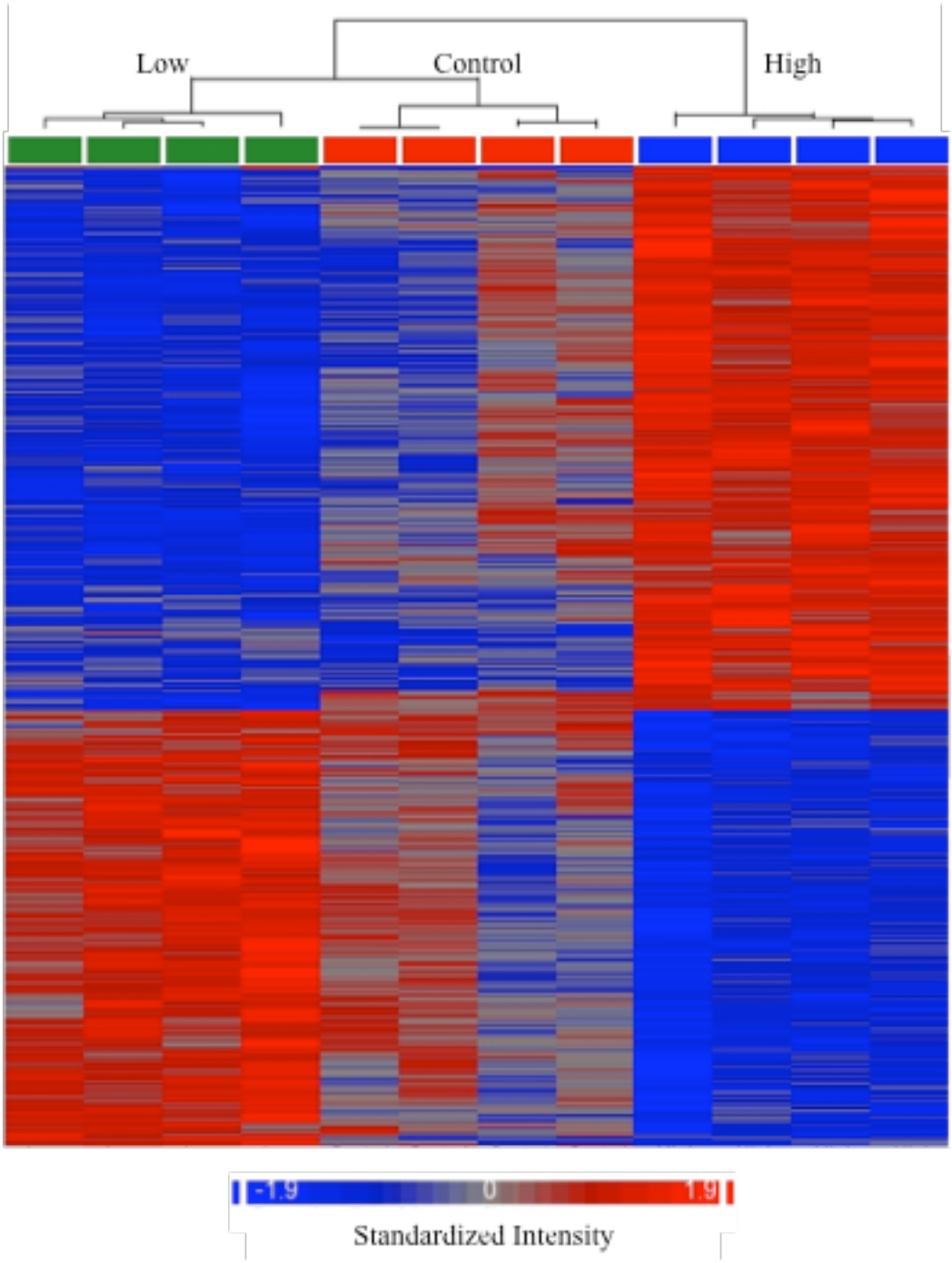

Figure 25. Hierarchical clustering of genes from mice fed *ad libitum* diet of low (10%), control (20%) or high (60%) protein.

mRNA from four mice/diet was collected and assayed for gene expression by microarray. Expression profiles of genes clustered by diet.

Changes in mRNA expression in mice fed a time-restricted diet

Microarrays were also conducted using mRNA extracted from livers of mice fed a high or control diet on a time-restricted 8 and 16 hour fasting-feeding schedule for 12 days. Livers were collected immediately prior to reintroduction of food, and 30, 60, and 120 minutes after reintroduction of food on the 12th day. Of the 21554 genes, a total of 3970 genes differed significantly ($p<0.05$) by both diet and time. Overall, the time-restricted feeding schedule led to reduced variability in expression of genes within each group as demonstrated by the total number of genes that differed between the high and control diet, 7000. This feeding method resulted in many more significantly different genes than the *ad libitum* high and control diets (5672) and subsequent analysis of the mRNA expression was based on the time-restricted diet data.

Within this group of 7000 genes, a greater difference was noted between diets than between samples collected at different time points following reintroduction of food. Accordingly, hierarchical clustering analysis of significant genes clustered first by diet and then by time point (n=3; Figure 26). This clustering was suggested by the number of significant genes that differed between diets at a particular time point including 2841 at 0 min, 3094 at time 30 min, 3982 at time 60 min, and 2590 at time 120 min. In contrast, there was a small number of genes whose expression differed between time points within a diet (Table 22), suggesting that these changes represent the pace of transcriptional activation following feeding. Table 23 shows the top 50 genes that differed between mice on the high and control diets and by time of collection.

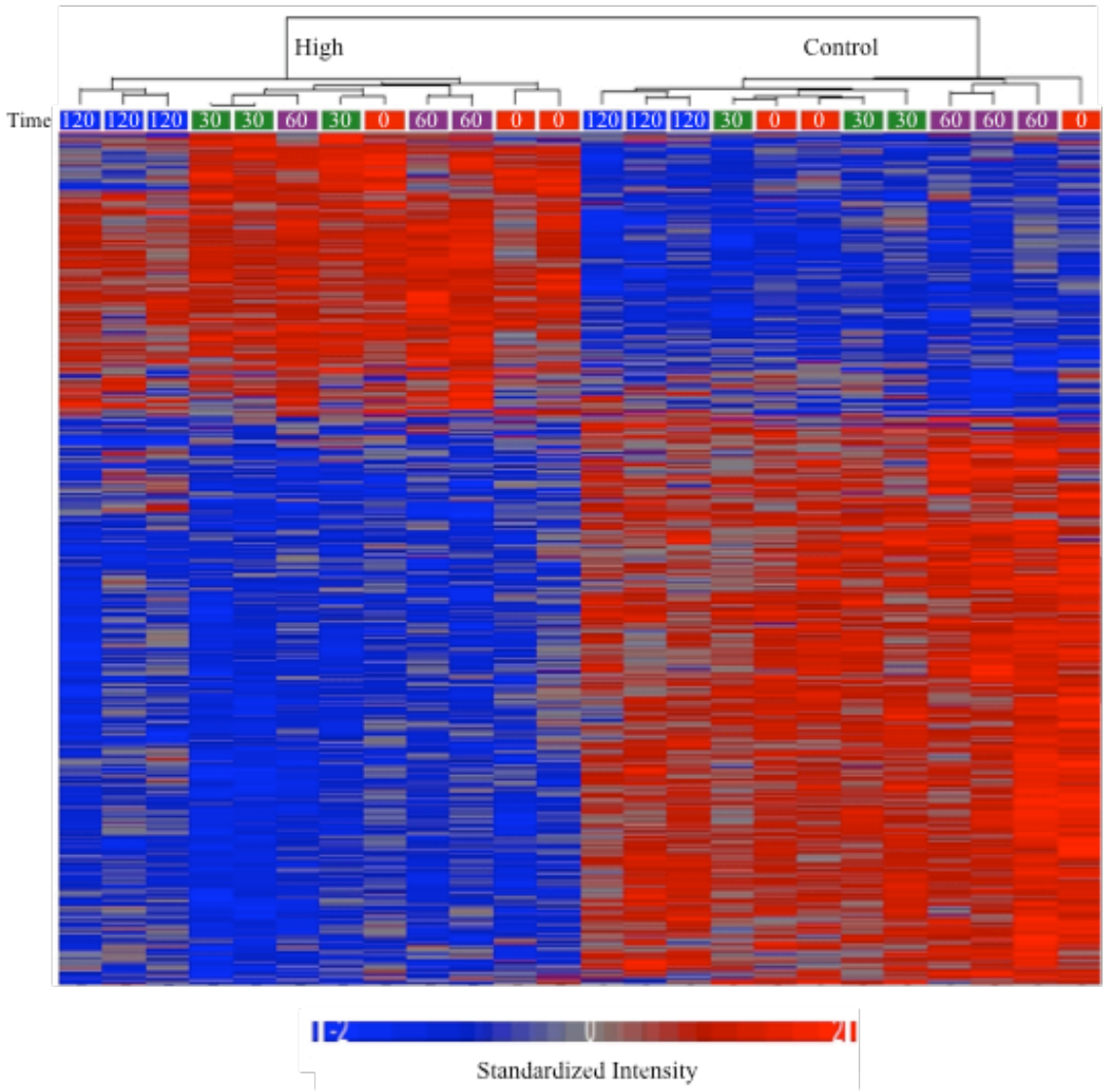

Figure 26. Hierarchical clustering of genes from mice fed a time-restricted diet of control (20%) or high (60%) protein with samples collected immediately before reintroduction of food, Time=0, or 30, 60, or 120 minutes after reintroduction of food.

mRNA from three mice/diet/time point was collected and assayed for gene expression by microarray. Expression profiles of genes clustered by diet and then by time.

Table 22. Changes by number of genes over time after reintroduction of food.

	20% Protein (control)	60% Protein (high)
T0 vs T30	1132	1141
T0 vs T60	1515	1405
T0 vs T120	2454	2307

Changes in the number of significantly different gene expression ($p<0.05$) over time, was reflected in both the control (20%) and high (60%) protein fed diet groups.

Table 23. Top 50 genes that differed between diets and time of collection.

Gene Symbol	Gene Title
Gldc	glycine decarboxylase
C4bp	complement component 4 binding protein
Antxr2	anthrax toxin receptor 2
Serpina12	serine (or cysteine) peptidase inhibitor, clade A
Nat8	N-acetyltransferase 8 (GCN5-related, putative)
Gm8615	glucosamine-6-phosphate deaminase 1 pseudogene
Nrp1	neuropilin 1
Rpl22	ribosomal protein L22
Orm2	orosomucoid 2
Gsto1	glutathione S-transferase omega 1
Eif3e	eukaryotic translation initiation factor 3, subunit E
Ces1e	carboxylesterase 1E
Gpr98	G protein-coupled receptor 98
Cyp4a31	cytochrome P450, family 4, subfamily a, polypeptide 31
Nlrp6	NLR family, pyrin domain containing 6
Ccl9	chemokine (C-C motif) ligand 9
Tardbp	TAR DNA binding protein
Clec2h	C-type lectin domain family 2, member h
Il1rap	interleukin 1 receptor accessory protein
Dpys	dihydropyrimidinase
Slc13a2	solute carrier family 13
Dnase2a	deoxyribonuclease II alpha
Ces1d	carboxylesterase 1D
Ikbkg	inhibitor of kappaB kinase gamma
Lbp	lipopolysaccharide binding protein
Rpl10	ribosomal protein 10
Ldlr	low density lipoprotein receptor
Tardbp	TAR DNA binding protein
Chka	choline kinase alpha
Agxt2l1	alanine-glyoxylate aminotransferase 2-like 1
Ftl1/Ftl2	ferritin light chain 1 / ferritin light chain 2
Gabarapl1	gamma-aminobutyric acid A receptor-associated protein-like 1
Fam55b	family with sequence similarity 55, member B
Tor1a	torsin A
Retsat	retinol saturase
Tnfrsf19	tumor necrosis factor receptor superfamily, member 19
Nt5e	5' nucleotidase, ecto
Apol9a / Apol9b	apolipoprotein L 9a / b
Slc22a18	solute carrier family 22
Ces1d	carboxylesterase 1D
Eef1d	eukaryotic translation elongation factor 1 delta
Suox	sulfite oxidase
Gamt	guanidinoacetate methyltransferase

Trfr2	transferrin receptor 2
Agxt	alanine-glyoxylate aminotransferase
1500017E21Rik	RIKEN cDNA 1500017E21 gene
Hmgn2 / Hmgn4	high mobility group nucleosomal binding domain 2 / 4
Rps3a	ribosomal protein S3A
Eps8l2	EPS8-like 2
Hmgcs2	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2

mRNA expression profiling of gene networks and pathways regulated by diet

mRNA expression profiles and observed expression changes in samples from the time-restricted diet study were analyzed to explore metabolic and signaling pathways that respond to diet using IPA software. The pathways in which the most genes changed expression between diets included lipid metabolism and small molecule biosynthesis, which were decreased in mice fed the high protein diet while amino acid metabolism was increased in mice fed the high protein diet. In addition, comparison of genes from the fasted mice, T0- before reintroduction of food, also identified energy production through fatty acid oxidation as the function in which the largest number of genes increased in expression in mice fed the control diet. Similarly, the pathway of valine, leucine, and isoleucine degradation was up-regulated in the high protein diet reflecting a mechanism for maintenance of the body's nitrogen balance following consumption of a high protein meal. In the comparisons between samples collected at different times from mice fed the same diet, genes involved in initiation and activation of transcription increased in expression 30, 60, and 120 minutes after reintroduction of food, thus reflecting increased gene transcription in response to consumption of either diet.

Changes in protein expression in mice fed *ad libitum*

Proteins were isolated from the livers of mice fed *ad libitum* high, control, or low protein diets (n=3) for 12 days, separated using denaturing conditions with SDS-Page electrophoresis, extracted using organic solvents, and subject to identification and spectral counting using mass spectrometry. A total of about 4000 of the most abundant liver proteins were identified using this method and the fold-change of protein expression was quantified between diets using spectral counting. Within this group of proteins, 1821

unique proteins were identified in all diets. Expression changes were determined using fold-change between two diets with cutoffs of ≤ -1.5, for down-regulation, or ≥ 1.5, for up-regulation. The liver protein profiles of mice fed the high protein diet showed that 262 proteins were up-regulated while 569 proteins were down-regulated, compared to mice fed the low protein diet. Fewer proteins were up- or down- regulated, 440 and 215 respectively, when livers of mice fed high protein diets were compared to those fed a control diet. Comparison of livers from mice fed the control protein diet to low, identified 238 up-regulated and 326 down-regulated proteins as well. Table 24 shows the top 50 proteins that differed between mice on the high and control diets and by time of collection.

Protein expression profiling identifies candidate pathways regulated by diet

Protein expression data from *ad libitum* fed mouse samples were analyzed to determine pathways that respond to diet using IPA software. Proteins whose expression significantly differed between diets were involved in up-regulation of protein synthesis, amino acid metabolism, and energy production in response to the high protein diet, and involved in up-regulation of lipid metabolism in response to the low protein diet. Within these functional groups, eIF2 signaling, valine, leucine, and isoleucine degradation, and fatty acid metabolism had the most significant p-values based on the number of genes and degree of expression changes. These pathways had greatest number of proteins whose expression changed when making comparisons between the low, control, and high protein diets.

Table 24. Top 50 proteins that differed between diets

Protein Symbol	Protein Title
CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial
FAS	Fatty acid synthase
MYH9	Myosin-9
DYHC1	Cytoplasmic dynein 1 heavy chain 1
ACLY	ATP-citrate synthase
FTHFD	10-formyltetrahydrofolate dehydrogenase
ACACA	Acetyl-CoA carboxylase 1
CH60	60 kDa heat shock protein, mitochondrial
TLN1	Talin-1
M2GD	Dimethylglycine dehydrogenase, mitochondrial
IQGA2	Ras GTPase-activating-like protein IQGAP2
PYC	Pyruvate carboxylase, mitochondrial
PYGL	Glycogen phosphorylase, liver form
ATPB	ATP synthase subunit beta, mitochondrial
C1TC	C-1-tetrahydrofolate synthase, cytoplasmic
FLNB	Filamin-B
GRP78	78 kDa glucose-regulated protein
HSP7C	Heat shock cognate 71 kDa protein
HSP72	Heat shock-related 70 kDa protein 2
MUG1	Murinoglobulin-1
CATA	Catalase
AASS	Alpha-aminoadipic semialdehyde synthase, mitochondrial
CLH	Clathrin heavy chain 1
SBP2	Selenium-binding protein 2
THIM	3-ketoacyl-CoA thiolase, mitochondrial
ACOX1	Peroxisomal acyl-coenzyme A oxidase 1
MTP	Microsomal triglyceride transfer protein large subunit
ATPA	ATP synthase subunit alpha, mitochondrial
LPPRC	Leucine-rich PPR motif-containing protein, mitochondrial
ECHA	Trifunctional enzyme subunit alpha, mitochondrial
ENPL	Endoplasmin
ACOC	Cytoplasmic aconitate hydratase
ENOA	Alpha-enolase
HYES	Epoxide hydrolase 2
GRP75	Stress-70 protein, mitochondrial
IDHC	Isocitrate dehydrogenase [NADP] cytoplasmic
RRBP1	Ribosome-binding protein 1
SARDH	Sarcosine dehydrogenase, mitochondrial
ACADV	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial
DHE3	Glutamate dehydrogenase 1, mitochondrial
SND1	Staphylococcal nuclease domain-containing protein 1
DHAK	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase

	(cyclizing)
ANXA6	Annexin A6
ASSY	Argininosuccinate synthase
EF2	Elongation factor 2
TERA	Transitional endoplasmic reticulum ATPase
PGK1	Phosphoglycerate kinase 1
SBP1	Selenium-binding protein 1
ACON	Aconitate hydratase, mitochondrial
HS90B	Heat shock protein HSP 90-beta

Nutrient Sensing pathways are regulated by protein content of diet

The main focus for these studies was to identify the signaling mechanism that senses intake of dietary amino acids. The mRNA expression data revealed that the AMPK signaling pathway (change in 13 genes, p-value 6.6×10^{-3}) is up-regulated in response to high protein diet while protein expression data revealed that the eIF2 (change in 105 proteins, p-value 5.1×10^{-52}) and mTOR (change in 67 proteins, p-value 1.09×10^{-28}) signaling pathways are also involved in amino acid sensing, and are up-regulated in response to high protein diet. The AMPK, mTOR, and eIF2 pathways share nutrient directed messaging, especially by insulin, and each pathway influences the other so that they should be considered as a single combined mechanism (Figure 27). Analysis of these pathways, which are integral to nutrient sensing and metabolism of fats and carbohydrates, identified 16 genes and 19 proteins that change in association with diet and time of tissue collection after feeding (Table 25). Within a treatment, the up- or down-regulated genes differed by diet and time of tissue collection (Table 25A, Figure 28A, B, C, and D) and the differentially regulated proteins differ by diet (Table 25B, Figure 29A, B, and C).

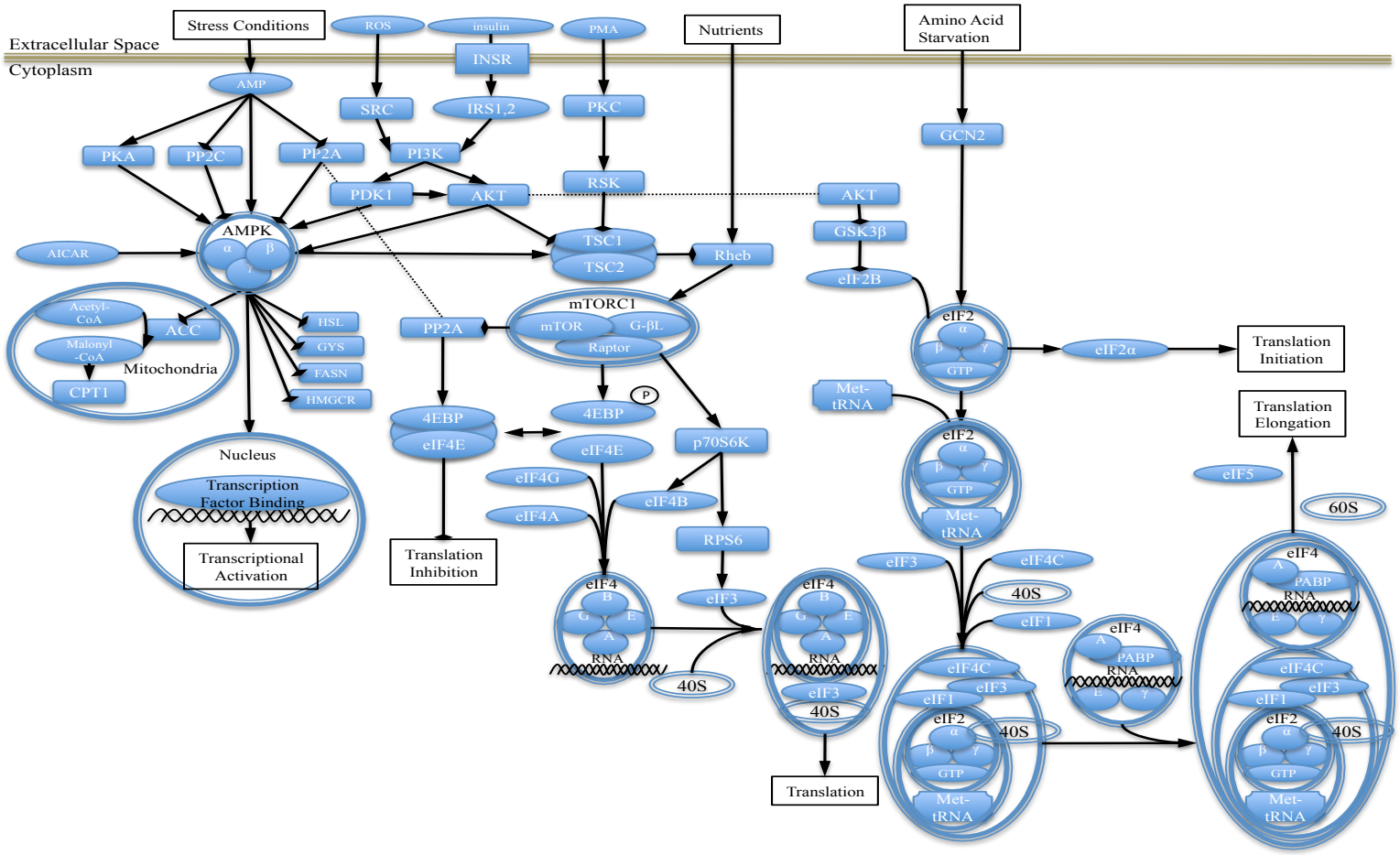

Figure 27. Pathways involved in sensing changes in protein content of diet.

The AMPK, mTOR, and eIF2 nutrient sensing pathways are integral to carbohydrate and fat metabolism and mRNA and protein profiling data suggests their importance in amino acid sensing.

Table 25. mRNA and protein expression changes associated with dietary change.

A.

gene name	High vs Control protein diets			
	T0	T30	T60	T120
PI3K	↓	↓	↑↓	↑↓
PP2A		↑	↑	↑
PP2C	↑	↑	↑	
PDK1				↑
PKA				↑
AMPKα				↑
ACC	↓	↓	↓	↓
HMGCR	↓	↓	↓	↓
FASN	↓	↓	↓	↓
GYS	↓	↓	↓	↓
TSC1		↑		
GSK3β		↑	↑	↑
eIF1				↓
eIF5				↑
40S ribosomal subunit		↓		
60S ribosomal subunit	↓	↓	↓	↓

B.

protein name	High vs Low	High vs. Control	Control vs Low
AMPKγ	↓	↑	↓
IRS1,2		↑	
PKA		↑	↓
PP2C	↓	↓	↓
PI3K	↑		
ACC	↓	↓	↓
GYS	↓	↓	↑
FASN	↓	↓	↓
RPS6		↓	
eIF2γ		↑	↑
eIF2β		↑	
eIF3	↑		↑
eIF4A			↓
eIF4B	↑	↑	
eIF4C	↓	↑	↓
eIF4E	↓	↓	↓
eIF4G	↓		↑
PABP		↓	↓
eIF5	↓		↑

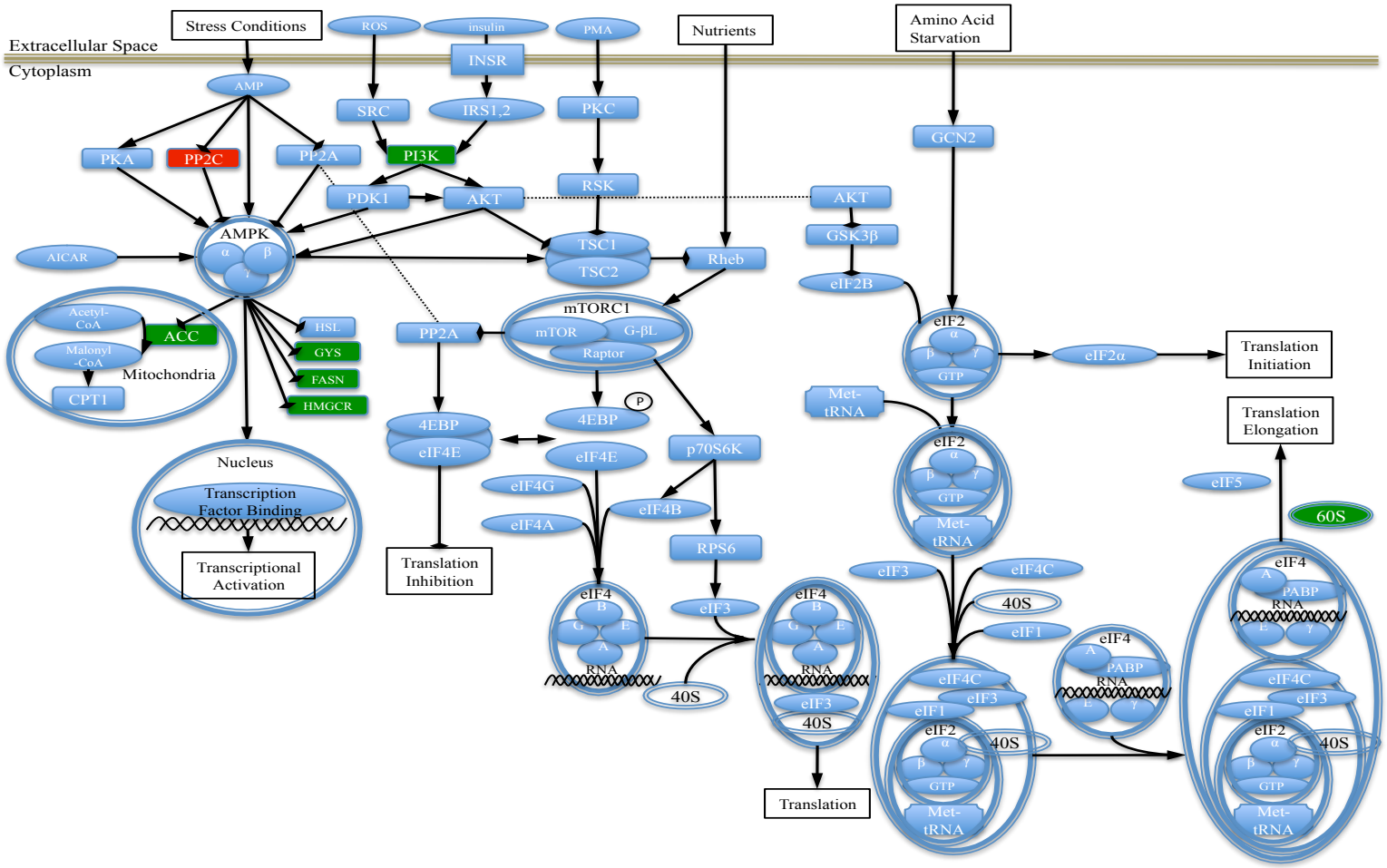

Figure 28A. mRNA expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of mRNA increased (red) or decreased (green) in mice fed the time-restricted high protein compared to the control protein diet with tissue collection at 0 min, immediately before reintroduction of food.

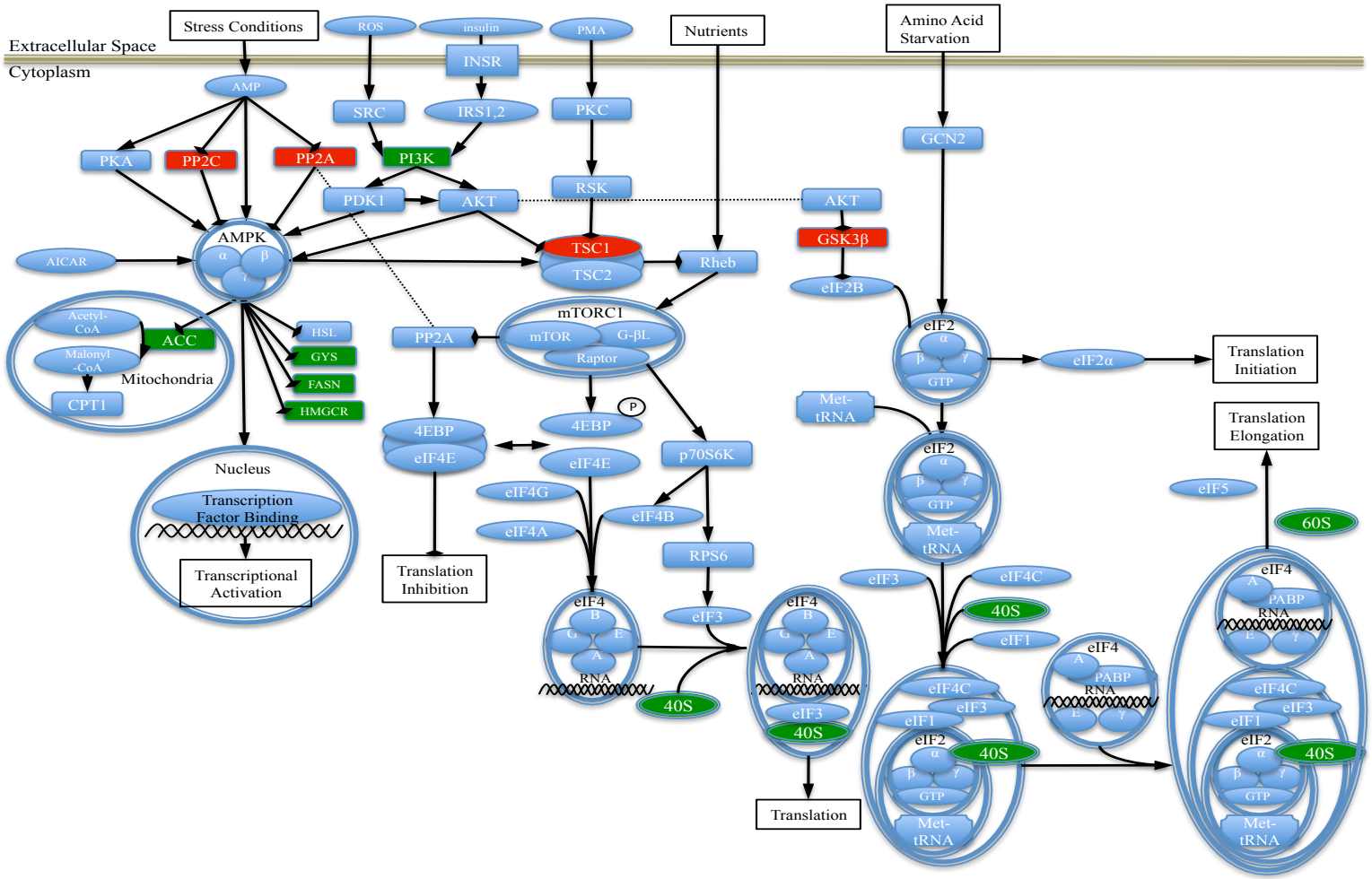

Figure 28B. mRNA expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of mRNA increased (red) or decreased (green) in mice fed the time-restricted high protein compared to the control protein diet with tissue collection at 30 min after reintroduction of food.

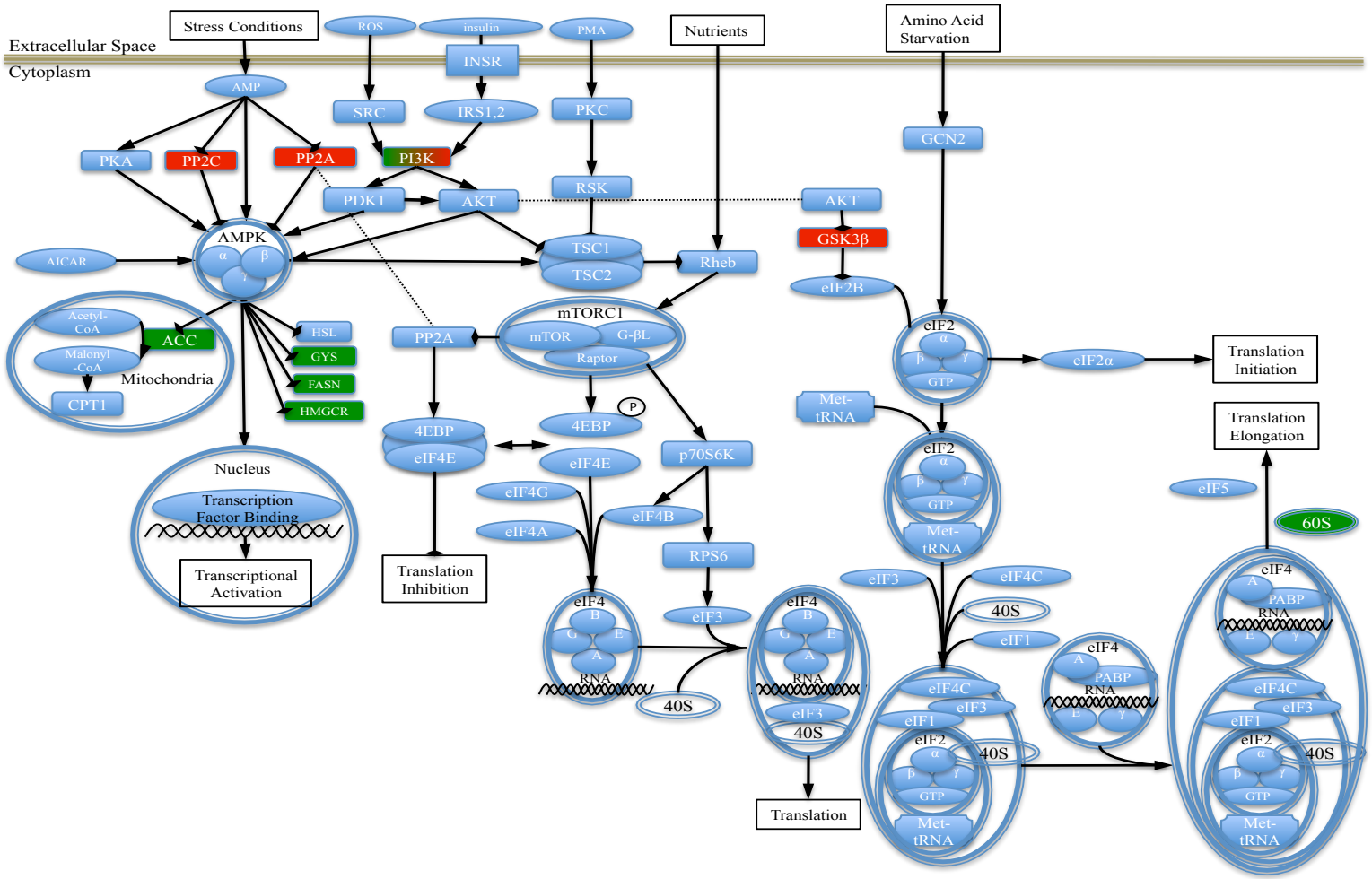

Figure 28C. mRNA expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of mRNA increased (red) or decreased (green) in mice fed the time-restricted high protein compared to the control protein diet with tissue collection at 60 min after reintroduction of food.

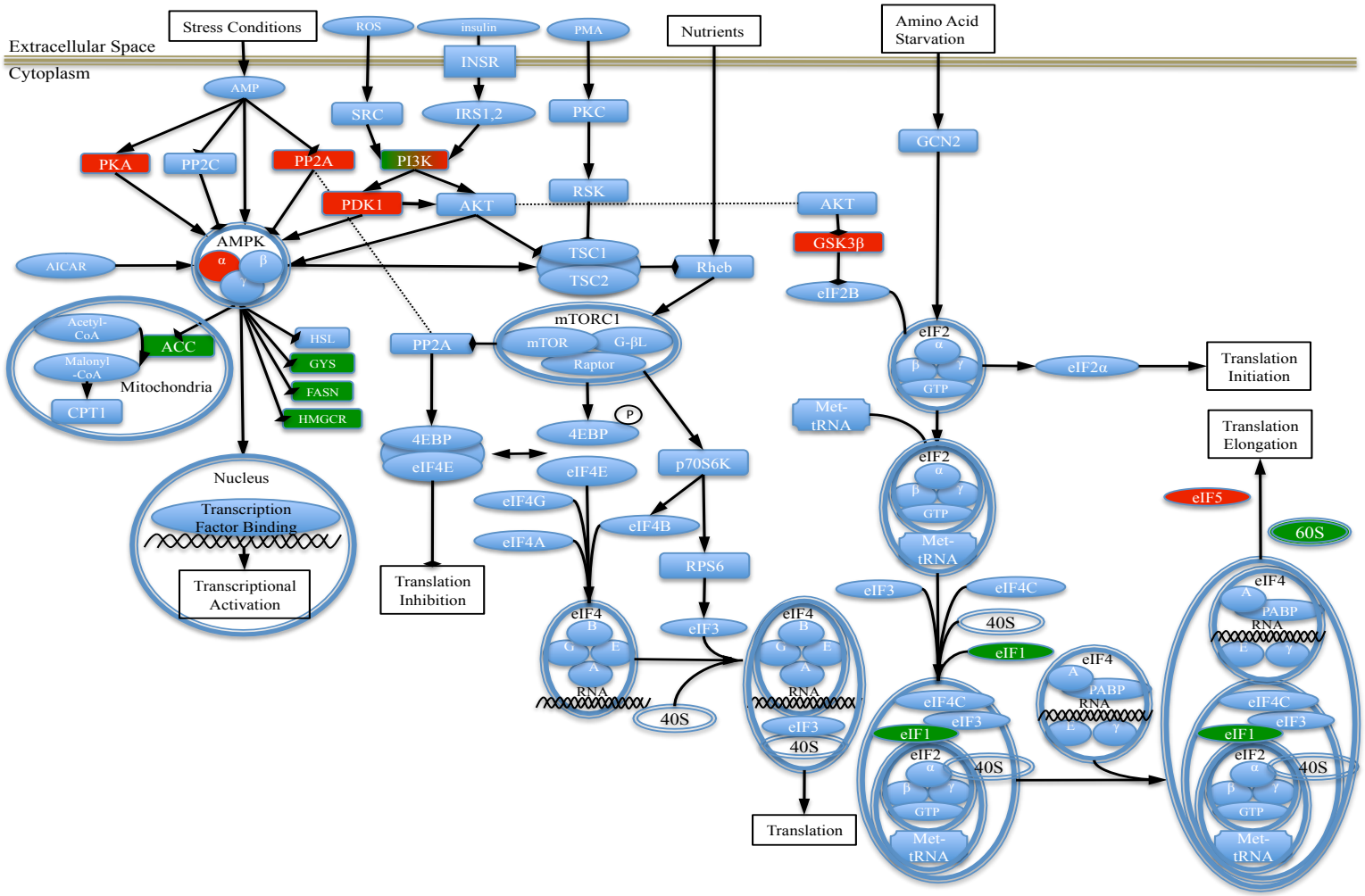

Figure 28D. mRNA expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of mRNA increased (red) or decreased (green) in mice fed the time-restricted high protein compared to the control protein diet with tissue collection at 120 min after reintroduction of food.

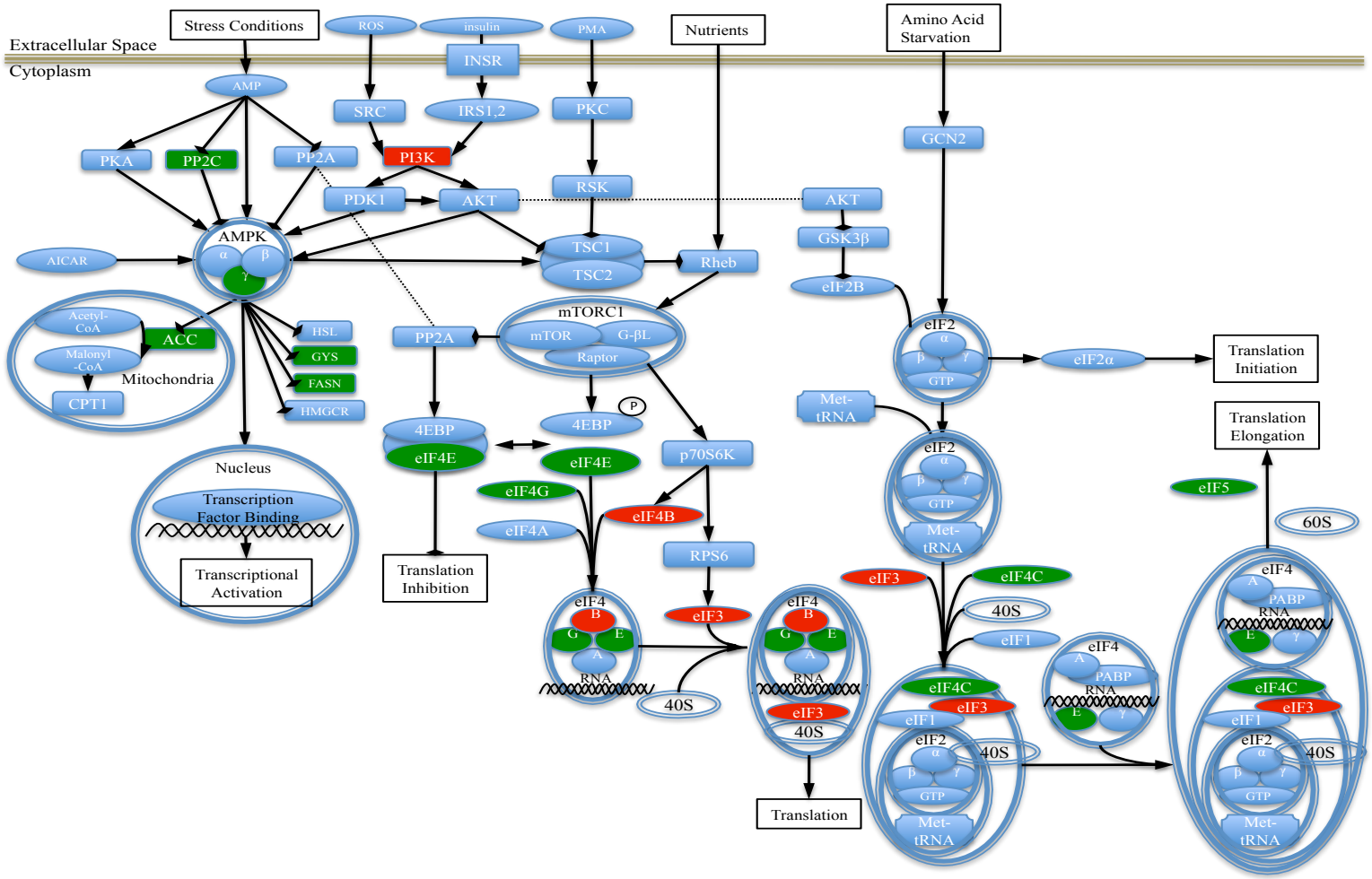

Figure 29A. Protein expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of protein changed by diet in mice fed *ad libitum*. Protein expression increased (red) or decreased (green) in mice fed the high protein diet compared to the low protein diet.

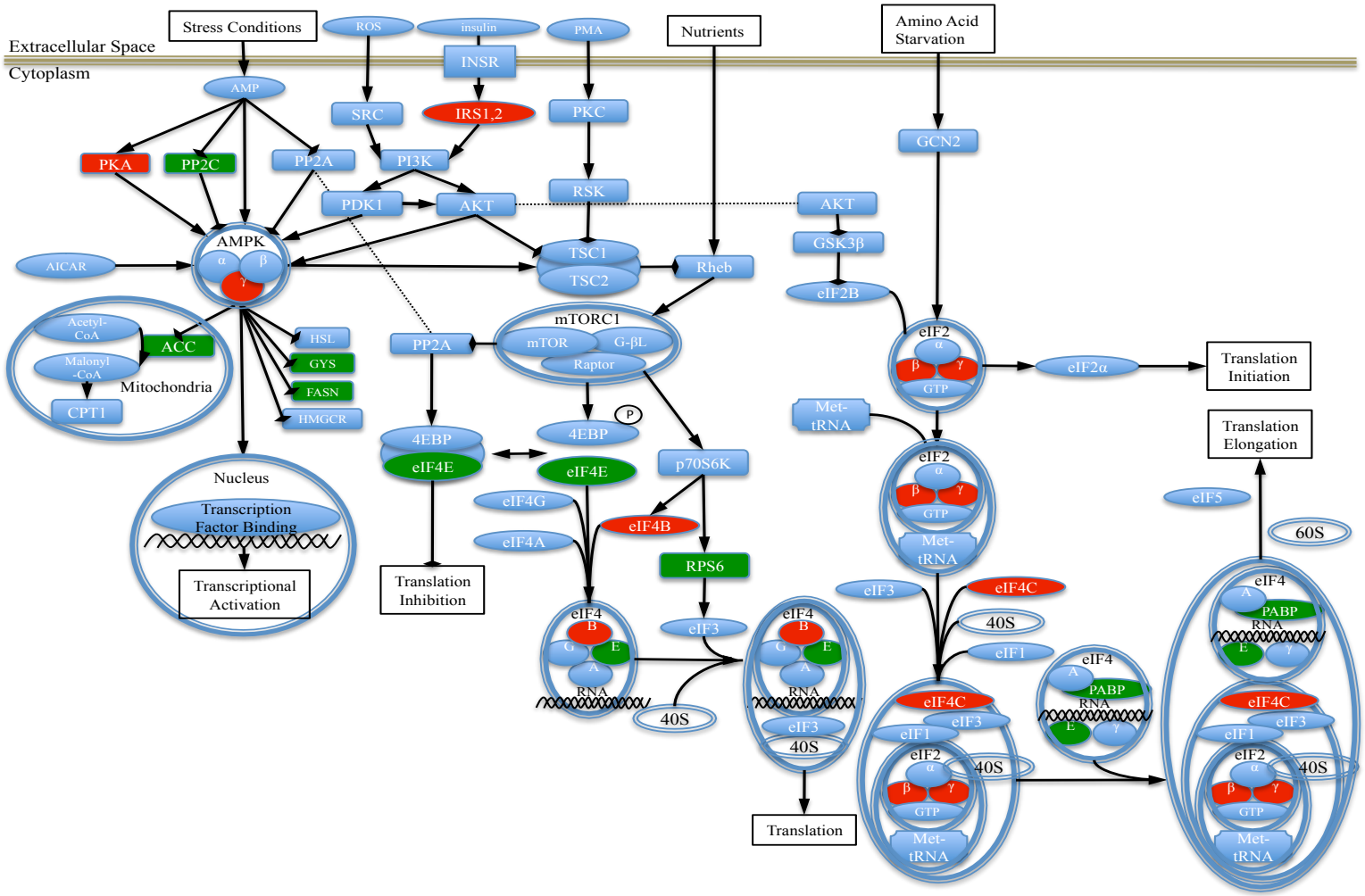

Figure 29B. Protein expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of protein changed by diet in mice fed *ad libitum*. Protein expression increased (red) or decreased (green) in mice fed the high protein diet compared to the control protein diet.

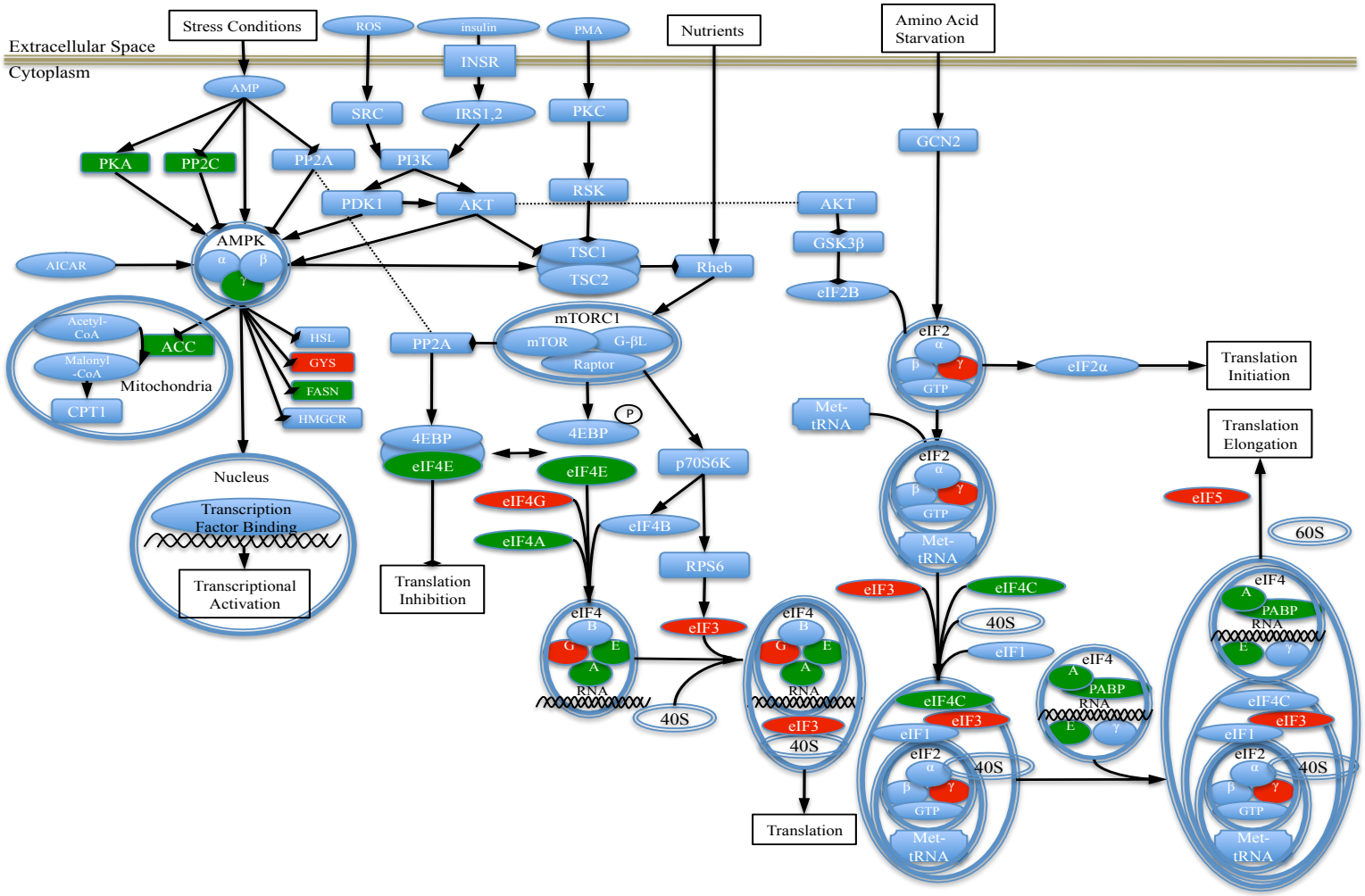

Figure 29C. Protein expression changes in the AMPK, mTOR, and eIF2 pathways by diet.

Expression of protein changed by diet in mice fed *ad libitum*. Protein expression increased (red) or decreased (green) in mice fed the control protein diet compared to the low protein diet.

At the mRNA level (Table 25A, Figure 28), several genes change expression over time including phosphatidyinositol-3-kinase (*PI3K*), protein phosphatase 2A (*PP2A*), and protein phosphatase 2C (*PP2C*), upstream of the AMPK, mTOR, and eIF2 pathways. In mice fed a high protein diet, expression of these genes in increased as demonstrated by greater expression in the high protein than the control protein diet by 60 minutes after reintroduction of food (Figure 28C). By 120 minutes after reintroduction of food, *AMPKα* and its activators protein kinase A *(PKA*), *PI3K* and 3-phosphoinositide dependent protein kinase-1 (*PDK1*) were all up-regulated in the high protein diet group (Figure 28D). In further support of the interpretation of the expression results suggesting the activation of *AMPK* in the high protein diet group, acetyl CoA carboxylase (*ACC*), hydroxymethylglutaryl-CoA reductase (*HMGCR*), glycogen synthase (*GYS*), and fatty acid synthase (*FASN*), which are all inhibited by AMPK, had decreased expression in the high protein diet at all time points (Figure 28A, B, C, and D). Upstream of mTOR, protein kinase C (*PKC*) represses the mTOR inhibiting complex tuberous sclerosis 1 & 2 (*TSC*). Expression of *TSC1* increased in the high protein diet 30 minutes after reintroduction of food (Figure 28B) in accordance with increased *mTOR* inhibition in the high protein diet group. The kinases upstream of eIF2, including PI3K, PDK1, and glycogen synthase kinase 3-beta (GSK3β), inhibit eIF2β by phosphorylation and are counteracted by insulin. mRNA expression of these kinases is higher in the high protein diet than the control diet 120 minutes after reintroduction of food (Figure 28D), reflecting the control diet's abundance of carbohydrates. Downstream, the eukaryotic initiation factor (eIF) 5 gene, which mediates release of the translation initiation complex to allow for elongation as the final step of eIF2 signaling, has increased expression in mice fed the

high protein versus the control diet after 120 minutes (Figure 28D), further suggesting that the high protein diet stimulates eIF2 signaling of translation. Overall the mRNA expression data indicates increased signaling through AMPK and eIF2 with inhibition of mTOR signaling.

Examination of the protein expression data for differences between high, control and low protein diets identified a possible response mechanism to long-term exposure to each diet (Table 25B, Figure 29). Expression of PP2C, the AMP sensitive AMPK inhibitor, is decreased in the high protein diet while IRS1/2, PI3K, and PKA, which activate AMPK in response to insulin and AMP, are increased in the high protein diet compared to the control diet (Figure 29A and B). Downstream of AMPK, ACC and FASN, proteins inhibited by AMPK, have the highest expression in the low protein diet (Figure 29). In the cascade of mTOR phosphorylation, its target 70-kDa ribosomal protein S6 kinase (p70S6K) activates eIF4B and phosphorylates ribosomal protein S6 (RPS6). Both eIF4B and eIF3, which is activated by RPS6, increased in expression in association with a high protein diet (Figure 29A). Additional eIF4 translation initiation complex members in the mTOR pathway, including eIF4A and eIF4G, have varying response to diet such that eIF4A only shows a slight increase following feeding of the low protein diet, and eIF4G decreases expression in the high or low protein diets compared to the control diet (Figure 29). eIF2β and eIF2γ, members of the eIF2 pathway translation initiation complex, increase expression when mice were fed the high or protein diet (Figure 29B and C). Further down the eIF2 pathway, members of the translation elongation complex eIF2β, eIF2γ, eIF4C, poly(A) binding protein (PABP), and eIF5 all have varying response to diet (Figure 29). Overall, the expression of eIF2β

and eIF2γ seemed to be decreased in the low protein diet group, eIF3 is increased in the high protein diet groups, eIF4C had increased expression in the high and low protein groups and PABP and eIF5 had lowest expression in the high protein group (Figure 29). These varying results suggest that the high protein diet leads to an inhibition of mTOR and eIF2 signaled translation elongation, while it also leads to an increase of eIF2 translation initiation.

Changes in urea cycle expression in response to diet

Expression of the urea cycle genes and enzymes was analyzed and, with the exception of NAGS, their response to diet matched the previously reported studies [4,5,84] such that their expression of mRNA (Figure 30) and protein (Figure 31) increased in a high protein diet and decreased in a low protein diet. NAGS mRNA expression, however, decreased in the high protein diet, but, due to very low abundance of the NAGS protein, it is undetectable by mass spectrometry. The cause for this disparate expression remains unknown. While additional mechanisms may be regulating the activity and/or availability of NAGS and all of these enzymes, overall, the urea cycle enzymes show marked response to dietary protein which is expected for handling the increased nitrogen load.

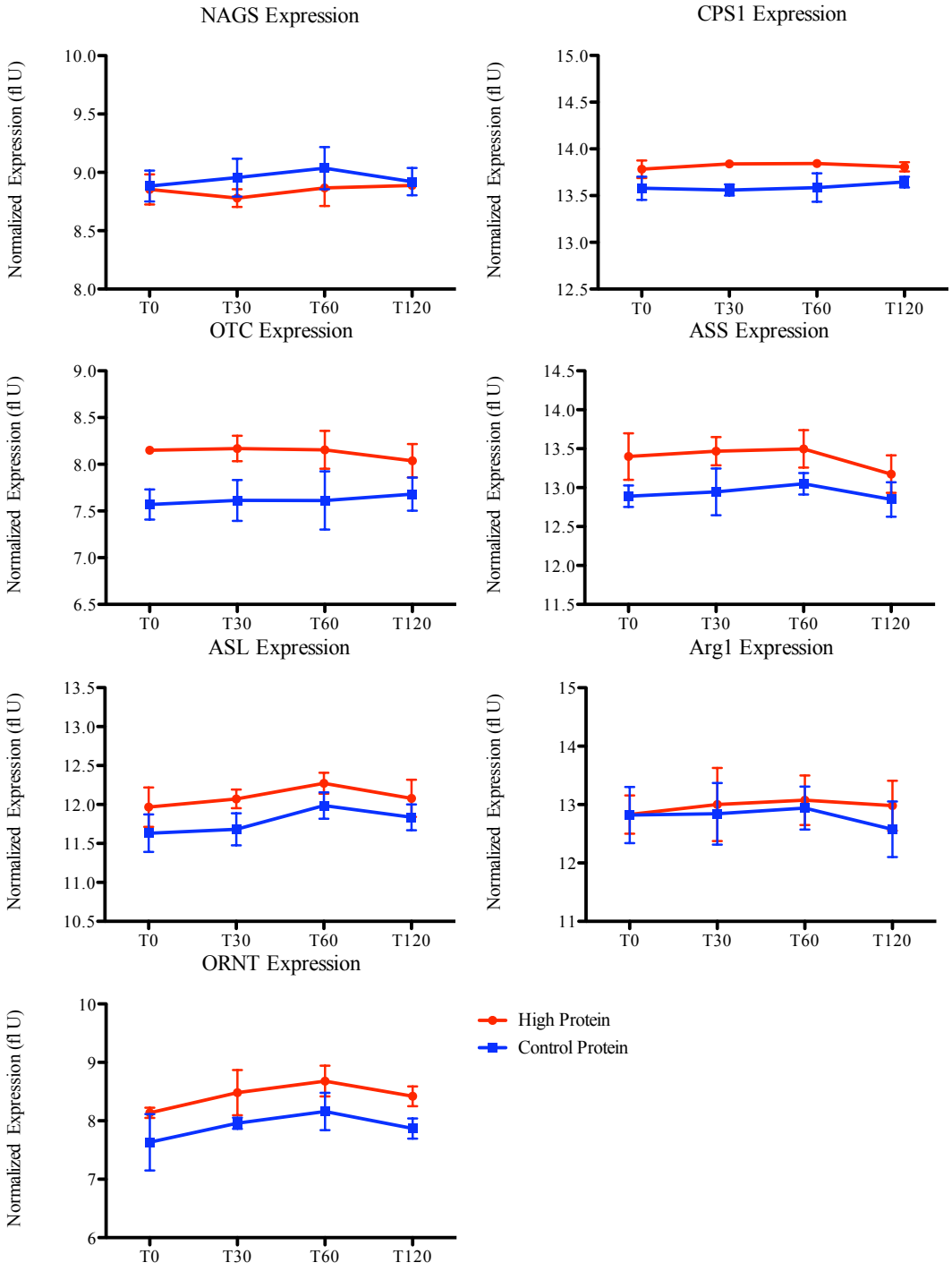

Figure 30. Urea cycle gene expression changes in mice fed a time-restricted diet.

mRNA expression of urea cycle genes was higher in the high protein diet for CPS1 (B), OTC (C), ASS (D), ASL (E), Arg1 (F), ORNT (G) while expression of NAGS (A) was decreased in the high protein diet. All points represent an averaged normalized expression as measured by fluorescent units (fl. U) with standard error (n=3).

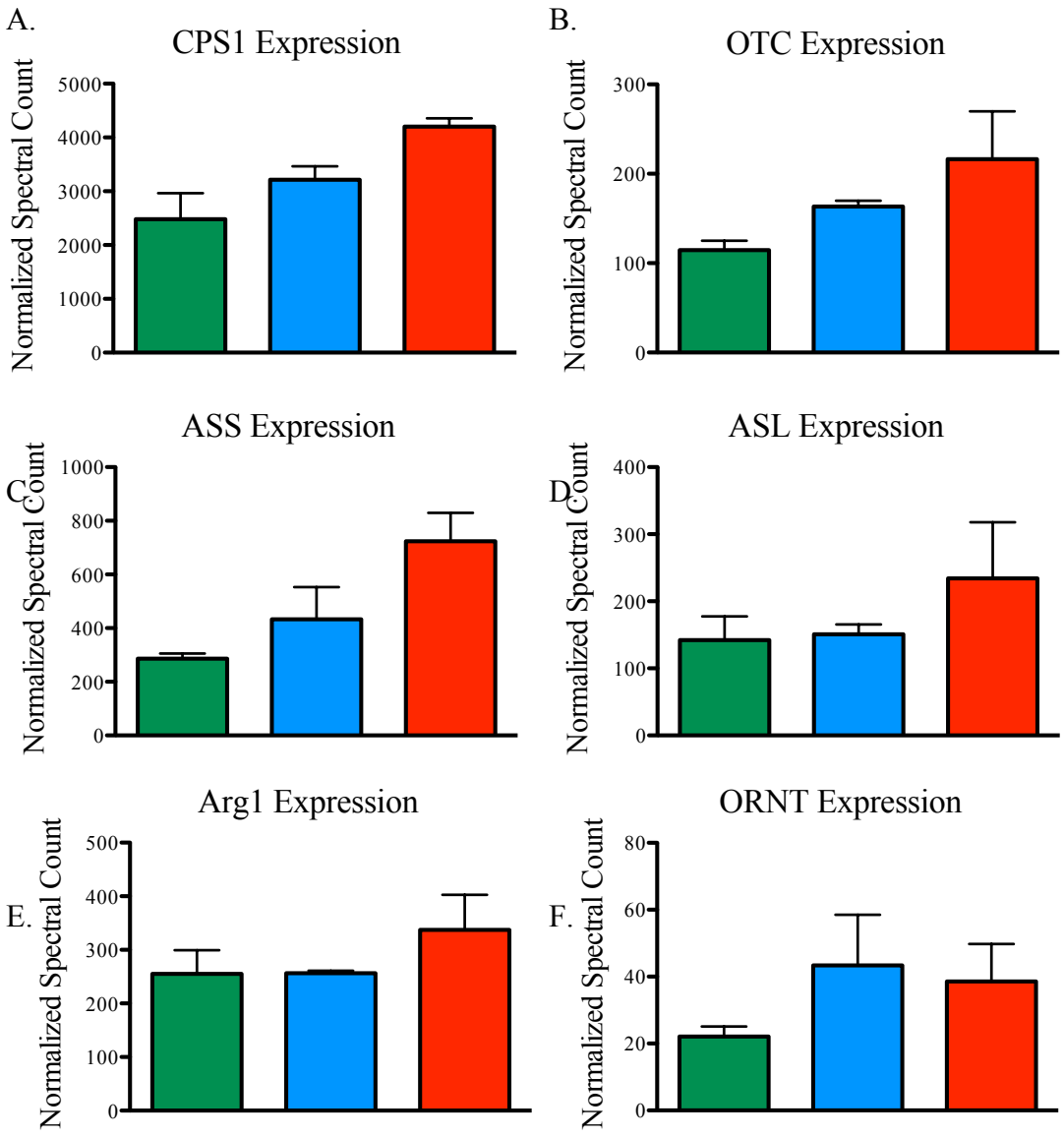

Figure 31. Urea cycle protein expression changes in mice fed an *ad libitum* diet.

Protein expression of urea cycle proteins CPS1 (A), OTC (B), ASS (C), ASL (D), Arg1 (E), ORNT (F) was highest in the high protein diet. NAGS was not detected in label-free mass spectrometry quantification. All points represent averaged normalized spectral count with standard error (n=3).

Discussion and Future Directions

In this study we examined the effect of dietary protein on and expression of mouse liver mRNA and protein. Syngeneic adult male mice were fed a diet of low (10%), control (20%) or high (60%) protein on either an *ad libitum* or a time-restricted feeding schedule for 12 days. No significant difference in weight was seen between diet groups for either feeding schedule. Mice on the high protein diet, however, consumed less food, as high protein signals satiety [266,267], and drank more water to excrete the excess urea [268].

In consideration of the physiological changes associated with dietary protein consumption, analysis of mRNA expression using the microarray platform identified thousands of genes whose expression significantly differed between diets and time points. Hierarchical clustering of probe sets by differing expression uniformly grouped samples by diet. Since transcriptional regulation of genes typically occurs on a time scale of minutes rather than days [269], it was important to analyze changes in gene expression at time points after reintroduction of feed. While replicate samples from each time point did not all cluster together, earlier time points of 0 and 30 minutes tended to cluster together, possibly reflecting the time lag from ingestion of food to signal transcription while and the 120 minute time point replicates consistently clustered together to reflect a fed state. Upon further examination, the time-restricted feeding schedule successfully synchronized expression of genes that respond to diet and twice the number of genes were identified as significantly different between high and control protein diets in the time-restricted diet vs. the *ad libitum* diet.

Protein expression may play a more important role than mRNA in the response to diet and therefore spectral counting analysis was used to identify and measure abundance of liver proteins. A limitation of label-free mass spectrometry for spectral counting and identification of peptides is a failure to identify low abundance proteins. Regardless, proteins involved in nutrient signaling leading to initiation of translation are abundant in the liver, as demonstrated by our analysis. The complementary nature of analysis on both mRNA and protein expression is reflected in pathway mapping that identified metabolism of fats, carbohydrates, and proteins as main functions of the significantly differing genes.

The main focus of this research was to identify signaling pathways responsive to dietary protein. In combination, the mRNA and protein expression analysis identified AMPK, mTOR, and eIF2 as the main signaling pathways that respond to diet. It is important to note that there are a number of shared members in each of these pathways, as shown in Figure 27, that reflect mechanisms for control of transcription, translation initiation, and translation elongation. As detailed in the previous section, gene and protein expression data suggest that AMPK and eIF2 pathways respond to high protein diet by increasing signaling through the pathway and increasing transcription or translation initation. mTOR signaling and translation elongation signaling through eIF2, however appear to decrease in response to a diet high in protein. It is important to note that within a pathway, complexes of proteins that do not respond to diet as a group may either adapt without increasing expression of each protein, because the complex can form and dissociate fast enough to accommodate the translation initiation and elongation load, or may adapt via a post-translational control mechanism such as phosporylation [23].

In conclusion, the results of this study strongly suggest that AMPK, mTOR, and eIF2 pathways play an important role in responding to dietary protein as they regulate translation in response to a high protein diet. The profiling data also show that urea cycle genes and enzymes respond to diet and suggests that these nutrient sensing pathways have a role in their regulation. To further confirm these analyses, the mRNA expression results will be validated using qPCR. qPCR will be performed on genes that change in expression due to diet, according to the microarray results, and are involved in the nutrient sensing pathways (Table 26). Validation of the mass spectrometry data will be conducted by immunoblot and should confirm that urea cycle proteins expression change in response to diet via upstream and/or downstream members of the nutrient sensing pathways as described in Table 27.

Table 26. Genes analyzed for mRNA expression changes using qPCR.

Gene Symbol	RefSeq ID
NAGS	NM_178053.4,NM_145829.1
CPS1	NM_001080809.1
OTC	NM_008769.3
ASS1	NM_007494.3
ASL	NM_133768.4
Arg1	NM_007482.3
ORNT	NM_181325.4
PI3K	NM_001024955.1,NM_001077495.1
PP2A	NM_016891.3
PP2C	NM_008910.3
Rheb	NM_053075.3
Tsc1	NM_022887.3
Tsc2	NM_001039363.1,NM_011647.2
Akt1	NM_001165894.1,NM_009652.3
PKA	NM_021880.2
PDK1	NM_172665.4
ACCa	NM_133360.2
ACCb	NM_133904.2
AMPKα1	NM_001013367.3
AMPKα2	NM_178143.2
PKC	NM_011103.3
GSK3b	NM_019827.6
mTOR	NM_020009.2
IRS1	NM_010570.4
IRS2	NM_001081212.1
eIF2α	NM_001177806.1,NM_013719.3
P70S6K	NM_028259.4,NM_001114334.1
RPS6	NM_009096.3
HMGCR	NM_008255.2
GYS1	NM_030678.3
FASN	NM_007988.3

Table 27. Proteins analyzed for expression and phosphorylation from mice fed high, control or low protein diet

	mTOR	AMPK	eIF2
Upstream	mTOR	AMPK	
Downstream	4EBP1 S6K	ACC	eIF2

As an indicator of pathway activity, phosphorylation status of the proteins in Table 25 will also be assessed. It is expected that mTOR pathway members will show increased phosphorylation following feeding of the high protein diet while AMPK and eIF2 pathway members will have reduced phosphorylation [23]. To directly show the effect of these pathways on the urea cycle genes mRNA and protein expression, additional studies will include targeted inhibition or stimulation of these pathways. Upon treatment of primary mouse hepatocytes with rapamycin, it binds the mTOR complex and inhibits formation of the mTORC1 complex, thereby inhibiting downstream signaling and phosphorylation of p70S6K and 4EBP1 [270,271]. This cascade should decrease expression of the urea cycle proteins. 5-aminoimidazole-4-carboxamide-1-beta-4-ribofuranoside (AICAR) treatment of primary hepatocytes activates AMPK through 5-amino-4-imidazolecarboxamide ribotide (ZMP) leading to inhibition of AMPK dephosphorylation [272,273] and should increase mRNA expression of the urea cycle genes. While ACC is used for fatty acid metabolism, and therefore does not directly effect the urea cycle, an increase in AMPK activation by AICAR should increase phosphorylation and inhibition of ACC. Finally, amino acid sufficiency causes phosphorylation of eIF2 by mTORC1 and thereby inhibits its ability to phosphorylate and inhibit eIF2α. Therefore, a high protein diet would lead to dephosphorylation of eIF2α and increased translation of the urea cycle proteins. Treatment of hepatocytes with leucine, the branched chain amino acid commonly thought to signal amino acid load to the metabolic machinery [23,274] should trigger the same response in the mTOR, AMPK, and eIF2 pathways as a high protein diet and increase mRNA and protein expression of the urea cycle enzymes.

Chapter 6: Conclusions and Future Directions

For this project the overall goal was to identify the mechanism that controls the urea cycle's response to diet. Our approach to answer this question was to first dissect the mechanism of transcriptional regulation for NAGS. We learned that NAGS possesses two regulatory regions upstream of the translational start site using a combination of bioinformatic analyses and *in vitro* functional assays. We termed these regions the promoter and enhancer and used bioinformatic software to identify transcription factor binding sites within each region. The possible *cis*-acting sites were tested and binding was verified using chromatin immunoprecipitation, DNA-protein pull-down assays, and *in vitro* functional assays in hepatoma and intestine cell lines. Using 5' RACE, we found that transcription initiates at multiple and different sites in the liver and small intestine, and that Sp1 and CREB sites within the promoter facilitate this transcription. The enhancer contains binding sites for HNF-1 and NF-Y, and the absence of HNF-1 in intestine cells suggests liver specific regulation through the NAGS enhancer. Overall, transcription of NAGS is regulated by factors that respond to glucocorticoid and cAMP stimulation, as well as through the hepatic nuclear factor (HNF) family of proteins. These proteins confer tissue specificity with feedback regulation, and are important for the regulation of many urea cycle enzymes.

Additional future studies could include analysis on differential regulation of transcription between the liver and small intestine cells, supported by the 5'RACE results and the *in vitro* liver-specific enhancer. These studies could be used to verify the role of NAGS in the small intestine, currently hypothesized to be for citrulline synthesis. To further discern the mechanism of transcriptional regulation, studies could investigate the

possibility of synergistic activation by the multiple Sp1 sites present, and the importance of interactions and feed-back regulation between the HNF factors. Finally, the relatively small increase of transcription by the enhancer suggests that another mechanism, perhaps a proximal enhancer (Figure 3, promoter B) may also be acting on transcription, and additional reporter assays could be conducted to further examine the regulatory regions.

As a result of the studies on NAGS transcriptional regulation, we were able to confirm the diagnosis of NAGS deficiency in a patient with a deleterious mutation in the enhancer of NAGS. This patient had the clinical signs of NAGSD, but as initial sequencing failed to identify a deleterious mutation she was treated with ammonia scavenging agents and a low protein diet for seven years. Clinical trial results showed NCG to be markedly effective and, at the same time, a mutation in the HNF-1 binding site within the enhancer of NAGS was found to decrease transcriptional activation, as tested using *in vitro* reporter assays, through decreased binding of HNF-1, as determined by DNA-protein pull-down assays. Following confirmation by functional studies, she was diagnosed with NAGSD, and her treatment was changed to NCG while the drug efficacy enabled her to consume protein *ad libitum.*

Following the success of the previous study involving NAGSD, a study was conducted to analyze the regulatory regions of previously undiagnosed patients with the clinical signs of OTC deficiency (OTCD), but lacking deleterious mutations in the exons or intron/exon boundries. Mutations in the promoter of 3 undiagnosed patients were determined to be uncommon, using SNP genotyping, and reduced binding of HNF-4 transcription factor, as determined used DNA-protein pull down assay. Future studies

include *in vitro* functional studies to determine the effect of these mutations on transcriptional regulation in hepatocytes.

The yield of the NAGSD and OTCD regulatory region mutation analysis suggests the global importance of this analysis for diagnosis of patients without disease causing mutations in the exons and intron/exon boundries, particularly in the advent of exome analysis for clinical indications. Ours are only the 2nd and 3rd papers on deleterious regulatory region mutations in the urea cycle genes, reflecting the less than 2% of currently reported regulatory region mutations in the Human Genome Mutation Database. Diagnosing disease caused by non-coding regions variants, however, requires functional assays, which are time consuming and impractical for rapid diagnosis. For future research, a high throughput method for non-coding region analysis should be developed. If and when this comes to be, I would expect the percentage of disease causing mutations in the regulatory regions to increase substantially.

Our research was the first to determine the factors involved in transcriptional regulation of NAGS and thereby has enabled further study of the nutrient sensing mechanisms that regulate all 6 enzymes in the cycle. Livers from mice in the *ad libitum* feeding study that received the low (10%), control (20%), or high (60%) protein diet and the time-restricted diet study that received the control or high protein diets and were sacrificed 0, 30, 60 or 120 minutes after reintroduction of food were analyzed for mRNA and protein expression using Affymetrix microarray and label-free mass spectrometry analysis, respectively. Through these analysis we identified that the AMPK, mTOR, and eIF2 nutrient sensing pathways are responsive to content of dietary protein. Our data showed an increase in expression of AMPK pathway members, suggesting that it

responds to diet by signaling transcription of factors that are known to bind to urea cycle enzymes. mTOR and eIF2 pathway members, which are involved in the initiation and elongation of translation, show decreased expression in response to the high protein diet, and yet urea cycle enzyme protein expression increases. These data will further investigated and be validated using qPCR and immunoblotting**.**

Future studies for this project could include identification of the amino acid(s) that these pathways detect, and we hypothesize that this may involve leucine or glutamate. While transcription and protein expression play an important role in responding to dietary protein, post-translational modification may provide the mechanism for a rapid-response. Additional analyses will include the phosphorylation state of AMPK, mTOR, and eIF2 pathway members as well as targeted inhibition or activation of the pathways followed by analysis of changes in urea cycle enzyme expression. By determining the mechanisms of regulation of the urea cycle, we aim to learn how diet affects patients with urea cycle disorders and hope to identify novel treatments based on these regulatory mechanisms. In addition, recent studies have demonstrated that acetylation could play a role in the regulation of CPS1, OTC and ASL [26,27,28] [264], and suggest a role for activation in the remaining urea cycle enzymes as well. Future studies could investigate whether this modification is present, which acetylases are responsible, and how it affects enzyme activity of urea cycle enzymes. We may find that, along with phosphorylation of the nutrient sensing pathways, acetylation is key to regulating the rapid response of the urea cycle enzymes to the diet and may provide a new treatment avenue for patients with urea cycle disorders.

Bibliography

1. Krebs HA, Henseleit K (1932) Studies on urea formation in the animal organism. Hoppe-Seylers Z Physiol Chem 210: 33-66.

2. Lightbody HD, Kleinman, A. (1939) Variations produced by food differences in the concentration of arginase in the livers of white rats. Journal of Biological Chemistry 129: 71-78.

3. Schimke RT (1963) Studies on factors affecting the levels of urea cycle enzymes in rat liver. J Biol Chem 238: 1012-1018.

4. Schimke RT (1962) Differential effects of fasting and protein-free diets on levels of urea cycle enzymes in rat liver. J Biol Chem 237: 1921-1924.

5. Schimke RT (1962) Adaptive characteristics of urea cycle enzymes in the rat. J Biol Chem 237: 459-468.

6. Schimke RT (1964) The Importance of Both Synthesis and Degradation in the Control of Arginase Levels in Rat Liver. J Biol Chem 239: 3808-3817.

7. Tsuda M, Shikata Y, Katsunuma T (1979) Effect of dietary proteins on the turnover of rat liver argininosuccinate synthetase. J Biochem 85: 699-704.

8. Tatibana M, Kawamoto S, Sonoda T, Mori M (1982) Enzyme regulation of n-acetylglutamate synthesis in mouse and rat liver. Adv Exp Med Biol 153: 207-216.

9. McGivan JD, Bradford NM, Beavis AD (1977) Factors influencing the activity of ornithine aminotransferase in isolated rat liver mitochondria. Biochem J 162: 147-156.

10. Kawamoto S, Ishida H, Mori M, Tatibana M (1982) Regulation of N-acetylglutamate synthetase in mouse liver. Postprandial changes in sensitivity to activation by arginine. Eur J Biochem 123: 637-641.

11. Shigesada K, Tatibana M (1978) N-Acetylglutamate synthetase from rat-liver mitochondria. Partial purification and catalytic properties. Eur J Biochem 84: 285-291.

12. Stewart PM, Walser M (1980) Short term regulation of ureagenesis. J Biol Chem 255: 5270-5280.

13. Cohen PP (1981) The ornithine-urea cycle: biosynthesis and regulation of carbamyl phosphate synthetase I and ornithine transcarbamylase. Curr Top Cell Regul 18: 1-19.

14. Meijer AJ, Lof C, Ramos IC, Verhoeven AJ (1985) Control of ureogenesis. Eur J Biochem 148: 189-196.

15. Banko G, Zollner H (1985) Does ornithine stimulate carbamoylphosphate synthetase? Int J Biochem 17: 503-507.

16. Kaloyianni M, Freedland RA (1989) A new approach of the estimation of Km of carbamyl phosphate synthetase for ammonia in isolated rat hepatocytes. Int J Biochem 21: 971-974.

17. Edmonds MS, Lowry KR, Baker DH (1987) Urea cycle metabolism: effects of supplemental ornithine or citrulline on performance, tissue amino acid concentrations and enzymatic activity in young pigs fed arginine-deficient diets. J Anim Sci 65: 706-716.

18. Jackson MJ, Allen SJ, Beaudet AL, O'Brien WE (1988) Metabolite regulation of argininosuccinate synthetase in cultured human cells. J Biol Chem 263: 16388-16394.

19. Morris SM, Jr. (2002) Regulation of enzymes of the urea cycle and arginine metabolism. Annu Rev Nutr 22: 87-105.

20. Morimoto BH, Brady JF, Atkinson DE (1990) Effect of level of dietary protein on arginine-stimulated citrulline synthesis. Correlation with mitochondrial N-acetylglutamate concentrations. Biochem J 272: 671-675.

21. Li Z, Yarmush ML, Chan C (2004) Insulin concentration during preconditioning mediates the regulation of urea synthesis during exposure to amino acid-supplemented plasma. Tissue Eng 10: 1737-1746.

22. Brosnan JT (2000) Glutamate, at the interface between amino acid and carbohydrate metabolism. J Nutr 130: 988S-990S.

23. Chotechuang N, Azzout-Marniche D, Bos C, Chaumontet C, Gausseres N, et al. (2009) mTOR, AMPK, and GCN2 coordinate the adaptation of hepatic energy metabolic pathways in response to protein intake in the rat. Am J Physiol Endocrinol Metab 297: E1313-1323.

24. Caldovic L, Ah Mew N, Shi D, Morizono H, Yudkoff M, et al. (2010) N-acetylglutamate synthase: structure, function and defects. Mol Genet Metab 100 Suppl 1: S13-19.

25. Caldovic L, Lopez GY, Haskins N, Panglao M, Shi D, et al. (2006) Biochemical properties of recombinant human and mouse N-acetylglutamate synthase. Mol Genet Metab 87: 226-232.

26. Ogura M, Nakamura Y, Tanaka D, Zhuang X, Fujita Y, et al. Overexpression of SIRT5 confirms its involvement in deacetylation and activation of carbamoyl phosphate synthetase 1. Biochem Biophys Res Commun 393: 73-78.

27. Nakagawa T, Lomb DJ, Haigis MC, Guarente L (2009) SIRT5 Deacetylates carbamoyl phosphate synthetase 1 and regulates the urea cycle. Cell 137: 560-570.

28. Yu W, Lin Y, Yao J, Huang W, Lei Q, et al. (2009) Lysine 88 acetylation negatively regulates ornithine carbamoyltransferase activity in response to nutrient signals. J Biol Chem 284: 13669-13675.

29. Hallows WC, Yu W, Smith BC, Devries MK, Ellinger JJ, et al. Sirt3 promotes the urea cycle and fatty acid oxidation during dietary restriction. Mol Cell 41: 139-149.

30. Ferre P, Azzout-Marniche D, Foufelle F (2003) AMP-activated protein kinase and hepatic genes involved in glucose metabolism. Biochem Soc Trans 31: 220-223.

31. Viollet B, Foretz M, Guigas B, Horman S, Dentin R, et al. (2006) Activation of AMP-activated protein kinase in the liver: a new strategy for the management of metabolic hepatic disorders. J Physiol 574: 41-53.

32. Proud CG (2004) Role of mTOR signalling in the control of translation initiation and elongation by nutrients. Curr Top Microbiol Immunol 279: 215-244.

33. Anthony TG, McDaniel BJ, Byerley RL, McGrath BC, Cavener DR, et al. (2004) Preservation of liver protein synthesis during dietary leucine deprivation occurs at the expense of skeletal muscle mass in mice deleted for eIF2 kinase GCN2. J Biol Chem 279: 36553-36561.

34. Dever TE, Hinnebusch AG (2005) GCN2 whets the appetite for amino acids. Mol Cell 18: 141-142.

35. Avruch J, Long X, Ortiz-Vega S, Rapley J, Papageorgiou A, et al. (2009) Amino acid regulation of TOR complex 1. Am J Physiol Endocrinol Metab 296: E592-602.

36. Ijichi C, Matsumura T, Tsuji T, Eto Y (2003) Branched-chain amino acids promote albumin synthesis in rat primary hepatocytes through the mTOR signal transduction system. Biochem Biophys Res Commun 303: 59-64.

37. Ma XM, Blenis J (2009) Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol 10: 307-318.

38. Ruderman NB, Park H, Kaushik VK, Dean D, Constant S, et al. (2003) AMPK as a metabolic switch in rat muscle, liver and adipose tissue after exercise. Acta Physiol Scand 178: 435-442.

39. Wong AK, Howie J, Petrie JR, Lang CC (2009) AMP-activated protein kinase pathway: a potential therapeutic target in cardiometabolic disease. Clin Sci (Lond) 116: 607-620.

40. Viollet B, Guigas B, Leclerc J, Hebrard S, Lantier L, et al. (2009) AMP-activated protein kinase in the regulation of hepatic energy metabolism: from physiology to therapeutic perspectives. Acta Physiol (Oxf) 196: 81-98.

41. Nebes VL, Morris SM, Jr. (1988) Regulation of messenger ribonucleic acid levels for five urea cycle enzymes in cultured rat hepatocytes. Requirements for cyclic adenosine monophosphate, glucocorticoids, and ongoing protein synthesis. Mol Endocrinol 2: 444-451.

42. Caldovic L, Morizono H, Gracia Panglao M, Gallegos R, Yu X, et al. (2002) Cloning and expression of the human N-acetylglutamate synthase gene. Biochem Biophys Res Commun 299: 581-586.

43. Caldovic L, Morizono H, Yu X, Thompson M, Shi D, et al. (2002) Identification, cloning and expression of the mouse N-acetylglutamate synthase gene. Biochem J 364: 825-831.

44. Heibel SK, Lopez GY, Panglao M, Sodha S, Marino-Ramirez L, et al. Transcriptional regulation of N-acetylglutamate synthase. PLoS One 7: e29527.

45. Abdullah Abu Musa DM, Kobayashi K, Yasuda I, Iijima M, Christoffels VM, et al. (1999) Involvement of a cis-acting element in the suppression of carbamoyl phosphate synthetase I gene expression in the liver of carnitine-deficient mice. Mol Genet Metab 68: 346-356.

46. Schoneveld OJ, Hoogenkamp M, Stallen JM, Gaemers IC, Lamers WH (2007) cyclicAMP and glucocorticoid responsiveness of the rat carbamoylphosphate synthetase gene requires the interplay of upstream regulatory units. Biochimie 89: 574-580.

47. Takiguchi M, Mori M (1995) Transcriptional regulation of genes for ornithine cycle enzymes. Biochem J 312 (Pt 3): 649-659.

48. Grompe M, Jones SN, Caskey CT (1990) Molecular detection and correction of ornithine transcarbamylase deficiency. Trends Genet 6: 335-339.

49. Jones SN, Grompe M, Munir MI, Veres G, Craigen WJ, et al. (1990) Ectopic correction of ornithine transcarbamylase deficiency in sparse fur mice. J Biol Chem 265: 14684-14690.

50. Kimura T, Chowdhury S, Tanaka T, Shimizu A, Iwase K, et al. (2001) CCAAT/enhancer-binding protein beta is required for activation of genes for ornithine cycle enzymes by glucocorticoids and glucagon in primary-cultured hepatocytes. FEBS Lett 494: 105-111.

51. Murakami T, Nishiyori A, Takiguchi M, Mori M (1990) Promoter and 11-kilobase upstream enhancer elements responsible for hepatoma cell-specific expression of the rat ornithine transcarbamylase gene. Mol Cell Biol 10: 1180-1191.

52. Sladek FM, Zhong WM, Lai E, Darnell JE, Jr. (1990) Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev 4: 2353-2365.

53. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, et al. (2004) Control of pancreas and liver gene expression by HNF transcription factors. Science 303: 1378-1381.

54. Spath GF, Weiss MC (1997) Hepatocyte nuclear factor 4 expression overcomes repression of the hepatic phenotype in dedifferentiated hepatoma cells. Mol Cell Biol 17: 1913-1922.

55. Kimura A, Nishiyori A, Murakami T, Tsukamoto T, Hata S, et al. (1993) Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). J Biol Chem 268: 11125-11133.

56. Kadonaga JT, Carner KR, Masiarz FR, Tjian R (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51: 1079-1090.

57. Anderson GM, Freytag SO (1991) Synergistic activation of a human promoter in vivo by transcription factor Sp1. Mol Cell Biol 11: 1935-1943.

58. Boyce FM, 3rd, Freytag SO (1989) Regulation of human argininosuccinate synthetase gene: induction by positive-acting nuclear mechanism in canavanine-resistant cell variants. Somat Cell Mol Genet 15: 113-121.

59. Boyce FM, Anderson GM, Rusk CD, Freytag SO (1986) Human argininosuccinate synthetase minigenes are subject to arginine-mediated repression but not to trans induction. Mol Cell Biol 6: 1244-1252.

60. Sunyakumthorn P, Boonsaen T, Boonsaeng V, Wallace JC, Jitrapakdee S (2005) Involvement of specific proteins (Sp1/Sp3) and nuclear factor Y in basal transcription of the distal promoter of the rat pyruvate carboxylase gene in beta-cells. Biochem Biophys Res Commun 329: 188-196.

61. Dorn A, Bollekens J, Staub A, Benoist C, Mathis D (1987) A multiplicity of CCAAT box-binding proteins. Cell 50: 863-872.

62. Hooft van Huijsduijnen R, Li XY, Black D, Matthes H, Benoist C, et al. (1990) Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits. EMBO J 9: 3119-3127.

63. Takiguchi M, Mori M (1991) In vitro analysis of the rat liver-type arginase promoter. J Biol Chem 266: 9186-9193.

64. Ringold GM, Yamamoto KR, Bishop JM, Varmus HE (1977) Glucocorticoid-stimulated accumulation of mouse mammary tumor virus RNA: increased rate of synthesis of viral RNA. Proc Natl Acad Sci U S A 74: 2879-2883.

65. Bronner M, Hertz R, Bar-Tana J (2004) Kinase-independent transcriptional co-activation of peroxisome proliferator-activated receptor alpha by AMP-activated protein kinase. Biochem J 384: 295-305.

66. Houde VP, Brule S, Festuccia WT, Blanchard PG, Bellmann K, et al. (2010) Chronic rapamycin treatment causes glucose intolerance and hyperlipidemia by upregulating hepatic gluconeogenesis and impairing lipid deposition in adipose tissue. Diabetes 59: 1338-1348.

67. Sengupta S, Peterson TR, Laplante M, Oh S, Sabatini DM (2010) mTORC1 controls fasting-induced ketogenesis and its modulation by ageing. Nature 468: 1100-1104.

68. Lin J, Puigserver P, Donovan J, Tarr P, Spiegelman BM (2002) Peroxisome proliferator-activated receptor gamma coactivator 1beta (PGC-1beta), a novel PGC-1-related transcription coactivator associated with host cell factor. J Biol Chem 277: 1645-1648.

69. Yoon JC, Puigserver P, Chen G, Donovan J, Wu Z, et al. (2001) Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. Nature 413: 131-138.

70. Christoffels VM, Grange T, Kaestner KH, Cole TJ, Darlington GJ, et al. (1998) Glucocorticoid receptor, C/EBP, HNF3, and protein kinase A coordinately activate the glucocorticoid response unit of the carbamoylphosphate synthetase I gene. Mol Cell Biol 18: 6305-6315.

71. Hoogenkamp M, Gaemers IC, Schoneveld OJ, Das AT, Grange T, et al. (2007) Hepatocyte-specific interplay of transcription factors at the far-upstream enhancer of the carbamoylphosphate synthetase gene upon glucocorticoid induction. Febs J 274: 37-45.

72. Inoue Y, Hayhurst GP, Inoue J, Mori M, Gonzalez FJ (2002) Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4alpha (HNF4alpha). HNF4alpha regulates ornithine transcarbamylase in vivo. J Biol Chem 277: 25257-25265.

73. Nishiyori A, Tashiro H, Kimura A, Akagi K, Yamamura K, et al. (1994) Determination of tissue specificity of the enhancer by combinatorial operation of tissue-enriched transcription factors. Both HNF-4 and C/EBP beta are required for

liver-specific activity of the ornithine transcarbamylase enhancer. J Biol Chem 269: 1323-1331.

74. Choudhury M, Qadri I, Rahman SM, Schroeder-Gloeckler J, Janssen RC, et al. (2011) C/EBPbeta is AMP kinase sensitive and up-regulates PEPCK in response to ER stress in hepatoma cells. Mol Cell Endocrinol 331: 102-108.

75. Thomson DM, Herway ST, Fillmore N, Kim H, Brown JD, et al. (2008) AMP-activated protein kinase phosphorylates transcription factors of the CREB family. J Appl Physiol 104: 429-438.

76. Guei TR, Liu MC, Yang CP, Su TS (2008) Identification of a liver-specific cAMP response element in the human argininosuccinate synthetase gene. Biochem Biophys Res Commun 377: 257-261.

77. Howell BW, Lagace M, Shore GC (1989) Activity of the carbamyl phosphate synthetase I promoter in liver nuclear extracts is dependent on a cis-acting C/EBP recognition element. Mol Cell Biol 9: 2928-2933.

78. Samadani U, Costa RH (1996) The transcriptional activator hepatocyte nuclear factor 6 regulates liver gene expression. Mol Cell Biol 16: 6273-6284.

79. Ktistaki E, Talianidis I (1997) Modulation of hepatic gene expression by hepatocyte nuclear factor 1. Science 277: 109-112.

80. Thomas H, Jaschkowitz K, Bulman M, Frayling TM, Mitchell SM, et al. (2001) A distant upstream promoter of the HNF-4alpha gene connects the transcription factors involved in maturity-onset diabetes of the young. Hum Mol Genet 10: 2089-2097.

81. Nagaki M, Moriwaki H (2008) Transcription factor HNF and hepatocyte differentiation. Hepatol Res 38: 961-969.

82. Peterson RS, Clevidence DE, Ye H, Costa RH (1997) Hepatocyte nuclear factor-3 alpha promoter regulation involves recognition by cell-specific factors, thyroid transcription factor-1, and autoactivation. Cell Growth Differ 8: 69-82.

83. Rausa FM, Tan Y, Costa RH (2003) Association between hepatocyte nuclear factor 6 (HNF-6) and FoxA2 DNA binding domains stimulates FoxA2 transcriptional activity but inhibits HNF-6 DNA binding. Mol Cell Biol 23: 437-449.

84. Morris SM, Jr., Moncman CL, Rand KD, Dizikes GJ, Cederbaum SD, et al. (1987) Regulation of mRNA levels for five urea cycle enzymes in rat liver by diet, cyclic AMP, and glucocorticoids. Arch Biochem Biophys 256: 343-353.

85. Morris SM, Jr. (1992) Regulation of enzymes of urea and arginine synthesis. Annu Rev Nutr 12: 81-101.

86. Hazra A. D, D. C., Almon, R. R., Snyder, G. H., Jusko, W. J. (2008) Pharmacodynamic Modeling of Acute and Chronic Effects of Methylprednisolone on Hepatic Urea Cycle Genes in Rats. Gene Regulation and Systems Biology 2: 1-19.

87. Snodgrass PJ, Lin RC, Muller WA, Aoki TT (1978) Induction of urea cycle enzymes of rat liver by glucagon. J Biol Chem 253: 2748-2753.

88. Altarejos JY, Montminy M CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. Nat Rev Mol Cell Biol 12: 141-151.

89. Kopperud R, Krakstad C, Selheim F, Doskeland SO (2003) cAMP effector mechanisms. Novel twists for an 'old' signaling system. FEBS Lett 546: 121-126.

90. Tasken K, Aandahl EM (2004) Localized effects of cAMP mediated by distinct routes of protein kinase A. Physiol Rev 84: 137-167.

91. Brusilow SW, Maestri NE (1996) Urea cycle disorders: diagnosis, pathophysiology, and therapy. Adv Pediatr 43: 127-170.

92. Summar M (2001) Current strategies for the management of neonatal urea cycle disorders. J Pediatr 138: S30-39.

93. Tuchman M, Lee B, Lichter-Konecki U, Summar ML, Yudkoff M, et al. (2008) Cross-sectional multicenter study of patients with urea cycle disorders in the United States. Mol Genet Metab 94: 397-402.

94. Ah Mew N, Caldovic L (2011) N-acetylglutamate synthase deficiency: an insight into genetics, epidemiology, pathophysiology, and treatment. App Clin Gen 4: 127-135.

95. Heibel SK, Ah Mew N, Caldovic L, Daikhin Y, Yudkoff M, et al. (2011) N-carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the NAGS gene and to effective therapy. Hum Mutat 32: 1153-1160.

96. Bachmann C, Colombo JP, Jaggi K (1982) N-acetylglutamate synthetase (NAGS) deficiency: diagnosis, clinical observations and treatment. Adv Exp Med Biol 153: 39-45.

97. Elpeleg O, Shaag A, Ben-Shalom E, Schmid T, Bachmann C (2002) N-acetylglutamate synthase deficiency and the treatment of hyperammonemic encephalopathy. Ann Neurol 52: 845-849.

98. Gessler P, Buchal P, Schwenk HU, Wermuth B (2010) Favourable long-term outcome after immediate treatment of neonatal hyperammonemia due to N-acetylglutamate synthase deficiency. Eur J Pediatr 169: 197-199.

99. Haberle J, Schmidt E, Pauli S, Kreuder JG, Plecko B, et al. (2003) Mutation analysis in patients with N-acetylglutamate synthase deficiency. Hum Mutat 21: 593-597.

100. Elpeleg ON, Colombo JP, Amir N, Bachmann C, Hurvitz H (1990) Late-onset form of partial N-acetylglutamate synthetase deficiency. Eur J Pediatr 149: 634-636.

101. Vockley J, Vockley CM, Lin SP, Tuchman M, Wu TC, et al. (1992) Normal N-acetylglutamate concentration measured in liver from a new patient with N-acetylglutamate synthetase deficiency: physiologic and biochemical implications. Biochem Med Metab Biol 47: 38-46.

102. Hinnie J, Colombo JP, Wermuth B, Dryburgh FJ (1997) N-Acetylglutamate synthetase deficiency responding to carbamylglutamate. J Inherit Metab Dis 20: 839-840.

103. Plecko B, Erwa W, Wermuth B (1998) Partial N-acetylglutamate synthetase deficiency in a 13-year-old girl: diagnosis and response to treatment with N-carbamylglutamate. Eur J Pediatr 157: 996-998.

104. Forget PP, van Oosterhout M, Bakker JA, Wermuth B, Vles JS, et al. (1999) Partial N-acetyl-glutamate synthetase deficiency masquerading as a valproic acid-induced Reye-like syndrome. Acta Paediatr 88: 1409-1411.

105. Schmidt E, Nuoffer JM, Haberle J, Pauli S, Guffon N, et al. (2005) Identification of novel mutations of the human N-acetylglutamate synthase gene and their functional investigation by expression studies. Biochim Biophys Acta 1740: 54-59.

106. Caldovic L, Morizono H, Panglao MG, Lopez GY, Shi D, et al. (2005) Late onset N-acetylglutamate synthase deficiency caused by hypomorphic alleles. Hum Mutat 25: 293-298.

107. Grody WW, Chang RJ, Panagiotis NM, Matz D, Cederbaum SD (1994) Menstrual cycle and gonadal steroid effects on symptomatic hyperammonaemia of urea-cycle-based and idiopathic aetiologies. J Inherit Metab Dis 17: 566-574.

108. Deignan JL, Cederbaum SD, Grody WW (2008) Contrasting features of urea cycle disorders in human patients and knockout mouse models. Mol Genet Metab 93: 7-14.

109. Summar M, Tuchman M (2001) Proceedings of a consensus conference for the management of patients with urea cycle disorders. J Pediatr 138: S6-10.

110. Herrin JT, McCredie DA (1969) Peritoneal dialysis in the reduction of blood ammonia levels in a case of hyperammonaemia. Arch Dis Child 44: 149-151.

111. Wiegand C, Thompson T, Bock GH, Mathis RK, Kjellstrand CM, et al. (1980) The management of life-threatening hyperammonemia: a comparison of several therapeutic modalities. J Pediatr 96: 142-144.

112. Kim S, Paik WK, Cohen PP (1972) Ammonia intoxication in rats: protection by N-carbamoyl-L-glutamate plus L-arginine. Proc Natl Acad Sci U S A 69: 3530-3533.

113. Guffon N, Schiff M, Cheillan D, Wermuth B, Haberle J, et al. (2005) Neonatal hyperammonemia: the N-carbamoyl-L-glutamic acid test. J Pediatr 147: 260-262.

114. Nordenstrom A, Halldin M, Hallberg B, Alm J (2007) A trial with N-carbamylglutamate may not detect all patients with NAGS deficiency and neonatal onset. J Inherit Metab Dis 30: 400.

115. Guffon N, Vianey-Saban C, Bourgeois J, Rabier D, Colombo JP, et al. (1995) A new neonatal case of N-acetylglutamate synthase deficiency treated by carbamylglutamate. J Inherit Metab Dis 18: 61-65.

116. Tuchman M, Caldovic L, Daikhin Y, Horyn O, Nissim I, et al. (2008) N-carbamylglutamate markedly enhances ureagenesis in N-acetylglutamate deficiency and propionic acidemia as measured by isotopic incorporation and blood biomarkers. Pediatr Res 64: 213-217.

117. Belanger-Quintana A, Martinez-Pardo M, Garcia MJ, Wermuth B, Torres J, et al. (2003) Hyperammonaemia as a cause of psychosis in an adolescent. Eur J Pediatr 162: 773-775.

118. Schubiger G, Bachmann C, Barben P, Colombo JP, Tonz O, et al. (1991) N-acetylglutamate synthetase deficiency: diagnosis, management and follow-up of a rare disorder of ammonia detoxication. Eur J Pediatr 150: 353-356.

119. Dionisi-Vici C, Rizzo C, Burlina AB, Caruso U, Sabetta G, et al. (2002) Inborn errors of metabolism in the Italian pediatric population: a national retrospective survey. J Pediatr 140: 321-327.

120. Nicolaides P, Liebsch D, Dale N, Leonard J, Surtees R (2002) Neurological outcome of patients with ornithine carbamoyltransferase deficiency. Arch Dis Child 86: 54-56.

121. Lindgren V, de Martinville B, Horwich AL, Rosenberg LE, Francke U (1984) Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. Science 226: 698-700.

122. McCullough BA, Yudkoff M, Batshaw ML, Wilson JM, Raper SE, et al. (2000) Genotype spectrum of ornithine transcarbamylase deficiency: correlation with the clinical and biochemical phenotype. Am J Med Genet 93: 313-319.

123. Yamaguchi S, Brailey LL, Morizono H, Bale AE, Tuchman M (2006) Mutations and polymorphisms in the human ornithine transcarbamylase (OTC) gene. Hum Mutat 27: 626-632.

124. Maestri NE, Lord C, Glynn M, Bale A, Brusilow SW (1998) The phenotype of ostensibly healthy women who are carriers for ornithine transcarbamylase deficiency. Medicine (Baltimore) 77: 389-397.

125. Ricciuti FC, Gelehrter TD, Rosenberg LE (1976) X-chromosome inactivation in human liver: confirmation of X-linkage of ornithine transcarbamylase. Am J Hum Genet 28: 332-338.

126. Schultz RE, Salo MK (2000) Under recognition of late onset ornithine transcarbamylase deficiency. Arch Dis Child 82: 390-391.

127. Gordon N (2003) Ornithine transcarbamylase deficiency: a urea cycle defect. Eur J Paediatr Neurol 7: 115-121.

128. Tuchman M, McCullough BA, Yudkoff M (2000) The molecular basis of ornithine transcarbamylase deficiency. Eur J Pediatr 159 Suppl 3: S196-198.

129. Busuttil AA, Goss JA, Seu P, Dulkanchainun TS, Yanni GS, et al. (1998) The role of orthotopic liver transplantation in the treatment of ornithine transcarbamylase deficiency. Liver Transpl Surg 4: 350-354.

130. Snyder RO, Miao CH, Patijn GA, Spratt SK, Danos O, et al. (1997) Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors. Nat Genet 16: 270-276.

131. Mount JD, Herzog RW, Tillson DM, Goodman SA, Robinson N, et al. (2002) Sustained phenotypic correction of hemophilia B dogs with a factor IX null mutation by liver-directed gene therapy. Blood 99: 2670-2676.

132. Nathwani AC, Gray JT, Ng CY, Zhou J, Spence Y, et al. (2006) Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. Blood 107: 2653-2661.

133. Nathwani AC, Gray JT, McIntosh J, Ng CY, Zhou J, et al. (2007) Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. Blood 109: 1414-1421.

134. Wang L, Morizono H, Lin J, Bell P, Jones D, et al. Preclinical evaluation of a clinical candidate AAV8 vector for ornithine transcarbamylase (OTC) deficiency reveals functional enzyme from each persisting vector genome. Mol Genet Metab 105: 203-211.

135. Goping IS, Shore GC (1994) Interactions between repressor and anti-repressor elements in the carbamyl phosphate synthetase I promoter. J Biol Chem 269: 3891-3896.

136. Lagace M, Goping IS, Mueller CR, Lazzaro M, Shore GC (1992) The carbamyl phosphate synthetase promoter contains multiple binding sites for C/EBP-related proteins. Gene 118: 231-238.

137. Schoneveld OJ, Gaemers IC, Hoogenkamp M, Lamers WH (2005) The role of proximal-enhancer elements in the glucocorticoid regulation of carbamoylphosphate synthetase gene transcription from the upstream response unit. Biochimie 87: 1033-1040.

138. Boyce FM, 3rd, Pogulis RJ, Freytag SO (1989) Paradoxical regulation of human argininosuccinate synthetase cDNA minigene in opposition to endogenous gene: evidence for intragenic control sequences. Somat Cell Mol Genet 15: 123-129.

139. Matsubasa T, Takiguchi M, Matsuda I, Mori M (1994) Rat argininosuccinate lyase promoter: the dyad-symmetric CCAAT box sequence CCAATTGG in the promoter is recognized by NF-Y. J Biochem 116: 1044-1055.

140. Santoro C, Mermod N, Andrews PC, Tjian R (1988) A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature 334: 218-224.

141. Ryall JC, Quantz MA, Shore GC (1986) Rat liver and intestinal mucosa differ in the developmental pattern and hormonal regulation of carbamoyl-phosphate synthetase I and ornithine carbamoyl transferase gene expression. Eur J Biochem 156: 453-458.

142. Hazra A, DuBois DC, Almon RR, Snyder GH, Jusko WJ (2008) Pharmacodynamic modeling of acute and chronic effects of methylprednisolone on hepatic urea cycle genes in rats. Gene Regul Syst Bio 2: 1-19.

143. Murakami T, Takiguchi M, Inomoto T, Yamamura K, Mori M (1989) Tissue- and developmental stage-specific expression of the rat ornithine carbamoyltransferase gene in transgenic mice. Dev Genet 10: 393-401.

144. Snodgrass PJ (1991) Dexamethasone and glucagon cause synergistic increases of urea cycle enzyme activities in livers of normal but not adrenalectomized rats. Enzyme 45: 30-38.

145. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

146. Frith MC, Fu Y, Yu L, Chen JF, Hansen U, et al. (2004) Detection of functional DNA motifs via statistical over-representation. Nucleic Acids Res 32: 1372-1381.

147. Marino-Ramirez L, Spouge JL, Kanga GC, Landsman D (2004) Statistical analysis of over-represented words in human promoter sequences. Nucleic Acids Res 32: 949-958.

148. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, et al. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31: 3497-3500.

149. Kim HJ, Ko MS, Kim HK, Cho WJ, Lee SH, et al. Transcription factor Sp1 regulates basal transcription of the human DRG2 gene. Biochim Biophys Acta 1809: 184-190.

150. Zhang W, Tian Z, Sha S, Cheng LY, Philipsen S, et al. Functional and sequence analysis of human neuroglobin gene promoter region. Biochim Biophys Acta 1809: 236-244.

151. Convertini P, Infantino V, Bisaccia F, Palmieri F, Iacobazzi V Role of FOXA and Sp1 in mitochondrial acylcarnitine carrier gene expression in different cell lines. Biochem Biophys Res Commun 404: 376-381.

152. Michels AJ, Hagen TM (2009) Hepatocyte nuclear factor 1 is essential for transcription of sodium-dependent vitamin C transporter protein 1. Am J Physiol Cell Physiol 297: C1220-1227.

153. Wang Z, Burke PA Hepatocyte nuclear factor-4alpha interacts with other hepatocyte nuclear factors in regulating transthyretin gene expression. FEBS J 277: 4066-4075.

154. Tue NT, Yoshioka Y, Yamaguchi M NF-Y transcriptionally regulates the Drosophila p53 gene. Gene 473: 1-7.

155. Pallai R, Simpkins H, Chen J, Parekh HK The CCAAT box binding transcription factor, nuclear factor-Y (NF-Y) regulates transcription of human aldo-keto reductase 1C1 (AKR1C1) gene. Gene 459: 11-23.

156. Xiang H, Wang J, Boxer LM (2006) Role of the cyclic AMP response element in the bcl-2 promoter in the regulation of endogenous Bcl-2 expression and apoptosis in murine B cells. Mol Cell Biol 26: 8599-8606.

157. Callens N, Baert JL, Monte D, Sunesen M, Van Lint C, et al. (2003) Transcriptional regulation of the murine brca2 gene by CREB/ATF transcription factors. Biochem Biophys Res Commun 312: 702-707.

158. Deng WG, Zhu Y, Montero A, Wu KK (2003) Quantitative analysis of binding of transcription factor complex to biotinylated DNA probe by a streptavidin-agarose pulldown assay. Anal Biochem 323: 12-18.

159. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20: 110-121.

160. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, et al. (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15: 1034-1050.

161. Christoffels VM, van den Hoff MJ, Lamers MC, van Roon MA, de Boer PA, et al. (1996) The upstream regulatory region of the carbamoyl-phosphate synthetase I gene controls its tissue-specific, developmental, and hormonal regulation in vivo. J Biol Chem 271: 31243-31250.

162. Christoffels VM, van den Hoff MJ, Moorman AF, Lamers WH (1995) The far-upstream enhancer of the carbamoyl-phosphate synthetase I gene is responsible for the tissue specificity and hormone inducibility of its expression. J Biol Chem 270: 24932-24940.

163. Klein H, Vingron M (2007) Using transcription factor binding site co-occurrence to predict regulatory regions. Genome Inform 18: 109-118.

164. Renga B, Mencarelli A, Cipriani S, D'Amore C, Zampella A, et al. The nuclear receptor FXR regulates hepatic transport and metabolism of glutamine and glutamate. Biochim Biophys Acta 1812: 1522-1531.

165. Emami KH, Burke TW, Smale ST (1998) Sp1 activation of a TATA-less promoter requires a species-specific interaction involving transcription factor IID. Nucleic Acids Res 26: 839-846.

166. Muckenfuss H, Kaiser JK, Krebil E, Battenberg M, Schwer C, et al. (2007) Sp1 and Sp3 regulate basal transcription of the human APOBEC3G gene. Nucleic Acids Res 35: 3784-3796.

167. Schibler U, Sierra F (1987) Alternative promoters in developmental gene expression. Annu Rev Genet 21: 237-257.

168. Hai T, Hartman MG (2001) The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis. Gene 273: 1-11.

169. De Cesare D, Sassone-Corsi P (2000) Transcriptional regulation by cyclic AMP-responsive factors. Prog Nucleic Acid Res Mol Biol 64: 343-369.

170. Matuoka K, Yu Chen K (1999) Nuclear factor Y (NF-Y) and cellular senescence. Exp Cell Res 253: 365-371.

171. Juang HH, Costello LC, Franklin RB (1995) Androgen modulation of multiple transcription start sites of the mitochondrial aspartate aminotransferase gene in rat prostate. J Biol Chem 270: 12629-12634.

172. Pave-Preux M, Aggerbeck M, Veyssier C, Bousquet-Lemercier B, Hanoune J, et al. (1990) Hormonal discrimination among transcription start sites of aspartate aminotransferase. J Biol Chem 265: 4444-4448.

173. Kadonaga JT, Courey AJ, Ladika J, Tjian R (1988) Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science 242: 1566-1570.

174. Li L, He S, Sun JM, Davie JR (2004) Gene regulation by Sp1 and Sp3. Biochem Cell Biol 82: 460-471.

175. Solomon SS, Majumdar G, Martinez-Hernandez A, Raghow R (2008) A critical role of Sp1 transcription factor in regulating gene expression in response to insulin and other hormones. Life Sci 83: 305-312.

176. Wierstra I (2008) Sp1: emerging roles--beyond constitutive activation of TATA-less housekeeping genes. Biochem Biophys Res Commun 372: 1-13.

177. Montminy M, Koo SH, Zhang X (2004) The CREB family: key regulators of hepatic metabolism. Ann Endocrinol (Paris) 65: 73-75.

178. Mayr B, Montminy M (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol 2: 599-609.

179. Kuo CJ, Conley PB, Chen L, Sladek FM, Darnell JE, Jr., et al. (1992) A transcriptional hierarchy involved in mammalian cell-type specification. Nature 355: 457-461.

180. Sladek FM (1993) Orphan receptor HNF-4 and liver-specific gene expression. Receptor 3: 223-232.

181. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, et al. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 99: 16899-16903.

182. Lomenick JP, Melguizo MS, Mitchell SL, Summar ML, Anderson JW (2009) Effects of meals high in carbohydrate, protein, and fat on ghrelin and peptide YY secretion in prepubertal children. J Clin Endocrinol Metab 94: 4463-4471.

183. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447: 799-816.

184. Mitchell S, Murdock D, Summar M (2008) Plasma peptide tyrosine tyrosine (PYY) levels are increased in urea cycle disorder patients. Mol Gen Metab 93: 258 (abstract).

185. Neill MA, Aschner J, Barr F, Summar ML (2009) Quantitative RT-PCR comparison of the urea and nitric oxide cycle gene transcripts in adult human tissues. Mol Genet Metab 97: 121-127.

186. Myrsen-Axcrona U, Ekblad E, Sundler F (1997) Developmental expression of NPY, PYY and PP in the rat pancreas and their coexistence with islet hormones. Regul Pept 68: 165-175.

187. Ekblad E, Sundler F (2002) Distribution of pancreatic polypeptide and peptide YY. Peptides 23: 251-261.

188. Lundberg JM, Tatemoto K, Terenius L, Hellstrom PM, Mutt V, et al. (1982) Localization of peptide YY (PYY) in gastrointestinal endocrine cells and effects on intestinal blood flow and motility. Proc Natl Acad Sci U S A 79: 4471-4475.

189. Lukinius AI, Ericsson JL, Lundqvist MK, Wilander EM (1986) Ultrastructural localization of serotonin and polypeptide YY (PYY) in endocrine cells of the human rectum. J Histochem Cytochem 34: 719-726.

190. Geng M, Li T, Kong X, Song X, Chu W, et al. Reduced expression of intestinal N-acetylglutamate synthase in suckling piglets: a novel molecular mechanism for arginine as a nutritionally essential amino acid for neonates. Amino Acids 40: 1513-1522.

191. Uchiyama C, Mori M, Tatibana M (1981) Subcellular localization and properties of N-acetylglutamate synthase in rat small intestinal mucosa. J Biochem 89: 1777-1786.

192. Dubois N, Cavard C, Chasse JF, Kamoun P, Briand P (1988) Compared expression levels of ornithine transcarbamylase and carbamylphosphate synthetase in liver and small intestine of normal and mutant mice. Biochim Biophys Acta 950: 321-328.

193. Ishihara K, Oshimura M, Nakao M (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. Mol Cell 23: 733-742.

194. Renda M, Baglivo I, Burgess-Beusse B, Esposito S, Fattorusso R, et al. (2007) Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. J Biol Chem 282: 33336-33345.

195. Majumder P, Gomez JA, Chadwick BP, Boss JM (2008) The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J Exp Med 205: 785-798.

196. Luksan O, Jirsa M, Eberova J, Minks J, Treslova H, et al. (2010) Disruption of OTC promoter-enhancer interaction in a patient with symptoms of ornithine carbamoyltransferase deficiency. Hum Mutat 31: E1294-1303.

197. Batshaw ML (1984) Hyperammonemia. Curr Probl Pediatr 14: 1-69.

198. Bachmann C, Krahenbuhl S, Colombo JP, Schubiger G, Jaggi KH, et al. (1981) N-acetylglutamate synthetase deficiency: a disorder of ammonia detoxication. N Engl J Med 304: 543.

199. Bachmann C, Brandis M, Weissenbarth-Riedel E, Burghard R, Colombo JP (1988) N-acetylglutamate synthetase deficiency, a second patient. J Inherit Metab Dis 11: 191-193.

200. Heckmann M, Wermuth B, Haberle J, Koch HG, Gortner L, et al. (2005) Misleading diagnosis of partial N-acetylglutamate synthase deficiency based on enzyme measurement corrected by mutation analysis. Acta Paediatr 94: 121-124.

201. Hwu WL, Chien YH, Tang NL, Law LK, Lin CY, et al. (2007) Deficiency of the carnitine transporter (OCTN2) with partial N-acetylglutamate synthase (NAGS) deficiency. J Inherit Metab Dis 30: 816.

202. Caldovic L, Morizono H, Panglao MG, Cheng SF, Packman S, et al. (2003) Null mutations in the N-acetylglutamate synthase gene associated with acute neonatal disease and hyperammonemia. Hum Genet 112: 364-368.

203. Hall LM, Metzenberg RL, Cohen PP (1958) Isolation and characterization of a naturally occurring cofactor of carbamyl phosphate biosynthesis. J Biol Chem 230: 1013-1021.

204. Caldovic L, Morizono H, Daikhin Y, Nissim I, McCarter RJ, et al. (2004) Restoration of ureagenesis in N-acetylglutamate synthase deficiency by N-carbamylglutamate. J Pediatr 145: 552-554.

205. Stenson PD, Mort M, Ball EV, Howells K, Phillips AD, et al. (2009) The Human Gene Mutation Database: 2008 update. Genome Med 1: 13.

206. Caldovic L, Morizono H, Tuchman M (2007) Mutations and polymorphisms in the human N-acetylglutamate synthase (NAGS) gene. Hum Mutat 28: 754-759.

207. Kuchler G, Rabier D, Poggi-Travert F, Meyer-Gast D, Bardet J, et al. (1996) Therapeutic use of carbamylglutamate in the case of carbamoyl-phosphate synthetase deficiency. J Inherit Metab Dis 19: 220-222.

208. O'Connor JE, Jorda A, Grisolia S (1985) Acute and chronic effects of carbamyl glutamate on blood urea and ammonia. Eur J Pediatr 143: 196-197.

209. Ah Mew N, McCarter R, Daikhin Y, Nissim I, Yudkoff M, et al. (2010) N-carbamylglutamate augments ureagenesis and reduces ammonia and glutamine in propionic acidemia. Pediatrics 126: e208-214.

210. Ah Mew N, Payan I, Daikhin Y, Nissim I, Tuchman M, et al. (2009) Effects of a single dose of N-carbamylglutamate on the rate of ureagenesis. Mol Genet Metab 98: 325-330.

211. Courtois G, Baumhueter S, Crabtree GR (1988) Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. Proc Natl Acad Sci U S A 85: 7937-7941.

212. Olsen J, Laustsen L, Troelsen J (1994) HNF1 alpha activates the aminopeptidase N promoter in intestinal (Caco-2) cells. FEBS Lett 342: 325-328.

213. Toniatti C, Demartis A, Monaci P, Nicosia A, Ciliberto G (1990) Synergistic trans-activation of the human C-reactive protein promoter by transcription factor HNF-1 binding at two distinct sites. EMBO J 9: 4467-4475.

214. Wu GD, Chen L, Forslund K, Traber PG (1994) Hepatocyte nuclear factor-1 alpha (HNF-1 alpha) and HNF-1 beta regulate transcription via two elements in an intestine-specific promoter. J Biol Chem 269: 17080-17085.

215. Wu KJ, Wilson DR, Shih C, Darlington GJ (1994) The transcription factor HNF1 acts with C/EBP alpha to synergistically activate the human albumin promoter through a novel domain. J Biol Chem 269: 1177-1182.

216. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, et al. (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29: 308-311.

217. Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, et al. (2002) A human genome diversity cell line panel. Science 296: 261-262.

218. Thompson PD, Moyna N, Seip R, Price T, Clarkson P, et al. (2004) Functional polymorphisms associated with human muscle size and strength. Med Sci Sports Exerc 36: 1132-1139.

219. Crossley M, Brownlee GG (1990) Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B. Nature 345: 444-446.

220. Perez-Tur J, Froelich S, Prihar G, Crook R, Baker M, et al. (1995) A mutation in Alzheimer's disease destroying a splice acceptor site in the presenilin-1 gene. Neuroreport 7: 297-301.

221. Sakai T, Ohtani N, McGee TL, Robbins PD, Dryja TP (1991) Oncogenic germ-line mutations in Sp1 and ATF sites in the human retinoblastoma gene. Nature 353: 83-86.

222. Coffee EM, Tolan DR Mutations in the promoter region of the aldolase B gene that cause hereditary fructose intolerance. J Inherit Metab Dis 33: 715-725.

223. Pandya AL, Koch R, Hommes FA, Williams JC (1991) N-acetylglutamate synthetase deficiency: clinical and laboratory observations. J Inherit Metab Dis 14: 685-690.

224. Coude FX, Rabier D, Cathelineau L, Grimber G, Parvy P, et al. (1982) A mechanism for valproate-induced hyperammonemia. Adv Exp Med Biol 153: 153-161.

225. de Meer K, Roef MJ, Kulik W, Jakobs C (1999) In vivo research with stable isotopes in biochemistry, nutrition and clinical medicine: an overview. Isotopes Environ Health Stud 35: 19-37.

226. Kalhan SC, Oliven A, King KC, Lucero C (1986) Role of glucose in the regulation of endogenous glucose production in the human newborn. Pediatr Res 20: 49-52.

227. Brusilow SW, Horwich AL (2001) Urea Cycle Enzymes. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The Metabolic & Molecular Bases of Inherited Disease: McGraw-Hill. pp. 1909-1963.

228. Hata A, Tsuzuki T, Shimada K, Takiguchi M, Mori M, et al. (1986) Isolation and characterization of the human ornithine transcarbamylase gene: structure of the 5'-end region. J Biochem 100: 717-725.

229. Horwich AL, Fenton WA, Williams KR, Kalousek F, Kraus JP, et al. (1984) Structure and expression of a complementary DNA for the nuclear coded precursor of human mitochondrial ornithine transcarbamylase. Science 224: 1068-1074.

230. Horwich AL, Kalousek F, Fenton WA, Pollock RA, Rosenberg LE (1986) Targeting of pre-ornithine transcarbamylase to mitochondria: definition of critical regions and residues in the leader peptide. Cell 44: 451-459.

231. Shi D, Morizono H, Ha Y, Aoyagi M, Tuchman M, et al. (1998) 1.85-A resolution crystal structure of human ornithine transcarbamoylase complexed with N-phosphonacetyl-L-ornithine. Catalytic mechanism and correlation with inherited deficiency. J Biol Chem 273: 34247-34254.

232. Numata S, Koda Y, Ihara K, Sawada T, Okano Y, et al. (2010) Mutant alleles associated with late-onset ornithine transcarbamylase deficiency in male patients have recurrently arisen and have been retained in some populations. J Hum Genet 55: 18-22.

233. Breningstall GN (1986) Neurologic syndromes in hyperammonemic disorders. Pediatr Neurol 2: 253-262.

234. Veres G, Craigen WJ, Caskey CT (1986) The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression. J Biol Chem 261: 7588-7591.

235. Takiguchi M, Murakami T, Miura S, Mori M (1987) Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-linked gene with an atypical promoter. Proc Natl Acad Sci U S A 84: 6136-6140.

236. Kimura T, Christoffels VM, Chowdhury S, Iwase K, Matsuzaki H, et al. (1998) Hypoglycemia-associated hyperammonemia caused by impaired expression of ornithine cycle enzyme genes in C/EBPalpha knockout mice. J Biol Chem 273: 27505-27510.

237. Quental R, Azevedo L, Rubio V, Diogo L, Amorim A (2009) Molecular mechanisms underlying large genomic deletions in ornithine transcarbamylase (OTC) gene. Clin Genet 75: 457-464.

238. Shchelochkov OA, Li FY, Geraghty MT, Gallagher RC, Van Hove JL, et al. (2009) High-frequency detection of deletions and variable rearrangements at the ornithine transcarbamylase (OTC) locus by oligonucleotide array CGH. Mol Genet Metab 96: 97-105.

239. Arranz JA, Madrigal I, Riudor E, Armengol L, Mila M (2007) Complete deletion of ornithine transcarbamylase gene confirmed by CGH array of X chromosome. J Inherit Metab Dis 30: 813.

240. Kim GH, Choi JH, Lee HH, Park S, Kim SS, et al. (2006) Identification of novel mutations in the human ornithine transcarbamylase (OTC) gene of Korean patients with OTC deficiency and transient expression of the mutant proteins in vitro. Hum Mutat 27: 1159.

241. Ogino W, Takeshima Y, Nishiyama A, Okizuka Y, Yagi M, et al. (2007) Mutation analysis of the ornithine transcarbamylase (OTC) gene in five Japanese OTC deficiency patients revealed two known and three novel mutations including a deep intronic mutation. Kobe J Med Sci 53: 229-240.

242. Balasubramaniam S, Rudduck C, Bennetts B, Peters G, Wilcken B, et al. (2010) Contiguous gene deletion syndrome in a female with ornithine transcarbamylase deficiency. Mol Genet Metab 99: 34-41.

243. Jakubiczka S, Bettecken T, Mohnike K, Schneppenheim R, Stumm M, et al. (2007) Symptoms of OTC deficiency but not DMD in a female carrier of an Xp21.1 deletion including the genes for dystrophin and OTC. Eur J Pediatr 166: 743-745.

244. Engel K, Nuoffer JM, Muhlhausen C, Klaus V, Largiader CR, et al. (2008) Analysis of mRNA transcripts improves the success rate of molecular genetic testing in OTC deficiency. Mol Genet Metab 94: 292-297.

245. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.

246. Wingender E (1988) Compilation of transcription regulating proteins. Nucleic Acids Res 16: 1879-1902.

247. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A (2010) Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20: 110-121.

248. Bolotin E, Schnabl JK, Sladek FM HNF4A. Transcription Factor Encyclopedia.

249. Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, et al. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol 9: e1001046.

250. Towle HC (2007) The metabolic sensor GCN2 branches out. Cell Metab 5: 85-87.

251. Dubbelhuis PF, Meijer AJ (2002) Hepatic amino acid-dependent signaling is under the control of AMP-dependent protein kinase. FEBS Lett 521: 39-42.

252. Ishiguro H, Katano Y, Nakano I, Ishigami M, Hayashi K, et al. (2006) Clofibrate treatment promotes branched-chain amino acid catabolism and decreases the phosphorylation state of mTOR, eIF4E-BP1, and S6K1 in rat liver. Life Sci 79: 737-743.

253. Carraro V, Maurin AC, Lambert-Langlais S, Averous J, Chaveroux C, et al. (2010) Amino acid availability controls TRB3 transcription in liver through the GCN2/eIF2alpha/ATF4 pathway. PLoS One 5: e15716.

254. Yokoyama Y, Iguchi K, Usui S, Hirano K (2011) AMP-activated protein kinase modulates the gene expression of aquaporin 9 via forkhead box a2. Arch Biochem Biophys 515: 80-88.

255. Prieur X, Schaap FG, Coste H, Rodriguez JC (2005) Hepatocyte nuclear factor-4alpha regulates the human apolipoprotein AV gene: identification of a novel response element and involvement in the control by peroxisome proliferator-activated receptor-gamma coactivator-1alpha, AMP-activated protein kinase, and mitogen-activated protein kinase pathway. Mol Endocrinol 19: 3107-3125.

256. Leclerc I, Lenzner C, Gourdon L, Vaulont S, Kahn A, et al. (2001) Hepatocyte nuclear factor-4alpha involved in type 1 maturity-onset diabetes of the young is a novel target of AMP-activated protein kinase. Diabetes 50: 1515-1521.

257. Bauer M, Hamm AC, Bonaus M, Jacob A, Jaekel J, et al. (2004) Starvation response in mouse liver shows strong correlation with life-span-prolonging processes. Physiol Genomics 17: 230-244.

258. Anthony JC, Anthony TG, Kimball SR, Vary TC, Jefferson LS (2000) Orally administered leucine stimulates protein synthesis in skeletal muscle of postabsorptive rats in association with increased eIF4F formation. J Nutr 130: 139-145.

259. Anthony JC, Anthony TG, Layman DK (1999) Leucine supplementation enhances skeletal muscle recovery in rats following exercise. J Nutr 129: 1102-1106.

260. Buse MG, Reid SS (1975) Leucine. A possible regulator of protein turnover in muscle. J Clin Invest 56: 1250-1261.

261. Li JB, Jefferson LS (1978) Influence of amino acid availability on protein turnover in perfused skeletal muscle. Biochim Biophys Acta 544: 351-359.

262. Schimke RT (1961) Studies on adaptation of urea cycle enzymes in the rat. Cold Spring Harb Symp Quant Biol 26: 363-366.

263. Seo J, Gordish-Dressman H, Hoffman EP (2006) An interactive power analysis tool for microarray hypothesis testing and generation. Bioinformatics 22: 808-814.

264. Zhao P, Seo J, Wang Z, Wang Y, Shneiderman B, et al. (2003) In vivo filtering of in vitro expression data reveals MyoD targets. C R Biol 326: 1049-1065.

265. Li J, Wong L (2001) Emerging patterns and gene expression data. Genome Inform 12: 3-13.

266. Paddon-Jones D, Westman E, Mattes RD, Wolfe RR, Astrup A, et al. (2008) Protein, weight management, and satiety. Am J Clin Nutr 87: 1558S-1561S.

267. Veldhorst M, Smeets A, Soenen S, Hochstenbach-Waelen A, Hursel R, et al. (2008) Protein-induced satiety: effects and mechanisms of different proteins. Physiol Behav 94: 300-307.

268. Petersen KF, Vilstrup H, Tygstrup N (1990) Effect of dietary protein on the capacity of urea synthesis in rats. Horm Metab Res 22: 612-615.

269. Hager GL, Nagaich AK, Johnson TA, Walker DA, John S (2004) Dynamics of nuclear receptor movement and transcription. Biochim Biophys Acta 1677: 46-51.

270. Brown EJ, Albers MW, Shin TB, Ichikawa K, Keith CT, et al. (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature 369: 756-758.

271. Wiederrecht GJ, Sabers CJ, Brunn GJ, Martin MM, Dumont FJ, et al. (1995) Mechanism of action of rapamycin: new insights into the regulation of G1-phase progression in eukaryotic cells. Prog Cell Cycle Res 1: 53-71.

272. Corton JM, Gillespie JG, Hawley SA, Hardie DG (1995) 5-aminoimidazole-4-carboxamide ribonucleoside. A specific method for activating AMP-activated protein kinase in intact cells? Eur J Biochem 229: 558-565.

273. Henin N, Vincent MF, Van den Berghe G (1996) Stimulation of rat liver AMP-activated protein kinase by AMP analogues. Biochim Biophys Acta 1290: 197-203.

274. Kimball SR (2001) Regulation of translation initiation by amino acids in eukaryotic cells. Prog Mol Subcell Biol 26: 155-184.

Curriculum Vitae

Sandra Kirsch Heibel

37C Ridge Rd, Greenbelt, MD 20770
sbk135@gmail.com
(202) 997-6306

EDUCATION

Doctor of Philosophy, Molecular and Cellular Biology Completion May 2012
University of Maryland, College Park, MD & Children's National Medical Cntr, Washington, DC
Dissertation Title: Transcriptional Regulation of N-Acetylglutamate Synthase with Implications for Patient Diagnosis.

Master of Science, Animal Science August 2007
University of Maryland, College Park, MD
Thesis Title: Characterization of Chicken CAT-2 Isoforms.

Bachelor of Science, with honors, Animal Bioscience May 2005
The Pennsylvania State University, University Park, PA
Thesis Title: The Effects of 15-HPETE on HIV Output from Microglial Cell Line CHME-3.

HIGHLIGHTED EXPERIENCE

Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC
Molecular and Cell Biology Program, University of Maryland, College Park, MD
Graduate Research Assistant August 2007-present

- Designed and conducted experiments, optimized methodologies, analyzed data, and evaluated results sufficient to complete a translational research doctoral program on the regulation of Urea Cycle enzyme N-Acetylglutamate Synthase and disease causation of N-Acetylglutamate Synthase Deficiency.
- Communicated experimental results in three first author publications, on panels at two national and two international conferences, and presented posters at seven national conferences.
- Wrote, implemented and managed a two-year federal grant worth $200,000 with project goals delivered on time and within budget (American Reinvestment and Recovery Act).
- Reviewed two rare disease clinical research studies leading to improved treatment of seven patients.
- Trained and supervised three laboratory technicians and four undergraduate research assistants.
- Recruited and evaluated prospective University of Maryland Molecular Biology Doctoral students.

Department of Animal and Avian Sciences, University of Maryland, College Park, MD
Graduate Teaching Assistant August 2005-August 2007

- Developed a research project, conducted experiments, and analyzed data sufficient to earn a master's of science degree for research on tissue specificity, cellular localization, and kinetic activity of the lysine and arginine transporters.
- Communicated experimental results in one first author and one second-author publication and presented posters at three national conferences.
- Facilitated classroom discussion and lead laboratory sessions in four science courses.

- Trained and supervised one undergraduate research assistant in experimental technique.

Department of Veterinary Medicine, The Pennsylvania State University, State College, PA
Undergraduate Research Assistant Sept 2003 – May 2005
- Performed experiments to assist a senior graduate student and to achieve Honors in animal bioscience for research on the regulation of HIV viral load by cytokines, chemokines, and anti-oxidants.

SPECIAL COMPETENCIES

Molecular biology: RNA, DNA and protein isolation, RT-PCR, PCR, qPCR, 5'RACE, Chromatin IP, DNA-protein pull-down assay, Microarray transcriptional profiling, DNA recombination, plasmid construction, site-directed mutagenesis, chemical/electroporetic cell transformation, subcloning, DNA sequencing, protein quantification, SDS-PAGE, Western transfer, protein phosphorylation analysis, in-gel protein digestion and extraction, column protein purification, proteomic profiling by Mass Spectrometry, luciferase reporter assays, SNP genotyping

Cell biology: Culture cell lines, stable/transient transfection, immunofluorescent staining, radio-isotope transport kinetic assays, cell viability assays, light, fluorescent, and laser confocal microscopy,

Bioinformatics software and databases: NCBI, UCSC genome browser, Cis-eLement OVER representation software, NCBI Blast, dbSNP, Partek, Ingenuity IPA, Ingenuity iReport, ProteoIQ

AFFILIATIONS

- Women in Bio: Member July 2011-present
- Society for Biochemistry and Molecular Biology: Member September 2008-present

LEADERSHIP AND SERVICE ROLES

Young Women in Bio (Women in Bio outreach program): Member July 2011-present
- Developed programs to introduce girls 12 to 18 years to careers in the biological sciences.
- Facilitated partnerships between Women in Bio and local bioscience companies.

Animal Science Graduate Student Association: Member August 2005-August 2007
- Advocated for improved dialogue between graduate students and administration.
- Planned and coordinated graduate student fundraising and community events.

Graduate Student Government: Animal Science Representative August 2006-August 2007
- Promoted initiatives to set a minimum graduate student salary throughout the University.

TEACHING EXPERIENCE

University of Maryland
- Bioscience 105- Principles of Biology Fall 2007
- Biometrics 301- Introduction to Biometrics Fall 2006
- Animal Science 315- Applied Animal Nutrition Spring 2006
- Animal Science 314- Comparative Animal Nutrition Fall 2005

HONORS AND AWARDS

- Society of Inherited Metabolic Disorders Meeting Travel Award March 2010
- Society for Biochemistry and Molecular Biology Graduate Travel Award April 2009
- Molecular and Cell Biology Recruitment Fellowship Fall 2007-Spring 2008
- Animal Science Annual Symposium Travel Award June 2007

- Richard F. Davis Memorial Academic Excellence Scholarship Fall 2006
- Schreyer Honors College Scholarship Fall 2001-Spring 2005
- Presidential Academic Excellence Scholarship Fall 2003-Spring 2005

PUBLICATIONS

SK Heibel, AL LaBella, TP Feeney, M Panglao, S Sodha, GY Lopez, M Tuchman, L Marino-Ramirez, H Morizono, L Caldovic. (2012) Disease-causing Mutations in the Promoter of the Ornithine Transcarbamylase Gene in Patients Without Molecular Diagnosis. *Ann Hum Genet.* [submitted for publication]

SK Heibel, G Lopez, M Panglao, S Sodha, L Mariño-Ramírez, M Tuchman, L Caldovic. (2012) Transcriptional regulation of N-Acetylglutamate synthase. *PLoS ONE* 7(2):e29527

SK Heibel, N Ah Mew, L Caldovic, Y Daikhin, M Yudkoff, M Tuchman. (2011) N-Carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the *NAGS* gene and to effective therapy. *Hum Mut*. 32(10):1153-60

MJ Harding, E Barnaeva, N Agladze, M Bustamante, K Reyes-Gordillo, CM Lepus, C Rahner, **SK Heibel**, H Morizono, JS Pober, BD Lindenbach, MH Sjogren, M Rojkind. 2011. Characterization of a novel human hepatocyte/ human hepatic stellate cell co-culture system. [pending]

S Kirsch, BD Humphrey. (2010) Functional characterization of the chicken cationic amino acid transporter-2 isoforms. *Comp Biochem Physiol B Biochem Mol Biol.* 156(4):279-86.

BD Humphrey, **S Kirsch**, D Morris. 2008. Molecular cloning and characterization of chicken cationic amino acid transporter-2 gene. *Comp Biochem Physiol B Biochem Mol Biol.* 150(3):301-11.

PRESENTATIONS

SK Heibel, *et al*. (2012) "Nitrogen Load Sensing Pathways." Society of Inherited Metabolic Disorders Meeting

SK Heibel, *et al.* (2011) "N-Carbamylglutamate Enhancement of Ureagenesis Leads to Discovery of a Mutation in the Enhancer of the NAGS Gene." Society of Inherited Metabolic Disorders Meeting; Experimental Biology Annual Meeting; Children's National Research Week (Best Poster).

SK Heibel, *et al.* (2010) "A Urea Cycle Nitrogen Load Sensor: Transcriptional Regulation of N-Acetylglutamate Synthase." International Conference on Arginine and Pyrimidines; Society of Inherited Metabolic Disorders Meeting (Travel Award Recipient).

SK Heibel, *et al.* (2008, 2009) "Transcriptional Regulation of N-Acetylglutamate Synthase." International Conference on Arginine and Pyrimidines; Experimental Biology Annual Meeting (Travel Award Recipient); Molecular and Cell Biology Annual Retreat (Best Poster)

Curriculum Vitae

Sandra Kirsch Heibel

37C Ridge Rd, Greenbelt, MD 20770
sbk135@gmail.com
(202) 997-6306

EDUCATION

Doctor of Philosophy, Molecular and Cellular Biology Completion May 2012
University of Maryland, College Park, MD & Children's National Medical Cntr, Washington, DC
Dissertation Title: Transcriptional Regulation of N-Acetylglutamate Synthase and Its Clinical Relevance .

Master of Science, Animal Science August 2007
University of Maryland, College Park, MD
Thesis Title: Characterization of Chicken CAT-2 Isoforms.

Bachelor of Science, with honors, Animal Bioscience May 2005
The Pennsylvania State University, University Park, PA
Thesis Title: The Effects of 15-HPETE on HIV Output from Microglial Cell Line CHME-3.

HIGHLIGHTED EXPERIENCE

Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC
Molecular and Cell Biology Program, University of Maryland, College Park, MD
Graduate Research Assistant August 2007-present

- Designed and conducted experiments, optimized methodologies, analyzed data, and evaluated results sufficient to complete a translational research doctoral program on the regulation of Urea Cycle enzyme N-Acetylglutamate Synthase and disease causation of N-Acetylglutamate Synthase Deficiency.
- Communicated experimental results in three first author publications, on panels at two national and two international conferences, and presented posters at seven national conferences.
- Wrote, implemented and managed a two-year federal grant worth $200,000 with project goals delivered on time and within budget (American Reinvestment and Recovery Act).
- Reviewed two rare disease clinical research studies leading to improved treatment of seven patients.
- Trained and supervised three laboratory technicians and four undergraduate research assistants.
- Recruited and evaluated prospective University of Maryland Molecular Biology Doctoral students.

Department of Animal and Avian Sciences, University of Maryland, College Park, MD
Graduate Teaching Assistant August 2005-August 2007

- Developed a research project, conducted experiments, and analyzed data sufficient to earn a master's of science degree for research on tissue specificity, cellular localization, and kinetic activity of the lysine and arginine transporters.
- Communicated experimental results in one first author and one second-author publication and presented posters at three national conferences.

- Facilitated classroom discussion and lead laboratory sessions in four science courses.
- Trained and supervised one undergraduate research assistant in experimental technique.

Department of Veterinary Medicine, The Pennsylvania State University, State College, PA
Undergraduate Research Assistant Sept 2003 – May 2005

- Performed experiments to assist a senior graduate student and to achieve Honors in animal bioscience for research on the regulation of HIV viral load by cytokines, chemokines, and anti-oxidants.

SPECIAL COMPETENCIES

Molecular biology: RNA, DNA and protein isolation, RT-PCR, PCR, qPCR, 5'RACE, Chromatin IP, DNA-protein pull-down assay, Microarray transcriptional profiling, DNA recombination, plasmid construction, site-directed mutagenesis, chemical/electroporetic cell transformation, subcloning, DNA sequencing, protein quantification, SDS-PAGE, Western transfer, protein phosphorylation analysis, in-gel protein digestion and extraction, column protein purification, proteomic profiling by Mass Spectrometry, luciferase reporter assays, SNP genotyping

Cell biology: Culture cell lines, stable/transient transfection, immunofluorescent staining, radio-isotope transport kinetic assays, cell viability assays, light, fluorescent, and laser confocal microscopy,

Bioinformatics software and databases: NCBI, UCSC genome browser, Cis-eLement OVER representation software, NCBI Blast, dbSNP, Partek, Ingenuity IPA, Ingenuity iReport, ProteoIQ

AFFILIATIONS

- Women in Bio: Chapter Lead July 2011-present
- Society for Biochemistry and Molecular Biology: Member September 2008-present

LEADERSHIP AND SERVICE ROLES

Young Women in Bio (Women in Bio outreach program): Member July 2011-present

- Developed programs to introduce girls 12 to 18 years to careers in the biological sciences.
- Facilitated partnerships between Women in Bio and local bioscience companies.

Animal Science Graduate Student Association: Member August 2005-August 2007

- Advocated for improved dialogue between graduate students and administration.
- Planned and coordinated graduate student fundraising and community events.

Graduate Student Government: Animal Science Representative August 2006-August 2007

- Promoted initiatives to set a minimum graduate student salary throughout the University.

TEACHING EXPERIENCE

University of Maryland

- Bioscience 105- Principles of Biology Fall 2007
- Biometrics 301- Introduction to Biometrics Fall 2006
- Animal Science 315- Applied Animal Nutrition Spring 2006
- Animal Science 314- Comparative Animal Nutrition Fall 2005

HONORS AND AWARDS

- Society of Inherited Metabolic Disorders Meeting Travel Award March 2010
- Society for Biochemistry and Molecular Biology Graduate Travel Award April 2009
- Molecular and Cell Biology Recruitment Fellowship Fall 2007-Spring 2008

- Animal Science Annual Symposium Travel Award — June 2007
- Richard F. Davis Memorial Academic Excellence Scholarship — Fall 2006
- Schreyer Honors College Scholarship — Fall 2001-Spring 2005
- Presidential Academic Excellence Scholarship — Fall 2003-Spring 2005

PUBLICATIONS

SK Heibel, AL LaBella, TP Feeney, M Panglao, S Sodha, GY Lopez, M Tuchman, L Marino-Ramirez, H Morizono, L Caldovic. (2012) Disease-causing Mutations in the Promoter of the Ornithine Transcarbamylase Gene in Patients Without Molecular Diagnosis. *Ann Hum Genet.* [submitted for publication]

SK Heibel, G Lopez, M Panglao, S Sodha, L Mariño-Ramírez, M Tuchman, L Caldovic. (2012) Transcriptional regulation of N-Acetylglutamate synthase. *PLoS ONE* 7(2):e29527

SK Heibel, N Ah Mew, L Caldovic, Y Daikhin, M Yudkoff, M Tuchman. (2011) N-Carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the *NAGS* gene and to effective therapy. *Hum Mut*. 32(10):1153-60

MJ Harding, E Barnaeva, N Agladze, M Bustamante, K Reyes-Gordillo, CM Lepus, C Rahner, **SK Heibel**, H Morizono, JS Pober, BD Lindenbach, MH Sjogren, M Rojkind. 2011. Characterization of a novel human hepatocyte/ human hepatic stellate cell co-culture system. [pending]

S Kirsch, BD Humphrey. (2010) Functional characterization of the chicken cationic amino acid transporter-2 isoforms. *Comp Biochem Physiol B Biochem Mol Biol.* 156(4):279-86.

BD Humphrey, **S Kirsch**, D Morris. 2008. Molecular cloning and characterization of chicken cationic amino acid transporter-2 gene. *Comp Biochem Physiol B Biochem Mol Biol.* 150(3):301-11.

PRESENTATIONS

SK Heibel, *et al*. (2012) "Nitrogen Load Sensing Pathways." Society of Inherited Metabolic Disorders Meeting, Children's National Medical Center Research Week

SK Heibel, *et al.* (2011) "N-Carbamylglutamate Enhancement of Ureagenesis Leads to Discovery of a Mutation in the Enhancer of the NAGS Gene." Society of Inherited Metabolic Disorders Meeting; Experimental Biology Annual Meeting; Children's National Research Week (Best Poster).

SK Heibel, *et al.* (2010) "A Urea Cycle Nitrogen Load Sensor: Transcriptional Regulation of N-Acetylglutamate Synthase." International Conference on Arginine and Pyrimidines; Society of Inherited Metabolic Disorders Meeting (Travel Award Recipient).

SK Heibel, *et al.* (2008, 2009) "Transcriptional Regulation of N-Acetylglutamate Synthase." International Conference on Arginine and Pyrimidines; Experimental Biology Annual Meeting (Travel Award Recipient); Molecular and Cell Biology Annual Retreat (Best Poster)