Molecular Phylogenetics and Evolution 70 (2014) 195–203

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Review

Evolution of glucose utilization: Glucokinase and glucokinase regulator protein

David M. Irwin [a,b,c,*], Huanran Tan [c]

[a] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada
[b] Banting and Best Diabetes Centre, University of Toronto, Toronto, Canada
[c] Department of Pharmacology, Peking University, Health Sciences Centre, Beijing, China

ARTICLE INFO

Article history:
Received 27 May 2013
Revised 14 September 2013
Accepted 17 September 2013
Available online 25 September 2013

Keywords:
Glucokinase (GCK)
Hexokinase (HK)
Glucokinase regulatory protein (GCKR)
Liver
Glucose

ABSTRACT

Glucose is an essential nutrient that must be distributed throughout the body to provide energy to sustain physiological functions. Glucose is delivered to distant tissues via be blood stream, and complex systems have evolved to maintain the levels of glucose within a narrow physiological range. Phosphorylation of glucose, by glucokinase, is an essential component of glucose homeostasis, both from the regulatory and metabolic point-of-view. Here we review the evolution of glucose utilization from the perspective of glucokinase. We discuss the origin of glucokinase, its evolution within the hexokinase gene family, and the evolution of its interacting regulatory partner, glucokinase regulatory protein (GCKR). Evolution of the structure and sequence of both glucokinase and GCKR have been necessary to optimize glucokinase in its role in glucose metabolism.

Contents

1. Mammalian glucose metabolism

Glucose is an essential energy molecule for many species as it can readily be metabolized to generate ATP by either aerobic on anaerobic respiration. In mammals, several tissues, such as the brain, depend upon glucose for the generation of ATP for cellular functions (Wasserman, 2009; Thorens, 2011). To supply these tissues, glucose is distributed by the blood stream, and complex systems have evolved to maintain appropriate blood glucose levels in the face of changes due to the availability of food and the expenditures of energy (Suh et al., 2007; Wasserman, 2009; Polakof et al., 2011; Thorens, 2011). The diet is the ultimate source of glucose in mammals, where it is either directly absorbed from the digested food in the intestine or is generated by gluconeogenesis (generally from amino acids) from precursors obtained from the diet. Whether glucose comes directly from the diet, or is synthesized via gluconeogenesis, depends upon the carbohydrate content of the diet. Counter-regulatory hormone systems have evolved to maintain blood glucose within a narrow range. Insulin promotes glucose uptake from the blood by diverse tissues when blood glucose levels are high; e.g., after feeding, while glucagon induces gluconeogenesis by the liver to release glucose and prevent hypoglycemia (Bansal and Wang, 2008; Wasserman, 2009). Additional

* Corresponding author. Address: Department of Laboratory Medicine and Pathobiology, Room 6211 Medical Sciences Building, 1 King's College Circle, University of Toronto, Toronto, Ont. M5S 1A8, Canada.
E-mail address: david.irwin@utoronto.ca (D.M. Irwin).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.09.016

physiological systems, including other hormones and the central nervous system, contribute to the regulation of blood glucose levels (Polakof et al., 2011; Thorens, 2011; Grayson et al., 2013). Excess energy obtained from the diet is stored as glycogen in the liver and muscle and as lipid in adipose tissue (Klover and Mooney, 2004). Glycogen in muscle can be rapidly broken down to release glucose when it is needed by muscle tissue (Jensen and Richter, 2012); similarly, glycogen is a source for the production of glucose that can be released by the liver to maintain blood glucose levels (Klover and Mooney, 2004).

While glucose can be dispersed to distant tissues of the body through the blood, it still must enter cells, as it cannot cross the lipid membranes of cells by simple diffusion, as it is hydrophilic. In mammals, glucose is transported across the cell membrane by transporters, which belong to one of two families: the glucose transporters (SLC2 or GLUT gene family) (Augustin, 2010; Wilson-O'Brien et al., 2010; Mueckler and Thorens, 2013) and the sodium-coupled glucose transporters (SGLT or SLC5 gene family) (Wright et al., 2011; Wright, 2013). Members of the SGLT are expressed in the kidney and intestine where they actively transport glucose and allow reabsorption against a concentration gradient (Wright et al., 2011; Wright, 2013). The SLC2 family members (or GLUT proteins) are facilitative transports and only move glucose in the direction of a concentration gradient (Augustin, 2010; Mueckler and Thorens, 2013). The diverse members of the SLC2 family have varying substrate specificities, kinetics, and expression profiles allowing cells to have tissue-specific differences in their glucose uptake (Augustin, 2010; Mueckler and Thorens, 2013).

2. Glucokinase and the vertebrate hexokinase gene family

Once glucose (or other simple sugars, e.g., fructose) enter cells it is phosphorylated. Phosphorylation has two purposes: (1) the first step in metabolism, and (2) decreasing the intracellular concentration of the unphosphorylated form of the sugar, thus driving the uptake of the sugar from the external environment. Enzymes, called hexokinases, phosphorylate glucose, and other six-carbon sugars. In mammals, and other vertebrates, four hexokinase isozymes have been identified (Ureta, 1982; Wilson, 1995, 1997, 2003, 2004; Cárdenas et al., 1998). The different isozymes of hexokinase were initially distinguished by letters (i.e., hexokinase (HK) A, B, C, and D) based on their elution time from DEAE cellulose columns (González et al., 1964), but subsequently given numbers (i.e., hexokinase I, II, III, and IV) based on their migration in electrophoretic gels (Katzen et al., 1965). Hexokinase IV (or D) is most often called glucokinase (GCK), although it is not specific for glucose (Cárdenas et al., 1998; Wilson, 2004). Genes encoding these hexokinases use Arabic numbers; e.g., HK1 encodes hexokinase I. Mammalian hexokinases have been extensively characterized, with possibly their most striking difference being their molecular weights (Ureta, 1982; Wilson, 1995, 1997, 2003, 2004; Cárdenas et al., 1998). Hexokinases I, II, and III have a molecular weight of approximately 100 kD, while glucokinase has a molecular mass of about 50 kD. Hexokinases having a molecular weight of 50 kD, but not 100 kD, have also been found, and characterized, in many other eukaryotes species, including non-vertebrate animals, plants, and yeast (Ureta, 1982; Wilson, 1995, 1997, 2003, 2004; Cárdenas et al., 1998). In some of these non-mammalian species multiple isozymes have been characterized (Wilson, 1995, 2004; Cárdenas et al., 1998). While enzymes that phosphorylate sugars have also found in bacteria, these sequences show no significant sequence similarity to the eukaryotic hexokinases. However there is some similarity in the three-dimensional structures of hexokinases from bacteria and eukaryotes, which has been used to suggest that they share a common ancestor (Bork et al., 1993; Cárdenas et al., 1998; Kawai et al., 2005).

Sequences of hexokinases were initially deduced and predicted based on cDNA clones (Andreone et al., 1989; Schwab and Wilson, 1989, 1991; Griffin et al., 1991; Thelen and Wilson, 1991). Analysis of genome sequence data identified, in contrast to the 4 expected known hexokinase enzymes (Ureta, 1982; Wilson, 1995, 1997, 2003, 2004; Cárdenas et al., 1998), a total of 6 hexokinase-like genes in the human genome, and 5 in the genomes of most other vertebrates (Irwin and Tan, 2008; González-Alvarez et al., 2009). In addition to the expected genes encoding the known hexokinases (HK1, HK2, HK3, and GCK), a fifth hexokinase-like gene was found in the genome searches that is conserved throughout vertebrates. The fifth hexokinase-like gene was named hexokinase domain containing 1 (HKDC1), however, the biological function of this gene is currently unknown. SNPs near the HKDC1 gene have been associated with attention-deficit/hyperactivity disorder (Neale et al., 2010) and the protein was found to potentially interact with iNOS2 (Foster et al., 2013). The human genome, as well as those of a few other primates, contains a sixth hexokinase-like sequence, which is a reverse-transcribed pseudogene copy of the HK2 gene (Ardehali et al., 1995; Irwin and Tan, 2008). The HK2 pseudogene was generated relatively recently during primate evolution (Ardehali et al., 1995; Irwin and Tan, 2008).

Characterization of the protein sequences of the hexokinases provided an explanation for the difference in the molecular weights of the 100 kD hexokinases (hexokinases I, II and III) and the 50 kD glucokinase – the larger hexokinases contain two kinases domains, while the smaller glucokinase has only a single kinase domain (Andreone et al., 1989; Schwab and Wilson, 1989, 1991; Griffin et al., 1991; Thelen and Wilson, 1991; Cárdenas et al., 1998). This conclusion is illustrated in Fig. 1, where similarity between the different kinase domains is shown as long diagonal lines in dotplots. Glucokinase contains only a single kinase domain while hexokinase I has two. When hexokinase I is compared to glucokinase (Fig. 1A) two diagonal lines are generated since the N-terminal half of hexokinase I is similar to all of glucokinase (lower diagonal, Fig. 1A), and the C-terminal half of hexokinase I also is similar to all of glucokinase (upper diagonal in Fig. 1A). If glucokinase is compared to itself (Fig. 1B), no repetitive structure is identified, only similarity to itself along the entire sequence (diagonal from lower left to upper right), although there are short sequences that show limited similarity (short lines that are off the central diagonal, Fig. 1, also seen in the other comparisons). If hexokinase I is compared to itself (Fig. 1C) or to hexokinase II (Fig. 1D) a total of three long diagonal lines are seen due to the repetitive structures in both proteins. The central diagonal line indicates that the to proteins are similar to each other over their entire lengths, while the diagonal lines in the upper left and lower right indicate that similarity between the N-terminal and C-terminal halves of the is observed. Similar results are observed with the hexokinase III and HKDC1 (which also predicts an approximately 100 kD protein) sequences (results not shown).

Isolation and partial characterization of the genes encoding hexokinase II and glucokinase (Magnuson et al., 1989; Thelen and Wilson, 1991; Kogure et al., 1993) strengthened the conclusion that the larger hexokinases were generated by duplication of the kinase domain. The intron–exon structures of the two halves of the HK2 gene are similar to each other and to the glucokinase gene, where the sizes of the exons are similar and the phase of the codons interrupted by the introns are identical (Fig. 2). These observations suggesting that the two halves of HK2 were generated by a duplication of glucokinase-like sequence (Magnuson et al., 1989; Thelen and Wilson, 1991; Kogure et al., 1993). The similarity in intron–exon structure is seen not only for HK2, but also for all other members of the mammalian hexokinase gene family