Human pancreatic glucokinase

Structural and physico-chemical studies related to catalytic activation, kinetic cooperativity and GCK-diabetes

Janne Molnes

Dissertation for the degree philosophiae doctor (PhD) at the University of Bergen

2012
Contents

Scientific environment ........................................................................................................ 4
Acknowledgements ............................................................................................................. 5
Abbreviations ..................................................................................................................... 6
Abstract ............................................................................................................................. 8
List of publications ............................................................................................................. 10

1. INTRODUCTION ........................................................................................................... 11

1.1. BACKGROUND .......................................................................................................... 11
1.2. MONOGENIC DIABETES ...................................................................................... 12
1.3. GLUCOKINASE GLYCEMIC DISEASES ............................................................ 14
   1.3.1. GCK-MODY .................................................................................................... 16
          Pathophysiology .................................................................................................. 16
          Clinical features .................................................................................................. 17
          Diagnosis and treatment .................................................................................... 17
          Prevalence .......................................................................................................... 17
   1.3.2. GCK-PNDM ..................................................................................................... 18
   1.3.3. GCK-HI .......................................................................................................... 19
1.4. THE ROLE OF PANCREATIC β-CELL GK IN GLUCOSE-STIMULATED
    INSULIN SECRETION AND GLYCEMIC DISEASE ........................................... 20
1.5. NATURALLY OCCURRING GCK MUTATIONS AND THE MECHANISM OF
    THEIR ASSOCIATED DISEASE .............................................................................. 22
1.6. GK – A MEMBER OF THE HEXOKINASE FAMILY OF ENZYMES WITH
    UNIQUE PROPERTIES ............................................................................................... 23
   1.6.1. Kinetic properties of GK ................................................................................ 24
   1.6.2. Gene structure and tissue-specific gene regulation ....................................... 25
1.7. THE 3D STRUCTURE OF GK AND THE GLUCOSE-INDUCED
    CONFORMATIONAL CHANGE .................................................................................. 26
   1.7.1. The active site ................................................................................................ 28
          Glucose-binding site .......................................................................................... 28
          ATP-binding site ............................................................................................... 29
1.7.2. The C-terminal α-helix ................................................................. 30
1.7.3. The allosteric activator site and pharmacological GK activator drugs .......... 32
1.7.4. Cysteine residues at the active site of GK .................................................. 33

1.8. KINETIC MODELS OF POSITIVE COOPERATIVITY .................................. 34
1.9. REGULATION OF hGK ACTIVITY .......................................................... 37
1.9.1. Post-translational regulation of hGK ......................................................... 38
  Covalent post-translational modifications ....................................................... 40
  The ubiquitin conjugating system and ubiquitin-mediated proteolytic pathway .... 41
  Role of the ubiquitin-proteasome pathway in ß-cell dysfunction and
  hyperglycemia ................................................................................................. 43

2. AIMS OF THE PRESENT STUDY ................................................................. 45
3. SUMMARY OF RESULTS .............................................................................. 46
  3.1. PAPER I ........................................................................................................ 46
  3.2. PAPER II ....................................................................................................... 47
  3.3. PAPER III ...................................................................................................... 48

4. METHODS AND METHODOLOGICAL ASPECTS ...................................... 49
  4.1. FUNCTIONAL COMPARISON OF GST-hGK AND NON-TAGGED hGK ...... 49
  4.2. HELIX NOMENCLATURE ........................................................................... 50
  4.3. ANALYSES OF PROTEIN DYNAMICS ..................................................... 50
    4.3.1. Intrinsic fluorescence spectroscopy ....................................................... 51
    4.3.2. Extrinsic fluorescence spectroscopy ..................................................... 53
    4.3.3. Computational methods ....................................................................... 54

5. GENERAL DISCUSSION .................................................................................. 55
  5.1. GK PROTEIN STABILITY .......................................................................... 55
  5.2. THE MULTIPHASIC GLOBAL CONFORMATIONAL TRANSITION AND
       PERTURBATIONS OF CONFORMATIONAL EQUILIBRIA ......................... 57
  5.3. ALLOSTERIC EFFECTORS OF hGK .......................................................... 62
    5.3.1. Glucokinase activators (GKAs) ............................................................ 63
    5.3.2. Glucokinase regulatory protein (GKR) ................................................ 64
5.3.3. The bifunctional enzyme (PFK2/FBPase2) ................................................................. 65
5.3.4. Polyubiquitin ........................................................................................................... 65

5.4. CATALYTIC MECHANISM ...................................................................................... 66
5.5. MOLECULAR MECHANISMS OF DISEASE ............................................................. 68

6. FUTURE PERSPECTIVES ............................................................................................ 70
REFERENCES .................................................................................................................. 72
PAPERS I-III


Scientific environment

This work was carried out at the Section for Pediatrics, Department of Clinical Medicine, University of Bergen, at the Department of Biomedicine, University of Bergen, and at the Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital. The financial support for this project was mainly obtained from Helse-Vest, but additional funding was provided by the University of Bergen, the Meltzer Foundation, the Norwegian Diabetes Association, the Research Council of Norway and the Novo Nordisk Foundation.
Acknowledgements

Many people have made valuable contributions to this thesis. First of all, I would like to express my gratitude to my main supervisor professor Pål R. Njølstad for giving me the opportunity to work with this interesting project, for sharing his expertise in diabetes genetics, and for his encouragement and support throughout these years. A special thanks goes to my co-supervisor professor emeritus Oddmund Søvik for his valuable help, guidance and for always taking interest in this work. My sincere gratitude goes to my co-supervisor professor emeritus Torgeir Flatmark for his invaluable contribution to this project, for sharing his great knowledge in biochemistry and enzymology, and for his encouragement and constant enthusiasm. Special thanks also to Lise Bjørkhaug Gundersen (co-supervisor) and Ingvild Aukrust for their friendship and contribution to an inspiring scientific environment, as part of the GK team. I have been privileged to work in the Bergen Diabetes Research Group, which consists of a number of positive and very skillful people. Thanks to all of you. Furthermore, I would like to thank Jørn Sagen for being a good colleague and friend over the years, to Bente B. Johansson for being such a nice office-mate and for her encouragement during the preparation of this thesis, Maria Negahdar for the late evening pep talks, and Christine Andersen for listening and for her good sense of humor. Also thanks to all the people at Forskningslabben and the Center for Medical Genetics and Molecular Medicine for creating such a friendly working environment. Moreover, thanks to all the co-authors for their contributions to the papers. In addition, I would like to thank Knut Teigen for reading the methodological section and for answering all my questions regarding MD simulations.

Sincere thanks go to my parents Guri and Sverre-Jan, for their encouragement and for always believing in me. I also thank my wonderful friends and my parents-in-law Jorunn & Ingmar for their support and interest in my work.

Finally, my warmest gratitude goes to my dear Roald for his love and constant support, and to Hannah and Sander for their encouragement, for inspiring me with their smiles and making me laugh. You’re the best!

Bergen, 2012 \hspace{1cm} \textit{Janne Molnes}
Abbreviations

ANS 8-anilino-1-naphthalenesulfonate
AdN adenine nucleotide
AMP-PNP adenosine-5’-(β,γ-imido)triphosphate
ATP adenosine-5’-triphosphate
CHI congenital hyperinsulinism of infancy
CD circular dichroism
DTT dithiothreitol
DUB deubiquitinating enzymes
EC_{50} half maximal effective concentration
GKA glucokinase activator
GKRP glucokinase regulatory protein
GSIS glucose-stimulated insulin secretion
G6P glucose-6-phosphate
GST glutathione-S-transferase
HbA1c glycosylated hemoglobin
H-bond hydrogen bond
hGK human glucokinase
HI hyperinsulinism of infancy
ITF intrinsic tryptophan spectroscopy
L-domain large domain
LIST ligand-induced slow transition
Lys lysine
[L]_{0.5} ligand concentration at half-maximal effect
MD molecular dynamics
MODY maturity-onset diabetes of the young
NDM neonatal diabetes mellitus
n_{H} hill coefficient
NMR nuclear magnetic resonance
NOS nitric oxide synthase
OGTT oral glucose tolerance test
PDB protein data bank
PFK2/FBPase2 phosphofructokinase-2/fructose-2,6-bisphosphatase
Phe phenylalanine
PHHI persistent hyperinsulinemic hypoglycemia of infancy
PNDM permanent neonatal diabetes mellitus
PQC protein quality control
PTM post-translational modification
RRL rabbit reticulocyte lysate
SAXS small angle x-ray scattering
S-domain small domain
SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis
[S]_{0.5} substrate concentration at half-maximal effect
T1D type 1 diabetes mellitus
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2D</td>
<td>type 2 diabetes mellitus</td>
</tr>
<tr>
<td>Trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Tyr</td>
<td>tyrosine</td>
</tr>
<tr>
<td>Ub</td>
<td>ubiquitin</td>
</tr>
<tr>
<td>UBA</td>
<td>Ub-association</td>
</tr>
<tr>
<td>UBD</td>
<td>ubiquitin-binding domain</td>
</tr>
<tr>
<td>UbL</td>
<td>Ub-like protein</td>
</tr>
<tr>
<td>UIM</td>
<td>ubiquitin-interacting motif</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
</tbody>
</table>
Abstract

Glucokinase (GK) functions as a glucose sensor in insulin-producing pancreatic β-cells and as a regulator of hepatic glycolysis, glycogen synthesis and gluconeogenesis. Its key role in glucose homeostasis is evidenced by naturally occurring GK gene mutations causing monogenic diabetes and hyperinsulinemic hypoglycemia and by the discovery of allosteric GK activators (GKA) that hold promise as new antidiabetic agents.

GK catalyzes the first step in glucose metabolism, i.e. the conversion of α-D-glucose to glucose-6-phosphate (G6P), using MgATP$^{2-}$ as the phosphoryl donor. Glucose activates GK on its binding to the active site by inducing a global conformational change. Using intrinsic tryptophan fluorescence (ITF) spectroscopy as a probe on the glucose-induced conformational change, we identified key residues in this process. The glucose-induced fluorescence increase was primarily determined by W99 and W167, and little affected by W257. Based on results from functional mutagenesis and structural dynamic analyses, we have proposed that three active site residues (N204, N231 and E256) in the L-domain function as primary contact residues for glucose binding to the super-open form. Moreover, local torsional stresses at N204 and D205 of the highly flexible connecting region II was important for the subsequent propagation of the conformational transition towards cleft closure.

No structural data have been available on ATP binding to the apoenzyme and how it possibly affects its conformation. Here, we provide the first experimental evidence for an equilibrium binding of ATP and its analogue AMP-PNP to the ligand-free enzyme. Moreover, ITF quenching analyses and molecular dynamics (MD) simulations indicated a significant conformational change upon nucleotide binding. This finding was supported by the protective effect of ATP on binding of the extrinsic fluorescence probe ANS and on limited proteolysis with trypsin. Furthermore, the modeled structure of the GK-ATP binary complex provided insight into the active site contact residues involved in the interaction with ATP.
The knowledge on covalent modifications of human GK (hGK) and their possible regulatory functions are limited, and the molecular and cellular mechanisms involved in its degradation/turnover are also poorly understood. Using the rabbit reticulocyte lysate (RRL) as an in vitro model system, we demonstrated that pancreatic β-cell (isoform 1) and liver (isoform 2) hGK are substrates for the ubiquitin-conjugating enzyme system, and that both isoforms are polyubiquitinated on at least two lysine residues. A putative ubiquitin interacting motif (UIM) site at the C-terminal end was identified by 3D structural analysis, and associated with polyubiquitination at one of the sites. Moreover, our results supported that poly/multiubiquitination of recombinant pancreatic hGK in vitro target the newly synthesized enzyme for proteasomal degradation. Interestingly, purified free pentaubiquitin chains were demonstrated to interact with and allosterically activate (~1.4-fold) recombinant hGK, assigned to their equilibrium binding to the UIM site. Both these ubiquitin-mediated processes represent potential physiological regulatory mechanisms of GK.
List of publications


Related articles not included in the thesis


1. Introduction

1.1. Background

Glucose is the primary source of energy in the cell and is essential for life. An adequate supply of glucose is for instance necessary for a normal function of the brain, and low blood glucose (hypoglycemia) is therefore associated with loss of consciousness, seizures and in the most severe cases - death. On the other side, chronic hyperglycemia, the key element of diabetes mellitus, is associated with dysfunction of organs like the cardiovascular system, kidneys and eyes. Normally, the blood glucose concentration varies within a narrow physiological range (4-8 mmol/l). This steady-state concentration is mediated by coordinated homeostatic mechanisms, involving the endocrine pancreas (insulin and glucagon), liver (glucose stores) and peripheral tissues (glucose stores and energy expenditure), as well as a balanced secretion of other hormonal effectors, for instance from the gut (incretins). The concept of a glucose sensor component in this homeostatic feedback loop originated already in the late 1960s [1]. Over the past three decades, the central role of glucokinase (GK) as a glucose sensor in the pancreatic β-cell and its impact on whole body glucose homeostasis has become increasingly evident and is today widely accepted (Table 1) [2-6]. GK plays a key role in glucose-stimulated insulin secretion in pancreatic β–cells and in the liver hepatocytes where it stimulates glucose uptake and glycogen synthesis [7-9]. In the 1990s and early 2000s, it was discovered that naturally occurring GK mutations can cause different forms of glycemic disorders. This new knowledge gave a considerable boost to the GK glucose sensor concept. The perception of the essential role of human GK (hGK) in glucose homeostasis culminated in 2003 with the discovery of a class of small synthetic organic compounds as potent allosteric activators of GK [10-13]. Recently, it has been demonstrated a potential application of these compounds in the treatment of type 2 diabetes mellitus (T2D) [9, 14]. In 2004, human liver GK was successfully crystallized in the unliganded and glucose-bound conformation [15]. The 3D structures represented a breakthrough in the research on GK and GK-linked glycemic disorders, opening up new and intriguing approaches to
Table 1. Historical milestones in GK research

<table>
<thead>
<tr>
<th>Year</th>
<th>Discovery</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>GK identified in rat liver</td>
<td>[16-19]</td>
</tr>
<tr>
<td>1968</td>
<td>GK identified in mouse pancreatic islets</td>
<td>[20]</td>
</tr>
<tr>
<td>1975/76</td>
<td>Sigmoidal glucose dependency</td>
<td>[21, 22]</td>
</tr>
<tr>
<td>1977/80</td>
<td>Mnemonic and slow transition mechanistic models of GK cooperativity</td>
<td>[21, 22]</td>
</tr>
<tr>
<td>1984/86</td>
<td>The GK glucose sensor paradigm</td>
<td>[2, 3]</td>
</tr>
<tr>
<td>1986</td>
<td>Differential regulation of GK activity in liver and pancreatic β-cells</td>
<td>[23]</td>
</tr>
<tr>
<td>1986</td>
<td>Detection of GK in human islets</td>
<td>[24]</td>
</tr>
<tr>
<td>1989</td>
<td>GKRPs identified</td>
<td>[25, 26]</td>
</tr>
<tr>
<td>1989/91</td>
<td>Cloning of rat and human liver GK cDNA</td>
<td>[27, 28]</td>
</tr>
<tr>
<td>1992</td>
<td>GK linkage to MODY</td>
<td>[29, 30]</td>
</tr>
<tr>
<td>1998</td>
<td>GCK-HI described for the first time</td>
<td>[31]</td>
</tr>
<tr>
<td>2001</td>
<td>GCK-PNDM described for the first time</td>
<td>[32]</td>
</tr>
<tr>
<td>2001/03</td>
<td>First reports on GKA</td>
<td>[10-13]</td>
</tr>
<tr>
<td>2004</td>
<td>Crystal structure of hGK solved and deposited to the PDB</td>
<td>[15]</td>
</tr>
<tr>
<td>2008/10</td>
<td>First reports on use of GKA in diabetic patients</td>
<td>[14, 33]</td>
</tr>
</tbody>
</table>

GKRPs, GK regulatory protein; GKA, GK activator; PDB, Protein Data Bank. The table is modified from [34], with kind permission from Springer Science and Business Media © 2011.

The functional characterization of the wild type (WT) enzyme, including its dynamic, catalytic and regulatory properties, and in particular to investigate disease-associated mutant forms. The 3D structures of hGK represent the fundamental basis of the present work.

1.2. Monogenic diabetes

Diabetes mellitus is a metabolic disorder of multiple etiologies. It is characterized by chronic hyperglycemia resulting from defects in insulin secretion, insulin action, or both [35]. Whereas, type 1 diabetes (T1D) and T2D, the two major forms of diabetes, are multi-factorial complex diseases caused by a combination of environmental, life-
style related and genetic risk factors, some rare forms of diabetes are monogenic, i.e. resulting from single-gene defects. About 1-3 % of all diabetes cases can be attributed to monogenic causes [36-38]. In recent years, there has been a substantial progress in defining the etiological genes for monogenic diabetes. So far, some 20 genes have been identified, involved in pancreas development/differentiation or normal β-cell physiology [39]. Identifying these genes has enabled a better understanding of the genetic and biochemical basis of β-cell function and, moreover, improved patient management in terms of precise diagnosis, appropriate treatment and prognostic information as well as genetic counseling and follow-up of family members [40]. Genetic testing for monogenic diabetes in Norway has so far enabled the diagnosis of the exact subtype in about 60% of the cases (Figure 1). Thus, there are still many cases in which a monogenic cause is suspected but a genetic diagnosis have not been made (MODYX), presumably due to the presence of mutations in as-yet-undetermined genes.

Monogenic diabetes is a clinically heterogeneous disease. It can be subdivided into two predominant forms, neonatal diabetes mellitus (NDM) and maturity-onset diabetes of the young (MODY). NDM is defined as insulin-requiring hyperglycemia that develops within the first 6-12 months of life [41] and occurs in ~1 of every 100,000 live births. The condition is associated with intrauterine growth retardation and low birth weight as a consequence of low fetal insulin levels [42]. NDM can follow one of two clinical courses, which differ in the duration of insulin dependence early in the disease. About 50-60% of the cases are transient (transient neonatal diabetes mellitus or TNDM (OMIM #601410)) and resolve within the first 18 months of life, but then frequently relapse as T2D later in life [43]. In the remaining cases of NDM, the condition is permanent (permanent neonatal diabetes mellitus or PNDM (OMIM #606176)) and require lifelong medical treatment. MODY is the most common form of monogenic diabetes. It is an autosomal, dominantly inherited disease that is characterized by an early age of onset (at least one affected family member with an onset before 25 years of age), non-ketotic diabetes mellitus and primary pancreatic β-cell dysfunction [44]. Hence, the majority of patients typically presents with
14

Figure 1. The relative prevalence of monogenic diabetes subgroups in Norway based on genetic analysis of 314 probands referred to the Norwegian MODY registry. Only the individuals that fulfill the conventional MODY/monogenic criteria are included. In the Norwegian MODY registry, GCK-MODY has a relative prevalence of approximately 15%. MODYX refers to the individuals that have tested negative for the known causes of monogenic diabetes. The data (unpublished) are based on numbers extracted from the Norwegian MODY Registry per September 2011.

hyperglycemia in childhood or adolescence and have a strong family history of diabetes.

1.3. Glucokinase glycemic diseases

GK was discovered almost 50 years ago, first as an enzyme of rat liver (1963) [16-19] and subsequently in mouse pancreatic islets (1968) [20]. With its central role in insulin secretion, it was early an obvious candidate gene for diabetes [45], and the GK-encoding gene (GCK) was in fact the first MODY gene to be identified [29, 30, 46-
Alterations in the catalytic activity of GK are associated with abnormal glycemia. Three different glycemic diseases, due to naturally occurring GCK mutations, have been described; GCK-MODY, GCK-PNDM and congenital hyperinsulinism. Hence, GK-associated diseases constitute a broad spectrum of clinical phenotypes, accounted for by the different nature of the causative mutations (summarized in Figure 2). Furthermore, common genetic variants of the GCK gene are associated with elevated fasting plasma glucose levels and T2D risk [49-51]. Altogether, these facts emphasize the significant role of GK in maintaining normoglycemia.

Figure 2. Schematic representation of the spectrum of clinical phenotypes associated with mutations in the GCK gene. See text for details.
1.3.1. GCK-MODY

Inactivating heterozygous mutations in \textit{GCK} that impair either the function or level of expression of the enzyme cause a subtype of MODY, formerly known as MODY2 and today more preferably designated GCK-MODY or GCK-monogenic diabetes (OMIM #125851) [29, 30].

\textbf{Pathophysiology}

The pathophysiological mechanism of the hyperglycemia in GCK-MODY patients is primarily related to a \(\beta\)-cell dysfunction characterized by reduced GK activity and an increased threshold for glucose-stimulated insulin secretion (GSIS) [52, 53]. The glucose threshold value is generally shifted from \(\sim 5\) to \(7\) mmol/l (mM), consistent with a defect in glucose sensing by pancreatic \(\beta\)-cells (see section 1.4). In normal physiology, the liver contributes to the control of blood-glucose homeostasis by net glucose uptake and storage after meals when postprandial blood glucose levels are high, and by net glucose production in fasted state [54, 55]. GK is central in this process, serving as a metabolic switch to shift hepatic carbohydrate metabolism between fed and fasting states. The hepatocytes of the liver contain \(\sim 99\%\) of the body’s total GK content [5]. Moreover, GK provides \(\sim 95\%\) of the glucose-phosphorylating (hexokinase) activity in these cells. Clinical investigations in patients with GCK-MODY have revealed disorders of liver metabolism, in addition to the altered \(\beta\)-cell function, such as reduced net hepatic glycogen synthesis and abnormal regulation of hepatic glucose output [56, 57]. Extensive studies on tissue-specific and global \textit{GCK} knock-out mice, as well as models of GCK overexpression, have confirmed the findings in humans [58-64]. Furthermore, the mechanisms of action and therapeutic applications of small molecule GK activators (GKAs), also support a role of liver GK in the pathophysiology of this disease (see section 1.7.3.). Mice lacking hepatic GK are mildly hyperglycemic. In contrast, a \(\beta\)-cell specific knock-out of GK results in severe hyperglycemia and death shortly after birth [27, 60]. Taken together, even though the reduced GK activity in the \(\beta\)-cell and the subsequent impairment in GSIS seem to be the dominant physiological effects, abnormalities in hepatic glucose
metabolism seems to be a contributing factor in the pathogenesis of GCK-MODY hyperglycemia.

**Clinical features**

The clinical phenotype of GCK-MODY is very homogenous, despite a large number of naturally occurring patient mutations with varying effects on enzyme kinetics (see section 1.5) [65]. This has, at least in part, been explained by a compensatory up-regulation of the remaining WT allele [66]. GCK-MODY is distinct from the other MODY subtypes. The disease is characterized by (i) mildly elevated fasting hyperglycemia of 5.5-8 mM (from a normal basal concentration of 4-5 mM), (ii) glycosylated hemoglobin (HbA1c) that is just above the upper limit of normal and rarely exceeds 7.5%, and (iii) a small 2-hour plasma glucose increment (≤4.6 mM) during an oral glucose tolerance test (OGTT) [67, 68]. Accordingly, many GCK mutation carriers do not develop manifest diabetes, but are diagnosed with impaired glucose tolerance or increased fasting plasma glucose. The fasting hyperglycemia is present from birth [69] and deteriorates only slightly with age. Due to the mild hyperglycemia patients are often asymptomatic and the hyperglycemia can easily be overlooked during childhood and young adulthood.

**Diagnosis and treatment**

Due to the mild nature of the disease, the diagnosis of GCK-MODY is most often made during routine testing, for instance in pregnancy. GCK mutations have been found in 2-5% of women diagnosed with gestational diabetes [70-73]. Except during pregnancy, GCK-MODY patients rarely require pharmacological treatment. The risk of developing diabetes-related chronic complications like cardiovascular disease, kidney failure, retinopathy and neuropathy is low [74, 75].

**Prevalence**

It is difficult to assess the true prevalence of GCK-MODY. In UK, about 2% of Caucasian pregnant women were diagnosed with gestational diabetes, and of these about 2-5% had a mutation in the GCK gene, suggesting a population prevalence of
GCK-MODY of 0.04-0.10% [72]. As the disease has a mild expression, patients are frequently not diagnosed and, hence, GCK-MODY is probably underestimated. Moreover, the estimated prevalence of the disease in European Caucasian MODY families varies considerably. Whereas mutations in the GCK gene seems to be the second most common cause of MODY in Northern Europe, including Norway (Figure 1) [75-79], it seems to be the most prevalent MODY subtype in countries of Southern Europe like France, Italy and Spain [48, 80-85]. This discrepancy is probably caused not only by the variations in the genetic background, but also, at least in part, by differences in the recruitment criteria [65]. To assess the exact prevalence of this disease, large-scale population studies are required.

1.3.2. GCK-PNDM

Mutations in the GCK gene are an infrequent cause of PNDM in the European population [86, 87]. GCK-PNDM is caused by complete deficiency of GK activity due to homozygous or compound heterozygous inactivating mutations in GCK [65, 88]. The condition should be considered primarily in cases of neonatal diabetes presenting with a family history of mild hyperglycemia or GCK-MODY, especially when consanguinity is suspected. The first case of GCK-PNDM was described in 2001 [32], and later 11 new cases have been reported [89-93] in addition to at least one unpublished case (Eltonbary et al, unpublished). Almost all cases are attributed to consanguinity. Clinically, the disease is characterized by severe intrauterine growth retardation, low birth weight, pronounced hyperglycemia and ketoacidosis within the first days of life [32, 89]. This is a life-threatening disease that requires insulin treatment. However, recent studies have presented evidence for the potential use of sulfonylureas (in addition to insulin) to augment improved glycemic control [92, 93]. It was further demonstrated that use of sulphonylureas stimulated endogenous insulin secretion from β-cells (albeit at very low levels) [92], and the combined treatment with sulfonylurea allowed management on a reduced insulin dose. It has therefore been speculated that PNDM cases due to mild GCK mutations may cause a less severe
phenotype that respond better to sulfonylureas [90]. The mechanism of sulfonylurea action in these patients is unclear.

1.3.3. GCK-HI

Hyperinsulinism of infancy (HI), with an estimated incidence of 1 in 40,000-50,000 live births [94, 95], is the most common cause of persistent hypoglycemia in early infancy [96]. The condition is also known as congenital hyperinsulinism of infancy (CHI), persistent hyperinsulinemic hypoglycemia of infancy (PHHI) and nesidioblastosis. HI is characterized by recurrent episodes of hyperinsulinemic hypoglycemia due to inappropriate secretion of insulin [94]. This is a life-threatening disorder where early diagnosis and treatment are essential to prevent potentially acute and severe complications like seizures, coma and, at worst, death. Several genetic causes of HI have been identified [97]. In 1998, activating heterozygous mutations in GCK were recognized as a cause of neonatal hypoglycemia [31]. GCK-HI (OMIM #602485) is a rare subtype, estimated to ~1.2% of all HI cases [98]. At least 15 activating GCK mutations have been reported to date [65, 99-101]. The primary mechanism in GCK-HI is a lowered threshold for GSIS, leading to a failure to suppress insulin secretion in the presence of hypoglycemia. The clinical phenotype covers a broad spectrum, even within the same family harboring the same mutation, both with regard to age of onset (from first day of life to adulthood), severity of the hyperinsulinism (from asymptomatic to unconsciousness and seizures), clinical course (a few cases have progressed to insulin resistance), as well as responsiveness to pharmacological treatment [102-104]. Even though a few cases with severe, medically unresponsive GCK-HI have been reported, most patients present with mild fasting hypoglycemia responsive to the sulfonylurea receptor agonist diazoxide. In two of the severe (medical unresponsive) cases of GCK-HI, pancreatic histology revealed abnormally large islets with some β-cells containing large nuclei [100, 105], which could be due to increased intracellular glucose flux and β-cell proliferation [100]. Thus, a complex mechanism for GK regulation is implicated in GCK-HI, and, to some
extent, the severity of hypoglycemia seems to correlate with the severity of the mutation/enzyme defect [65, 104].

1.4. The role of pancreatic β-cell GK in glucose-stimulated insulin secretion and glycemic disease

The body’s sole glucose-lowering hormone, insulin, is produced, stored and secreted by the pancreatic β-cells of the islets of Langerhans [106]. To maintain normoglycemia, these cells correlate their insulin secretion with changes in plasma glucose concentration. A postprandial fall in blood glucose leads to a decrease in insulin secretion. However, the β-cells maintain a basal insulin secretion which allows the cells of the body to utilize glucose also between meals and during night. Furthermore, the β-cells respond quickly (within minutes) to postprandial spikes in blood glucose by secreting increased amounts of insulin. Insulin subsequently acts to reduce blood glucose levels by stimulating glucose uptake in insulin-sensitive cells of liver, muscle and fat, and by inhibiting glucose production in the liver [107].

In healthy individuals, the physiological threshold for GSIS in the pancreatic β-cell is a result of a finely tuned coordinated interaction between glucose transport, oxidative glucose metabolism, the K\textsubscript{ATP}-channel and the Ca\textsuperscript{2+}-channel (Figure 3). GK is a key component in this glucose-sensing machinery and is frequently referred to as the glucose sensor of the pancreatic β-cell [108]. The activity of GK is directly coupled to the blood glucose concentration, and rises and falls accordingly. Initially, glucose enters the β-cells via the GLUT family of transporters [110]. The elevated intracellular concentration of glucose, in turn, activates GK, which catalyzes the phosphorylation of glucose to G6P. Glycolysis and further energy production in the Krebs cycle and by oxidative phosphorylation in mitochondria generate ATP. The resulting increase in ATP/MgADP ratio stimulates the closure of the ATP-sensitive potassium (K\textsubscript{ATP}) channels, depolarization of the plasma membrane, subsequent opening of the L-type
Figure 3. A schematic model of the pancreatic β-cell illustrating the main sequence of cellular events and the role of GK in glucose-stimulated insulin secretion (GSIS). Blood glucose levels are the major determinant of the rate of insulin secretion. As the plasma glucose level rises above basal levels (~5 mM), glucose is transported into the β-cell via a GLUT carrier. Increased intracellular glucose levels activate GK which, in turn, catalyzes the phosphorylation of glucose to G6P. Metabolism of glucose leads to an increase in the cellular ATP/MgADP ratio, which triggers insulin secretion. An inactivating mutation in the GCK gene, the cause of GCK-MODY, leads to an impairment of GK enzymatic activity or protein stability, resulting in decreased glucose phosphorylation and an increased threshold for GSIS, causing mildly elevated fasting blood glucose levels (6-8 mM). In the case of GCK-PNDM, both alleles are affected, resulting in severe insulin deficiency of infancy due to a complete inactivation of GK activity. In contrast, in GCK-HI, activating mutations that increase the catalytic efficiency of GK cause a reduction in the threshold for GSIS (in some cases to as low as 1.5 mM). Thus, in both GCK-MODY and GCK-HI the GSIS is still intact, however induced from a shifted glucose threshold [109]. The figure is reprinted from [44], with permission from Massachusetts Medical Society © 2001.
voltage-dependent calcium (Ca\(^{2+}\)) channels, influx of extracellular calcium and mobilization of intracellular calcium stores. The following rise in intracellular calcium triggers the exocytosis of insulin from granules docked at the plasma membrane (Figure 3) [111]. In humans, the physiological β-cell glucose threshold for GSIS is maintained close to 5 mM. As glucose phosphorylation is a key point of control for glycolytic flux in the β-cell [112, 113], small changes in GK activity or stability can be physiologically significant since it will directly affect the threshold for GSIS. This is the mechanism by which many naturally occurring mutations in GCK cause glycemic disease in humans (see figure 3 legends). Mathematical modeling of the effect of naturally occurring GCK mutations on GSIS has indeed demonstrated a tight relationship between GK activity and the threshold for GSIS [114].

**1.5. Naturally occurring GCK mutations and the mechanism of their associated disease**

The strongest evidence for the GK glucose-sensor concept derives from the functional consequences of GK gene mutations in humans. From the first identified mutations in 1992 until today, > 630 GCK mutations in more than 1440 families have been reported [65]. Missense, nonsense, small deletion/insertion (frameshift) and splice site mutations have been identified, and are distributed throughout the gene [65]. There are no mutational “hot spots” in the GCK gene. Recently, partial and whole gene deletions have been identified, but are probably a rare cause of GCK-MODY [115, 116].

Of the ~630 GCK mutations described, around 65% are missense mutations, of which most (~97%) are inactivating mutations causing hyperglycemia [65]. Many of the missense mutations have only been described in a single family [65]. In many cases co-segregation has not been established. Despite the importance of ascribing pathogenicity to these variants, less than 20% have been functionally characterized [65, 114], most of them as E. coli expressed recombinant pancreatic hGK-glutathione-S-transferase (GST) fusion proteins. The vast majority of the functionally characterized GK mutants demonstrate alteration in one or more enzyme kinetic
parameters \((k_{\text{cat}}, \text{glucose } S_{0.5}, n_{H}, K_{m} \text{ for MgATP}^{2-})\). For inactivating mutations (MODY, PNDM), the mutant proteins show reduced catalytic efficiency due to decreased catalytic activity \((k_{\text{cat}})\) and/or increased glucose \(S_{0.5}\). The Hill coefficient \((n_{H})\) and the MgATP\(^{2-}\) \(K_{m}\) can be increased or decreased [65]. For GK-activating mutations (HI), the opposite effect on the kinetic parameters is seen, with an overall increase in catalytic efficiency, mainly due to a reduction in glucose \(S_{0.5}\) and/or increased \(k_{\text{cat}}\). The Hill coefficient is usually decreased [65]. Furthermore, it has been shown that in some cases other mutational mechanisms are in play. For instance, some GK mutants demonstrate altered affinity for known interaction partners (see section 1.9.) or reduced thermostability [65]. Interestingly, careful kinetic analyses of recombinant GK enzymes have uncovered several missense variants not detected in healthy individuals, that demonstrate normal or near normal enzyme kinetic parameters as well as normal interaction with known regulatory proteins. These are potentially interesting since novel mutational mechanisms may be involved. In these cases, co-segregation with fasting hyperglycemia in the family will strengthen the interpretation of the variants as pathogenic. The importance of combining functional studies and segregation analyses has recently been emphasized [117].

1.6. GK - a member of the hexokinase family of enzymes with unique properties

GK belongs to the hexokinase family of enzymes as one of four glucose-phosphorylating isozymes (ATP:D-hexose 6-phosphotransferase (EC 2.7.1.1)) found in mammalian tissues, designated hexokinase I-IV [118] or A-D [119]. The hexokinases represent key metabolic enzymes, catalyzing the MgATP-dependent phosphorylation of glucose to G6P, which is the first and rate-limiting step in glycolysis (Figure 4).

*Glucokinase* (hexokinase IV or D) has derived its name from its preference for \(\alpha\)-D-glucose as a substrate under physiological conditions. The enzyme is distinguished
Figure 4. The reaction catalyzed by GK. GK catalyzes the phosphorylation of α-D-glucose (on carbon 6*) to form G6P, the entry point of glucose into glycolysis. ATP participates as the phosphoryl donor (cosubstrate) in the reaction in a form complexed to Mg$^{2+}$ (MgATP$^{2-}$).

from the other mammalian hexokinases in molecular mass, structure, kinetic properties, tissue distribution and physiological function [120, 121].

1.6.1. Kinetic properties of GK

GK operates as a monomer and has a molecular mass of about 52 kDa, as opposed to ~100 kDa for hexokinase I-III. It is the principal enzyme responsible for glucose phosphorylation in hepatocytes and pancreatic β-cells [24, 122, 123], the tissues important for the regulation and clearance of blood glucose. In these cells, GK has unique catalytic and kinetic properties to serve this physiological function [6, 124].

GK is kinetically distinct from hexokinase I-III: (i) GK has a low affinity for glucose, i.e. a glucose concentration at half-maximal velocity ($S_{0.5}$) of ~8 mM, (ii) GK is not inhibited by the product (G6P), (iii) GK displays positive cooperativity with respect to glucose ($n_H$ of ~1.8), and (iv) GK is physiologically saturated with its cosubstrate MgATP ($K_m$ of 0.15-0.4 mM). The enzyme displays classical hyperbolic Michaelis-Menten kinetics with respect to the concentration of MgATP [120].

Liver and β-cell GK share the same unique kinetic properties [125]. Most importantly, the low affinity for glucose and the positive cooperativity with the substrate (see
section 1.8) allow GK to maintain a high sensitivity and responsiveness to variations in the glucose level within the physiological concentration range (4-10 mM).

1.6.2. Gene structure and tissue-specific gene regulation

The tissue distribution of GK and the ‘low-\(K_m\)’ (20-130 \(\mu\)M) hexokinases are different. Hexokinases I-III are involved in energy metabolism in most tissues, as ubiquitous housekeeping enzymes. In contrast, GK is expressed with stringent tissue specificity in accordance with its highly specialized physiological functions.

The \(GCK\) gene is located on chromosome 7p15.3-15.1 and consists of 12 exons that span a region of 45,168 bps. \(GCK\) transcription is regulated by two alternative tissue specific promoters which enable differential regulation of GK in \(\beta\)-cells and liver in a manner consistent with the different function of these two tissues. Whereas the downstream promoter operates exclusively in the hepatocytes, the upstream, neuroendocrine promoter is functional in pancreatic \(\beta\)-cells, but also in pancreatic \(\alpha\)-cells and some rare glucose-sensing neuroendocrine cells of the brain (hypothalamus and anterior pituitary gland) and enteroendocrine cells of the gut [28, 126-132]. Thus, there is a complex network of GK-expressing glucose-sensing cells in the body, important for maintenance of whole body glucose homeostasis.

Three different GK transcripts are produced by alternative splicing, encoding isoform 1 (neuroendocrine) and isoform 2 and 3 (liver). The isoforms differ in the first 14-16 N-terminal amino acids which constitute exon 1 (a, b, c) of the protein. Exon 1a is expressed in the pancreatic \(\beta\)-cell whereas exons 1b and 1c are expressed in the liver [133, 134]. Exons 2-10 are identical in the three isoforms.
1.7. The 3D structure of hGK and the glucose-induced conformational change

GK is a monomeric enzyme composed of two domains; a small (S) and a large (L) domain, linked by three connecting regions (Figure 5). The active site resides in a cleft between the two domains [15, 135-137]. GK is highly conserved with 95% sequence similarity between human and rat (HomoloGene). Before the crystal structure of hGK was solved, structural predictions were based on homology modeling. The initial models were made using the X-ray crystal structure of yeast hexokinase B [135, 136, 138, 139]. However, the accuracy of these models was limited by the rather low amino acid sequence similarity (~30%) to hGK. Models of human brain hexokinase I (sequence similarity of 54%) provided the basis for more accurate structural predictions of GK with its substrates, and have been used to locate some of the disease-causing GK mutations and to interpret their functional implications [137, 140-147]. Due to its high structural flexibility, GK has proved difficult to crystallize. In 2004, Kamata et al overcame this problem by deleting the 11-15 N-terminal residues of hGK. These deletion mutants demonstrated similar \textit{in vitro} kinetic properties as WT hGK [15]. Two high resolution crystal structures of liver hGK were obtained, one in its unliganded apo form (3.4 Å, PDB entry 1v4t) and the other in a complex with glucose and a synthetic GK activator compound (N-thiazol-2-yl-2-amino-4-fluoro-5-(1-methylimidazol-2-yl)thiobenzamide) (2.3 Å, PDB entry 1v4s) [15]. It should be noted that residues 157–179 were unassigned in the electron density map of the unliganded hGK due to a disorder of this region. According to these structures [15], the sequences 1–64 and 206–439 constitute the L-domain, the sequences 72–201 and 445–465 constitute the S-domain, and the sequences 65–71, 202–205 and 440–444 represent the connecting regions I-III, respectively.

Structural analyses of the two crystallized conformations of hGK confirmed previous biochemical and biophysical studies [150, 151], providing structural evidence that hGK undergoes a reversible, large-scale global conformational transition (isomerization) upon binding glucose, from a ‘super-open’ (inactive) conformation to a ‘closed’ (active) conformation (Figure 5). The spatial arrangement of the S- and L-
domains in the apo form of hGK is different (more extended) from the open form of hexokinase I, and, hence, the structure of apo GK has been referred to as the ‘super-open’ form [15]. Although the structure and spatial relationship of the closed conformation of GK and the C-terminal half of the closed form of hexokinase I are very similar, GK undergoes a much larger domain motion. This is in part due to a more flexible structure of the interconnecting region I in GK, which facilitates the movement of the C-terminal helix (see section 1.7.2) [15]. The glucose-induced conformational rearrangement of GK involves a large-scale hinge bending/sliding
motion, which requires the breakage and reformation of numerous interactions. The final result is a reorientation of the S-domain, involving a \( \sim 104^\circ \) rotation (as compared to \( \sim 12^\circ \) for hexokinase I) toward the L-domain, which remains largely stationary (Figure 5) [15]. Hence, the enzyme adopts a more compact (‘closed’) structure involving a 96% closure of the active site cleft.

1.7.1. The active site
The active site of GK resides in a channel-shaped cleft between the S- and L-domain, and provides a favorable microenvironment for the phosphorylation of the substrate \( \alpha \)-D-glucose. The large-scale domain movement induced by binding of glucose (Figure 5), closes the active site cleft and creates the stereochemical environment for binding of the cosubstrate MgATP. In the ternary GK-glucose-MgATP catalytic complex, the ATP \( \gamma \)-phosphate is in close spatial proximity to the 6-hydroxyl group of glucose [152].

Glucose-binding site
In the super-open conformation of GK, the glucose-binding site is exposed to the solvent and important residues for substrate interaction and catalytic activity are missing from the site (Figure 6) [15]. Hence, even if glucose is able to bind at this site with low affinity, the super-open state of the enzyme is considered catalytically inactive. Glucose binding induces closure of the interdomain cleft and displacement of the flexible S151-C181 loop structure from the protein surface (Paper III). The loop, which is part of the S-domain, closes over the incomplete active site as a lid (Figure 5), protecting the active site from solvent. This active site loop forms one rim of the glucose-binding site, and contributes to the catalytic environment in the hGK ternary complex.

The residues that form hydrogen (H) bonds with the oxygen atoms of glucose have been defined crystallographically [15], and consist of T168 and K169 in the S-domain and N204, D205, E256 and E290 in the L-domain (Figure 6). The residues N204 and
Figure 6. Structure of the glucose-binding site. (A) The glucose-binding site in the closed, binary GK-glucose complex. E256 and E290 of the L-domain (blue stick model), T168 and K169 of the S-domain (cyan stick model), and N204 and D205 of connecting region II (green stick model) form hydrogen bonds with glucose (pink stick model). (B) The glucose-binding site in the super-open, apo form of hGK is exposed to solvent. The residues T168 and K169 of the S-domain (see A) are part of a disordered structure, and are not assigned in the electron density map of the super-open structure (1v4t). The color coding is as described in (A). The figures are reprinted from [15], with permission from Elsevier © 2004.

D205 are part of connecting region II. Mutational analyses have confirmed the importance of several of these amino acids [138, 146, 153]. In general, mutations in glucose binding residues have a dramatic effect on the catalytic activity [137]. Most of these residues have been found mutated in GCK-MODY patients, emphasizing their critical role for enzyme function [65].

It should be mentioned that in 2004 the structure of hGK with glucose alone (i.e. without activator) was not solved and, hence, it was not clear whether and to what extent the active conformation of GK as well as the structure and interactions around the glucose-binding site are altered by binding of the activator compound.

ATP-binding site

Yeast hexokinase and hGK have a low ATPase activity representing 1/10000 and 1/2000 of its kinase activity, respectively [136]. The higher ATPase activity of GK
indicated that water molecules are more accessible to the γ-phosphate of ATP in the absence of glucose [136]. Interestingly, an ATPase domain with the same basic structural fold is shared between hexokinases from bacteria, yeast and plants, to humans and other vertebrates, as well as other less related protein families as actin and the hsp70 family of heat shock proteins [154, 155]. This actin fold consists of five well conserved sequence motifs involved in the interaction with ATP; phosphate 1, connect 1, phosphate 2, adenosine and connect 2 (Figure 7) [154].

At the time of conducting this work no crystal structure of a ternary GK-glucose-ATP complex was available and the most accurate structural predictions on the ATP binding site were based on homology modeling using human hexokinase type I as template [137, 140, 141, 143, 146, 147]. Based on these models and ATPase domain sequence conservation, the residues T82, R85, T228, K296, S336, S411 and K414 were predicted to form H-bond interactions with the ribose moiety or phosphate oxygens of ATP, placing the nucleotide in the correct orientation and conformation to interact with glucose. GK mutations that perturb or eliminate the interaction with ATP generally lead to enzymes with reduced affinity for ATP and severely impaired catalytic activity [137, 146]. Most of these critical contact residues have been found mutated in GCK-MODY patients [65].

1.7.2. The C-terminal α-helix

The C-terminal α-helix (residues 447-460 (helix 17) in the super-open conformation and residues 443-461 (helix 19) in the closed conformation) plays a crucial role in the global conformational transition of hGK [15, 157]. Upon domain closure, the C-terminal helix (referred to as ‘α13’ in the literature), which is located adjacent to connecting region III (residues 440-444), moves from a solvent exposed interdomain position in the super-open form to a sequestered internal location within the S-domain in the closed conformation [15, 157, 158]. The helix contains several small residues (glycine and alanine) which assist with the movement around the connecting region. In both conformations, the C-terminal helix specifically interacts with helix 6 (α5) of the
Figure 7. The five sequence ATP binding motifs of hGK. Multiple alignment of hGK isoform 1 and yeast hexokinases isozymes PI and PII was created using ClustalX v.2 [156] and the reference sequences P35557, P04806 and P04807 (UniProt), respectively. (*) indicates positions which have a single, fully conserved residue; (:) indicates conservation between groups of strongly similar properties, and (.) indicates conservation between groups of weakly similar properties. Yellow boxes denote the most evolutionary conserved residues. Residue numbers and secondary structures of the motifs are given (arrow, beta strand; cylinder, helix; half circle, helix-turn or bend). Annotations of the secondary structure elements are derived from the Protein Data Bank using the PDB entry 1v4t.
L-domain [15]. During the conformational transition, changes in the domain interface take place, with the relative orientation of the helices changing from a parallel to a perpendicular orientation with a subsequent change in main residue interactions [15]. Recent MD simulations support a model in which the ‘release’ of the C-terminal helix from the S-domain is the final step in the slow conformational transition from the closed to the super-open state [158]. Moreover, evidence has been presented that this substrate-induced repositioning of the C-terminal helix is essential for GK kinetic cooperativity [157]. Interestingly, several GK activating mutations map to this helix or neighboring interacting structures [15, 65, 157-160].

1.7.3. The allosteric activator site and pharmacological GK activator drugs

The crystal structure of hGK complexed with glucose and a small molecule synthetic GK activator (GKA), revealed a hydrophobic binding pocket for the GKA molecule at the domain interface, 20 Å remote from the glucose and ATP binding sites [15]. The surface of the activator binding site is formed by the flexible loop of connecting region I (V62-G72), helix 6 of the L-domain (D205-Y215) and the C-terminal helix (E443-C461) of the S-domain [15, 161, 162]. The search for small molecules that were capable of activating GK began already in the 1990s [5, 13], and today a wide range of GK activator compounds have been discovered as a result of intensive medicinal chemistry efforts. The GKAAs have diverse chemical structures, but display a similar pharmacophore [34, 163]. For some of the activators the binding site and its contact residues have been defined crystallographically [15, 162, 164-167]. In the unliganded (glucose-free) form of hGK, the loop of connecting region I is occluded and the C-terminal helix is released from the small domain, thus, the allosteric pocket is distorted or absent.

GKAs bind close to one of the hinge regions involved in interdomain communication and propagation of conformational transitions. By binding to this regulatory site GKAs are able to allosterically enhance GK activity [13]. In vitro, these molecules increase the catalytic activity of GK, predominantly by lowering the $[S]_{0.5}$ value for glucose (5-10 fold). Many GKAs increase (up to twofold) the turnover rate ($k_{cat}$) of the enzyme,
but in a few instances the $k_{cat}$ is slightly decreased. The Hill coefficient ($n_H$) is lowered to a varying extent, and in some cases approaching unity [34]. The potential application of GKAs in the treatment of T2D is currently under investigation [9, 34, 168, 169]. So far, the results from clinical trials (and animal studies) have demonstrated a dual mechanism of the activator molecules by potentially acting on both the β-cell, by improving GSIS, and the liver, by reducing uncontrolled glucose output and restoring postprandial glucose uptake and glycogen synthesis [34]. Some GKAs are believed to carry a risk of inducing hypoglycemia. However, perhaps of greatest concern are the potential long-term adverse effects of GKAs. As 99% of the body’s GK is present in the liver, it is possible that enhanced GK activity may lead to increased de novo lipogenesis and plasma triglyceride levels [9, 170, 171]. Moreover, the potential effect of GKAs on the function of other GK-expressing cells and organs must be assessed.

The existence of a physiological, endogenous GKA molecule has been postulated. Interestingly, the functional effects of synthetic GKAs mimic the kinetic effects of naturally occurring activating mutations causing HI [31, 65, 98, Barbetti, 2009 #177, 101, 103, 105, 144, 172]. In contrast to inactivating $GCK$ mutations, which are distributed throughout the 3D structure of hGK, almost all the HI-associated mutations colocalize to a common region in the 3D structure, which coincides with the binding site of the GKAs [15]. The overlap of these sites may suggest that common molecular mechanisms of activation are involved.

1.7.4. **Cysteine residues at the active site of GK**

GK contains 12 cysteine residues, including five that appear as a conserved ring motif close to the active centre [135, 153]. The close vicinity and spatial localization of the Cys residues, together with the high structural flexibility of GK, makes the enzyme very vulnerable towards oxidative formation of intrachain disulfide bridges, and a concomitant inactivation of enzymatic activity [173-175]. GK is one of the most sensitive thiol enzymes in the pancreatic β-cell, demonstrating high sensitivity towards
SH group oxidizing compounds such as the glucose analogue alloxan [176-179]. Glucose protects the enzyme against alloxan-induced inhibition [175]. The inhibitory effect of alloxan on GK catalytic activity can be prevented and frequently reversed by sulfhydryl group reducing agents, such as dithiothreitol (DTT), and the presence of thiol agents has demonstrated to be mandatory during purification of the enzyme [173]. In the absence of thiol agents, GK experienced a constant decay in activity [173]. When recombinant GK proteins are subjected to non-reducing SDS-PAGE, a double-band pattern can be observed, corresponding to the 52 kDa native GK protein and a faster migrating, more compact, oxidized form of ~49 kDa [176, 180]. The intensity of the 49 kDa band increased significantly upon alloxan treatment [181]. Moreover, freshly purified recombinant GK also displays a characteristic electrophoretic double-band pattern, indicating that the enzyme naturally exists in at least two different conformations [181], dependent upon the redox status of the sulfhydryl groups. The oxidized form represents ~1-2% of the total protein (Molnes et al, unpublished observations). No disulfide bonds have been observed in the crystal structure of pancreatic hGK [152]. In summary, these aspects emphasize that GK, with its high susceptibility to sulfhydryl oxidation at low glucose concentrations, is very vulnerable towards oxidative stress, especially in the pancreatic β-cell having low enzymatic antioxidative defense mechanisms [179].

1.8. Kinetic models of positive cooperativity

The positive kinetic cooperativity of GK with respect to glucose was discovered 35 years ago [182, 183]. This property is unique since GK functions exclusively as a monomer [184] with a single substrate binding site [15]. Two models have appeared particularly appropriate to explain the observed cooperative kinetics of GK; the mnemonic and ligand-induced slow transition (LIST) models [21, 185-191]. Both models attribute the positive cooperativity to a slow, reversible, glucose-dependent conformational transition between a low affinity and a high affinity form of GK, and the conversion between the two conformations is slower than the catalytic cycle ($k_{cat}$). Accordingly, the existence of different conformational states of the enzyme, combined
with their failure to reach conformational equilibria during the course of catalysis, is the fundamental basis for the sigmoidicity in the reaction rate profile of GK [22, 192, 193].

The mnemonic mechanism [186] is based on the concept that the enzyme retains a “memory” of the active conformation for some time after product conversion and release. According to this model, the unliganded enzyme alternates between two distinct conformational species, a low affinity state (E*) and a high-affinity state (E) (Figure 8A). Only the low-affinity conformation is postulated to exist in the absence of glucose. Both forms can bind glucose to form the same binary (E·G) complex, which reacts rapidly with MgATP to generate the ternary complex. Catalysis occurs from the high-affinity E·G state. If glucose is abundant, the high-affinity form (E) can rapidly go through another catalytic cycle, whereas if glucose levels are low, the enzyme slowly relaxes back to the low-affinity state (E*). Thus, the mnemonic model postulates the existence of a thermodynamically favored enzyme conformation in the absence of glucose and involves one catalytically active enzyme species [187]. In contrast, the LIST mechanism postulates the existence of two catalytic cycles. The model assumes that a pre-existing equilibrium exists between the two conformational states (E* and E) in the absence of glucose, and that each can accomplish a separate catalytic cycle, involving different enzyme states (E·G or E*·G) and kinetics (Figure 8B) [21, 189, 194]. The overall steady-state rate is the sum of the rates for the two cycles. Furthermore, the equilibrium between the low-affinity (E*) vs. the high-affinity (E) form is governed by the concentration of glucose. Thus, as the glucose level increases the equilibrium is shifted towards the high-affinity form. Evidence in support for both mechanistic models can be found within the experimental data collected on GK over the last 35 years [21, 120, 150, 151, 182, 187, 188, 194, 195].

Initial support for the existence of a slow glucose-induced isomerization involving different conformational states of GK came from various observations, e.g. the lag-phase (~5-10 min) observed in kinetic experiments when rat liver GK was stored/preincubated at glucose concentrations lower than in the assay [194], and the
Figure 8. Alternative mechanisms proposed to explain the positive cooperativity of monomeric GK. (A) The mnemonic and (B) LIST models of positive cooperativity. E and E* represent two conformations of the enzyme. Red arrows represent the slow conformational transition (isomerization) between conformational states. (C) The “Kamata” model based on the crystallographic structures and analysis of domain movement [15]. See the text for details. The figures are reprinted from [15, 193], with permission from Elsevier © 2004/2012.

A time-dependent increase in intrinsic tryptophan fluorescence (ITF) of rat liver GK upon addition of glucose [151]. Structural evidence for the large-scale conformational alterations that accompany glucose binding was provided by the crystallization of the unliganded and glucose-bound liver hGK in 2004 [15]. Moreover, based on analysis of domain movement, Kamata et al predicted the occurrence of an intermediate “open” conformational form of hGK, in addition to the structurally solved “super-open” and “closed” forms, and, accordingly, devised a kinetic model that can account for the cooperative properties of GK (Figure 8C) [15]: (i) The super-open form is thermodynamically favorable in the absence of glucose; (ii) Upon binding glucose, a
slow transition from the “super-open” (low-affinity, inactive) form to the “closed” (high-affinity, active) form is triggered, that is slower than the catalytic reaction; (iii) After glucose phosphorylation, a fast “closed—open” transition takes place. A large proportion of the enzyme stays in the “open” conformation for some time, and if glucose binds during this time, the enzyme re-enters the catalytic cycle. This “fast cycle” is favored if the glucose concentration is sufficiently high. At low glucose concentrations, GK will return to the “super-open” form; (iv) GK has two catalytic cycles, a slow “super-open—closed” cycle and a fast “open—closed” cycle. The ratio between these two catalytic cycles, which is glucose-dependent, is responsible for the characteristic sigmoidal nature of GK and its low affinity for glucose [15].

1.9. Regulation of hGK activity

GK is regulated both at the transcriptional and post-translational level by a complex network of mechanisms which are fundamentally different in the hepatocytes and pancreatic β–cells [6, 34, 196]. As already mentioned (Section 1.6.2), GK gene transcription is driven by two tissue-specific promoters [133, 134, 197]. The hepatic promoter, which is primarily controlled by insulin levels, operates exclusively in the liver, whereas the upstream or neuroendocrine promoter controls GK expression in pancreatic β–cells as well as in glucose sensor cells of the gut, hypothalamus and anterior pituitary gland [5, 6, 127, 129-132, 196, 198]. In β-cells, GCK expression is constitutive at a relatively low basal activity but subject to direct regulation by glucose levels [23, 112, 124, 199, 200]. However, some studies suggest that the glucose-stimulated up-regulation of β-cell GCK transcription is mediated, in part, via insulin secreted in response to glucose [201, 202].

The adaptive response of GK to glucose also occurs at a post-translational level by means of substrate activation and stabilization [200]. The focus in the rest of this section and in the discussion will be on post-translational regulatory mechanisms of GK.
1.9.1. Post-translational regulation of hGK

In the pancreatic β-cell, the GK protein levels show only marginal variations under physiological conditions, and the enzyme is considered to be regulated at the post-translational level [196]. However, post-translational regulation of GK in both β-cells and hepatocytes is complex and only partially understood. It is established that GK activity, its subcellular localization and cellular stability are regulated by a spectrum of non-covalent GK-protein interactions that are different in β–cells and hepatocytes. In the hepatocytes, the interaction with the GK regulatory protein (GKRP) is a key short-term regulatory mechanism of GK activity [203, 204]. The 68 kDa GKRP, first discovered in rat liver, is an allosteric inhibitor of GK that binds preferentially to the super-open form of GK which predominates when glucose is low [15, 25, 203]. GKRP binds GK competitively with respect to glucose, and glucose binding releases GK from the GK-GKRP complex by inducing a conformational transition to the closed form. Physiologically, the GK-GKRP interaction and the subsequent inhibition of GK activity are promoted by fructose-6-phosphate, and suppressed by fructose-1-phosphate, compounds that bind to GKRP. Thus, the GK-GKRP interaction is modulated in response to fasting and feeding states [205, 206]. Importantly, GKRP provides a regulated translocation of GK between the cytosol and nucleus. As the glucose supply declines during periods of fasting, GKRP binds free cytoplasmic GK and transports it to the nucleus where GK is sequestered in an inactive state [207, 208]. Postprandially elevated glucose levels dissociate the GK-GKRP complex, and the active form of GK is translocated to the cytosol. This enables a rapid increase in GK activity and stimulation of glucose phosphorylation. Moreover, the GKRP-mediated compartmental redistribution of GK to the nucleus may serve to maintain a functional reserve of GK that can be quickly mobilized after a meal, in addition to stabilize and protect the enzyme from degradation by cytoplasmic proteolytic mechanisms [209, 210]. In GKRP deficient mice, the disruption of this regulation and the subsequent decrease in GK activity leads to altered glucose metabolism and impaired glycemic control [209]. Furthermore, functional studies on recombinant hGK enzymes have demonstrated that some GCK mutations cause a loss of regulation by GKRP which
may contribute to glucose intolerance in patients with GCK-MODY [65, 211-213]. The GKRP is not present or detectable in the pancreatic islets [214-216]. However, the presence of an inhibitory protein distinct from the liver GKRP has been suggested [215].

A second important regulator of GK activity is the bifunctional enzyme phosphofructokinase-2/fructose-2,6-bisphosphatase (PFK2/FBPase2). The enzyme is expressed in both hepatocytes and β-cells where it is involved in the regulation of the glycolytic (β-cell and hepatocytes) and gluconeogenic (hepatocytes) pathways [217]. PFK-2/FBPase-2 is a cytoplasmic binding partner of GK, and in insulin-producing cells this interaction has been demonstrated to enhance the catalytic activity of GK [217-219]. This effect may be, at least in part, due to stabilization of a catalytically favorable (closed) enzyme conformation at elevated glucose concentrations [218-220].

Beside its localization in the cytosol (hepatocytes and β-cells) and in the nucleus in complex with GKRP (hepatocytes), GK has also been found to bind to subcellular structures such as mitochondria (hepatocytes and β-cells) [221-223] and insulin-containing secretory granules (β-cells) [216, 224, 225]. In mitochondria, GK is part of a regulatory multiprotein complex, and association of this complex with the outer mitochondrial membrane is dependent on the presence of the proapoptotic protein BAD [221, 223, 226]. Moreover, the phosphorylation status of BAD helps regulate the catalytic activity of GK. The detailed molecular basis for the GK/BAD interaction and its physiological significance for glucose metabolism, GSIS and apoptosis are, so far, not fully understood. Another potential mechanism for post-translational modulation of GK function, specific for the pancreatic β-cell, arose from the observations that GK is part of the outer structure of insulin secretory granules in islet β-cells and insulin-secreting β-cell lines [216, 224, 225, 227, 228]. The granule↔cytoplasm translocation of GK is regulated by insulin and, moreover, the release of GK from the granule-bound state was accompanied by an increase in enzyme activity [228]. Hence, it was suggested that changes in GK activity induced by association/dissociation from insulin granules may be implicated in the regulation of GSIS in pancreatic β-cells [216].
Covalent post-translational modifications

Reversible post-translational modifications (PTMs) are used to dynamically modulate protein activity and stability. PTMs can occur at any step in the life-cycle of a protein serving various purposes, e.g. to mediate proper folding, cellular stability/turnover, subcellular localization, allosteric activation/inactivation, alter protein-protein interactions as well as to target a protein for degradation. There are > 200 different PTMs of which the majority occur by enzyme-mediated covalent attachment of a small chemical group, sugar, lipid or protein to one or more of the amino acid side chains in a target protein (e.g. glycosylation, acetylation, methylation, phosphorylation, S-nitrosylation, oxidation, ubiquitination, SUMOylation etc.). As already described, post-translational processes are important in the regulation of the cellular activity and stability of GK in hepatocytes and β-cells, but the knowledge of covalent PTMs of GK in target cells and their possible regulatory functions has been very limited. However, studies on cultured β-cells have demonstrated that the association of GK with insulin granules [216, 225] was dependent on its interaction with nitric oxide synthase (NOS) and that the localization and activity of GK was regulated by post-translational S-nitrosylation of the enzyme [216, 229]. Furthermore, it was suggested that defects in site-specific cysteine S-nitrosylation of GK are associated with naturally occurring GCK-MODY mutations in humans [230].

Ubiquitin (Ub) is a 76-amino acid globular protein (~8.5 kDa) that is highly conserved in eukaryotic cells. Ubiquitination (or ubiquitylation) is a reversible cellular process that involves the covalent attachment of one or several Ub proteins to a target protein [120, 231, 232]. Protein ubiquitination is an elegant example of how a single protein can regulate an array of diverse cellular processes such as cell cycle progression, regulated cell proliferation, cellular differentiation, apoptosis, transcriptional regulation and protein quality control (PQC) [233-236]. Given the central role of Ub in these processes, dysregulation of the ubiquitination machinery has been found associated (directly or indirectly) with the pathogenesis of many human diseases [237-241]. In addition to Ub, there is a growing family of Ub-like proteins (UbLs) which,
similar to Ub, covalently modify proteins on lysines by related enzymatic pathways, but with distinct functional implications [242].

The ubiquitin conjugating system and ubiquitin-mediated proteolytic pathway

Conjugation of Ub to a protein substrate or to itself usually involves the sequential action of three enzymatic reactions (Figure 9) [237, 243]. In the first step, Ub is activated for transfer by ubiquitin-activating enzyme E1 in an ATP-dependent reaction: E1 catalyzes the adenylation of Ub at the C-terminal glycine residue (G76), followed by the formation of an E1-S-Ub thioester intermediate [244]. In the second (ubiquitin-conjugating) step, activated Ub is transiently transferred from E1 onto an active site cysteine of the E2 enzyme via a trans(thio)esterification process to form another thioester bond [245]. To complete the enzymatic sequence, E2-S-Ub interacts with the E3 ubiquitin-protein ligase, which recognizes and associates with the substrate, promoting the transfer and conjugation of Ub [246, 247]. The Ub moiety is generally conjugated to target proteins through the formation of an isopeptide bond between a lysine (ε-amino group) on the protein and the C-terminal carboxyl group of Ub. The completion of one ubiquitination cycle results in a monoubiquitinated substrate. However, most often the cycle is repeated to form polyubiquitinated or multiubiquitinated substrates. The ubiquitin-chain is lengthened by the E3 ligase in collaboration with E1 and E2, sometimes with the help of an additional conjugating factor E4, specifically required for efficient multiubiquitin chain extension [248-250]. Eukaryotes are today estimated to have two E1 enzymes, ~40 E2 enzymes and > 600 E3s or E3 multiprotein complexes [251]. The E2-E3 pair is the primary determinant of substrate specificity [243]. Ubiquitination is reversible through the action of a large family of deubiquitinating enzymes (DUBs) (isopeptidases) which releases and disassembles polyubiquitin chains, enabling recycling of ubiquitin [252-254]. The human genome is predicted to encode ~95 DUBs which fall into at least five different classes [255].
Several ubiquitin-binding domains (UBDs), including UIM (Ub-interacting motifs) and UBA (Ub-association), can form transient, non-covalent interactions with either mono-ubiquitin or polyubiquitin chains [249, 250]. Cellular proteins containing one or several UBDs (often called ubiquitin receptors) are the immediate decoders of ubiquitination, being responsible for transmitting specific Ub signals to downstream cellular events. Importantly, the selective preferences of UBDs for ubiquitin chains of specific length and linkage are central to the versatile functions of Ub.

The topology of the polyubiquitin chain appears to be a function of specific E2s and the E2-E3 combinations involved [251-253]. All seven internal Lys residues (K6, K11, K27, K29, K33, K48 and K63) of Ub are believed to contribute to the assembly of polyubiquitin chains [254, 255]. This ability to form a variety of structures with diverse lengths and linkages distinguishes ubiquitination from UbL modification of proteins. It is generally assumed that the formation of polyubiquitin chains of different linkage types provides functional specificity that determines the fate of the modified protein [254].

Ub is best known for its function in targeting proteins for controlled degradation by the 26S proteasome [256, 257]. Most proteins degraded by the ubiquitin-proteasome pathway are linked to a polyUb chain in which the Ubs are coupled through isopeptide

![Figure 9. The ubiquitin-proteasome degradation pathway.](image_url)

*Figure 9. The ubiquitin-proteasome degradation pathway.* See text for details. The figure is reprinted from [248], with permission from Nature Publishing Group, a division of Macmillan Publishers Ltd © 2010.
linkages to K48 (or in some cases K11 or K29) on the preceding Ub. At least four sequentially added Ub moieties are believed to be required for efficient recognition and binding of the modified target protein to the proteasome [258]. In contrast, monoubiquitination and K63-linked polyubiquitin chains serve non-proteolytic functions in various intracellular pathways [234-236, 254, 259, 260]. Once a target protein is marked by K48-linked polyubiquitination, it appears to have a short half-life in cells, as it is rapidly degraded by the proteasome. Proteins can be targeted directly to the proteasome through recognition of polyubiquitin by the 26S complex or, alternatively, it can occur indirectly, mediated by various proteasomal shuttle factors [255]. The proteasome is a large (26S or ~2400 kDa) ATP-dependent multicatalytic protease that is present both in the nucleus and cytosol of eukaryotic cells [237, 261]. The 26S proteasome is composed of two subcomplexes, a 20S proteolytic core that provides the proteolytic activity needed to degrade modified substrates, and a 19S regulatory complex that confers the ability to recognize and unfold polyubiquitinated protein substrates and insert them into the proteolytic chamber [261]. The degradation through the proteasome is an irreversible process and the proteins are degraded to small peptides, most of which are hydrolyzed to amino acids by the sequential action of endo- and exopeptidases in the cytosol and nucleus. Proteins destined for degradation need to be deubiquitinated to ensure efficient substrate degradation as well as recycling of Ub, and tightly regulated deubiquitination is accomplished by proteasome-associated DUBs [262, 263]. The main steps in the ubiquitin-proteasome pathway are illustrated in Figure 9.

**Role of the ubiquitin-proteasome pathway in β–cell dysfunction and hyperglycemia**

Chronic hyperglycemia has been associated with oxidative stress and subsequent defective insulin secretion [264]. As previously mentioned, the pancreatic β-cell is particularly vulnerable towards oxidative stress [179], which in turn may trigger the ubiquitin-proteasome pathway [265], suggesting a potential involvement of this pathway in hyperglycemia and β-cell dysfunction [266]. Interestingly, the ubiquitin-proteasome degradation pathway has been ascribed a regulatory role in the pancreatic
β-cell in glucose-stimulated (pro)-insulin synthesis, biogenesis and surface expression of the \( K^+_{\text{ATP}} \) channel and in maintaining the normal function of the \( Ca^{2+} \) channel – all essential components of GSIS [267-269]. However, little is known about covalent post-translational modifications of GK and their possible regulatory functions in the target cells, and the molecular and cellular mechanisms involved in the degradation/turnover of GK are also poorly understood. The possible implications of Ub (or UbLs) and the ubiquitin-proteasome pathway in this regard have not yet been subject for investigation and, hence, remain to be elucidated.
2. AIMS OF THE PRESENT STUDY

Two 3D structures of hGK were solved in 2004 [15]. This represented a milestone in the research on hGK and hGK-linked glycemic disorders opening up new and intriguing possibilities for structure-function studies of WT hGK and its disease-associated mutant forms. The two structures also represented the starting point of the present study, with the main objective to characterize structural, dynamic, catalytic and regulatory properties of WT hGK as a basis for studies on disease-related mutant forms in which the molecular mechanism(s) of the disease remained to be established.

The specific aims were:

1. To characterize steady-state kinetic properties of WT hGK and some selected mutant forms.
2. To study the reversible glucose-induced global conformational transitions in WT hGK and Trp mutant forms using ITF spectroscopy.
3. To provide new insight into the mechanism of substrate (glucose) activation:
   a. Identify the active site residues involved in the initial recognition and binding of glucose to the super-open (apoenzyme) conformational form.
   b. Determine the site of initiation of the glucose-induced conformational transition.
4. To provide new insight into the interaction of the cosubstrate (Mg)ATP with hGK:
   a. Investigate whether the cosubstrate can bind to the apoenzyme and the possibility of a related conformational change.
   b. Identify active site residues involved in ATP recognition and binding to the super-open (apoenzyme) conformational form as well as related structural changes.
5. To gain further insight into the catalytic function of WT hGK and its multiple disease-related mutant forms.
6. To examine ubiquitination as a potential regulatory mechanism of hGK, in particular with respect to enzyme activity and stability (degradation/turnover).
3. SUMMARY OF RESULTS

3.1. Paper I

In this work, we have studied binding of α-D-glucose and the associated glucose-induced global conformational transition of pancreatic hGK in real-time by ITF spectroscopy. By a combined approach using functional mutagenesis and structural dynamic analyses we aimed to identify key residues involved in this process. A hyperbolic binding isotherm for α-D-glucose was observed at 25 °C with a $K_d$ value of 4.8 ± 0.1 mM. Single-site W→F mutation analyses revealed a contribution of all three tryptophans in the overall glucose-induced ITF enhancement, with major contributions of W99 and W167, which both are located in highly flexible loop structures. The measured biphasic time-course of the fluorescence enhancement is influenced to a variable extent by all three tryptophans. To identify the residues involved in the initial binding of glucose to the unliganded “super-open” conformation and subsequent conformational transitions, all interacting residues defined in the 3D structure of the closed conformation were individually mutated. The mutations N204A, D205A and E256A/K of the L-domain resulted in enzyme forms which did not bind α-D-glucose at 200 mM as measured by ITF and they were essentially catalytically inactive. In contrast, in the mutants of the S-domain (T168G and K169N), glucose induced a significant fluorescence enhancement, but with increased $K_d$ values. Overall, our data support a molecular dynamic model in which initial binding of α-D-glucose to residues N204, N231 and E256 of the L-domain and subsequent interaction with D205 trigger local structural changes that propagate in a concerted fashion toward a closed conformation. This involves the highly flexible interdomain connecting region II (R192-N204), helix 5 (V181-R191), helix 6 (D205-Y215) and the C-terminal helix 17 (R447-K460).
3.2. Paper II

Catalytic activation by α-D-glucose binding is a key regulatory mechanism of hGK. The mechanism of activation is reasonable well understood based on the structural analyses of two crystallized conformations of the enzyme. By contrast, there are no similar structural data on ATP binding to the ligand-free enzyme and how it affects its conformation. Using ITF spectroscopy and the catalytically inactive mutant form T228M as a reference to correct for the inner filter effect, we demonstrated that binding of adenine nucleotides to the unliganded form of pancreatic hGK resulted in a conformational change. Furthermore, we observed that AMP-PNP and ATP bind to WT hGK in the absence of glucose with apparent $[L]_{0.5}$-values of 0.27 ± 0.02 mM and 0.78 ± 0.14 mM, respectively. Moreover, the binding of nonhydrolysable ATP analogues to the apoenzyme increased its affinity for α-D-glucose. The nucleotide-induced change in conformation was further supported by the significant protective effect of ATP on the binding of the extrinsic fluorescence probe 8-anilino-1-naphthalenesulfonate (ANS) and on the limited proteolysis by trypsin. Subsequently, the biochemical and biophysical data were confirmed by MD simulations conducted with a modeled structure, providing the first insights into the dynamics of the binary complex with ATP including a motion of the flexible surface/active site loop. In this complex, the adenosine moiety is packed between the α-helices 12 and 15 and stabilized by H-bonds (to T228, T332 and S336) and hydrophobic interactions (to V412 and L415). The binding resulted in a conformational change with domain motions and a partial closure of the active site cleft. Overall, our data indicate that ATP-induced changes in the conformation of hGK may have implications for its kinetic cooperativity with respect to α-D-glucose.
3.3. Paper III

In this study, we investigated the role of ubiquitination as a potential post-translational regulatory mechanism of hGK. Using the rabbit reticulocyte lysate (RRL) as an *in vitro* model system, we demonstrated that pancreatic β-cell (isoform 1) and liver (isoform 2) hGK serve as substrates for the ubiquitin-conjugating enzyme system. Moreover, we found that both isoforms are polyubquitinated on at least two lysine residues. Mutational analysis was used to possibly detect the target residue(s), and the results indicated that multiple lysine residues function as redundant ubiquitin acceptor sites. Deletion of the C-terminal α-helix, containing a putative ubiquitin-interacting motif (UIM), resulted in loss of polyubiquitination at one site, suggesting that one of the target lysines is associated with the UIM motif. Proteasome-dependent degradation of hGK was studied in a proteasome and ATP-enriched RRL system and demonstrated that poly-/multi-ubiquitination of hGK *in vitro*, at least in part, served as a signal for proteasomal degradation of the newly synthesized protein. Moreover, we demonstrated that free polyubiquitin chains (Ub₅, K48-linked) interacted with and allosterically activated recombinant hGK at low nanomolar concentrations (1.4 fold at ~100 nM Ub₅), assigned to their equilibrium binding to the C-terminal UIM site. Furthermore, the affinity of Ub₅ binding to hGK was regulated by the glucose-dependent conformational state of the site. We suggest that both these ubiquitin-mediated processes represent potential physiological regulatory mechanisms for hGK as a glucose sensor in pancreatic β-cells.
4. METHODS AND METHODOLOGICAL ASPECTS

4.1. Functional comparison of GST-hGK and non-tagged hGK

For detailed biochemical and physico-chemical studies on hGK, milligram quantities of highly purified recombinant enzyme are imperative. In this study, efficient and high-level expression of recombinant pancreatic hGK was provided by expression in *E. coli* as GST fusion proteins [270]. Rapid and gentle purification by affinity chromatography ensured the recovery of pure (>98% as determined by SDS-PAGE) and enzymatically active, full-length GST-hGK (76 kDa). After cleavage of the fusion protein at the restriction site by factor Xa, non-tagged hGK (52 kDa) was isolated with a recovery of ~50 % and no detectable loss in enzyme activity. The relative low yield of isolated, tag-free hGK can be explained by incomplete cleavage and loss in the final repurification step.

Prior to the outset of this project, the GST fusion proteins of hGK were mostly used in the kinetic characterization of mutant forms, due to the potentially stabilizing effect of the fusion partner. In order to validate the utility of GST-hGK in kinetic and fluorescence studies, we compared the functional properties of tagged and non-tagged WT enzyme. They demonstrated similar steady-state kinetic parameters and comparable apparent binding constants for glucose (Paper I) and AMP-PNP (Paper II) in equilibrium binding studies by ITF spectroscopy. These findings, in good agreement with other reports on the subject [271-274], substantiate that the fusion partner at the N-terminal does not affect the kinetics of WT hGK or perturb the substrate-induced conformational changes. Moreover, the increase in ITF signal response when glucose is added at saturating concentrations was comparable without \((F_{eq}-F_o \sim 33)\) and with \((F_{eq}-F_o \sim 30)\) the fusion partner (Paper I), considering the experimental error in determining the absorption coefficients at 280 nm for the two proteins. However, the relative fluorescence increase \((\Delta F_{max,eq}/F_o)\) of WT GST-hGK was \(\sim 27\%\), as compared with a \(\sim 60\%\) increase for WT hGK. This is explained by GST containing four Trp residues [275], contributing to the background fluorescence \((F_o)\). These quantitative differences in the ITF response to glucose between non-tagged and tagged enzyme
were also observed in the binding studies with AdNs, although the $\left( \Delta F_{\text{max,eq}}/F_o \right)$ value of the AdN-induced fluorescence quenching was much smaller.

Concurrent studies of more than 30 mutant forms of hGK have provided a detailed comparison of the kinetic parameters of tagged and non-tagged hGK, which further support a limited interference of the GST tag [273, 274]. However, care should be taken when characterizing more mutant proteins since the GST moiety may not always be inert [273]. In fact, recent data are consistent with GST acting as a transiently stabilizing partner for some mutant forms (Negahdar et al, unpublished and [273]). This finding is not surprising since recombinant, non-tagged WT hGK has a marginal thermal stability and a propensity to aggregate at temperatures $\geq 40$ °C (Paper II and [276]).

### 4.2. Helix nomenclature

The helices in hGK are designated according to the crystal structures resolved in 2004 [15]. A total of 17 helices were identified in the unliganded, super-open state (PDB identity: 1v4t) versus 19 helices in the glucose and allosteric activator bound, closed state (PDB identity: 1v4s) (see Paper I). In this study, we have numbered the helices successively irrespective of helix type ($\alpha$- or $3_{10}$ helix). This differs from Kamata et al who generally refer only to $\alpha$-helices (one exception). Thus, in hGK the C-terminal helix corresponds to helix 17/19 or $\alpha$13. The information on secondary structure elements were generally retrieved from the MolMovDB of the Yale Morph Server [277, 278].

### 4.3. Analyses of protein dynamics

Although the static 3D crystal structures of GK have been determined for the apoenzyme and different complexes with substrate(s) and low molecular weight activators (see Section 1.7), the function of the protein in solution at physiological temperatures is governed by its dynamic properties. In the present study, we have used
fluorescence spectroscopy and computational methods to gain insights into different conformational states and dynamics of hGK.

4.3.1. Intrinsic fluorescence spectroscopy

Intrinsic fluorescence spectroscopy is a highly sensitive biophysical tool for the structural and dynamic characterization of proteins. The method relies on the natural (intrinsic) fluorescence of proteins in the near UV region, which mainly arises from the aromatic amino acids tryptophan (Trp) and tyrosine (Tyr). Trp is by far the most frequently used fluorescence probe due to its high extinction coefficient ($\varepsilon$), large Stokes’ shift ($h_{\text{EX}} - h_{\text{EM}}$) and quantum yield ($\Phi_T$), coupled with the unique environmental sensitivity of its indole ring. Moreover, Trp is a relatively rare amino acid in proteins, and therefore more specific information on the local conformation and environment of individual Trp residues can be attained. Its fluorescence is characterized by a high signal to noise ratio with need of relatively low concentrations ($\mu$M) of pure protein when reagents of high analytical grade are used. In intrinsic tryptophan fluorescence (ITF) spectroscopy the fluorescence from Trp is selectively measured using excitation wavelengths in the range 295-305 nm. Both the fluorescence emission intensity and maximum emission wavelength ($\lambda_{\text{max}}$) vary depending upon the polarity and dynamics of the molecular environment, which make ITF a useful conformational probe. The $\lambda_{\text{max}}$ of Trp residues in proteins is highly related to their degree of solvent exposure. In general, the $\lambda_{\text{max}}$ of the indole ring is blue-shifted if the group is buried within the hydrophobic core of the native protein, and its emission shifts to longer wavelengths (red shift) upon exposure to polar amino acid groups or solvent water (following, for instance, ligand-induced conformational changes). Trp residues buried in apolar core regions of a protein have a blue emission maximum, as low as ~310 nm, whereas Trp fully exposed to an aqueous environment can have a red emission maximum of ~355 nm. The fluorescence spectra of multi-Trp proteins are generally broader than that of single-Trp proteins due to the contribution of multiple emitters with different, but overlapping spectra.
GK has three Trp residues (W99, W167 and W257), and the 3D structures of the protein have provided valuable information on their backbone conformation and microenvironment in different conformational states. Thus, ITF has been found to be a sensitive indicator of even minor structural changes of the enzyme and its dynamic properties on the ms-to-min time scale. As first reported by Lin and Neet [151], binding of glucose results in an increase in the ITF signal response and a shift in $\lambda_{\text{max}}$ of the emission spectrum. In order to assess the contribution of the individual Trp residues to the glucose-induced fluorescence increase, we made (Paper I) single-Trp mutants (retaining two Trp residues) by site-directed mutagenesis, where the Trp residues were substituted one at the time. Trp was replaced by phenylalanine (Phe), which does not contribute to the fluorescence when excited at 295 nm. The mutations did not significantly affect the backbone structure or conformational stability of the enzyme (Paper I and [273]). However, the mutations affected the steady-state kinetic parameters of the enzyme. The W167F mutant in the active site loop showed a reduction in catalytic efficiency ($k_{\text{cat}}/[S]_{0.5}$) and kinetic cooperativity with respect to glucose, which probably can be explained by its localization in the glucose binding pocket adjacent to the glucose-interacting residues T168 and K169. The W257F mutant revealed a ~50% reduction in catalytic efficiency, which may be due to its position in a rigid backbone structure near the glucose binding site, conferring minor local structural perturbations affecting the glucose affinity. Interestingly, the W99F mutant in a highly flexible loop structure displayed a small increase in substrate affinity, and thus, catalytic efficiency. W99 is located in close vicinity to connecting region I and the GK activator (GKA) binding site, and it has been suggested that mutations at this site may stabilize the closed form of the enzyme [273].

In order to determine the contribution of the three Trp residues to the glucose-induced fluorescence enhancement of hGK, their +glucose/-glucose fluorescence difference spectra were generated. Mutations of W99 and W167 resulted in a large decrease (~70 and ~80 %, respectively) in the fluorescence response to glucose, whereas mutation of the “buried” W257 gave only a moderate decrease of ~35%. Furthermore, the W99F mutant demonstrated a ~11 nm blue shift in $\lambda_{\text{max}}$ in the difference spectrum compared
to WT hGK, whereas only a ~2 nm shift was observed for W167F and W257F (towards the blue and red, respectively). This indicates that W99 is responsible for the more red emission of WT hGK, in agreement with its high degree of exposure to aqueous solvent in both the super open (27%) and closed state (45%) of hGK. The fluorescence spectrum of WT hGK apoenzyme (pH 7.0) displayed an emission maximum of ~340 nm. Upon binding of glucose a ~2 nm blue shift was observed. This shift is most likely explained by changes in the fluorescence of W167, experiencing a large reduction in solvent accessibility upon cleft closure and desolvation of the active site (Paper I). Recently, the relative solvent accessibility of the Trp residues was determined by acrylamide quenching analyses (W99 (11.6/13.1 %), W167 (14.4/5.8 %) and W257 (6.2/5.2 %)), providing experimental evidence that the solvation of W167 is significantly decreased by glucose [279]. Thus, the magnitude of the glucose-induced fluorescence increase is primarily determined by W99 and W167, and little affected by W257. Complementary studies on double-Trp mutants (retaining only one single Trp residue) have recently been performed [273, 279], and the results are consistent with the findings presented herein for the single-Trp mutants.

4.3.2. Extrinsic fluorescence spectroscopy

A number of polyaromatic hydrocarbons or heterocycles have been used as extrinsic fluorescent probes in the study of protein structure and ligand-induced conformational changes. Here, we have used 8-anilinonaphthalene-1-sulfonate (ANS) as an extrinsic fluorophore with affinity for hydrophobic clusters in GK. The weak fluorescence of ANS was greatly enhanced upon binding to ligand-free WT hGK. Both ATP and glucose binding significantly reduced the ANS fluorescence signal (glucose>ATP), compatible with a decrease in accessible hydrophobic clusters as compared with the ligand-free enzyme.
4.3.3. Computational methods

X-ray-diffraction data on GK contains information not only about the average 3D structure but also about the spatial distribution around this state. This mean-square atomic displacement [280] is commonly expressed as the B-factor or temperature factor. The isotropic B-factor values calculated for the Cα carbons (Paper I) demonstrated the freedom and restriction for various sites, with low values (≤ 30 Å²) for the glucose-interacting residues in the unliganded state, except for T168 and K169. Complementary global mode analysis by the Gaussian Network Model (GNM) [281, 282] revealed similar sites (minima) of low translation mobility.

Among the different theoretical methods available for description of protein flexibility, molecular dynamics (MD) is probably the most powerful [283-285]. In this study (Paper II), we have used state of the art atomistic MD simulations, close to the physiological conditions, by the widely used AMBER program package [286]. Based on the atomic coordinates for unliganded hGK as a starting point, MD simulations (ps-to-2 ns time scale) of the binary complexes with ATP and glucose were carried out in explicit water (that is, individual water molecules were included in the simulations). The simulations indicated a significant conformational change of hGK upon ATP binding, including motion of the flexible surface/active site loop and partial closure of the active site cleft, and also showed the residues directly contacting the nucleotide at the active site. In contrast, when the MD simulations were performed with glucose in the super-open conformation, there was no measurable closure of the active site cleft during the 2-ns MD simulations. In this case, the simulation time was too short to observe the large glucose-induced conformational change of hGK, characteristic of a hysteretic enzyme, which occurs on a ms-to-min time scale in the ITF studies (Paper I and [157, 287]).
5. GENERAL DISCUSSION

5.1. GK protein stability

GK plays a central role in maintaining glucose homeostasis, serving as a glucose sensor in the pancreatic β-cells and as a regulator of hepatic glucose uptake and glycogen synthesis [7]. Alterations in GK activity and regulation are associated with abnormal glycemia, as illustrated by naturally occurring GCK mutations causing diabetes or hyperinsulinemic hypoglycemia. This emphasizes the importance of preserving the stability and activity of this enzyme, also under conditions of cellular stress.

Recombinant GK without a fusion partner is only marginally stable under physiological conditions. It has a low intrinsic thermal stability (apparent $T_m = 42.4 \pm 0.2 \, ^\circ\text{C}$) and a propensity to aggregate at $T_p \geq 40 \, ^\circ\text{C}$ (Paper II and [276]). Glucose and ATP stabilize the enzyme through a ligand-induced conformational change to a more compact conformation, as demonstrated by the protective effect of the ligands (glucose > ATP) on limited proteolysis (Paper II). A stabilizing effect of glucose on the thermal stability of GK has also been observed in ITF spectroscopy [279] and differential scanning calorimetry studies [279, 287]. Moreover, glycerol has been demonstrated to increase the thermostability of GK and promote glucose binding (lowering its $K_d$) presumably by reducing/removing water interactions with the protein [279]. However, the stabilization by glucose and glycerol was limited as they were unable to prevent irreversible heat denaturation above 40 °C [279]. Interestingly, in contrast to guanidine-Cl (Paper II), urea prevented irreversible aggregation and preserved reversibility of the thermal unfolding process, probably explained by its ability to bind directly to the protein backbone and the hydrophobic side chains of GK [279]. Overall, these studies have contributed to an increased understanding of the biophysical basis of WT GK thermostability and established a foundation for studies on disease-associated mutant forms.
In silico calculation of the free energy of folding/unfolding (± ΔAG) [288, 289], predict that several misfolded missense GCK-MODY mutations harbor a reduced thermal stability that could be related to an increased rate of degradation when expressed in mammalian cells (Negahdar et al, unpublished). To date, ~15 GCK-MODY mutations have been identified with an apparently reduced in vitro thermal stability of the recombinant enzymes as fusion proteins, and has been proposed to explain (at least in part) their “loss-of-function” [114, 139, 211, 212, 271, 272, 290-295]. However, it can be anticipated that many more of the >600 GCK mutations identified so far will demonstrate reduced in vitro stability. Although such in vitro stability studies can provide valuable information on mutant forms, the kinetic and thermodynamic parameters for folding-unfolding processes in the cell are different from the “ideal” in vitro conditions. For instance, macromolecular crowding affects the structure and function of proteins under physiological conditions and favors intermolecular interactions with proteins and other cellular components [296-299]. Recently, we have investigated the effect of the catalytically near-normal mutations S263P, G264S and R275C on the cellular stability of GK (Negahdar et al, unpublished). Evidence was provided that their “loss-of-function” is, at least in part, a result of protein misfolding and destabilization, causing dimerization/aggregation and an enhanced rate of cellular degradation.

A complex network of mechanisms is involved in the regulation of GK catalytic activity in pancreatic β–cells and hepatocytes. However, little is known regarding the cellular protein quality control (PQC) machinery involved in the recognition and targeted degradation of GK and its many disease-associated mutant forms [65]. In Paper III, we have demonstrated that recombinant pancreatic and liver WT hGK are covalently modified by ubiquitination, and that poly/multi-ubiquitination of hGK in vitro can serve as a signal for proteasomal cotranslational degradation of the newly synthesized protein. Protein synthesis, folding and degradation are closely coupled processes, and a significant fraction of newly synthesized proteins are degraded cotranslationally [300-302]. This may represent an early recognition by the PQC machinery of misfolded nascent proteins [301, 303]. Furthermore, recombinant hGK
was found to interact with and to be allosterically activated up to ~1.4-fold by purified free polyubiquitin chains at low nanomolar concentrations, assigned to their equilibrium binding to a C-terminal UIM site (Paper III). Together, these findings suggest a potential role of ubiquitination in the cellular regulation of GK catalytic activity and stability/turnover, as well as in the interaction with other cellular proteins, including ubiquitin-conjugated proteins and free/conjugated poly-ubiquitin chains. Because ubiquitination of misfolded proteins associated with cytoplasmic chaperones are mostly degraded through the ubiquitin-proteasome system [304], further studies are warranted to investigate this possibility for selected loss-of-function mutants of GK.

5.2. The multiphasic global conformational transitions and perturbations of conformational equilibria

In steady-state kinetics, GK shows non-hyperbolic (sigmoidal) dependence on glucose concentration [150, 187], and the underlying mechanism for the cooperative behavior of GK has been of primary interest in the study of this enzyme. The classical concerted model of cooperative kinetics [305] does not apply to the monomeric GK with a single substrate binding site. The possibility of two glucose binding sites in hGK has been proposed [147], but only a single glucose binding site was observed in the crystal structure [15, 152]. Furthermore, equilibrium binding of glucose to unliganded WT hGK displays a hyperbolic (non-cooperative) binding isotherm ($K_d = 4.9 \pm 0.1$ mM) (Paper I and [273, 287]), consistent with a single binding site. It is now widely accepted that the positive cooperativity of GK with respect to glucose is kinetic in origin, and monomeric GK has emerged as a model system for understanding this type of cooperative response.

GK is activated by binding of glucose, and this process has been described as a reversible, slow transition from a super-open, inactive (low affinity) state to a closed, high activity (high affinity) state [151, 287]. Crystal structure analyses of the unliganded and glucose-bound hGK [15, 152] have confirmed these observations by
demonstrating that binding of glucose at the active site induces a large-scale domain movement that closes the active site cleft and creates the stereochemical environment for binding of the cosubstrate MgATP$^{2-}$ and catalysis. The three residues N204, N231 and E256 in the L-domain was proposed to function as primary contact residues (Paper I and [15]). The hGK-glucose association is driven by a favorable entropy change ($\Delta S = 150 \pm 10$ J·mol$^{-1}$·K$^{-1}$) (Paper I and [279]), in keeping with the fact that an increase in protein dynamics plays a dominant role in the interaction. In the closed conformation, precise alignment of additional substrate interacting residues (notably T168 and K169 in the flexible surface/active site loop of the S-domain) and the subsequent higher affinity for glucose, efficiently accelerate the chemical reaction on binding of MgATP$^{2-}$.

Recent results from x-ray crystallography, high-resolution NMR, targeted MD simulations, pre-steady state and steady-state glucose binding studies have shown that hGK is an intrinsically mobile enzyme that can sample multiple distinct conformational states both in its unliganded and binary enzyme-glucose complex [151, 158, 160, 276, 287, 306-308]. Furthermore, its positive kinetic cooperativity with respect to glucose is mediated by the glucose-induced conformational transition, which probably includes interconvertible intermediate states exhibiting variable degree of cleft closure and different affinities for glucose. Interestingly, Liu et al [309] have assessed conformations of GK in solution by small angle x-ray scattering (SAXS) experiments, demonstrating that glucose in solution dose-dependently converts GK from an apo (super-open) conformation to an intermediate active open conformation, distinct from the active, closed conformation observed in crystals [15]. The glucose-bound open conformation was predicted with a cleft opening angle of $\sim 17^\circ$ relative to the closed conformation obtained for the crystal structure [309]. The authors suggest that the active open conformation represents a physiological intermediate in the reaction cycle of GK, in agreement with a previous proposal [15].

Increasing evidence substantiate that also the ligand-free enzyme in solution is in a preexisting equilibrium between at least two conformers, i.e. the super-open conformation and an alternative (presumably less open) conformation with a higher
affinity for glucose [276, 306-308]. However, the apoenzyme is considered to be dominated by the thermodynamically favored super-open conformation [158, 306, 307, 309], as suggested based on 3D structure analyses [15, 309]. In order to gain insight into the free-energy landscape of proteins, temperature perturbations can be used to shift the population equilibrium [310]. In unliganded WT hGK the observed temperature induced (1 °C to 39 °C) reversible quenching of the ITF fluorescence is consistent with a slow conformational isomerization (thermal hysteresis), reminiscent of the global glucose-induced conformational change (Paper I). Moreover, the observed biphasic time course suggests the presence of a relatively stable intermediate in the thermal transition, and gives support to the existence of an equilibrium between conformational substates of hGK in the absence of ligand.

In Paper II, we have provided the first experimental evidence for an equilibrium binding of ATP and the analogue AMP-PNP to the ligand-free enzyme and a subsequent nucleotide-dependent conformational change, including motion of the flexible surface/active site loop and partial closure of the active site cleft. In addition, MD simulations provided insights into the active site contact residues involved in binding ATP. The ligand-free and glucose-bound enzymes are dynamic entities that can sample multiple conformational substates [276, 308], and we hypothesize that nucleotide binding may shift the equilibrium between different conformations (see Figure 10), as previously demonstrated for glucose (Paper I and [151, 160, 287, 306, 307]) and allosteric effectors (see below). In fact, results from steady-state kinetic analyses (Paper II) gave support to a certain conformational control of GK catalytic activity by binding of (Mg)ATP, with possible implications for the kinetic cooperativity with respect to glucose. Moreover, using ITF fluorescence titrations it was demonstrated that binding of ATP analogues to the ligand-free enzyme resulted in an increased equilibrium binding affinity for glucose, both in the WT enzyme and the low affinity GCK-MODY mutant L146R (Paper II). We speculate that these effects may be related to a partial catalytic activation of GK following ATP binding, and that similar or possible larger effects of ATP may occur in other GK disease-associated mutant forms.
Figure 10. Schematic representation of the standard free energy landscape for the catalytic network of the GK reaction. Conformational changes occur along both axes. The changes occurring along the Reaction Coordinate axis correspond to the environmental reorganization at the active site that facilitates the chemical reaction (Papers I and II, and references therein). It includes molecular motions related to substrate binding (Section 4 and [158]). In contrast, the conformational changes occurring along the Ensemble Conformations axis represent the ensembles of conformational substates existing at all stages along the reaction coordinate [276, 307, 308], possibly leading to a number of parallel catalytic pathways [310]. A state is defined as a minimum in the energy surface, whereas a transition state is the maximum between the wells [311]. A change in the system (ligand binding, mutation) shift equilibrium between states. The figure and portions of the caption are adapted from a recent model on the catalytic network of the enzyme dihydrofolate reductase [312].

It has been suggested that the allosteric GK regulators GKA and GKRP modulate the activity of GK by finely tuning its conformational landscape [307]. In other words, they act as reversible conformational traps, by binding and stabilizing a specific GK conformation [307]. According to this model, GKRP stabilizes open GK conformers that prevail at low glucose concentrations, whereas increased glucose concentration and/or addition of GKAAs shift the equilibrium towards the closed conformers and induce a dissociation of the GK-GKRP complex. SAXS experiments have demonstrated that the GKAAs shift the conformational equilibrium of GK to the active conformation at lower glucose concentrations [309]. The SAXS profiles revealed small conformational differences between the various GK-glucose-GKA complexes,
which was consistent with the inherent ability of the individual GKA to activate GK (GKAs that suppressed $[S]_{0.5}$ for glucose mildly gave a GK-glucose-GKA complex that was less compact than those GKAs with a more dramatic impact on the $[S]_{0.5}$ -value) [309].

Similarly, it has been proposed that mutations in hGK may shift its conformational equilibria by stabilizing or destabilizing one conformer, or facilitate or slow the super-open→closed transition. In the same way as for GKAs, activating GK mutations, most of which occur at or near the allosteric GKA site, are predicted to shift the enzyme towards the active conformation [158, 160, 309]. This can be exemplified by the activating mutations Y214C/S, V455M and A456V, which have been predicted to perturb the electrostatic and hydrophobic interactions between helix 6 (D205-Y215) and the C-terminal helix (helix 17, R447-K460) which stabilize the inactive, super-open conformation (Paper I). These helices are structurally important in the allosteric regulation of GK. As a result, the structural perturbation of the apoenzyme favor glucose binding and the transition to the active, closed conformation (Paper I and [158, 160, 309]). Furthermore, transient and steady-state kinetic studies and/or SAXS analyses on the mutations Y214C/S, Y215A and A456V have demonstrated that the mutated enzymes seem to adopt a more compact conformation in the apo form, being favorable for glucose binding but perturbing the allosteric regulation of the enzyme by GKA and, in some cases, also by GKRP [160, 273, 274, 309] (see also section 5.5). Further mutational studies, notably on GCK-MODY associated mutations, may contribute to a better insight into the mechanistic and functional implications of the multiple conformational equilibria and ligand-induced conformational transitions of hGK.

To summarize, both substrate binding and catalysis are associated with substantial conformational changes in GK. The catalytic pathway of the enzyme involves a glucose-induced cooperative conformational transition and multiple intermediates and conformations along the reaction coordinate (Figure 10) [151, 158, 160, 276, 287, 306-309]. Furthermore, the binding of various ligands and effectors (glucose, ATP, GKA, GKRP etc.) as well as mutations in GK may modulate the enzyme activity by shifting
the conformational equilibria. We propose a simplified reaction scheme for mammalian GK (Scheme 1), based on the existence of a pre-existing equilibrium between at least two (a super-open (GK\textsuperscript{‡}) and a presumably less open (GK\textsuperscript{z})) conformers with different affinity for glucose (low and higher affinity, respectively).

**Scheme 1: Reaction scheme for mammalian GK.** GK\textsuperscript{‡} and GK\textsuperscript{z} are the proposed apo conformers in a two-state conformational model, whereas GK* is the active, closed conformation of the enzyme (Paper II).

### 5.3. Allosteric effectors of hGK

Allosteric effectors are commonly small molecules or proteins. However, mutations and covalent modifications such as phosphorylation and formation/breaking of disulfide bonds, can also act as allosteric triggers [314].

In addition to the basic kinetic cooperativity of GK with respect to glucose, the catalytic activity of GK is modulated by its interaction with physiological and pharmacological allosteric effectors, including proteins and small molecules. The discovery of small molecules as GK activators, and the identification of their binding site in the 3D structure [15, 152], represents a promising development in the treatment of T2D [9, 13, 34, 168, 169, 315-318]. Against this background it has become important to fully understand the complex allosteric regulation of the enzyme and the possible cross-talk between the allosteric effectors.
The crystal structures of GK have revealed that its surface properties are very dissimilar in the open and closed conformations [15, 152]. Moreover, a number of related algorithms have revealed that the location of potential binding sites of small molecules in surface motifs (pockets) is very different in the two conformational states [319]. This difference explains the biochemical finding that the function of the allosteric effectors is dependent on the conformational state of the enzyme.

5.3.1. Glucokinase activators (GKAs)

The small molecule GKAs have diverse chemical structures, and varies in their activating potency depending on how they modulate the kinetic parameters ($[S]_{0.5}$ for glucose, $V_{\text{max}}$ and $n_H$) of GK, and thus their overall effect on its catalytic efficiency [9, 34, 309]. Today, the binding site of several GKAs has been defined crystallographically [15, 152, 162, 164-167], and they all bind to a common hydrophobic pocket [9, 15, 34, 152, 161, 162, 318]. The site is located near the hinge region between the S- and L-domains in the closed (active) conformation, but is in general inaccessible in apo GK (even though some GKAs have been reported to bind with low affinity to the unliganded form) [15, 152, 307].

In general, many allosteric effector molecules exert their effect by causing a conformational change in the target protein that affects its ability to bind, or properly interact with, the substrate(s) [314]. However, no conformational changes were observed in the crystal structure of GK upon binding of GKA, and no apparent effect on the molecular flexibility could be detected [152]. Therefore, the authors proposed an allosteric mechanism in which GKAs, on binding to the closed form of GK, cause a shift in the equilibrium of conformational ensembles (Section 5.2) towards the active closed (high-affinity) form, also supported by kinetic studies [307]. However, this allosteric mechanism does not sufficiently account for the variable effects of GKAs upon enzyme kinetics. The allosteric activator site includes the connecting region I and the C-terminal helix, involved in propagation of the glucose-induced conformational change. It is therefore reasonable to hypothesize that structural changes are closely linked also to the allosteric modulation of GK activity, but presumably too small to be
seen in the crystal structures [313]. In that case, a fundamental question is how the allosteric effect/signal is communicated from the allosteric to the active site 20 Å away. Recently, Larion and Miller have proposed a model for the structural coupling of the two sites, involving a network of interconnected amino acids that may have significant impact on glucose binding and/or catalysis [193]. According to their model, helix 6 is involved in this site-to-site communication, with the glucose interacting residues N204 and D205 (catalytic base) located at its N-terminal end and the residues Y214 and Y215, forming hydrophobic interactions with the activator, at the C-terminal end of the helix [193].

5.3.2. Glucokinase regulatory protein (GKRP)

One of the endogenous modifiers of GK is the GKRP, an allosteric inhibitor that modulate GK activity in hepatocytes. GKRP binds preferentially to the super-open conformation, and sequesters the inhibited enzyme in the nucleus [15, 25, 203, 204]. The crystal structure of GKRP has not been solved, and therefore the current knowledge on the structural interface of the GK-GKRP interaction is not established. Several motifs of GK have been proposed to be involved, including two asparagine-leucine motifs, located in the hinge region near the substrate binding site (L58/N204) and in the large domain (L355/N350) [320]. Interestingly, L58 is part of an in silico predicted disordered region in GK [321]. Moreover, the L58/N204 motif is partially solvent exposed in the super-open conformation, but inaccessible in the closed glucose-bound form (data not shown), consistent with the finding that the interaction of GKRP with GK is competitive with glucose [203]. Recently, a common allosteric regulator region with non-overlapping binding sites has been proposed for GKAs and GKRP [274], that encompass different residues than the previous proposal. Two separate contact patches for GKRP (E51 and E52 & K140-L144 and M197) were proposed, also consistent with a previous study [322]. A concerted motion of the GKA and GKRP subdomains was proposed, in which glucose binding causes the two GKRP binding regions to separate (~ 20 Å → 40 Å), causing dissociation of the GK-GKRP complex, and opening of the GKA binding site [274, 309]. This model is consistent
with the view that binding of GKA and GKRP are mutually exclusive, and that GKA binding is glucose-dependent and GKRP inhibition of GK is competitive with glucose [274].

5.3.3. The bifunctional enzyme (PFK2/FBPase2)

In contrast to GKRP, the bifunctional enzyme is expressed in both pancreatic β-cells and hepatocytes, and it acts as an allosteric activator of GK [217-220]. The activation increases the $V_{\text{max}}$ value of GK, presumably by binding to and stabilizing its active closed form [218]. The interaction between GK and PFK2/FBPase2 is mediated by a binding motif in the bisphosphatase domain [217], whereas the interaction site on GK is currently not known.

5.3.4. Polyubiquitin

In Paper III, we have presented a potential novel mechanism for the allosteric regulation of GK. It was demonstrated that recombinant hGK interacts with and is allosterically activated by purified free pentaubiquitin chains (with an apparent $EC_{50}$ of 93 nM), and possibly also by unidentified polyubiquitinated proteins. The interaction with pentaubiquitin affected mainly $V_{\text{max}}$ (>40 % enhancement at ~100 nM Ub5), but with only a modest increase in the affinity for glucose. By 3D structural analysis we identified a classical UIM site in the highly mobile C-terminal part of GK including the C-terminal helix, and mutational removal of this site almost abolished the binding of free polyubiquitin to GST-hGK. Moreover, polyubiquitination of GK appeared to be coupled to this site. The UIM site (residues 442-464) interacts specifically with helix 6 in both the super-open and closed conformations. Interestingly, in the closed conformation three residues of the C-terminal helix (V452, V455 and A456) form one face of the GKA binding site, whereas residues M210, I211, Y214 and Y215 in helix 6 forms another face of the site. Thus, the proximity of the UIM and GKA site is striking. Moreover, the maximum stimulation of $V_{\text{max}}$ by pentaubiquitin was in the same order of magnitude as that reported for many GKAs (~1.5-fold) [9, 34, 323]. But
in contrast to GKAs, polyubiquitin binds preferentially to GST-hGK in the absence of glucose, explained by a high accessibility of key interacting residues of the UIM site in the 3D structure of its unliganded form, whereas the accessibility of some contact residues is lower in the glucose-bound form (Paper III). As far as the molecular mechanism of the catalytic activation is concerned, the binding at the UIM site may perturb the equilibria between conformational substates, as has been proposed for GKAs and GKRPs (see Section 5.2).

5.4. Catalytic mechanism

Enzyme kinetics and crystallographic studies have revealed that GK and other hexokinases share a common catalytic strategy that involves the initial activation of the 6'-OH group of glucose, followed by a nucleophilic attack on the γ-phosphate of ATP [152, 309, 324]. A conserved aspartate, whose side-chain carboxylate group is in close proximity to the 6'-OH group of glucose at the active site, serves as a general base catalyst, extracting a proton, thereby activating the 6'-oxygen to act as a nucleophile in the reaction. In GK, this catalytic residue is D205 [15, 152, 325], located in helix 6 of the L-domain and part of a classical ATP-binding motif (Connect 1 site) [154]. In the super-open conformation (PDB i.d. 1v4t), the D205 carboxylate is locked into a position outside the active site by forming a salt-bridge with the side-chain of R447 in the C-terminal helix (Paper I and [152]). In the glucose-bound closed conformer (PDB i.d. 1v4s), the D205 carboxylate group is rotated 177.7° about the Cα–Cβ bond to form a H-bond to glucose (Paper I). This repositioning of the catalytic base is coordinated with the glucose-induced closure of the active site cleft [15]. On binding MgATP and formation of the ternary complex, D205 forms a H-bond interaction also with Mg²⁺ (Figure 11), as part of the hexahedral coordination of the ion in the catalytic complex [152]. Moreover, K169 in the active site (activation) loop plays an essential role in both glucose and ATP binding as well as in catalysis (Paper I and [152, 325]). In the closed binary GK-glucose enzyme, K169 forms H-bond interaction with glucose, and hence, contributes to the stabilizing of this complex. In
Figure 11. The catalytic scaffold of GK. The figure highlights the most important H-bond interactions involved in the stabilization of the catalytic complex. See text for details. The figure is based on the coordinates given in [152].

In the crystal structure of the ternary complex, K169 interacts with the $\gamma$-phosphate of AMP-PNP (Figure 11), and is predicted to directly participate in glucose phosphorylation by acting as a general acid catalyst, providing a proton to protonate the $O_{3}^{\gamma}$ atom of the nucleotide [152, 325]. Moreover, K169N is a naturally occurring mutation in the GCK gene, associated with familial mild fasting hyperglycemia [80]. Consistent with its critical role in both ligand binding (glucose and ATP) and catalysis, the mutation of lysine results in a partial loss of glucose binding and an almost complete loss of catalytic activity (Paper I). N169 is not able to act as a general acid catalyst because its side-chain is incapable of providing a proton to ATP [325].

In the crystal structure of the ternary complex the 6'-OH group of glucose forms an important H-bond interaction with the $O_{3}^{\gamma}$ atom of AMP-PNP (Figure 11), and additional H-bonds between GK and its substrates stabilize the catalytic complex. T228 is a highly conserved residue at the active site of the hexokinase family of enzymes, and mutations of this residue (T228M) are associated with GCK-MODY in
the heterozygous state [46], and with GCK-PNDM in the homozygous state [32]. The catalytic activity of GK is also dependent upon Mg\(^{2+}\), which coordinate the \(\gamma\)- and \(\beta\)- phosphates of ATP (Figure 11). According to the crystal structure, the ion is held in a hexahedral coordination, directly interacting with the side chain of D205, and the oxygens of the AMP-PNP \(\beta\)-phosphate (O1 and O2) and \(\gamma\)-phosphate (O3), and indirectly with the side chains of S151 and D78, mediated through water molecules [152].

5.5. Molecular mechanisms of disease

Functional characterization of recombinant mutant enzymes has demonstrated that GK mutations can be divided into groups of different molecular mechanisms, or a combination:

1. **Mutations that affect one or more enzyme kinetic parameters** \((k_{\text{cat}}, [S]_{0.5} \text{ for glucose}, n_H, K_m \text{ for MgATP}^{2+})\). The majority of GCK mutations can be explained by an alteration in one or more kinetic parameters of the recombinant enzymes. This includes mutations that perturb the interactions with glucose/ATP or transmit structural changes to the substrate/cosubstrate binding site, and mutations that affect the activator (GKA) binding site or the connecting regions linking the L- and S-domains (Paper I and [65, 138, 146, 153, 325]). A subgroup of (1) is

2. **Mutations that affect the mechanism of GK activation**, either by promoting glucose binding to the apoenzyme (T65I, Y214C and A456V) or by facilitating enzyme isomerization to the active form (W99R, Y214C and V455M) [160]. It is possible that this form of structural perturbation, that alters the kinetic basis of enzyme activation, is applicable also to inactivating GK mutations, but with an opposite effect on catalytic efficiency. Some of these mutations also seem to cause a structural compaction and partial closure of the apoenzyme, which mimics the activated (closed) form of hGK [160, 274, 309].
3. Mutations that shift the equilibrium between conformational substates of GK. See Section 5.2 for details.

4. Mutations that interfere with the normal interactions of GK with known regulatory proteins. See Section 1.9 for details.

5. Mutations that cause an increased susceptibility to protein misfolding and destabilization. Enzyme kinetics alone cannot always explain the clinical and metabolic phenotypes of GK mutations. In fact, a number of mutations have been identified that co-segregate with GCK-MODY in the family, and are regarded as pathogenic, but when expressed as recombinant proteins they demonstrate near-normal or even slightly increased catalytic efficiency [65, 114]. Instead, a reduced protein stability can be observed in vitro (e.g. thermal instability, aggregation and increased susceptibility to limited proteolysis) and/or in vivo (e.g. oligomerization/aggregation, a susceptibility to inactivation by oxidation [326] and increased turnover). See Section 5.1 for details.

6. Mutations in the regulatory region of pancreatic GCK. Up to date, only a single mutation has been reported in the regulatory region of GCK. The mutation -71G>C, located in the β-cell promoter, was demonstrated to significantly reduce gene expression in vitro through a loss of regulation by the transcription factor Sp1 [327].

This classification highlights the variety of molecular mechanisms that contribute to the reduced (MODY, PNDM) or increased (HI) phosphorylation capacity of GK, and emphasizes the importance of combining biochemical, biophysical and structural approaches in the study of naturally occurring GCK mutations. Moreover, cell-biological approaches are warranted to fully characterize the mutant enzymes and to gain a deeper knowledge of the complex molecular regulatory mechanisms of GK.
6. FUTURE PERSPECTIVES

The 3D structures of hGK [15] represented the starting point of the present study (Papers I-III), with the main objective to study structure-function relationships and regulatory properties of WT hGK. The unique kinetic properties of WT GK studied in Papers I and II represent the basis for its physiological role as a glucose sensor in pancreatic β-cells. Further kinetic, biophysical and molecular dynamics studies are, however, needed to fully understand the basis of its positive kinetic cooperativity and the molecular mechanism of the glucose-induced conformational changes. Moreover, the extent of conformational heterogeneity of GK in its unliganded and complexed states is only partly understood (see Section 5.2).

The new insights obtained in the current study with respect to the function of WT GK constitute a valuable basis for further studies on disease-related mutant forms. Mutations in GCK are the second most frequent cause of MODY in Norway. Except during pregnancy, patients with GCK-MODY rarely need pharmacological treatment. Recognizing GCK-MODY in these patients is therefore important since a correct diagnosis may change management [75]. Moreover, identification and functional studies on naturally occurring GCK mutations have proven important to provide insight into the biochemical basis of glucose sensor regulation, and a variety of cellular and molecular regulatory mechanisms are involved in the control of GK activity and stability (summarized in Section 5.5). WT hGK is only marginally stable under physiological conditions (Paper II), and we recently demonstrated that the catalytically near-normal mutations S263P, G264S and R275C cause protein misfolding, aggregation and reduced protein stability when expressed in HEK293 and MIN6 β-cells (Negahdar et al, unpublished). The possibility that this also may induce a cellular stress in β-cells, with its low antioxidative defense system, should be studied. It has recently been demonstrated that some GCK-MODY associated mutants show an increased susceptibility towards oxidative stress and/or protein instability [326].

The discovery that small molecule glucokinase activators (GKAs) allosterically activate GK by increasing its substrate affinity and maximal catalytic rate, represent a
promising development in the treatment of T2D (www.ClinicalTrials.gov). In clinical trials the GKAs favorably influence glucose homeostasis by a dual action, affecting GK activity in both the liver and pancreatic \( \beta \)-cells [9, 34]. Endogenous compounds have been proposed to play a similar activating/regulatory role, but they remain to be discovered. Some GKAs might induce hypoglycemia, which may narrow the therapeutic window. In order to reduce this risk, current efforts are focused on partial GK agonists that reduce \( [S]_{0.5} \) for glucose moderately, i.e. mimicking mild activating GK mutations, and hepatoselective compounds [9, 34]. As summarized in Section 5.3, the molecular mechanisms of allosteric regulation of GK are complex and only partly understood, and studies on a possible cross-talk between the different allosteric effectors should have a high priority. In paper III, we demonstrated that polyubiquitin allosterically enhances the catalytic activity of recombinant GK. We also found that WT hGK is polyubiquitinated \textit{in vitro}, and evidence was presented that this may serve as a signal for proteasomal degradation of the newly synthesized protein. Cellular studies are warranted to examine whether these ubiquitin-mediated processes represent potential physiological regulatory mechanisms affecting the role of GK in glucose metabolism and insulin secretion. Moreover, a 3D structure of the GK-(Ub)\(_n\) complex would define the molecular mechanism of the catalytic activation.
References


77. Lindner TH, Cockburn BN & Bell GI (1999) Molecular genetics of MODY in Germany. Diabetologia 42, 121-123.


