Acta Biochim Biophys Sin 2011, 43: 618–629 | © The Author 2011. Published by ABBS Editorial Office in association with Oxford University Press on behalf of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. DOI: 10.1093/abbs/gmr049.
Advance Access Publication 23 June 2011

Original Article

Leptospira interrogans encodes an ROK family glucokinase involved in a cryptic glucose utilization pathway

Qing Zhang[1], Yunyi Zhang[1,2], Yi Zhong[1], Jun Ma[1], Nanqiu Peng[1], Xingjun Cao[3], Chen Yang[1], Rong Zeng[3], Xiaokui Guo[2], and Guoping Zhao[1,4*]

[1]Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China
[2]Department of Microbiology and Parasitology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
[3]Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[4]Department of Microbiology, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China
*Correspondence address. Tel: +86-21-54924002; Fax: +86-21-64837495; E-mail: gpzhao@sibs.ac.cn

Although *Leptospira interrogans* is unable to utilize glucose as its carbon/energy source, the *LA_1437* gene of *L. interrogans* serovar Lai potentially encodes a group III glucokinase (GLK). The *L. interrogans* GLK (LiGLK) heterogeneously expressed in *Escherichia coli* was purified and proved to be a homodimeric enzyme with its specific activity of 12.3 ± 0.6 U/mg·protein determined under an improved assay condition (pH 9.0, 50°C), 7.5-fold higher than that assayed under the previously used condition (pH 7.3, 25°C). The improved sensitivity allowed us to detect this enzymatic activity of $(5.0 \pm 0.6) \times 10^{-3}$ U/mg·protein in the crude extract of *L. interrogans* serovar Lai cultured in standard Ellinghausen–McCullough–Johnson–Harris medium. The k_{cat} and K_m values for D-glucose and ATP were similar to those of other group III GLKs, although the K_m value for ATP was slightly higher. Site-directed mutagenesis analysis targeting the conserved amino acid residues in the potential ATP-binding motif hinted that a proper array of Gly residues in the motif might be important for maintaining the conformation that was essential for its function. Gene expression profiling and quantitative proteomic data mining provided preliminary evidence for the absence of efficient systems involved in glucose transport and glycolysis that might account for the failure of glucose utilization in *L. interrogans*.

Keywords *Leptospira interrogans*; glucokinase; glucose utilization

Received: February 25, 2011 Accepted: April 18, 2011

Introduction

Species of the genus *Leptospira* are divided into two groups: the pathogenic *L. interrogans* sensu lato and the saprophytic *L. biflexa* sensu lato [1]. Previous studies have indicated that neither pathogenic nor saprophytic leptospires are able to utilize glucose [2]. They only use fatty acids as their main energy and carbon source when cultured *in vitro* [3].

Baseman and Cox failed to detect any of the hexokinase/glucokinase activity in crude cell extract of three representative strains of *Leptospira* comprising a water isolate and two pathogenic serotypes, *Pomona* and *Schueffneri*. Therefore, he proposed that the absence of hexokinase/glucokinase might account for the inability of *Leptospira* to utilize glucose [2]. In 2003, initial annotation of the genome of *L. interrogans* strain 56601 failed to identify the gene encoding hexokinase/glucokinase, which echoed the rationale accounting for incapable of utilizing glucose as carbon or energy source in *Leptospira* species [4]. Later, by comparative genomic analysis, Nascimento *et al.* identified gene *LA_1437* of strain 56601 potentially encoding a glucokinase (GLK; EC 2.7.1.2), which may function in the first step of glycolytic pathway as that of hexokinase [5]. A gene potentially encoding GLK was also annotated in the genome of *L. biflexa* strain Patoc 1 [6,7]. On the other hand, based on protein sequence similarity comparisons, Bulach *et al.* proposed that LA_1437 could more likely function as a regulatory protein rather than a GLK [8]. Therefore, experimental characterization of LA_1437 becomes critical in understanding the biology of *Leptospira* species.

GLK is responsible for catalyzing the ATP/ADP-dependent phosphorylation of glucose to form glucose 6-phosphate. Based on their protein sequence variations, microbial GLKs can be divided into three families. The group I is ADP-dependent GLK found in archaea [9–12]. On the other hand, both groups II and III GLKs

are ATP dependent. The difference is that group III GLKs possess the so-called ROK (**r**epressor, **o**pen reading frame, sugar **k**inase) sequence signature [13,14]. A previous study also found a conserved CXCGX(2)GCXE motif in group III proteins, which is important for the enzymatic activity of GLKs [15].

In this study, we experimentally characterized the GLK of *L. interrogans* encoded by the gene *LA_1437* of strain 56601. It showed similar biochemical properties as other bacterial group III GLKs *in vitro* and its expression in *L. interrogans* cells under laboratory culture condition was detected. Site-directed mutagenesis analysis improved our knowledge about the sequence content of the apparently conserved ATP-binding domain. Combining the gene expression and proteomic profiling data of *L. interrogans* together, we formulated a hypothesis to account for the failure of glucose utilization of the bacterium.

Materials and Methods

Bacterial strains, plasmids and culture conditions

The virulent strain 56601 of *L. interrogans* serogroup Icterohaemorrhagiae serovar Lai was grown in the Ellinghausen–McCullough–Johnson–Harris (EMJH) liquid medium [16,17] at 28°C. *Escherichia coli* was routinely grown in Luria Bertani (LB) medium at 37°C [18] except where indicated otherwise. Antibiotics, when required, were used at the following concentrations: 100 μg/ml of ampicillin and 50 μg/ml of kanamycin.

Bioinformatic analysis

To detect the candidate gene(s) possibly encoding hexokinase or glucokinase, we used all of the *in silico*-predicted 3702 protein coding sequences (CDSs) of the updated annotation of the Lai serovar type strain 56601 genome composed of chromosome I (CI) and chromosome II (CII) (GenBank accession number AE010300 for CI and AE010301 for CII) [19] to search against the local NCBI non-redundant database [20] and KEGG (Kyoto Encyclopedia of Genes and Genomes) database (BLASTP search, $e < 10^{-3}$) [21]. Multiple sequences were compared using the software BioEdit (Tom Hall, Carlsbad, USA). Conserved protein structural domains and motifs were searched in the NCBI conserved domains database (CDD). To analyze the phylogeny of ROK protein sequences in relation to their domain characterizations, we blasted the Swiss-Prot [22] database using LA_1437 as the query and the resulted protein subjects for further analysis were selected based on the criteria of alignment length >250 aa; identity >25%; and subject protein length <400 aa. The phylogeny tree was built based on Neighbor-Joining method using the MEGA4 software [23], and the reliability of each branch was tested by 1000 bootstrap replications. The characteristic domains/motifs were identified by vision observation based on sequence homology alignment.

DNA manipulation

Genomic DNA of *L. interrogans* was extracted using Bacteria DNA Mini Kit (Watsonbiot, Shanghai, China). Polymerase chin reaction (PCR) fragments were purified from agarose gel with QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Restriction endonucleases were purchased from local suppliers and were used according to their specifications. Plasmids DNA from *E. coli* were prepared using the Plasmid Mini Kit (Tiangen, Beijing, China).

Overexpression and purification of recombinant LiGLK protein

LA_1437 gene was amplified using *Pfu* polymerase. The primers designed with *Nde*I and *Not*I restriction sites incorporated are: P1, 5′-GGCCAGCATATGAAAATGAAATC CTTTCTTG-3′ (*Nde*I, underlined) and P2, 5′-AAAAAA GCGGCCGCTCATTCTAAATTCTCCA-3′ (*Not*I, underlined). The PCR product was ligated to the pET28b (Novagen, Gibbstown, USA) after digestion with *Nde*I and *Not*I, creating the plasmid pET-LA_1437. This plasmid was introduced into *E. coli* BL21 (DE3) [24]. The transformants were cultivated at 37°C and induced at 30°C by 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation for 10 min at 6000 *g* at 4°C and the supernatant was removed.

Cell pellets were resuspended in a lysis buffer [50 mM Tris-HCl, pH 7.5, 300 mM NaCl and 1 mM phenylmethylsulfonyl fluoride (PMSF)] and disrupted by ultrasonication. Cell debris and unbroken cells were removed by centrifugation for 30 min at 13,000 *g* at 4°C. The recombinant protein was purified using Ni^{2+}-NTA (nitrilotriacetic acid) super flow column (Qiagen). The purity of the purified protein was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 12% polyacrylamide gels followed by staining with Coomassie brilliant blue R-250 according to standard procedures [25]. Protein concentrations were determined by the method of Bradford [26] using bovine serum albumin to generate a standard curve.

Analysis of the LiGLK conformation with gel filtration chromatography

Purified LiGLK was applied to gel filtration chromatography to measure its oligomeric state in solution. The analysis was carried out on an AKTA FPLC workstation (GE Healthcare, Wisconsin, USA). One hundred and twenty microliters of the purified LiGLK samples (3 mg/ml) were loaded into a 200-μl sample loop and injected into a Superdex™ 200 10/300 GL chromatographic

separation column (GE Healthcare). The High Molecular Weight (mass) Calibration Kit (GE Healthcare) was used to calibrate the molecular mass of LiGLK. The apparent molecular mass of the protein sample was calculated according to the protocol provided in this kit.

Glucokinase activity assay

The glucokinase activity was spectrophotometrically measured by monitoring the reduction of NADP in a glucose-6-phosphate dehydrogenase (G6PDH) coupled reaction [15,27]. Both the purified LiGLK protein and *L. interrogans* crude cell extract were used to assay the glucokinase activity. The standard assay mixture (500 μl) contains 50 mM Tris-HCl (pH 9.0), 10 mM D-glucose, 5 mM ATP and 10 mM $MgCl_2$. After pre-incubation at 50°C for 30 min, the reaction was started with an aliquot of appropriate amount of either the purified LiGLK protein (1–2.5 μg) or *L. interrogans* crude cell extract (2–10 mg). The mixture was then incubated at 50°C for 7 min, and the reaction was stopped by adding EDTA to a final concentration of 10 mM and cooling on ice for 2 min. Then 2 U of G6PDH (glucose-6-phosphate dehydrogenase from baker's yeast, purchased from Sigma, St. Louis, USA) and $NADP^+$ (final concentration of 1 mM, purchased from Sigma) was added to the mixture with 50 mM Tris-HCl buffer (pH 7.5) supplemented to a final volume of 1 ml and then incubated at 30°C for 30 min to ensure the total oxidation of glucose 6-phosphate to generate equal molar of NADPH measured by the optical density at 340 nm. The GLK activity was calculated by dividing the total μmol of NADPH generated by complete consumption of glucose-6-phosphate in the G6PDH catalyzed second reaction by the time used for the first reaction and the unit of the GLK activity was defined as μmol of $NADP^+$ consumed per minute (first reaction) as measured by the coupled reaction methodology [27]. This procedure was used as the standard assay system for all enzymatic measurements, except as otherwise stated. Besides, we also used another condition (pH 7.3, 25°C), which was used by Baseman and Cox to compare the enzymatic activity assays [2].

Biochemical properties of LiGLK

In order to determine the optimum reaction temperature, the enzymatic activity assay was carried out under different temperatures ranging from 4 to 60°C. For the determination of thermal stability, the purified enzyme was pre-incubated for 30 min at various temperatures from 4 to 60°C. After cooling down on ice for 10 min, the remaining enzymatic activity was assayed using a standard assay condition at 50°C.

The optimum pH was determined at 50°C in 50 mM citric acid buffer covering a pH range of 3.6–6.0, 50 mM Tris-HCl buffer covering a pH range of 7.0–9.1 and 50 mM CAPS buffer covering a pH range of 9.7–11.0. All the pH levels are measured at 50°C. The enzyme activity was tested using a standard assay system at 50°C. In order to determine the pH stability, the purified enzyme was exposed to different pH buffer at room temperature for 1 h and then supernatant was assayed using a standard assay system at pH 9.0/50°C after proper centrifugation.

To test the specificity of LiGLK for sugars, the substrate was substituted with fructose, mannose and galactose. ATP is replaced by ADP to test the phosphoryl donor. The enzyme activities were defined as above.

Kinetic studies of LiGLK

Kinetic studies for glucokinase were carried out in a standard assay system described above. To determine the effect of glucose, the concentration of ATP was fixed at 10 mM (about 10 folds of K_m) and the concentration of glucose was varied from 0.05 to 3 mM. To determine the effect of ATP, the concentration of glucose was fixed at 10 mM (about 25 folds of K_m) and the concentration of ATP was varied from 0.2 to 12 mM. All of the kinetic measurements were repeated at least twice under the same conditions. The kinetic data were analyzed using Prism 4.0 for Windows (GraphPad) and the parameters (V_{max}, K_m and k_{cat}) were calculated from the Michaelis–Menten plot.

Glucokinase mutants

Two single mutants A11G, G12T, and a double mutant A11G–G12T of the LiGLK were introduced in the cloned *LA_1437* gene (pET-LA_1437) employing the QuickChange® site-directed mutagenesis kit (Stratagene, Cedar Creek, USA). The constructed plasmids containing the corresponding mutated genes were designated pET-LA_1437 [A11G], pET-LA_1437 [G11T] and pET-LA_1437 [A11G–G12T], respectively. All of the mutations were confirmed by sequencing. The mutant clones were transformed into *E. coli* BL21 (DE3) individually for protein expression and purification as that of the wide-type enzyme described above.

Microarray and quantitative proteomic data mining

The potential glucose transmembrane transporters were predicted by TransportDB [28]. The genes encoding metabolic enzymes involved in the glycolysis pathway were identified by homologous search against KEGG database. All these candidate genes were subsequently analyzed for their expression and translation profiles by using our microarray [29] and LC-MS/MS [19] data. The average expression values for each gene detected by microarray were calculated from the data derived from three independent biological samples. In mass spectrometry analysis of protein profiling, all the available MS/MS spectra were searched

against the database containing the candidate glucose-utilizing proteins using the TurboSEQUEST program and the protein abundance were evaluated by calculating the APEX-index [30] using data derived from two independent biological samples.

Results

Bioinformatic analysis indicated that LA_1437 encoded a bacterial group III GLK

Aligning all the CDSs of *L. interrogans* strain 56601 with the local NCBI NR database [20] and the KEGG database [21], we found that, as indicated by Nascimento *et al.* [5], *LA_1437* was the only candidate gene which likely encodes a potential sugar kinase. By searching the Swiss-Prot database [22] employing the same set of parameters, LA_1437 protein exhibited a great sequence similarity to the group III GLK with ROK family protein sequence signature and a conserved CXCGX(2)GCXE motif. On the other hand, it is clearly distinct from that of the ROK family regulators. It is particularly significant that the characteristic N-terminal DNA binding domain is missing in LA_1437 but is exhibited in all of the seven homologous ROK family regulators (**Fig. 1**). Therefore, LA_1437 protein is designated *L. interrogans* glucose kinase (LiGLK).

Biochemical assay for glucose kinase activity of the heterogeneously purified LiGLK and detection of its activity in the *in vitro* cultured *L. interrogans* cells

The LiGLK heterogeneously expressed in *E. coli* was purified and its molecular mass was determined to be about 35 kDa [**Fig. 2(A)**], which is consistent with the molecular weight (34,084 Da) calculated from the protein sequence including the His_6 tag residues. Gel filtration chromatographic analysis showed that the heterogeneously expressed and purified LiGLK is a homodimeric protein with its apparent native molecular weight calculated to be 67.143 kDa [**Fig. 2(B)**].

Several assay conditions that may affect the enzymatic activity of LiGLK were tested. Although the thermal stability analysis showed that with a 30-min pre-incubation condition, the enzyme was stable only up to 42°C and about 20% of the activity was lost at 50°C [**Fig. 3(B)**], the enzyme was stable and presented its most optimal activity at 50°C as long as the incubation time was <15 min [**Fig. 3(A,C)**]. Therefore, the condition of 50°C with 7 min reaction was adapted for the standard assay.

Three types of buffer were used to test the pH dependence of the enzyme. As shown in **Fig. 3(D)**, LiGLK exhibited its maximal GLK activity at pH 9.0 in Tris-HCl buffer. The pH stability analysis showed that the enzyme was stable under the pH values ranging from 5.0 to 10.0. However, pre-incubating the enzyme at pH 11.0 for 1 h, the loss of activity was obvious [**Fig. 3(E)**].

The LiGLK was shown to be strictly specific to D-glucose as its substrate and exhibited no activity against other sugars, such as fructose, galactose or mannose [**Fig. 3(F)**]. Replacement of ATP by ADP revealed undetectable activity (data not show), which means that the LiGLK only uses ATP rather than ADP as the phosphoryl donor.

The enzyme-specific activities were assayed with purified LiGLK or *L. interrogans* strain 56601 crude extract under either of the two conditions, i.e. pH 9.0, 50°C defined by this work or pH 7.3, 25°C defined by Baseman and Cox [2]. The ATP-dependent GLK-specific activity of the purified enzyme was 12.3 ± 0.6 U/mg·protein and 1.6 ± 0.2 U/mg·protein, respectively [**Fig. 3(G)**]. The LiGLK-specific activity in the crude cell extract of *L. interrogans* strain 56601 [**Fig. 3(H)**] was $(5.0 \pm 0.6) \times 10^{-3}$ U/mg· protein under the pH 9.0 at 50°C, while no GLK activity was detected under the assay condition as previously illustrated [2].

Kinetics analysis of LiGLK

The kinetic parameters of LiGLK were determined employing the purified enzyme assayed under our optimized condition. The k_{cat} values for glucose and ATP were 17.8 ± 1.1 and 17.5 ± 1.4 s^{-1}, respectively. The K_m values for glucose and ATP were 430 ± 79 and 1011 ± 230 μM, respectively (**Table 1**).

Mutational analysis for the potential ATP-binding motif

In group III GLKs, the ATP-binding motif D[I/L/V]GG[S/T] was known to be located near the N-terminus [31]. We noticed that LiGLK possessed a potential ATP-binding motif, 8DIGAG12 (**Fig. 1**), which differed from the proposed conserved ATP-binding motif with the second Gly of the conserved 'GG' being replaced by an Ala while the last 'S/T' residue being replaced by a Gly.

Single mutations of A11G, G12T, and a double mutation of A11G–G12T were constructed in the LiGLK. All three purified mutated LiGLKs were soluble and, therefore, their kinetic properties were determined and compared with those of the wild type. As shown in **Table 1**, the kinetic data indicated that both the G12T single mutant and the double mutant only resulted in minor variations on K_m and k_{cat}. The A11G mutant, on the contrary, showed a decrease in the affinity for LiGLK toward ATP.

The LiGLK can be detected by MS with low abundance in *L. interrogans*

MS/MS technology-based proteomic analysis employing the crude cell extracts of *L. interrogans* cultured in EMJH

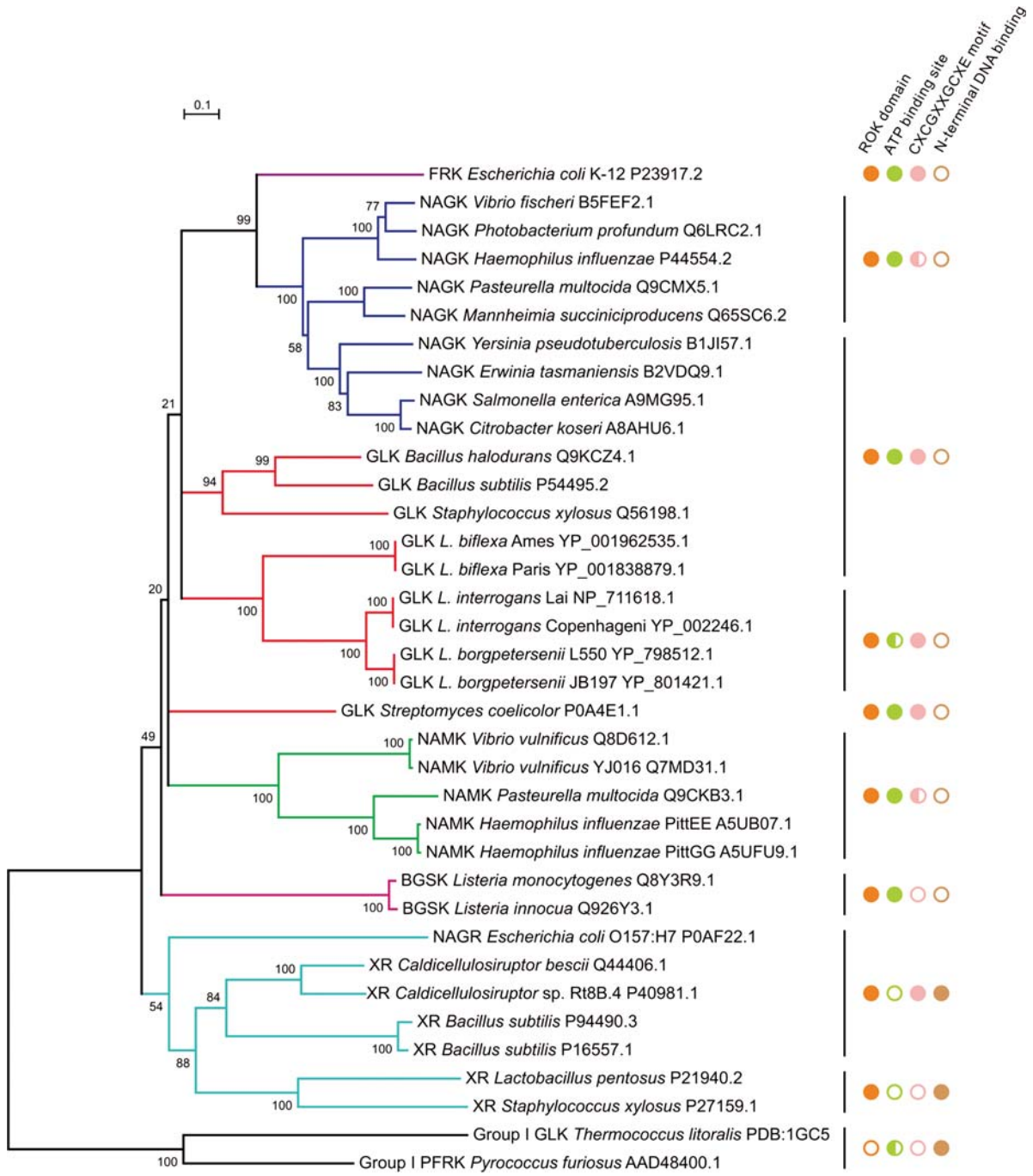

Figure 1 Phylogenetic analysis of LiGLK and its homologous proteins collected in Swiss-Prot database The abbreviations for the proteins are: GLK, glucokinase; NAMK, N-acetylmannosamine kinase; NAGK, N-acetyl-D-glucosamine kinase; BGSK, β-glucoside kinase; FRK, fructokinase; PFRK, phosphofructokinase; XR, xylose operon repressor; NAGR, N-acetylglucosamine operon repressor. The sequence accession numbers were labeled at the end of each protein. The domain information for each protein was indicated by the color-filled circles: orange, ROK-conserved sequence; green, ATP-binding site; pink, CXCGXXGCXE motif; brown, N-terminal residue of repressor. Hollow circles represent the domains do not exist in the corresponding proteins, whereas the half solid circles represent the cases where conserved amino acid(s) were replaced by others. Examples of group I sugar kinases were used as the out group of the phylogenetic tree. Bootstrap values are labeled on the major nodes of the branches.

medium at 28°C [19] totally detected 15 peptides from LA_1437, which can be grouped into six unique peptides (**Table 2**) covering 33.89% of the protein coding region. This result confirmed that LiGLK was indeed expressed in the laboratory culture. Furthermore, by using the simplified APEX-index approach [30] to evaluate the protein quantity, the abundance for LiGLK is 0.1419. It is a relative low value comparing with the average abundance (1.0877) for whole cell extracts. The same strategy was also used to calculate other enzymes involved in glycolysis pathway as well as the candidate glucose transporters predicted by TransportDB [28]. Notably, we could not find the gene encoding

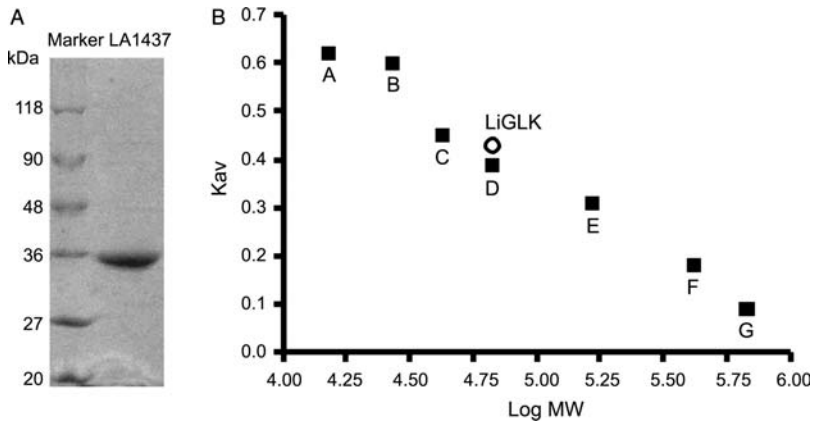

Figure 2 Purified LiGLK protein and the determination of its native molecular weight (A) SDS-PAGE of purified LiGLK was shown to have molecular weight of 35 kDa, with the standard molecular size marker. (B) The native molecular weight of purified LiGLK was determined by gel filtration chromatography on a platform of Superdex^{TM} 200. The marker proteins used for molecular weight calibration are as follows: A, thymoglobulin (669 kDa); B, ferritin (440 kDa); C, aldolase (158 kDa); D, conalbumin (75 kDa); E, ovalbumin (43 kDa); F, chymotrypsinogen A (25 kDa); G, ribonuclease A (13.7 kDa). The $MW_{LiGLK}=10^{4.827}=67{,}143$ Da. MW, molecular weight.

6-phosphofructokinase either. However, the diphosphate-fructose-6-phosphate 1-phosphotransferase (LB_111), which can also convert D-fructose 6-phosphate to D-fructose 1,6-biphosphate in certain species [32], existed in the *L. interrogans* genome. For each protein involved in glucose utilization, the calculated abundance and the expression data from previous microarray result [29] were listed in **Table 3**. It demonstrated that the transporter candidates and the genes encoding enzymes responsible for the first three steps (LA_1437, glucose-6-phosphate isomerase LA_3888 and LB_111) as well as the last step (pyruvate kinase LA_2924) in glycolysis were expressed low compared with the average amount. Other genes involved in this pathway are functionally shared with gluconeogenesis, and thus maintained relatively high-expression levels (**Table 3**).

Discussion

As designated, the ROK family proteins include not only sugar kinases but also transcription repressors and yet uncharacterized open reading frames. Although these three kinds of functionally distinct proteins share high similarity in their protein sequences [33], the repressors so far have always been found to bear the characteristic N-terminal helix-turn-helix motif related to DNA binding [34–36] while the GLKs should bear an ATP-binding motif [31]. As re-annotated by Bulach *et al.* [8] and in our recently revised NCBI genome sequence registry (GenBank accession number AE010300) [19], the N-terminal amino acid sequence of LA_1437 of *L. interrogans* strain 56601 does not contain the DNA-binding motif but has the ATP-binding motif (**Fig. 1**), which is validated by three peptides detected in our MS data (**Table 2**). Therefore, we infer that LA_1437 is more likely an enzyme rather than a regulatory protein albeit it requires further investigation about whether it might simultaneously bear the functions of both.

LiGLK is a homodimeric protein (**Fig. 2**), same as most of the ATP-GLKs of group III except that from *Bacillus subtilis*, which occurs in both monomeric and homodimeric forms or those from *Streptomyces coelicolor* and *Streptomyces peucetius* var. *caesius*, which constitute tetrameric forms [15,27,37]. The spectrum of sugar substrates for the bacterial GLK was known to be narrow [27], which was also shown in the case of LiGLK [**Fig. 3(F)**]. It used only ATP rather than ADP as the phosphoryl donor (data not shown), consistent with the known properties of the bacterial GLKs [27].

The LiGLK showed some special biochemical properties. In addition to its optimal reaction pH of 9.0 in Tris-HCl buffer, at the highest pH margin for GLKs from other bacteria is ranging from 6.2 to 9.0 [27,37,38]; its optimum reaction temperature (about 50°C) is higher than most of the GLKs from mesophilic bacteria, ranging from 30 to 42°C [37–39]. Although higher optimal temperature (60–93°C) was detected for those from thermophilic bacteria [40–42], high sequence similarity between LA1437 and the GLKs from mesophilic bacteria (data not shown) was observed, inferring that this property merely sets up a new top margin of the optimal reaction temperature for this kind of GLKs evolved via adapting to the living environment of the bacteria.

Previously, the GLK activity in *Leptospira* crude extracts was undetectable under the assay condition of pH 7.3, 25°C with direct coupling to the G6PDH-catalyzed oxidation–reduction reaction [2]. However, we showed that, it was about 7.5-fold decrease of the sensitivity comparing with that assayed under our improved condition (pH 9.0, 50°C) [**Fig. 3(G)**]. Although the temperature we used for assay, 50°C is within the sharp transition region of the

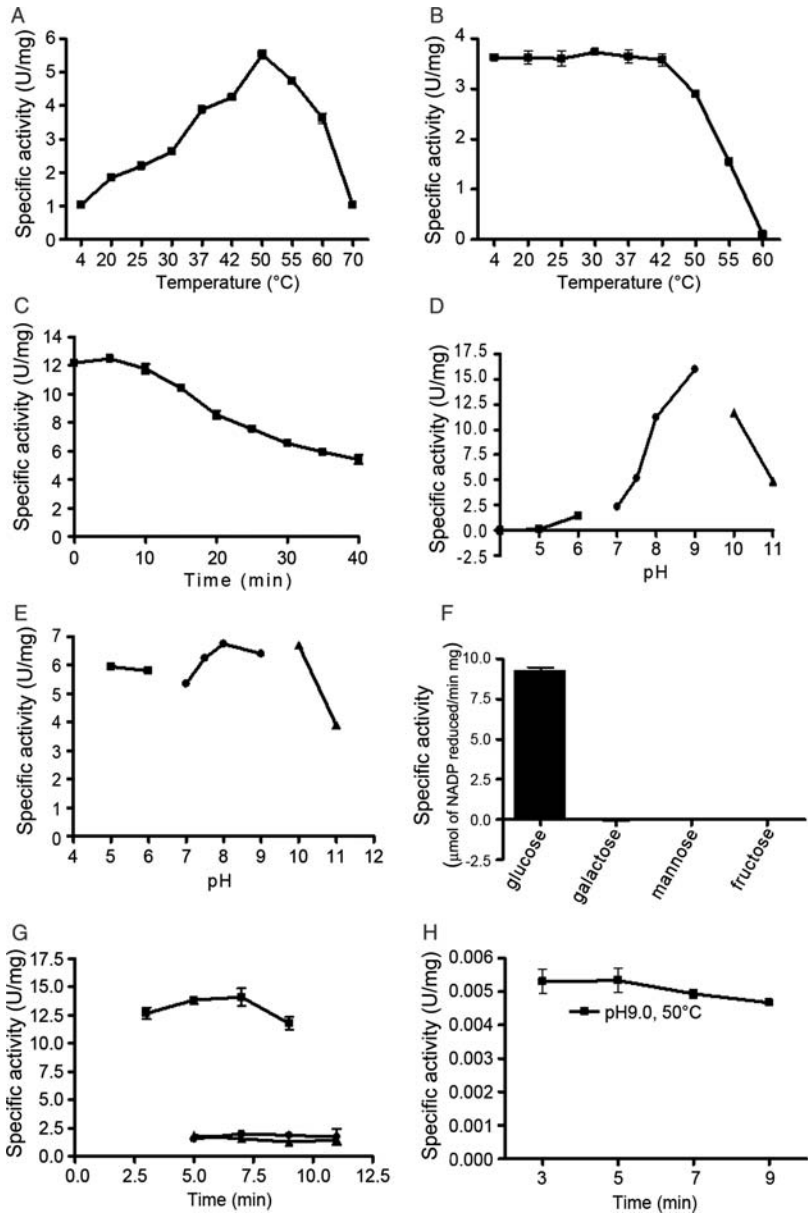

Figure 3 Biochemical characteristics of LiGLK All of the assays were carried out as described in Materials and Methods. (A) Optimum reaction temperature assayed under the standard condition with varied reaction temperatures. (B) Thermal stability under different temperature with 30 min pre-incubation. (C) Thermal stability at 50°C with different length of pre-incubation periods as indicated. (D) Optimum reaction pH assayed under three buffer systems (■: citric acid buffer; •: Tris-HCl buffer; ▲: CAPS buffer). (E) pH stability assayed under three buffer systems (■: citric acid buffer; •: Tris-HCl buffer; ▲: CAPS buffer). (F) Substrates specificity tests. (G) Specific activity comparison of purified LiGLK under the two assay conditions (■: 50°C, pH 9.0, two-step assay; •: 25°C, pH 7.3, two-step assay; ▲: 25°C, pH 7.3, one-step assay.). (H) Comparison of specific activities of LiGLK in *L. interrogans* crude extracts under the two assay conditions as illustrated.

temperature stability profile. We noticed that the enzyme activity assayed under the condition of pH 9.0, 50°C within 7 min was generally stable with much higher activity than that assayed under the condition of pH 7.3, 25°C. We should emphasize that the 50°C optimal reaction temperature for LiGLK, observed for both purified recombinant protein and the crude cell extracts of *L. interrogans* cultured under laboratory condition, was achieved by employing a shortened reaction time to compromise its thermal instability at 50°C. Therefore, it is unlikely related to better folding of the recombinant protein at 50°C but is simply an improved condition to enhance the sensitivity of the assay. In addition, because of the G6PDH used in this experiment not being stable at 50°C, we performed the coupled oxidation reaction after the GLK-catalyzed phosphorylation was terminated. On careful comparison of the two methods under the 25°C condition, no significant difference could be observed [**Fig. 3(G)**]. Thus, there is no surprise that this kind of two-step reaction methodology has been widely accepted for GLK assays [27]. The

Table 1 Kinetic data of wild-type and mutant LiGLKs

		LiGLK	LiGLK-A11G	LiGLK-G12T	LiGLK-A11G-G12T
ATP	K_m (μM)	1011 ± 229	1572 ± 347*	806 ± 142**	807 ± 173**
	k_{cat} (s^{-1})	17.5 ± 1.4	14.1 ± 1.3*	18.4 ± 1.6**	17.5 ± 1.2**
Glucose	K_m (μM)	430 ± 79	483 ± 87**	531 ± 102**	504 ± 134**
	k_{cat} (s^{-1})	17.8 ± 1.1	16.4 ± 1.0**	18.6 ± 1.4**	19.3 ± 2.0**

The kinetic analysis experiment was described in Materials and Methods. The kinetic parameters were calculated with the computer program Graphpad Prism 4.0 and the values shown are mean ± SD. The statistic significance of the differences in the values of kinetic parameters between the wild-type and the mutated enzymes are indicated by P values. *$P < 0.05$; **$P > 0.05$.

Table 2 The information of MS-detected peptides for LiGLK

Sequence[a]	MH^{+}[b]	Charge[c]	Xcorr[d]	DeltaCn[e]	Sp[f]	RSp[g]	Ions[h]	Position[i]
K.ASLIDSNGNVLK.S*	1231.381	2	2.0539	0.1404	692.5	1	14|22	15
K.GSPNLVILTLGTGLGGGWVYQGK.L	2288.6307	2	2.8124	0.6146	228.2	1	15|44	122
K.GSPNLVILTLGTGLGGGWVYQGK.L	2288.6307	2	1.9554	0.2141	87.4	20	9|44	122
K.QFLDSICDIVSEMK.N*	1685.9142	2	1.9466	0.2468	258.1	1	11|26	40
K.QFLDSICDIVSEMK.N*	1685.9142	2	2.208	0.2523	249.1	2	10|26	40
K.TASILLNEGIDALAQLCR.N	1959.2272	2	4.2836	0.7031	1819.1	1	22|34	213
K.TASILLNEGIDALAQLCR.N	1959.2272	2	3.3339	0.6302	1067.8	1	20|34	213
K.TASILLNEGIDALAQLCR.N	1959.2272	3	4.5751	0.5742	2020.2	1	33|68	213
K.TASILLNEGIDALAQLCR.N	1959.2272	2	3.1942	0.575	653.9	1	16|34	213
K.TGEVLDSAEEFFEK.T	1601.6925	2	3.4888	0.6257	1416.6	1	19|26	193
K.TGEVLDSAEEFFEK.T	1601.6925	2	3.7179	0.5478	1740.6	1	19|26	193
K.TGEVLDSAEEFFEK.T	1601.6925	2	1.9309	0.2684	379.4	2	12|26	193
K.TGEVLDSAEEFFEK.T	1601.6925	2	2.2277	0.389	248.6	4	12|26	193
K.TGEVLDSAEEFFEK.T	1601.6925	2	1.9663	0.2227	81.6	49	9|26	193
R.ELIFPIFR.T	1035.2633	2	2.1056	0.333	375.4	1	10|14	264

[a]Peptides labeled with asterisks are those mapped onto the ATP-binding motif.
[b]MH^{+}, the molecular weight of the peptide.
[c]Charge, the charge state of the parent ion.
[d]Xcorr, the cross-correlation score for candidate peptides.
[e]DeltaCn, [Xcorr(top hit)—Xcorr(n)]/Xcorr(top hit).
[f]Sp, the preliminary score for each candidate peptide.
[g]RSp, the preliminary score rank for each candidate peptide.
[h]Ions, the number of ions for each peptide.
[i]Position, the position of the first amino acid of the peptide on the LiGLK protein.

LiGLK-specific activity of $(5.0 \pm 0.6) \times 10^{-3}$ U/mg· protein in the crude protein was estimated to account for ~0.4‰ of the total crude extract proteins. We can further estimate that, if assayed under the condition of pH 7.3, 25°C, the LiGLK-specific activity in crude cell extract would be $(6.5 \pm 0.6) \times 10^{-4}$ U/mg· protein. It is so close to the lower detection limit (2×10^{-4} U/mg· protein) provided by Baseman and Cox [2] that might account for their failure of detecting any glucokinase activity in the cell extract.

Thus, the ATP-dependent glucose kinase activity of *L. interrogans* serovar Lai strain 56601 (LiGLK) was detected, not only for the purified exogenously expressed enzyme but also in the crude cell extracts prepared from the EMJH culture of the bacterium. In addition, we also detected the glucokinase activities in the crude cell extracts from both the pathogenic *Leptospira* species (such as *L. javanica*, *L. canicola*, *L. pyrogenes* and *L. grippotyphosa*, etc.) and the saprophytic *Leptospira* species (such as *L. semaranga*, *L. pulpudeva* and *L. anhui*, etc.) under our assay conditions (data not shown).

MS/MS technology-based proteomic analysis employing the crude cell extracts of *L. interrogans* strain 56601 cultured in EMJH medium at 28°C detected 15 peptides

Table 3 Semi-quantitative assessment of enzymes and the corresponding mRNAs of the encoding genes involved in the glucose utilizing pathway via microarray and mass spectrometry data mining

	Gene ID	Annotation	EC number	Gene	COGs	Microarray data[a]	MS data[a]
Candidates for Glc transport	LA0694	Na^+/glucose symporter				349.57	≈ 0
	LA2133	Sodium:solute symporter family protein		*putP*	COG0591ER	502.88	≈ 0
	LB245	Probable sodium:solute symporter		*putP*	COG0591ER	294.77	≈ 0
Glycolysis and/or Gluconeogenesis	LA1437	ROK family protein/glucose kinase	EC:2.7.1.2	*glk*	COG1940KG	419.87	0.1419
	LA3888	Glucose-6-phosphate isomerase	EC:5.3.1.9	*pgi*	COG0166G	NA	0.1950
	LB111[b]	Diphosphate-fructose-6-phosphate 1-phosphotransferase	EC:2.7.1.90	*pfk*	COG0205G	384.87	0.1073
	LA1532	Fructose-bisphosphate aldolase class I	EC:4.1.2.13	*fbaB*	COG1830G	3517.33	6.6277
	LA1696	Triose phosphate isomerase	EC:5.3.1.1	*tpiA*	COG0149G	445.77	1.6582
	LA1704	Glyceraldehyde 3-phosphate dehydrogenase	EC:1.2.1.12	*gapA*	COG0057G	1971.67	4.8279
	LA1703	Phosphoglycerate kinase	EC:2.7.2.3	*pgk*	COG0126G	1873.00	2.0537
	LA0439	Phosphoglycerate mutase	EC:5.4.2.1	*gpmI*	COG0696G	566.27	0.2707
	LA1951	Phosphopyruvate hydratase/enolase	EC:4.2.1.11	*eno*	COG0148G	2426.33	3.2029
	LA2924	Pyruvate kinase	EC:2.7.1.40	*pykF*	COG0469G	NA	0.3855
	LA2226	Fructose-1,6-bisphosphatase	EC:3.1.3.11	*fbp*	COG0158G	376.40	1.6549
	LA0251	Phosphoenolpyruvate carboxykinase	EC:4.1.1.49	*pck*	COG1866C	4666.67	4.8346
TCA	LA0671	Citrate synthase	EC:2.3.3.1	*gltA*	COG0372C	1200.50	1.3157
	LA0790	Citrate synthase	EC:2.3.3.1	*gltA*	COG0372C	1735.50	3.3318

[a]The numbers in microarray and MS data column are the transcriptional signals and the calculated protein abundance, respectively. The average expression signal of all the genes on the microarray was 817.53, and the average protein abundance of whole cell extracts was 1.0877. NA, not available due to lack of oligonucleotide probes for the corresponding genes on the microarray. Citrate synthase, the rate-limited enzyme in TCA cycle, were used as the positive control for expression levels.

[b]The 6-phosphofructokinase (EC 2.7.1.11) is absent in *Leptospira interrogans*. The conversion of fructose 6-phosphate to fructose 1,6-bisphosphate is presumably catalyzed by diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90).

derived from the LA_1437 protein (**Table 2**). This also demonstrates that LiGLK is expressed in *L. interrogans* strain 56601 in the laboratory culture.

The k_{cat} values for glucose and ATP of the LiGLK (**Table 1**) were similar to other group III GLKs [27]. The K_m values of LiGLK for glucose were also similar to those of other GLKs in group III. However, the K_m value for ATP (1.0 mM) was slightly higher than that reported for other group III GLKs (0.13–0.78 mM) [15,37,38,43–46], indicating a weaker affinity of LiGLK toward ATP than that of other members of the family. We noticed that LiGLK possessed a potential ATP-binding motif: DIGAG (**Fig. 1**), which is different from the conserved ATP-binding motif D[I/L/V]GG[S/T] as the DIGGT sequence identified in the *E. coli* group III GLK, located in the conserved β-strand-loop-β-strand motif with the conserved continuous two Gly residues contributing to the formation of a loop and to form the main chain hydrogen bonds with the ATP phosphates [13]. We thus hypothesized that the weak affinity of LiGLK toward ATP might result from the variation of the amino acid residues in this proposed ATP-binding motif. The kinetic properties of three recombinant enzymes with single or double mutations to match with the consensus sequence in the motif were compared with that of the wild-type enzyme (**Table 1**). It was unexpected that both the G12T single mutant and the A11G–G12T double mutant resulted in slight decrease of their K_m values without statistical significance, while the A11G mutant showed a significant decrease in the affinity for LiGLK toward ATP (**Table 1**). In contrast to other variants with one or two Gly residues in the motifs, including the wild-type LiGLK (DIGAG), the G12T mutant, and the A11G–G12T double mutant (DIGGT, identical to the ATP-binding motif of the *E. coli* GLK), there were three glycine residues in a row in the A11G mutant motif. It has been known that the glycine-rich sequences of protein kinases participate in nucleotide binding, substrate recognition, and enzyme catalysis [13,47]. Although the glycine residues were shown to be very important for the recognition or binding of the phosphoryl moiety [47], the number of glycine residues critical for maintaining the conformation of the motif was

never studied in the group III GLKs. In the study of the homodimeric protein Rop from *E. coli* with typical helix–loop–helix motif, Nagi *et al.* [48] noticed that when the native 'Asp–Ala' loop was mutated to 'Gly–Gly', it gained much more flexibility that could make the transition state more accessible and lower the energy barrier for the protein unfolding and folding processes. However, after the loop length was systematically increased to three glycine residues and beyond, the refolding rates decreased while the unfolding rates increased. It was shown that the longer loops had more degrees of freedom, and thus required more energy to constrain them on protein folding. The presence of a flexible loop destabilizes the folded native state and stabilizes the unfolded state, without a substantial change in the transition state [49]. Similarly in this case, we may hypothesize that in order to maintain the proper conformation of the loop [13], the number of Gly in the motif is critical. Three Gly residues in a row may significantly alter the conformation of the loop (e.g. to make it more flexible than the wild-type enzyme) and thus account for the decrease in the binding affinity of the enzyme against ATP.

In this study, we have biochemically characterized a GLK (LA_1437) belonging to the ROK family in *L. interrogans* serovar Lai strain 56601. Therefore, the failure of *L. interrogans* in utilizing glucose as its carbon/energy source [3] cannot be explained simply by the hypothesis of missing hexose kinase that we once proposed [3–5]. To further explore the hypothesis of limited glucose transportation [5], TransportDB, a database of cellular membrane transport systems [28], was used to analyze all types of possible glucose transport systems of *L. interrogans*. First, for the type of Secondary Transporter, although the **s**olute:-**s**odium **s**ymporter (SSS) family protein, LA_0694, may act as a Na^+/glucose symporter, its expressed protein detected in the *in vitro* cultivated cells was quite low according to our microarray and MS data (**Table 3**). Low transcription/expression was also observed in other two orthologous transporters of the SSS family (LA_2133 and LB_245), with their substrate specificities yet to be determined. Second, for the only possible phosphoenolpyruvate-dependent phosphotransferase system (PTS) encoded by *L. interrogans*, the *LA_0632*-encoded membrane binding permease EIIA component was obviously incomplete, lacking the EII B/C component for specific sugar-uptake, and thus it ought to be functionally deficient. Finally, due to the lack of experimental or bioinformatic evidence, it is uncertain as to whether any protein of the ATP-binding cassette (ABC) transporters in *L. interrogans* strain 56601 may function as glucose transporters. Therefore, limited glucose transportation hypothesis might be the first probable but uncertain scenario accounting for the failure of glucose utilization of *L. interrogans* [5].

Another possible scenario attributable to the failure of glucose utilization in *L. interrogans* is its potential lack of existence of Embden–Meyerhof pathway (glycolysis). First of all, the optimal *in vitro* assay condition of LiGLK (pH 9.0, 50°C) that we used to detect its activity in *L. interrogans* differed significantly from the physiological growth condition of the bacterium, particularly in the natural environment represented by the laboratory culture in EMJH at 28°C, pH 7.0. Although it is close to that of the parasitic host (e.g. human or pig and rodents, 37–40°C, pH 7.0), the real total enzymatic activity of LiGLK *in vivo* should be lower than what we calculated based on its *in vitro* assays. In addition, comparing the K_m value of LiGLK toward ATP with that of other GLKs of the ROK family, it is the highest, which might further infer its lower catalytic activity under physiological conditions. Second, although we failed to identify, by homology search, the gene encoding 6-phosphofructokinase (EC 2.7.1.11), a generally rate-limiting step in glycolysis [50], *L. interrogans* harbors a diphosphate-fructose-6-phosphate 1-phosphotransferase (LB_111, EC 2.7.1.90). In several species such as *Entamoeba histolytica*, this enzyme catalyzes the formation of D-fructose 1,6-bisphosphate from D-fructose 6-phosphate with the phosphate donor diphosphate instead of ATP used by 6-phosphofructokinase [32]. However, this putative function of LB_111 is yet to be confirmed biochemically. Finally, given LB_111 could be responsible for this reaction in *L. interrogans*, another rate-limited enzyme involved in glycolysis, pyruvate kinase LA_2924, together with the GLK (LA_1437), the diphosphate-fructose-6-phosphate 1-phosphotransferase (LB_111), all presented low transcription/expression levels in comparison with other reversible enzymes also responsible for gluconeogenesis (**Table 3**). In summary, both the limited glucose transportation and the deficiency in enzymes catalyzing glycolysis of *L. interrogans* are likely the multiple factors accounting for the incapability of glucose utilization.

Leptospira belongs to the order of *Spirochaetales*, which is primitive in prokaryotes' phylogeny [51]. The genome of *L. interrogans* bears a few genetic characters similar to either *Archea* or *Eukaryota*, signifying its close relationship with these distinct domains of life [4]. However, it is still unclear as to how this seemingly complete but actually silent glucose-utilization pathway (both transport and glycolysis) that we observed in *L. interrogans* today was evolved away from the glucose dominated catalytic physiology that we observed in most of the bacteria. This open question should be an interesting research topic for future studies.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (30770111,

30830002, 30900051 and 30970125), the National Key Program for Infectious Diseases of China (2008ZX10004 and 2009ZX10004), the Program of Shanghai Subject Chief Scientist (09XD1402700), and the Program of Shanghai Research and Development (10JC1408200).

References

1 Levett PN. Leptospirosis. Clin Microbiol Rev 2001, 14: 296–326.
2 Baseman JB and Cox CD. Intermediate energy metabolism of *Leptospira*. J Bacteriol 1969, 97: 992–1000.
3 Stalheim OH. Leptospiral selection, growth, and virulence in synthetic medium. J Bacteriol 1966, 92: 946–951.
4 Ren SX, Fu G, Jiang XG, Zeng R, Miao YG, Xu H and Zhang YX, *et al*. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. Nature 2003, 422: 888–893.
5 Nascimento AL, Ko AI, Martins EA, Monteiro-Vitorello CB, Ho PL, Haake DA and Verjovski-Almeida S, *et al*. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. J Bacteriol 2004, 186: 2164–2172.
6 Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ and Creno S, *et al*. Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. PLoS ONE 2008, 3: e1607.
7 Kefford B, Humphrey BA and Marshall KC. Adhesion: a possible survival strategy for leptospires under starvation conditions. Curr Microbiol 1986, 13: 247–250.
8 Bulach DM, Seemann T, Zuerner RL and Adler B. The organization of *Leptospira* at a genomic level. Bacterial Genomes Infect Dis 2007, 7: 109–123.
9 Tuininga JE, Verhees CH, van der Oost J, Kengen SW, Stams AJ and de Vos WM. Molecular and biochemical characterization of the ADP-dependent phosphofructokinase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Biol Chem 1999, 274: 21023–21028.
10 Ronimus RS, de Heus E and Morgan HW. Sequencing, expression, characterisation and phylogeny of the ADP-dependent phosphofructokinase from the hyperthermophilic, euryarchaeal *Thermococcus zilligii*. Biochim Biophys Acta 2001, 1517: 384–391.
11 Verhees CH, Tuininga JE, Kengen SW, Stams AJ, van der Oost J and de Vos WM. ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea. J Bacteriol 2001, 183: 7145–7153.
12 Ito S, Fushinobu S, Yoshioka I, Koga S, Matsuzawa H and Wakagi T. Structural basis for the ADP-specificity of a novel glucokinase from a hyperthermophilic archaeon. Structure 2001, 9: 205–214.
13 Lunin VV, Li Y, Schrag JD, Iannuzzi P, Cygler M and Matte A. Crystal structures of *Escherichia coli* ATP-dependent glucokinase and its complex with glucose. J Bacteriol 2004, 186: 6915–6927.
14 Concha MI and Leon G. Cloning, functional expression and partial characterization of the glucose kinase from *Renibacterium salmoninarum*. FEMS Microbiol Lett 2000, 186: 97–101.
15 Mesak LR, Mesak FM and Dahl MK. *Bacillus subtilis* GlcK activity requires cysteines within a motif that discriminates microbial glucokinases into two lineages. BMC Microbiol 2004, 4: 6.
16 Johnson RC and Harris VG. Differentiation of pathogenic and saprophytic letospires. I. Growth at low temperatures. J Bacteriol 1967, 94: 27–31.
17 Dong K, Li Q, Liu C, Zhang Y, Zhao G and Guo X. Cloning and characterization of three cheB genes in *Leptospira interrogans*. Acta Biochim Biophys Sin 2010, 42: 216–223.
18 Sambrook J, Fritsch EF and Maniatis T. eds. Molecular Cloning: A Laboratory Manual. New York: Spring Harbor Laboratory Press, 1989.
19 Zhong Y, Chang X, Cao X, Zhang Y, Zheng H, Zhu Y and Cai C, *et al*. Comparative proteogenomic analysis of the *Leptospira interrogans* virulence attenuated strain IPAV against the pathogenic strain 56601. Cell Res 2011, doi:10.1038/cr2011.
20 Pruitt KD, Tatusova T and Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 2005, 33: D501–D504.
21 Kanehisa M, Goto S, Furumichi M, Tanabe M and Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 2010, 38: D355–D360.
22 Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E and Martin MJ, *et al*. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 2003, 31: 365–370.
23 Tamura K, Dudley J, Nei M and Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007, 24: 1596–1599.
24 Chung CT, Niemela SL and Miller RH. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci USA 1989, 86: 2172–2175.
25 Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970, 227: 680–685.
26 Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976, 72: 248–254.
27 Hansen T, Reichstein B, Schmid R and Schonheit P. The first archaeal ATP-dependent glucokinase, from the hyperthermophilic crenarchaeon *Aeropyrum pernix*, represents a monomeric, extremely thermophilic ROK glucokinase with broad hexose specificity. J Bacteriol 2002, 184: 595559–595565.
28 Ren Q, Kang KH and Paulsen IT. TransportDB: a relational database of cellular membrane transport systems. Nucleic Acids Res 2004, 32: D284–D288.
29 Qin JH, Zhang Q, Zhang ZM, Zhong Y, Yang Y, Hu BY and Zhao GP, *et al*. Identification of a novel prophage-like gene cluster actively expressed in both virulent and avirulent strains of *Leptospira interrogans* serovar Lai. Infect Immun 2008, 76: 2411–2419.
30 Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S and Zimmermann P, *et al*. Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. Science 2008, 320: 938–941.
31 Hsieh PC, Shenoy BC, Samols D and Phillips NF. Cloning, expression, and characterization of polyphosphate glucokinase from *Mycobacterium tuberculosis*. J Biol Chem 1996, 271: 4909–4915.
32 Reeves RE, South DJ, Blytt HJ and Warren LG. Pyrophosphate: D-fructose 6-phosphate 1-phosphotransferase. A new enzyme with the glycolytic function of 6-phosphofructokinase. J Biol Chem 1974, 249: 7737–7741.
33 Titgemeyer F, Reizer J, Reizer A and Saier MH, Jr. Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. Microbiology 1994, 140: 2349–2354.
34 Berens C, Altschmied L and Hillen W. The role of the N terminus in Tet repressor for tet operator binding determined by a mutational analysis. J Biol Chem 1992, 267: 1945–1952.
35 Steitz TA. Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. Q Rev Biophys 1990, 23: 205–280.
36 Struhl K. Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. Trends Biochem Sci 1989, 14: 137–140.
37 Imriskova I, Arreguin-Espinosa R, Guzman S, Rodriguez-Sanoja R, Langley E and Sanchez S. Biochemical characterization of the glucose kinase from *Streptomyces coelicolor* compared to *Streptomyces peucetius* var. caesius. Res Microbiol 2005, 156: 361–366.
38 Han B, Liu H, Hu X, Cai Y, Zheng D and Yuan Z. Molecular characterization of a glucokinase with broad hexose specificity from *Bacillus sphaericus* strain C3–41. Appl Environ Microbiol 2007, 73: 3581–3586.

39 Skarlatos P and Dahl MK. The glucose kinase of *Bacillus subtilis*. J Bacteriol 1998, 180: 3222–3226.
40 Goward CR, Hartwell R, Atkinson T and Scawen MD. The purification and characterization of glucokinase from the thermophile Bacillus stearothermophilus. Biochem J 1986, 237: 415–420.
41 McCarthy JK, O''Brien CE and Eveleigh DE. Thermostable continuous coupled assay for measuring glucose using glucokinase and glucose-6-phosphate dehydrogenase from the marine hyperthermophile Thermotoga maritima. Anal Biochem 2003, 318: 196–203.
42 Hansen T and Schonheit P. ATP-dependent glucokinase from the hyperthermophilic bacterium *Thermotoga maritima* represents an extremely thermophilic ROK glucokinase with high substrate specificity. FEMS Microbiol Lett 2003, 226: 405–411.
43 Tomita K, Ikeda T and Takahashi T. Synthesis of adenosine 5′-triphosphate derivatives and their substrate activities to thermostable glucokinases. Ann NY Acad Sci 1998, 864: 548–551.
44 Miller BG and Raines RT. Identifying latent enzyme activities: substrate ambiguity within modern bacterial sugar kinases. Biochemistry 2004, 43: 6387–6392.
45 Klein DP and Charles AM. Purification and physical properties of glucokinase from Thiobacillus versutus (A2). Can J Microbiol 1986, 32: 937–941.
46 Sakuraba H, Mitani Y, Goda S, Kawarabayasi Y and Ohshima T. Cloning, expression, and characterization of the first archaeal ATP-dependent glucokinase from aerobic hyperthermophilic archaeon Aeropyrum pernix. J Biochem 2003, 133: 219–224.
47 Bossemeyer D. Loss of kinase activity. Nature 1993, 363: 590.
48 Nagi AD, Anderson KS and Regan L. Using loop length variants to dissect the folding pathway of a four-helix-bundle protein. J Mol Biol 1999, 286: 257–265.
49 Banner DW, Kokkinidis M and Tsernoglou D. Structure of the ColE1 rop protein at 1.7 A resolution. J Mol Biol 1987, 196: 657–675.
50 Pollack JD and Williams MV. PPi-dependent phosphofructotransferase (phosphofructokinase) activity in the mollicutes (mycoplasma) *Acholeplasma laidlawii*. J Bacteriol 1986, 165: 53–60.
51 Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN and Kunin V, *et al*. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature 2009, 462: 1056–1060.