

Christina Rother, BSc

# **Mechanistic considerations on levoglucosan kinase and its irreversible reaction**

MASTER'S THESIS  
to achieve the university degree of  
Diplom Ingenieurin (MSc)

Master's degree programme  
Biotechnology

Submitted to  
**Graz University of Technology**

Supervisor  
Univ.-Prof. Dipl.-Ing. Dr.techn. Bernd Nidetzky

Institute of Biotechnology and Biochemical Engineering

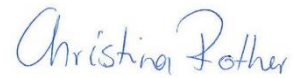
Graz, September 2016

## AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

20.09.2016

Date



Signature

## Acknowledgements

I sincerely thank my supervisor Bernd Nidetzky for his patience, his never ending advice and the numerous fruitful discussions. He has always taken the time to help me with ideas how to proceed with my experiments and to explain the rationale behind many things. I really appreciate his ability to ask the right questions at the right time to stimulate further thoughts and ideas about my work.

Thanks also to Rama Gudiminci for guiding me through my thesis, for his time to read through and correct my drafts for presentations, summaries and reports. Thanks for helping me get back any motivation I lost and getting access to many equipment or help I needed.

Alexander Gutmann was always there to help me out when I was struggling with difficulties of any kind, starting from planning of experiments to data interpretation and evaluation or error messages from equipment. Although working on something completely else, he always found time to think about my project and give me helpful advice. Thanks for your valuable support!

Many others from our research group have also contributed to this work fundamentally. Thanks to Manuel Eibinger for his advice concerning the Dionex measurements, to Alexander Lepak for helping me throughout the experiments with glucose-fluoride and to Martin Pfeiffer for his general interest and advice. Of course I would like to thank all the others from the institute their support and for making my time there more enjoyable.

I would also like to thank Jörg Weber for taking the time to plan and carry out the NMR measurements and their evaluation. He was never tired of helping me although we often needed several attempts to take the best spectra. Gernot Strohmeier was very accommodating during sample analysis by LC-MS, taking very much time for sample preparation as well as for measurements, data evaluation and analysis.

Most importantly, I would like to thank Christopher Ludwig for his priceless support, for his understanding concerning bad moods or delays, his ability to cheer me up and for his patience to answer my endless row of chemical questions or just listen to my considerations about the thesis. Thanks for making my life happier each and every day.

## Contents

AFFIDAVIT .....	2
Acknowledgements .....	3
<b>Chapter 1 – Manuscript in preparation .....</b>	<b>6</b>
1. Abstract .....	6
2. Graphical abstract .....	7
3. Introduction.....	7
4. Experimental .....	10
4.1. Cloning.....	10
4.2. Enzymatics .....	10
4.3. Mechanistics.....	12
5. Results and Discussion .....	14
5.1. Levoglucosan phosphorylation by LGK.....	14
5.2. Reversion of the natural reaction: Enzymatic synthesis of levoglucosan from glucose-6-phosphate.....	19
5.3. Overcoming energy barriers by using more reactive substrates .....	23
6. Conclusions.....	27
References.....	28
<b>Chapter 2 – Supporting information for the manuscript .....</b>	<b>31</b>
1. Experimental .....	31
2. Results and Discussion .....	32
2.1. Purification of LGK with N- and C-terminal His-tag .....	32
2.2. Further enzyme characterization .....	35
2.3. Synthesis of substrate analogues .....	38
References.....	40
<b>Chapter 3 – Additional experiments for characterization of levoglucosan kinase .....</b>	<b>42</b>
Supporting Information 2.....	42
1. Introduction.....	42
2. Experimental .....	42
3. Results and Discussion .....	45
3.1. Expression and analysis in small scale.....	45
3.2. Characterization of a side-product formed from glucose-6-phosphate by LGK enzyme preparations .....	47
4. Conclusion .....	52

<b>Appendix</b> .....	52
1. <i>Lipomyces starkeyi</i> strain YZ-215 levoglucosan kinase mRNA, complete cds.....	52
2. <i>Aspergillus niger</i> levoglucosan kinase mRNA, complete sequence .....	53
3. Primers for cloning of <i>L. starkeyi</i> LGK into pET 41 .....	53
4. Protein sequence of <i>L. starkeyi</i> LGK with C-terminal His-tag.....	53
5. Additional figures .....	54
List of figures .....	54

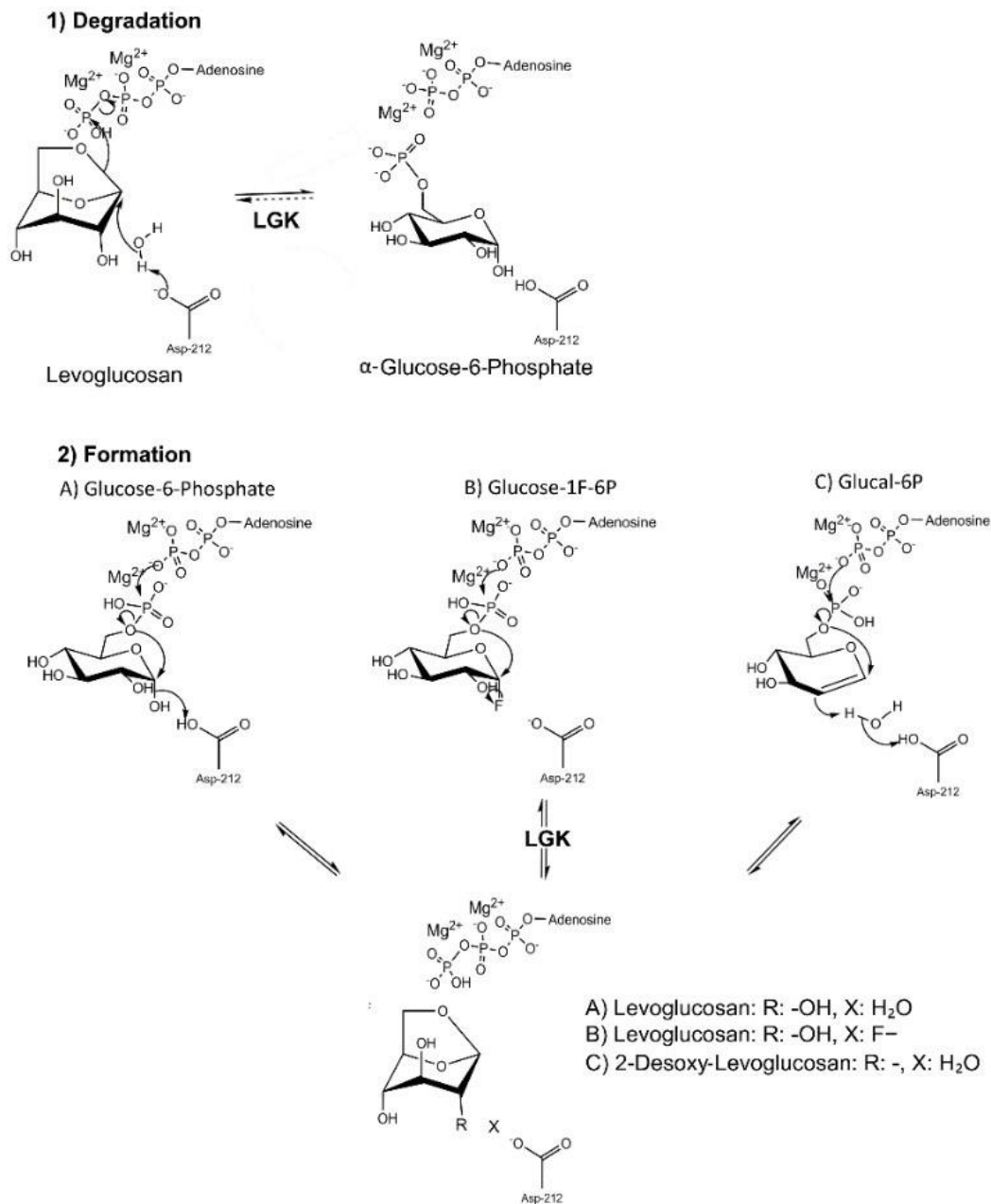
## Chapter 1 – Manuscript in preparation

### Mechanistic considerations on levoglucosan kinase and its irreversible reaction

#### 1. Abstract

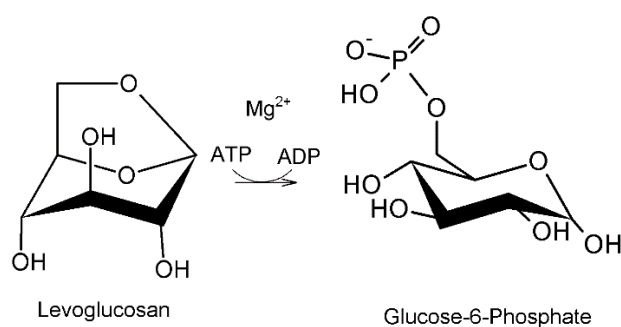
Levoglucosan kinase (LGK) catalyzes the simultaneous anhydro-ring opening and phosphorylation of levoglucosan (1,6-anhydro- $\beta$ -D-glucopyranose) in the presence of ATP. For the *Lipomyces starkeyi* LGK, we show here with real-time *in situ* NMR spectroscopy at 10 °C and pH 7.0 that the enzymatic reaction proceeds in a manner reminding of inverting glycoside hydrolases, resulting in the formation of the  $\alpha$ -anomer of D-glucose-6-phosphate. Kinetic characterization revealed the necessity of at least two-fold excess of the cofactor  $Mg^{2+}$  over ATP for optimum activity, the apparent binding of levoglucosan ( $K_M = 150$  mM) and ATP ( $K_M = 0.86$  mM) as well as the inhibition by ADP ( $K_I = 1.7$  mM) and D-glucose-6-phosphate ( $K_I = 54$  mM). The enzyme was highly specific for levoglucosan and exhibited weak ATPase activity in the absence of substrate. The equilibrium conversion of levoglucosan and ATP lay far on the product side and no enzymatic back reaction from D-glucose-6-phosphate and ADP could be demonstrated over a wide range of conditions. 6-Phospho- $\alpha$ -D-glucose-1-fluoride and 6-phospho-D-glucal were synthesized as probes of the enzymatic mechanism but proved inactive with the enzyme in the presence of ADP. The pyranose ring-flip  ${}^4C_1 \rightarrow {}^1C_4$  required for 1,6-anhydro-product synthesis from D-glucose-6-phosphate presents a major thermodynamic restriction to the back reaction, which due to 6-phospho-product binding tighter than substrate ( $K_M/K_I = 3$ ) may be difficult to overcome by simple mass action.

## 2. Graphical abstract



## 3. Introduction

Levoglucosan kinase (LGK) catalyzes the phosphorylation of the anhydrosugar levoglucosan (1,6-anhydro- $\beta$ -D-glucopyranose) in presence of ATP and  $\text{Mg}^{2+}$ . This reaction involves simultaneous opening of the anhydro-ring and phosphoryl-transfer from ATP to the C6 position of the sugar (Fig. 1). Together with AnmK (1,6-anhydro-N-acetyl muraminic acid kinase), which phosphorylates 1,6-anhydro-N-acetyl muraminic acid, levoglucosan kinase forms the sub-family of anhydrosugar kinases within the hexokinase family <sup>1,2</sup>



**Fig. 1: Phosphorylation of levoglucosan by LGK.** The reaction requires the presence of ATP and  $\text{Mg}^{2+}$ . The sugar stereochemistry changes from the  ${}^1\text{C}_4$  conformation in the anhydrosugar, to the  ${}^4\text{C}_1$  conformation in the phosphorylated product.

Levoglucosan itself is formed during burning of lignocellulosic materials at temperatures higher than  $300\text{ }^\circ\text{C}$ <sup>3</sup>. It is thus used as a biomarker for combustion events or human-made air pollution in aerosol samples<sup>3</sup>. Industrially, levoglucosan plays a role in the non-enzymatic depolymerization of biomass into fermentable sugars. Pyrolysis is considered as an efficient and faster alternative to enzymatic and acid hydrolysis of the raw materials<sup>4</sup>. In this process, biomass particles are decomposed at  $250 - 550^\circ\text{C}$  in the absence of oxygen<sup>5</sup>. Levoglucosan is the most abundant sugar in the liquid fraction of the pyrolysates, representing up to about 30 wt%<sup>4,6</sup>. For conversion of levoglucosan into sugars that can be further metabolized, prokaryotic organisms as *Arthrobacter* sp. use a reductive, three step reaction<sup>7</sup>. In contrast, LGK enables many eukaryotes to degrade the stable levoglucosan in one reaction to universally usable glucose-6-phosphate. LGKs from three different organisms (*Lipomyces starkeyi*<sup>8,9</sup>, *Aspergillus niger*<sup>10,11</sup>, *Sporobolomyces salmonicolor*<sup>12</sup>) have been purified and characterized until now and many putative enzymes with the same function have been proposed from sequence homology data<sup>8,11</sup>.

Levoglucosan is further considered as a highly valuable synthon for the generation of structurally diverse products. Prominent examples are macrolide antibiotics such as rifamycin S, or modified sugars<sup>13</sup> such as alpha-glycosyl azides<sup>14</sup> and the galactosidase inhibitor neurostegine<sup>15</sup>. Other biologically active molecules that have been synthesized from levoglucosan comprise the ionophore indanomycin, thromboxane B2, (+) biotin, the neurotoxin tetrodotoxin<sup>13</sup>, or naphthoquinone antibiotics<sup>16</sup>. The advantage of levoglucosan as an anhydrosugar compared to conventional hexoses is the fact that the pyranose ring is locked in the otherwise energetically highly unfavorable  ${}^1\text{C}_4$  conformation. This results in the stabilization of an inverted chair conformation compared to the more abundant  ${}^4\text{C}_1$  counterparts<sup>13</sup>, which opens new synthetic possibilities. Therefore, attempts have been made to generate large quantities of this compound in a rapid and reproducible manner. The isolation of levoglucosan from pyrolysis oil<sup>13</sup> is the most obvious approach thereof. Other preparation methods involve microwave-assisted pyrolysis<sup>17,18</sup> or chemical synthesis<sup>13,19</sup>. However, high temperatures or



harsh chemicals often make these processes environmentally unfriendly. Further, the high purity required for pharmaceutical use is still difficult to achieve because of the high concentrations of side products which often have to be removed. An efficient enzymatic synthesis of levoglucosan at ambient temperatures and under mild conditions would be thus highly desirable. Conversion of maltose to levoglucosan by a fungal glycosyltransferase has been proposed with 1 % substrate conversion<sup>20</sup>. However, this has not been investigated further.

Some specific kinases are known to be capable of conversion of the phosphorylated product and ADP to ATP, or even proved to be evolved for that purpose by nature. Pyruvate kinase, phosphoglycerate kinase<sup>21</sup> or creatin kinase<sup>22</sup> play a detrimental role in cellular metabolism for physiological ATP synthesis by glycolysis and rapid regeneration of energy reservoirs in muscle cells respectively. These reactions have in common that they start from a molecule with a high energy level, which shifts the reaction towards ATP and the dephosphorylated product<sup>21,22</sup>. For the purpose of organic synthesis, kinases have been mainly used for ATP dependent phosphorylation of substrates or co-factor recycling of nucleotides<sup>23</sup>, e.g. in the enzymatic synthesis of nucleotide sugars<sup>24</sup> or phosphorylated carbohydrates<sup>25</sup>. A reversed kinase reaction for synthesis, as in the use of pyruvate kinase<sup>25</sup> for NTP preparation, is more difficult to be found due to unfavorable energy demands, resulting in low product yields. Nevertheless, reversion of kinase reactions with unfavorable  $\Delta G$  were proposed to be possible as for glycerol kinase from *Trypanosoma brucei*, although activity in the reversed reaction by the purified enzyme could not be detected<sup>26</sup>.

The important question is whether the reaction of levoglucosan kinase can be reversed towards the synthesis of levoglucosan from glucose-6-phosphate for the preparation of levoglucosan in a highly pure form. On closer examination of the otherwise excellent study of Bacik et al.<sup>1</sup>, the mechanistic explanation of sugar ring opening for the simultaneous synthesis of both  $\alpha$ - and  $\beta$ - anomers of glucose-6-phosphate was not convincing. Therefore, before attempting to reverse the reaction towards the synthesis of levoglucosan, thorough investigations were carried out to understand the reaction mechanism and kinetics of levoglucosan kinase in the forward direction, i.e. synthesis of glucose-6-phosphate from levoglucosan. The effect of different parameters on enzyme stability and on the rate and completeness of levoglucosan phosphorylation was elucidated. The reversibility of the reaction was examined using not only the natural substrate, but also enzymatically synthesized substrate analogues. New kinetic values have been estimated in this study which are significantly different from the earlier studies because of a systematic error that has been frequently made in the previous studies<sup>8,9</sup>.

Both published sequences of LGK, *Aspergillus niger* (GenBank DB: AY987027<sup>11</sup>) and *Lipomyces starkeyi* (GenBank DB: EU751287<sup>8</sup>) LGK were used for cloning and expression. However, the *A. niger* sequence

shows some irregularities concerning the sequence length (800 bp), the longest translated peptide sequence possible (170 aa) and the described size of the putative LGK (75 kDa)<sup>11</sup>. Also, no overexpression of any peptide could be detected in the soluble and insoluble fraction of *E. coli* BL21 after expression. For this reason, the focus of this work will be on the research done on *Lipomyces starkeyi* LGK.

## 4. Experimental

### 4.1. Cloning

The LGK gene from *Lipomyces starkeyi* (GenBank DB: EU751287) was ordered as gBlocks® Gene Fragment in the codon optimized version for *Escherichia coli* from IDT, Iowa, USA. Enzymes and buffers for DNA manipulation, as well as PCR and agarose gel purification kits were obtained from Thermo Fisher Scientific. For cloning of *L. starkeyi* LGK into pET 41b(+), the sequence was amplified by PCR from its pQE-30 construct, introducing restriction sites for *NdeI* and *NotI* and simultaneously removing the stop codon for in-frame addition of the C-terminal His-tag. Ligation of the digested and dephosphorylated fragments was carried out at 4 °C overnight. The correct sequence of the amplified plasmid was confirmed by sequencing (Microsynth) prior to its use for transformation of *E. coli* BL21 Gold DE3.

The expression host was cultivated from an  $A_{600}$  of 0.01 in LB selection media (50 µg/mL kanamycin) with shaking at 37 °C. At a maximum  $A_{600}$  of 0.8, IPTG was added to 1 mM and expression was induced for 3 h at 30 °C. Cells were harvested by centrifugation (20 min, 5000 rpm, 4°C) and the pellet was dissolved in 50 mM HEPES pH 7.8 and frozen at -20 °C. Thawed cells were sonicated for cell disruption. The cleared lysate was mixed with ¼ volume of binding buffer (20 mM Na-Phosphate pH 7.4, 500 mM NaCl, 20 mM imidazole) and applied to HisTrap HP Ni Sepharose columns (GE Healthcare) using ÄKTA LC systems (GE Healthcare). After washing with binding buffer, bound protein was eluted using a gradient of elution buffer supplemented with 500 mM imidazole. The elution fractions were dialyzed with 20 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM DTT and aliquots were stored at 4°C or -20°C respectively. The purity and correct size of the construct was confirmed by SDS-PAGE (NuPAGE® Bis-Tris Precast Gels, Thermo Fisher Scientific). The distance of active-site residues was determined using PyMOL<sup>27</sup>.

### 4.2. Enzymatics

**Assays.** Activity assays for phosphorylation of levoglucosan (LG) to glucose-6-phosphate (G6P) were carried out according to literature<sup>8,10</sup>, with some adaptations. For conversion studies, a non-continuous assay was established where samples were removed from the reaction solution and inactivated by

heat (5 min at 95 °C) or by addition of ACN to 50 vol%. Adenosine nucleotides in the samples were measured by monitoring  $A_{259}$  with a MWD on HPLC using Phenomenex Kinetex C18 HPLC column with 2 mL/min flow (40 mM TBAB, 20 mM  $\text{KH}_2\text{PO}_4$  pH 5.9). Glucose-6-phosphate was quantified in a spectrophotometric dehydrogenase (DH) assay in 96-well plates.  $\frac{1}{10}$  sample volume was mixed with  $\frac{9}{10}$  test buffer (50 mM  $\text{KH}_2\text{PO}_4$  pH 7.0, 10 mM  $\text{MgCl}_2$ , 10 mM EDTA, 0.9 mg/mL  $\text{NAD}^+$ ). The reaction was started by addition of 2 U G6P-DH (Sigma, *Leuconostoc mesenteroides*) and  $A_{340}$  was followed over time. G6P concentration was determined using a calibration curve.

For determination of enzyme activity, a continuous LGK-assay was used, containing the components for the reaction from levoglucosan to G6P as well as the components for the G6P-DH assay. Optimized reaction conditions contained 50 mM HEPES pH 8, 300 mM LG (Sigma or Carbosynth), 2.5 mM ATP, 20 mM  $\text{MgCl}_2$ , 0.9 mg/mL  $\text{NAD}^+$ , 2 U G6P-DH, 1 mg/mL BSA, 0.02 mg/mL LGK. The reaction was started by levoglucosan addition and change of  $A_{340}$  was observed at 30 °C over a maximum of 20 min. For determination of G6P-concentration, the molar extinction coefficient for NADH of  $6.22 \text{ mM}^{-1}\text{cm}^{-1}$  and a pathway length of 0.588 for the 96-deep well plates were used. LGK activity was calculated from the linear range of the increase in G6P concentration over time.

*Basic characterization.* The stability of LGK was investigated by incubating the enzyme in a concentration of 1 mg/mL in dialysis buffer pH 7.5, containing BSA in concentrations from 0 – 5 mg/mL. pH-dependent stability was investigated by incubation of LGK (1 mg/mL) at 25 °C with 1 mg/mL BSA in 50 mM buffer at respective pH (MES pH 6.4, Gly-HCl pH 7, HEPES pH 7.9, Gly-NaOH pH 9.6). Residual activity was tested after various time spans in continuous mode. The  $K_M$ -value of LGK for levoglucosan was determined by measuring enzyme activities at LG concentrations from 10 – 750 mM. From the reaction mixture (2.5 mM ATP, 10 mM  $\text{MgCl}_2$ , 5 mg/mL BSA, 50 mM HEPES pH 7.8, 0.02 mg/mL LGK, 10 – 750 mM LG), the extent of ADP increase within 2 min of reaction was determined by Kinetex C18 column.  $K_M$  for ATP was determined using 500 mM LG, varying ATP concentrations from 0.02 – 20 mM.  $\text{MgCl}_2$  concentrations were 10 mM for 0.02 – 1 mM ATP, 20 mM for 2 – 20 mM ATP and 50 mM for 5 – 20 mM ATP. Adding increasing amounts of ADP (0 – 20 mM) and G6P (0 – 190 mM) to the reaction mixture (2.5 mM ATP, 300 mM LG, 50 mM HEPES pH 8, 5 mg/mL BSA, 20 mM  $\text{MgCl}_2$ , 0.02 mg/mL LGK), the inhibiting effects of the products on the reaction rate were determined. The apparent  $K_i$  was established by fitting a sigmoidal curve to the data for determining the concentration of substrate where half-maximal velocity was reached.

Alternative anhydrosugars were ordered from Carbosynth, UK. Substrate specificity of LGK was tested using 100 mM sugar (D-glucose, levoglucosan, 1,6-anhydro- $\beta$ -D-mannopyranose, 1,6-anhydro- $\beta$ -D-galactopyranose, 1,6-anhydro- $\beta$ -D-cellobiose, maltosan) with 2 mM ATP, 10 mM  $\text{MgCl}_2$ , 50 mM HEPES

pH 7.8, 1 mg/mL LGK. The changes in adenosine nucleotides as well as the released phosphates as described by Saheki et al.<sup>28</sup> were measured in aliquots within 24 h.

*Reaction reversibility.* For determination of the reaction equilibrium, various ratios of ATP and levoglucosan were used in a reaction containing 100 mM buffer, 20 mM MgCl<sub>2</sub>, 1 mg/mL BSA and 3 mg/mL LGK. Aliquots were removed from the reaction mixture for 24 h and changes in adenosine nucleotides and G6P were quantified. At equimolar substrate concentrations, the effect of pH variations from pH 7 to pH 10 was investigated. With the aim of reversing the LGK reaction, the substrates ADP and G6P were used in twofold K<sub>i</sub>-concentrations (108 mM G6P, 3.4 mM ADP) and totally inhibiting concentrations (150 mM G6P, 10 mM ADP). Further reaction components were 50 mM buffer pH 6, 7 or 8, 20 mM MgCl<sub>2</sub> and 1 mg/mL BSA. The reaction was carried out at 25 °C, 300 rpm for increased stability of LGK. Samples were analyzed using HPLC (HPX-87H column, BioRad, with 0.8 mL/min 5 mM H<sub>2</sub>SO<sub>4</sub> and RID). For very sensitive detection of possibly formed levoglucosan, samples were measured on high performance anion exchange chromatography with a pulsed amperometric detector (HPAEC-PAD, Dionex) with Carbopac PA10, 0.8 mL/min H<sub>2</sub>O, 15 % 200 mM NaOH. After each run, the column was washed with 1 M sodium acetate (10 min), followed by washing with 50 % (10 min) and 15 % (20 min) 200 mM NaOH. Sample preparation was carried out before analysis for removal of residual substrates to enable maximum loading capacity. Anion exchange using TOYOPEARL SUPER Q-650 M (TOSOH) column with 1 mL/min H<sub>2</sub>O was used on ÄKTA FPLC system. Samples were loaded twice and the column was washed with 2 M NaCl in between. The fractions showing increased conductivity during sample loading were collected as levoglucosan is the only uncharged molecule apart from buffer in this reaction mixture.

### 4.3. Mechanistics

*In situ* <sup>1</sup>H NMR was used for determination of the anomeric state of the glucose-6-phosphate formed by LGK. As the half-life of glucose-6-phosphate at 25 °C and pH 7 was described to be just 5 s ( $\alpha$ ) or 10 s ( $\beta$ )<sup>29</sup>, experiments were carried out at pD 5 first, but product formation was too slow and enzyme inactivation was occurring fast. Therefore, pD 7.0 and 10 °C were used as well as high enzyme concentrations of 2 mg/mL. Under these conditions, the half-life of  $\beta$ -glucose-6-phosphate should be around 3.5 min (2.1 min at pH 7, 0 °C<sup>30</sup>, kH<sub>2</sub>O/D<sub>2</sub>O=1.7<sup>31</sup>). 100 mM levoglucosan, 10 mM ATP, 20 mM MgCl<sub>2</sub> and 1 mg/mL BSA were used to reach a more balanced spectrum of the substrate signals and still stay in high enzymatic activity ranges. Spectra were recorded on a Varian (Agilent) INOVA 500-MHz spectrometer (Agilent Technologies, Santa Clara, United States) using VNMRJ 2.2D software. <sup>1</sup>H NMR spectra were measured at 499.98 MHz on a 5 mm indirect detection PFG-probe. The water signal was pre-saturated by a shaped pulse, using standard pre-saturation sequence with relaxation delay 2 s; 90° proton pulse; acquisition time 2.048 s; spectral width 8 kHz; number of points 32 k.

$\alpha$ -D-Glucopyranosyl fluoride (Glc-1F) was synthesized from 2,3,4,6-Tetra-O-acetyl- $\alpha$ -D-glucopyranosyl fluoride (Carbosynth), using a slightly adapted protocol from Steinmann et al.<sup>32</sup>. Shortly, the peracetylated substrate was deprotected by Zemplén deacetylation<sup>33</sup> using sodium methoxide in methanol under stirring at 0 °C. After neutralization to pH 7 using Amberlite IR-120 (10 min presoaked in MeOH), the solution was filtered through cotton before removal of solvents in oil pump vacuum. For preparation of 6-phospho- $\alpha$ -D-glucose-1-fluoride (Glc-1F-6P), 100 mM Glc-1F were used in a reaction with 100 mM ATP, 100 mM MgCl<sub>2</sub> and 4 mg/mL hexokinase type III from *Saccharomyces cerevisiae* (Sigma, 25 U/mg) in D<sub>2</sub>O. pD was adjusted to 7.0 with concentrated NaOH. <sup>19</sup>F and <sup>1</sup>H spectra were recorded 0 h and 24 h after hexokinase addition. At the latter time point, levoglucosan kinase was added to a final concentration of 1.5 mg/mL and again, <sup>19</sup>F and <sup>1</sup>H spectra were recorded 24 h after the second enzyme addition.

Enzymatic synthesis of 6-phospho-D-glucal was described using yeast hexokinase, which converts glucal in a rate of 0.2 % compared to the activity with D-glucose<sup>34</sup>. Phosphorylation of D-glucal (Sigma) was carried out using two different reaction setups. In the first, glucal, ATP and MgCl<sub>2</sub> were mixed in equimolar concentrations (50 mM). In the second reaction setup, 100 mM glucal were used with 10 mM ATP and an ATP-recycling system including 90 mM PEP, 12.5 U/mL pyruvate kinase (PK/LDH, Sigma) and 20 mM MgCl<sub>2</sub>. Both reaction setups contained 50 mM HEPES, 1 mg/mL BSA and 1 mg/mL hexokinase. 1 mg/mL levoglucosan kinase was each added 24 h after hexokinase addition. At this time, pH in the solution was 6.8 without ATP-recycling and 7.3 with recycling respectively. The reaction mixture was stopped by ACN 24 h after addition of the levoglucosan kinase. Samples were analyzed by HPX-87H column and TLC. Merck silica plates were used with 1-butanol:1-propanol:ethanol:H<sub>2</sub>O in a ratio of 3:3:1:1 as a mobile phase and orcinol stain (40 mg 5-methylresorcin, Aldrich, 80 mL ACN, 4 mL H<sub>2</sub>SO<sub>4</sub> conc.). As glucal and levoglucosan could not be separated on Dionex, LC-MS was used as sensitive detection method of possibly formed 2-desoxy levoglucosan. Without derivatization, product and substrate could not be discriminated without a standard for 2-desoxy-levoglucosan, as 2-desoxy levoglucosan and glucal have exactly the same mass. Peracetylation was therefore carried out, as available hydroxyl groups differ in number in the two compounds. An adapted protocol from Zottola et al.<sup>13</sup> was used for this purpose. After removal of water, pyridine (at least 80 eq) and acetic anhydride (at least 45 eq) were added in excess. The reaction was incubated at room temperature for 4 h, followed by incubation at 50 °C for another 14 h. After treatment with nitrogen stream for the removal of volatiles, the reaction was quenched by successive addition of one volume of water and three volumes of methanol. Filtered samples were analyzed on Agilent 1200 LC-MS equipped with multiple wavelength detector (MWD) and 6120 quadrupole mass spectrometer. Poroshell 120 SB-C18 column (100 × 3 mm, 2.7  $\mu$ m, Agilent, Vienna) was used for the analyses at a flow rate of 0.7 mL/min

H<sub>2</sub>O/HCOOH (0.1 %). A gradient of ACN/HCOOH (0.1 %) from 30 to 70 % within 5 minutes and from 70 to 100 % within one minute was applied. Each sample was analyzed at 210 nm and in MS scan mode (API-ES positive) using a mass range setting (m/z) between 100 and 500. Additionally, single ion monitoring (SIM) was set to 253, 269 and 211, correlating to the masses of sodium- or potassium adducts of di- and monoacetylated 2-desoxy levoglucosan.

The pK<sub>A</sub>-value of the catalytic base was estimated in the structure with PDB-code 4ZLU using PDB2PQR<sup>35</sup> with PROPKA<sup>36</sup> service using Parse or Amber force field and default settings. Further, YASARA (YASARA Biosciences GmbH, Vienna, Austria) was used for pK<sub>A</sub>-prediction with AMBER99 forcefield at pH 7.

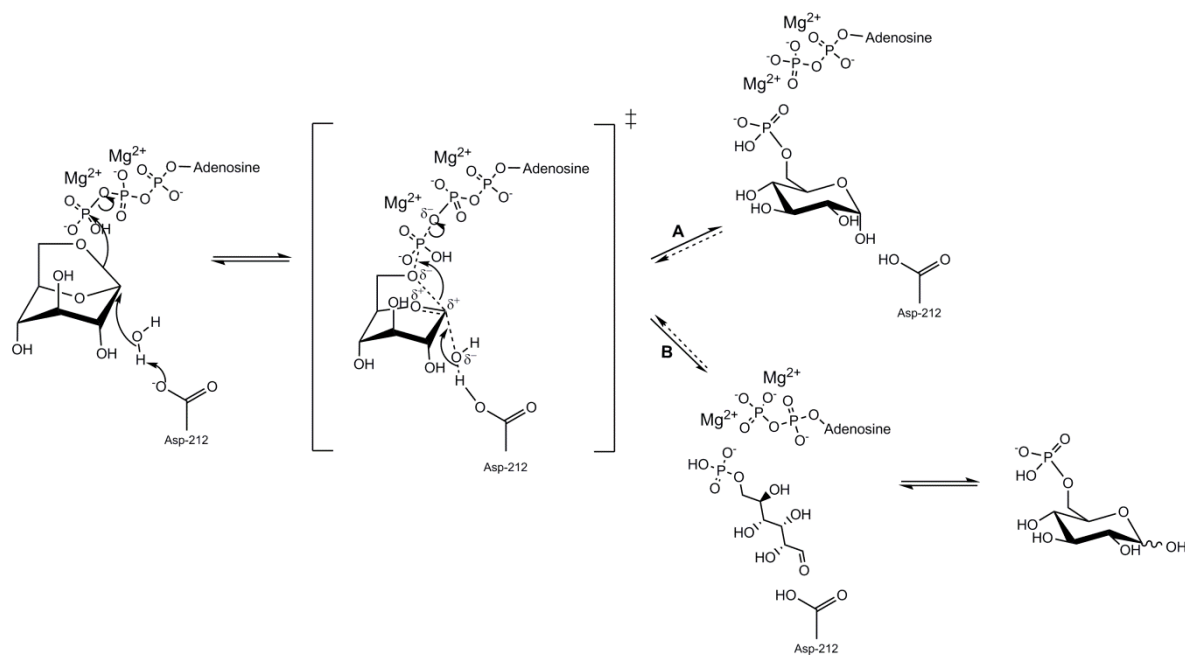
## 5. Results and Discussion

### 5.1. Levoglucosan phosphorylation by LGK

Levoglucosan kinase was described to be a homodimer with a molecular weight of 48 kDa for each subunit<sup>1</sup>. It was described to be stable in a wide pH range of 7 – 10 and temperature range of 4 – 35°C, with a maximum activity at 30°C and pH 9. The K<sub>M</sub> values for levoglucosan and ATP are 105 mM and 0.2 mM respectively<sup>8</sup>. For *A. niger* LGK, linear competitive inhibition of levoglucosan phosphorylation by ADP with a K<sub>i</sub> of 0.2 mM was reported<sup>10</sup>. Further, the buffer substance Tris binds to the active site and inhibits with a K<sub>i</sub> of 48 mM<sup>1</sup>. The maximum activity of crude cell extract of *Lipomyces starkeyi* LGK overexpressing *E. coli* BL21 DE3 was reported to be 7.9 U/mg, but no activity of the purified enzyme was stated<sup>8</sup>. Another group reported maximum activities of about 30 U/mg purified recombinant protein<sup>1</sup>.

*Mechanistic investigations.* For the phosphorylation of levoglucosan, a detailed mechanism has been proposed, based on the crystal structure<sup>1</sup>. The same mechanism was reported for the structurally and sequentially highly similar 1,6-anhydro-N-acetylmuramic acid kinase (AnmK) from *Pseudomonas aeruginosa*<sup>37</sup>. The reaction is initiated by subtraction of a proton from a water molecule by the catalytic base (Asp-212), followed by a nucleophilic attack of the activated water on C1. This results in the opening of the anhydro-ring and the nucleophilic attack of the O6 to the phosphoryl atom of the  $\gamma$ -phosphate from ATP. From the positioning of the Asp-212, it can be assumed that this mechanism is of inverting nature and thus results in the generation of  $\alpha$ -glucose-6-phosphate. This anomer can quickly isomerize to the  $\beta$ -anomer in aqueous solution (Fig. 2A). Alternatively, the opening of the intramolecular sugar hemiacetal is proposed in addition to the opening of the anhydro-bond. Due to circularization of the open-chain sugar in aqueous solution, this results in both anomers appearing in equilibrium concentrations<sup>1</sup> (Fig. 2B) of 62 %  $\beta$ -glucose-6-phosphate and 38 %  $\alpha$ -glucose-6-

phosphate<sup>29</sup>. Tensions during the conformational change of the sugar from the <sup>1</sup>C<sub>4</sub> conformation in levoglucosan to the <sup>4</sup>C<sub>1</sub> conformation in glucose-6-phosphate are proposed to be the reason for formation of the open-chain G6P<sup>1</sup>. In accordance with experimental findings, the levoglucosan phosphorylation reaction works best in high pH ranges, as the catalytic base gets deprotonated quickly, which is critical for further catalysis.



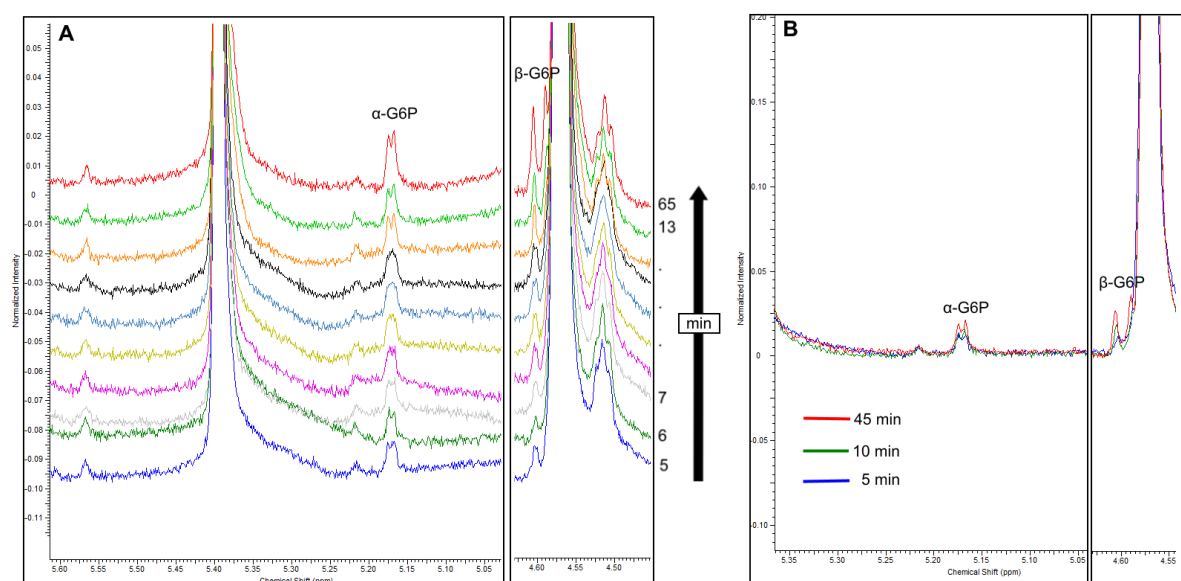
**Fig. 2: Proposed mechanism of levoglucosan phosphorylation by LGK. A)** A water molecule is activated by the catalytic base, followed by a nucleophilic attack on C1. This results in opening of the anhydro-ring, phosphoryl transfer to C6 and formation of α-glucose-6-phosphate. **B)** Tensions during re-arrangement of the sugar conformation are proposed to result in opening of the sugar semi-acetal. The open chain glucose-6-phosphate is then converted into its cyclic form in aqueous solution<sup>1</sup>.

Using *in situ* <sup>1</sup>H NMR studies, simultaneous formation of both anomers in equilibrium ratio was observed previously, which would support the alternative mechanism of sugar ring opening<sup>1</sup> (Fig. 2B). However, the formation rate of the product in the linear phase shown in this study was 0.0013 mmol/s which indicates that the reaction turnover rate might have been more slowly than mutarotation under used conditions. The half-life of glucose-6-phosphate anomers at 25 °C in water at pH 7 was described to be 5 (α) and 10 s (β)<sup>29</sup> and slightly higher in deuterated solvents with a k<sub>H<sub>2</sub>O/D<sub>2</sub>O</sub> of 1.7<sup>31</sup>.

In a repetition of these experiments, pH 7.0 and 10 °C were used. Under these conditions, the half-life of β-glucose-6-phosphate should be around 3.5 min (2.1 min at pH 7, 0°C<sup>30</sup>, k<sub>H<sub>2</sub>O/D<sub>2</sub>O</sub>=1.7). Additionally, high enzyme concentrations of 2 mg/mL and high substrate concentrations (100 mM levoglucosan, 10 mM ATP) were chosen to enable enzymatic rates to be as high as possible. We found an initial build-up of α-glucose-6-phosphate in an excess of 56 % compared to the β-anomer<sup>1</sup> with 44

% in the first spectrum, which was taken 5 minutes after reaction start. Afterwards, the ratio of  $\beta$ -glucose-6-phosphate compared to  $\alpha$ -glucose-6-phosphate increases until the equilibrium ratio is reached (Fig. 3 and Fig. 4). This is due to a slow-down of the reaction itself as a result of ATP-depletion, whereas the mutarotation rate stays constant. Thus, the area of the  $\alpha$ -anomer of G6P does not increase significantly in later spectra, whereas the signal for the  $\beta$ -anomer increases more predominantly compared to that. The results of this experiment support the original theory of the reaction mechanism that enables glucose-6-phosphate formation without opening of the sugar ring and results in inversion of the conformation of the anomeric carbon (Fig. 2A). This shows that levoglucosan kinase combines features of both kinase<sup>38</sup> and inverting glycoside hydrolase<sup>39</sup> enzymes in its catalytic properties.

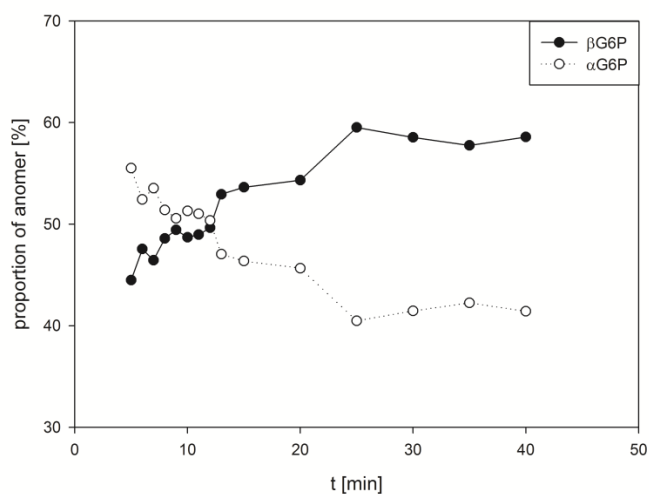
For integration of the signal for  $\beta$ -glucose-6-phosphate, the clearly separated half of the doublet signal was integrated and the resulting area was multiplied by two. The values in Fig. 4 serve for visualization of the trend of the ratios of glucose-6-phosphate anomers, but absolute values should be taken with care. Not only that integration had to be solved suboptimal, but also pre-saturation with the water signal and overlay with the neighboring levoglucosan signal could have an influence on the absolute values. However, ratios of the anomers to each other are not influenced by these factors.



**Fig. 3: Anomeric preferences of levoglucosan kinase.**  $^1\text{H}$  NMR (500 M Hz) spectra of the reaction in  $\text{D}_2\text{O}$  were recorded for 65 min at  $10^\circ\text{C}$ . Reaction conditions: 100 mM LG, 10 mM ATP, 20 mM  $\text{MgCl}_2$ , 1 mg/mL BSA, 2 mg/mL LGK, pD 7.0. Under these conditions, the half-life of  $\beta$ -glucose-6-phosphate was reported to be around 3.5 min. The pre-acquisition delay was 5 min, equilibrium of the anomers by mutarotation was reached after around 15 minutes. **A)** Overlay of the spectra 5 – 13 min and 65 min after reaction start. After an initial build-up of  $\alpha$ -G6P, the area of  $\beta$ -G6P increases significantly, whereas the signal for  $\alpha$ -G6P stays very much the same. Due to the decrease of reaction rate, all  $\alpha$ -G6P formed is converted to  $\beta$ -G6P

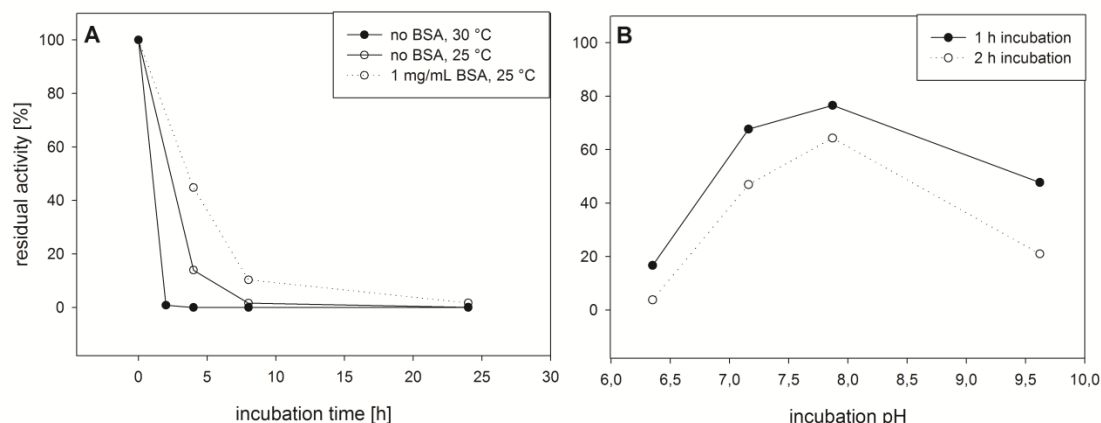


afterwards. **B)** Overlay of spectra taken 5, 10 and 45 min after reaction start show the excess of  $\alpha$ -G6P after 5 min, which is slowly decreasing because of anomerization.



**Fig. 4: Formation of G6P anomers by LGK.**  $^1\text{H}$  NMR (500 MHz) was recorded using 100 mM LG, 10 mM ATP, 20 mM  $\text{MgCl}_2$ , 1 mg/mL BSA, 2 mg/mL LGK, pD 7.0 at 10 °C. 5 min after reaction start, spectra were recorded every minute for 60 min. Areas were integrated by ACD/NMR processor and show an initial built-up of  $\alpha$ -glucose-6-phosphate in an excess of 56 % compared to the  $\beta$ -anomer with 44 % in the first spectrum. As the reaction rate slows down, the equilibrium ratio is approached.

*Basic properties.* LGK was reported to be 100 % stable within a pH range from 7 – 9 at 30 °C within 30 min of incubation<sup>8,9</sup>. While this might be true for short term exposure to the mentioned conditions, incubation of the enzyme for longer time spans significantly decreases its activity. Incubation at pH 7.5 at the optimum reaction temperature of 30 °C leads to loss of 99 % activity within 2 h (Fig. 5A). A decrease of the reaction temperature to 25 °C and addition of BSA enables improved stability. 45 % of activity was retained within 4 h under these conditions. The stabilizing effect of BSA was found to be independent of BSA concentrations between 0.5 and 2 mg/mL. However, after 24 h of incubation, only 1.7 % activity could be measured despite initial stabilization. This activity decrease is more significant when the buffer pH varies from optimum conditions (Fig. 5B). At pH 6.4 and pH 9.6, more than half of the activity was already lost within 1 h of incubation, whereas only 23 % of activity was lost at pH 7.8. As decreased activity due to storage at 4 °C was also observed within 2 weeks, storage at -20 °C was favored over the higher temperature. Apart from Tris, LGK was found to be inhibited by other buffering molecules like citric acid and sodium acetate.

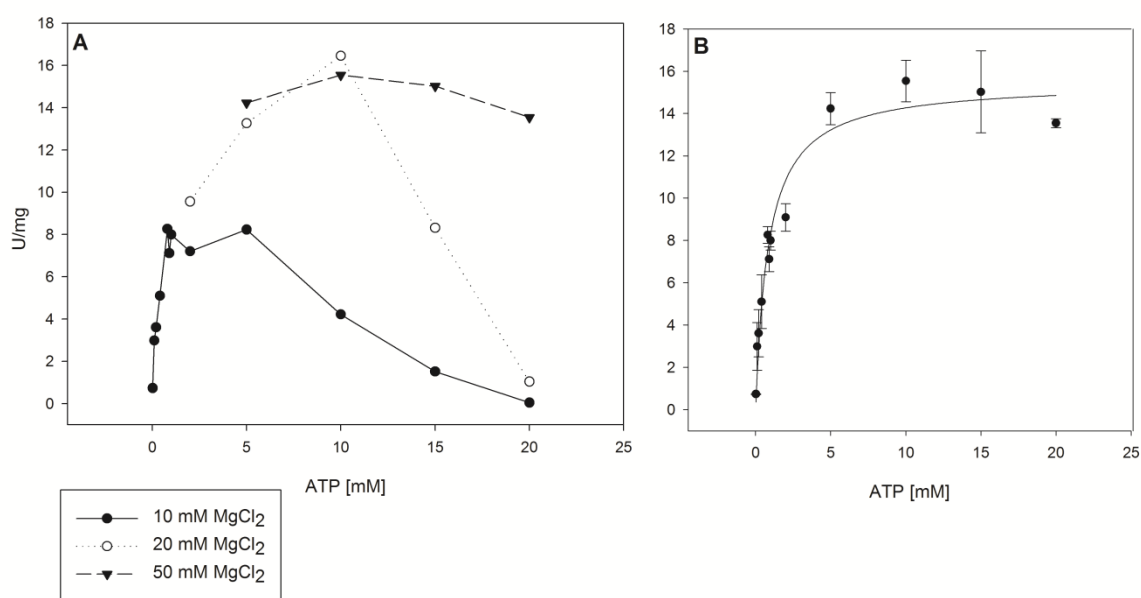


**Fig. 5 Stability of LGK.** A) Long term stability. 1 mg/mL enzyme was incubated in dialysis buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM DTT) at 30 °C or 25 °C with or without BSA. Residual activity was determined by continuous assay after indicated time spans. B) pH dependent stability. 1 mg/mL LGK was incubated at 25 °C with 1 mg/mL BSA in 50 mM buffer at respective pH (MES pH 6.4, Gly-HCl pH 7, HEPES pH 7.9, Gly-NaOH pH 9.6) for 1 or 2 h before measurement of residual activity by continuous assay.

Substrate specificity of levoglucosan kinase was investigated using various sugars and anhydrosugars, namely glucose, levoglucosan, anhydro-mannose, anhydro-galactose, anhydro-maltose and anhydro-cellobiose. As the enzyme was found to be inhibited by many buffering substances, a widely open active site can be assumed that allows binding of different substrates. In former studies, LGK substrate specificity has been investigated by *in vivo*<sup>8</sup> or spectrophotometric<sup>12</sup> assays with ambiguous results. Therefore, LGK was incubated with the tested sugars and change in adenosine nucleotides within 24 h was observed by HPLC for a more sensitive detection of sugar phosphorylation. Among the tested sugars, only the reaction with levoglucosan showed substantial changes in adenosine nucleotides. Within 30 minutes, 100 % of 2 mM ATP was consumed, without considerable release of phosphate. In all other sugars, as well as in the reaction mixture without sugar, decrease of ATP by about 0.6 mM was observed within 24 h, going along with the formation of the same amount of phosphate (Fig. S 5). This suggests that the changes in adenosine nucleotides are due to unproductive hydrolysis of ATP by LGK. The strict substrate specificity is especially unexpected as levoglucosan was reported to form only one hydrogen bond with the enzyme in the active site, which is the bond to the catalytic aspartate (Asp-212). Additionally, a hydrogen bond interaction between a water molecule, O2 of levoglucosan and Thr-116 was proposed<sup>1</sup>, which would suggest that C2-epimers of anhydro-glucose just as anhydro-mannose might not be positioned correctly and thus should not be accepted as substrates.

$K_M$  of LGK for levoglucosan was  $147 \pm 39$  mM, which is slightly higher compared to literature values (69 - 119 mM<sup>1,8,9</sup>).  $K_M$  for ATP is also higher than in literature with  $0.86 \pm 0.14$  mM compared to 0.2 – 0.68 mM<sup>8,9</sup>. The reason therefore lies in a systematic error that has been frequently made before. Using 10 mM MgCl<sub>2</sub>, the LGK activity increases with increasing ATP concentrations and seems to level

off at 2 mM ATP, which is the maximum concentration of ATP used in former measurements for determination of the  $K_M$ . However, increasing the ATP concentration further leads to an activity drop (Fig. 6A). Using higher  $MgCl_2$  concentrations, it can be demonstrated that this is due to depletion of free  $Mg^{2+}$ , which was already proposed by former studies<sup>1</sup> but not taken into account during  $K_M$  measurements. Using at least twofold concentrations of  $MgCl_2$  compared to ATP, LGK activity increases up to 10 mM ATP. On the other hand, too high excess of  $Mg^{2+}$  compared to ATP ( $> 10x$ ) seems to inhibit the reaction rate slightly (Fig. S 4), in a manner reminiscent of cyclin dependent kinases which also show dependence on two magnesium ions<sup>40</sup>. This supports the theory that a second  $Mg^{2+}$  promotes catalysis of levoglucosan phosphorylation by electrostatic and probably structural effects, whereas the stable binding of two  $Mg^{2+}$  to ADP limits the rate of product release at higher concentrations of the ion<sup>1,40</sup>. The  $v_{max}$  of LGK determined in this experiment lies at  $15.5 \pm 0.7$  U/mg (Fig. 6B).

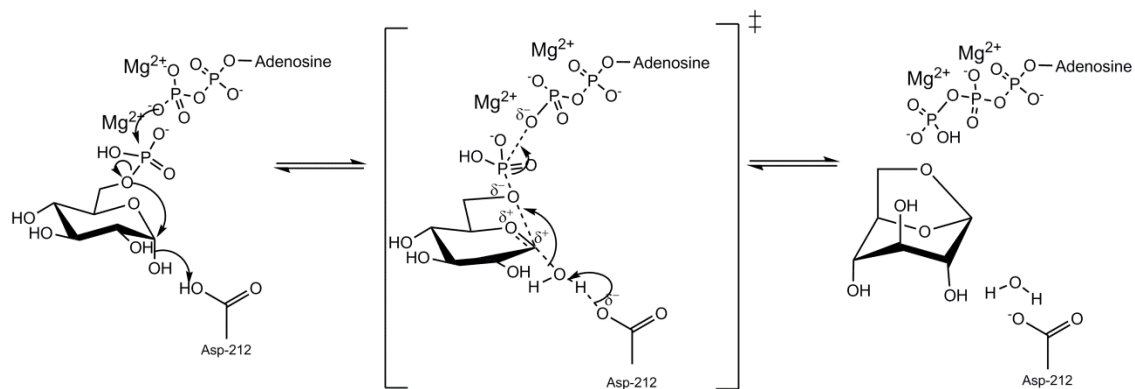


**Fig. 6:  $K_M$  of LGK for ATP. A) Dependency of LGK activity on ATP and  $MgCl_2$  concentrations.** ATP consumption within 2 min was detected by HPLC. Reaction conditions: 500 mM LG, 1 mg/mL BSA, 50 mM HEPES pH 7.8, 0.02 mg/mL LGK, 10 – 50 mM  $MgCl_2$ , 0.02 – 20 mM ATP. **B) Combined graph for determination of  $K_M$ .** The determined  $K_M$ -value of LGK for ATP is  $0.86 \pm 0.14$  mM.

## 5.2. Reversion of the natural reaction: Enzymatic synthesis of levoglucosan from glucose-6-phosphate

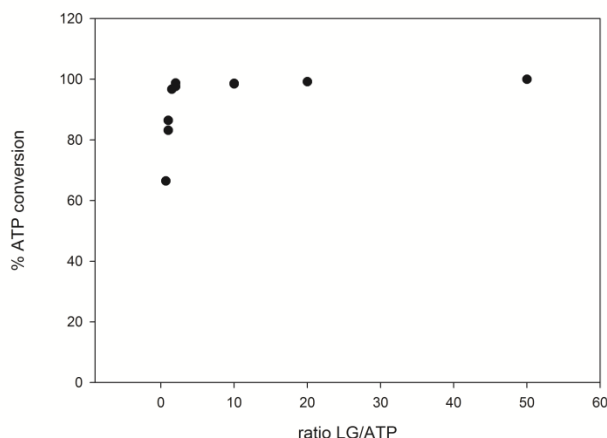
Having an environmentally friendly synthesis of levoglucosan as a valuable building block in mind, reversion of the reaction naturally catalyzed by LGK was addressed. The formation of levoglucosan from glucose-6-phosphate involves a nucleophilic attack of the ADP  $\beta$ -phosphoryl-oxygen to the phosphorous atom of glucose-6-phosphate and detachment of the phosphate from G6P. As a consequence, the nucleophilic O6 of the sugar substrate can attack its C1. Protonation of the anomeric

OH by Asp-212 enables formation of a positively charged oxocarbenium intermediate and the detachment of H<sub>2</sub>O as a leaving group (Fig. 7). As the pK<sub>A</sub> of aspartic acid is 3.9 in solution, this reaction should work best in acidic conditions, where the catalytic aspartate can act as a proton donor. However, the enzyme does not tolerate pH values below 7 very well, which is why it has to be assumed that the surrounding in the active site enables protonation of the Asp-212 already in mild pH ranges for this reaction to be possible.



**Fig. 7: Proposed reaction mechanisms for levoglucosan formation from glucose-6-phosphate.** This reaction requires the catalytic base to be in protonated form to be able to protonate the anomeric hydroxyl as a leaving group. Nucleophilic attack of the G6P phosphorous by the  $\beta$ -phosphoryl-oxygen of ADP leaves a nucleophilic oxygen on C6 of the sugar that can attack C1 for anhydro-bond formation.

To test the general feasibility of the reversion of the LGK reaction, experiments concerning the reaction equilibrium and its pH dependency were carried out. When the two substrates were used in equimolar concentrations, around 80 % conversion to ADP and glucose-6-phosphate and a  $K_{eq}$  of around 30 were observed. Between pH 7 and 10, no pH dependency of  $K_{eq}$  could be shown under equimolar substrate concentrations. However, the more levoglucosan was used in relation to the ATP concentration, the higher conversion and  $K_{eq}$  got until full consumption of ATP was observed at 50 times excess of levoglucosan over ATP (Fig. 8).  $K_{eq}$  thus is as high as 100 or even higher, but cannot be reached at lower levoglucosan concentrations due to high  $K_M$  for levoglucosan, inhibition by ADP and slow reaction rates which allows enzyme inactivation before full conversion can be achieved. These results show that the equilibrium of the LGK reaction, if there is any back-reaction at all, lies far on the side of ADP and glucose-6-phosphate. The independency of the reaction equilibrium on pH is highly unexpected, as the protonation state of the catalytic aspartate should vary between these conditions and plays an important role in the preferred direction of the catalyzed reaction. Either, Asp-212 is so much buried inside the protein that solvent pH does not have a huge impact, or the reaction equilibrium lies that far on the side of levoglucosan that it does not allow any detectable back-reaction at all. Another possibility is, that pH-values low enough to allow protonation of the Asp-212 cannot be reached because of the protein's inherent instability.

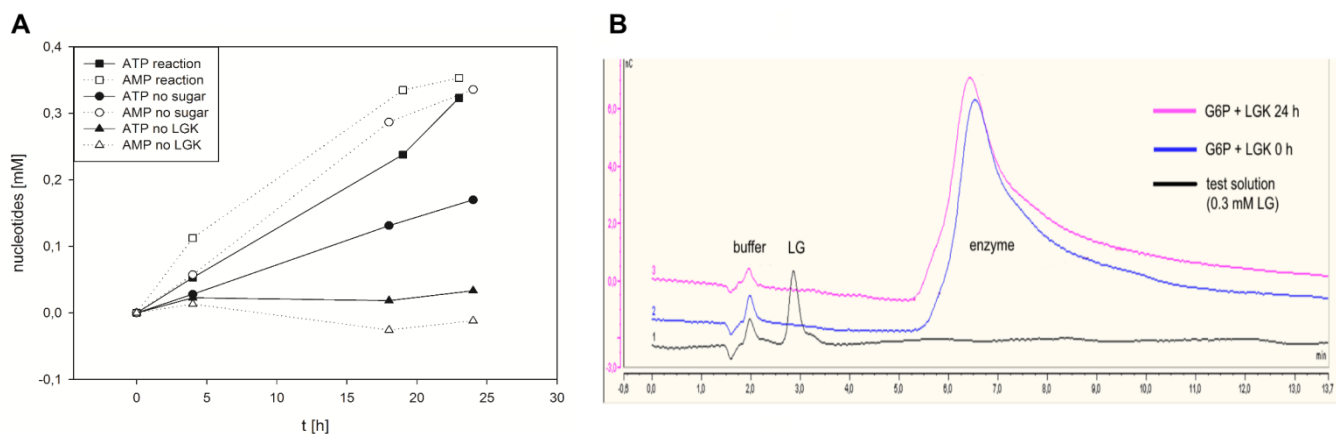


**Fig. 8: Dependency of levoglucosan conversion on LG/ATP ratios.** Almost complete conversion of ATP is only observed if levoglucosan concentration exceeds ATP concentration at least 2-fold. Different ratios of LG/ATP were used with 100 mM buffer, 20 mM MgCl<sub>2</sub>, 1 mg/mL BSA and 3 mg/mL LGK. ATP and G6P concentrations were determined within 4 h until at least two constant values after each other could be observed.

As no experimental approximation of the pK<sub>A</sub> of the catalytic base could be obtained, software tools were used for that task. However, estimation of the pK<sub>A</sub> results in very different values, depending on the tool that is used. PDB2PQR<sup>35</sup> reports a pK<sub>A</sub> of 8.85 – 9 for Asp-212 with both PROPKA and AMBER forcefield. This seems to be a very high value considering the pH optimum of 8 for the phosphorylation of levoglucosan, where the catalytic base should be deprotonated. In turn, the pK<sub>A</sub>-utility from YASARA indicates a pK<sub>A</sub> of 3.8 – 4.3 for Asp-212, which is an enormous difference in the logarithmic pH-scale. Therefore, it can be hardly predicted which protonation state the catalytic aspartate has at neutral pH and whether it is possible for it to act as a proton donor as well. Anyway, it is favorable to use solvent pH-values as low as possible for the enzymatic synthesis of levoglucosan for an efficient supply of the Asp-212 with H<sup>+</sup> ions.

For determination of optimal substrate concentrations for the reversed LGK reaction, the inhibiting effects of the products on the rate of levoglucosan phosphorylation were determined. Increasing amounts of ADP (0 – 20 mM) and G6P (0 – 190 mM) were added to the reaction and remaining activity was determined in the continuous assay. The apparent K<sub>i</sub> value for ADP was 1.7 mM and 54 mM for G6P (Fig. S 7). As shown for the previously mentioned CDK<sup>40</sup>, the complex of 2 Mg<sup>2+</sup> and ADP seems to be more tightly bound to the enzyme than the corresponding ATP-complex, resulting in inhibition of the levoglucosan phosphorylation reaction at high concentrations of ADP and Mg<sup>2+</sup>. For levoglucosan synthesis, ADP and G6P were therefore used first in twofold K<sub>i</sub> concentrations (108 mM G6P, 3.4 mM ADP), then at totally inhibiting concentrations (150 mM G6P, 10 mM ADP) in buffer with pH 6, 7 and 8. A slight increase of ATP was observed in all reactions, with a maximum of 0.3 mM at pH 6 using 150 mM glucose-6-phosphate and 10 mM ADP within 24 h. At higher pH and lower substrate

concentrations, the increase was even less. However, the increase in AMP was about the same as the ATP increase. Similar changes were observed in the reaction mixture without sugar, indicating disproportionation of the ADP to ATP and AMP. Thus, ATP formation might not be correlated to levoglucosan formation (Fig. 9A).



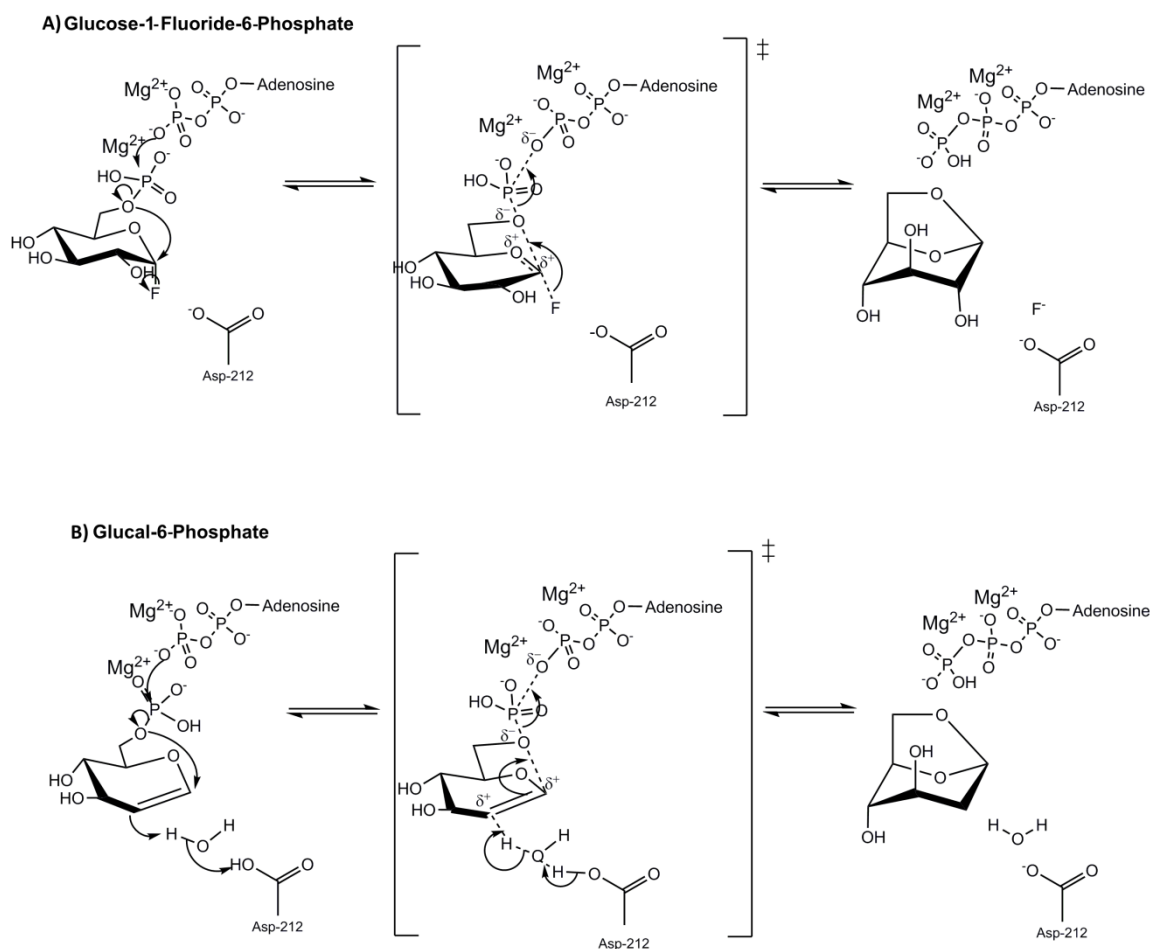
**Fig. 9: Investigation of reversibility of LGK reaction. A) Change in ATP and AMP during incubation of G6P with LGK.** In the reaction (squares), ATP concentration increases by 0.32 mM within 24 h. However, AMP increases by a very similar extent (0.35 mM). Without sugar (circles), ATP and AMP increase are also observed, although the ATP increase is less substantial. No changes are observed in the reaction mixture without enzyme (triangles). Reaction conditions: 150 mM G6P, 10 mM ADP, 50 mM MES pH 6, 1 mg/mL BSA, 20 mM MgCl<sub>2</sub>, 1.13 mg/mL LGK, 25 °C, 300 rpm. **B) Sensitive detection of levoglucosan by high performance anion exchange chromatography with a pulsed amperometric detector (Dionex).** All samples were prepared by two passes through anion exchange. In the test solution, 0.3 mM levoglucosan was added to prove the applicability of the sample preparation method. Nearly 100 % were recovered. No levoglucosan could be detected in the reaction solution.

No levoglucosan formation could be detected using HPX-87H column or TLC. To confirm the independency of ATP formation from product formation, high performance anion exchange chromatography with a pulsed amperometric detector (HPAEC-PAD) was chosen as a very sensitive detection method. Anion exchange chromatography was used beforehand for removal of residual glucose-6-phosphate and ADP in order not to overload the column and still avoid diluting the sample too much for maximum sensitivity. Besides the actual reaction mixture, a test solution spiked with 0.3 mM levoglucosan was prepared and treated the same way to test the applicability of the preparation method. Nearly 100 wt% of the initially loaded levoglucosan was recovered, though the solution was twofold diluted. Nevertheless, an initial concentration of levoglucosan as small as 0.3 mM could be clearly detected. However, no signal was observed from the reaction mixture, confirming that no levoglucosan formation had occurred (Fig. 9B). Further, no additional signal of uncharged sugars appears, speaking against a simple dephosphorylation of glucose-6-phosphate and e.g. glucose formation. This supports the idea of a disproportionation reaction of ADP in high concentrations. The

active site of LGK seems to be even open enough to enable binding of a second ADP molecule and the phosphoryl transfer from one ADP to the other in very low rates. The formation of each one AMP and ATP is the resulting outcome. This shows that glucose-6-phosphate is not converted to levoglucosan by LGK. Multiple reasons can be found as an explanation. One of them is the inability of the enzyme to subtract the OH-group on C1 as a leaving group. Corresponding to that is probably the impossibility to use pH-values that are low enough to ensure protonation of the catalytic amino acid residue.

### 5.3. Overcoming energy barriers by using more reactive substrates

Since glucose-6-phosphate as a natural substrate for LGK is not accepted,  $\alpha$ -D-Glucopyranosyl fluoride (Glc-1F) and 6-phospho-D-glucal (glucal-6P) as alternative substrates have been chosen based on the mechanistic considerations. Both should overcome the limitations concerning activation of the anomeric carbon for nucleophilic attack of the O6. The formation of levoglucosan from Glc-1F-6P follows the same reaction course as the formation from the non-fluorinated substrate in principle, including the nucleophilic attack of the  $\beta$ -phosphoryl-oxygen of ADP on the G6P phosphorous and attack of the dephosphorylated sugar O6 on the anomeric carbon (Fig. 10A). However, fluoride is readily detached from glucose-6-phosphate which means that protonation of Asp-212 would support the subtraction of fluoride, but is not absolutely necessary. Thus, low pH would be desirable but not substantial and the reaction could take place also in case the Asp-212 is not protonated. The use of glucal-6P as a substrate for enzymatic levoglucosan synthesis circumvents the need of a leaving group on C1 during nucleophilic attack of the dephosphorylated C6. However, pH-dependency is again more relevant than with Glc-1F-6P as the C2 in glucal-6P has to be protonated to create a positively charged carbonium ion during opening of the C-C double bond. As the distance of the sugar C2 to Asp-212 is around 5.4 Å, this is more likely to occur over a water atom (Fig. 10B).

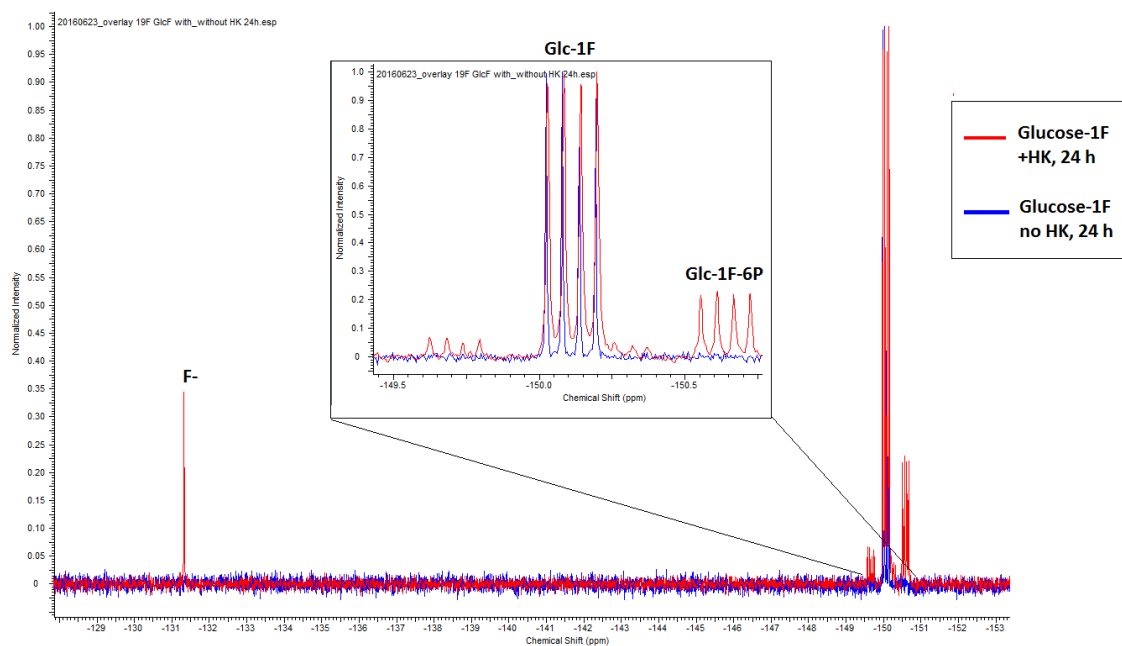


**Fig. 10: Proposed reaction mechanisms for the formation of levoglucosan from alternative sugar substrates. A) From Glc-1F-6P.** This reaction follows the same principle as the reaction from glucose-6-phosphate, just that the fluoride could theoretically be detached non-enzymatically. However, protonation of Asp-212 would facilitate this process. **B) From glucal-6P.** Protonation of C2 of glucal-6P is necessary in order for the dephosphorylated, nucleophilic O6 to attack C1. Due to spatial conditions within the active site, this most likely occurs involving a water molecule.

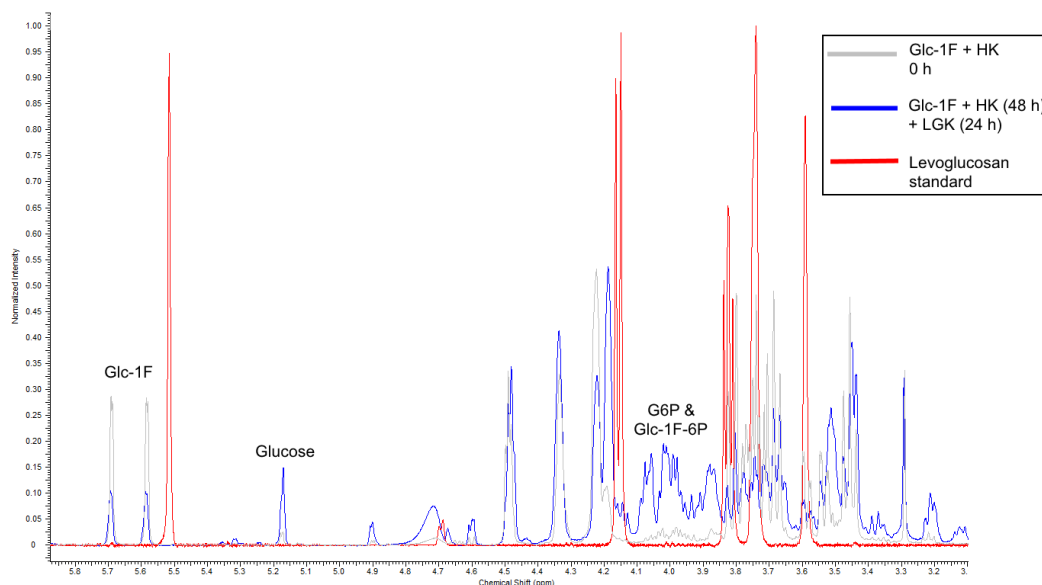
For the preparation of Glc-1F-6P, chemically synthesized glucose-fluoride was enzymatically phosphorylated in C6 position. The successful phosphorylation of Glc-1F by hexokinase was shown using  $^{19}\text{F}$  NMR. In the reaction mixture after 24 h, two major signals were visible in addition to the Glc-1F signal at -150.2 ppm (Fig. 11). The first at -131.4 ppm belongs to the free fluoride and the other one at -150.7 ppm belongs to Glc-1F-6P. The area of the phosphorylated glucose-fluoride amounts to 23 % of the remaining glucose-fluoride substrate. In the control reaction without enzyme, none of these signals except the substrate signal are visible after 24 h of incubation. This also demonstrates that free fluoride is more readily released from the sugar in presence of hexokinase than in its absence. After 24 h reaction of Glc-1F with hexokinase, LGK was added to 1.5 mg/mL and the reaction mixture was incubated for another 24 h.  $^1\text{H}$  spectra taken afterwards do not indicate formation of levoglucosan (Fig. 12). This might also partly be due to inhibition by remaining ATP, which was not completely



consumed during glucose-fluoride phosphorylation. Further, the low concentration of Glc-1F-6P (around 9 mM) could be a limiting factor for levoglucosan formation because of the high  $K_M$  of LGK for its sugar substrate.



**Fig. 11: Phosphorylation of Glc-1F by hexokinase.** 100 mM Glc-1F were incubated in  $D_2O$  with 100 mM ATP, 100 mM  $MgCl_2$  and 4 mg/mL hexokinase.  $^{19}F$  spectra were recorded after 24 h of reaction (red) and compared to the control sample omitting enzyme after 24 h (blue). The generation of two new fluorinated molecular species could be observed, one of which is proposed to be Glc-1F-6P.



**Fig. 12:  $^1H$  NMR of Glc-1F-6P with LGK after 24 h.** 100 mM Glc-1F were incubated in  $D_2O$  with 100 mM ATP, 100 mM  $MgCl_2$  and 4 mg/mL hexokinase (grey signal after 0 h). 24 h after hexokinase addition, LGK was added to 1.5 mg/mL and the solution was incubated for another 24 h, before the last spectrum was recorded (blue). The signals that should appear in case of levoglucosan presence can be seen in red.

In the enzymatic preparation of glucal-6-phosphate, the phosphorylation of the substrate glucal was neither complete, as was demonstrated by the residual substrate on TLC and by analysis of the adenosine nucleotides for the reaction without ATP recycling. 54 % of ATP were consumed in this reaction within 24 h, indicating the formation of about 27 mM glucal-6-phosphate. No levoglucosan could be detected after 24 h of incubation with LGK on TLC with a sensitivity of about 10 mM for levoglucosan (Fig. S 8). On different available columns for HPAEC-PAD, no separation of glucal and levoglucosan could be achieved. For LC-MS, sample preparation was carried out for introducing differences in mass and polarity to glucal and 2-desoxy-levoglucosan, which otherwise both have a molecular weight of 146.04 g/mol.

From the UV-chromatogram and the mass spectrometry data, one can see that the peracetylation of the remaining glucal was complete in both samples, containing the reaction of LGK with glucal-6-phosphate with and without ATP recycling. A strong UV and mass signal (TIC) at 3.62 min retention time with a main mass of 295 ( $[M+23]^+$ ) is visible, corresponding to the sodium adduct of the peracetylated glucal (Fig. S 9A, S 10A). To enable higher sensitivity for the detection of minor compounds, increased sample volumes were used in repetition measurements. Therefore, mass analyses were carried out until a cut-off time of 3.5 min to allow a maximum loading for the low concentrated samples, but avoiding contamination of the MS-detector by large amounts of peracetylated glucal. The chemical properties of peracetylated 2-desoxy-levoglucosan are believed to be similar to peracetylated glucal, but a shorter elution time than the main compound must be expected because of its higher polarity. However, in none of the samples, significant differences between the signals after 0 h and 24 h of LGK incubation could be detected in all chromatography traces monitored (Fig. S 9B, S 10B). Additional mass analyses during the whole analysis runs using lower injection amounts did not confirm any formation of 2-desoxy-levoglucosan. As the mass response of glucal in the peracetylated form was very high and consequently a similar behavior for peracetylated 2-desoxy-levoglucosan is assumed, even minor amounts of the latter compound should have been detectable. Thus, one must conclude that 2-desoxy-levoglucosan was not formed by LGK from glucal-6-phosphate in detectable amounts under the used conditions. Residual adenosine nucleotides measured on HPLC before LGK addition were 27 mM ATP and 23 mM ADP in the reaction without ATP recycling and 6.9 mM ATP and 3.1 mM ADP in the reaction with recycling system. Although this seems not optimal for an efficient drag of the reaction to the product side, the fact that product formation was not detectable at all with high sensitivity measurements proves that glucal-6-phosphate is no suitable substrate for LGK in the synthesis of levoglucosan.

The possible rationale behind the unsuccessful reversion of the levoglucosan phosphorylation may be the high difference in free enthalpy of substrates and products. Not only the formation of levoglucosan

from glucose-6-phosphate, but also the phosphorylation of ADP to ATP is thermodynamically highly unfavorable. The phosphorylation of ADP in aqueous solution requires 26.4 kJ/mol energy<sup>41</sup>. The free enthalpy of levoglucosan is even around 40 kJ/mol higher than that of the substrate glucose-6-phosphate<sup>41,42</sup>. Additionally, the energy of a potential transition state is likely to be even higher than that. The free-energy barrier for formation of levoglucosan at 600 K was estimated to be 288 kJ/mol<sup>43</sup> and even higher at ambient temperatures. Another reason for the irreversibility of the LGK reaction might be wrong positioning of the substrate within the active site. As the enzyme is not promiscuous at all regarding the acceptance of substrates for phosphorylation, it can be assumed that it acts the same way in the reversed reaction. The fluorinated or dehydrated substrates might have hampered correct positioning of the sugar which is the uppermost prerequisite for the anhydro-ring to form.

## 6. Conclusions

Levoglucosan kinase is a mechanistically highly interesting enzyme due to its ability to open the unusual anhydro-ring in 1,6-anhydro- $\beta$ -D-glucopyranose. However, the application of the purified enzyme for industrial use in conversion of levoglucosan to fine or bulk chemicals is restricted not only by the unfavorably high  $K_M$  for its sugar substrate, but also by its strict substrate specificity and the insufficient long term stability in respect to low pH and moderate temperature. The sum of these properties, together with the high energy level of levoglucosan, makes the *L. starkeyi* LGK also unsuitable for its use in the synthesis of anhydrosugars in its natural form. Directed evolution or immobilization of the enzyme for increased tolerance of acidic conditions might enable the use of reactions at low pH for protonation of the catalytic base. Introduction of specific mutations within the active site for optimal positioning of the activated sugar O6 could further be advantageous. For example, amino acids that are positioned above the anhydro-bond to be formed could be exchanged to bulkier residues to allow less motion of the C6 of the sugar substrate. Asn186, Ile190 or Asn191 could be promising candidates therefore. If the enzyme was to be used in industry for fermentation of levoglucosan to ethanol in its purified form, substantial stabilization had to be done e.g. by immobilization or addition of stabilizing co-solvents. Further, high excess of the sugar substrate as well as ATP recycling systems would have to be used to make the process economically viable.

## References

1. Bacik, J.-P. *et al.* Producing glucose 6-phosphate from cellulosic biomass: structural insights into levoglucosan bioconversion. *J. Biol. Chem.* **290**, 26638–48 (2015).
2. Uehara, T. *et al.* Recycling of the anhydro-N-acetylmuramic acid derived from cell wall murein involves a two-step conversion to N-acetylglucosamine-phosphate. *J. Bacteriol.* **187**, 3643–9 (2005).
3. Hu, Q.-H., Xie, Z.-Q., Wang, X.-M., Kang, H. & Zhang, P. Levoglucosan indicates high levels of biomass burning aerosols over oceans from the Arctic to Antarctic. *Sci. Rep.* **3**, 3119 (2013).
4. Islam, Z. U., Zhisheng, Y., Hassan, E. B., Dongdong, C. & Hongxun, Z. Microbial conversion of pyrolytic products to biofuels: a novel and sustainable approach toward second-generation biofuels. *J. Ind. Microbiol. Biotechnol.* **42**, 1557–79 (2015).
5. Kersten, S. & Garcia-Perez, M. Recent developments in fast pyrolysis of ligno-cellulosic materials. *Curr. Opin. Biotechnol.* **24**, 414–20 (2013).
6. Jarboe, L. R., Wen, Z., Choi, D. & Brown, R. C. Hybrid thermochemical processing: fermentation of pyrolysis-derived bio-oil. *Appl. Microbiol. Biotechnol.* **91**, 1519–23 (2011).
7. Nakahara, K., Kitamura, Y., Yamagishi, Y., Shoun, H. & Yasui, T. Levoglucosan dehydrogenase involved in the assimilation of levoglucosan in *Arthrobacter* sp. I-552. *Biosci. Biotechnol. Biochem.* **58**, 2193–6 (1994).
8. Dai, J. *et al.* Cloning of a novel levoglucosan kinase gene from *Lipomyces starkeyi* and its expression in *Escherichia coli*. *World J. Microbiol. Biotechnol.* **25**, 1589–1595 (2009).
9. Ning, J. *et al.* Purification and characterization of levoglucosan kinase from *Lipomyces starkeyi* YZ-215. *World J. Microbiol. Biotechnol.* **24**, 15–22 (2007).
10. Zhuang, X. & Zhang, H. Identification, characterization of levoglucosan kinase, and cloning and expression of levoglucosan kinase cDNA from *Aspergillus niger* CBX-209 in *Escherichia coli*. *Protein Expr. Purif.* **26**, 71–81 (2002).
11. Xie, H., Zhuang, X., Zhang, H., Bai, Z. & Qi, H. Screening and identification of the levoglucosan kinase gene (lgk) from *Aspergillus niger* by LC-ESI-MS/MS and RT-PCR. *FEMS Microbiol. Lett.* **251**, 313–9 (2005).
12. Kitamura, Y. & Yasui, T. Purification and Some Properties of Levoglucosan (1,6-Anhydro- $\beta$ -D-glucopyranose) Kinase from the Yeast *Sporobolomyces salmonicolor*. *Agric. Biol. Chem.* **55**, 523–529 (2014).
13. Zottola, M. A., Alonso, R., Vite, G. D. & Fraser-Reid, B. A practical, efficient large-scale synthesis of 1,6-anhydrohexopyranoses. *J. Org. Chem.* **54**, 6123–6125 (1989).
14. Lepage, M. L., Bodlenner, A. & Compain, P. Stereoselective Synthesis of  $\alpha$ -Glycosyl Azides by TMSOTf-Mediated Ring Opening of 1,6-Anhydro Sugars. *European J. Org. Chem.* **2013**, 1963–1972 (2013).
15. Salamone, S. *et al.* Synthesis and evaluation of galacto-noeurostegine and its 2-deoxy analogue as glycosidase inhibitors. *Org. Biomol. Chem.* **13**, 7979–92 (2015).
16. Freskos, J. N. & Swenton, J. S. Annelation reaction of levoglucosenone. Chiral intermediates for the synthesis of naphtho[2.3-c]pyran-5,10-quinone antibiotics. *J. Chem. Soc. Chem. Commun.* 658 (1985).

doi:10.1039/c39850000658

17. Al Shra'ah, A. & Helleur, R. Microwave pyrolysis of cellulose at low temperature. *J. Anal. Appl. Pyrolysis* **105**, 91–99 (2014).
18. Hoai, N. T. *et al.* Selective synthesis of 1,6-anhydro- $\beta$ -d-mannopyranose and -mannofuranose using microwave-assisted heating. *Carbohydr. Res.* **346**, 1747–51 (2011).
19. Ohara, M., Takagaki, A., Nishimura, S. & Ebitani, K. Syntheses of 5-hydroxymethylfurfural and levoglucosan by selective dehydration of glucose using solid acid and base catalysts. *Appl. Catal. A Gen.* **383**, 149–155 (2010).
20. Pazur, J. H., Tominaga, Y., DeBrosse, C. W. & Jackman, L. M. The synthesis of 1,6-anhydro- $\beta$ -d-glucopyranose and d-glucosyl oligosaccharides from maltose by a fungal glucosyltransferase. *Carbohydr. Res.* **61**, 279–290 (1978).
21. Berg, J. M., Tymoczko, J. L. & Stryer, L. in *Stryer Biochemie* 456–500 (Springer Berlin Heidelberg, 2013).  
doi:10.1007/978-3-8274-2989-6\_16
22. Szasz, G., Gruber, W. & Bernt, E. Creatine kinase in serum: 1. Determination of optimum reaction conditions. *Clin. Chem.* **22**, 650–6 (1976).
23. Koeller, K. M. & Wong, C.-H. Enzymes for chemical synthesis. *Nature* **409**, 232–240 (2001).
24. Bülter, T. & Elling, L. Enzymatic synthesis of nucleotide sugars. *Glycoconj. J.* **16**, 147–159 (1999).
25. Wong, C.-H. & Whitesides, G. M. *Enzymes in synthetic organic chemistry*. (Academic Press, 1994).
26. Krakow, J. L. & Wang, C. C. Purification and characterization of glycerol kinase from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **43**, 17–25 (1990).
27. Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.8. (2015).
28. Saheki, S., Takeda, A. & Shimazu, T. Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. *Anal. Biochem.* **148**, 277–281 (1985).
29. Schray, K. J. & Benkovic, S. J. Anomerization rates and enzyme specificity for biologically important sugars and sugar phosphates. *Acc. Chem. Res.* **11**, 136–141 (1978).
30. Gernert, E. & Keston, A. S. Mutarotase (aldose-1-epimerase) catalyzed anomerization of glucose-6-phosphate. *Arch. Biochem. Biophys.* **161**, 420–425 (1974).
31. Schray, K. J. & Howell, E. E. Anomerization of glucose 6-phosphate. *Arch. Biochem. Biophys.* **189**, 102–105 (1978).
32. Steinmann, A., Thimm, J., Matwiejuk, M. & Thiem, J. Formation of Homooligosaccharides Using Base-Promoted Glycosylation of Unprotected Glycosyl Fluorides. *Macromolecules* **43**, 3606–3612 (2010).
33. Zemplén, G. & Pacsu, E. Über die Verseifung acetylierter Zucker und verwandter Substanzen. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.)* **62**, 1613–1614 (1929).
34. Chenault, H. K. & Mandes, R. F. Selective inhibition of metabolic enzymes by enzymatically synthesized D-glucal-6-phosphate. *Bioorg. Med. Chem.* **2**, 627–629 (1994).

35. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–7 (2004).
36. Holst, M. & Saied, F. Multigrid solution of the Poisson-Boltzmann equation. *J. Comput. Chem.* **14**, 105–113 (1993).
37. Bacik, J.-P. *et al.* Molecular basis of 1,6-anhydro bond cleavage and phosphoryl transfer by *Pseudomonas aeruginosa* 1,6-anhydro-N-acetylmuramic acid kinase. *J. Biol. Chem.* **286**, 12283–91 (2011).
38. Pollard-Knight, D. & Cornish-Bowden, A. Mechanism of liver glucokinase. *Mol. Cell. Biochem.* **44**, 71–80 (1982).
39. Davies, G. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859 (1995).
40. Jacobsen, D. M., Bao, Z.-Q., O'Brien, P., Charles L. Brooks, I. & Young, M. A. Price To Be Paid for Two-Metal Catalysis: Magnesium Ions That Accelerate Chemistry Unavoidably Limit Product Release from a Protein Kinase. (2012).
41. Flamholz, A., Noor, E., Bar-Even, A. & Milo, R. eQuilibrator--the biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770–5 (2012).
42. Zhang, X., Yang, W. & Dong, C. Levoglucosan formation mechanisms during cellulose pyrolysis. *J. Anal. Appl. Pyrolysis* **104**, 19–27 (2013).
43. Zheng, M. *et al.* Initial reaction mechanisms of cellulose pyrolysis revealed by ReaxFF molecular dynamics. *Fuel* **177**, 130–141 (2016).

## Chapter 2 – Supporting information for the manuscript

### Mechanistic considerations on levoglucosan kinase and its irreversible reaction

#### 1. Experimental

For cloning of LGK with N-terminal His-tag, the gBlocks® Gene Fragments were amplified in pJET 1.2/blunt (Thermo Fisher Scientific). After restriction digestion with *Bam*HI and *Hind*III, the LGK gene was ligated into pQE-30 with T4 DNA ligase. The correct assembly was confirmed by sequencing (Microsynth) and the vector was used for transformation of *E. coli* BL21 DE3. Shaken flask cultivation of the expression host was carried out as described in the main part, using 100 µg/mL ampicillin as selection marker. The harvested cell pellet was dissolved in lysis buffer (50 mM HEPES pH 7.8, 50 mM NaCl, 1 mM EDTA, 0.5 mM DTT), followed by cell disruption by two passes through French press. The centrifuged lysate was applied for purification by Ni-affinity chromatography with the same protocol as LGK with C-terminal His-tag. The elution fractions were concentrated using Vivaspin Turbo 15 columns (10 kDa MWCO) and washed with 50 mM HEPES pH 7.8. This buffer was also used for storage. Size exclusion chromatography was used for further purification of the protein with N-terminal His-tag, using HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) with 50 mM HEPES, 150 mM NaCl pH 7.2. Fractions showing increased UV-absorbance were analyzed by SDS-PAGE. For analysis of inclusion bodies, the disrupted cells were centrifuged and the remaining pellet was washed carefully with  $\frac{1}{5}$  of the original volume 50 mM HEPES for dissolution of the upper grey layer. The remaining white pellet was dissolved in the same buffer. Both pellet fractions were used for SDS-PAGE or activity assays.

The stabilizing effect of substrates and products in different concentrations and combinations on LGK was tested using differential scanning fluorimetry. Therefore, CyproOrange was added to  $\frac{1}{1000}$  conc. to the solutions containing 50 mM buffer pH 7.8, 0.75 mg/mL LGK and various additives (Table 1). Using a BioRad RT-PCR machine, the temperature was increased in steps of 0.5 °C every 5 seconds from 20 °C to 95 °C. The fluorescence (extinction<sub>492</sub>/emission<sub>610</sub>) was continuously measured.

**Table 1: Additives tested for their effect on LGK thermal stability by thermofluor differential scanning fluorimetry.** Every sample contained 0.75 mg/mL LGK, 10 mM MgCl<sub>2</sub> (unless otherwise stated) and 50 mM buffer.

<b>HEPES pH 7.8</b>	20 mM G6P, 20 mM ADP
	100 mM LG
	2 mM ATP
	2 mM ATP, no MgCl <sub>2</sub>
	100 mM G6P
	2 mM ADP
<b>TAPS pH 9</b>	2 mM ATP

Circular dichroism (CD) analysis was carried out for determination of a possible effect of ATP binding on the secondary structure of LGK. Samples contained 20 mM HEPES pH 7.5, 0.5 mg/mL LGK, with or without 2 mM ATP and were measured against a blank containing NaCl-free dialysis buffer instead of LGK. Secondary structures were calculated using dichroweb<sup>1</sup> with CDSSTR method using SP 175 and SMP 180 as a reference data set.

For determination of dependence of LGK reaction on Mg<sup>2+</sup>, increasing amounts of MgCl<sub>2</sub> were added to the reaction mixture (2.5 mM ATP, 500 mM LG, 5 mg/mL BSA, 50 mM HEPES pH 8, 0.02 mg/mL LGK) and the ADP formation within 2 min was detected on HPLC.

The inhibiting effect of fructose, glucose and mannose on LGK was determined using the sugars in concentrations that exceed their K<sub>M</sub>-concentration for *Saccharomyces cerevisiae* type III hexokinase tenfold. 3.3 mM fructose, 1.2 mM glucose or 0.5 mM mannose were added to the reaction in continuous mode assay and the remaining activity was determined as described previously.

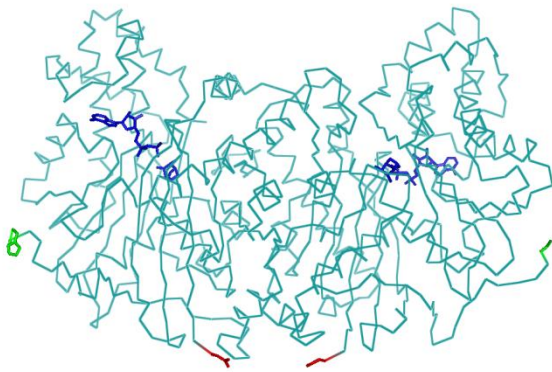
## 2. Results and Discussion

### 2.1. Purification of LGK with N- and C-terminal His-tag

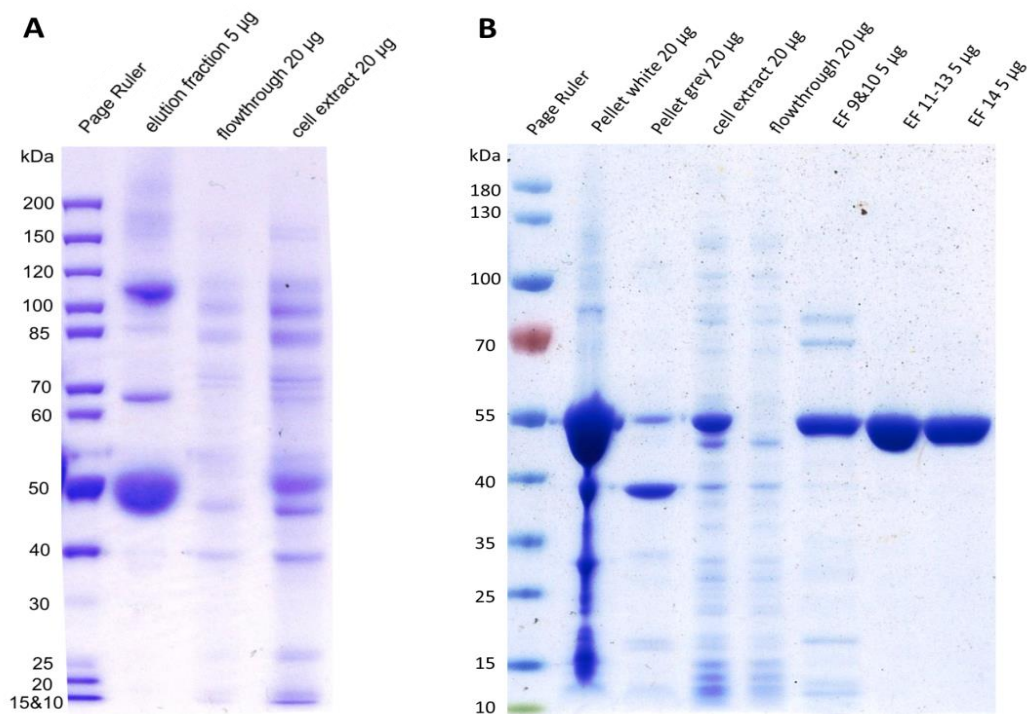
The purification of LGK by Ni affinity chromatography was shown to be less reproducible when the His-tag was placed on the N-terminus than when a C-terminal His-tag was used. Especially when using decreased temperatures of 18 °C instead of 30 °C for expression of the enzyme with N-terminal tag, no active LGK could be recovered. Instead, most of the enzyme was found in the flowthrough. No



activity could be found in the grey and white portion of the pellet from this cultivation, although a significant amount of LGK was observed in these fractions on SDS-PAGE (not shown). The difficulties in purification lead to the assumption that the N-termini are more prone to be buried inside the enzyme, especially when expression is carried out at lower temperature. This is supported by the crystal structure which shows that the two N-termini of the dimer are pointing towards each other and might easily interact, preventing efficient binding to the column (Fig. S 1). Further, several bands of contaminating proteins can be seen on the SDS-gel of the N-terminally tagged enzyme after purification, whereas the C-terminal tag enabled purification to 100 % apparent purity (Fig. S 2). Overexpression was always less predominant using the N-terminal His-tag than with C-terminal tag. For LGK with the C-terminal His-tag, it can be also seen that the inclusion bodies contained a high proportion of LGK with 48 kDa. However, also in the cell extract, overexpression is clearly visible (Fig. S 2B).

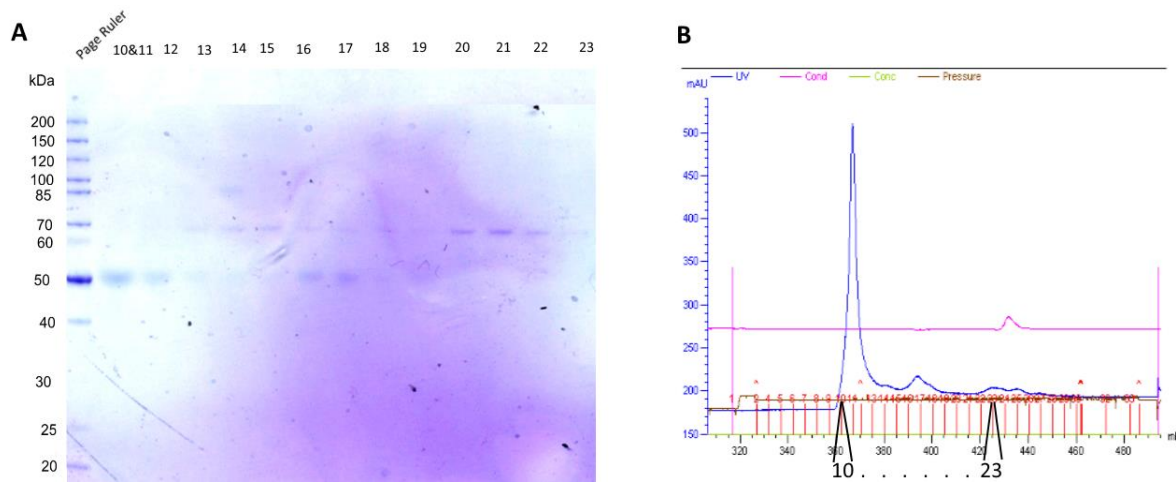


**Fig. S 1: Crystal structure of *L. starkeyi* LGK dimer bound to ADP, magnesium and levoglucosan (4ZLU).** The N-termini are marked in red. The more exposed C-termini are marked in green. The co-crystallized components in the active site (ADP, levoglucosan, MgCl<sub>2</sub>) are highlighted in dark blue. The graphical representation was prepared using PyMOL<sup>2</sup>.



**Fig. S 2: Purification of LGK (48 kDa) by HisTrap affinity chromatography. A) N-terminal His-tag.** Several contaminating proteins can be seen when N-terminally tagged protein was used. Further, purification was not reproducible. **B) C-terminal His-tag.** Elution fractions (EF) 11 – 14 do not show any contaminating proteins and were pooled for use as a biocatalyst. The inclusion bodies (pellet white) contain a great proportion of LGK. In both flowthrough fractions, no LGK is visible. Standards: PageRuler™ Unstained (A) and Prestained (B) Protein Ladder

Size exclusion chromatography was applied to the preparation of LGK with N-terminal His-tag after affinity chromatography purification to remove residual impurities. On SDS-gel, it can be seen that LGK was mainly not eluted according to its size (48 kDa) or the size of its dimer but at the very beginning of the UV-signal, where proteins of much greater size would have been expected. This indicates formation of large agglomerates that could not diffuse into the pores of the SEC-column (Fig. S 3). This also confirms an interaction of the His-tags with each other, resulting in the protein eluting earlier than it would be expected.

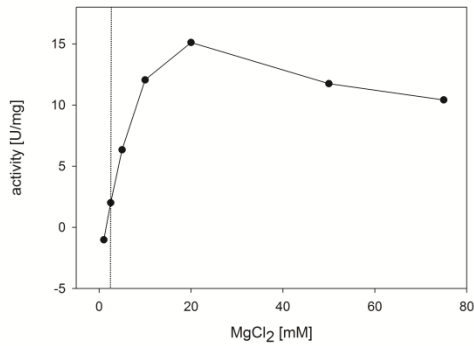


**Fig. S 3: Purification of N-terminally tagged LGK by size exclusion chromatography.** A) SDS-PAGE, B) UV-Signal (blue). Numbers shown are the elution fractions in their order of collection. LGK (48 kDa) is eluted in the very first fractions showing a UV-signal, which does not correspond to its size. A second signal for LGK is visible in the fractions 16 and 17, which might correspond to the actual size of the dimer. All other contaminants were diluted beyond detection range, except for a protein of around 65 kDa, which shows its predominant signal in fractions 20 – 23, after the signal for LGK dimer. In later fractions showing a UV-signal during purification, no proteins could be visualized on SDS-PAGE.

## 2.2. Further enzyme characterization

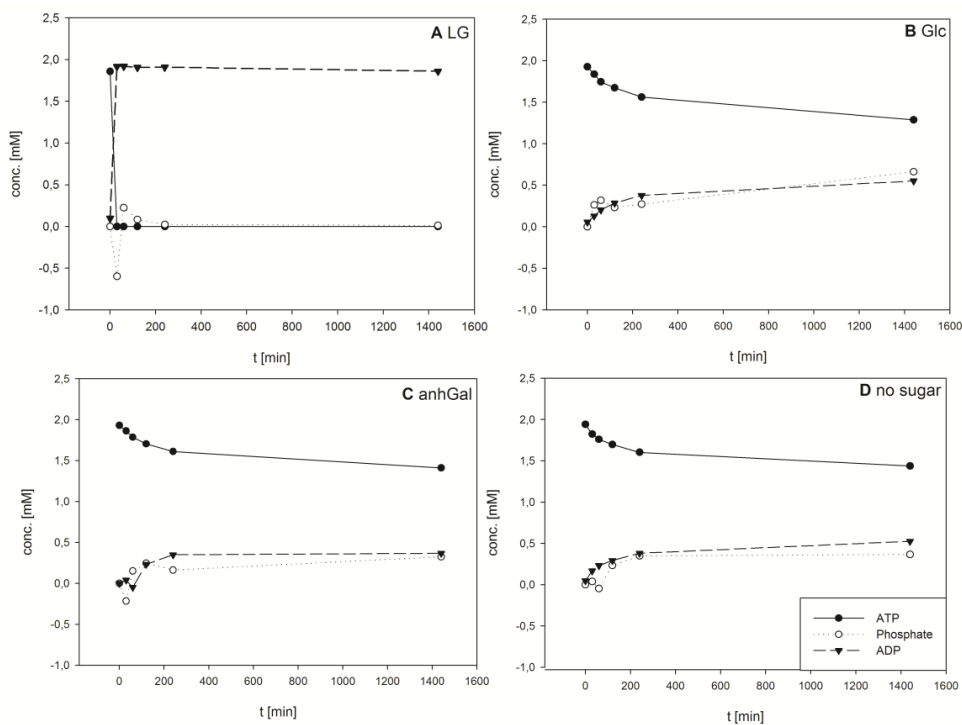
The  $T_m$  of LGK without additives in HEPES was determined to be 37 °C. Only the addition of 2 mM ATP and 10 mM  $MgCl_2$  had a stabilizing effect on the enzyme, increasing the  $T_m$  to 40 °C. All other combinations and concentrations of substrates and products had no significant effect on the enzyme's melting temperature. Circular dichroism measurements did not show any changes in secondary structure due to binding of 2 mM ATP.

Using 2.5 mM ATP, LGK activity increases in dependence on  $MgCl_2$  concentrations until almost tenfold excess (20 mM) is reached. At higher  $MgCl_2$  concentrations, LGK activity is slightly inhibited, likely due to the tight binding of the complex of ADP with two  $Mg^{2+}$  ions to the active site, resulting in slow product release<sup>3,4</sup>.



**Fig. S 4: LGK activity in dependence of MgCl<sub>2</sub> concentrations.** The dotted line indicates equimolar concentrations of ATP and MgCl<sub>2</sub>, where saturation of one magnesium binding site should be almost given, assuming a K<sub>D</sub> of 50 - 100 μM for the ATP-Mg complex<sup>5</sup>. However, the rate of levoglucosan phosphorylation increases significantly up to nearly 10 x excess of MgCl<sub>2</sub> over ATP and slows down at higher concentrations. An inhibiting effect by the stable binding of ADP and Mg<sup>2+</sup> as for cyclin dependent kinases<sup>4</sup> might be the underlying rationale behind this observation.

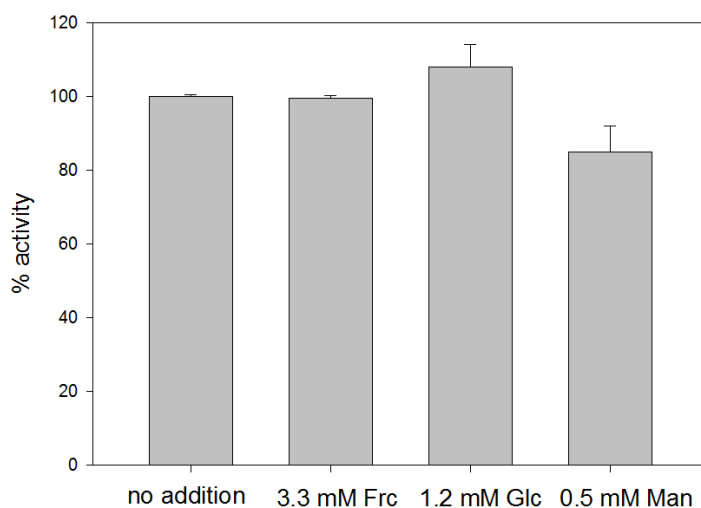
As already described in the main part, LGK exhibits strict substrate specificity. Only with levoglucosan, substantial change in adenosine nucleotides could be observed without massive phosphate release. With all other sugars, as well as in the control reaction without any sugar, unproductive ATP hydrolysis was observed, while phosphate concentrations increase to the same extent as ATP decreases (Fig. S 5).



**Fig. S 5: Substrate specificity of LGK.** Different sugars were tested for their acceptance by LGK. Changes in adenosine nucleotides and phosphate were determined by HPLC and Saheki molybdate assay respectively. Tested sugars were glucose, levoglucosan, anhydro-mannose, anhydro-galactose, anhydro-maltose and anhydro-cellobiose. Apart from levoglucosan (A),

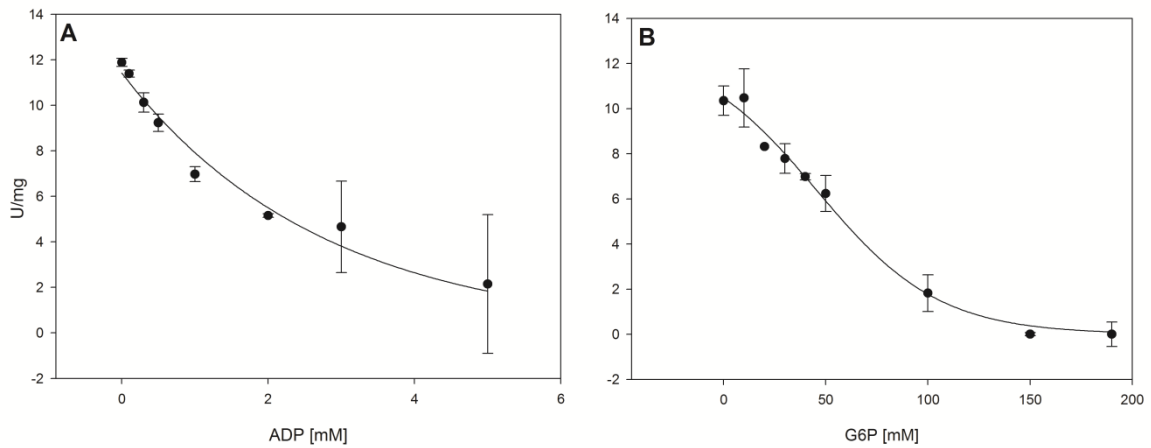
none of the tested alternative substrates was used by the enzyme. Reaction courses of glucose (B) and anhydro-galactose (C) are shown as an example of substrates which are not accepted by LGK. Unspecific hydrolysis of ATP occurs in presence of these sugars like it does in the absence of any sugar (D). No change in adenosine nucleotides and phosphate was observed in the negative control without enzyme (not shown). Reaction conditions: 100 mM sugar, 2 mM ATP, 10 mM MgCl<sub>2</sub>, 50 mM HEPES pH 7.8 and 1 mg/mL LGK.

LGK was not inhibited by the addition of different hexoses in concentrations that exceed their respective  $K_M$  for *S. cerevisiae* hexokinase tenfold. Only the addition of 0.5 mM mannose resulted in a slight decrease of LGK reaction rate by 16 % (Fig. S 6). This would allow the use of these sugars in a recycling system for ADP. In case of successful levoglucosan synthesis from glucose-6-phosphate and ADP, the concentration of the nucleotide product should be kept low to avoid the back-reaction. Using glucose and hexokinase, the simultaneous recycling of ADP and glucose-6-phosphate would be possible. As no productive ATP-formation was observed, this system was not applied, but is still worth mentioning for future considerations.



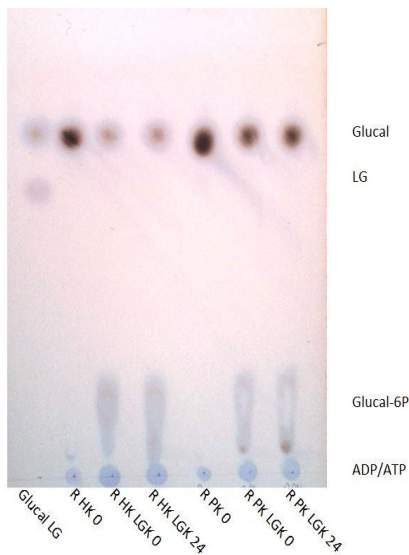
**Fig. S 6: Inhibition of LGK by various sugars.** Fructose (Frc), glucose (Glc) and mannose (Man) were added in concentrations that exceed their  $K_M$  for *S. cerevisiae* hexokinase tenfold. Only mannose shows slightly inhibiting effects at this concentration, decreasing LGK activity by 16 %.

Inhibition of LGK by the sugar and nucleotide products was examined to determine their optimum concentrations in the reversion of the natural reaction. The apparent  $K_I$  value for ADP was 1.7 mM and 54 mM for G6P (Fig. S 7).



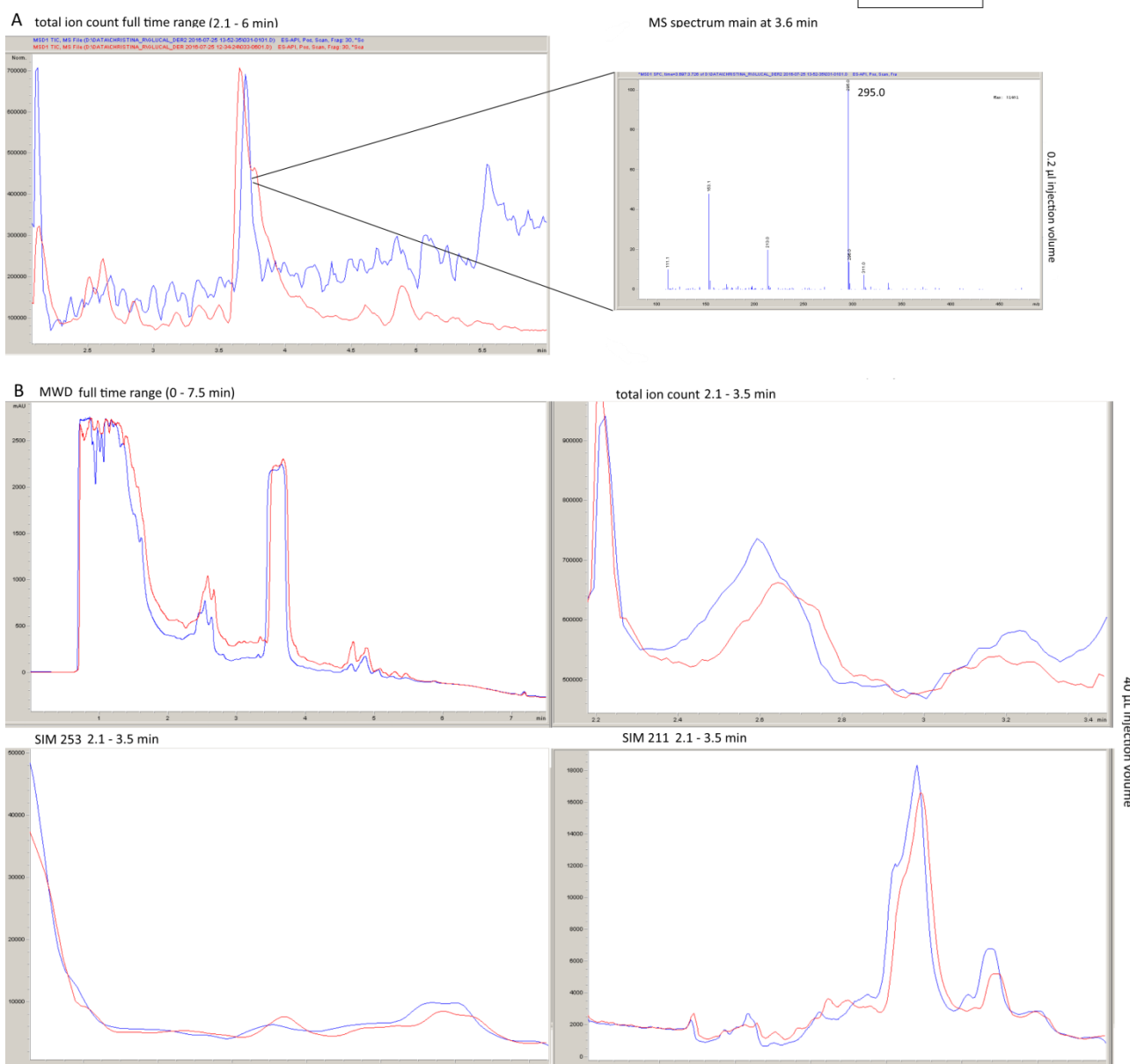
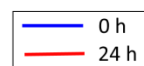
**Fig. S 7: Inhibiting effect of A) ADP and B) G6P on LGK activity.** The apparent  $K_i$ -values are 1 mM for ADP and 54 mM for G6P. Reaction conditions: 0 – 20 mM ADP or 0 – 190 mM G6P, 2.5 mM ATP, 300 mM LG, 50 mM HEPES pH 8, 5 mg/mL BSA, 20 mM  $MgCl_2$ , 0.02 mg/mL LGK.

### 2.3. Synthesis of substrate analogues



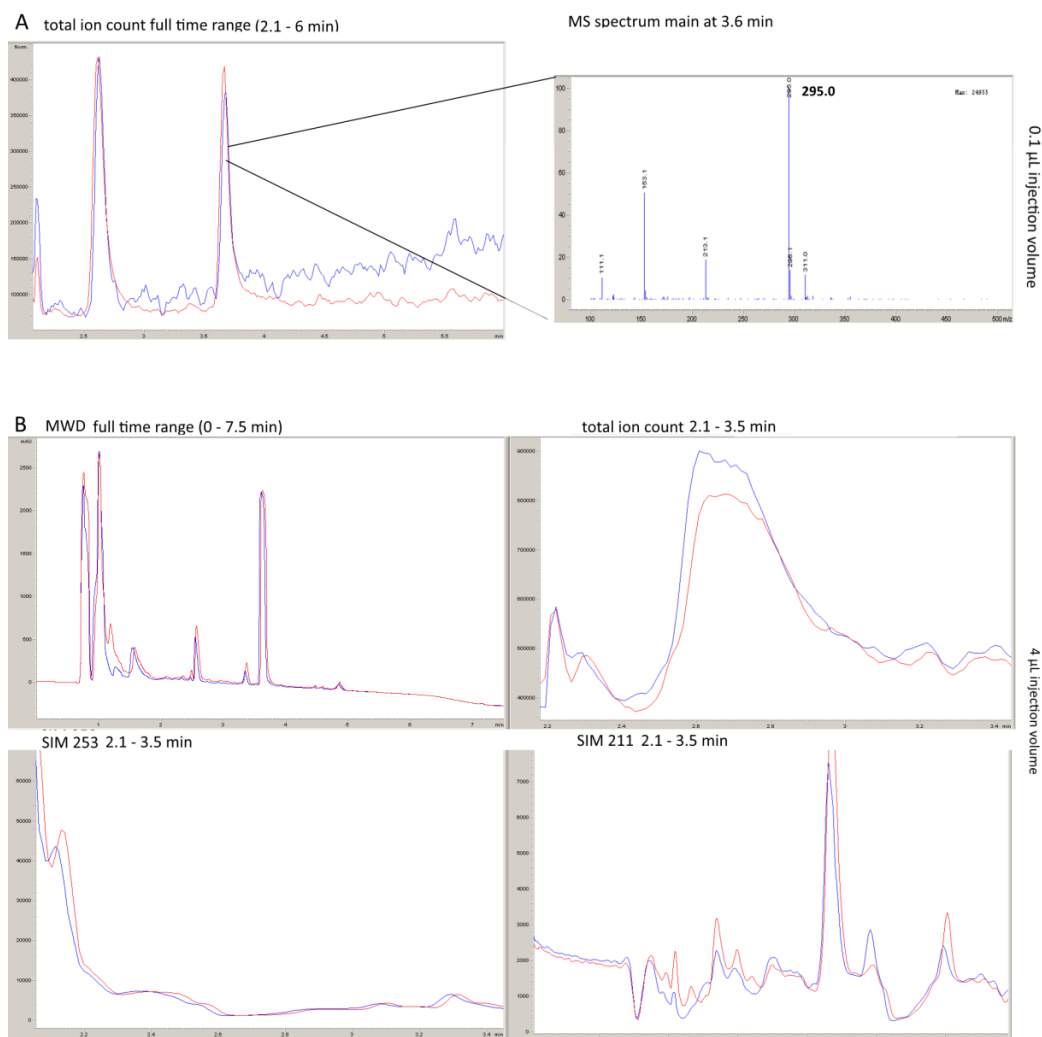
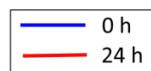
**Fig. S 8: TLC showing conversion of Glucal with hexokinase and levoglucosan kinase.** 24 h after incubation of glucal with hexokinase (HK), the substrate concentration is significantly decreased. Instead, new spots are visible with lower  $R_f$ -values that belong to the phosphorylated glucal. However, within 24 h of incubation with LGK, no substance resulting in a spot with the retention of levoglucosan is produced. The clearly visible standard contained each 20 mM glucal and levoglucosan. 1  $\mu$ L of standard and 1.5  $\mu$ L of samples were loaded on the TLC silica plates. R HK 0: reaction without ATP recycling at t0 after HK addition, R HK LG 0: reaction without ATP recycling 24 h after HK addition, 0 h after LGK addition, R HK LGK 24: reaction without ATP recycling 48 h after HK addition, 24 h after LGK addition. R PK: corresponding to R HK, but each with ATP recycling using pyruvate kinase.

### Reaction from Glucal-6 Phosphate to Levoglucosan, without recycling system



**Fig. S 9: MS analysis of LGK reaction with glucal-6-phosphate without ATP recycling system.** Sugars were peracetylated using acetic anhydride/pyridine. **A)** Complete peracetylation of residual glucal can be seen from the mass spectrum of the predominant mass signal at 3.6 min retention time. **B)** To enable higher injection volumes, mass analysis was stopped before elution of peracetylated glucal. Neither MWD nor SIM signals for different ion adducts of mono- or diacetylated 2-deoxy-levoglucosan show any differences between samples after 0 and 24 h of incubation with LGK. 253 [M+23]<sup>+</sup>: Na<sup>+</sup>-adduct of diacetylated 2-deoxy-LG; 211 [M+23]<sup>+</sup>: Na<sup>+</sup>-adduct of monoacetylated 2-deoxy-LG, 269 [M+39]<sup>+</sup> (not shown): K<sup>+</sup>-adduct of diacetylated 2-deoxy-LG.

## Reaction from Glucal-6 Phosphate to Levoglucosan, with Pyruvate Kinase ATP recycling



**Fig. S 10: MS analysis of LGK reaction with glucal-6-phosphate with ATP recycling by pyruvate kinase.** Sugars were peracetylated using acetic anhydride/pyridine. **A)** Complete peracetylation of residual glucal can be seen from the mass spectrum of the predominant mass signal at 3.6 min retention time. **B)** To enable higher injection volumes, mass analysis was stopped before elution of peracetylated glucal. Neither MWD nor SIM signals for different adducts of mono- or diacetylated 2-deoxy-levoglucosan show any differences between samples after 0 and 24 h of incubation with LGK. 253 [M+23]<sup>+</sup>: Na<sup>+</sup>-adduct of diacetylated 2-deoxy-LG; 211 [M+23]<sup>+</sup>: Na<sup>+</sup>-adduct of monoacetylated 2-deoxy-LG, 269 [M+39]<sup>+</sup> (not shown): K<sup>+</sup>-adduct of diacetylated 2-deoxy-LG.

## References

- Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–73 (2004).
- Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.8. (2015).
- Bacik, J.-P. *et al.* Producing glucose 6-phosphate from cellulosic biomass: structural insights into levoglucosan bioconversion. *J. Biol. Chem.* **290**, 26638–48 (2015).



4. Jacobsen, D. M., Bao, Z.-Q., O'Brien, P., Charles L. Brooks, I. & Young, M. A. Price To Be Paid for Two-Metal Catalysis: Magnesium Ions That Accelerate Chemistry Unavoidably Limit Product Release from a Protein Kinase. (2012).
5. O'Sullivan, W. J. & Smithers, G. W. Stability constants for biologically important metal-ligand complexes. *Methods Enzymol.* **63**, 294–336 (1979).

## Chapter 3 – Additional experiments for characterization of levoglucosan kinase

### Supporting Information 2

#### 1. Introduction

This chapter covers experiments that were carried out during the master thesis but were not considered relevant enough to be described in the manuscript. This includes small scale expression studies of LGK for determination of highly overexpressing colonies and the optimal time of cultivation until cell harvest. Further, the efforts to analyze the degradation product formed in the LGK reactions within long incubation times are described.

#### 2. Experimental

Cloning and expression of *Lipomyces starkeyi* LGK with N- or C-terminal His-tag have already been described in the main part and supporting information of the manuscript. After transformation of the expression host with pQE-30\_LGK and pET-41\_LGK, different colonies were examined for their overexpressing capacities in small scale (10 mL). For C-terminally tagged LGK, samples were taken at 0, 3 and 16 h after induction and analyzed on SDS-PAGE and dot blot. BugBuster solution (9 mM HEPES pH 7, 1 x BugBuster®, Merck, 1000 x diluted Benzonase) was used for cell disruption. Therefore, the cell pellet was mixed with 10 µL BugBuster solution per mg pellet, incubated at 24 °C and 900 rpm and centrifuged at 4 °C, 13000 rpm for 15 min. After removal of the supernatant (soluble fraction), the pellet was dissolved in 6 M urea and the supernatant after another 5 min of centrifugation was used as the insoluble fraction. Both fractions were analyzed on SDS-PAGE or dot blot. For dot blot analysis, samples were additionally cooked in 4 % SDS, 200 mM NaOH and 10 % DTT and spotted on Roth Roti-NC Nitrocellulose membranes. The dried membrane was soaked in 5 % BSA/TBS-T (20 mM Tris-HCl, 150 mM NaCl, 0.05 % Tween 20), incubated with anti-His-tag (D3I10)XP(R) rabbit mAB (Cell Signaling Technology) in BSA/TBS-T and HRP-linked goat anti-rabbit IgG (Cell Signaling Technology). Thorough washing with TBS-T after each incubation step was carried out, followed by a last washing step with TBS without Tween 20 before application of the detection reagent. Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific) was used for generation of chemiluminescence which was detected in a G:Box gel imaging system (Syngene).

Initial reactions for formation of levoglucosan were set up with lower substrate concentrations compared to experiments described in chapter 1. 20 mM G6P, 20 mM ADP, 10 mM MgCl<sub>2</sub> and 28 mM HEPES pH 7.8 were incubated with 1 mg/mL LGK for 72 h. Concentrations of substrates and co-factors, as well as buffer pH were varied afterwards in many different ways (Table S2. 1).

**Table S2. 1: Variations in the conditions tested for enzymatic synthesis of levoglucosan by LGK.** Compounds printed in bold are the ones that vary in comparison to the original version (version 1). Version 0 was used in initial experiments and contains suboptimal buffer and enzyme concentrations.

	<b>version</b>	<b>G6P</b> [mM]	<b>ADP</b> [mM]	<b>MgCl<sub>2</sub></b> [mM]	<b>Buffer</b> [mM]	<b>Enzyme</b> [mg/mL]
	0	20	20	10	14	0,44
<b>HEPES pH 7.8</b>	1	20	20	10	50	0,5
	2	<b>100</b>	<b>2</b>	10	50	0,5
	3	<b>10</b>	<b>10</b>	10	50	0,5
	4	<b>5</b>	<b>5</b>	10	50	0,5
	5	<b>1</b>	<b>1</b>	10	50	0,5
	6	<b>20</b>	<b>0,2</b>	10	50	0,5
	7	<b>5</b>	<b>0,2</b>	10	50	0,5
	9	20	20	<b>1</b>	10	0,5
	10	20	20	<b>50</b>	10	0,5
<b>TAPS pH 9</b>	11	20	20	10	50	0,5
<b>MES pH 6.5</b>	12	20	20	10	10	0,5

As explained in chapter 1 for experiments carried out after more thorough characterization of the enzyme, various components of the reaction mixture were measured in different ways. HPX-87H column (BioRad) was used with 0.8 mL/min 5 mM H<sub>2</sub>SO<sub>4</sub> on HPLC system for detection of G6P, phosphate and a possible product on RID and detection of ATP and ADP by UV-detector at 259 nm. Further, Phenomenex Kinetex C18 HPLC column was used with 2 mL/min flow (40 mM TBAB, 20 mM KH<sub>2</sub>PO<sub>4</sub> pH 5.9) for detection and quantification of the different adenosine nucleotides using the MWD at 259 nm. Spectrophotometric assays used were the non-continuous G6P-DH assay for quantification

of G6P and the Saheki molybdate assay for phosphate quantification. Detection of glucose in the reaction mixture was carried out using a dehydrogenase assay similar to the non-continuous G6P-DH assay described in chapter 1. Therefore,  $\frac{1}{10}$  volume of the sample was added to  $\frac{9}{10}$  volume of assay mixture containing 100 mM Tris-HCl pH 8, 1.5 mg/mL NAD<sup>+</sup> and 2 U glucose dehydrogenase ('Amano' 2, Amano Enzyme Inc., Japan). A<sub>340</sub> was measured over time and compared to standards containing 1-10 mM glucose. Further, samples were spotted onto TLC silica plates and placed into a saturated chamber of different mobile phases. Butanol:ethanol:propanol:H<sub>2</sub>O in a ratio of 3:3:2:2 proved to be the best choice for separation of all predominant compounds. For detection, concentrated H<sub>2</sub>SO<sub>4</sub> and thymole-solution (5 mg/mL thymole in 5% H<sub>2</sub>SO<sub>4</sub>, 95% EtOH) were sprayed onto the plates and the plates were heated to 80 – 100°C until the spots became visible.

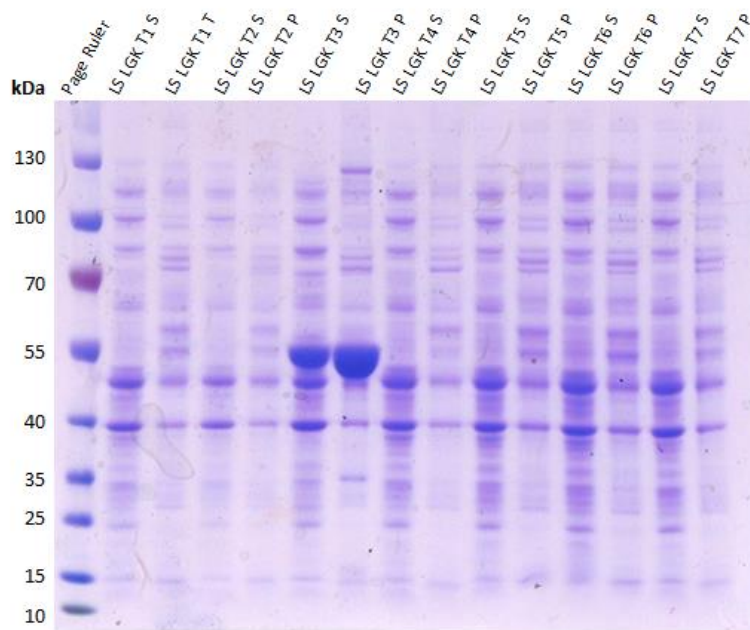
For identification of the degradation product formed in the LGK reactions within 72 h, samples were stopped by heat and partly purified for MS-analysis. First, purification by SuperQ anion exchange as described in the main part was carried out by passing the solutions through the column manually. All fractions starting from sample loading until washing with 2 M NaCl were collected and analyzed on HPX-87-H-column. For separation of the charged PY from the other compounds, adenosine nucleotides were digested by addition of 5 U CIP (Calf intestinal alkaline phosphatase, BioLabs) in a maximum of 200 µL sample. Complete digestion was confirmed by analysis of an aliquot on Kinetex column. Afterwards, each 60 µL sample were loaded onto HPX-87H column and the fractions where elution of the degradation products was observed on RID were collected. To remove H<sub>2</sub>SO<sub>4</sub> prior to MS analysis, samples were concentrated on SpeedVac System (ISS110, Thermo Fisher Scientific) and additionally purified with Phenomenex Kinetex 5 µ PFP 100A column using 1 mL/min 0.1 % formic acid, 5 % ACN. Collected fractions were again analyzed by HPX-87H column to confirm the presence of the desired compounds and used for LC-MS. In positive mode, Agilent Poroshell 120 SB-C18 was used at 40 °C, 0.7 mL/min H<sub>2</sub>O. In negative mode, the same column was used with 0.7 mL/min 10 mM NH<sub>4</sub>HCO<sub>2</sub>.

Reactions were further set up in D<sub>2</sub>O for *in situ* NMR-measurements using decreased buffer concentrations of 10 mM HEPES pH 7.8 or Tris pH 7.5 to minimize buffer effects. As in the initial reaction setup, samples further contained 20 mM G6P, 10 mM MgCl<sub>2</sub> and 20 mM ADP. Control reactions without ADP were additionally set up. Over 124 h, samples were kept at 30 °C and <sup>1</sup>H spectra were taken once a day, using the same equipment as described in the main part. Additionally, several 2D spectra (COSY, HSQC, HMBC) were additionally recorded at the last time point of measurements.

### 3. Results and Discussion

#### 3.1. Expression and analysis in small scale

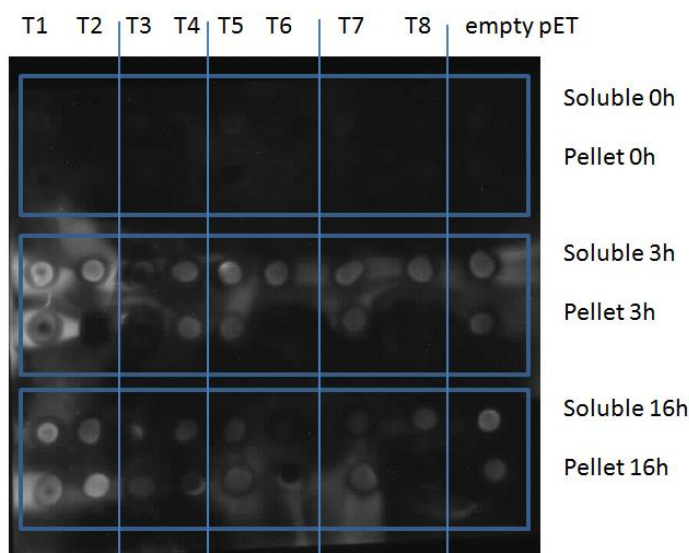
Before expression in large scale, different colonies after transformation were grown in small scale and tested for their expression of LGK in the soluble and insoluble fraction. Great differences could be seen for the LGK with N-terminal His-tag. Only one colony showed overexpression of the protein of interest on SDS-PAGE (Fig. S2. 1). This colony was used for large scale expression afterwards.



**Fig. S2. 1: Small scale expression of LGK from different colonies after transformation.** Only one colony (T3) showed significant overexpression and was chosen for further cultivation.

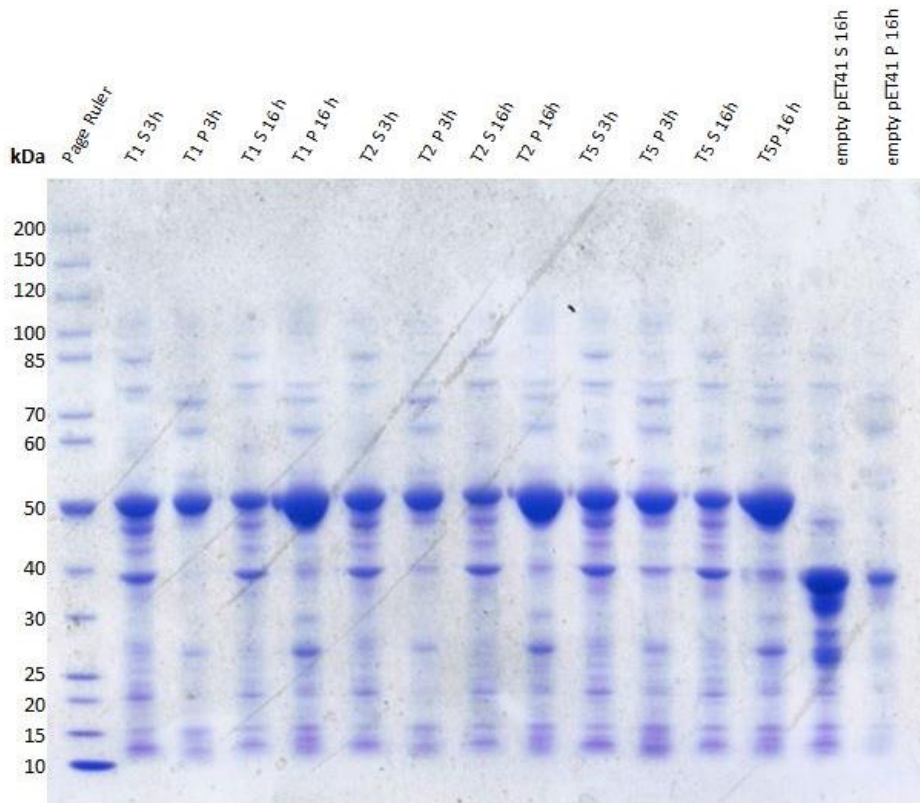
For LGK with C-terminal His-tag, differences in overexpression of the single colonies after transformation were not as significant as for the kinase with N-terminal tag. In the dot blot, all of the cell lysates from small scale LGK expression show clear signals of an overexpressed protein with a His-tag. Although the dot blot does not seem to be perfectly reliable for quantification of expression, it indicates that the LGK signals in the soluble fractions are higher in the samples 3 h after induction than in the samples 16 h after induction. The cultivation of empty pET 41-vector also shows a signal on dot blot, which results from the number of tags, including a C-terminal His-tag, which are placed after the promotor. This leads to the expression of a small peptide that is responsible for the signal on the blot (Fig. S2. 2). To prove the higher solubility of LGK after 3 h of induction, soluble and insoluble fractions of the cell lysates of three colonies 3 h and 16 h after induction were analyzed on SDS-PAGE. The same volume of protein preparation was always loaded on the gel, ensuring comparability between the different samples. It was shown that the concentration of LGK in the soluble fraction is in fact highest

3 h after induction. 16 h after induction, the band of the protein of interest is a bit weaker in the soluble fraction compared to the samples 3 h after induction, but the signal in the insoluble fraction is significantly higher (Fig. S2. 3). It has to be taken into account though that the sample preparation was normalized to the weight of the pellet – the more mg of pellet was measured after cell harvest, the more cell disruption solution was added. The higher the mass of cells expressing proteins, the more the samples were thus diluted. The yield of soluble protein was probably higher after 16 h of expression as the mass of the pellet was increased due to larger amounts of enzyme entrapped within the cells. Still, expression was stopped 3 h after induction as the proportion of well-folded protein was higher compared to longer incubation, where more inclusion bodies were formed. This avoids problems in purification as were observed with N-terminally tagged enzyme and still provides sufficiently high yields in highly active LGK. 45 mg of purified protein with 15 U/mg were recovered after buffer exchange by dialysis, resulting in 657 U per L fermentation broth. Only 38 mg LGK of lower purity and (specific) activity could be obtained during purification of LGK with N-terminal His-tag, resulting in a yield of 8.55 U per L fermentation broth. However, this value might have been higher initially, as activity of the purified protein was just measured after 3 weeks of storage at 4 °C. At this temperature, decreased LGK activities by about a half of the initial value were observed within two weeks. Taking the lower purity of the enzyme preparation with N-terminal His-tag compared to the one with C-terminal tag and activity loss during storage into account, the difference in activity yields is still substantial. Thus, it must be assumed that the His-tag on the N-terminus also interferes with the enzyme's activity.



**Fig. S2. 2: Dot blot of cell lysates from LGK-cultivation different time points after induction.** Eight colonies after transformation of LGK with C-terminal His-tag were cultivated in small scale. The cell lysates were spotted on a nitrocellulose

membrane and enzymes were detected by Anti-His goat antibody and HRP-coupled Anti-goat rabbit IgG. T1 – T3 were further analyzed on SDS-PAGE.

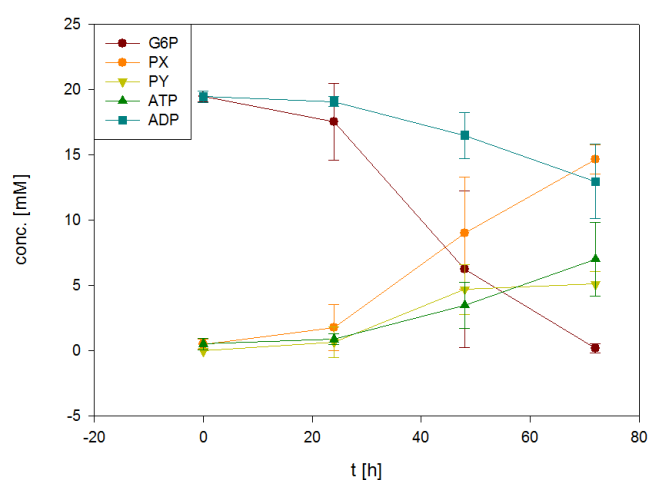


**Fig. S2. 3: SDS-PAGE of soluble and insoluble fraction of cell lysates from LGK expression.** Soluble and insoluble fractions from the cultivation of three single colonies after transformation (T1 – T3) were analyzed 3 and 16 h after induction. The same volume of cell extract was each loaded for comparison of LGK concentration in the different samples. After 16 h, the concentration of LGK (48 kDa) in the insoluble fraction each increases, whereas the concentration in the soluble fraction decreases slightly compared to the corresponding sample 3 h after induction. In the cells harboring empty vector, a smaller peptide is transcribed, which consist of the different tags that have not been removed.

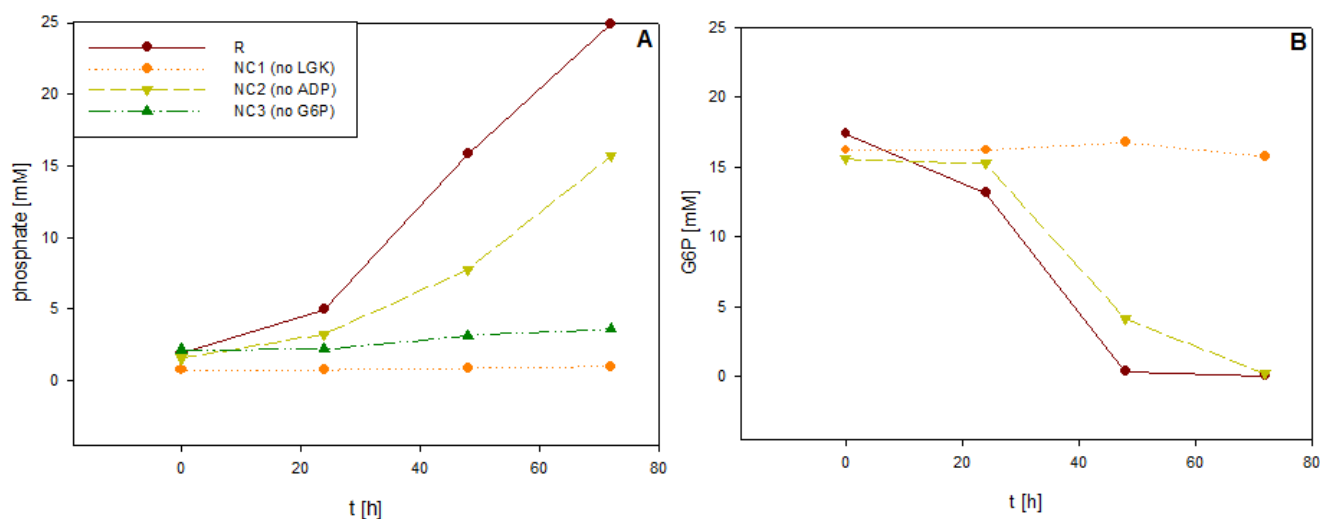
### 3.2. Characterization of a side-product formed from glucose-6-phosphate by LGK enzyme preparations

During incubation of LGK with ADP and G6P in equimolar, low concentrations (20 mM), almost complete depletion of G6P from the reaction mixture was observed within 72 h using HPX-87H column with RID. Simultaneously, conversion of ADP to ATP was observed on the UV-detector, and the formation of two new products (named PX and PY) was detected on RID (Fig. S2. 4). As retention of G6P on HPX-87H column was very low, G6P concentrations were also confirmed by G6P-DH assay (Fig. S2. 5B). Similar profiles of substrate consumption and formation of PX and PY were also observed in the reaction mixtures omitting ADP, although in even lower rates (Fig. S2. 5). As PY had very similar retention times on HPX-87H column as levoglucosan, 1:10 diluted reactions of LGK with G6P without

ADP were spiked with 2 mM levoglucosan and analyzed with 0.4 mL/min flow of the mobile phase instead of 0.8 mL/min. A shoulder in the resulting signal at around 19.5 min shows that PY is not levoglucosan (Fig. S2. 6). Comparison of the retention time of PY on HPX-87H column with the elution profile of glucose, glucal and gluconic acid also showed that none of these compounds had the same retention as the unknown degradation product. Further, glucose could not be detected by a dehydrogenase assay, which was sensitive enough to detect 1 mM of glucose standard. PX was confirmed by a spiking experiment to be free phosphate, which also correlates with the increasing phosphate concentrations that were observed using Saheki assay (Fig. S2. 5A). No changes were observed in the reaction mixture omitting enzyme.



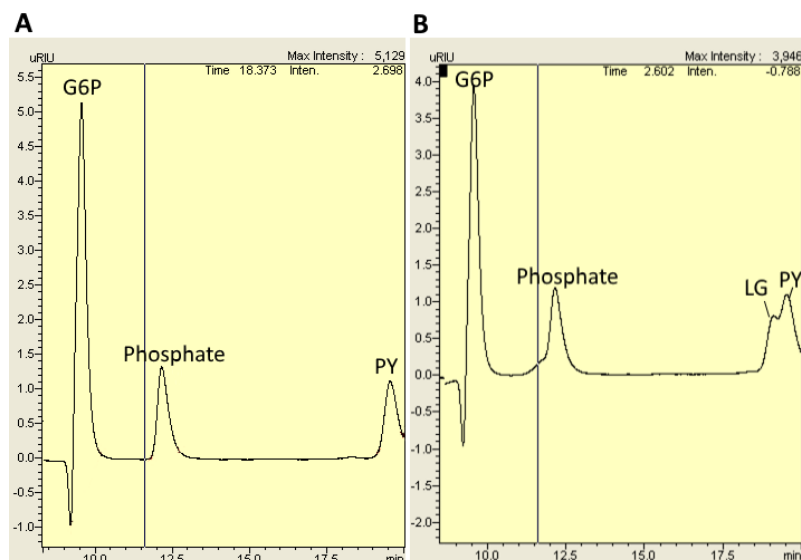
**Fig. S2. 4: Conversion of G6P during 72 h of incubation with LGK.** Samples were measured with HPLC using HPX-87H column equipped with RID and MWD (259 nm). Reaction conditions: 20 mM G6P, 20 mM ADP, 10 mM MgCl<sub>2</sub>, 22 mM HEPES pH 7.8, 1 mg/mL LGK.



**Fig. S2. 5: Incubation of G6P and ADP with LGK. A)** Phosphate concentrations in the reaction solution and the negative controls without enzyme (NC1), without ADP (NC2) and without G6P (NC3) were determined by Saheki molybdate assay. **B)**

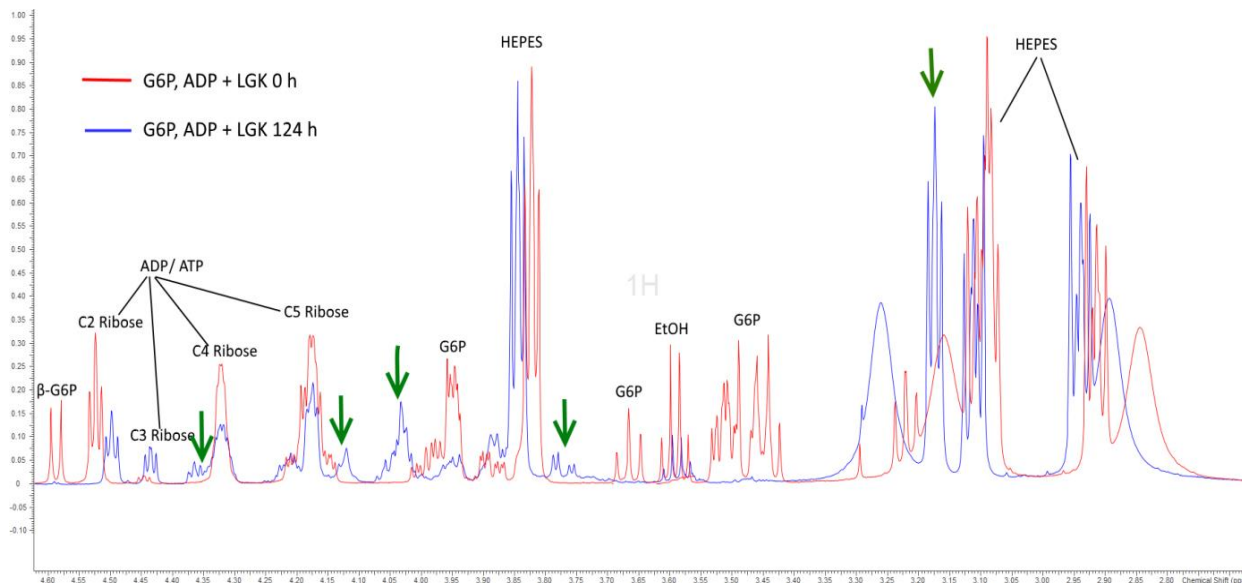


Conversion of G6P in the reaction and NC1 and 2, determined by G6P-DH assay. In the reaction mixture, as well as in the control reaction without ADP, an increase in phosphate and a decrease in G6P were observed.



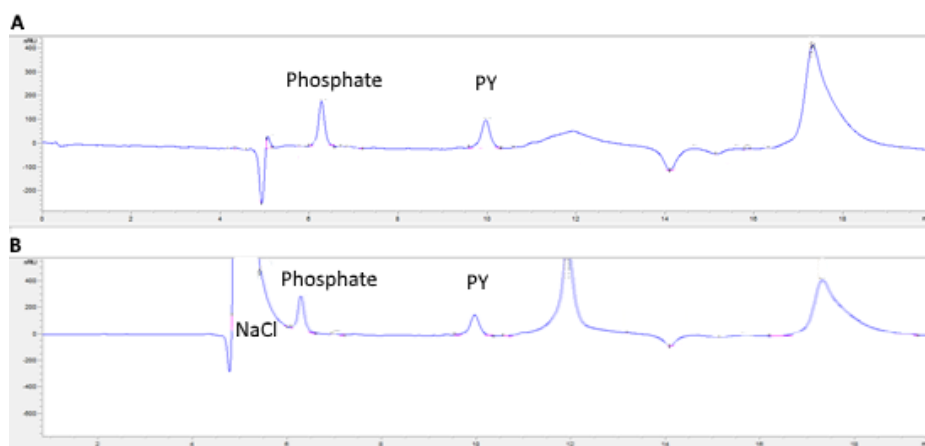
**Fig. S2. 6: Reaction of LGK with G6P (without ADP).** Samples were analyzed on HPX-87H column with 0.4 mL/min  $\text{H}_2\text{SO}_4$ . **A)** The 1:10 diluted reaction mixture shows a signal for G6P, PX (phosphate) and PY (not identified), which has a similar retention time to levoglucosan. **B)** Spiking of the sample with 2 mM levoglucosan shows that PY is not the desired compound.

NMR and MS-analysis were carried out for identification of PY.  $^1\text{H}$  Spectra of the reaction 5 h and 124 h after addition of LGK show complete depletion of G6P and formation of several new signals (Fig. S2. 7). The same observations were again made in the reaction solution with and without ADP. However, neither 1D nor 2D measurements (HSQC, COSY, HMBC) could unambiguously solve the structure of the unknown side-product. Several signals were found to overlay with signals from the buffering substance HEPES. Replacing HEPES by the Tris as an alternative buffer with less signals in  $^1\text{H}$  spectra did not fully solve this problem.



**Fig. S2. 7:** <sup>1</sup>H NMR of the reaction of G6P and ADP at low concentrations with LGK. Both substrates were used in concentrations of 20 mM. <sup>1</sup>H spectra were taken after 0 h (red) and 124 h (blue) incubation. The decrease of glucose-6-phosphate in D<sub>2</sub>O was even more slowly than in non-deuterated solvents. Still, after 124 h, substrate concentrations are significantly decreased. Newly formed signals are marked with green arrows. Several shifts of the HEPES signals are visible, which is partly due to change of pH during the measurements. However, some new signals are also formed in this region. The nature of the product could not be solved from these measurements. EtOH..contaminating ethanol in the sample.

Purification of PY by anion exchange was not possible, as this compound bound to the column and was eluted during washing with NaCl along with the other charged substances including the adenosine phosphates and the free phosphate (Fig. S2. 8). This suggests that PY is a charged molecule as well.

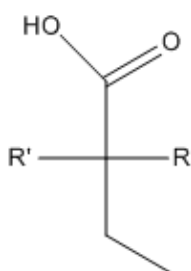


**Fig. S2. 8:** Purification of PY with SuperQ anion exchange on HPX-87H column. A) RID signal of the untreated sample, B) Washing fraction with NaCl. PY is eluted along with the charged molecules as phosphates, ADP, ATP (MWD not shown) when the column was washed with increased salt concentrations. This shows that PY is also a charged molecule at neutral pH.

Alternative sample preparation for MS-analysis for separation of PY from the other charged components included CIP-digestion of residual adenosine nucleotides, fractionation by HPX-87H

column and removal of  $\text{H}_2\text{SO}_4$  by fractionation using PFP-column. Considerable amounts of glycerol were introduced into the samples by the enzyme preparation of CIP, which is why the purified samples could not be used for NMR analysis. In positive mode, no predominant mass for PY could be determined. However, in negative mode measurements, 196 and 172 g/mol were found to be the most abundant masses.

What can be said about the unknown molecule now is that it is a C4 molecule with 196 or 172 g/mol. According to 2D NMR measurements, C1 is most likely a charged carbonyl group and C2 is proposed to have two substituents of unknown origin (Fig. S2. 9).



**Fig. S2. 9: Proposed structure of PY at acidic pH.** From NMR-spectra, the nature of R and R' could not be solved.

The time courses of PY formation within 72 h were not reproducible in their rate and in the extent of product formation. As significant precipitates were also observed, it is assumed that non-enzymatic degradation of the substrate occurs predominantly. The low stability of the enzyme under the used conditions, which has not been known at the time of these experiments, supports this theory. Changes within 24 h are minor except for a small release of phosphate. This is the maximum time span, after which LGK still shows residual activity despite decreased reaction temperature and addition of stabilizing BSA. However, non-enzymatic ATP formation is not likely to occur, which is why contamination of the enzyme preparation by a more stable enzyme cannot be ruled out. When the C-terminally tagged enzyme preparation, which was of higher purity was incubated with this reaction mixture, the same changes were observed within 72 h, although G6P consumption was not as complete as observed previously when LGK preparation with N-terminal His-tag was used. Therefore, the nature of PY was not investigated further. Sample preparation of a larger volume of sample by fractionated HPLC without addition of CIP might have enabled the analysis of PY in the presence of ADP, which is simultaneously eluted from HPX-87H column, by NMR, but as the reaction was so slow and unpredictable, the formation of the partly identified molecule was considered to be of no synthetic or scientific value.

#### 4. Conclusion

It was shown that the proportion of soluble LGK is higher when expression is induced for just 3 h compared to longer post-induction times. A negative effect of the N-terminal His-tag on the enzyme's activity cannot be ruled out. Within 72 h incubation of the LGK preparation with G6P with or without ADP, the substrate concentration decreases significantly. The formation of a new, unidentified compound is accompanied by the release of phosphate and, in reactions with ADP, the increase of ATP concentration. However, these changes cannot be attributed to LGK activity, but more likely to non-enzymatic conversion and activities of protein impurities respectively.

#### Appendix

##### 1. *Lipomyces starkeyi* strain YZ-215 levoglucosan kinase mRNA, complete cds

GenBank: EU751287.1

The part that is marked in grey represents the longest open reading frame in the sequence and was ordered in a codon optimized version for *E. coli* from IDT.

```
AACATCATTCCATCATGCCCATCGCGACTTCCACTGGCGACAATGTGCTCGACTTCACCGTGCTCGGCCTCAACT
CGGGGACGAGTATGGACGGCATCGACTGTGCGCTATGCCACTTTACCAAAAAACTCCCGACGCGCCCATGGA
GTTTGAGCTGCTCGAGTATGGAGAGGTCCCGCTTGCCAGCCCATCAAGCAGCGAGTCATGCGGATGATCTTG
GAGGACACGACATCGCCGTGAGAGCTGTCCGAGGTCAACGTCAATTCTCGGGGAGCACTTTGCCGATGCTGTTG
GACAGTTTGCGGCCGAGCGCAACGTGGACTTGAGCACTATCGACGCGATTGCAAGCCACGGTCAGACGATCT
GGCTGCTGTCCATGCCGGAGGAGGGACAGGTCAAGTCGGCTCTGACCATGGCGGAAGGCGCGATCCTCGCAT
CTCGCACCGGCATCACGTCCATCACCGACTTCGAATCTCCGACCAGGCCGCCGGTCGTCAGGGTGCTCCGCTG
ATTGCCTTCTTCGACGCTCTGCTCCTTACCACCCGACCAAGCTGCGTGCGTGCCAGAACATCGGTGGTATCGC
AAACGTCTGCTTCATCCCTCCCGACGTTGATGGCCGACGCACCGACGAGTACTACGACTTTGACACGGGACCA
GGCAATGTCTTCATAGATGCGGTCGTCCGACACTTCACCAACGGGGAGCAGGAGTACGACAAGGATGGAGCG
ATGGGGAAGCGAGGCAAGGTGGACCAGGAGCTCGTGGATGATTTCTTGAAGATGCCATACTTCCAACCTGGAC
CCTCCCAAGACTACCGGTCCGGGAGGTCTCCGTGATACTCTGGCTCACGACTTGATCCGTCGCGCTGAGGCGA
AAGGACTGTCCCCGATGACATCGTTGCGACGACCACCAGGATTACCGCGCAAGCCATTGTTGACCACTACCG
GCGCTACGCTCCTAGCCAAGAGATCGACGAGATCTTCATGTGCGGCGGAGGCGCGTACAACCCGAACATCGTC
GAGTTCATTAGCAAAGCTACCCTAACACCAAGATCATGATGCTCGACGAGGCTGGGGTCCCCGCTGGAGCAA
AGGAGGCCATCACGTTGCTTGGCAAGGAATGGAAGCCCTTGTGGCCGATCCATCCCTGTCCCACCCGCGT
GGAGACGCGACAACACTACGTGTTGGCAAGGTGTCCCCGGGACTGAACTACCGCAGCGTGATGAAGAAGG
GTATGGCGTTCGGCGGAGACGCGCAGCAGCTGCCGTGGGTGACGAGATGATTGTGAAGAAAAAGGGCAAG
```

GTCATTACCAACAACACTGGGCTTAAATGACATGCGCATTGCAGTTGTGCCACCATTGTACCTCCTCCCCACACATAC  
ACACATCATGACGTGAATGCATTGCTTGAGTAACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## 2. *Aspergillus niger* levoglucosan kinase mRNA, complete sequence

GenBank: AY987027.1

The part that is marked in grey represents the longest open reading frame in the sequence and was ordered in a codon optimized version for *E. coli* from IDT. The codon for Ser22 was changed from TCC to TCG to prevent the appearance of a *Bam*HI restriction site within the gene. However, apart from many unidentified bases on the 3' end, the proposed sequence shows some irregularities concerning the sequence length (800 bp), the longest translated peptide sequence possible (167 aa) and the described size of the putative LGK (75 kDa)<sup>11</sup>. Not even the tetramer of this 17.9 kDa peptide could reach the observed size of the *A. niger* LGK.

AATATATGGCTCCAGAACAACCATGCCACTCCGCCTCTTTGCCAACCGCACCTCCCTAGCAGCCTTCACCCTCA  
CCTTCCTCCACGGCATGCTCCTCTACTGGACTATCTACTTTCTCCCGTCTACTTCCAAGGGGTAAAAGCAGCT  
CCCCTATCCGCTCCGGCGTGCAGCTTCTCCAACAGTCATCGTCACGGTACCAGCGGCCATCGTCGCAGGTGTA  
GTTCTTACGAAGACCGGACGCTACAAACCCATTCAAATTGCCGTTTCGTCTTCATGGCTCTCGGCATGGGCT  
ATTTAGTCTCTCCGATAAAAACCTCCAGCACAGGCGCCTGGACCGGCTTCCAGCTCCTCGCAGGCATCGGATCCG  
GCTTTGTTATCACCTCCACTCTCCCGCCGCACAGGCTGAACTAGCCGAGACTGACGTGCGCCGCTTCNCACTGC  
CACATGGGCTTCTGCGCTCCTTGGGCTCCGTGTGGGGCGTGCCTATCCCCGCTGCTATCTTTAATAATCAGTT  
AGCTTCCCAGTGGCGGCTGTCGGGGTTGGACAAGGGACCCGCGCATCGCNGGTATCCTGGGGTTCTAGTGG  
GGCGNATGANCAAGCATCCGGGGGNCTTTTGNNGNTACGNTCCCGGTGGNTGNNCANCCGCTNNTATCGA  
GAGCTTATCTGGGGNCCATAATANTNTACCCGGAGNCNTGCAAGNGNGGGNACTGGTCCGTCNGTTNNANA  
NCTCTGGGNNTGGGAAAACCTGNTGNTNCCGCACNTAATNNCATTNGNGNNACATNTCCCTCTC

## 3. Primers for cloning of *L. starkeyi* LGK into pET 41

forward primer, addition of *Nde*I: AGTAACATATGCCCATCGCGACTTCCACTG

reversed primer, addition of *Not*I, removal of Stop-Codon:

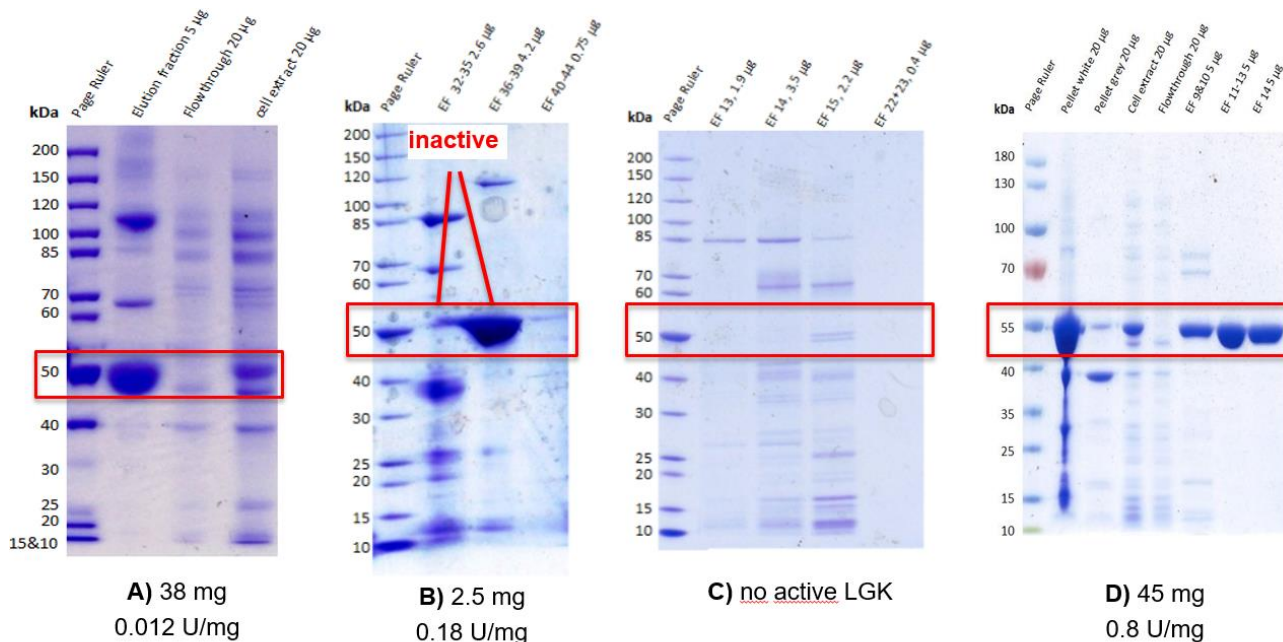
TACTAGCGGCCGAGCCAGTTGTTGGTAATGACC

## 4. Protein sequence of *L. starkeyi* LGK with C-terminal His-tag

MDGIDCALCHFYQKTPDAPMEFELLEYG EVPLAQPIKQRV MRMILED TTSPEL  
SEVNVILGEHFADAVRQFAAERNVDLSTIDA IASHGQTIWLLSMPEEGQVKSA L  
TMAEGAILASRTGITSITDFRISDQAAGRQGAPLIAFFDALLLHPTKLRACQNI

GGIANVCFIPDPVDGRRTDEYYDFDTGPGNVFIDAVVRHFTNGEQEYDKDGAM  
 GKRKVDQELVDDFLKMPYFQLDPPKTTGREVFRDTLAHDLIRRAEAKGLSPDD  
 IVATTTRITAQAIVDHYRRYAPSQEIDEIFMCGGGAYNPNIVEFIQQSYPN TKIM  
 MLDEAGVPAGAKEAITFAWQGMEALVGRSIPVPTRVETRQHYVLGKVPGLNY  
 RSVMKKGM AFGGDAQQLPWVSEMIVKKKGKVITNNWAAAAL EHHHHHHH

## 5. Additional figures



**Fig. S2. 10: SDS-PAGE of LGK purifications by Ni-affinity chromatography. A) – C) LGK with N-terminal His-tag, D) LGK with C-terminal His-tag.** The activity was measured with a not optimized, non-continuous mode assay. Activities in the text (Supporting Information 2, Chapter 3) have been normalized based on the 18.75 times improvement of activity of D) by changed assay conditions. The purification as seen in A) could not be reproduced with LGK with N-terminal His-tag (B, C). In B), most purified protein was inactive. In C), expression was carried out at 18 °C instead of 30 °C which resulted in no active protein to be recovered. In D), C-terminally tagged LGK could be obtained in high purity and yield.

## List of figures

### Chapter 1 – Manuscript in progress

Fig. 1: Phosphorylation of levoglucosan by LGK.....	8
Fig. 2: Proposed mechanism of levoglucosan phosphorylation by LGK.....	15
Fig. 3: Anomeric preferences of levoglucosan kinase.....	16
Fig. 4: Formation of G6P anomers by LGK.....	17
Fig. 5 Stability of LGK.....	18
Fig. 6: $K_M$ of LGK for ATP.....	19
Fig. 7: Proposed reaction mechanisms for levoglucosan formation from glucose-6-phosphate. ....	20

Fig. 8: Dependency of levoglucosan conversion on LG/ATP ratios. ....	21
Fig. 9: Investigation of reversibility of LGK reaction.....	22
Fig. 10: Proposed reaction mechanisms for the formation of levoglucosan from alternative sugar substrates. ....	24
Fig. 11: Phosphorylation of Glc-1F by hexokinase.....	25
Fig. 12: <sup>1</sup> H NMR of Glc-1F-6P with LGK after 24 h.....	25

## Chapter 2 – Supporting information for the manuscript

Fig. S 1: Crystal structure of <i>L. starkeyi</i> LGK dimer bound to ADP, magnesium and levoglucosan (4ZLU). ....	33
Fig. S 2: Purification of LGK (48 kDa) by HisTrap affinity chromatography. ....	34
Fig. S 3: Purification of N-terminally tagged LGK by size exclusion chromatography.....	35
Fig. S 4: LGK activity in dependence of MgCl <sub>2</sub> concentrations.....	36
Fig. S 5: Substrate specificity of LGK.....	36
Fig. S 6: Inhibition of LGK by various sugars.....	37
Fig. S 7: Inhibiting effect of A) ADP and B) G6P on LGK activity.. ....	38
Fig. S 8: TLC showing conversion of Glucal with hexokinase and levoglucosan kinase .....	38
Fig. S 9: MS analysis of LGK reaction with glucal-6-phosphate without ATP recycling system.....	39
Fig. S 10: MS analysis of LGK reaction with glucal-6-phosphate with ATP recycling by pyruvate kinase. ....	40

## Chapter 3 – Additional experiments for characterization of levoglucosan kinase

Fig. S2. 1: Small scale expression of LGK from different colonies after transformation.....	45
Fig. S2. 2: Dot blot of cell lysates from LGK-cultivation different time points after induction .....	46
Fig. S2. 3: SDS-PAGE of soluble and insoluble fraction of cell lysates from LGK expression.....	47
Fig. S2. 4: Conversion of G6P during 72 h of incubation with LGK.....	48
Fig. S2. 5: Incubation of G6P and ADP with LGK.....	48
Fig. S2. 6: Reaction of LGK with G6P (without ADP).....	49
Fig. S2. 7: <sup>1</sup> H NMR of the reaction of G6P and ADP at low concentrations with LGK. ....	50
Fig. S2. 8: Purification of PY with SuperQ anion exchange on HPX-87H column. ....	50
Fig. S2. 9: Proposed structure of PY at acidic pH.....	51
Fig. S2. 10: SDS-PAGE of LGK purifications by Ni-affinity chromatography. ....	54