Establishment of a five-enzyme cell-free cascade for the synthesis of uridine diphosphate *N*-acetylglucosamine

Reza Mahour[1*], Jan Klapproth[2*], Thomas F. T. Rexer[1#], Anna Schildbach[2], Steffen Klamt[3], Markus Pietzsch[2], Erdmann Rapp[3], Udo Reichl[1,4]

[*] RM and JK contributed equally to this work

[#] Corresponding author

[1] Max Planck Institute for Dynamics of Complex Technical Systems, Bioprocess Engineering, Magdeburg, Germany

[2] Martin Luther University Halle-Wittenberg, Institute of Pharmacy, Department of Downstream Processing, Halle (Saale), Germany

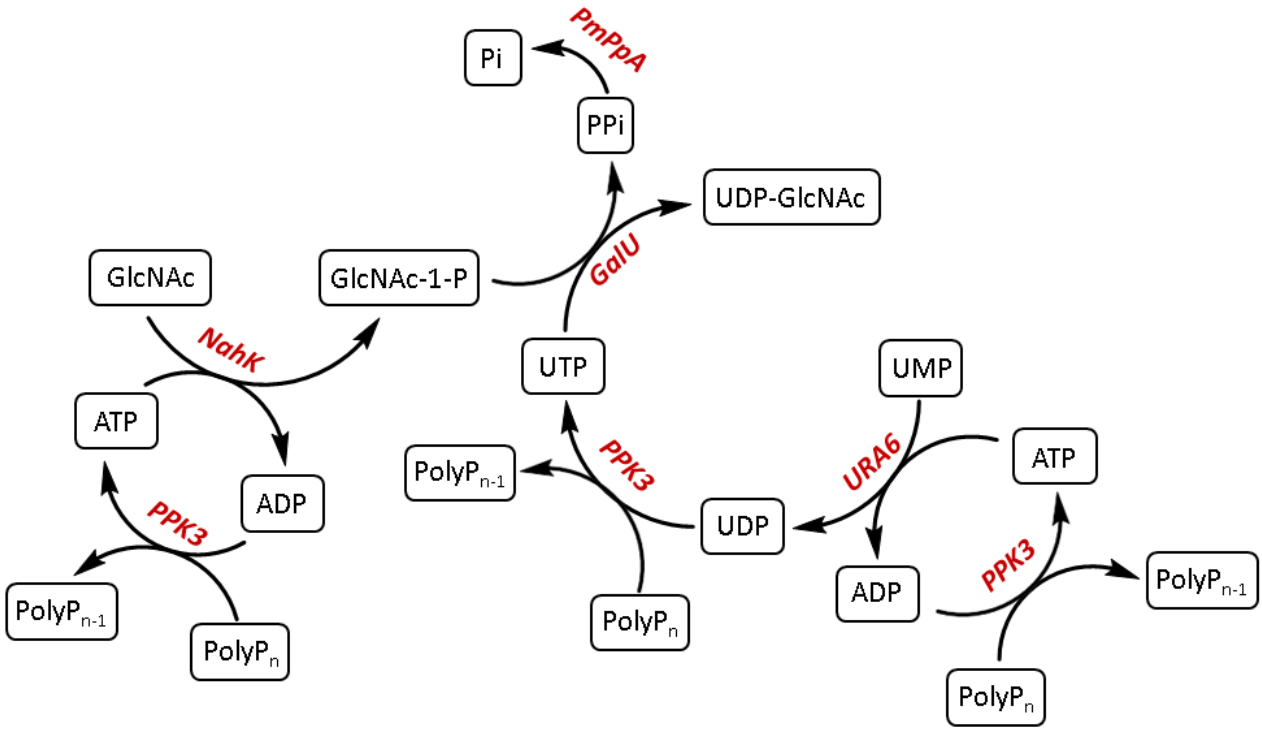
[3] Max Planck Institute for Dynamics of Complex Technical Systems, Analysis and Redesign of Biological Networks, Magdeburg, Germany

[4] Otto-von-Guericke University Magdeburg, Chair of Bioprocess Engineering, Magdeburg, Germany

Reza Mahour	mahour@mpi-magdeburg.mpg.de +49 391 6110 371
Jan Klapproth	jan.klapproth@pharmazie.uni-halle.de + 49 345 55 25 941
Thomas F. T. Rexer, *corresponding author*	rexer@mpi-magdeburg.mpg.de +49 391 6110 375
Anna Schildbach	anna.schildbach@pharmazie.uni-halle.de + 49 345 55 25 941
Steffen Klamt	klamt@mpi-magdeburg.mpg.de +49 391 6110 480
Markus Pietzsch	markus.pietzsch@pharmazie.uni-halle.de + 49 345 55 25 949
Erdmann Rapp	rapp@mpi-magdeburg.mpg.de +49 391 6110 314
Udo Reichl	reichl@mpi-magdeburg.mpg.de +49 391 6110 200

Keywords: Cell-free systems, UDP-GlcNAc, kinetic modeling, metabolic design equation, nucleotide sugar synthesis

Graphical Table of Contents



A cell-free cascade consisting of five enzymes expressed in *E.coli* for the synthesis of UDP-GlcNAc, an essential substrate for *in vitro* glycoengineering of proteins, was developed. UDP-GlcNAc was synthetized from low-cost substrates with a yield approaching 100 %. The design of the cascade is complemented by steady state and kinetic analysis

Abstract

In spite of huge endeavors in cell line engineering to produce glycoproteins with desired and uniform glycoforms, it is still not possible *in vivo*. Alternatively, *in vitro* glycoengineering can be used for the modification of glycans. However, *in vitro* glycoengineering relies on expensive nucleotide sugars, such as uridine 5′-diphospho-*N*-acetylglucosamine (UDP-GlcNAc) which serves as GlcNAc donor for the synthesis of various glycans. In this work, we present a systematic study for the cell-free *de novo* synthesis and regeneration of UDP-GlcNAc from polyphosphate, UMP and GlcNAc by a cascade of five enzymes (*N*–acetylhexosamine kinase (NahK), Glc–1P uridyltransferase (GalU), uridine monophosphate kinase (URA6), polyphosphate kinase (PPK3), and inorganic diphosphatase (PmPpA). All enzymes were expressed in *E. coli* BL21 Gold (DE3) and purified using immobilized metal affinity chromatography (IMAC). Results from one-pot experiments demonstrate the successful production of UDP-GlcNAc with a yield approaching 100 %. The highest volumetric productivity of the cascade was about 0.81 g L^{-1} h^{-1} of UDP-GlcNAc. A simple model based on mass action kinetics was sufficient to capture the dynamic behavior of the multienzyme pathway. Moreover, a design equation based on metabolic control analysis was established to investigate the effect of enzyme concentration on the UDP-GlcNAc flux and to demonstrate that the flux of UDP-GlcNAc can be controlled by means of the enzyme concentrations. The effect of temperature on the UDP-GlcNAc flux followed an Arrhenius equation and the optimal co-factor concentration (Mg^{2+}) for high UDP-GlcNAc synthesis rates depended on the working temperature. In conclusion, the study covers the entire engineering process of a multienzyme cascade, i.e. pathway design, enzyme expression, enzyme purification, reaction kinetics and investigation of the influence of basic parameters (temperature, co-factor concentration, enzyme concentration) on the synthesis rate. Thus, the study lays the foundation for future cascade optimization, preparative scale UDP-GlcNAc synthesis and for *in situ* coupling of the network with UDP-GlcNAc transferases to efficiently regenerate UDP-GlcNAc. Hence, this study

provides a further step towards cost-effective *in vitro* glycoengineering of antibodies and other glycosylated proteins.

1. Introduction

Post-translational modifications of proteins take place ubiquitously in higher cells. One of the most important post-translational modification is glycosylation (Hebert et al., 2005; Varki et al., 2017). The presence and structures of glycans plays a crucial role in cellular life and functioning of glycoproteins, e.g. cell-cell recognition, pharmacokinetics, physical stability and immunogenicity (Bailey et al., 2002). However, there is a lack of thorough understanding of the structure-function relationships of glycans due to macro- and micro-heterogeneities of biological samples (Lalonde and Durocher, 2017; Walsh and Jefferis, 2006). *In vitro* biosynthetic techniques have the potential to overcome challenges associated with cell culture methods, e.g. no by-product formation or substrate consumption for cell maintenance, control of reaction conditions, and potentially easier scale up (Dudley et al., 2015; Hodgman and Jewett, 2012; Karim and Jewett, 2016). These advantages can be exploited for the modification of glycoproteins to produce defined glycan structures (Bülter and Elling, 1999; Hanson et al., 2004; Song et al., 2006).

To build oligosaccharides with defined structures one needs building blocks. Sugar nucleotides such as uridine 5′-diphospho-*N*-acetylglucosamine (UDP-GlcNAc) and guanosine 5′-diphospho-mannose (GDP-mannose) are activated sugars that act as the sugar donor for the synthesis of oligosaccharides by Leloir-type glycosyltransferases (Cai, 2012; Rexer et al., 2018)

From a commercial point of view, sugar nucleotides are expensive sugar donor compounds (Rodríguez-Díaz et al., 2012). Accordingly, their synthesis is crucial for the development of an efficient platform for the cost-effective production and regeneration of sugar nucleotides with the capability of coupling the platform to glycosyltransferases (Chung et al., 2006; Raju et al., 2001; Wang et al., 1993).

The biosynthetic pathways of high energy sugar nucleotides are well known (Bülter and Elling, 1999; Cai, 2012). Sugar nucleotides, except cytidine 5′-monophosphate *N*-acetylneuraminic acid (CMP-Sialic Acid), such as UDP-Galactose, UDP-Glucose, GDP-mannose, and UDP-GlcNAc share similar pathways described in the following (Bülter and Elling, 1999; Cai, 2012; Rexer et al., 2018; Yu and Chen, 2016). The direct reaction of GDP and UDP, respectively, (sugar carriers) and monosaccharides (sugar) is thermodynamically unfavorable ($\Delta G>0$) (Flamholz et al., 2011; Noor et al., 2014). Thus, to increase the energy level of monosaccharides, they are phosphorylated by ATP in a kinase-catalyzed reaction. In the next step, a monosaccharide–1 phosphate guanyl– or uridyltransferase is required to produce a high energy sugar nucleotide by the reaction of monosaccharide 1-phosphate and UTP/GTP (Zhao et al., 2010). UDP-GlcNAc is one of the sugar nucleotides that is *in vivo* produced in the hexosamine pathway (Hanover et al., 2010). It is the GlcNAc donor in the *O*- and *N*-glycosylation pathways and for *O*-GlcNAcylation (Hanover et al., 2010; Varki et al., 2017; Yang and Qian, 2017). UDP-GlcNAc is also the first sugar nucleotide that is required in the assembly of lipid-linked oligosaccharides in the endoplasmic reticulum (ER) of eukaryotes (Aebi, 2013). Moreover, in order to *in vitro* build the core structure of monoclonal antibodies and other recombinant proteins, UDP-GlcNAc is extensively needed (Sha et al., 2016).

Chemical synthesis of UDP-GlcNAc in a five-step process has reached a yield of 15 % (Heidlas et al., 1992). Microbial fermentation has also not resulted in sufficient yields (Heidlas et al., 1992; Rodríguez-Díaz et al., 2012; Tabata et al., 2000; Ying et al., 2009). Conversely, enzymatic synthesis has shown much higher yields. For example, Zhao et al. used three enzymes *N*–acetylhexosamine kinase (NahK), UDP-N-acetylglucosamine diphosphorylase (GlmU) and inorganic diphosphatase (PmPpA) to produce UDP-GlcNAc/UDP-GalNAc and their derivatives at preparative scale with a yield of 10 % to 65 % (Zhao et al., 2010). Chen et al. managed to obtain a yield of 81 % with the same enzymes (Chen et al., 2011). Shao et al. used five immobilized enzymes to produce

UDP-GlcNAc with a maximum yield of 78 % (Shao et al., 2002). In their study, AGX1 (the mammalian type of GlmU) and GlmU were used together to increase the yield of GlcNAc-1-phosphate to UDP-GlcNAc. The regeneration of ATP from ADP was conducted by pyruvate kinase using phosphoenolpyruvate (Shao et al., 2002).

In the study presented here, we used a novel set of recombinant enzymes - NahK, PmPpA as well as Glucose–1-phosphate uridyltransferase (GalU), uridine monophosphate kinase (URA6) and polyphosphate kinase (PPK3) - all produced in *E. coli* and purified using immobilized metal affinity chromatography (IMAC) to synthesize UDP-GlcNAc (Figure 1). To add versatility to the cascade, GalU was utilized to convert GlcNAc-1-phosphate to UDP-GlcNAc. The polyphosphate kinase PPK3 was used to convert diphosphate nucleotides into triphosphates with low-cost polyphosphate as a co-substrate (Nahálka and Pätoprstý, 2009). In addition, URA6 was used to produce UDP from UMP. Diphosphate is known to inhibit GalU and, thus, PmPpA was used to catalyze the diphosphate hydrolysis to increase the UDP-GlcNAc yield. To our knowledge, this study is the first which demonstrates that UDP-GlcNAc can be produced with a yield approaching 100 % using a minimal number of enzymes for the *in situ* regeneration of mono-, di-, and triphosphate nucleotides. The study is complemented by a systematic investigation of the effect of temperature and co-factor concentration. A design equation based on metabolic control analysis is established to investigate the effect of enzyme concentration on the productivity of the cascade. Moreover, the kinetics of the cascade are studied by a simple kinetic model based on mass action kinetics. Motivation for the systematic investigation of the cascade was to provide a basis for one-pot *in vitro* glycoengineering of proteins with *in situ* regeneration of UDP-GlcNAc and/or cost-effective preparative scale synthesis of UDP-GlcNAc (Rexer et al., 2018; Thomann et al., 2015).

2. Material and methods

For a detailed list with chemicals and purification grades used see the supplementary information.

2.1. Analytics: HPAEC – UV/CD

All measurements for quantification of analytes were carried out by high performance anion exchange chromatography (HPAEC) with UV and conductivity detection (CD) (BioLCType DX320 – Idstein, Germany). The eluent gradient was adopted from the work of Ritter and co-workers and optimized for our purpose (Ritter et al., 2006). The optimized eluent gradient (5 –100 mM KOH) was evaluated statistically (Figure 2). For additional information on the HPAEC method and parameters see the supplementary information.

2.2. Pathway design

The engineered cell-free synthetic metabolic pathway consists of five enzymes (Figure 1) all produced by *E. coli* BL21 Gold (DE3). Enzymes were chosen according to literature: NahK (EC 2.7.1.162) from *Bifidobacterium longum* to phosphorylate GlcNAc (Nishimoto and Kitaoka, 2007); GalU (EC 2.7.7.9) from *E. coli* K-12 MG1655 as a GlcNAc-1P uridyltransferase (Thoden & Holden, 2007); URA6 (EC 2.7.4.14) from *Arabidopsis thaliana* for *in situ* regeneration of UDP from UMP (Zhou et al., 1998; Zhou and Thornburg, 1998); PPK3 (EC 2.7.4.1) from *Ruegeria pomeroyi* for *in situ* recovery of energy carriers, ADP and UDP, to their tri–phosphate conjugates (Nahálka and Pätoprstý, 2009); and PmPpA (EC 3.6.1.1) from *Pasteurella multocida* pm70 for the decomposition of GalU inhibiting pyrophosphate (Lau et al., 2010).

2.3. Protein synthesis

2.3.1. Gene synthesis

The host strain *E. coli* BL21 Gold (DE3) was purchased from Stratagene Corp. (Amsterdam, Netherlands) and the plasmid pET-28a(+) was purchased from Invitrogen (Carlsbad, USA). The gene and corresponding protein sequences were obtained from the UniProt database: PmPpA (P57918),

NahK (E4R3E3), GalU (P0AEP3), PPK3 (Q5LSN8), and URA6 (O04905). Gene Designer 2.0 software (Gene Designer, DNA 2.0, Menlo Park, California) was used for optimizing the codon usage of nucleotide sequences for expression in *E. coli*. The resulting sequences were synthesized *de novo* and cloned by GeneArt™ (Thermo Fisher Scientific, Regensburg, Germany). The following restriction sites for subcloning into vector pET-28a(+) were used: NcoI and XhoI for GalU, NahK and PmPpA (enzymes carrying a C-terminal hexahistidin-tag (His-tag)), NdeI and XhoI with PPK3 and URA6 (for an N-terminal His-tag). After transformation of the plasmids into *E. coli*, the DNA was isolated and the accuracy of the constructs was checked by gene sequencing (Eurofins Genomics, Ebersberg, Germany).

2.3.2. Cultivation of transformed *E. coli* BL21 Gold (DE3)

Plasmids were transferred into chemically competent *E. coli* BL21 Gold (DE3) cells (Sambrook et al., 1989). Aliquots of 150 µL were transferred to LB-agar plates to establish colonies for pre-cultures of 50 mL LB/TB-media in 100 mL shaker flasks. Pre-cultures were grown for 18 h at 37°C and 180 rpm. Main cultures of 500 mL LB/TB-media in 2000 mL shaker flasks were then inoculated to an OD_{600} of 0.1 and grown at 37°C and 120 rpm. Induction was carried out by the addition of 1 mM Isopropyl-β-D-thiogalactopyranosid (IPTG) to the culture when the OD_{600} reached values of about 0.5. All cultures were supplemented with 50 µg mL^{-1} kanamycin. Biomass was harvested at about 10 h after induction (6 h for PmPpA harboring cells) by centrifugation at 10.000 x g for 15 min. Overexpression was monitored by SDS-PAGE following a standard protocol (Laemmli, 1970). Bio wet mass (BWM) was stored at -20°C until purification. Media compositions are described in supplementary information (Table SI 1).

2.3.3. Enzyme purification

For enzyme purification, immobilized metal affinity chromatography (IMAC) was used (ÄKTA explorer 100, Amersham Bioscience plc, Uppsala, Sweden). 4 g (2 g PmPpA) of *E. coli* BWM were

suspended in 40 mL of equilibration buffer, containing 50 mM Tris/HCl, 300 mM NaCl and 5 mM imidazole at pH 7.5. Cells were disintegrated with a high pressure homogenizer (Emulsiflex C5, Avestin Inc., Ottawa, Canada) at 1000 bar. Cell debris was separated by centrifugation (45 min; 17.700 x g) and the supernatant was applied to an IMAC set-up equipped with an pre-equilibrated column containing Ni^{2+}-STREAMLINE chelating™ material (GE Healthcare, Freiburg, Germany) as a stationary phase (column volume: 13 mL). After washing the column with equilibration buffer, elution was carried out using a buffer containing 50 mM Tris/HCl, 300 mM NaCl and 500 mM imidazole at pH 7.5. Fractions of 1 mL were collected. Protein concentration was determined according to Bradford using bovine serum albumin as a standard (Bradford, 1976). After addition of glycerol to a final concentration of 50 %, enzyme stock solutions were stored at -20°C.

2.4. Multienzymatic experiments for kinetic analysis

Reactions took place in reaction volumes of 1 mL in buffered (50 mM Tris/HCl, pH 7.5) aqueous solutions at 30°C and a co-factor concentration of 45 mM $MgCl_2$ - unless stated otherwise. Initial concentrations of GlcNAc and UMP were 1 mM and 0.8 mM, and initial concentrations of ATP and $PolyP_{14}$ were 2.5 and 2 mM, respectively. For more information on the reaction set-up see supplementary information.

2.5. Mathematical modeling

Metabolic Control Analysis (MCA): steady state approach

For the rational design of an enzymatic pathway metabolic control analysis (MCA) can be applied (Stephanopoulos et al., 1998). MCA is based on the concept of sensitivity analysis allowing the quantification of a perturbation (e.g., changes in enzyme concentrations) on metabolic system properties, e.g. flux (see (Heinrich and Schuster, 1996)) for a detailed theoretical description). Flux control coefficients (FCC) quantify the normalized changes in flux upon changes in enzymes levels in a

pathway. To mathematically describe the concept of FCCs, we have Visser and Heijnen (Visser and Heijnen, 2002):

$$C_i^{J_0} = \frac{e_i^0}{J^0}\frac{dJ}{de_i} \qquad \textbf{Equation 1}$$

In Equation 1, $C_i^{J_0}$ indicates the FCC of an enzyme e_i on the flux J. The parameters e_i^0 and J^0 show the concentration of enzyme i and flux J of the product at reference state, respectively. The response coefficient quantifies flux changes towards changes in independent metabolites (substrates) and is defined as (Visser and Heijnen, 2002):

$$R_i^{J_0} = \frac{c_i^0}{J^0}\frac{dJ}{dc_i} \qquad \textbf{Equation 2}$$

In Equation 2, $R_i^{J_0}$ is the response coefficient of substrate I; J and J^0 is the flux at non-reference and reference state, respectively. c_i and c_i^0 denote the concentration of substrate at non-reference and reference state, respectively.

Derivation of the design equation is comprehensively explained in literature and demonstrated on an *in vitro* pathway (Visser and Heijnen, 2003; Wu et al., 2004). Since controlling the flux by varying independent parameters (enzymes and substrates) was of interest, the design equation reads (Wu et al., 2004)

$$C^{J_0} \cdot \left(\frac{J}{J^0} \cdot \frac{e^0}{e}\right) = i + R_c^{J_0} \ln\left(\frac{c}{c^0}\right) \qquad \textbf{Equation 3}$$

where i is unity.

Design equations allow a fast identification of key control points (substrate and enzyme concentrations) in the system where small changes may lead to large increases in certain (product) fluxes.

In this work, the concentration of substrates remained constant. Therefore, it was not required to calculate response coefficients, and the design equation reduces to:

$$C^{J_0} \cdot \left(\frac{J}{J^0} \cdot \frac{e^0}{e}\right) = i \qquad \textbf{Equation 4}$$

For the whole pathway it reads:

$$\frac{J^0}{J} = C_{NahK}^{J_0}\left(\frac{e_{NahK}^0}{e_{NahK}}\right) + C_{GalU}^{J_0}\left(\frac{e_{GalU}^0}{e_{GalU}}\right) + C_{PPK3}^{J_0}\left(\frac{e_{PPK3}^0}{e_{PPK3}}\right) + C_{URA6}^{J_0}\left(\frac{e_{URA6}^0}{e_{URA6}}\right) \qquad \textbf{Equation 5}$$

Based on Equation 5, sets of different experiments were conducted to calculate the FCCs based on multiple linear regressions.

Mass action kinetics: Dynamic approach

In order to elucidate the dynamic behavior of the cascade under investigation, we used simple mass action kinetics to describe the concentration time course of compounds. Parameter estimation was carried out using the systems biology toolbox developed for MATLAB (Schmidt and Jirstrand, 2006).

3. Results and discussion

3.1. Enzyme expression and purification

SDS-PAGE gel of purified enzymes (Figure 3) confirmed the successful gene expression and subsequent protein purification. LB media was substituted for TB media for the expression of URA6 and PPK3, under the same condition, TB media provided a 1.5 fold higher concentration of BWM. It was found that NahK precipitated in the IMAC purified fractions. Dilution of the purified NahK fractions with elution buffer prevented enzyme precipitation. Enzyme yields and concentration are detailed in Table I.

3.2. Multienzyme reactions for the synthesis of UDP-GlcNAc

To verify the reproducible synthesis of UDP-GlcNAc by the multienzyme cascade, a one-pot reaction containing all five enzymes (0.1 mg mL^{-1} Nahk, URA6, PPK3, GalU; 0.03 mg mL^{-1} PmPpa) at conditions specified in M&M was carried out in biological triplicates and compared against a set of negative control experiments (see supplementary information Table SI 2), where one enzyme each was

missing with the exception of PmPpa. The enzyme concentrations are constrained by the enzyme stock solutions (see Table I). Initial enzyme and substrate concentrations were selected empirically and based on our findings can be optimized in future. In the positive controls, UDP-GlcNAc was detected after about 10 min (Figure 4). The concentration of UDP-GlcNAc after 120 min was about 750 µM, and concentration time series for all metabolites showed good reproducibility (Figure 4).

In the following, the effects of temperature and co-factor concentration are investigated. Moreover, using mathematically approaches, the enzyme kinetics are studied and a steady state design equation is established to investigate the effect of enzyme concentration on the flux of UDP-GlcNAc. Both models are useful for coupling the cascade to glycosyltransferases or upscaling UDP-GlcNAc synthesis (Rexer et al., 2018). Coupling the cascade directly to UDP-GlcNAc-transferases for *in vitro* glycoengineering of proteins can circumvent cost-intensive downstream processing of UDP-GlcNAc (Lemmerer et al., 2016).

As mentioned in M&M the initial concentration of substrates, GlcNAc (1 mM), UMP (0.8 mM), ATP (2.5 mM) and $PolyP_{14}$ (2 mM), were kept constant for all reactions as the focus was on studying the impact of temperature and co-factor concentration as well as the control the various enzymes exert on the flux of UDP-GlcNAc of the cascade. However, for preparative scale synthesis higher concentrations of substrates could further increase the productivity.

3.2.2. Influence of temperature on the cascade

Generally speaking, the temperature stability of enzymes heavily depends on their source. For instance, GalU from the thermophile bacterium *Sulfolobus tokodaii* was stable at 80°C (Zhang et al., 2005); but GalU from *Homo sapiens* lost its activity after 15 min of incubation at 52°C (Chacko et al., 1972). GalU from *E.coli* is reported to be stable at 37°C (Weissborn et al., 1994). NahK from *Bifidobacterium longum* lost half of its activity after 30 min of incubation at 50°C (Nishimoto and Kitaoka, 2007). URA6 from *Arabidopsis thaliana* (Zhou et al., 1998; Zhou and Thornburg, 1998) also

lost half of its activity when it was incubated for 10 min at 58°C (Zhou et al., 1998; Zhou and Thornburg, 1998). Enzyme assays using PPK3 from *Ruegeria pomeroyi* were conducted at 30°C (Nahálka and Pätoprstý, 2009). No information is available about the temperature stability of PmPpA from *Pasteurella multocida*.

To gain insight into the temperature dependency of the complete pathway, experiments were carried out at four different temperatures (20, 30, 40, and 50°C). The effect of temperature on the pH value of Tris/HCl at pH 7.5 was negligible (data not shown).

The flux of UDP-GlcNAc was considered as the kinetic representative for evaluating the performance of the cascade at different temperatures. The flux increased from 20 to 40°C, and decreased at 50°C (Figure 5A). Similar behavior was observed in other biological systems and can be explained by enzyme deactivation (Huang et al., 2011; Swarup et al., 2014). The trend of increasing flux with temperature can be described by an Arrhenius model (Figure 5B):

$$J = A\, exp\left(\frac{-E_a}{RT}\right) \qquad \textbf{Equation 6}$$

In Equation 6, J is the flux of UDP-GlcNAc (μM h^{-1}); the pre-exponential constant A is 9.178 x 10^{14} (μM h^{-1}); R, is the universal gas constant (8.314 J K^{-1} mol^{-1}); T is the temperature (K), and the activation energy E_a is 71.186 (kJ mol^{-1}). A similar increase in volumetric productivity through increasing the temperature has been observed in other multienzymatic reactions as well (Eixelsberger and Nidetzky, 2014). Moreover, the multienzyme cascade experiment at 40°C showed that UDP-GlcNAc yields approaching 100 % (after 70 min at 40°C) are feasible (Figure 5A).

3.2.3. Co-factor concentration

The role of co-factors on enzyme activity has been described extensively in literature (Beard and Qian, 2008). Magnesium ions are co-factors for all the enzymes in the cascade. It has been shown that NahK from *Bifidobacterium longum* needs Mg^{+2} for its catalytic activity (Li et al., 2011; Nishimoto and Kitaoka, 2007). The presence of magnesium ion is also necessary for the catalytic activity of GalU

from *E. coli* (Turnquist and Hansen, 1973).The polyphosphate kinase homologues derived from *Ruegeria pomeroy* have different behaviors regarding magnesium ions (Achbergerová and Nahálka, 2014). Moreover, URA6 and PmPpA both require Mg^{+2} for their catalytic activity (Schomburg et al., 2002).

The role of divalent magnesium ions as a co-factor can be classified into two general categories: magnesium-substrate complex formation and subsequent binding and turnover and magnesium-enzyme complex formation (Cowan, 2002). In the second case Mg^{2+} either alters the enzyme structure and/or is directly involved in the catalysis (Cowan, 2002). In a study on uridine phosphates it was found that binding of Mg^{2+} to form phosphate metal complexes is endothermic except for pyrophosphate (Zea et al., 2008).

In order to determine the effect of various Mg^{+2} ion concentrations on the productivity of the cascade, two different sets of experiments were carried out. At first, the role of magnesium ions on the productivity was evaluated by conducting multienzyme cascade reactions at 30°C with Mg^{2+} concentrations of 20, 45, 100 and 150 mM (Figure 6A). In the second set of experiments, the productivity was investigated at 40°C with Mg^{+2} concentrations of 10, 15, 20 and 45 mM (Figure 6B). The results indicate that the optimal co-factor concentration is temperature-dependent. It is equal or below 20 mM at 30°C, and equal or above 45 mM at 40°C. The temperature-dependent cofactor optimum is possibly due to changed concentration levels of phosphate metal complexes. Moreover, Mg^{2+} may also play a role in structural stabilization of enzymes against thermal inactivation (Huang and Cowan, 1994; Liu et al., 2007). For definitive conclusions, each enzyme and reaction needs to investigated individually by isothermal titration calorimetry and enzyme crystallography.

3.3. Mathematical modeling

To systematically analyze the kinetics and the impact of enzyme concentration on the UDP-GlcNAc synthesis rate seven independent experiments were conducted in which one enzyme

concentration was changed at a time (Table II). The activity of PmpPA with respect to its mass exceeds that of all of the other four enzymes by far. Thus, the impact of PmPpA on the UDP-GlcNAc synthesis was not further investigated, i.e. the PmPpA concentration (0.03 mg mL^{-1}) was kept constant throughout the experiments. At this concentration, the conversion of diphosphate to phosphate was instant. UDP-GlcNAc fluxes were calculated by linear regression from the linear concentration-time curves between 25 and 90 min and were between 300 and 426 µM h^{-1} (Table II). The accumulation of GlcNAc-1-phosphate up to an apparent threshold of about 500 µM did not have any specific effect on the cascade.

Steady state approach: Design equation

FCCs were calculated by conducting multiple linear regressions using Equation 5. Potentially, each of the seven experiments could be taken as reference state and consequently values for FCCs could vary based on the selection of the corresponding reference state. Results of the steady state approach are shown in Table III. The high FCC for GalU highlights the control that the GalU concentration exerts over the UDP-GlcNAc flux, i.e. the UDP-GlcNAc flux is most sensitive to the GalU concentration. Consequently, the highest UDP-GlcNAc flux (442 µM h^{-1}) in the seven experiments is observed with the highest GalU concentration (0.13 mg mL^{-1}) (Table II). The second highest FCC was calculated for NahK, followed by PPK3 and URA6.

Dynamic kinetic modelling

To elucidate the underlying reactions in the cascade, i.e. the dynamics of the network, a kinetic model based on mass action kinetics was established (Table IV). The following simplifications were conducted to reduce the complexity of the kinetic model: the concentration of $PolyP_{14}$ was assumed to be constant, and the reactions of GalU and PmPpA were lumped as one reaction (the activity of PmPpA is much higher than the activity of GalU – unpublished data). The mechanism of $PolyP_{14}$ consumption – processive or non-processive – is not known. However, it can be assumed that an excessive amount

of phosphate for the phosphorylation reactions is present. The kinetic parameters were estimated taking all seven experiments (one fit) into account. Results of the parameter estimation are shown in Table IV.

In general, a good fit for all investigated reactions was achieved (Figure 8). The developed model captures the dynamics of the UMP/UDP, UDP/UTP, and UTP/UDP-GlcNAc conversion well. However, for the first minutes of the reaction the model fails to reflect the dynamics of the ADP/ATP conversion (Figure 4). This implies that the regeneration from polyphosphate by PPK3 using ATP, and probably also UTP, follows a more complex reaction mechanism. Moreover, according to the model there is no indication of any inhibition and, thus, under the conditions tested the network will always result in a yield of 100% of UDP-GlcNAc.

Cross validation and discussion

It has been suggested that the best way to check the predictability of a model is to compare additional independent experiments performed with simulated results (Broadhurst and Kell, 2006). Therefore, two experiments were carried out with various enzyme concentrations within the previously investigated range of concentrations (denoted "CV_1" and "CV_2"; Table II). The design equation (Equation 5) was used to predict the UDP-GlcNAc flux for both reactions (Figure 8A). In addition, the concentration over time of UDP-GlcNAc was simulated by the kinetic model (Figure 8B).

Regardless of the reference state, the design equation could predict the flux with a very good accuracy (deviation between experimental and simulated flux < 1 %) for CV_2 and fairly well for CV_1 (deviation of 11 %) (Figure 8A and Figure SI 1).

The kinetic model can simulate the concentration over time for CV_1 and CV_2 well. Differences in UDP-GlcNAc concentrations between experimental and simulated data after 90 min are less than 10 % for CV_1 and CV_2 (Figure 8B).

In conclusion, the models perform equally well in predicting the synthesis of UDP-GlcNAc within the range of conditions investigated. The advantage of design equations is that it predicts the

flux by means of dimensionless parameters and, thus, has potential to be used in scale up studies, e.g. cost analysis. The kinetic model can be used to study reaction inhibition when the cascade is coupled to UDP-GlcNAc transferases for assembly of lipid-linked oligosaccharides or *in vitro* glycoengineering of proteins (Rexer et al., 2018; Thomann et al., 2015).

4. Conclusion

A cell-free enzyme cascade for the *de novo* synthesis of UDP-GlcNAc and the regeneration of nucleotide mono- and diphosphates was established. It consists of five enzymes expressed in *E. coli* BL21 Gold (DE3), which were purified using IMAC. Under the chosen reaction conditions, a UDP-GlcNAc yield of 100 % was obtained at 40°C and 45 mM Mg^{2+} from the low cost substrates, GlcNAc, polyphospate, and UMP. ATP was continuously regenerated. The maximum volumetric productivity observed was 0.81 g L^{-1} h^{-1} at 40°C and 45 mM Mg^{2+}.

A set of multienzyme reactions with various enzyme concentrations was carried out to investigate the enzyme kinetics as well as the impact of enzyme concentration on the productivity. The metabolites were quantified by HPAEC using an optimized gradient. By means of a design equation based on metabolic control analysis it was found that GalU exerts the highest control over the UDP-GlcNAc synthesis rate. In addition a simple model based on mass action kinetics could describe the concentrations over time for all measured metabolites. Both models, the design equation and the kinetic model, can predict the UPD-GlcNAc flux and concentration over time, respectively, within the range of conditions investigated.

Overall, the results obtained in this study lay the groundwork for cost-effective synthesis of UDP-GlcNAc at preparative scale. Moreover, the results and both models are useful, for the coupling of the cell-free network to UDP-GlcNAc transferases for *in vitro* glycoengineering of proteins. By coupling multienzyme reactions for the synthesis and regeneration of nucleotide sugars to

glycosyltransferases in one-pot reactions cost-intensive downstream processing of nucleotide sugars can be avoided.

Acknowledgement

The authors would like to thank Silvana Fischer and Vivien Winkler for technical support. Thomas Bissinger is thanked for his excellent support in analytics. The project has been funded by the Federal Ministry of Education and Research (grant no. 031A156A and B). The authors declare no conflict of interest.

References

Achbergerová, L., Nahálka, J., (2014) Degradation of polyphosphates by polyphosphate kinases from Ruegeria pomeroyi. Biotechnology Letters 36, 2029-2035.

Aebi, M., (2013) N-linked protein glycosylation in the ER. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1833, 2430-2437.

Bailey, J.E., Prati, E., Jean-Mairet, J., Sburlati, A., Umaña, P., (2002) Engineering Glycosylation in Animal Cells. In: Merten, O.-W., Perrin, P., Griffiths, B. (Eds.), New Developments and New Applications in Animal Cell Technology: Proceedings of the 15th ESACT Meeting. Springer Netherlands, Dordrecht, pp. 5-23.

Beard, D.A., Qian, H., (2008) Chemical biophysics: quantitative analysis of cellular systems. Cambridge University Press.

Broadhurst, D.I., Kell, D.B., (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics 2, 171-196.

Bülter, T., Elling, L., (1999) Enzymatic synthesis of nucleotide sugars. Glycoconjugate journal 16, 147-159.

Cai, L., (2012) Recent Progress in Enzymatic Synthesis of Sugar Nucleotides. Journal of Carbohydrate Chemistry 31, 535-552.

Chacko, C.M., McCrone, L., Nadler, H.L., (1972) Uridine diphosphoglucose pyrophosphorylase and uridine diphosphogalactose pyrophosphorylase in human skin fibroblasts derived from normal and galactosemic individuals. Biochimica et Biophysica Acta (BBA)-Enzymology 268, 113-120.

Chen, Y., Thon, V., Li, Y., Yu, H., Ding, L., Lau, K., Qu, J., Hie, L., Chen, X., (2011) One-pot three-enzyme synthesis of UDP-GlcNAc derivatives. Chemical Communications 47, 10815-10817.

Chung, S.-W., Joo, H.-S., Jang, K.-S., Lee, H.-J., Lee, S.-G., Kim, B.-G., (2006) Galactosylation and sialylation of terminal glycan residues of human immunoglobulin G using bacterial glycosyltransferases with in situ regeneration of sugar-nucleotides. Enzyme and microbial technology 39, 60-66.

Cowan, J.A., (2002) Structural and catalytic chemistry of magnesium-dependent enzymes. Biometals 15, 225-235.

Dudley, Q.M., Karim, A.S., Jewett, M.C., (2015) Cell-free metabolic engineering: Biomanufacturing beyond the cell. Biotechnology Journal 10, 69-82.

Eixelsberger, T., Nidetzky, B., (2014) Enzymatic Redox Cascade for One-Pot Synthesis of Uridine 5'-Diphosphate Xylose from Uridine 5'-Diphosphate Glucose. Advanced synthesis & catalysis 356, 3575-3584.

Flamholz, A., Noor, E., Bar-Even, A., Milo, R., (2011) eQuilibrator—the biochemical thermodynamics calculator. Nucleic Acids Research 40, D770-D775.

Gottwald, W., (2000) Statistik für Anwender. Wiley-VCH.

Hanover, J.A., Krause, M.W., Love, D.C., (2010) The hexosamine signaling pathway: O-GlcNAc cycling in feast or famine. Biochimica et Biophysica Acta (BBA)-General Subjects 1800, 80-95.

Hanson, S., Best, M., Bryan, M.C., Wong, C.-H., (2004) Chemoenzymatic synthesis of oligosaccharides and glycoproteins. Trends in Biochemical Sciences 29, 656-663.

Hebert, D.N., Garman, S.C., Molinari, M., (2005) The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. Trends in Cell Biology 15, 364-370.

Heidlas, J.E., Lees, W.J., Pale, P., Whitesides, G.M., (1992) Gram-Scale Synthesis of Uridine 5'-Diphospho-N-acetylglucosamine: Comparison of Enzymic and Chemical Routes. Journal of Organic Chemistry 57, 146-151.

Heinrich, R., Schuster, S., (1996) The regulation of cellular systems. Springer Science & Business Media.

Hodgman, C.E., Jewett, M.C., (2012) Cell-free synthetic biology: Thinking outside the cell. Metabolic Engineering 14, 261-269.

Huang, H.W., Cowan, J.A., (1994) Metallobiochemistry of the magnesium ion: Characterization of the essential metal-binding site in Escherichia coli ribonuclease H. European Journal of Biochemistry 219, 253-260.

Huang, L., Hwang, A., Phillips, J., (2011) Effect of temperature on microbial growth rate–mathematical analysis: the Arrhenius and Eyring–Polanyi connections. Journal of food science 76, E553-E560.

Karim, A.S., Jewett, M.C., (2016) A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. Metabolic Engineering 36, 116-126.

Laemmli, U.K., (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. nature 227, 680-685.

Lalonde, M.-E., Durocher, Y., (2017) Therapeutic glycoprotein production in mammalian cells. Journal of Biotechnology 251, 128-140.

Lau, K., Thon, V., Yu, H., Ding, L., Chen, Y., Muthana, M.M., Wong, D., Huang, R., Chen, X., (2010) Highly efficient chemoenzymatic synthesis of [small beta]1-4-linked galactosides with promiscuous bacterial [small beta]1-4-galactosyltransferases. Chemical Communications 46, 6066-6068.

Lemmerer, M., Schmölzer, K., Gutmann, A., Nidetzky, B., (2016) Downstream Processing of Nucleoside-Diphospho-Sugars from Sucrose Synthase Reaction Mixtures at Decreased Solvent Consumption. Advanced Synthesis & Catalysis 358, 3113-3122.

Li, Y., Yu, H., Chen, Y., Lau, K., Cai, L., Cao, H., Tiwari, V.K., Qu, J., Thon, V., Wang, P.G., (2011) Substrate promiscuity of N-acetylhexosamine 1-kinases. Molecules 16, 6396-6407.

Liu, W.-F., Zhang, A., Cheng, Y., Zhou, H.-M., Yan, Y.-B., (2007) Effect of magnesium ions on the thermal stability of human poly(A)-specific ribonuclease. FEBS Letters 581, 1047-1052.

Mandel, J., (1964) The statistical analysis of experimental data. Wiley, New York.

Nahálka, J., Pätoprstý, V., (2009) Enzymatic synthesis of sialylation substrates powered by a novel polyphosphate kinase (PPK3). Organic & biomolecular chemistry 7, 1778-1780.

Nishimoto, M., Kitaoka, M., (2007) Identification of N-acetylhexosamine 1-kinase in the complete lacto-N-biose I/galacto-N-biose metabolic pathway in Bifidobacterium longum. Applied and environmental microbiology 73, 6444-6449.

Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., Milo, R., (2014) Pathway thermodynamics highlights kinetic obstacles in central metabolism. PLoS Comput Biol 10, e1003483.

Raju, T.S., Briggs, J.B., Chamow, S.M., Winkler, M.E., Jones, A.J., (2001) Glycoengineering of therapeutic glycoproteins: in vitro galactosylation and sialylation of glycoproteins with terminal N-acetylglucosamine and galactose residues. Biochemistry 40, 8868-8876.

Rexer, T.F.T., Schildbach, A., Klapproth, J., Schierhorn, A., Mahour, R., Pietzsch, M., Rapp, E., Reichl, U., (2018) One pot synthesis of GDP-mannose by a multi-enzyme cascade for enzymatic assembly of lipid-linked oligosaccharides. Biotechnol Bioeng 115, 192-205.

Ritter, J.B., Genzel, Y., Reichl, U., (2006) High-performance anion-exchange chromatography using on-line electrolytic eluent generation for the determination of more than 25 intermediates from energy metabolism of mammalian cells in culture. Journal of Chromatography B 843, 216-226.

Rodríguez-Díaz, J., Rubio-del-Campo, A., Yebra, M.J., (2012) Metabolic engineering of Lactobacillus casei for production of UDP-N-acetylglucosamine. Biotechnology and Bioengineering 109, 1704-1712.

Rodríguez-Díaz, J., Rubio-del-Campo, A., Yebra, M.J., (2012) Metabolic engineering of Lactobacillus casei for production of UDP-N-acetylglucosamine. Biotechnology and bioengineering 109, 1704-1712.

Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.

Schmidt, H., Jirstrand, M., (2006) Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. Bioinformatics 22, 514-515.

Schomburg, I., Chang, A., Schomburg, D., (2002) BRENDA, enzyme data and metabolic information. Nucleic acids research 30, 47-49.

Sha, S., Agarabi, C., Brorson, K., Lee, D.-Y., Yoon, S., (2016) N-Glycosylation Design and Control of Therapeutic Monoclonal Antibodies. Trends in biotechnology.

Shao, J., Zhang, J., Nahálka, J., Wang, P.G., (2002) Biocatalytic synthesis of uridine 5'-diphosphate N-acetylglucosamine by multiple enzymes co-immobilized on agarose beads. Chemical Communications, 2586-2587.

Song, J., Zhang, H., Li, L., Bi, Z., Chen, M., Wang, W., Yao, Q., Guo, H., Tian, M., Li, H., (2006) Enzymatic biosynthesis of oligosaccharides and glycoconjugates. Current Organic Synthesis 3, 159-168.

Stephanopoulos, G.N., Aristidou, A.A., Nielsen, J., (1998) CHAPTER 11 - Metabolic Control Analysis. Metabolic Engineering. Academic Press, San Diego, pp. 461-533.

Swarup, A., Lu, J., DeWoody, K.C., Antoniewicz, M.R., (2014) Metabolic network reconstruction, growth characterization and 13 C-metabolic flux analysis of the extremophile Thermus thermophilus HB8. Metabolic engineering 24, 173-180.

Tabata, K., Koizumi, S., Endo, T., Ozaki, A., (2000) Production of UDP-N-acetylglucosamine by coupling metabolically engineered bacteria. Biotechnology Letters 22, 479-483.

Thomann, M., Schlothauer, T., Dashivets, T., Malik, S., Avenal, C., Bulau, P., Rüger, P., Reusch, D., (2015) In vitro glycoengineering of IgG1 and its effect on Fc receptor binding and ADCC activity. PloS one 10, e0134949.

Turnquist, R.L., Hansen, R.G., (1973) 2 Uridine Diphosphoryl Glucose Pyrophosphorylase. In: Boyer, P.D. (Ed.), The Enzymes. Academic Press, pp. 51-71.

Varki, A., Cummings, R.D., Esko, J.D., Hart, G.W., (2017) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press.

Visser, D., Heijnen, J.J., (2002) The Mathematics of Metabolic Control Analysis Revisited. Metabolic Engineering 4, 114-123.

Visser, D., Heijnen, J.J., (2003) Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. Metabolic Engineering 5, 164-176.

Walsh, G., Jefferis, R., (2006) Post-translational modifications in the context of therapeutic proteins. Nat Biotech 24, 1241-1252.

Wang, P., Shen, G.J., Wang, Y.F., Ichikawa, Y., Wong, C.H., (1993) Enzymes in oligosaccharide synthesis: active-domain overproduction, specificity study, and synthetic use of an .alpha.-1,2-mannosyltransferase with regeneration of GDP-Man. The Journal of Organic Chemistry 58, 3985-3990.

Weissborn, A.C., Liu, Q., Rumley, M.K., Kennedy, E.P., (1994) UTP: alpha-D-glucose-1-phosphate uridylyltransferase of Escherichia coli: isolation and DNA sequence of the galU gene and purification of the enzyme. Journal of bacteriology 176, 2611-2618.

Wu, L., Wang, W., van Winden, W.A., van Gulik, W.M., Heijnen, J.J., (2004) A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. European Journal of Biochemistry 271, 3348-3359.

Yang, X., Qian, K., (2017) Protein O-GlcNAcylation: emerging mechanisms and functions. Nature Reviews Molecular Cell Biology 18, 452.

Ying, H., Chen, X., Cao, H., Xiong, J., Hong, Y., Bai, J., Li, Z., (2009) Enhanced uridine diphosphate N-acetylglucosamine production using whole-cell catalysis. Applied microbiology and biotechnology 84, 677-683.

Yu, H., Chen, X., (2016) One-pot multienzyme (OPME) systems for chemoenzymatic synthesis of carbohydrates. Organic & Biomolecular Chemistry 14, 2809-2818.

Zea, C.J., Camci-Unal, G., Pohl, N.L., (2008) Thermodynamics of binding of divalent magnesium and manganese to uridine phosphates: implications for diabetes-related hypomagnesaemia and carbohydrate biocatalysis. Chemistry Central Journal 2, 15-15.

Zhang, Z., Tsujimura, M., Akutsu, J.-i., Sasaki, M., Tajima, H., Kawarabayasi, Y., (2005) Identification of an extremely thermostable enzyme with dual sugar-1-phosphate nucleotidylyltransferase activities from an acidothermophilic archaeon, Sulfolobus tokodaii strain 7. Journal of Biological Chemistry 280, 9698-9705.

Zhao, G., Guan, W., Cai, L., Wang, P.G., (2010) Enzymatic route to preparative-scale synthesis of UDP–GlcNAc/GalNAc, their analogues and GDP–fucose. Nature protocols 5, 636.

Zhou, L., Lacroute, F., Thornburg, R., (1998) Cloning, expression in Escherichia coli, and characterization of Arabidopsis thaliana UMP/CMP kinase. Plant physiology 117, 245-254.

Zhou, L., Thornburg, R., (1998) Site-Specific Mutations of Conserved Residues in the Phosphate-Binding Loop of theArabidopsisUMP/CMP Kinase Alter ATP and UMP Binding. Archives of Biochemistry and Biophysics 358, 297-302.

List of Figures

List of Tables

List of Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BWM	Bio wet mass
CMP-sialic	Cytidine-5'-monophospho-N-acetylneuraminic acid
FCC	Flux control coefficient
GalU	Bifunctional Glc–1P uridyltransferase
GDP	Guanosine diphosphate
GlcNAc	N-Acetylglucosamine
GTP	Guanosine triphosphate
HCl	Hydrochloride
His-tag	Hexahistidine-tag
HPAEC-UV/CD	High-performance anion-exchange chromatography with ultraviolet and conductivity detection
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MCA	Metabolic Control Analysis
NahK	N–acetylhexosamine kinase
OD_{600}	Absorbance at wavelength 600 nm
PmPpA	Inorganic pyrophosphatase
PPK3	Bifunctional polyphosphate kinase
$PolyP_{14}$	Polyphosphate, average chain length 14 phosphate units
rpm	Rounds per min
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
UMP	Uridinphosphate
UDP	Uridindiphosphate
UTP	Uridintriphosphate
UDP-GlcNAc	Uridindiphosphate-*N*-acetylglucosamine
URA6	Uridine monophosphate kinase

Figures

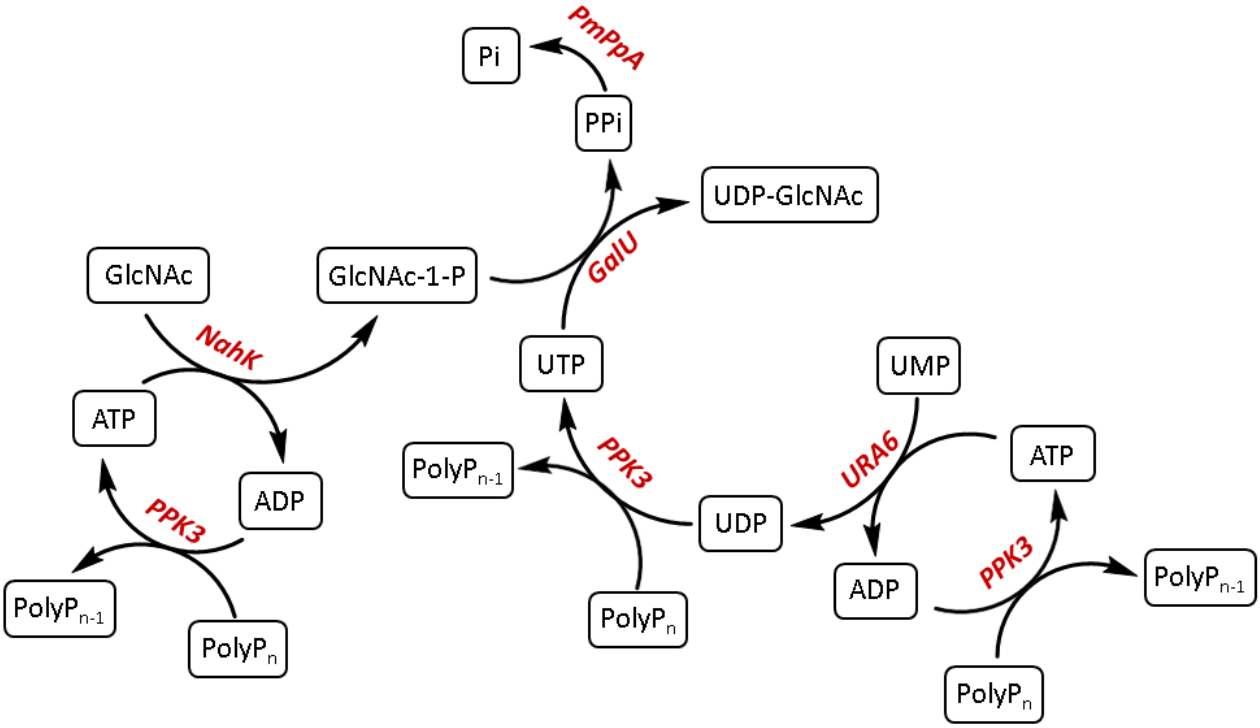


Figure 1: Engineered cell-free metabolic pathway for the continuous synthesis of UDP-GlcNAc from low-cost substrates using a minimal number of enzymes expressed in *E. coli*. *N*–acetylhexosamine kinase (NahK, EC 2.7.1.162; bifunctional Glc-1 phosphate uridyltransferase (GalU, EC 2.7.7.9), uridine monophosphate kinase (URA6, EC 2.7.4.14), bifunctional polyphosphate kinase (PPK3, EC 2.7.4.1), and inorganic diphosphatase (PmPpA, EC 3.6.1.1).

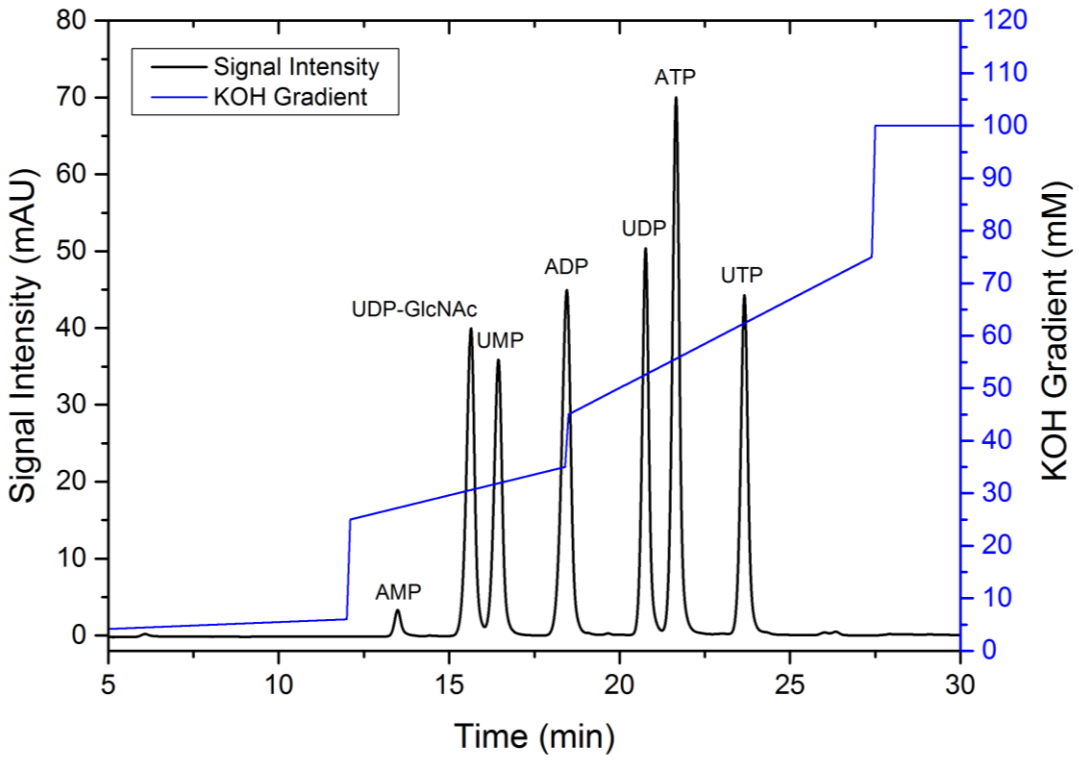


Figure 2: UV-chromatogram of a 170 µM standard sample for the optimized and validated eluent (KOH) gradient. The UV wavelength for detection is 260 nm. The first 5 min (column equilibration) of the chromatogram are not shown. GlcNAc-1-phosphat was measured and quantified by conductivity detection (CD) (data not shown).

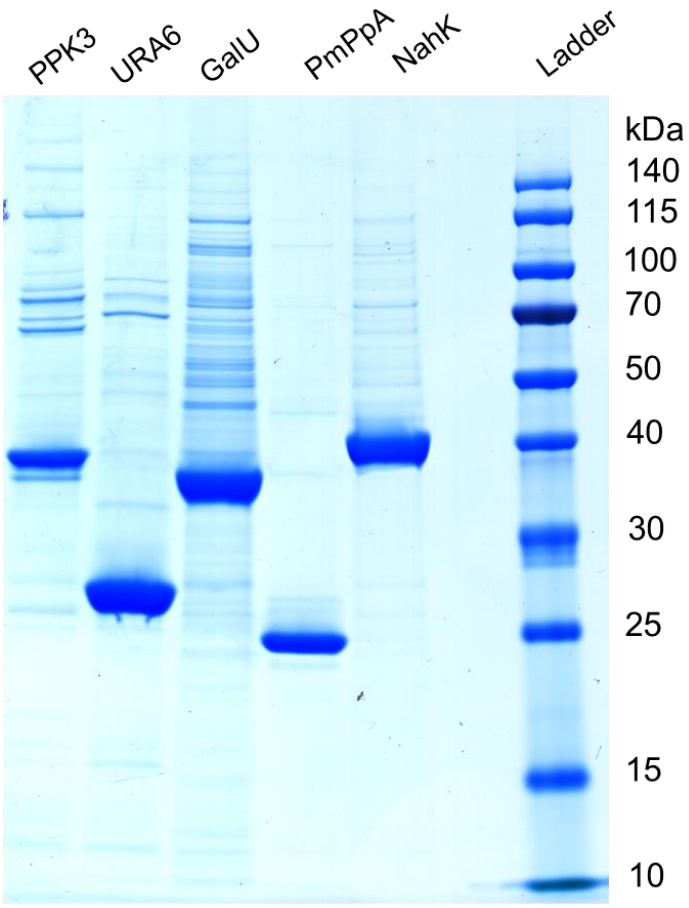


Figure 3: SDS-PAGE analysis after purification of the five recombinant enzymes used in the study: All enzymes were purified by immobilized metal ion affinity chromatography (IMAC). 5 µg of each enzyme loaded; staining with Coomassie blue.

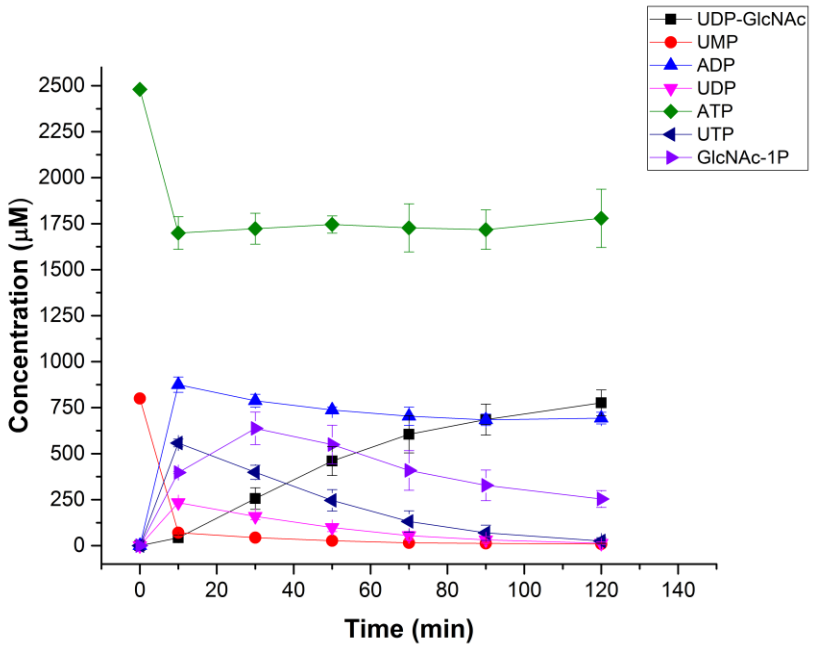


Figure 4: One pot multienzymatic production of UDP-GlcNAc. Each sampling point was analyzed in triplicate by UV and CD (GlcNAc-1-phosphat). Error bars represent standard deviation of biological triplicates. Enzyme concentrations: NahK, GalU, PPK3, and URA6 all at 0.1 mg mL^{-1}; PmPpA 0.03 mg mL^{-1}. Connecting lines are just for illustration.

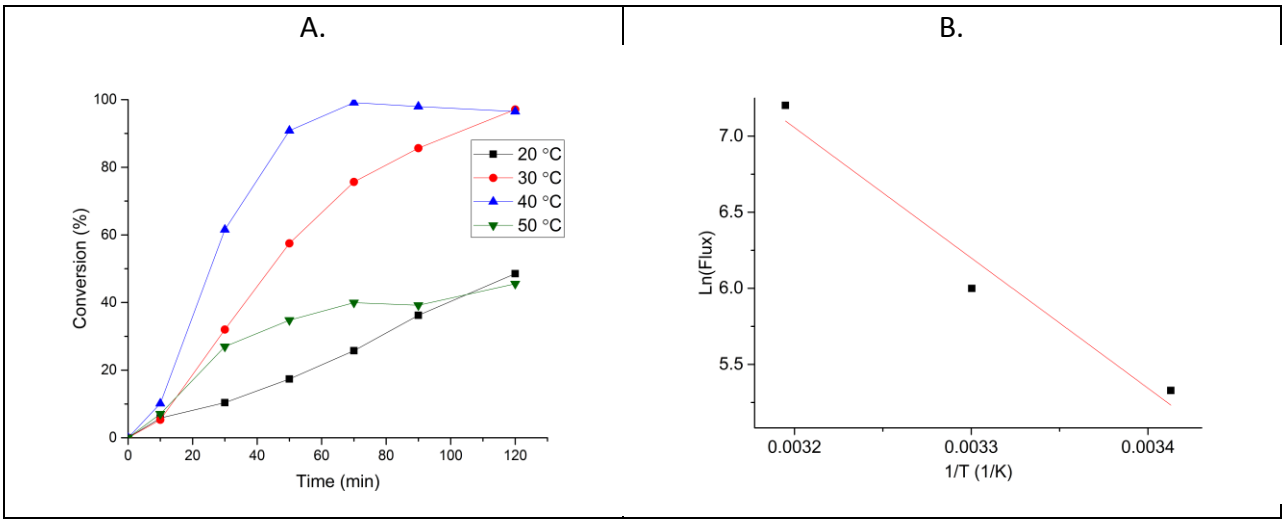


Figure 5: (A) Conversion of UMP to UDP-GlcNAc over time for different temperatures. Enzyme concentrations: NahK, GalU, PPK3, and URA6 all at 0.1 mg mL^{-1}; PmPpA 0.03 mg mL^{-1}. Connecting lines are just for illustration. (B) Arrhenius plot for the reaction at 20, 30, and 40°C. The R^2 for linear regression is 0.97. The derived equation is shown in Equation 6.

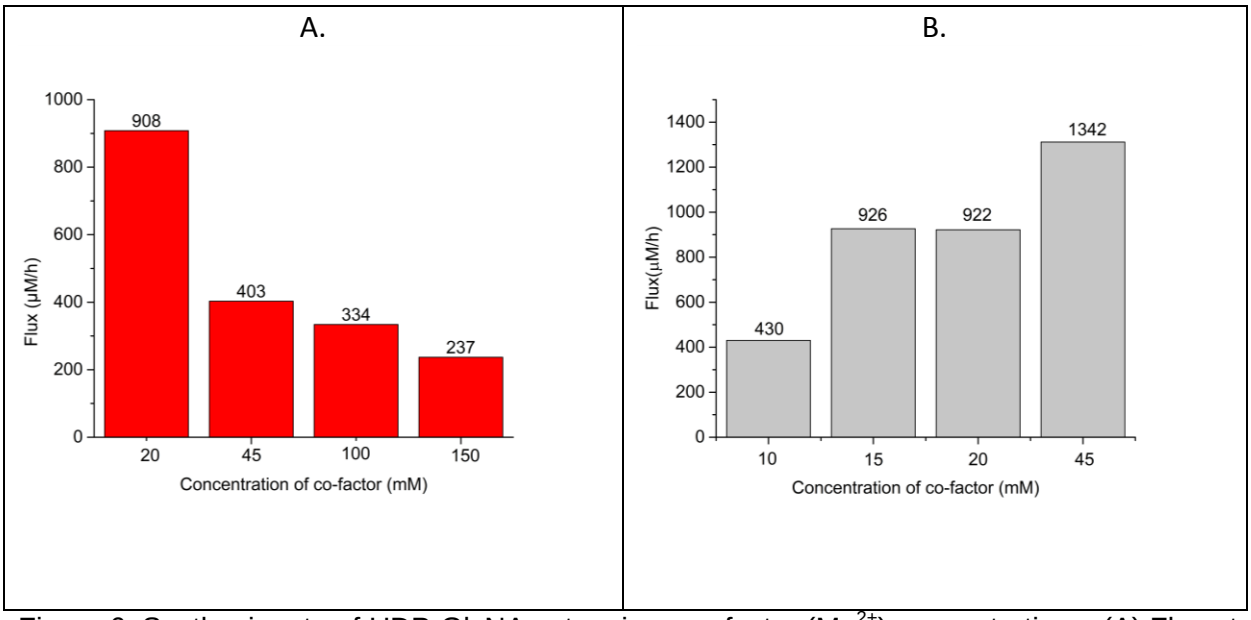
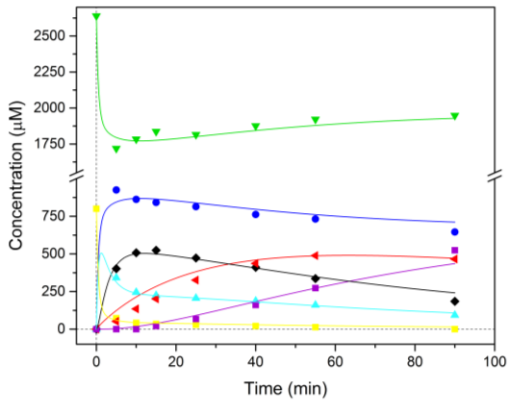
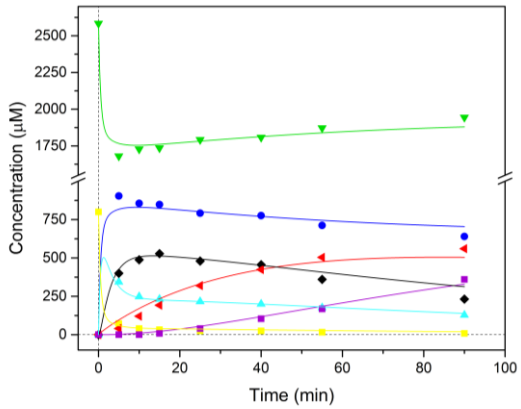


Figure 6: Synthesis rate of UDP-GlcNAc at various co-factor (Mg^{2+}) concentrations. (A) Flux at 30°C and (B) Flux at 40°C. Enzyme concentrations: NahK, GalU, PPK3 and URA6 all at 0.1 mg mL^{-1}; PmPpA 0.03 mg mL^{-1}.

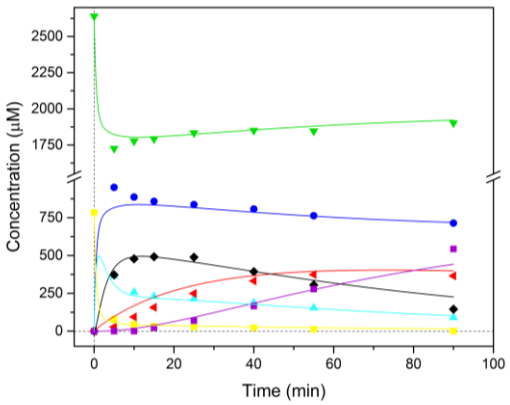
Experiment 1



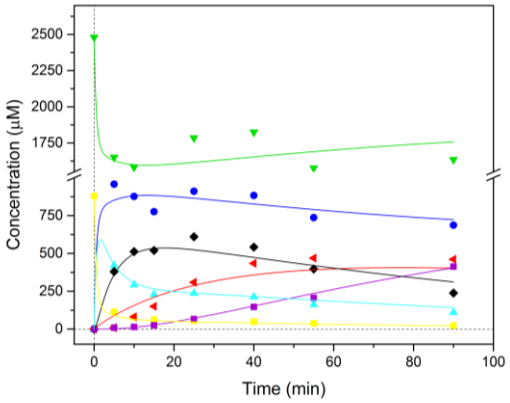
Experiment 2



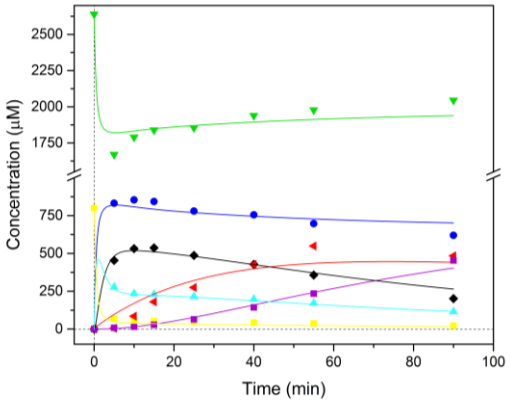
Experiment 3



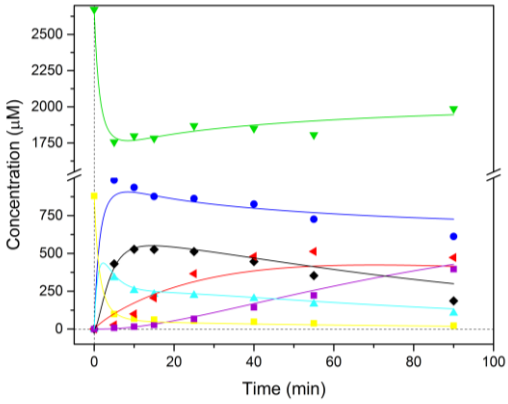
Experiment 4



Experiment 5



Experiment 6



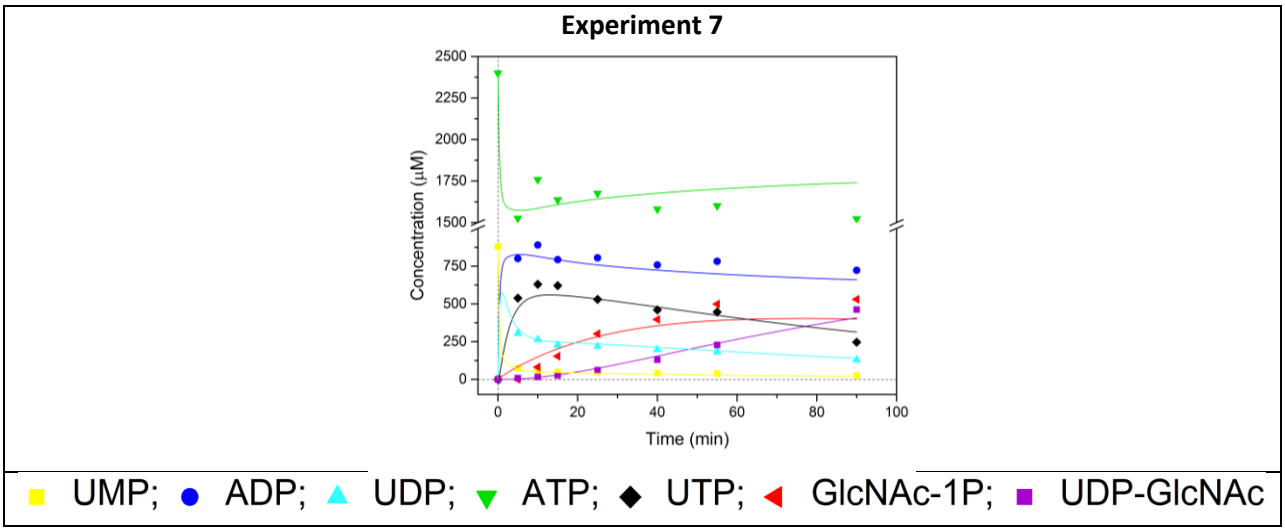


Figure 7: Dynamic kinetic modeling of the cascade. Seven independent multienzyme reactions (experiment 1-7) with various concentrations of NahK, PPK3, UAR6 and GalU were carried out and fitted (one-fit) to the model detailed in Table IV . Lines show the fitted model and dots show the experimental data; numbers in the graphs refer to experiment number (Table II)

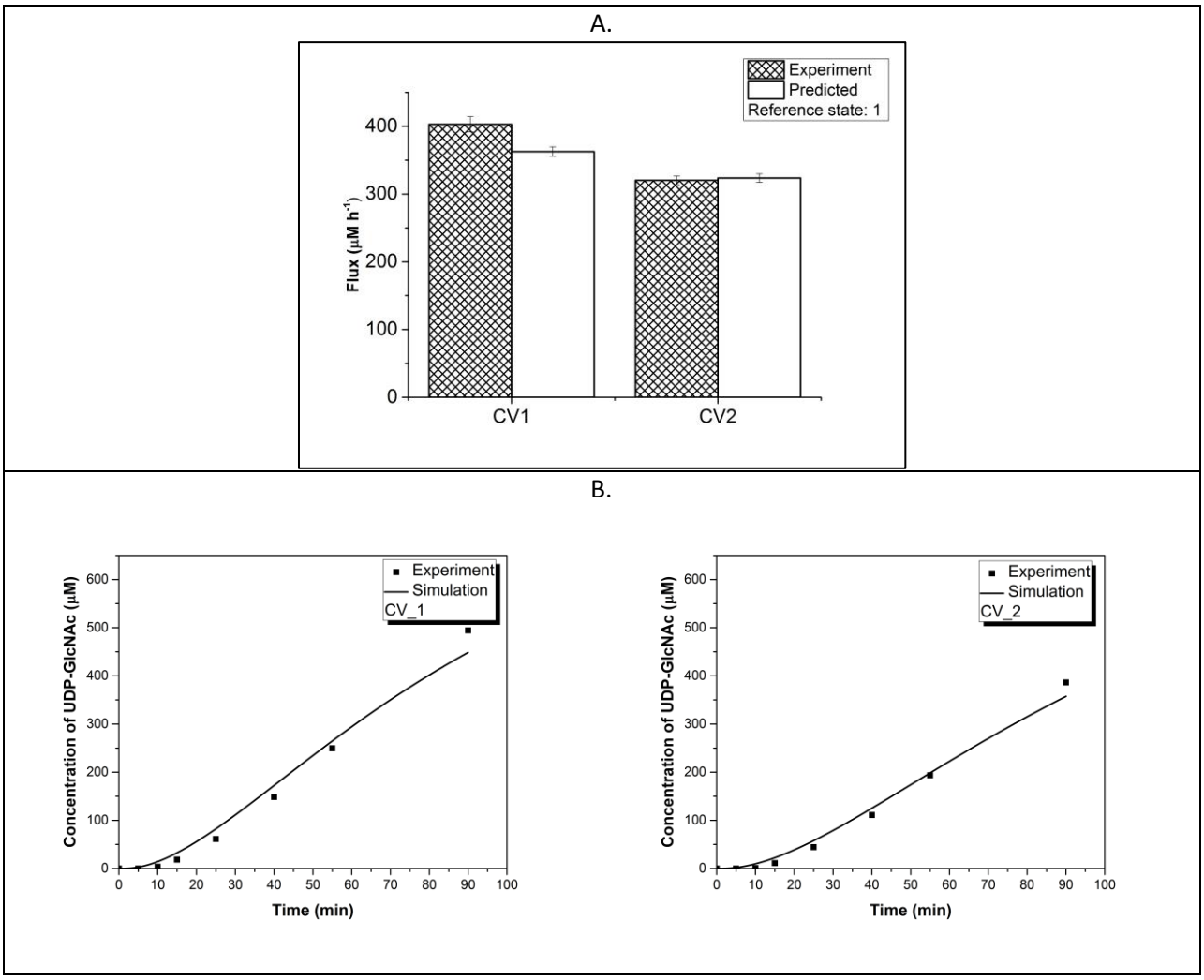


Figure 8: Comparison of experimental versus simulated UDP-GlcNAc concentration by the design equation (A) and based on dynamic mass action kinetics model (B). For the flux prediction by the dynamic model FCCs based on reference state 1 were used. For the prediction based on other reference states see supplementary information Figure SI 1. Error bars representing the standard error of regression were too small to be visualized.

Tables

Table I: Bio wet mass (BWM) and protein yields from *E.coli* cultivation in 500 mL media at 37°C as specified in 2.3.2.

Enzyme	**Media**	**BWM / cultivation ($g\ L^{-1}$)**	**Protein / BWM ($mg\ g^{-1}$)**	**Protein conc. after IMAC ($mg\ mL^{-1}$)**
NahK	LB	4.2	46.0	2.3
GalU	LB	4.8	32.6	5.0
PmPpa	LB	5.3	13.3	4.1
PPK3	TB	6.5	17.8	7.4
URA6	TB	6.8	12.6	3.2

Table II: Multienzyme cascade reactions performed. Reactions 1-7: Perturbation of enzyme concentrations to estimate flux control coefficients. Reactions CV_1 and CV_2: Control experiments for the cross validation of the kinetic model.

Experiment	**NahK (mg ml^{-1})**	**GalU (mg ml^{-1})**	**PPK3 (mg ml^{-1})**	**URA6 (mg ml^{-1})**	**PmPpA (mg ml^{-1})**	**Flux (µM h^{-1})**
1	0.13	0.1	0.1	0.1	0.03	426
2	0.1	0.07	0.1	0.1	0.03	300
3	0.1	0.13	0.1	0.1	0.03	442
4	0.1	0.1	0.06	0.1	0.03	318
5	0.1	0.1	0.15	0.1	0.03	362
6	0.1	0.1	0.1	0.04	0.03	303
7	0.1	0.1	0.1	0.15	0.03	376
CV_1	0.1	0.1	0.1	0.1	0.03	403
CV_2	0.07	0.1	0.1	0.1	0.03	320

Table III: Calculated flux control coefficients. Values of FCCs differ depending on the selected reference state. The first column states which experiment was selected as a reference state (see Table II) for the calculation of the FCCs. For any reference state, FCCs are statistically significant ($p<0.05$).

Ref. State	C^J_{NahK}	C^J_{GalU}	C^J_{PPK3}	C^J_{URA6}	$\sum C^J$	R^2_{adj}	**SSE**
Exp. 1	0.26	0.53	0.18	0.13	1.09	0.58	0.83
Exp. 2	0.24	0.54	0.13	0.09	0.99	0.58	0.66
Exp. 3	0.35	0.43	0.19	0.14	1.1	0.6	1.37
Exp. 4	0.25	0.40	0.23	0.10	0.98	0.27	2.23
Exp. 5	0.28	0.45	0.10	0.11	0.96	0.46	1.45
Exp. 6	0.24	0.38	0.13	0.24	0.99	0.59	0.67
Exp. 7	0.30	0.47	0.16	0.08	1.01	0.3	2.96

Table IV: Dynamic kinetic model based on mass action kinetics with estimated parameters. The parameters were estimated by performing one-fit to the experiments detailed in Table II. The standard deviations for all parameters are below 1%. For more details, see Appendix B of the supplementary information.

Enzyme	**EC No.**	**Catalyzed reaction**	**Reaction rate equation**	**Estimated parameters**
NahK	2.7.1.2	GlcNAc + ATP → GlcNAc-1P + ADP	$r_1 = k_1\, e_{NahK}[GlcNAc][ATP]$	k_1 = 109.08 $L^2\ min^{-1}\ g^{-1}\ mol^{-1}$
GalU - PmPpA	2.7.7.9 3.6.1.1	UTP + GlcNAc-1P → UDP-GlcNAc + 2 Pi	$r_2 = k_2\, e_{GalU}[GlcNAc1P][UTP]$	k_2 = 330.52 $L^2\ min^{-1}\ g^{-1}\ mol^{-1}$
PPK3	2.7.4.1	$PolyP_n$ + ADP ↔ $PolyP_{n-1}$ + ATP $PolyP_n$ + UDP ↔ $PolyP_{n-1}$ + UTP	$r_3 = k_3\, e_{pk3}\left([ADP] - \frac{[ATP]}{K_{eq1}}\right)$ $r_4 = k_4\, e_{pk3}\left([UDP] - \frac{[UTP]}{K_{eq2}}\right)$	k_3 = 0.77 L $min^{-1}\ g^{-1}$ K_{eq1} = 2.90 - k_4 = 2.90 L $min^{-1}\ g^{-1}$ K_{eq2} =2.33 -
URA6	2.7.4.22	UMP + ATP ↔ UDP + ADP	$r_5 = k_5\, e_{URA6}\left([UMP][ATP] - \frac{[UDP][ADP]}{K_{eq3}}\right)$	k_5 = 6889.92 $L^2\ min^{-1}\ g^{-1}\ mol^{-1}$ K_{eq3} =2.72 -

Supplementary information

Material & Methods

Chemicals

The following chemicals were purchased from Sigma Aldrich Co. LLC (Munich, Germany): Adenosine 5′-diphosphate sodium salt (A2754, ≥ 95 %), adenosine 5′-triphosphate di-sodium salt hydrate (A7699, ≥ 99 %), *N*-acetyl-D-glucosamine 1-phosphate di-sodium salt (A2142, ≥ 95 %), uridine 5′-monophosphate (U1752, ≥ 99 %), uridine 5′-diphosphate disodium salt hydrate (94330**,** ≥ 96 %), uridine 5′-diphospho-*N*-acetylglucosamine sodium salt (U4375**,** ≥ 98 %). *N*-acetyl-D-glucosamine (8993.2, ≥ 99 %); uridine-5'-triphosphate tri-sodium salt tri-hydrate (K055.2, ≥ 90 %); HCl (4025, 37 %); imidazole (3899.3, ≥ 99 %); tetra-sodium diphosphate decahydrate (T883.1, ≥ 99 %), and substances for the LB- and TB-media (tryptone (8952.4), yeast extract (2363.2), glycerol (3783.2, ≥ 99,5 %), monopotassium phosphate (3904.1, ≥ 99 %), dipotassium phosphate (P749.2, ≥ 99 %), tetra-sodium pyrophosphate deca-hydrate (0269.1) and ammonia heptamolybdate tetra hydrat (7311.1) were purchased from Carl Roth GmbH&Co.KG (Karlsruhe, Germany). Kanamycin sulfate (A1493), isopropyl-β-D-thiogalactopyranosid (A1008.0025, ≥ 99 %), magnesium chloride (A2264, > 99 %) and Tris (hydroxymethyl)-aminomethan-Buffer (A2264, > 99,9 %) were bought from AppliChem GmbH (Darmstadt, Germany). The sodium hexametaphosphate ($PolyP_{14}$) with a chain length of 14 phosphate units on average was a kind gift from Thermphos International B.V. (Wittenberg, Germany).

Table SI 1: Media composition for protein expression.

	LB-media	TB-media
Tryptone	10 g L^{-1}	12 g L^{-1}
Yeast extract	10 g L^{-1}	24 g L^{-1}
NaCl	5 g L^{-1}	-
Glycerol	-	0.4% (v/v)
KH_2PO_4	-	2.31 g L^{-1}
K_2HPO_4	-	12.54 g L^{-1}

Analytics: HPAEC-UV/CD

The AG11 (50 x 2mm) column was applied as a guard column, followed by two in-series analytical columns AS11 (250 x 2 mm). All separations were performed at 25°C. All samples and standards were injected three times with a volume of 12.5 µL. Eluent flow was 0.35 mL min^{-1}. All columns, components, and the software were purchased from Thermo Scientific (Darmstadt, Germany). Data logging and chromatogram analysis was performed with Chromeleon 6.60 software. The statistical evaluation of the quantification of metabolite concentrations by linear regression is shown in Appendix A.

Multienzymatic experiments for kinetic analysis

The experiments were conducted in 1.5 mL vials (Eppendorf AG, Hamburg, Germany) rotated at 300 rpm in a thermomixer (Eppendorf AG, Hamburg, Germany). Reactions were started by addition of GlcNAc, UMP, ATP (1 mM) to pre-mixed reaction volumes containing enzymes, buffer, co-factor and PolyP14 (2 mM) (PmPpA is not able to hydrolyse PolyP14 – unpublished data). To obtain reaction time series, 100 µL of samples were aliquoted and diluted in 700 µL of MilliQ water, preheated to 90°C, followed by another 10 min of heating at 90°C. The quenching protocol was individually tested on each enzyme (except PmPpA) to ensure

enzyme inactivation. Heating did not have significant effects on the concentration of the compounds (data not shown).

Results

Negative control experiments

In each reaction of negative control experiments, one enzyme of the cascade was not included in the reaction mixture. The concentration of UDP-GlcNAc was monitored over 120 minutes. None of negative control sets showed production of UDP-GlcNAc (see Table SI 2).

Table SI 2: Set of negative control experiments.

Negative control set	NahK (mg mL^{-1})	GalU (mg mL^{-1})	PPK3 (mg mL^{-1})	URA6 (mg mL^{-1})	Flux (µM h^{-1})
I	0	0.1	0.1	0.1	0
II	0.1	0	0.1	0.1	0
III	0.1	0.1	0	0.1	0
IV	0.1	0.1	0.1	0	0

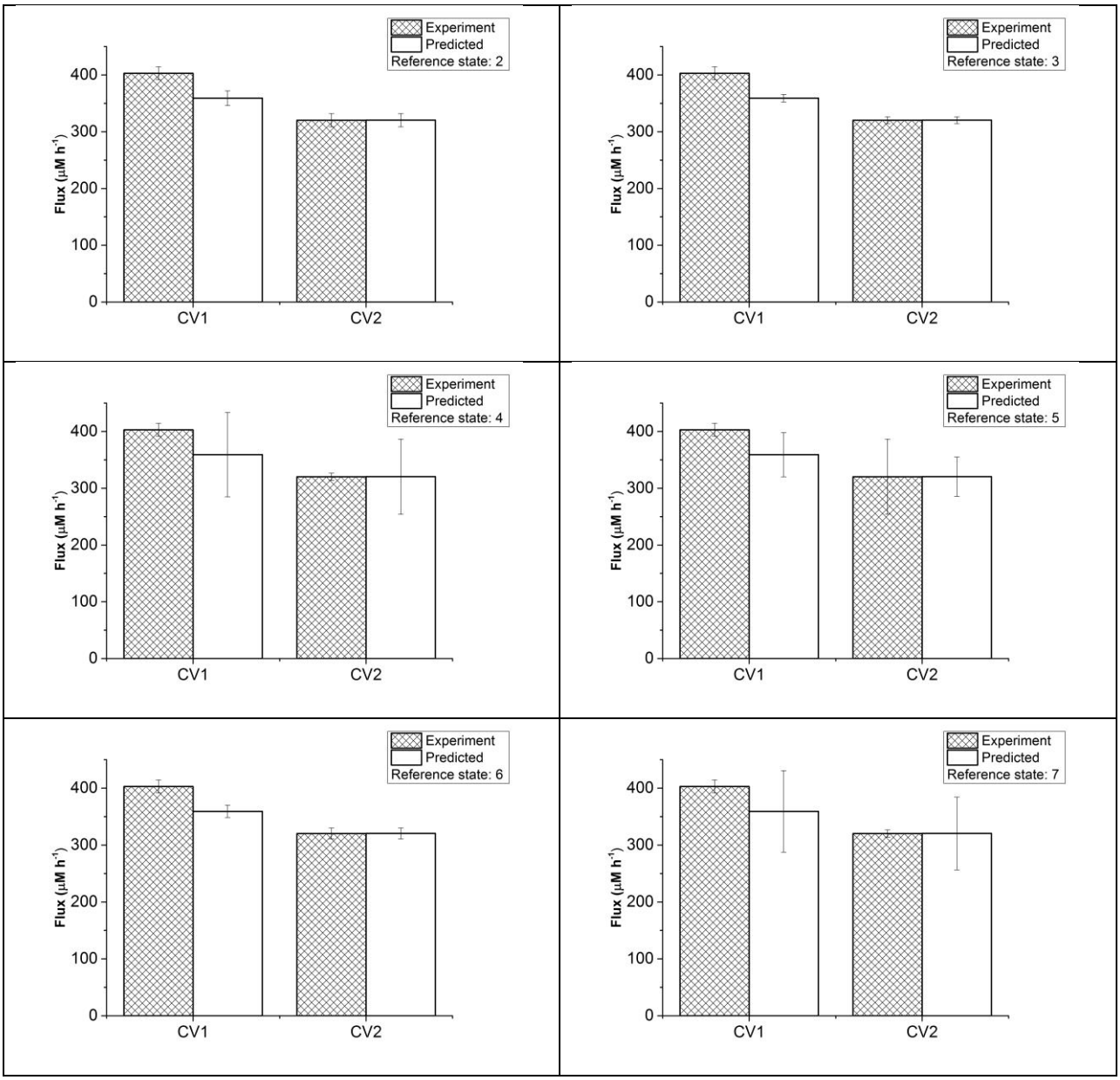


Figure SI 1: Comparison of experimental versus predicted UDP-GlcNAc flux by the design equation (various references states; see Table III). Error bars are the standard error of regression. For the simulated values the standard error of linear regression was taken into account.

Appendix A

Method validation of the HPAEC-UV/CD measurement for the quantification of metabolites

The validation was carried out according to Gottwald (Gottwald, 2000). Standard samples containing all metabolites were prepared in concentrations ranging from 5 to 170 µM for method validation and as internal standards for kinetic measurements. For the method validation lower (5 µM) and upper limits (170 µM) were measured at least 6 times and all other concentrations, with some exceptions, 3 times with the established gradient. GlcNAc-1P was measured by conductivity – all other metabolites by UV detection. For each kinetic measurements, internal calibration standards were used.

Test for linearity- Mandel test

The mandel test was applied to test for linearity (Gottwald, 2000) (Mandel, 1964). The results are shown in Table SI A I. UV/CD signals are linear over the range of concentrations (5 to 170 µM – except UDP-GlcNAc (130 µM)) investigated for all metabolites (see Table SI A I). This is as the test statistic for each metabolite is smaller than the corresponding f-value.

Table SI A I: The HPAEC-UV/CD method was tested for linearity of UV/CD signals over concentration by the mandel test. It was found that signals are linear in the range from 5 to 170 µM (130 µM for UDP-GlcNAc). For calculation of the parameters, we refer to literature.

Metabolite	Variance Difference	Test Statistic	F-Value
ATP	0.1015	0.857	34.12
ADP	0.0203	0.409	34.12.
UMP	0.0011	0.138	98.5
UDP	0.0007	0.428	98.5
UTP	0.3584	0.795	34.12
UDP-GlcNAc	0.0034	0.806	98.5
GlcNAc-1P	0.0006	7.100	34.12

Test for homgeneity of variances

Variances were tested for homogeneity to decide on whether a linear or a weighed linear regression applies. While for most metabolites variances determined from standard samples were inhomogeneous, there was no significant difference between linear and weighed linear regression for the calculation of concentration. Thus, we used linear regressions for the calculation of metabolite concentrations.

Table SI A II: Linear regression of the UV/CD signals over the standard sample concentrations. RSD = Residual standard deviation.

Linear regression: y = b0 + b1 · xi							
Compound	**ATP**	**ADP**	**UMP**	**UDP**	**UTP**	**UDP-GlcNAc**	**GlcNAc-1P**
Slope b1	0.098	0.084	0.054	0.066	0.061	0.062	0.002
Intercept b0	-0.096	-0.039	0.126	-0.070	-0.162	-0.021	0.032
RSD	0.3380	0.2055	0.0759	0.0358	0.1488	0.0631	0.0146
Variation of the procedure	3.4631	2.4433	1.4152	0.5426	2.4559	1.0245	6.4752
Coeff. variation of the procedure [%]	4.57	3.22	2.18	0.83	3.78	1.80	8.54
Coeff. of determination	0.9978	0.9989	0.9997	1.0000	0.9990	0.9997	0.9924

Appendix B

Error estimation of the kinetic model

The standard deviations were estimated by running a fit analysis using the systems biology toolbox of MATLAB (Schmidt and Jirstrand, 2006). The fit analysis was conducted by running 100 estimations with randomly perturbed parameter values. Perturbed parameters were obtained from an exponential distribution of 0.5 around the values stated in Table IV. The results of the analysis are depicted in the boxplot below (see Figure SI B 1).

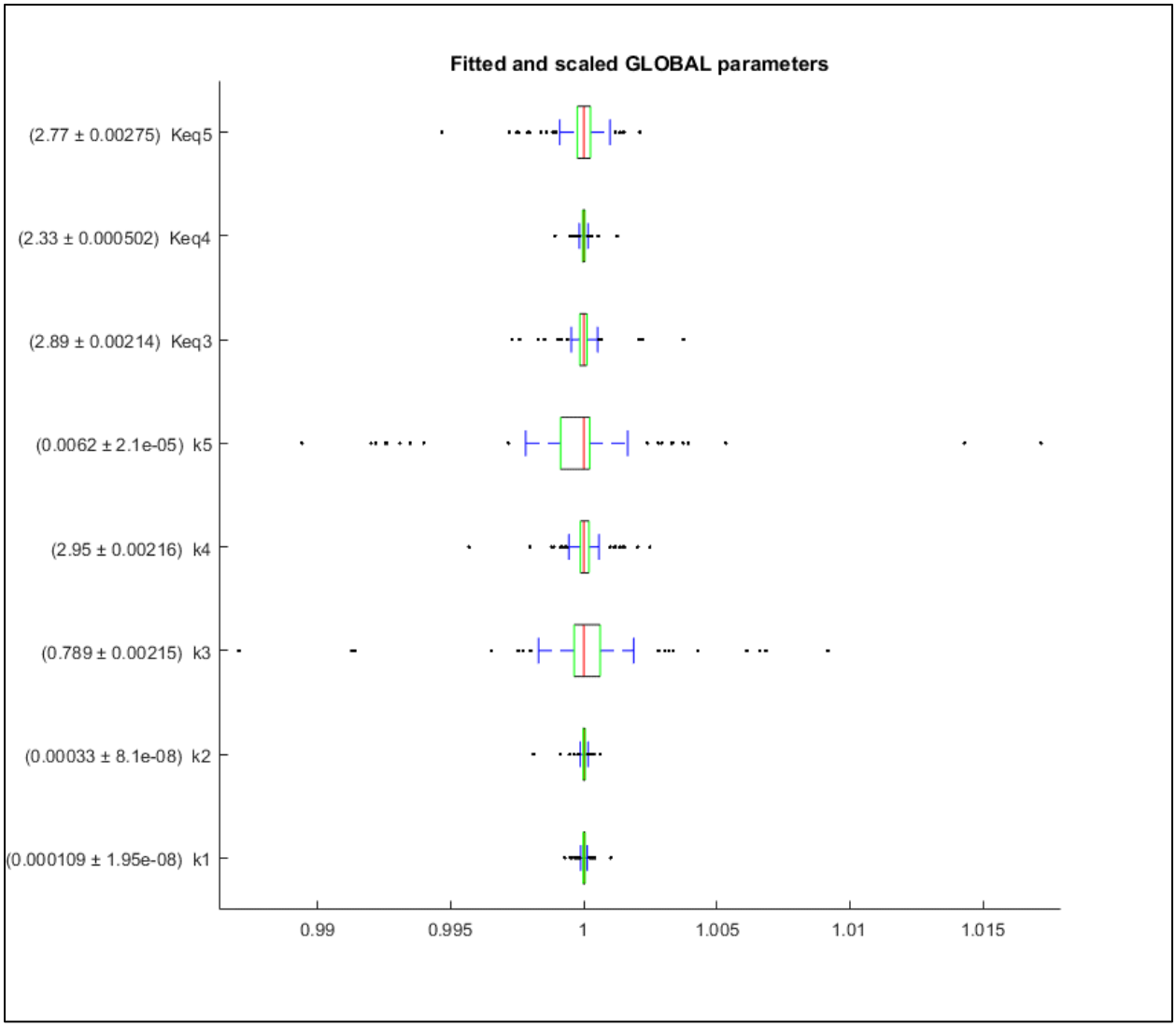


Figure SI B 1: Fitted and scaled global parameters. The data is normalized such that the median (red) of each parameter is equal to one.