Graz University of Technology

Chraracterization and application of three distinct *E. coli* sugar phosphate phosphatases

Masterthesis

by

Martin Pfeiffer

Institute of Biotechnology and Biochemical Engineering
Graz University of Technology

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

Graz, im Mai 2014

STATUTORY DECLARATION

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 marked all material which has been quoted either literally or by content from the used sources.

…… ……..

date (signature)

Acknowledgements

First of all, I want to thank my parents for their great support on so many different levels, whiteout you, I would not have come so far. Furthermore, I want to say “Thank you” to my beloved girlfriend Rahel, which shared with me the ups and downs, which I experienced during the work on the master thesis.
For her great supervision I want to thank Patricia Wildberger. During the time that we were working together on the project, I learned so much from you and I also enjoyed working together with you as a team.
A big “thank you” also goes to my main supervisor Univ. –Prof. Dipl. –Ing. Dr. Bernd Nidetzky for great supervision in general.

Furthermore I want to thank all the people in the Bioprocess Engeneering Department for the great working atmosphere and for all the help that they provided. In Addition I also want to thank Gerald N. Rechberger, Ruth Birner-Gruenberger und Lothar Brecker for their help with MS- and NMR Analysis.

Content

1 Phosphoryl transfer from α-D-glucose 1-phosphate catalyzed by *Escherichia coli* sugar-phosphate phosphatases of two protein-superfamily types

Patricia Wildberger[a], Martin Pfeiffer[a], Lothar Brecker[b], Gerald N. Rechberger[c,e], Ruth Birner-Gruenberger[d,e] and Bernd Nidetzky[a,f]#

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

[b] Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, A-1090 Vienna, Austria

[c] Institute of Molecular Biosciences, University of Graz, Humboltstraße 50/2, A-8010 Graz, Austria

[d] Institute of Pathology, and Center of Medical Research, Medical University of Graz, Stiftingtalstrasse 24, A-8036 Graz, Austria

[e] Omics Center Graz, Stiftingtalstrasse 24, A-8036 Graz, Austria

[f] Austrian Centre of Industrial Biotechnology (ACIB), Petersgasse 14, A-8010 Graz, Austria

Running Head: Enzymatic phosphoryl transfer from glucose 1-phosphate

Corresponding author; Tel. +43-316-873-8400; e-mail: bernd.nidetzky@tugraz.at

1.1 Abstract

The Cori ester α-D-glucose 1-phosphate (αGlc 1-*P*) is a high-energy intermediate of cellular carbohydrate metabolism. Its glycosidic phosphomonoester moiety primes αGlc 1-*P* for flexible exploitation in glucosyl and phosphoryl transfer reactions. Two structurally and mechanistically distinct sugar-phosphate phosphatases from *Escherichia coli* were characterized in this study for utilization of αGlc 1-*P* as phosphoryl donor substrate. The *agp* gene encodes a periplasmic αGlc 1-*P* phosphatase belonging to the histidine acid phosphatase family. Had13 is from the haloacid dehydrogenase-like phosphatase family. First-time cytoplasmic expression of αGlc 1-*P* phosphatase (in *E. coli* Origami B) gave a functional enzyme preparation (k_{cat} for hydrolysis of αGlc 1-*P* = 40 s^{-1}) that was shown by mass spectrometry to exhibit the native intramolecular disulfide bond between Cys^{189}-Cys^{195}. Enzymatic hydrolysis of αGlc 1-*P* in $H_2{}^{18}O$ solvent proceeded with complete ^{18}O-label incorporation into the phosphate released, consistent with catalytic reaction through O1-P, but not C1-O bond cleavage. Hydrolase activity of both enzymes was not restricted to a glycosidic phosphomonoester substrate, and D-glucose 6-phosphate was converted with similar k_{cat} as αGlc 1-*P*. By examining phosphoryl transfer from αGlc 1-*P* to an acceptor substrate other than water (D-fructose, D-glucose), we discovered αGlc 1-*P* phosphatase to exhibit pronounced *synthetic* activity, contrary to Had13, which utilized αGlc 1-*P* mainly for hydrolysis. By applying D-fructose in 10-fold molar excess over αGlc 1-*P* (20 mM), enzymatic conversion furnished D-fructose 1-phosphate as the main product in 55% overall yield. αGlc 1-*P* phosphatase is the first biocatalyst reported for use in transphosphorylation from αGlc 1-*P*.

1.2 Introduction

Phosphorylation of sugar substrates is a common biochemical transformation of central importance to cellular metabolism (1–3). It usually involves phosphoryl transfer from a phospho-activated donor substrate such as ATP to an acceptor group, typically a hydroxyl, on the sugar backbone (4–7). Various phosphotransferases (EC 2.7) catalyze sugar phosphorylation (8–12). In an alternative reaction catalyzed by glycoside phosphorylases (EC 2.4), where phosphorylation occurs exclusively at the sugar's anomeric position, a glycosyl residue is transferred from a sugar donor substrate to phosphate (Scheme 1A) (13, 14). The phosphomonoester moiety attached to sugars is a key element of biological recognition, across all steps of glycolysis for example, and it serves to prime sugars for further conversion in different biochemical pathways (15–18). It is known from intracellular metabolite profiling studies that concentration changes of common sugar phosphates (e.g. D-glucose 6-phosphate, Glc 6-*P*; D-fructose 6-phosphate, Fru 6-*P*) are often linked to major alterations in cellular physiology (19–22). Due to the requirement of authentic reference material in different biological investigations, there is considerable interest in the synthetic preparation of sugar phosphates (20). Technologically, sugar phosphates are applied as nutritional supplements and taste enhancers in food and feed products; moisturizing ingredients in cosmetics; ionic surface-active reagents in detergents; and building blocks of new polymers (23–25).

Naturally known phosphorylation reactions have been exploited with mixed success for the biocatalytic synthesis of sugar phosphates. Glycosyl transfer to phosphate is fundamentally limited in application to hexose 1-phosphate products, and the relatively narrow substrate scope of glycoside phosphorylase reactions restricts one to just a few glycosyl phosphates for which effective production has been demonstrated at preparative scale (e.g. α- and βGlc 1-*P*

(26, 27), α-D-galactose 1-phosphate (28)). Phosphotransferase reactions by contrast offer convenient access to a large diversity of sugar phosphate products, as shown for nucleoside triphosphate dependent phosphorylation of various hexose substrates by sugar kinases for example (29–33). However, high costs of the phospho-activated donor substrate and enzyme inhibition by the resulting dephosphorylation product (e.g. ADP from ATP) necessitate that the phosphoryl donor is supplied in only catalytic amounts and therefore regenerated continuously during the reaction (Scheme 1B) (34, 35). However, this makes the biocatalytic phosphoryl transfer a technically complex overall transformation of currently limited use in sugar phosphate synthesis.

Transphosphorylation catalyzed by phosphatases in an alteration of their natural phosphomonoester hydrolysis reaction was therefore considered as alternative route towards sugar phosphates (Scheme 1C) (36–39). Under conditions where sugar was present in a concentration high enough to outcompete effectively the reaction with water, some phosphatases (e.g. acid phosphatase) promoted sugar phosphate formation in moderate (e.g. D-mannose 6-phosphate; 15%) to excellent yields (e.g. Glc 6-*P*; up to 95%) (37). Usage of expedient phosphoryl donor substrates such as inorganic pyro- or oligo-phosphate was advantageous. However, high preponderance of donor substrate hydrolysis, fast secondary hydrolysis of sugar phosphate product, and combined inhibition and mass action effects of the released phosphates (Scheme 1C) were critical issues of phosphatase-catalyzed synthesis (40, 41). Identification of transphosphorylation systems with improved synthetic efficacy would therefore be important. Herein a new approach of biocatalytic transphosphorylation from the Cori ester αGlc 1-*P* was examined.

αGlc 1-*P* is a central intermediate of cellular carbohydrate metabolism. Its free energy of hydrolysis (Δ*G*°_hydrolysis) at 25 °C is −5.0 kcal/mol, which is similar to pyrophosphate (Δ*G*°_hydrolysis = −4.6 kcal/mol) and places αGlc 1-*P* slightly below ATP (Δ*G*°_hydrolysis = −7.3 kcal/mol) in terms of energy content (2). αGlc 1-*P* is produced conveniently from different saccharides, especially from sucrose (26). Its glycosidic phosphomonoester group primes αGlc 1-*P* for flexible exploitation in glucosyl and phosphoryl transfer reactions, however up to now, biocatalytic use of αGlc 1-*P* was restricted entirely to glucoside synthesis (e.g. (42, 43)). Considering that phosphorylation of sugar substrates was desired, we reasoned that αGlc 1-*P* active phosphatases might also exhibit suitable preference for sugars as phosphoryl acceptors. The notion was put to critical test by studying two structurally and mechanistically distinct sugar-phosphate utilizing phosphatases that we selected from *Escherichia coli*.

αGlc 1-*P* phosphatase is the *agp* gene product, and the enzyme is located to the *E. coli* periplasm (44). Crystal structure of the enzyme complex with glucosyl phosphate ligand reveals a two-domain protein topology typical of members of the high-molecular-weight histidine acid phosphatase family (45). The protein fold comprises a discrete α-helical domain next to a α/β domain and a catalytic center located in a deep cleft between the two domains (Figure 1A) (45). The αGlc 1-*P* phosphatase structure exhibits 3 disulfide bridges between Cys^{94} and Cys^{125}, Cys^{189} and Cys^{195}, as well as Cys^{384} and Cys^{392} (45). The active site contains a highly conserved histidine (His^{18}, substituted by Ala in Figure 1B) that is thought to function as catalytic nucleophile. Asp^{290} is the likely general acid-base catalyst, and the phosphomonoester group of αGlc 1-*P* is held tightly in place through a cluster of positively charged residues (Arg^{17}, Arg^{21}, Arg^{94}, His^{289}) (45). The proposed catalytic reaction of αGlc 1-

P phosphatase follows a double displacement-like mechanism via a covalent phospho-histidine intermediate, as shown in Scheme 2A.

The second phosphatase studied was Had13, which belongs to the haloacid dehalogenase-like phosphatase family (46). The Had13 structure, which was solved in the absence of ligands except for bound metal ion (Mg^{2+}), comprises a modified Rossmann fold-like α/β-core domain that has a smaller "cap" domain on its top (Figure 1C) (47). The active site is located at the interface of the two domains and contains two key Asp residues (Figure 1D), which are highly conserved and where one of which is believed to function as the catalytic nucleophile (Asp^{10}) and the other to function as the general acid-base (Asp^{12}). Conserved Asp residues (Asp^{221}, Asp^{225}) function in the coordination of Mg^{2+} (48, 49). By similarity to other Had-type enzymes (50), the substrate's phosphoryl group is probably accommodated in Had13 through Thr^{44} and Lys^{198}. Catalytic reaction of Had13 is proposed to occur in two steps via a covalently modified aspartyl-phosphate enzyme (Scheme 2B).

In both αGlc 1-*P* phosphatase and Had13, the phospho-enzyme intermediate is the potential source of transphosphorylation activity, considering that dephosphorylation of the enzyme could occur to a sugar acceptor rather than water. Herein we report recombinant production and isolation of both enzymes, and their detailed characterization in the conversion of αGlc 1-*P* under hydrolysis-only and transphosphorylation conditions.

1.3 Materials and methods

1.3.1 Chemicals and reagents.

Unless stated, all chemicals were of highest purity from Sigma-Aldrich (Vienna, Austria) or Roth (Karlsruhe, Germany). GeneJET genomic DNA purification kit was from Thermo Fisher Scientific (Waltham, USA). Oligonucleotides were from Sigma-Aldrich. DNA sequencing was performed at LGC Genomics (Berlin, Germany). Electrocompetent *E. coli* Origami B (DE3) cells were from Novagen (Merck KgaA, Darmstadt, Germany). $H_2{}^{18}O$ (97% isotopic purity) was from Sigma-Aldrich. The sodium salts of αGlc 1-*P*, Glc 6-*P*, Fru 6-*P*, and pyrophosphate were from Sigma-Aldrich. The barium salt of Fru 1-*P* was also from Sigma-Aldrich. Phytate (sodium salt) was from Roth. The potassium salt of βGlc 1-*P* (27) was from Prof. Tom Desmet (Ghent University, Belgium). N,O-bis(trimethylsilyl)trifluoroacetamide with 1% trimethylchlorosilane (BSTFA/1%TMCS) and pyridine were from Sigma-Aldrich. Purified preparations of sucrose phosphorylase from *Leuconostoc mesenteroides* (SPase) (51) and mannitol 1-phosphate dehydrogenase from *Aspergillus fumigatus* (52) were obtained by reported procedures.

1.3.2 Molecular cloning, expression and purification of αGlc 1-*P* phosphatase and Had13.

The *agp* gene was amplified from genomic DNA of *E. coli* BL21-Gold (DE3). The native signal sequence was removed, and an *N*-terminal *Strep*-tag was added. The final construct was cloned into linearized pMS470_dsbC vector via Gibson assembly (53), and the plasmid vector was then transformed into *E. coli* Origami B (DE3) cells. The detailed experimental procedure

is given in the Appendix. The vector p15TvL-Had13, encoding an *N*-terminally *His*-tagged Had13 (46), was transformed into electrocompetent *E. coli* BL21-Gold (DE3) cells.

Recipient *E. coli* strains were cultivated in 1-L baffeled shaken flasks at 37°C and 110 rpm using Lennox-medium containing 0.115 mg/mL ampicillin. When OD_{600} had reached 0.8, temperature was decreased to 18 °C prior to induction with isopropyl-β-D-thiogalactopyranoside (0.01 mM, αGlc 1-*P* phosphatase; 0.4 mM, Had13). After 20 h, cells were centrifuged at 4°C and 4,420 *g* for 30 min (Sorvall RC-5B Refrigerated Superspeed centrifuge; Du Pont Instruments, Newtown, USA). The pellet was resuspended in 50 mM Mes, pH 7.0 (αGlc 1-*P* phosphatase) or 50 mM Hepes (Had13), pH 7.0, and frozen at –20°C. Thawed cell suspension was passed twice through a French pressure cell press (American Instruments, Silver Springs, USA) at 150 bar, and cell debris was removed by centrifugation at 4°C, 20,000 *g* for 30 min.

αGlc 1-*P* phosphatase was isolated from the crude extract using a *Strep*-Tactin Sepharose column (IBA, Göttingen, Germany), using a general protocol described elsewhere (51). Had13 was isolated using a Cu^{2+}-loaded IMAC Sepharose High Performance column (GE Healthcare, Little Chalfont, U.K.) operated according to standard procedures. Pooled fractions containing αGlc 1-*P* phosphatase or Had13 were loaded on a Fractogel EMD-DEAE column (Merck, Darmstadt, Germany) and purified according to standard protocol. Buffer exchange to 50 mM Mes, pH 7.0 (αGlc 1-*P* phosphatase) or 50 mM Hepes, pH 7.0 (Had13) was performed using Amicon Ultra-15 Centrifugal Filter Units (Millipore, Billerica, USA). *Note:* Unless mentioned, all further experiments were done in 50 mM Mes (αGlc 1-*P* phosphatase) or 50 mM Hepes supplemented with 5.0 mM $MgCl_2$ (Had13) buffer, each pH 7.0. Purification was monitored by SDS-PAGE; protein bands were visualized by silver staining.

1.3.3 CD spectroscopy.

Far-UV CD spectra of αGlc 1-*P* phosphatase (0.4 mg/mL) and Had13 (0.7 mg/mL) were recorded at 25 °C on a Jasco J715 spectro-polarimeter (JASCO Inst., Gross-Umstadt, Germany). CD spectra were collected in the wavelength range 190 – 260 nm at a scan speed of 100 nm/min using a bandwidth of 2.0 nm and a response time of 1.0 s. All spectra were recorded in a 0.01 cm cuvette. For each sample, 10 spectra were recorded, averaged, and buffer signal substracted. CD spectra were evaluated by using Dichroweb (54) with reference database No. 4.

1.3.4 Mass spectrometry.

For sample preparation, purified αGlc 1-*P* phosphatase (5 µg) was incubated with 20 mM iodoacetamide (IAA) solution for 30 min at 37°C and digested with modified trypsin (Promega, Madison, USA) (55), and/or with 0.5 µg chymotrypsin (Roche Applied Sciences, Penzberg, Germany) in 50 mM ammonium bicarbonate and 10 mM $CaCl_2$. For MS analysis, digests were acidified to 0.1% formic acid and separated by nano-HPLC (Dionex Ultimate 3000) equipped with a µ-precolumn (C18, 5 µm, 100 Å, 5 × 0.3 mm) and a Acclaim PepMap RSLC nanocolumn (C18, 2 µm, 100 Å, 150 × 0.075 mm) (all Thermo Fisher Scientific). About 0.5 µg of digested protein were injected and concentrated on the enrichment column for 2 min at a flow rate of 20 µL/min with 0.5 % trifluoroacetic acid as isocratic solvent. Separation was carried out on the nanocolumn at a flow rate of 300 nL/min using the following gradient, where solvent A is 0.05% trifluoroacetic acid in water and solvent B is 0.05% trifluoroacetic acid in 80% acetonitrile: 0-2 min: 4% B; 2-70 min: 4-28% B; 70-94 min: 28-50% B, 94-96 min: 50-95% B; 96-116 min: 95% B; 116-116.1 min: 95-4% B; 116.1-

140 min: re-equilibration at 4% B. The sample was ionized in the nanospray source equipped with stainless steel emitters (ES528, Thermo Fisher Scientific) and analyzed in a Orbitrap velos mass spectrometer (Thermo Fisher Scientific) operated in positive ion mode, applying alternating full scan MS (m/z 400 to 2000) in the ion cyclotron and MS/MS by higher-energy collisional dissociation of the 20 most intense peaks with dynamic exclusion enabled. The LC-MS/MS data were analyzed by searching a database containing the protein sequences of αGlc 1-*P* phosphatase and known background proteins with Mascot 2.3 (MatrixScience, London, UK). Search criteria were charge 2+ or 3+, precursor mass error 0.05 Da, and product mass error 0.7 Da, and carbamidomethylation, oxidation on methionine, -2H on cysteine (disulfide) as variable modifications. A maximum false discovery rate of 0.05 using decoy database search, an ion score cut off of 20 and a minimum of 2 identified peptides were chosen as protein identification criteria.

1.3.5 Assays.

Specific activity of αGlc 1-*P* phosphatase and Had13 was determined at 37 °C and pH 7.0 using 20 mM αGlc 1-*P* as substrate. Reaction was started by addition of αGlc 1-*P* phosphatase (0.1 µM) and Had13 (0.2 µM), and release of free phosphate was measured over 75 min. Inorganic phosphate was determined colorimetrically at 850 nm (56).

αGlc 1-*P* and Glc 6-*P* were measured in a coupled enzymatic system with phosphoglucomutase and Glc 6-*P* dehydrogenase (57). Fru 6-*P* was measured using an assay with mannitol 1-phosphate dehydrogenase (52). Fru 1-*P* was measured indirectly from the phosphate mass balance for substrate consumed (Δ[αGlc 1-*P*]) and products formed (equation 1), or directly using HPAEC-PAD (see later).

$$\Delta[\alpha\text{Glc 1-}P] = [\text{Fru 1-}P] + [\text{Glc 6-}P] + [\text{Fru 6-}P] + [\text{phosphate}] \quad (1)$$

The total protein concentration was measured using Roti-Quant assay referenced against BSA (58).

1.3.6 Inorganic phosphate-water medium ^{18}O exchange during hydrolysis of αGlc 1-*P*.

All reaction mixtures were prepared from $H_2{}^{18}O$ (97%) to give a final ^{18}O isotopic purity of the water solvent of 90%. Phosphatase reactions were done using 2.0 mM αGlc 1-*P* as the substrate. The standard Mes and Hepes buffers were used. Phosphorylase reaction was done in 50 mM Mes buffer, pH 7.0, using 1.0 mM αGlc 1-*P*. Reactions were started by adding enzyme (αGlc 1-*P* phosphatase: 0.1 µM, Had13: 1.0 µM, sucrose phosphorylase: 14 µM) to substrate solution at 37 °C. Incubations proceeded for 2 h using agitation at 650 rpm in an Eppendorf thermomixer. Reactions were stopped by heating (99°C, 5 min). Control reactions were carried out in exactly the same way using normal $H_2{}^{16}O$ instead of $H_2{}^{18}O$. Conversion of αGlc 1-*P* substrate was 90% or higher in each sample.

Incorporation of ^{18}O-label into phosphate was analyzed by GC-MS. Samples were dried in a SpeedVac for 3 h and derivatized with BSTFA/1%TMCS, in pyridine (1:2, by volume). Analysis was performed using a Trace DSQ Single Quadrupole GC–MS instrument (Thermo Scientific). The following GC parameters were used: injection volume, 1 µL; injector temperature, 250 °C; carrier gas, He; carrier gas flow, 1mL/min; column, HP-5MS (60 m, ID 0.250 mm, film thickness 0.25 µm) from Agilent (Waldbronn, Germany). The temperature gradient was as follows: starting temperature, 110 °C for 4 min; ramp to 300 °C at a heating rate of 20 °C/min; final hold time, 10 min. The MS was operated in EI mode (source

temperature: 280 °C) and the detected mass range was 50 - 700 m/z. The extracted ion chromatograms of m/z 299 and m/z 301 were integrated using the Xcalibur 1.4 software (Thermo Scientific).

1.3.7 Phosphatase kinetics.

Hydrolysis reactions were carried out at an initial substrate concentration of 20 mM. The different substrates used are summarized in Table 1. Reactions were started by addition of enzyme (αGlc 1-*P* phosphatase, 0.1 µM; Had13, 0.2 µM), and incubations continued for up to 150 min at 37°C and a thermomixer agitation rate of 650 rpm. Samples were taken in 15 min (αGlc 1-*P* phosphatase) or 30 min (Had13) intervals, heat-treated (99°C, 5 min), and centrifuged at 20,000 g for 5 min. The release of phosphate was measured, and the initial rate V (mM min^{-1}) was determined from the linear relationship between phosphate concentration and time. Apparent turnover frequencies (k_{cat_app}; s^{-1}) were calculated with equation 2, where [E] is the molar enzyme concentration, determined from the protein concentration assuming a molecular mass of 31 kDa for Had13 and 45 kDa for αGlc 1-*P* phosphatase.

$$k_{cat\_app} = V/[E] \quad (2)$$

1.3.8 Phosphoryl transfer studies.

Reactions were performed as described above (Phosphatase kinetics) using 20 mM αGlc 1-*P* as the donor substrate. Fru (100, 200, 400 or 600 mM) or Glc (200 mM) was used as the acceptor substrate. Samples taken at certain times were analyzed for phosphate, αGlc 1-*P*, and Glc 6-*P*. When Fru was the acceptor, Fru 1-*P* and Fru 6-*P* were additionally measured in each sample. Sampling was done to allow for determination of initial rates, but the full reaction

time course was also recorded. Relevant k_{cat_app} values for substrate consumption, phosphate release, and phosphoryl transfer were obtained from the corresponding V values using equation 2. Internal consistency of the data was always checked by mass balance.

1.3.9 pH dependence of phosphoryl transfer.

Reactions (0.1 µM enzyme) contained 20 mM of each αGlc 1-*P* and Fru, and k_{cat_app} was determined from V for consumption of αGlc 1-*P* or formation of Fru 1-*P*. The pH range analyzed was 4.0 – 8.0. A 50 mM Mes buffer was used, except at pH ≤ 6.0 where 50 mM Mes and 20 mM sodium acetate were used, and at pH 8.0 where 50 mM Mes and 50 mM Tes were used. Buffer pH values were adjusted at 37 °C and controlled before and after recording each enzyme-catalyzed reaction. Sampling and analyses were as described above.

For data analysis by non-linear least-squares regression, SigmaPlot 2004 version 9.0 was used. Equation 3 describes a pH dependence where activity (expressed as logarithmic k_{cat_app}) is constant at low pH and decreases above pK_a. C is the pH-independent value of k_{cat_app} at the optimum state of protonation, K is the proton dissociation constant and $[H^+]$ is the proton concentration.

$$\log(k_{cat\_app}) = \log(C/(1 + K/[H^+])) \quad (3)$$

1.3.10 High performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD).

Selected samples were analyzed on a Dionex BioLC system (Dionex Corporation, Sunnyvale, USA) equipped with a CarboPac PA10 column (4 × 250 mm) and an Amino Trap guard column (4 × 50 mm) thermostated at 30 °C. Glc, Fru, αGlc 1-*P*, Glc 6-*P*, Fru 1-*P* and Fru 6-*P* were detected with an ED50A electrochemical detector using a gold working electrode and a silver/silver chloride reference electrode, applying the predefined waveform for carbohydrates. Elution was carried out at a flow rate of 0.9 mL/min with the following method: isocratic flow of 52 mM NaOH for 20 min, followed by a linear gradient from 100 mM NaOAc to 400 mM NaOAc, applied within 25 min in an isocratic background of 100 mM NaOH. The column was washed for 5 min with 52 mM NaOH. Under the conditions applied, Glc eluted after 10.2 min, Fru after 11.8 min, αGlc 1-*P* after 30.3 min, Glc 6-*P* after 36.6 min, Fru 1-*P* after 37.2 min and Fru 6-*P* after 39.1 min.

1.3.11 NMR spectroscopic measurements.

For sample preparation, 20 mM αGlc 1-*P* and 200 mM Fru were incubated with 0.1 µM αGlc 1-*P* phosphatase at 37°C for 75 min. The reaction was stopped by heating (99°C, 5 min). After centrifugation at 20,000 *g* for 5 min, the supernatant was applied on a DEAE FF column (GE Healthcare, Little Chalfont, U.K.), pre-equilibrated with deionized water. Unbound, non-charged monosaccharides were removed by washing with deionized water. Elution of phosphorylated reaction products was accomplished by using 50 mM NaCl. Fractions containing the phosphorylated product were pooled and concentrated by lyophilisation, prior to NMR analysis.

For NMR analysis, the isolated compound was dissolved in D_2O (~5 mg in 0.7 mL) and transferred into 5 mm high precision NMR sample tubes. Measurements were performed on a Bruker DRX-400 at 400.13 MHz (^1H) using the Topspin 1.3 software. 1D ^1H NMR spectra were recorded by acquisition of 64 k data points and Fourier transformation resulting in spectra with a range of 14 ppm. To determine the 2D COSY, TOCSY, and NOESY spectra 128 experiments with 2048 data points each were recorded and Fourier transformed to 2D-spectra with a range of 10 ppm. Measurement temperature was 298 K +/- 0.05 K and external acetone was used as shift reference standard (δ_H 2.225).

1.3.12 Isolated compounds.

β-D-fructopyranose 1-phosphate - ^1H NMR (D_2O): δ = 3,99 (1H, m, H-6a), 3.93 (1H, m, H-3), 3.84 (1H, m, H-4*), 3,82 (1H, m, H-1b), 3,81 (1H, m, H-1a), 3.77 (1H, m, H-5*), 3.65 (1H, m, H-6b); ^{13}C NMR (D_2O): δ = 96.3 (s,C-2), 71.9 (d, C-3), 71.8 (d, C-4*), 70.5 (d, C-5*), 68.6 (t, C-6), 66.2 (t, C-1); ^{31}P NMR (D_2O): δ = 5.5 ($RO\text{-}POH_2^-$); other isomers regarding anomeric form and pyrano/furano form are present in negligible concentrations; [*] shifts are not unequivocally assignable.

1.3.13 Molecular docking studies.

AutoDock 4.2 as implemented in Yasara V 11.11.21 was used for enzyme-ligand docking. The AMBER03 force field and the default parameters provided by the standard docking macro were used, except that the number of runs was increased to 50. The structure of αGlc 1-*P* phosphatase ($His^{18}\rightarrow Ala$ mutant; pdb entry 1nt4) was used as macromolecule in molecular docking experiments that employed αGlc 1-*P* or αGlc 6-*P*, each as di- or mono-anion, as the

ligand. 3D coordinates of the ligands were generated from SMILES strings using Chimera (http://www.cgl.ucsf.edu/chimera). Docking was performed with Asp^{290} protonated or unprotonated. The ligand was placed flexible into the enyzme active site that was covered completely by the applied search space of 15 × 10 × 10 Å. Docking poses were evaluated by their associated free energy and mechanistic plausibility. PyMOL (http://www.pymol.org) was used for visualization.

1.4 Results

1.4.1 Enzyme preparation and characterization.

Homologous overexpression of the *agp* gene in *E. coli* results in accumulation of recombinant enzyme in the periplasmic space whereby the 22 amino acid-long *N*-terminal targeting sequence appears to be processed off correctly (59, 60). We considered the possibility of keeping the recombinant αGlc 1-*P* phosphatase in the *E. coli* cytoplasm, but the presence of disulfide bonds in the protein structure presented critical conjuncture for protein production under these conditions. *E. coli* Origami B was therefore selected as expression host, and the endogenous disulfide bond isomerase (DsbC) was co-overexpressed from the plasmid vector used (61). The final αGlc 1-*P* phosphatase construct had the targeting sequence deleted and an *N*-terminal *Strep*-tag added for facilitated recovery. Recombinant enzyme was isolated from induced cells in a yield of about 1 mg/L culture, which is significantly lower than reported protein yield from periplasmic expression (40 mg/L) (60). Cytoplasmic production of αGlc 1-*P* phosphatase was useful nonetheless, because absence of *N*-terminal processing in the cytoplasm offered flexibility in modifying the *N*-terminus with a functional tag.

Had13 was obtained as *N*-terminally *His*-tagged protein (46) in a yield of about 500 µg protein/L cell culture. SDS-PAGE showed high purity of both enzymes (Figure A1), and it also confirmed that αGlc 1-*P* phosphatase and Had13 had been isolated in the expected molecular size of 45 and 31 kDa, respectively. Purified proteins were used in all experiments to be reported later.

Phosphatase activity was determined in standard assays measuring release of phosphate from αGlc 1-*P* (20 mM). The specific activity of αGlc 1-*P* phosphatase was 58 U/mg; that of Had13 was 24 U/mg, measured in the presence of 5.0 mM $MgCl_2$. In the absence of Mg^{2+}, the specific activity of Had13 was lowered by about 150-fold.

Far-UV circular dichroism (CD) spectra from solutions of αGlc 1-*P* phosphatase and Had13 are shown in Figure A2. Estimates of the relative content of secondary structural elements in each protein are summarized in Table A1 where they are compared to data derived from the experimental protein structures. There is variation in the mixed α-β compositions determined from CD spectrum and crystal structure; however, the results agree in showing that α-helices were present in slight excess over β-strands and that a substantial portion of the structure was classified as unordered. While it is impractical to conclude correctness of folding from the data presented, there is no hint at major protein misfolding in the recombinant enzyme preparations used.

Liquid chromatography-tandem mass spectrometry was applied to the determination of disulfide bond(s) in the as-isolated preparation of recombinant αGlc 1-*P* phosphatase. Protein digests with trypsin or chymotrypsin were resolved to 62% overall sequence coverage, 214 peptides, 37 unique peptides, and a Mascot score of 6046 (Figure A3). Identity of one of the

three native disulfide bonds (Figure 1A) between Cys^{189} and Cys^{195} was confirmed unambiguously from the peptide KDSPA**C**KEKQQ**C**SLVDGKNTF (ion score: 34; expect 0.00042). Figure A4 shows the corresponding mass spectrum and ion table. The two remaining disulfide bridges could not be detected directly, probably due to unsuitable cleavage sites producing too large crosslinked peptides outside of the detectable mass range. However, the results are clear in showing that disulfide bond formation has been possible in principle under the expression conditions used.

1.4.2 Mechanistic characterization of αGlc 1-*P* hydrolysis using ^{18}O labeling.

The proposed catalytic reactions of αGlc 1-*P* phosphatase and Had13 (Scheme 2) share the mechanistic features that the O1–P bond of substrate is cleaved, and that water (from solvent) attacks a phospho-enzyme intermediate in the dephosphorylation step of the overall hydrolytic conversion of αGlc 1-*P* (Scheme 2). Contrary to uncatalyzed and glycosidase-like catalyzed hydrolysis of αGlc 1-*P* where the C1–O bond is broken and water is incorporated to the anomeric glucosyl carbon, the phosphatase reaction is expected to result in water addition to phosphorus. Measurement of ^{18}O label incorporation from an isotopically enriched water solvent is therefore useful to distinguish between the two mechanistic possibilities. Enzymatic conversions of 2.0 mM αGlc 1-*P* were carried out in $H_2{}^{18}O$ (90% ^{18}O content in the final reaction mixture) and in normal water (control). αGlc 1-*P* was depleted fully in each reaction. GC-MS analysis was used to determine the isotopic composition of the phosphate released. A tris(trimethylsilyl) phosphate (TMS-phosphate) species was analyzed. Under electron ionization conditions as described in literature (62) and confirmed in our experiments, the TMS-phosphate molecular ion at *m/z* 314 fragments due to loss of one methyl group and

major ion at *m/z* 299 is formed. Results are summarized in Figure 2 in a form where the peak area ratio at *m/z* 301 (TMS-phosphate containing one atom of ^{18}O next to three ^{16}O atoms) and *m/z* 299 (unlabeled TMS-phosphate containing only ^{16}O atoms) is plotted. Full analytical details along with the experimental GC-MS chromatograms are provided in Figure A5. Reactions in normal water gave a peak area ratio of 0.13, which is consistent with the natural $^{18}O/^{16}O$ isotope ratio in phosphate calculated with IsoPro 3.1. Reactions in ^{18}O-enriched solvent gave much larger peak area ratios (up to 65-fold increase) than the corresponding controls, clearly indicating ^{18}O label incorporation from solvent to phosphate under these conditions. We also performed a mechanistic control where sucrose phosphorylase was incubated with αGlc 1-*P* in $H_2{}^{18}O$ solvent. αGlc 1-*P* was hydrolyzed by the phosphorylase without detectable labeling of the phosphate released (Figure 2), as expected from the catalytic mechanism. Therefore, these results strongly support the notion that enzymatic hydrolysis of αGlc 1-*P* by αGlc 1-*P* phosphatase or Had13 proceeds via the canonical phosphomonoester hydrolase mechanism where bond cleavage and formation occurs at phosphorus, as shown in Scheme 2.

1.4.3 Phospho-donor substrate utilization by αGlc 1-*P* phosphatase and Had13 in hydrolysis.

A series of phosphorylated sugars (Table 1) were tested as substrates for hydrolysis by each phosphatase. Apparent turnover frequencies (k_{cat_app}) determined at 20 mM substrate concentration are summarized in Table 1. Both enzymes exhibited high selectivity for hydrolyzing αGlc 1-*P* as compared to βGlc 1-*P*, which was completely inactive towards αGlc 1-*P* phosphatase and hardly converted by Had13. Phospho-sugars harboring the phosphomonoester group at a primary hydroxyl (Glc 6-*P*; Fru 6-*P*; D-fructose 1-phosphate,

Fru 1-*P*) were utilized by both enzymes with high activity, the respective k_{cat_app} being comparable within a 2-fold range to the k_{cat_app} for conversion of αGlc 1-*P*. With all sugar phosphates tested, αGlc 1-*P* phosphatase was the faster phosphatase than Had13. Pyrophosphate and phytate (*myo*-inositol hexakisphosphate) were also examined as enzyme substrates, and reactions were performed at pH 7.0 and pH 4.5 (50 mM sodium acetate). Pyrophosphate was not hydrolyzed which is interesting because it is an excellent substrate for many acid phosphatases (39, 63). αGlc 1-*P* phosphatase was weakly active with phytate, but only at the low pH of 4.5 where the phytate is no longer strongly complexed with metal ions (64).

1.4.4 Enzymatic transphosphorylation: αGlc 1-*P* phosphatase catalyzes efficient phosphoryl transfer from αGlc 1-*P* to the 6-hydroxyl of an external Glc acceptor.

During time-course studies of hydrolysis of αGlc 1-*P* (20 mM) by αGlc 1-*P* phosphatase we noticed that in a late phase of the reaction (≥ 70% substrate conversion) the molar concentration of phosphate released did not match up exactly (≥ 10%) to the αGlc 1-*P* consumed. Analysis of the reaction mixture by HP anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) revealed the presence of a new phosphorylated sugar, which was found to co-elute with an authentic Glc 6-*P* standard. The presence of Glc 6-*P* in the reaction sample was confirmed unambiguously using an enzymatic assay based on selective NAD^+ dependent oxidation of Glc 6-*P* by Glc 6-*P* dehydrogenase. We also determined that the amount of Glc 6-*P* formed accounted precisely for the phosphate missing in the balance with αGlc 1-*P* converted. Interestingly, samples from Had13 reaction carried out under otherwise identical conditions did not contain Glc 6-*P*, and there was close balance

between phosphate formation and αGlc 1-*P* consumption. The synthesis of Glc 6-*P* during conversion of αGlc 1-*P* was suggested to have resulted from an enzymatic transphosphorylation reaction in which αGlc 1-*P* was the donor, and the Glc formed in prior hydrolysis was the acceptor. Mechanistically, the substrate's phosphoryl group would be transferred to the catalytic His^{18} of the enzyme, and dephosphorylation of αGlc 1-*P* phosphatase could then occur by reaction with water (*hydrolysis*) or Glc (*phosphoryl transfer*). Because phosphoryl transfer would take place in competition with hydrolysis, the Glc concentration was expected to be decisive for efficient utilization of αGlc 1-*P* for synthesis of Glc 6-*P*.

We therefore repeated enzymatic conversion of αGlc 1-*P* (20 mM), but now added 200 mM Glc as external phosphoryl acceptor at reaction start. Figure A6 shows Glc 6-*P* production in relation to αGlc 1-*P* consumption along time courses of reaction catalyzed by αGlc 1-*P* phosphatase. Formation of free phosphate is also shown. Using αGlc 1-*P* phosphatase, Glc 6-*P* was formed in large amounts, accounting for nearly all of the αGlc 1-*P* cleaved in the enzymatic reaction. Compared to the control reaction lacking Glc, the phosphate release rate was suppressed to a large extent (~10-fold). The k_{cat_app} of phosphate formation was just 4 s^{-1} under these conditions. Data in Figure A6 were furthermore used to calculate k_{cat_app} values of 34 s^{-1} and 40 s^{-1} for Glc 6-*P* synthesis and αGlc 1-*P* consumption, respectively. Apparent turnover frequencies for the overall conversion were therefore similar in the absence and presence of Glc. However, it must be emphasized that the observable k_{cat_app} (αGlc 1-*P*) in the presence of Glc sets only a lower limit to the true turnover frequency of the enzyme under the transphosphorylation conditions used because there is also the possibility of a futile re-synthesis of αGlc 1-*P* that cannot be detected with the applied analytical methods. Conversion

of αGlc 1-*P* into Glc 6-*P* is formally equivalent to the phosphoglucomutase reaction, which however involves positional rearrangement of the phosphomonoester group *within the same glucose molecule* between O6 and α-O1 (7). Equilibrium of the phosphoglucomutase reaction strongly favors Glc 6-*P* (65). Synthesis of Glc 6-*P* by enzymatic transphosphorylation proceeded through a kinetic optimum of about 10 mM Glc 6-*P* (Figure A6). At longer reaction times, Glc 6-*P* was hydrolyzed completely.

Contrary to αGlc 1-*P* phosphatase, Had13 produced only minute amounts of Glc 6-*P* when Glc was added to the reaction with αGlc 1-*P*. The k_{cat_app} for Glc 6-*P* synthesis was determined as 0.3 s^{-1}, which is 22-fold lower than the hydrolysis k_{cat_app} under these conditions. Further studies of enzymatic transphosphorylation therefore focused mainly on αGlc 1-*P* phosphatase.

1.4.5 Synthesis of Fru 1-*P* by enzymatic transphosphorylation from Glc 1-*P*.

When Fru (200 mM) was added to conversions of αGlc 1-*P* (20 mM) by αGlc 1-*P* phosphatase, there was the interesting effect that compared to the control reaction performed in the absence of Fru the αGlc 1-*P* consumption rate was enhanced about 3-fold (k_{cat_app} = 137 s^{-1}) whereas the phosphate release rate was decreased about 2-fold at the same time (k_{cat_app} = 18 s^{-1}). Therefore, this provided clear indication that phosphoryl transfer from αGlc 1-*P* to an acceptor other than water, likely Fru, constituted the major route of phosphoryl donor substrate utilization under the conditions used. Using analysis by HPAEC-PAD, formation of new sugar phosphate products was confirmed (see later). Comparison of k_{cat_app} under hydrolysis-only and phosphoryl transfer conditions revealed that αGlc 1-*P* phosphatase was somehow "activated" in the presence of external acceptor. In the proposed scenario of

enzymatic transphosphorylation where an otherwise hydrolyzed phospho-enzyme intermediate is intercepted by acceptor, enhancement of k_{cat_app} in the presence of Fru is possible when enzyme dephosphorylation is a slow step of the overall hydrolysis process, and is accelerated by the reaction with Fru. Consistent with this notion, we showed that k_{cat_app} (αGlc 1-*P*) increased about 5-fold in hyperbolic dependence of the Fru concentration (100 – 600 mM; Table 2), approaching a calculated maximum value of 254 s^{-1} at saturating concentrations of the phosphoryl acceptor (Figure A7). An apparent K_M for Fru of 209 mM was determined from the data. Table 2 summarizes results of comprehensive kinetic analysis of enzymatic transphosphorylations at different Fru acceptor concentrations. Expressed in apparent first-order rate constants, the phosphate release rate was suppressed strongly in the presence of acceptor, to less than one-tenth the αGlc 1-*P* consumption rate under conditions of 600 mM Fru. Formation of Glc 6-*P* occurred at a very low rate. Synthesis of Fru 6-*P* took place at about one-fourth the rate of Fru 1-*P* formation, irrespective of the Fru concentration used. We plotted in Figure 3A the ratio of the total transphosphorylation rate (Fru 1-*P*, Fru 6-*P*, Glc 6-*P*) and the αGlc 1-*P* consumption rate and show its dependence of the Fru concentration. The ratio approached a value of unity at high Fru, consistent with the mechanistic notion that externally added (Fru) and *in situ* formed (Glc) acceptors compete effectively with water for reaction with the phospho-enzyme intermediate. Various phosphatases related to αGlc 1-*P* phosphatase by common membership to the histidine acid phosphatase family of proteins exhibit their optimum activity in the acidic pH range (≤ 5.0) (66, 67). We therefore determined pH dependencies of k_{cat_app} for αGlc 1-*P* consumption and Fru 1-*P* synthesis by αGlc 1-*P* phosphatase in the pH range 4.0 – 8.0. Activity was constant in the pH range 4.0 – 6.0 and gradually decreased at higher pH. Activity at the standard pH of 7.0 was still 88% of the activity at optimum pH. Interestingly, the pH-rate profiles for αGlc 1-*P* consumption and

Fru 1-*P* synthesis were identical (Figure 3B), indicating that partitioning of the phospho-enzyme intermediate between reaction with Fru and reaction with water is not dependent on pH.

A complete time course of conversion of αGlc 1-*P* in the presence of 200 mM Fru is shown in Figure 3C. From the phosphate concentration present at the time when αGlc 1-*P* was completely exhausted we determined the total concentration of transphosphorylation product(s) as 16 mM, corresponding to a yield of 80% based on donor substrate utilized. Using HPAEC-PAD analysis referenced against authentic standards, we showed that Fru 1-*P* was the main product (12 mM), and Fru 6-*P* (2.7 mM) and Glc 6-*P* (1.0 mM) were formed as by-products. However, Glc 6-*P* appeared mainly in the late phase of the reaction when substantial concentrations of Glc (≥ 10 mM) had already accumulated. The molar ratio of Fru 1-*P* to Glc 6-*P* decreased during the reaction from an initial value of 20 to about 10 at conversion end (Figure A7). Phosphorylation of Glc occurred despite Fru being present in about 20-fold molar excess, reflecting the acceptor substrate selectivity of αGlc 1-*P* phosphatase. We also noticed that in spite of the pronounced hydrolase activity of αGlc 1-*P* phosphatase towards each sugar phosphate synthesized in the reaction (Table 1), no loss of transphosphorylation product to secondary hydrolysis occurred in the timespan of the experiment. Unlike other phosphatase-catalyzed transphosphorylation reactions in which product kinetic stability presented a problem to synthetic application of the biocatalytic transformations (37, 38), sugar phosphates were formed from αGlc 1-*P* as if they were real equilibrium products (Figure 3C), thus enabling their convenient production. Figure 3D shows the composition of phosphorylated products obtained from conversion of αGlc 1-*P* at different Fru concentrations.

Identity of Fru 1-*P* was further confirmed through 1D and 2D NMR analysis of the product mixture. Fru 1-*P* was present in four different anomeric forms, whereby the β-D-fructopyranose 1-phosphate had highest abundance (~80 %). For this anomer, the proton signals of H-1a&b resonated as separated signal group in an ABX at 3.82 ppm and 3.81 ppm with a $^{3}J_{H\text{-}P}$ heteronuclear coupling to the phosphate group. All other proton signals belonged to spin system, as detected in the TOCSY spectrum. The proton signals of the methylen group in position 6 were present at 3.99 and 3.65 ppm, both showing $^{3}J_{H\text{-}H}$ couplings to H-5, which resonated at 3.77 ppm. Further coupling of H-5 to to H-4 (3.84 ppm) and consecutively of H-4 to H-3 (3.95) were also visible in COSY spectra. The large shift difference between signals of H-6a and H-6b indicated the anomeric form. Furthermore, NMR data of the enzymatically synthesized Fru 1-*P* were in exact agreement with reference spectra recorded from commercial Fru 1-*P*. The NMR data combined with evidence from HPAEC-PAD analysis and enzymatic assays confirmed Fru 6-*P* and Glc 6-*P* to be transphosphorylation products next to Fru 1-*P*.

Contrary to αGlc 1-*P* phosphatase, Had13 did not catalyze phosphorylation of Fru. Glc 1-*P* was hydrolyzed completely by the enzyme, and the rate of donor substrate consumption was the same in the presence and absence of 200 mM Fru.

1.4.6 Binding of αGlc 1-*P* and Glc 6-*P* by αGlc 1-*P* phosphatase analyzed by molecular docking.

The x-ray crystal structure of αGlc 1-*P* phosphatase having the catalytic His^{18} substituted by Ala contains a bound ligand reported to be αGlc 1-*P* (45). However, the phospho-sugar represented by the atomic coordinates of the enzyme-ligand complex is clearly βGlc 1-*P*, not

αGlc 1-*P*. An experimental electron density map was not released with the protein structure and cannot therefore be used for clarification. We performed docking studies in which αGlc 1-*P* was placed flexible into the substrate binding pocket of the enzyme (pdb entry 1nt4). Binding of singly or doubly negatively charged αGlc 1-*P* was examined, and the putative catalytic acid-base Asp^{290} was analyzed protonated or unprotonated. The best-fit docking poses received under the different conditions were independent of the charge state of αGlc 1-*P* and Asp^{290}. Figure 4A depicts the best-fit binding mode of singly charged αGlc 1-*P* to αGlc 1-*P* phosphatase having Asp^{290} protonated. αGlc 1-*P* is accommodated in an elongated conformation with the phosphoryl group placed *anti* relative to the glucosyl moiety. The substrate's glycosidic oxygen O1 is brought into a position that would allow protonated Asp^{290} to provide Brønsted catalytic assistance to O1–P bond fission and thus to the departure of the Glc leaving group. Figure 4B shows that Asp^{290} could adopt an analogous catalytic role during hydrolysis of βGlc 1-*P*, even though the distance between the catalytic enzyme and the reactive substrate groups is substantially larger (3.4 Å) than in the αGlc 1-*P* docking pose (2.9 Å; Figure 4A). We note that βGlc 1-*P* is not a substrate of αGlc 1-*P* phosphatase (Table 1), and the binding mode of βGlc 1-*P* in the protein crystal structure may actually be nonproductive for catalysis to O1–P bond cleavage (Figure 4B). Evidence from the docking analysis indicates that substrate binding recognition by αGlc 1-*P* phosphatase is mainly through numerous strong interactions with the phosphoryl group (Figure 4A) whereas the glucosyl moiety of αGlc 1-*P* is bound only weakly by comparison. The docking experiments resulted in up to 6 docking poses of similar calculated free energies. The different docking poses had the phosphoryl group bound identically, but featured slight variation in the alignment of the glucosyl moiety (Figure A8). Size and topology of the binding pocket appear

adequate for accommodation of different sugar structures in multiple orientations, as shown for binding of Glc 6-*P* (Figure 4C) compared to binding of αGlc 1-*P* (Figure 4A), thus explaining the relaxed donor and acceptor substrate specificity of the enzyme. It is not practical to examine transphosphorylation compared to hydrolysis reaction selectivity using molecular docking. Therefore, ligand docking to the Had13 structure was left for consideration in the future.

1.5 Discussion

First-time synthesis of sugar phosphates by enzymatic transphosphorylation from αGlc 1-*P* is reported, and *E. coli* αGlc 1-*P* phosphatase has been identified as efficient catalyst of the reaction. Fru 1-*P* and Glc 6-*P* were obtained in good yield (≥ 55%) based on the limiting phosphoryl donor substrate utilized in a one-pot single-step biotransformation. Cytoplasmic expression of αGlc 1-*P* phosphatase in *E. coli* Origami B enabled recombinant production of functional enzyme. "Re-localization" of the naturally periplasmic enzyme into the cytosol circumvented the need to adapt *N*-terminal fusion strategies (e.g. attachment of *Strep*-tag) to protein trimming events in the periplasm.

1.5.1 Molecular and kinetic properties of αGlc 1-*P* phosphatase underlying the enzyme's remarkable transphosphorylation activity.

Both αGlc 1-*P* phosphatase and Had13 catalyzed hydrolysis of different sugar phosphates with relaxed specificity for the structure of the leaving group (Glc, Fru) and for the position of the phosphoryl group on the sugar moiety (αGlc 1-*P*, Glc 6-*P*). In a comparison of the α and β-anomers of Glc 1-*P*, however, both enzymes discriminated strongly against hydrolysis of

βGlc 1-*P*. Docking analysis carried out on the x-ray crystal structure of αGlc 1-*P* phosphatase suggested a molecular basis for the anomeric selectivity of the enzyme (Figure 4, Figure A8). It also provided interpretation of the broad donor and acceptor substrate specificity of αGlc 1-*P* phosphatase.

Incorporation of ^{18}O label from $H_2{}^{18}O$ solvent into the phosphate released during hydrolysis of αGlc 1-*P* by αGlc 1-*P* phosphatase and Had13 supports enzymatic reactions through nucleophilic substitution at the phosphorus. According to mechanistic proposals for histidine acid phosphatases and Had-like phosphatases (Scheme 2), the overall conversion of αGlc 1-*P* is expected to proceed in a two-step double displacement-like catalytic process via a covalent phospho-histidine and phospho-aspartate enzyme intermediate, respectively. Phosphoryl transfer to sugar acceptors occurs from the phosphorylated enzyme, probably after dissociation of the Glc, in direct competition to hydrolysis. Efficient interception of the phospho-enzyme intermediate to prevent its reaction with water will therefore be key characteristic of synthetically useful acceptor substrates. Partitioning of the phosphorylated enzyme to catalytic breakdown via phosphoryl transfer and hydrolysis is adjustable by the acceptor substrate concentration, as shown for the phosphorylation of Fru by αGlc 1-*P* phosphatase (Figure 3A). Using the data from conversion of αGlc 1-*P* in the presence of 200 mM Glc, we can apply the experimental k_{cat_app} ratio for formation of Glc 6-*P* and phosphate to calculate that αGlc 1-*P* phosphatase (k_{cat_app} ratio = 34/4.0 = 8.5) surpassed Had13 (k_{cat_app} ratio = 0.3/7.0 = 0.04) by more than 200-fold in partition efficiency to favor phosphoryl transfer to the 6-OH of Glc. Differences between the two enzymes are even larger when phosphoryl transfer to Fru is considered. With current bioinformatic methodology it is not possible to infer the catalytic ability of transphosphorylation from a phosphatase protein

structure alone, therefore necessitating the dedicated biochemical characterization. Acceptor substrate selectivity and phosphorylation site selectivity of transphosphorylation will be fine-tuned by positioning of the sugar acceptor substrate in the binding pocket on the phospho-enzyme intermediate. αGlc 1-*P* phosphatase accommodated Fru in two distinct binding modes, where the preferred binding mode resulted in formation of Fru 1-*P*, and the other resulted in formation of Fru 6-*P*.

1.5.2 Synthetic use of transphosphorylation from αGlc 1-*P*.

Fru 1-*P* is a high-value metabolite that is difficult to synthesize chemically (68). Biocatalytic synthetic routes to Fru 1-*P* are aldolase (EC 4.1.2.13)-catalyzed carbon-carbon coupling between glyceraldehyde and dihydroxyacetone phosphate (69), or 1-phosphorylation of Fru from ATP by ketohexokinase (EC 2.7.1.3) (62, 63). Transphosphorylation from αGlc 1-*P* presents an interesting alternative that offers the advantage of a convenient and cheap phosphoryl donor substrate compared with ATP or dihydroxyacetone phosphate. Moreover, αGlc 1-*P* phosphatase showed excellent turnover frequency (≥ 100 s^{-1}; Table 2) for transphosphorylation between αGlc 1-*P* and Fru. By employing Fru at 600 mM, utilization of αGlc 1-*P* as donor substrate for transphosphorylation was almost complete in a sense that substrate hydrolysis was prevented effectively (≤ 10%; see Figure A7). Based on αGlc 1-*P* converted, Fru 1-*P* was obtained in 70% yield (14 mM). Remainder products were Fru 6-*P* (4 mM) and Glc 6-*P* (0.5 mM). It would be interesting to increase the initial αGlc 1-*P* concentration in a next step of optimization of the biocatalytic phosphorylation of Fru for synthesis of Fru 1-*P*. We clearly recognize the problem of formation of other phosphorylated sugars next to Fru 1-*P*. However, efficient methodology for chromatographic separation of phospho-sugars has been developed (70, 71) so that capture of the Fru 1-*P* from product

mixture with Fru 6-*P* and Glc 6-*P* is certainly possible. Alternatively, selective hydrolysis of the sugar 6-phosphates using a primarily hydrolytic phosphatase (72) might be useful.

Following the early discovery of phosphoryltransferase activity in certain phosphatases (36), recent work has demonstrated the synthetic usefulness of biocatalytic transphosphorylation reactions in the preparation of different sugar phosphates. Bacterial phosphatases from the class-A nonspecific acid phosphatase family were mostly used, and *Shigella flexneri* (37, 73), *Salmonella enterica* (73) and *Morgenella morganii* (74) were prominent sources of the enzyme. A key paper of Wever and colleagues showed phosphorylation of a series of aldohexose, aldopentose, and ketohexose acceptors from a pyrophosphate donor substrate using acid phosphatase from *S. flexneri* (37). Using donor and acceptor at 100 mM each, the yield of phosphorylated product was typically 15% or lower, except for the 6-phosphates of Glc and 5-thio-D-glucose that were obtained in substantially higher yields. Interestingly, Fru was a rather poor acceptor substrate for the *Shigella* phosphatase (5% phosphorylation yield), suggesting useful complementarity in the acceptor substrate specificities of αGlc 1-*P* phosphatase and the previously used transphosphorylation catalyst. Moreover, whereas the nonspecific acid phosphatases were used at pH 4.0, αGlc 1-*P* phosphatase can also be applied in the neutral pH range. Using aldohexose acceptors (e.g. Glc, D-mannose, D-galactose), it was shown in van Herk et al. (37) that the maximum concentration of phosphorylated product obtained from 100 mM pyrophosphate was strongly dependent on the acceptor concentration used; and that it went through a distinct kinetic optimum due to the pronounced effect of secondary hydrolysis at extended reaction times. Glc 6-*P* was an exception and it was obtained as a kinetically stable transphosphorylation product. Problems with secondary hydrolysis were not encountered during αGlc 1-*P* phosphatase-catalyzed phosphorylations of

Fru. However, Glc 6-*P* was hydrolyzed. Finally, even though pyrophosphate is generally viewed as a highly expedient phosphoryl donor substrate for enzymatic transphosphorylations, it must not be overlooked that even in the absence of hydrolysis, the pyrophosphate conversion results in the release of inorganic phosphate, which can inhibit the phosphatase (40). The use of αGlc 1-*P* in combination with a dedicated αGlc 1-*P* phosphatase might help minimizing the effect of product inhibition by phosphate in phosphatase-catalyzed transphosphorylation reactions. The enzyme from *E. coli* αGlc 1-*P* phosphatase is a powerful biocatalyst for this type of transformation.

1.6 Acknowledgements

We acknowledge financial support from the Austrian Science Funds FWF (project L586-B03); Alexander Yakunin (Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada) for providing p15TvL_yida; Harald Pichler (Institute of Molecular Biotechnology, Graz University of Technology, Austria) for providing pMS470_dsbC; Tom Desmet (Centre for Industrial Biotechnology and Biocatalysis, Ghent University, Belgium) for synthesis of βGlc 1-*P*; and Christian Fercher (Institute for Molecular Biosciences, Karl-Franzens University Graz, Austria) for help in recording CD spectra.

1.7 References

1. **Dzeja PP, Terzic A**. 2003. Phosphotransfer networks and cellular energetics. J. Exp. Biol. **206**:2039–2047.
2. **Berg JM, Tymoczko JL, Stryer L**. 2002. Biochemistry, 5th ed. W H Freeman, New York, USA.
3. **Wünschiers R**. 2012. Carbohydrate metabolism and citrate cycle, p. 37–58. *In* Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, 2nd ed. John Wiley & Sons, Inc., Hoboken, USA.
4. **Knowles JR**. 1980. Enzyme-catalyzed phosphoryl transfer reactions. Annu. Rev. Biochem. **49**:877–919.
5. **Lassila JK, Zalatan JG, Herschlag D**. 2011. Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. Annu. Rev. Biochem. **80**:669–702.
6. **Frey PA, Hedgeman AD**. Phosphotransfer and nucleotidyltransfer, p. 476–546. *In* Enzymatic reaction mechanisms. Oxford University Press, Oxford, UK.
7. **Allen KN, Dunaway-Mariano D**. 2004. Phosphoryl group transfer: evolution of a catalytic scaffold. Trends Biochem. Sci. **29**:495–503.
8. **Granot D, Kelly G, Stein O, David-Schwartz R**. 2014. Substantial roles of hexokinase and fructokinase in the effects of sugars on plant physiology and development. J. Exp. Bot. **65**:809–819.
9. **Agius L**. 2008. Glucokinase and molecular aspects of liver glycogen metabolism. Biochem. J. **414**:1–18.
10. **Kawai S, Mukai T, Mori S, Mikami B, Murata K**. 2005. Hypothesis: structures, evolution, and ancestor of glucose kinases in the hexokinase family. J. Biosci. Bioeng. **99**:320–330.

11. **Wilson JE**. 2003. Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J. Exp. Biol. **206**:2049–2057.
12. **Kornberg HL**. 2001. Routes for fructose utilization by *Escherichia coli*. J. Mol. Microbiol. Biotechnol. **3**:355–359.
13. **Luley-Goedl C, Nidetzky B**. 2010. Carbohydrate synthesis by disaccharide phosphorylases: reactions, catalytic mechanisms and application in the glycosciences. Biotechnol. J. **5**:1324–1338.
14. **Nakai H, Kitaoka M, Svensson B, Ohtsubo K**. 2013. Recent development of phosphorylases possessing large potential for oligosaccharide synthesis. Curr. Opin. Chem. Biol. **17**:301–309.
15. **Westheimer FH**. 1987. Why nature chose phosphates. Science **235**:1173–1178.
16. **Kamerlin SCL, Sharma PK, Prasad RB, Warshel A**. 2013. Why nature really chose phosphate. Q. Rev. Biophys. **46**:1–132.
17. **Morrow JR, Amyes TL, Richard JP**. 2008. Phosphate binding energy and catalysis by small and large molecules. Acc. Chem. Res. **41**:539–548.
18. **Nagano N, Orengo CA, Thornton JM**. 2002. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. J. Mol. Biol. **321**:741–765.
19. **Klimacek M, Krahulec S, Sauer U, Nidetzky B**. 2010. Limitations in xylose-fermenting *Saccharomyces cerevisiae*, made evident through comprehensive metabolite profiling and thermodynamic analysis. Appl. Environ. Microbiol. **76**:7566–7574.
20. **Zhang G-F, Sadhukhan S, Tochtrop GP, Brunengraber H**. 2011. Metabolomics, pathway regulation, and pathway discovery. J. Biol. Chem. **286**:23631–23635.

21. **Büscher JM, Czernik D, Ewald JC, Sauer U, Zamboni N**. 2009. Cross-platform comparison of methods for quantitative metabolomics of primary metabolism. Anal. Chem. **81**:2135–2143.

22. **McCloskey D, Palsson BØ, Feist AM**. 2013. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. Mol. Syst. Biol. **9**:661.

23. **Auriol D, Lefevre F, Nalin R, Redziniak G**. April 2011. Cosmetic and pharmaceutical composition comprising *N*-acetylglucosamine-6-phosphate. US 20130012475.

24. **Auriol D, Nalin R, Lefevre F, Ginolhac A, Guembecker DD, Zago C**. November 2008. Method for preparing C-6 phosphorylated D-aldohexoses and C-6 phosphorylated D-aldohexose derivatives. EP 2150620.

25. **Shin W-J, Kim B-Y, Bang W-G**. 2007. Optimization of ascorbic acid-2-phosphate production from ascorbic acid using resting cell of *Brevundimonas diminuta*. J. Microbiol. Biotechnol. **17**:769–773.

26. **Goedl C, Schwarz A, Minani A, Nidetzky B**. 2007. Recombinant sucrose phosphorylase from *Leuconostoc mesenteroides*: characterization, kinetic studies of transglucosylation, and application of immobilised enzyme for production of α-D-glucose 1-phosphate. J. Biotechnol. **129**:77–86.

27. **Van der Borght J, Desmet T, Soetaert W**. 2010. Enzymatic production of β-D-glucose-1-phosphate from trehalose. J. Biotechnol. **5**:986–993.

28. **De Groeve MRM, De Baere M, Hoflack L, Desmet T, Vandamme EJ, Soetaert W**. 2009. Creating lactose phosphorylase enzymes by directed evolution of cellobiose phosphorylase. Protein Eng. Des. Sel. **22**:393–399.

29. **Lange CF, Kohn P**. 1961. Substrate specificity of hexokinases. J. Biol. Chem. **236**:1–5.

30. **Nishimasu H, Fushinobu S, Shoun H, Wakagi T**. 2006. Identification and characterization of an ATP-dependent hexokinase with broad substrate specificity from the hyperthermophilic archaeon *Sulfolobus tokodaii*. J. Bacteriol. **188**:2014–2019.
31. **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*. Appl. Environ. Microbiol. **73**:6444–6449.
32. **Cai L, Guan W, Kitaoka M, Shen J, Xia C, Chen W, Wang PG**. 2009. A chemoenzymatic route to *N*-acetylglucosamine-1-phosphate analogues: substrate specificity investigations of *N*-acetylhexosamine 1-kinase. Chem. Commun. 2944–2946.
33. **Chen M, Chen L, Zou Y, Xue M, Liang M, Jin L, Guan W, Shen J, Wang W, Wang L, Liu J, Wang PG**. 2011. Wide sugar substrate specificity of galactokinase from *Streptococcus pneumoniae* TIGR4. Carbohydr. Res. **346**:2421–2425.
34. **Zhao H, van der Donk WA**. 2003. Regeneration of cofactors for use in biocatalysis. Curr. Opin. Biotechnol. **14**:583–589.
35. **Berke W, Schüz HJ, Wandrey C, Morr M, Denda G, Kula MR**. 1988. Continuous regeneration of ATP in enzyme membrane reactor for enzymatic syntheses. Biotechnol. Bioeng. **32**:130–139.
36. **Wilson IB, Dayan J, Cyr K**. 1964. Some properties of alkaline phosphatase from *Escherichia coli*: transphosphorylation. J. Biol. Chem. **239**:4182–4185.
37. **Van Herk T, Hartog AF, van der Burg AM, Wever R**. 2005. Regioselective phosphorylation of carbohydrates and various alcohols by bacterial acid phosphatases; probing the substrate specificity of the enzyme from *Shigella flexneri*. Adv. Synth. Catal. **347**:1155–1162.

38. **Babich L, Hartog AF, van der Horst MA, Wever R**. 2012. Continuous-flow reactor-based enzymatic synthesis of phosphorylated compounds on a large scale. Eur. J. Chem. **18**:6604–6609.

39. **Babich L, Hartog AF, van Hemert LJ, Rutjes FP, Wever R**. 2012. Synthesis of carbohydrates in a continuous flow reactor by immobilized phosphatase and aldolase. ChemSusChem **5**:2348–53.

40. **Fernley HN, Walker PG**. 1967. Studies on alkaline phosphatase - inhibition by phosphate derivatives and substrate specificity. Biochem. J. **104**:1011–1018.

41. **O'Brien PJ, Herschlag D**. 2002. Alkaline phosphatase revisited: hydrolysis of alkyl phosphates. Biochemistry **41**:3207–3225.

42. **Goedl C, Sawangwan T, Wildberger P, Nidetzky B**. 2010. Sucrose phosphorylase: a powerful transglucosylation catalyst for synthesis of α-D-glucosides as industrial fine chemicals. Biocatal. Biotransformation **28**:10–21.

43. **Renirie R, Pukin A, van Lagen B, Franssen MCR**. 2010. Regio- and stereoselective glucosylation of diols by sucrose phosphorylase using sucrose or glucose 1-phosphate as glucosyl donor. J. Mol. Catal. B Enzym. **67**:219–224.

44. **Pradel E, Marck C, Boquet PL**. 1990. Nucleotide sequence and transcriptional analysis of the *Escherichia coli* agp gene encoding periplasmic acid glucose-1-phosphatase. J. Bacteriol. **172**:802–807.

45. **Lee DC, Cottrill MA, Forsberg CW, Jia Z**. 2003. Functional insights revealed by the crystal structures of *Escherichia coli* glucose-1-phosphatase. J. Biol. Chem. **278**:31412–31418.

46. **Kuznetsova E, Proudfoot M, Gonzalez CF, Brown G, Omelchenko MV, Borozan I, Carmel L, Wolf YI, Mori H, Savchenko AV, Arrowsmith CH, Koonin EV, Edwards**

AM, Yakunin AF. 2006. Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. J. Biol. Chem. **281**:36149–36161.

47. **Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L**. 2006. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. J. Mol. Biol. **361**:1003–1034.
48. **Collet JF, Stroobant V, Pirard M, Delpierre G, Van Schaftingen E**. 1998. A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. J. Biol. Chem. **273**:14107–14112.
49. **Lahiri SD, Zhang G, Dunaway-Mariano D, Allen KN**. 2006. Diversification of function in the haloacid dehalogenase enzyme superfamily: The role of the cap domain in hydrolytic phosphoruscarbon bond cleavage. Bioorganic Chem. **34**:394–409.
50. **Morais MC, Zhang W, Baker AS, Zhang G, Dunaway-Mariano D, Allen KN**. 2000. The crystal structure of *Bacillus cereus* phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. Biochemistry **39**:10385–10396.
51. **Wildberger P, Todea A, Nidetzky B**. 2012. Probing enzyme-substrate interactions at the catalytic subsite of *Leuconostoc mesenteroides* sucrose phosphorylase with site-directed mutagenesis: the roles of Asp49 and Arg395. Biocatal Biotransformation **30**:326–337.
52. **Krahulec S, Armao GC, Weber H, Klimacek M, Nidetzky B**. 2008. Characterization of recombinant *Aspergillus fumigatus* mannitol-1-phosphate 5-dehydrogenase and its application for the stereoselective synthesis of protio and deuterio forms of D-mannitol 1-phosphate. Carbohydr. Res. **343**:1414–1423.

53. **Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods **6**:343–345.

54. **Whitmore L, Wallace BA**. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers **89**:392–400.

55. **Shevchenko A, Wilm M, Vorm O, Mann M**. 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. **68**:850–858.

56. **Saheki S, Takeda A, Shimazu T**. 1985. Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. Anal. Biochem. **148**:277–281.

57. **Eis C, Nidetzky B**. 1999. Characterization of trehalose phosphorylase from *Schizophyllum commune*. Biochem. J. **341**:385–393.

58. **Bradford MM**. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.

59. **Jia Z, Cottrill M, Pal GP, Lee D, Sung M, Forsberg CW, Phillips JP**. 2001. Purification, crystallization and preliminary X-ray analysis of the *Escherichia coli* glucose-1-phosphatase. Acta Crystallogr. Sect. D **57**:314–316.

60. **Cottrill MA, Golovan SP, Phillips JP, Forsberg CW**. 2002. Inositol phosphatase activity of the *Escherichia coli* agp-encoded acid glucose-1-phosphatase. Can. J. Microbiol. **48**:801–809.

61. **Kietzmann M, Schwab H, Pichler H, Ivancic M, May O, Luiten RGM**. July 2011. Preparation of an esterase. WO 2009004093.

62. **Graff G, Krick TP, Walseth TF, Goldberg ND**. 1980. The use of [$^{18}O_4$]phosphoric acid in the quantitation of phosphate by gas-liquid chromatography-mass spectrometry analysis. Anal. Biochem. **107**:324–331.

63. **Pradines A, Klaebe A, Perie J, Paul F, Monsan P**. 1991. Large-scale enzymatic synthesis of glycerol 1-phosphate. Enzyme Microb. Technol. **13**:19–23.

64. **Oh B-C, Choi W-C, Park S, Kim Y–., Oh T-K**. 2004. Biochemical properties and substrate specificities of alkaline and histidine acid phytases. Appl. Microbiol. Biotechnol. **63**:362–372.

65. **Ray WJ Jr, Roscelli GA**. 1964. A kinetic stuy of the phosphoglucomutase pathway. J. Biol. Chem. **239**:1228–1236.

66. **Mitchell DB, Vogel K, Weimann BJ, Pasamontes L, van Loon APGM**. 1997. The phytase subfamily of histidine acid phosphatases: isolation of genes for two novel phytases from the fungi *Aspergillus terreus* and *Myceliophthora thermophila*. Microbiology **143**:245–252.

67. **Tomschy A, Brugger R, Lehmann M, Svendsen A, Vogel K, Kostrewa D, Lassen SF, Burger D, Kronenberger A, van Loon APGM, Pasamontes L, Wyss M**. 2002. Engineering of phytase for improved activity at low pH. Appl. Environ. Microbiol. **68**:1907–1913.

68. **Pogell BM**. 1953. A new synthesis of fructose-1-phosphate with phosphorus pentoxide. J. Biol. Chem. **201**:645–649.

69. **Fessner W-D, Sinerius G, Schneider A, Dreyer M, Schulz GE, Badia J, Aguilar J**. 1991. Diastereoselective enzymatic aldol additions: L-rhamnulose and L-fuculose 1-phosphate aldolases from E. coli. Angew. Chem. Int. Ed. Engl. **30**:555–558.

70. **Lefebvre MJ, Gonzalez NS, Pontis HG**. 1964. Anion-exchange chromatography of sugar phosphates with triethylammonium borate. J. Chromatogr. **15**:495–500.

71. **Smits HP, Cohen A, Buttler T, Nielsen J, Olsson L**. 1998. Cleanup and analysis of sugar phosphates in biological extracts by using solid-phase extraction and anion-exchange chromatography with pulsed amperometric detection. Anal. Biochem. **261**:36–42.

72. **Lu Z, Dunaway-Mariano D, Allen KN**. 2005. HAD superfamily phosphotransferase substrate diversification: structure and function analysis of HAD subclass IIB sugar phosphatase BT4131. Biochemistry **44**:8684–8696.

73. **Tanaka N, Hasan Z, Hartog AF, van Herk T, Wever R, Sanders RJ**. 2003. Phosphorylation and dephosphorylation of polyhydroxy compounds by class A bacterial acid phosphatases. Org. Biomol. Chem. **1**:2833–2839.

74. **Mihara Y, Utagawa T, Yamada H, Asano Y**. 2001. Acid phosphatase/phosphotransferases from enteric bacteria. J. Biosci. Bioeng. **92**:50–54.

75. **Koerner TAW, Voll RJ, Cary LW, Younathan ES**. 1980. Carbon-13 nuclear magnetic resonance studies and anomeric composition of ketohexose phosphates in solution. Biochemistry **19**:2795–2801.

76. **Lu Z, Dunaway-Mariano D, Allen KN**. 2008. The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state. Proc. Natl. Acad. Sci. **105**:5687–5692.

1.8 Legend to Schemes and Figures

Scheme 1. Reactions catalyzed by disaccharide phosphorylases (A), kinases (B) and phosphatases (C).

Scheme 2. Proposed double displacement-like mechanism of hydrolysis of αGlc 1-*P* catalyzed by αGlc 1-*P* phosphatase (A) and Had13 (B). (C) Phosphoryl group-transfer from αGlc 1-*P* to Fru catalyzed by αGlc 1-*P* phosphatase.

Note: The theoretical anomeric composition of Fru 1-*P* in aqueous solution is α-furanose 5%, β-furanose 16%, α-pyranose 5% and β-pyranose 73% (75). Fru 1-*P* is shown here in the β-pyranose form.

Figure 1. αGlc 1-*P* phosphatase (A, B) and Had13 (C, D) represent different phosphatase protein families. (A) Histidine acid phosphatase fold of αGlc 1-*P* phosphatase (pdb entry 1nt4, chain A) showing the α-domain in violet, the α/β domain in cyan, and the central catalytic cleft in wheat. Disulfide bonds are indicated in yellow. (B) Close-up view of the active site having glucosyl phosphate ligand bound. (C) Haloacid dehydrogenase fold of Had13 (pdb entry 1rkq) showing the cap domain in violet and the α/β domain in cyan. Unique structural features of Had13, namely the squiggle (an almost complete α-helical turn; light brown) and

the flap (a β-hairpin turn; yellow) are also indicated (47, 76). (D) Close-up view of the active site having Mg^{2+} bound.

Figure 2. Incorporation of ^{18}O isotope into the phosphate released during hydrolysis of αGlc 1-*P* by phosphatase or phosphorylase in $H_2{}^{18}O$ solvent. The ratio m/z 301/299 (singly ^{18}O-labeled phosphate/non-labeled phosphtate) was measured by GC-MS. Black bars show reactions in normal water; grey bars show reactions in ^{18}O-enriched solvent. The m/z 301/299 at natural abundance is expected to be 0.13, corresponding to measurements in H_2O. For further explanations, see text.

Figure 3. Kinetic study of the hydrolysis-transphosphorylation reaction catalyzed by αGlc 1-*P* phosphatase. (A) The ratio of the total transphosphorylation rate (Fru 1-*P*, Fru 6-*P*, Glc 6-*P*) and the αGlc 1-*P* consumption rate in dependency on the Fru concentration (100 mM – 600 mM; solid line). The ratio approaches a value of unity at high Fru concentrations, as indicated by the dashed line. (B) pH dependencies of logarithmic k_{cat_app} for αGlc 1-*P* consumption (○) and Fru 1-*P* synthesis (●). Lines are fits of equation 3 to the data. (C) A complete time course of conversion of 20 mM αGlc 1-*P* in the presence of 200 mM Fru. Symbols indicate: αGlc 1-*P* (▼), sum of phosphorylated products (■) and phosphate (○). (D) Phosphorylated products synthesized in transphosphorylation reactions performed at different initial concentrations of Fru at 50 min. Fru 1-*P* (grey), Glc 6-*P* (black), Fru 6-*P* (grey dashed), phosphate (white).

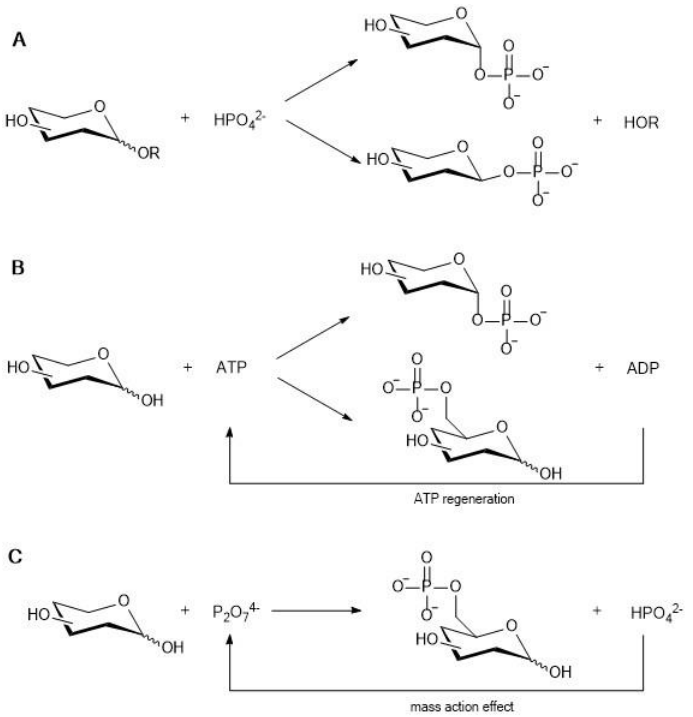
Figure 4. Best-fit molecular docking poses for αGlc 1-*P* (green, A) and Glc 6-*P* (magenta, C) in the substrate binding pocket of αGlc 1-*P* phosphatase (pdb entry 1nt4, molecule A). Figure B shows the βGlc 1-*P* ligand (blue) present in the published crystal structure of the enzyme. Hydrogen bonds (≤ 3.4 Å) between Ser^{291} and the C4-OH and/or C6-OH of the glucosyl moiety and/or Asp^{290} and the glycosidic oxygen atom are shown as black-dashed lines. The side-chains of Arg^{17}, Arg^{21}, Arg^{94} and His^{289} interact with the phosphate group of the ligand; the glucosyl moiety is coordinated by the side-chains of Asp^{290} and Ser^{291}.

Table 1. Apparent turnover frequencies for hydrolysis of phosphorylated sugars. The S.D. for k_{cat_app} was equal to or smaller than 10% of the reported value, except [a]14%, [b]18%, and [c]24%. n.t., not tested; n.d., no activity above assay detection limit.

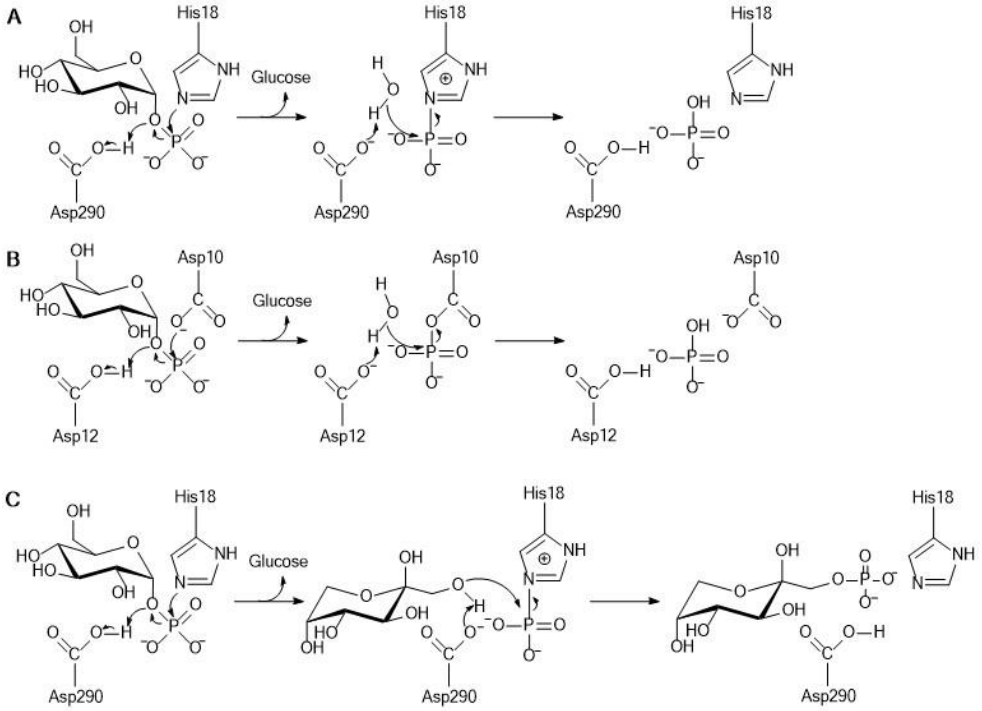
Substrate	**αGlc 1-*P* phosphatase**	**Had13**
	k_{cat_app} [s^{-1}]	k_{cat_app} [s^{-1}]
αGlc 1-*P*	40	12
βGlc 1-*P*	n.d.	$6.0 \cdot 10^{-2}$
Glc 6-*P*	56	8.6[b]
Fru 1-*P*	22	n.t.
Fru 6-*P*	22[a]	4.4[c]
Phytate	2.1	n.d.
Pyrophosphate	n.d.	n.d.

Table 2. Kinetic analysis of transphosphorylation by αGlc 1-*P* phosphatase analyzed at different Fru concentrations. The S.D. for k_{cat_app} was ≤ 10% of the reported value, except [a]11%, [b]16%, and [c]20%. Not applicable, n.a.

Fru **[mM]**	**Glc 1-*P*** k_{cat_app} **[s^{-1}]**	**Fru 1-*P*** k_{cat_app} **[s^{-1}]**	**Glc 6-*P*** k_{cat_app} **[s^{-1}]**	**Fru 6-*P*** k_{cat_app} **[s^{-1}]**	**Phosphate** k_{cat_app} **[s^{-1}]**
0	40	n.a.	n.a.	n.a.	40
100	73	39	5	11[b]	21
200	137	88	6	20	18
400	162	111[a]	4	28	14
600	189	135	5	36[c]	13



Scheme 1. Reactions catalyzed by disaccharide phosphorylases (A), kinases (B) and phosphatases (C).



Scheme 2. Proposed double displacement-like mechanism of hydrolysis of αGlc 1-*P* catalyzed by αGlc 1-*P* phosphatase (A) and Had13 (B). (C) Phosphoryl group-transfer from αGlc 1-*P* to Fru catalyzed by αGlc 1-*P* phosphatase.

Note: The theoretical anomeric composition of Fru 1-*P* in aqueous solution is α-furanose 5%, β-furanose 16%, α-pyranose 5% and β-pyranose 73% (75). Fru 1-*P* is shown here in the β-pyranose form.

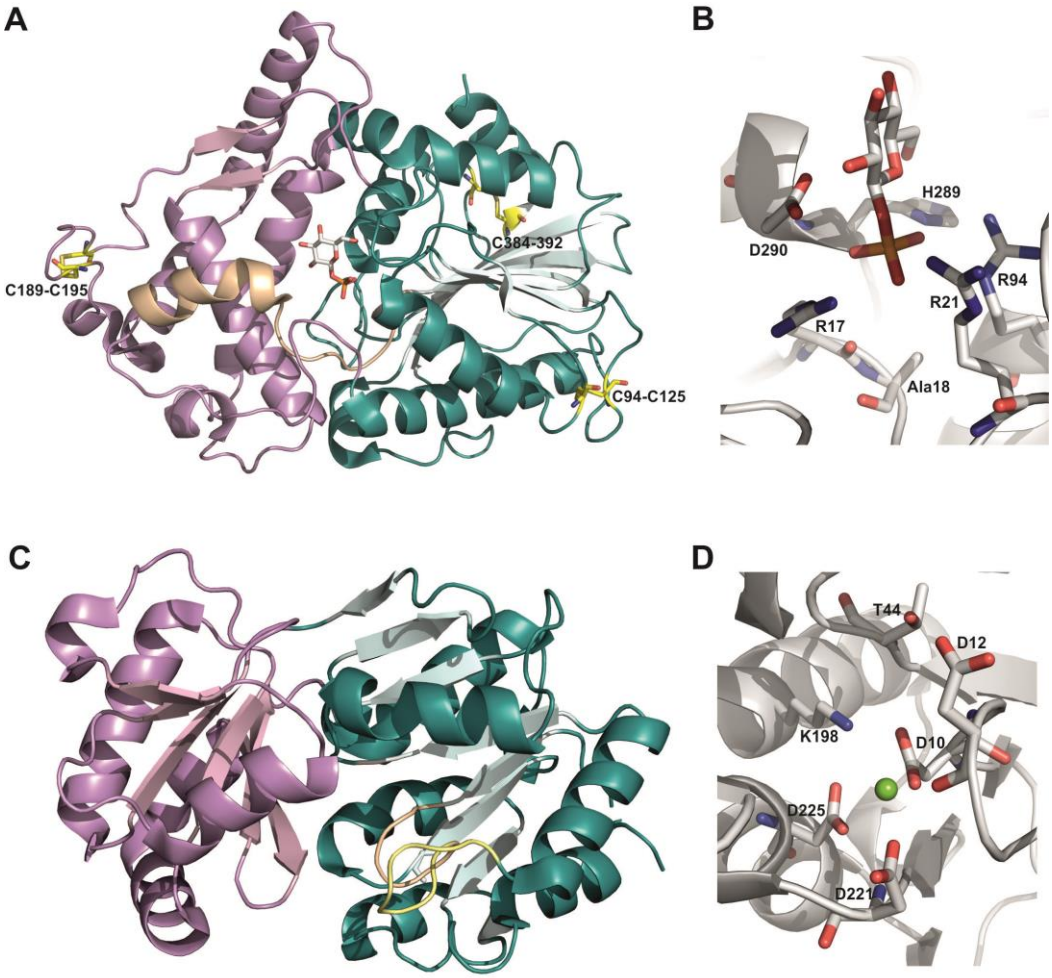


Figure 1. αGlc 1-*P* phosphatase (A, B) and Had13 (C, D) represent different phosphatase protein families. (A) Histidine acid phosphatase fold of αGlc 1-*P* phosphatase (pdb entry 1nt4, chain A) showing the α-domain in violet, the α/β domain in cyan, and the central catalytic cleft in wheat. Disulfide bonds are indicated in yellow. (B) Close-up view of the active site having glucosyl phosphate ligand bound. (C) Haloacid dehydrogenase fold of Had13 (pdb entry 1rkq) showing the cap domain in violet and the α/β domain in cyan. Unique structural features of Had13, namely the squiggle (an almost complete α-helical turn; light brown) and the flap (a β-hairpin turn; yellow) are also indicated (47, 76). (D) Close-up view of the active site having Mg^{2+} bound.

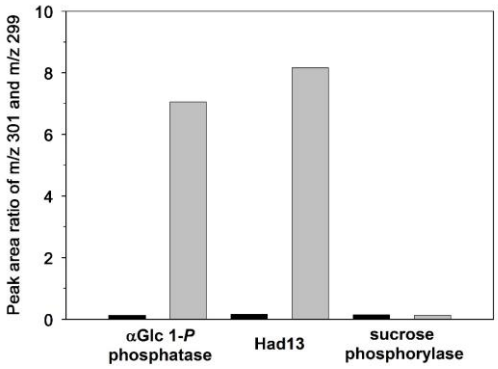


Figure 2. Incorporation of ^{18}O isotope into the phosphate released during hydrolysis of αGlc 1-*P* by phosphatase or phosphorylase in $H_2{}^{18}O$ solvent. The ratio m/z 301/299 (singly ^{18}O-labeled phosphate/non-labeled phosphtate) was measured by GC-MS. Black bars show reactions in normal water; grey bars show reactions in ^{18}O-enriched solvent. The m/z 301/299 at natural abundance is expected to be 0.13, corresponding to measurements in H_2O. For further explanations, see text.

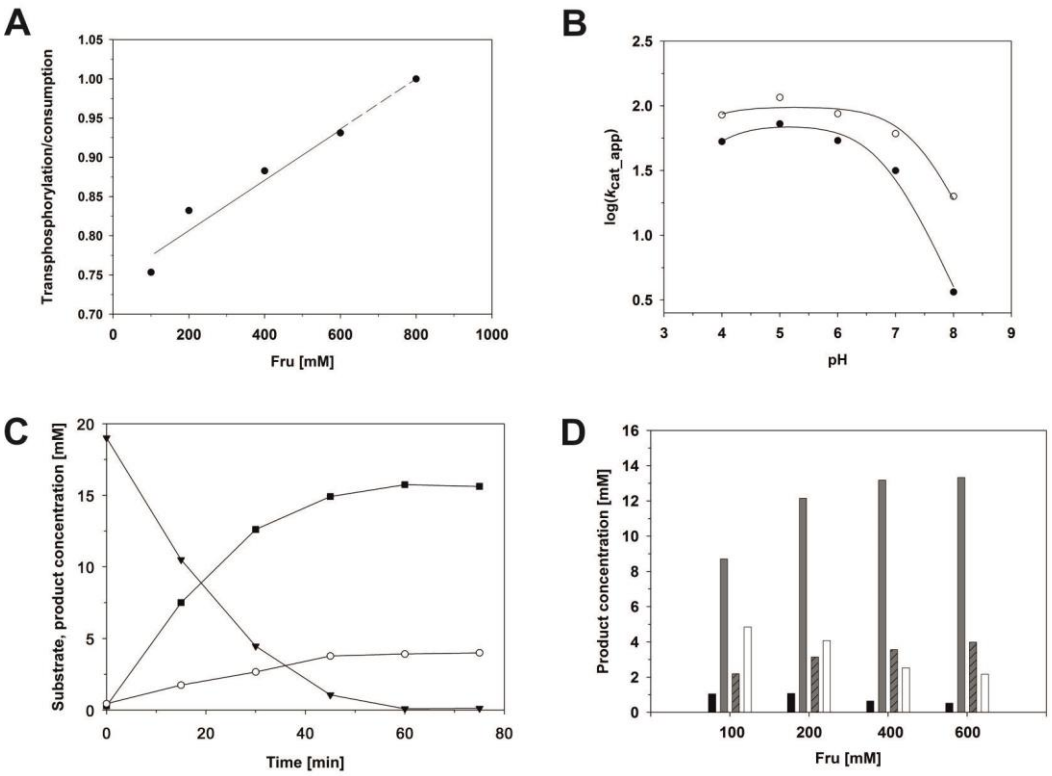


Figure 3. Kinetic study of the hydrolysis-transphosphorylation reaction catalyzed by αGlc 1-*P* phosphatase. (A) The ratio of the total transphosphorylation rate (Fru 1-*P*, Fru 6-*P*, Glc 6-*P*) and the αGlc 1-*P* consumption rate in dependency on the Fru concentration (100 mM – 600 mM; solid line). The ratio approaches a value of unity at high Fru concentrations, as indicated by the dashed line. (B) pH dependencies of logarithmic k_{cat_app} for αGlc 1-*P* consumption (○) and Fru 1-*P* synthesis (●). Lines are fits of equation 3 to the data. (C) A complete time course of conversion of 20 mM αGlc 1-*P* in the presence of 200 mM Fru. Symbols indicate: αGlc 1-*P* (▼), sum of phosphorylated products (■) and phosphate (○). (D) Phosphorylated products synthesized in transphosphorylation reactions performed at different initial concentrations of Fru at 50 min. Fru 1-*P* (grey), Glc 6-*P* (black), Fru 6-*P* (grey dashed), phosphate (white).

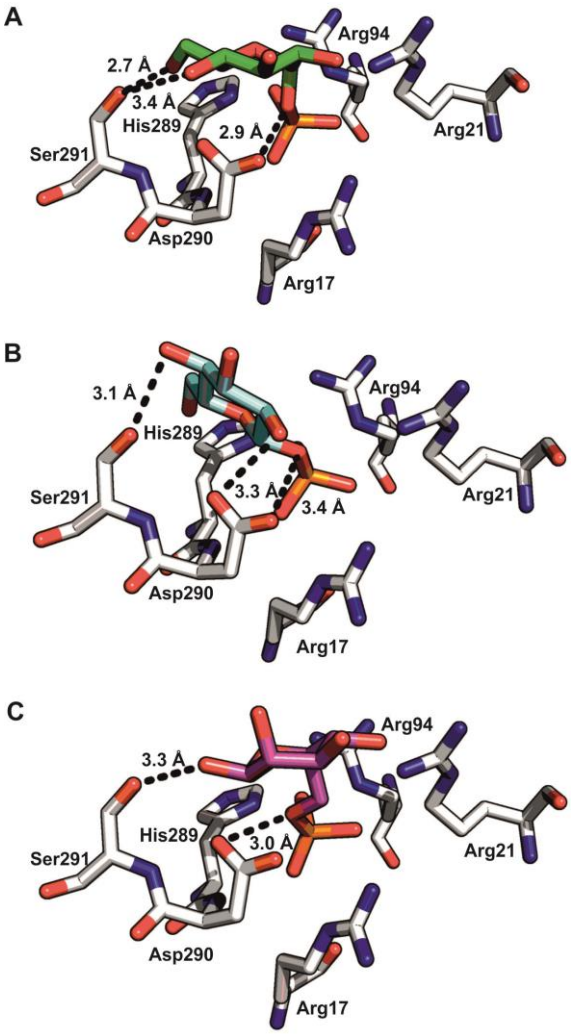
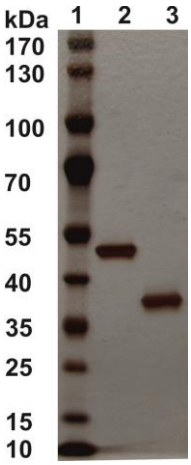


Figure 4. Best-fit molecular docking poses for αGlc 1-*P* (green, A) and Glc 6-*P* (magenta, C) in the substrate binding pocket of αGlc 1-*P* phosphatase (pdb entry 1nt4, molecule A). Figure B shows the βGlc 1-*P* ligand (blue) present in the published crystal structure of the enzyme. Hydrogen bonds (≤ 3.4 Å) between Ser^{291} and the C4-OH and/or C6-OH of the glucosyl moiety and/or Asp^{290} and the glycosidic oxygen atom are shown as black-dashed lines. The side-chains of Arg^{17}, Arg^{21}, Arg^{94} and His^{289} interact with the phosphate group of the ligand; the glucosyl moiety is coordinated by the side-chains of Asp^{290} and Ser^{291}.

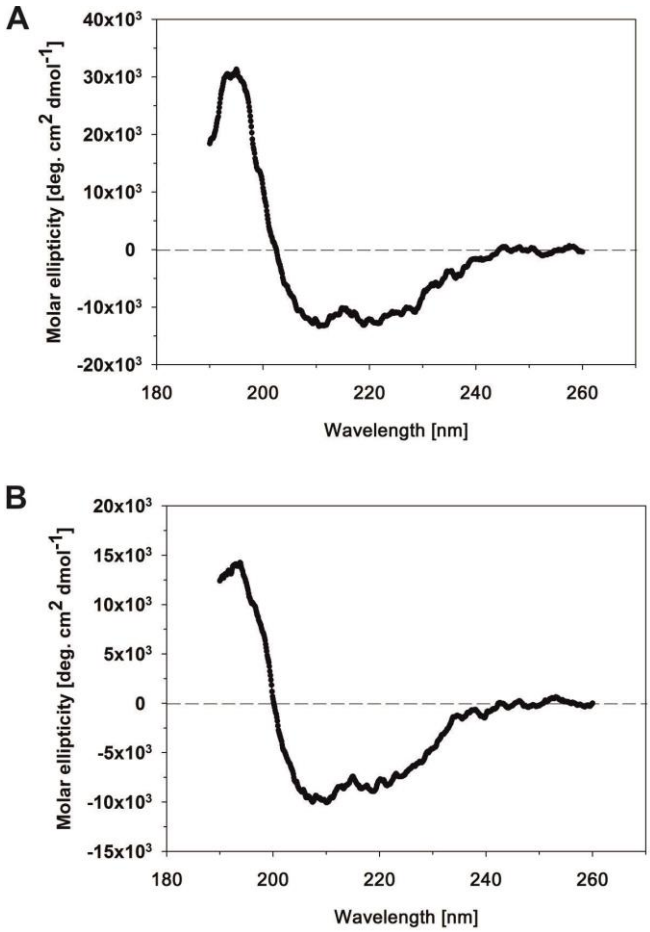
1.9 SUPPLEMENTARY INFORMATION

1.9.1 Molecular cloning of αGlc 1-*P* phosphatase – detailed experimental protocol.

E. coli BL21-Gold (DE3) cells were grown overnight at 30°C in Lennox-medium. The cells were centrifuged at 20,000 g and 4°C for 30 min. Pellet was used for preparation of genomic DNA. The gene encoding αGlc 1-*P* phosphatase was amplified by a PCR that used *Phusion* DNA polymerase (Thermo Fisher Scientific, Waltham, USA) in combination with oligonucleotide primers 1 and 2 (Supplementary Table S2). PCR consisted of a pre-heating step at 98°C for 5 min, followed by 30 reaction cycles of denaturation at 98°C for 30 s, annealing at 70°C for 30 s, and elongation at 72°C for 1.5 min. The final extension step was carried out at 72°C for 5 min. The presence of an *N*-terminal, 66 bp long signal sequence was predicted using SignalIP-4.1 at the Center of Biological Sequence Analysis (http://www.cbs.dtu.dk/services/SignalP-4.1/) (1). To remove the *N*-terminal signal sequence, to add an *N*-terminal *Strep*-tag, and to also extend the gene end with overlapping regions to the vector, we performed a PCR in which oligonuecleotide primers 3 and 4 were used (Supplementary Table S2). To extend the gene start with overlapping regions to the vector, a PCR with primers 5 and 3 was performed. The final amplification product was treated by *Dpn*I to subjected to digest the parental template. The final construct was cloned into the linearized pMS470_dsbC vector via Gibson assembly (2). Sequenced plasmid vectors harboring the structural gene were transformed into *E. coli* Origami B cells.



Supplementary Figure S1. SDS polyacrylamide gel showing purified enzymes. Lane 1: molecular mass standard; 2: αGlc 1-*P* phosphatase; 3: Had13. Silver staining was used.



Supplementary Figure S2. CD spectra of αGlc 1-*P* phosphatase (A) and Had13 (B). Spectra are averaged from ten individual wavelength scans.

1 M**ASWSHPQFE KIEGRQTVPE GY**QLQQVLMM SRHNLR**APLA NNGSVLEQST**

51 **PNKWPEWDVP GGQLTTK**GGV LEVYMGHYMR **EWLAEQGMVK SGECPPPDTV**

101 **YAYANSLQRT VATAQF**FITG AFPGCDIPVH HQEKMGTMDP TFNPVITDDS

151 AAF**SEKAVAA MEKEL**SK**LQL TDSYQLLEK**I VNY**KDSPACK EKQQCSLVDG**

201 **K**NTFSAK**YQQ EPGVSGPLK**V GNSLVDAFTL QYY**EGFPMDQ VAWGEIKSDQ**

251 **QW**KVLSKLK**N GYQDSLFTSP EVARNVAKPL VSYIDK**AL**VT DRTSAPKITV**

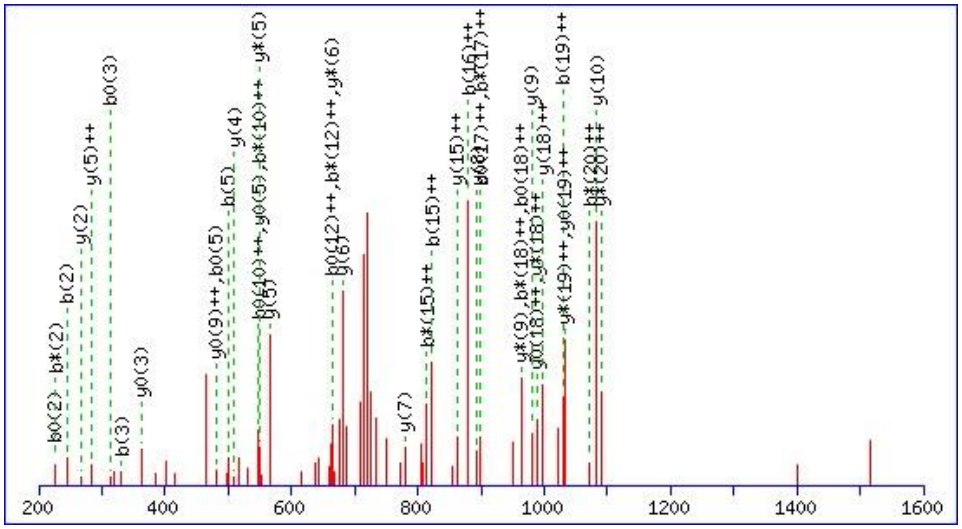
301 **LVGHDSNIAS L**LTALDFKPY QLHDQNERTP IGGKIVFQRW HDSKANRDLM

351 K**IEYVYQSAE QLRNADALTL QAPAQRVTL**E LSGCPIDANG F**CPMDKFDSV**

401 **LNEAVK**

Supplementary Figure S3. Sequence coverage obtained by LC-MS/MS analysis of tryptic and chymotryptic peptides generated from purified αGlc 1-*P* phosphatase. Matched peptides are bold.

A

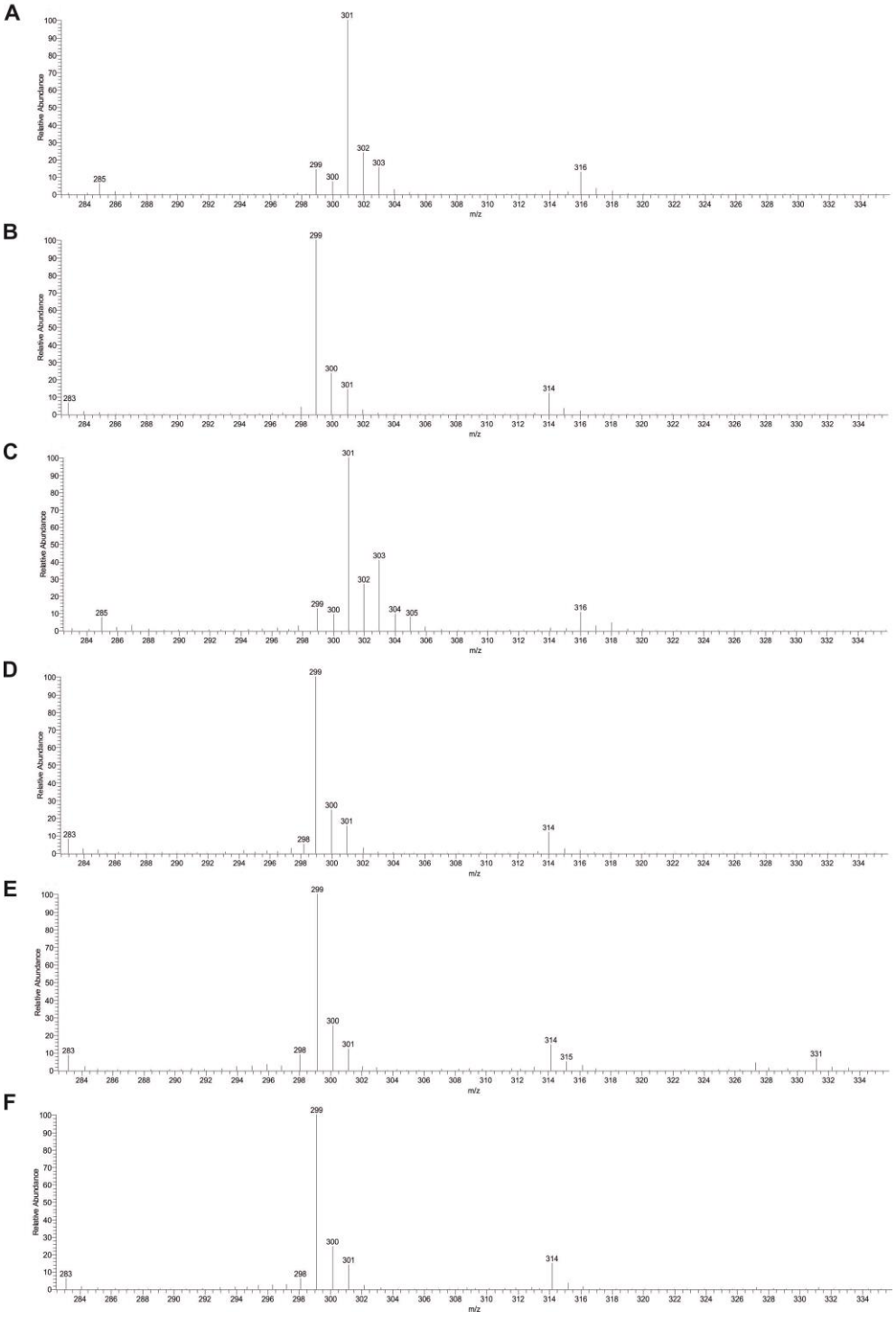


B

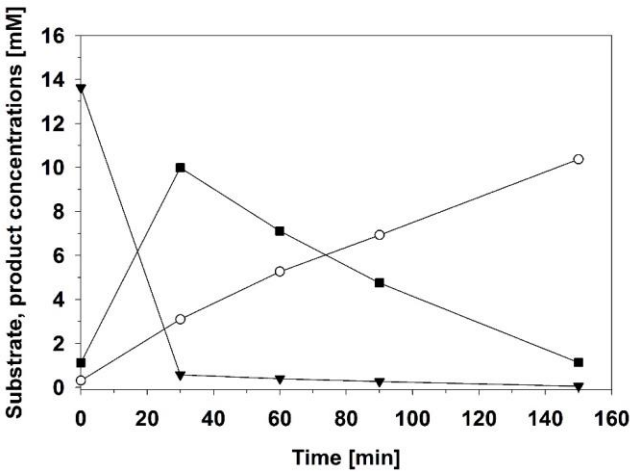
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	129.1022	65.0548	112.0757	56.5415			K							21
2	***244.1292***	122.5682	227.1026	114.0550	226.1186	113.5629	D	2196.0060	1098.5066	2178.9794	1089.9934	2177.9954	1089.5013	20
3	***331.1612***	166.0842	314.1347	157.5710	313.1506	157.0790	S	2080.9790	1040.9932	2063.9525	1032.4799	2062.9685	1031.9879	19
4	428.2140	214.6106	411.1874	206.0974	410.2034	205.6053	P	1993.9470	997.4771	1976.9205	988.9639	1975.9364	988.4719	18
5	***499.2511***	250.1292	482.2245	241.6159	481.2405	241.1239	A	1896.8942	948.9508	1879.8677	940.4375	1878.8837	939.9455	17
6	600.2446	300.6259	583.2181	292.1127	582.2341	291.6207	C	1825.8571	913.4322	1808.8306	904.9189	1807.8466	904.4269	16
7	728.3396	364.6734	711.3130	356.1602	710.3290	355.6681	K	1724.8636	862.9354	1707.8370	854.4222	1706.8530	853.9302	15
8	857.3822	429.1947	840.3556	420.6815	839.3716	420.1894	E	1596.7686	798.8880	1579.7421	790.3747	1578.7581	789.8827	14
9	985.4771	493.2422	968.4506	484.7289	967.4666	484.2369	K	1467.7260	734.3667	1450.6995	725.8534	1449.7155	725.3614	13
10	1113.5357	557.2715	1096.5092	548.7582	1095.5252	548.2662	Q	1339.6311	670.3192	1322.6045	661.8059	1321.6205	661.3139	12
11	1241.5943	621.3008	1224.5677	612.7875	1223.5837	612.2955	Q	1211.5725	606.2899	1194.5460	597.7766	1193.5619	597.2846	11
12	1344.6035	672.8054	1327.5769	664.2921	1326.5929	663.8001	C	***1083.5139***	542.2606	1066.4874	533.7473	1065.5034	533.2553	10
13	1431.6355	716.3214	1414.6090	707.8081	1413.6249	707.3161	S	***980.5047***	490.7560	963.4782	482.2427	962.4942	481.7507	9
14	1544.7196	772.8634	1527.6930	764.3502	1526.7090	763.8581	L	***893.4727***	447.2400	876.4462	438.7267	875.4621	438.2347	8
15	1643.7880	**822.3976**	1626.7614	813.8844	1625.7774	813.3924	V	***780.3886***	390.6980	763.3621	382.1847	762.3781	381.6927	7

16	1758.8149	**879.9111**	1741.7884	871.3978	1740.8044	870.9058	**D**	***681.3202***	341.1638	664.2937	332.6505	663.3097	332.1585	**6**
17	1815.8364	908.4218	1798.8098	899.9086	1797.8258	899.4166	**G**	***566.2933***	283.6503	549.2667	275.1370	548.2827	274.6450	**5**
18	1943.9314	972.4693	1926.9048	963.9560	1925.9208	963.4640	**K**	***509.2718***	255.1395	492.2453	246.6263	491.2613	246.1343	**4**
19	2057.9743	**1029.4908**	2040.9477	1020.9775	2039.9637	1020.4855	**N**	381.1769	191.0921	364.1503	182.5788	363.1663	182.0868	**3**
20	2159.0220	1080.0146	2141.9954	1071.5013	2141.0114	1071.0093	**T**	***267.1339***	134.0706			249.1234	125.0653	**2**
21							**F**	166.0863	83.5468					**1**

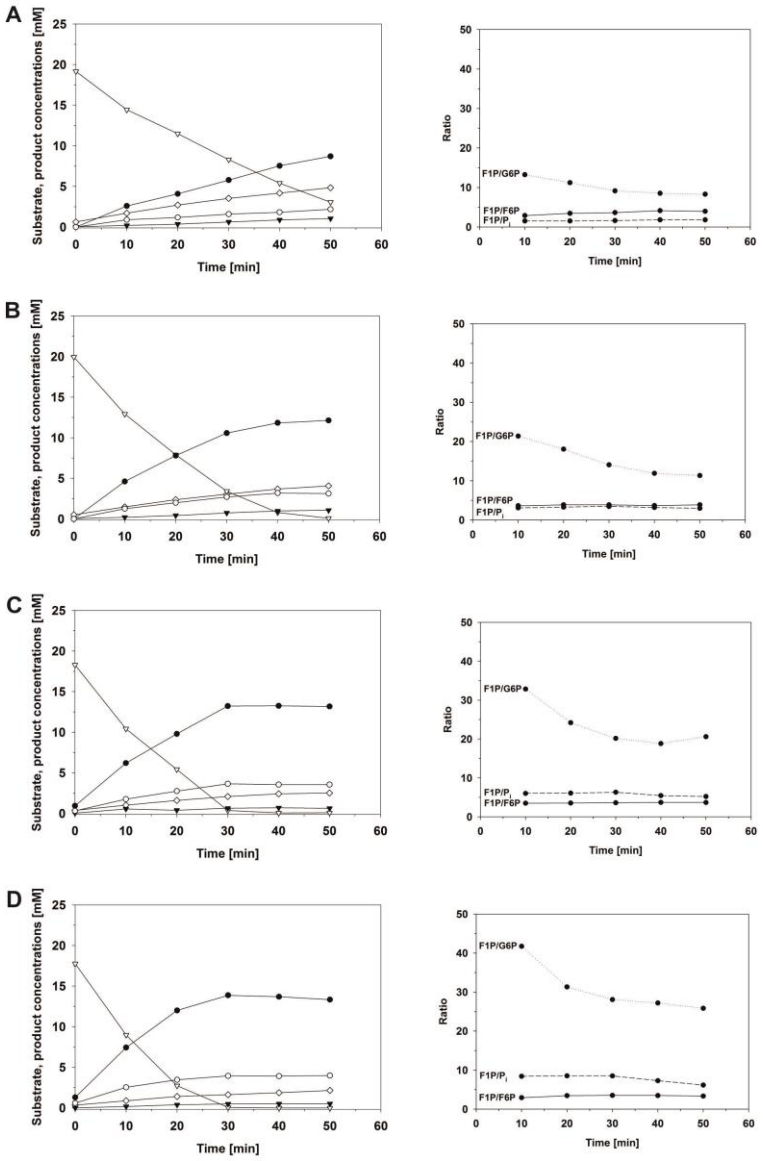
Supplementary Figure S4. Annotated MS-MS spectra (A) and detected b and y ions (B) from Mascot search result of the peptide KDSPA**C**KEKQQ**C**SLVDGKNTF containing the disulfide bond between Cys^{189} and Cys^{195} are shown. Monoisotopic mass of neutral peptide Mr(calc): 2323.0937 Da. Mass error: -0.0017 Da. Mascot ions score: 34. Expect: 0.00042. Matches: 42/232 fragment ions using 59 most intense peaks.



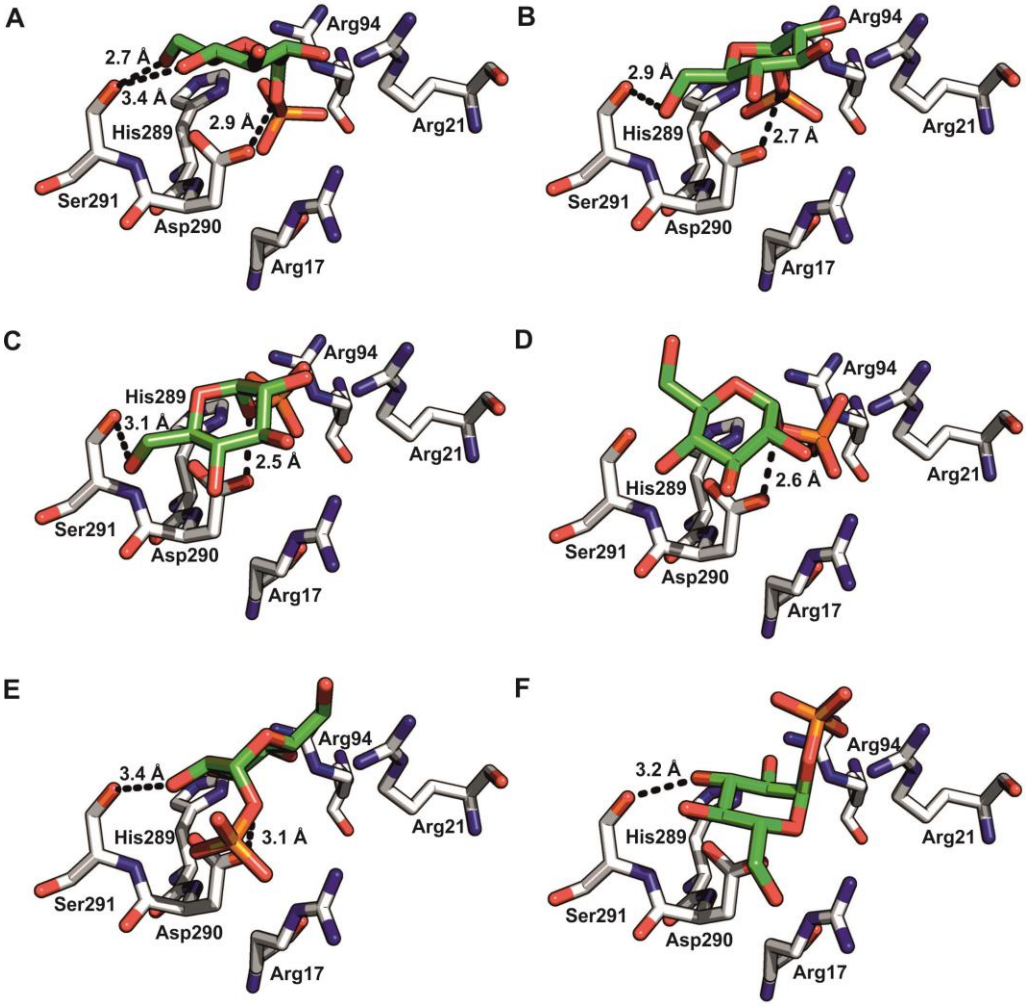
Supplementary Figure S5. GC-MS chromatograms showing the isotopic composition of the phosphate released during hydrolysis of αGlc 1-*P* catalyzed by αGlc 1-*P* phosphatase in normal H_2O and $H_2{}^{18}O$ solvent. (A: $H_2{}^{18}O$; B: H_2O), Had13 (C: $H_2{}^{18}O$; D: H_2O) and sucrose phosphorylase (E: $H_2{}^{18}O$; F: H_2O).



Supplementary Figure S6. Time-course for the synthesis of Glc 6-*P* from αGlc 1-*P* and Glc catalyzed by αGlc 1-*P* phosphatase. The reaction mixture contained 20 mM αGlc 1-*P*, 200 mM Glc and 0.1 μM αGlc 1-*P* phosphatase in 50 mM Mes, pH 7.0. Symbols indicate: αGlc 1-*P* (▼), Glc 6-*P* (■) and phosphate (○).



Supplementary Figure S7. Time-courses for Fru 1-*P* (●) production from 20 mM αGlc 1-*P* (▽) in the presence of 100 mM (A), 200 mM (B), 400 mM (C), and 600 mM Fru (D), respectively. The formation of side products, Glc 6-*P* (▼), Fru 6-*P* (○), and phosphate (◊) is indicated. The ratios of Fru 1-*P* to Glc 6-*P* (F1P/G6P), Fru 1-*P* to Fru 6-*P* (F1P/F6P) and Fru 1-*P* to hydrolysis (F1P/Pi), displayed on the right panel, are extracted from the corresponding time courses displayed on the left.



Supplementary Figure S8. Close-up view of molecular docking poses of αGlc 1-*P* (green) in the active site of αGlc 1-*P* phosphatase (pdb entry 1nt4, molecule A). The docking algorithm resulted in 6 docking poses, with the best-fit docking pose illustrated in panel A. Panels B-F display the remaining docking poses in descending order based on the calculated free energies. Hydrogen bonds (≤ 3.4 Å) between Ser^{291} and the C4-OH and/or C6-OH of the glucosyl moiety and/or Asp^{290} and the glycosidic oxygen atom are shown as black-dashed lines.

Supplementary Table S1. Relative content of secondary structural elements in αGlc 1-*P* phosphatase and Had13 obtained from the CD spectra and from the crystal structure (in parenthesis).

Secondary structure	**αGlc 1-*P* phosphatase [%]**		**Had13 [%]**	
Alpha helices	51	(39)	28	(39)
Beta strands	22	(14)	25	(22)
Turns	11	(47)	19	(39)
Unordered	16		28	

Supplementary Table S2: Oligonucleotide primers used for molecular cloning of αGlc 1-*P* phosphatase.

Primer	**Sequence (5' → 3')**
Primer 1	ATGAACAAAACGCTAATCACC
Primer 2	TTATTTCACCGCTTCATTCAAC
Primer 3	ATGGCTAGCTGGAGCCACCCGCAGTTCGAAAAAATCGAAGGGCGCCA AACCGTACCGGAAGGCTATCAGC
Primer 4	CATCCGCCAAAACAGCCAAGCTTATTATTTCACCGCTTCATTC
Primer 5	GTTTAACTTTAAGAAGGAGATATACATATGGCTAGCTGGAGCCACC

1.9.2 Supplementary References

1. **Petersen TN, Brunak S, von Heijne G, Nielsen H**. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8**:785–786.
2. **Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods **6**:343–345.

2 *Yihx*-encoded haloacid dehalogenase-like phosphatase HAD4 from *Escherichia coli* is a specific α-D-glucose 1-phosphate hydrolase useful for substrate-selective sugar phosphate transformations

Martin Pfeiffer[1], Patricia Wildberger[1] and Bernd Nidetzky[1,2*]

[1]Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

[2] acib - Austrian Centre of Industrial Biotechnology, Petersgasse 14, A-8010 Graz, Austria

* Corresponding author; e-mail: bernd.nidetzky@tugraz.at

Running title: Substrate-selective sugar phosphate hydrolysis by HAD

2.1 Abstract

Phosphomonoester hydrolases (phosphatases; EC 3.1.3.-) often exhibit extremely relaxed substrate specificity which limits their application to substrate-selective biotransformations. In search of a phosphatase catalyst specific for hydrolyzing α-D-glucose 1-phosphate (αGlc 1-*P*), we selected haloacid dehalogenase-like phosphatase 4 (HAD4) from *E. coli* and obtained highly active recombinant enzyme through a fusion protein (Z_{basic2}_HAD4) that contained Z_{basic2}, a strongly positively charged three α-helical bundle module, at its N-terminus. Highly pure Z_{basic2}_HAD4 was prepared directly from *E. coli* cell extract using capture and polishing combined in a single step of cation exchange chromatography. Kinetic studies showed Z_{basic2}_HAD4 to exhibit 565-fold preference for hydrolyzing αGlc 1-*P* (k_{cat}/K_M = 1.87 ± 0.03 $mM^{-1}s^{-1}$; 37°C, pH 7.0) as compared to D-glucose 6-phosphate (Glc 6-*P*). Due to substrate inhibition by αGlc 1-*P* with a K_{iS} 21-fold smaller than K_M, maximum activity was observed at a substrate concentration of about 10 mM where an apparent k_{cat} of about 10.7 s^{-1} was measured. Using mixture of αGlc 1-*P* and Glc 6-*P* (20 mM each) as the substrate, Z_{basic2}_HAD4 could be used to selectively convert the αGlc 1-*P* present, leaving back all of the Glc 6-*P* for recovery. Selective removal of αGlc 1-*P* from sugar phosphate preparations could be an interesting application of Z_{basic2}_HAD4 for which readily available broad-spectrum phosphatases are unsuitable.

Key words: Phosphatase; HAD superfamily; sugar phosphate; substrate selectivity; hydrolysis

2.2 Introduction

Approximately half of the intracellular carbohydrates in the *E. coli* metabolome contain a phosphate moiety [1]. A plethora of enzymes is involved in the dynamic regulation and homeostasis of sugar phosphate metabolite levels. One key enzyme subset are the phosphatases [2]. They catalyze phosphomonoester group hydrolysis to release inorganic phosphate from phosphorylated substrate.

The known sugar phosphate phosphatases represent a large diversity in protein structures, catalytic mechanisms, and substrate specificity. However, many of them are evolutionary related to haloacid dehalogenase (HAD)-like proteins [3-5]. Originally named after the activity of its first structurally characterized member (L-2-HAD from *Pseudomonas* sp. YL), the HAD-like protein superfamily currently comprises mainly phosphatases, with about 79% of the classified and biochemically studied proteins exhibiting this activity [6].

Common structural element of HAD-like phosphatases is a Rossmann fold-like α/β core domain that contains the active site (Figure 1A). According to the presence and structure of an additional domain, the so-called CAP, HAD-like phosphatases are subdivided into subfamilies CI (α-helical CAP), CII (mixed α/β CAP), and CIII (no CAP) [7]. The catalytic centre is highly conserved. It features a key Asp that is proposed to function as catalytic nucleophile in a double displacement-like enzyme reaction mechanism. Phospho-monoester hydrolysis would accordingly proceed in two catalytic steps via a covalent aspartyl-phosphate enzyme intermediate [6-8] (Figure 1B). HAD-like phosphatases require divalent metal ion (typically Mg^{2+}) bound in the active site for their full activity [9]. Phosphatase substrate specificity is dictated by flexible loop structures on the α/β core domain and on the CAP domain [10]. The

spectrum of phosphoryl substrates hydrolyzed is usually very broad, and it includes both sugar and non-sugar phosphates. Relationships between structure and specificity are only weakly defined for HAD-like phosphatases. Substrate binding and product release involve domain-closing and opening movements of the CAP [11-13] (Figure 1C). Protein conformational flexibility further complicates efforts to infer substrate specificity from sequence and three-dimensional structure information alone. Purposeful selection or molecular design of a HAD-like phosphatase for biocatalytic conversion of specific target substrates is therefore difficult. Evidence from biochemical characterization is vital for efficient development.

Selectivity is a prime feature of many enzymatic transformations [14]. A growing number of examples show that in a comparison of conventional-chemical and biocatalytic process options in organic synthesis, it is usually the bio-based selectivity that provides a decisive advantage for process development. In this paper, we address the problem of substrate-selective hydrolysis of α-D-glucose 1-phosphate (αGlc 1-*P*) by phosphatases and report identification of HAD4 (*yihx* gene product) from *E. coli* for that purpose. Major demand for the reaction is in reactive processing of mixtures of sugar phosphates to eliminate all of the αGlc 1-*P* present in the starting material. Sample work-up for analytics, in the field of sugar phosphate metabolomics for example, and facilitated sugar phosphate product recovery constitute interesting applications of selective αGlc 1-*P* converting phosphatases. The readily available broad-spectrum phosphatases are however not suitable for these applications, and discovery of new phosphatase catalysts is therefore required. HAD4 was obtained as highly active recombinant enzyme through a specially designed fusion protein (Z_{basic2}_HAD4) that contained Z_{basic2}, a strongly positively charged three α-helical bundle module, at its N-terminus [15]. Z_{basic2} facilitated functional expression of soluble protein in *E. coli* and enabled

protein capture and polishing efficiently combined in a single step of cation exchange chromatography. Kinetic characterization of Z_{basic2}_HAD4 with αGlc 1-*P* and the competing sugar phosphate substrate D-glucose 6-phosphate (Glc 6-*P*) is reported. Clear-cut separation of Glc 6-*P* from αGlc 1-*P* through selective hydrolysis of the sugar 1-phosphate is described.

2.3 Materials and methods

2.3.1 Chemicals and reagents.

Unless stated otherwise, all chemicals were of highest purity available from Sigma-Aldrich (Vienna, Austria) or Roth (Karlsruhe, Germany). Oligonucleotide primers were from Sigma-Aldrich (Vienna, Austria). DNA sequencing was done at LGC Genomics (Berlin, Germany).

2.3.2 Attempted expression and purification of His_6-HAD4

Expression vector pCA24N-yihx encoding HAD4 equipped with N-terminal hexahistidine tag was kindly provided by Dr. Alexander Yakunin. The vector was transformed into electro-competent cells of *E. coli* BL21 (DE3), and single-colony transformants were selected on agar plates containing 0.025 mg/mL chloramphenicol. Recombinant protein was produced in 1-L baffled shaking flasks at 37°C using an agitation rate of 110 rpm (Certomat BS-1 shaking incubator, Sartorius, Göttingen, Germany). Flasks contained 250 mL Lennox-medium supplemented with 0.025 mg/mL chloramphenicol. At OD_{600} of 0.8, temperature was decreased to 18°C, 0.01 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added, and cultivation was continued for 20 h. Centrifuged (4°C, 30 min; 4,420 g) and washed cells were suspended in 10 mL of 50 mM Mes pH 7.0 buffer and disrupted by triple passage through

French pressure cell at 150 bar (American Instruments, Silver Springs, USA). After removal of cell debris (4°C, 30 min; 20,000 g) and filtration through a 1.2 µm cellulose-acetate syringe filter (Sartorius, Göttingen, Germany). The pre-treated cell lysate (350 mg; 8 mL) was loaded on a 14 mL copper-loaded affinity column (XK 16/20, GE Healthcare, Little Chalfont, U.K.), beforehand equilibrated with buffer A (50 mM Tris HCl, pH 7.0 containing 300 mM NaCl) and mounted on an BioLogic DuoFlow™ system (Biorad, Hercules, California USA). Differential elution was performed with buffer B (50 mM Tris HCl, pH 7.0 containing 300 mM NaCl and 400 mM imidazole) at 10°C and a flow rate of 3 ml/min. The elution protocol comprised 5 steps, whereas concentration of buffer B was stepwise increased to 0 %, 10 %, 30 %, 60 % and 100 % in a volume of 100 mL, 90 mL, 60 mL, 30 mL and 75 mL, respectively. All buffers were degassed and filtered using 0.45 µm cellulose-acetate and 0.2 µm polyamide filters. Protein elution was monitored spectrophotometrically at 280 nm, and collected fractions were assayed for protein concentration and phosphatase activity against *p*-nitrophenyl phosphate (pNPP; see Assays). Active fractions were pooled and buffer was exchanged to buffer A_2 (50 mM KH_2PO_4-buffer, pH 7.0) with Amicon Ultra-15 Centrifugal Filter Units (Millipore, Billerica, USA). Pooled fractions were further purified using a 50 mL Fractogel EMD-DEAE column (XK 26/20, Merck, Darmstadt, Germany) mounted on a BioLogic DuoFlow™ system. Protein elution was performed with buffer B_2 (50 mM KH2PO4, 1M NaCl, pH 7.0) applying a five-step purification protocol at a constant flow rate of 5 mL/ min: step 1, isocratic flow of 100% buffer A_2 for 50 mL; step 2, linear increase to 15% buffer B_2 in 40 mL; step 3, isocratic flow of 85 % buffer A_2 and 15 % buffer B_2 for 80 mL; step 4, linear increase to 100 % buffer B_2 in 180 mL; step 5, isocratic flow of 100 % buffer B_2 for 70 mL. Active fractions were pooled and buffer was exchanged to 50 mM Hepes pH 7.0 with

Amicon Ultra-15 Centrifugal Filter Units. Protein purification was monitored by SDS PAGE and phosphatase activity measurements.

2.3.3 Molecular cloning, expression and purification of Z_{basic2}-HAD4

The HAD4 gene was amplified from pCA24N-yihx by PCR using Phusion DNA polymerase (Thermo Scientific, Waltham, USA) and the following pair of oligonucleotide primers:

5'-GAAGCTCTGTTCCAGGGTCCGCTCTATATCTTTGATTTAG-3', and

5'-CTTTGTTAGCAGCCGGATCTCTTAGCATAACACCTTCG-3'.

PCR consisted of a pre-heating step at 98°C for 5 min and was followed by 30 reaction cycles of denaturation at 98°C for 30 s, annealing at 70°C for 30 s, and elongation at 72°C for 1.5 min. The final extension step was carried out at 72°C for 5 min. The resulting PCR product comprised 5'- and 3'- overhangs complementary to the target sequence of the pT7ZbQGKlenow destination vector [15]. Vector pT7ZbQGKlenow was amplified by PCR using Phusion DNA polymerase and the following pair of oligonucleotide primers:

5'-CGGACCCTGGAACAGAGC-3', and

5'-GAGATCCGGCTGCTAACAAAG-3'.

PCR protocol was the same as above, except that elongation at 72°C was extended to 3 min. Amplification products of the two PCRs were individually subjected to parental template digest by *Dpn*I and then recombined via Gibson assembly [16]. The resulting construct encodes a fusion protein where the small module Z_{basic2} is joined by a flexible linker (11

amino acids) to the N-terminus of HAD4. Sequenced plasmid vector encoding the Z_{basic2}_HAD4 fusion protein was transformed into electrocompetent cells of *E. coli* BL21-Gold (DE3), and single-colony transformants were selected on agar plates containing 0.05 mg/mL kanamycin.

Protein production was done as described above, except for the use of 0.05 mg/mL kanamycin. Harvested cells were resuspended in the 10 ml of buffer A_3 (50 mM MES buffer, pH 7.0, containing 100 mM NaCl). Preparation of cell extract was done exactly as already described.

Protein purification was performed at 10°C using 2 x 5 ml HiTrap SP FF column (16/25 mm, GE Healthcare, Uppsala, Sweden) mounted on an ÄKTA prime plus (GE Healthcare, Uppsala, Sweden) system. The column was equilibrated with buffer A_3 and 8 ml of the pre-treated cell extract were loaded onto the column. Protein elution was performed using a continuous salt gradient ranging from 0% to 100% of buffer B_3 (50 mM HEPES buffer supplemented with 2 M NaCl, pH 7.0) in 150 mL at a flow rate of 4 mL/min. The target protein eluted at 50% buffer-B, corresponding to a salt concentration of 1 M NaCl. Pooled fractions were collected and buffer was exchanged to 50 mM MES, pH 7.0, containing 250 mM NaCl and 50 mM $MgCl_2$, as described before. The concentrated preparation was aliquoted and stored at -20°C. Protein purification was monitored by SDS PAGE and phosphatase activity measurements.

2.3.4 Assays

Protein was measured with Roti-Quant reagent (Roth, Karlsruhe, Germany) referenced against BSA. Purified Z_{basic2}_HAD4 was measured by UV absorbance at 280 nm using a molar extinction coefficient of 25440 $M^{-1}cm^{-1}$, calculated from the sequence with ProtParam analysis tool from the expasy server [17]. Standard phosphatase activity was measured with pNPP (4.0 mM), the release of p-nitrophenol (pNP) was recorded continuously at 405 nm with a Beckman Coulter DU 800 UV/VIS spectrophotometer (Krefeld, Germany).

Reaction was performed at 30°C using 50 mM MES, pH 7.0, containing 25 mM $MgCl_2$ and 100 mM NaCl. One unit of activity (U) is the amount of enzyme releasing 1 µmol of pNP under the conditions used. Specific activity is expressed in U/mg protein.

Inorganic phosphate was measured with a spectrophotometric assay described by Saheki et al [18]. Shortly, free phosphate forms a complex with molybdate and the phosphomolybdate complex is reduced by ascorbic acid in the presence of Zn^{2+}, which produces a chromophore that can be measured at 850 nm.

Reported coupled enzyme assays were used to measure αGlc 1-*P* and Glc 6-*P* [19]. In contrast to Eis et al no trehalose phosphorylase was present in our reaction mixture. Glucose 6-*P* dehydrogenase (G6PDH) from *Leuconostoc mesenteroides* (G8404; Sigma Aldrich) and α-phosphoglucomutase (PGM) from rabbit muscle (P3397; Sigma Aldrich) were added sequentially. After addition of G6PDH or PGM, formation of NADH was monitored at 340 nm and the NADH end concentration was used to calculate the corresponding Glc 6-*P* or αGlc 1-*P* concentration.

2.3.5 Kinetic characterization.

All reactions were performed in 50 mM MES buffer, pH 7.0, containing 25 mM $MgCl_2$ and 100 mM NaCl. Incubations were done at 37°C using agitation at 650 rpm on a Thermomixer comfort (Eppendorf, Hamburg, Germany.). αGlc 1-*P* or Glc 6-*P* was used as substrate in the concentration range 0.1 - 30 mM and 0.1 - 700 mM, respectively. To account for the lower activity of Z_{basic2}_HAD4 with Glc 6-*P* as compared to αGlc 1-*P*, the corresponding enzyme concentration was 0.39 µM and 16 µM, respectively. Reactions were started by addition of enzyme and allowed to run for up to 75 min while taking samples every 5 min. Heating (99°C, 5 min) was used to stop reactions. Free phosphate concentration was measured in supernatants of centrifuged samples.

It was confirmed that phosphate release was linear with incubation time. Phosphate concentration after 15 min or 120 min of reaction were used to determine the initial rate for Glc 1-P and Glc 6-P conversions, respectively. Phosphate blanks from reactions lacking enzyme were recorded.

Kinetic parameters were obtained from non-linear least squares fits (SigmaPlot, Systat, Erkrath, Germany) of initial rates to Michaelis-Menten equation expanded to include substrate inhibition (1). V is the initial rate, V_{max} is the maximum initial rate, K_M is the Michaelis constant, K_{iS} is the substrate inhibition constant, and [S] is the initial substrate concentration. Under conditions where $K_{iS} < K_M$ (see later under Results) fitting of equation 1 did not succeed and individual determination of all three parameters was not possible. However, V_{max}/K_m could be determined from the initial part of the V-[S] curve where V is linearly dependent on [S].

$$V = \frac{V_{max}\, S}{K_M + S\,(1+\frac{S}{K_{iS}})} \quad (1)$$

2.3.6 Substrate-selective hydrolysis of αGlc *1-P* in the presence of Glc 6-P

Reaction conditions were the same as above under Kinetic characterization. The substrate solution contained 20 mM of each αGlc 1-*P* and Glc 6-*P*. Reaction was started with 3.2 µM Z_{basic2}_HAD4. Samples were taken every 15 min and processed as already described. Samples were assayed for αGlc 1-*P*, Glc 6-*P*, and phosphate.

2.4 Results and discussion

2.4.1 Functional expression and purification of HAD4

Kuznetsova et al. reported isolation of µg amounts of His_6-HAD4 using automated high-throughput work-up procedures performed at µL scale [20]. We attempted to scale up the purification of His_6-HAD4 to a processing capacity of about 400 mg total protein from the crude *E. coli* cell extract. Contrary to the earlier study where Ni-NTA beads were used, Cu^{2+} loaded immobilized metal affinity chromatography in 14 mL column format (16 × 20 cm) was used herein. Even though about one-third of the applied protein was bound to the column, differential elution with an imidazole step gradient did not result in fractionation of protein(s) exhibiting clear phosphatase activity toward pNPP substrate. Protein pools potentially containing phosphatase activity showed a prominent band of correct size (25 kDa) next to

other major contaminants in the SDS polyacrylamide gel (Supporting Figure S1). They were therefore further fractionated by anion exchange chromatography. Collected fractions showed only trace amounts of phosphatase activity, and their analysis by SDS PAGE indicated that purification had not been successful (Supporting Figure S1).

To establish a more efficient procedure of isolation of HAD4, we chose not to optimize protocols for His_6-HAD4, but opted for an alternative fusion protein approach. A chimeric enzyme termed Z_{basic2}_HAD4 was constructed where the so-called Z_{basic2} module was attached to the N-terminus of HAD4. Z_{basic2} is an engineered arginine-rich variant of the Z domain, a 58 amino acid (7 kDa) three-helix bundle obtained from the B domain of staphylococcal protein A [21]. Fusion to Z_{basic2} was known from previous studies to enhance solubility of proteins otherwise poorly soluble under conditions of recombinant production in *E. coli*. Z_{basic2} was furthermore used as affinity module for protein purification by cation exchange chromatography. Finally, it was also applied for non-covalent immobilization of different enzymes on negatively charged carrier surfaces [22, 23] .

Z_{basic2}_HAD4 was isolated by cation exchange chromatography of the *E. coli* cell extract on HiTrap SP FF resin. Purified protein exhibited the expected size of 31 kDa. Capture of target protein, from the complex protein matrix and its purification to apparent homogeneity (Figure 1A) were combined efficiently in a single work-up step. Z_{basic2}_HAD4 eluted from the column as discrete protein peak at high salt concentration, as shown in Figure 1B. Analysis by SDS PAGE (Figure 1A) revealed that the flow-through fraction did not contain Z_{basic2}_HAD4, demonstrating that all of the applied target protein had bound to the column. Figure 1A also shows that insoluble protein recovered from the *E. coli* cells contained significant amounts of

Z_{basic2}_HAD4, indicating that fusion to Z_{basic2} was not fully sufficient to prevent aggregation of HAD4 during recombinant production in *E. coli*. About 11 mg of pure Z_{basic2}_HAD4 were recovered from each L of bacterial cell culture. Purification with HiTrap SP FF resin was extremely effective, considering that the total protein amount was almost 270 mg/L. The specific activity of Z_{basic2}_HAD4 was determined as 0.06 U/mg.

Stock solutions of purified Z_{basic2}_HAD4 were prepared by concentrating desalted protein in 50 mM MES buffer, pH 7.0, to about 10 mg/mL. After frozen storage at –20 °C, we observed massive protein precipitation in thawed samples. Addition of 250 mM NaCl resulted in partial re-solubilization of the precipitate concomitant with recovery of some of the original phosphatase activity in solution, suggesting effect of ionic strength on solubility and stability of Z_{basic2}_HAD4. Supplementing the storage buffer with a combination of salts (250 mM NaCl, 50 mM $MgCl_2$) prevented precipitation of Z_{basic2}_HAD4 completely and enzyme activity was stable over weeks of frozen storage. Tentative explanation for the pronounced tendency of Z_{basic2}_HAD4 to form insoluble aggregates is the large difference in the calculated isoelectric points for Z_{basic2} (pI = 11.5) and HAD4 (pI = 5.18). Intramolecular and intermolecular domain-domain interactions driven by charge complementarity might promote a native-like protein aggregation that is partly reversible at high salt concentration. Mg^{2+} ions could have two effects in this context. First of all, Mg^{2+} is an active-site ligand of HAD4. Structure of the catalytic centre (and perhaps that of the whole protein) might be stabilized when Mg^{2+} is bound. Secondly, adsorption of Z_{basic2} fusions to negatively charged surfaces is strongly weakened in the presence of Mg^{2+} ions. Capacity of Mg^{2+} to perturb charged interactions of the Z_{basic2} module are much more pronounced than that of Na^{+} (each as chloride salt) [24].

2.4.2 Kinetic analysis of hydrolysis of αGlc 1-*P* and Glc 6-*P* catalyzed by Z_{basic2}_HAD4.

Initial reaction rates at steady state were obtained from time-course measurements of substrate consumption and phosphate release. Disparity of initial rates determined from the different measurements would have indicated a phosphoryl transfer side reaction occurring next to the hydrolysis. Both the phosphorylated substrate and the glucose product could in principle function as acceptor of the phosphoryl group and thereby compete with water in the overall reaction. However, it was confirmed that initial rates of *substrate utilization* and *substrate hydrolysis* were identical within limits of experimental error.

Figure 2 shows initial rates determined at different substrate concentrations for hydrolysis of αGlc 1-*P* (Figure 2 A) and Glc 6-*P* (Figure 2 B). For each substrate, the *V*-[S] curve increased at low [S] to reach a distinct maximum value of *V* at intermediate [S], only to decrease again in the high-[S] range. The kinetic behavior indicates inhibition of Z_{basic2}_HAD4 by the phosphorylated substrate. It is impractical to speculate about the mechanism underlying the substrate inhibition, but clearly there must be an enzyme form occurring in the normal reaction that is capable of binding a second substrate molecule. The resulting complex may be inactive, so that active enzyme can only be regenerated by dissociation, or it may turn over at slower rate than would the uncomplexed enzyme. The results emphasize that phosphatases are not only subject to inhibition by inorganic phosphate [25], but can also be inhibited strongly by their (preferred) substrate.

αGlc 1-*P* differed from Glc 6-*P* in affinity and reactivity for Z_{basic2}_HAD4-catalyzed hydrolysis. Maximum value of *V* was observed at around 10 mM for αGlc 1-*P* while for Glc

6-*P* the corresponding substrate concentration was about 100 mM (Figure 2). Substrate inhibition also occurred at lower concentrations of αGlc 1-*P* as compared to Glc 6-*P*. An apparent k_{cat} was calculated from the maximum *V* value ($k_{cat\_app} = V_{max\_app}/[E]$), and αGlc 1-*P* (10.24 s^{-1}) was about 110-fold more active in k_{cat_app} terms than Glc 6-*P* (0.09 s^{-1}). By way of comparison, k_{cat_app} for hydrolysis of pNPP (4 mM) was 0.032 s^{-1}.

Equation 2 could be used for data fitting by nonlinear least squares regression, as shown in Figure 2; however, the resulting kinetic parameters (k_{cat}, K_M, K_{iS}) were poorly defined statistically. Covariance analysis revealed strong interdependence of all parameters, but especially of K_M and K_{iS}, implying that independent determination of each single parameter was not possible from the data available. An approximate ratio between K_M/K_{iS} can be given however, and this was 21 (60 ± 112 / 2.82 ± 5.41) for αGlc 1-*P* and 0.15 (51 ± 13 / 351 ± 80) for Glc 6-*P*. To nevertheless obtain parameter of enzyme specificity, we determined k_{cat}/K_M from the part of the *V*-[S] curve where *V* was linearly dependent on [S]. The relationship $V = (k_{cat}/K_M)$ [S] [E] holds under these conditions. k_{cat}/K_M for hydrolysis of αGlc 1-*P* (1.87 ± 0.03 $mM^{-1}s^{-1}$) exceeded the k_{cat}/K_M for hydrolysis of Glc 6-*P* (0.0032 ± 8 × 10^{-5} $mM^{-1}s^{-1}$) by 565-fold. Z_{basic2}_HAD4 strongly discriminates between αGlc 1-*P* and Glc 6-*P* as substrate for hydrolysis.

Despite significant variation in experimental condition, it is still interesting to compare kinetic observations made herein for Z_{basic2}_HAD4 with data reported by Kuznetsova et al. on His_6-HAD4. First of all, substrate inhibition was not mentioned in the study of His_6-HAD4. k_{cat}/K_M agrees for the two enzymes within a factor of 3.3, even though k_{cat_app} of Z_{basic2}_HAD4 (10.24 s^{-1}) exceeds the k_{cat} of His_6-HAD4 (1.4 s^{-1}) by an order of magnitude. Considering

applications of HAD4 in biocatalysis, the elevated turnover number could be a decisive advantage. Findings of this study on Z_{basic2}_HAD4 are consistent with Kuznetsova et al. working with His_6-HAD4 that αGlc 1-*P* is strongly preferred substrate for hydrolysis by HAD4. Other phosphorylated substrates exhibit very much lowered reactivity. In addition to the Glc 6-*P* examined here and the pNPP used in both studies, Kuznetsova et al. have applied a series of other phosphorylated compounds, including several sugar 1-phosphates other than αGlc 1-*P*, and showed that they were all hardly active in comparison to αGlc 1-*P*.

2.4.3 Substrate-selective hydrolysis of αGlc 1-*P* in a mixture of αGlc 1-*P* and Glc 6-*P*

High selectivity of Z_{basic2}_HAD4 for hydrolyzing αGlc 1-*P* as compared to Glc 6-*P*, as implied by the ratio of the corresponding k_{cat}/K_M values, was put to critical test in a biotransformation experiment in which both sugar phosphates were offered as substrate, each at 20 mM. Time course of the enzymatic reaction is shown in Figure 3. αGlc 1-*P* was rapidly depleted from the mixture, resulting in complete conversion of this substrate after about 70 min. The concentration of Glc 6-*P* did not change in the same time span. The αGlc 1-*P* conversion rate in the mixed sugar phosphate substrate was lowered about 2-fold in comparison to the rate expected from the initial rate measurements performed in the absence of Glc 6-*P* (Figure 2). It is known that two competing substrates act kinetically as if they were competitive inhibitors one of another, that is, they affect their respective counterpart substrate's K_M. A Glc 6-*P* concentration roughly half of its K_M in enzymatic hydrolysis ($K_M \approx 50 \pm 13$ mM) is expected to cause about 1.5-fold elevation in the apparent K_M for αGlc 1-*P*, according to the relation

$K_{M\_app}^{Glc\ 1\text{-}P} = K_M^{Glc\ 1\text{-}P}\ (1 + [Glc\ 6\text{-}P]/K_M^{Glc\ 6\text{-}P})$. Effect of Glc 6-*P* as inhibitor of αGlc 1-*P* utilization is therefore explained.

Z_{basic2}_HAD4-catalyzed hydrolysis resulted in complete removal of αGlc 1-*P* while the Glc 6-*P* present was retained fully at the same time, thus providing effective mean for separation of the two sugar phosphates. Specialized separation procedures such as borate complex ion exchange chromatography [26] would otherwise be needed to isolate the Glc 6-*P*. In the case of the alternative separation problem to remove Glc 6-*P* from αGlc 1-*P*, another HAD-like phosphatase, namely the enzyme BT4131 from *Bacteroides thetaiotaomicron*, could be of interest [12]. The phosphatase is highly active in hydrolysis of hexose 6-phosphates and pentose 5-phosphates, but exhibits barely any activity with hexose 1-phosphates. Taken together, these results suggest that HAD-like phosphatases present a potentially valuable source of selective sugar phosphate hydrolases that could be applied to resolution of mixtures of sugar phosphates that are otherwise very difficult to separate. In addition to phosphatases that exhibit broad substrate spectrum and are already available commercially, development of a toolbox of selective phosphomonoester hydrolases could provide interesting opportunities for biocatalysis.

2.5 Conclusions

First-time functional expression, isolation, and characterization of *E. coli* HAD4 as fusion protein harboring the charged module Z_{basic2} at the N-terminus is reported. Single-step capture and polishing of Z_{basic2}_HAD4 by cation exchange chromatography was notably efficient to obtain multi-mg amounts of highly purified enzyme. Kinetic evidence and results from mixed

sugar phosphate conversion experiments reveal that Z_{basic2}_HAD4 is highly selective for hydrolysis of αGlc 1-*P*. Z_{basic2}_HAD4 is proposed to be a useful catalyst for hydrolytic transformation (e.g. removal) of αGlc 1-*P* from complex substrate solutions containing multiple sugar phosphates. Z_{basic2}_HAD4 complements the current portfolio of phosphatases for use in biocatalysis by adding new characteristics of specificity and selectivity.

2.6 Acknowledgement

We would like to acknowledge financial support from the Austrian Science Funds FWF (project L586-B03); Alexander Yakunin (Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada) for plasmid vector pCA24N-yihx; and Sophia Hober (Department of Biotechnology, Alba Nova University Center, Stockholm, SE) for the plasmid vector containing the Z_{basic2} gene.

2.7 References

[1] I. Nobeli, H. Ponstingl, E.B. Krissinel, J.M. Thornton, J. Mol. Biol., 334 (2003) 697-719.
[2] K. Papenfort, Y. Sun, M. Miyakoshi, C.K. Vanderpool, J. Vogel, Cell, 153 (2013) 426-437.
[3] T. Hisano, Y. Hata, T. Fujii, J.Q. Liu, T. Kurihara, N. Esaki, K. Soda, J. Biol. Chem., 271 (1996) 20322-20330.
[4] I.S. Ridder, H.J. Rozeboom, K.H. Kalk, D.B. Janssen, B.W. Dijkstra, J. Biol. Chem., 272 (1997) 33015-33022.
[5] E.V. Koonin, R.L. Tatusov, J. Mol. Biol., 244 (1994) 125-132.
[6] K.N. Allen, D. Dunaway-Mariano, Curr. Opin. Struct. Biol., 19 (2009) 658-665.
[7] A.M. Burroughs, K.N. Allen, D. Dunaway-Mariano, L. Aravind, J. Mol. Biol., 361 (2006) 1003-1034.
[8] S.D. Lahiri, G. Zhang, D. Dunaway-Mariano, K.N. Allen, Biochemistry, 41 (2002) 8351-8359.
[9] M.C. Morais, W. Zhang, A.S. Baker, G. Zhang, D. Dunaway-Mariano, K.N. Allen, Biochemistry, 39 (2000) 10385-10396.
[10] H.H. Nguyen, L.B. Wang, H. Huang, E. Peisach, D. Dunaway-Mariano, K.N. Allen, Biochemistry, 49 (2010) 1082-1092.
[11] S.D. Lahiri, G.F. Zhang, J.Y. Dai, D. Dunaway-Mariano, K.N. Allen, Biochemistry, 43 (2004) 2812-2820.
[12] Z.B. Lu, D. Dunaway-Mariano, K.N. Allen, Biochemistry, 44 (2005) 8684-8696.
[13] G. Zhang, J. Dai, L. Wang, D. Dunaway-Mariano, L.W. Tremblay, K.N. Allen, Biochemistry, 44 (2005) 9404-9416.
[14] T. van Herk, A.F. Hartog, A.M. van der Burg, R. Wever, Adv. Synth. Catal., 347 (2005) 1155-1162.
[15] M. Hedhammar, S. Hober, J. Chromatogr. A, 1161 (2007) 22-28.
[16] D.G. Gibson, L. Young, R.Y. Chuang, J.C. Venter, C.A. Hutchison, 3rd, H.O. Smith, Nat. Methods, 6 (2009) 343-345.
[17] E. Gasteiger, C. Hoogland, A. Gattiker, D. S., M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, in: J.M. Walker (Ed.) The Proteomics Protocols Handbook, Humana Press, 2005, pp. 571-607.
[18] S. Saheki, A. Takeda, T. Shimazu, Anal. Biochem., 148 (1985) 277-281.
[19] C. Eis, B. Nidetzky, Biochem. J, 341 (1999) 385-393.
[20] E. Kuznetsova, M. Proudfoot, C.F. Gonzales, G. Brown, M.V. Omelchenko, I. Borozan, L. Carmel, Y.I. Wolf, H. Mori, A.V. Savchenko, C.H. Arrowsmith, E.V. Koonin, A.M. Edwards, A.F. Yakunin, J. Biol. Chem., 281 (2006) 36149-36161.
[21] T. Graslund, G. Lundin, M. Uhlen, P.A. Nygren, S. Hober, Protein Eng., 13 (2000) 703-709.
[22] J.M. Bolivar, B. Nidetzky, Biotechnol. Bioeng., 109 (2012) 1490-1498.
[23] J. Wiesbauer, J.M. Bolivar, M. Mueller, M. Schiller, B. Nidetzky, ChemCatChem, 3 (2011) 1299-1303.
[24] J.M. Bolivar, B. Nidetzky, Langmuir, 28 (2012) 10040-10049.
[25] S.P. Coburn, J.D. Mahuren, M. Jain, Y. Zubovic, J. Wortsman, J. Clin. Endocrinol. Metabol., 83 (1998) 3951-3957.
[26] M.J. Lefebvre, N.S. Gonzalez, H.G. Pontis, J. Chromatogr., A, 15 (1964) 495-500.

2.8 Legend to Schemes and Figures

Figure 1. (A) Purification of Z_{basic2}_HAD4 followed by SDS-PAGE. 1: PAGE-ruler protein ladder, 2: Cell extract (1.5 µg), 3: Insoluble protein fraction from *E. coli* expression culture (1.5 µg), 4: Flow-through (proteins that did not bind to the column) (1.5 µg), 5: Fraction with minor phosphatase activity (1.0 µg), 6: Fraction with minor phosphatase activity (0.5 µg), 7: Purified Z_{basic2}_HAD4 (1.5 µg); (**B)** Chromatogram of Z_{basic2}_HAD4 purification performed with 2 × 5 mL HiTrap SP FF column. The absorbance trace at 280 nm is shown as solid line. The salt gradient used is shown as dashed line, and the measured conductivity is shown as dotted line.

Figure 2. Z_{basic2}_HAD4 exhibits pronounced substrate inhibition. (A) Michaelis-Menten plot for the hydrolysis of αGlc 1-*P*. (B) Michaelis-Menten plot for the hydrolysis of Glc 6-*P*. See the Methods for details.

Figure 3. Selective hydrolysis of αGlc 1-*P* by Z_{basic2}_HAD4 in a reaction mixture containing 20 mM αGlc 1-*P* (●) and Glc 6-*P* (○). The buffer was 50 mM MES, pH 7.0. See the Methods for details.

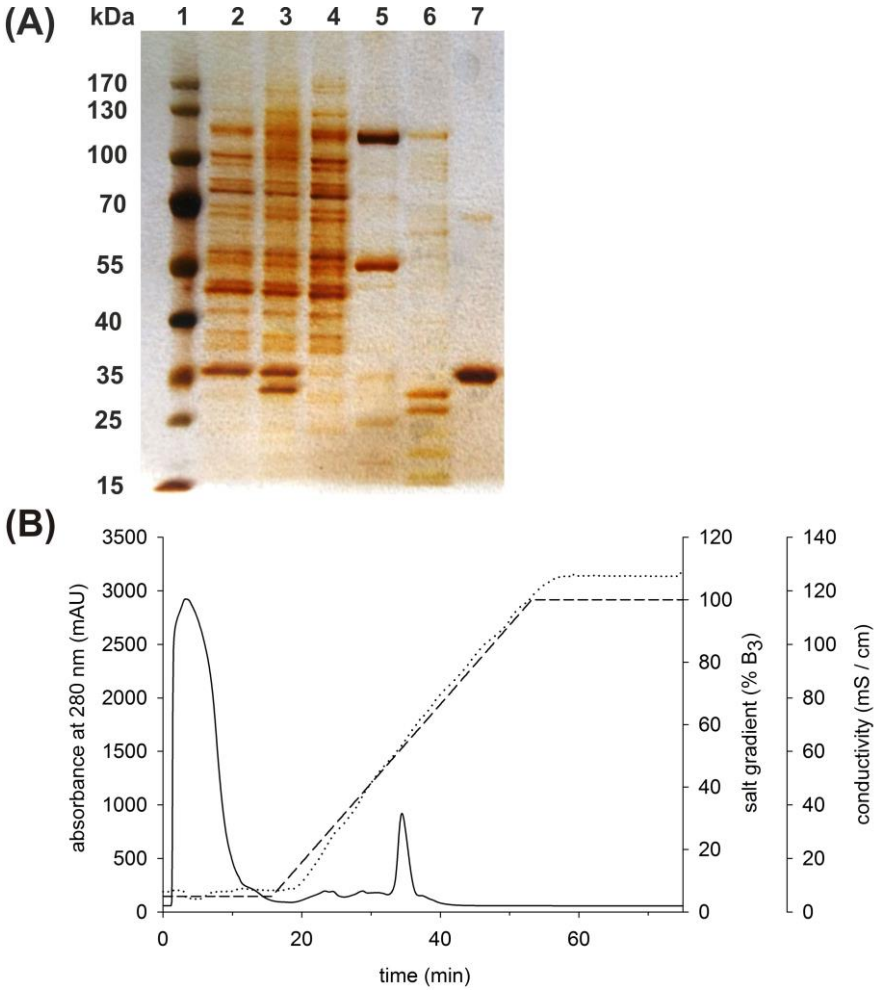


Figure 1. (A) Purification of Z_{basic2}_HAD4 followed by SDS-PAGE. 1: PAGE-ruler protein ladder, 2: Cell extract (1.5 µg), 3: Insoluble protein fraction from *E. coli* expression culture (1.5 µg), 4: Flow-through (proteins that did not bind to the column) (1.5 µg), 5: Fraction with minor phosphatase activity (1.0 µg), 6: Fraction with minor phosphatase activity (0.5 µg), 7: Purified Z_{basic2}_HAD4 (1.5 µg); (**B)** Chromatogram of Z_{basic2}_HAD4 purification performed with 2 × 5 mL HiTrap SP FF column. The absorbance trace at 280 nm is shown as solid line. The salt gradient used is shown as dashed line, and the measured conductivity is shown as dotted line.

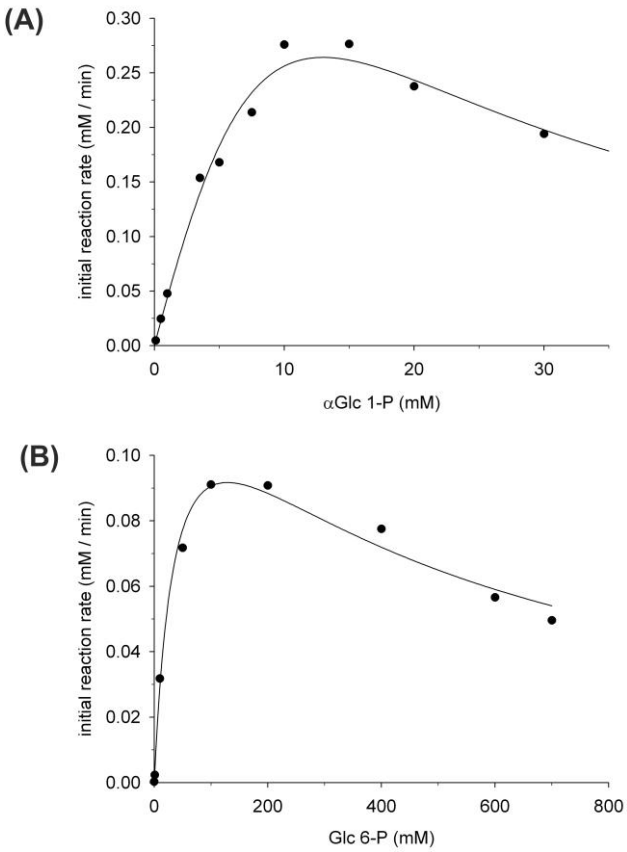


Figure 2. Z_{basic2}_HAD4 exhibits pronounced substrate inhibition. (A) Michaelis-Menten plot for the hydrolysis of αGlc 1-*P*. (B) Michaelis-Menten plot for the hydrolysis of Glc 6-*P*. See the Methods for details.

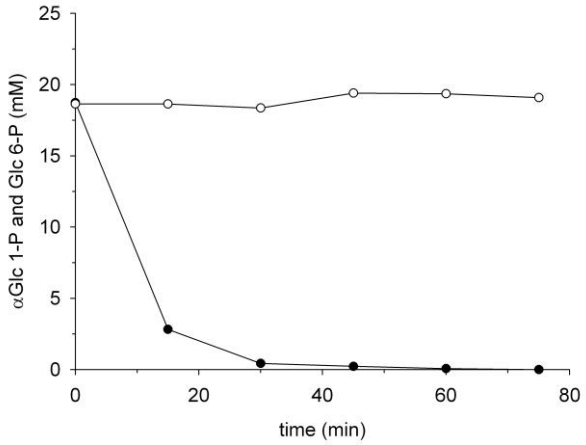
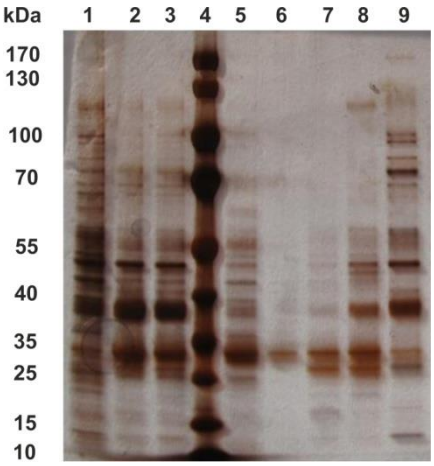


Figure 3. Selective hydrolysis of αGlc 1-*P* by Z_{basic2}_HAD4 in a reaction mixture containing 20 mM αGlc 1-*P* (●) and Glc 6-*P* (○). The buffer was 50 mM MES, pH 7.0. See the Methods for details.

2.9 Supplementary information



Supplementary Figure S1. His_6_HAD4 purification was followed by SDS-PAGE. 1: Filtered *E. coli* crude extract loaded onto the Cu^{2+}-chelate column (2 µg), 2-3: Fractions after Cu^{2+}-chelate column purification containing phosphatase activity loaded onto the DEAE-anion exchange column (2 µg each), 4: Page-Ruler protein ladder, 5: Flow-through (proteins that did not bind to the column) (1 µg), 6-9: Fractions of individual peaks which eluted at different concentrations of elution buffer during purification using the DEAE-anion exchange column (0.2 - 1 µg). Expected size of His_6_HAD4: 25 kDa.

3 Appendix:

Analysis of active-site residues in αG1Pase: effects of site-directed substitutions of the nucleophile His18, the general acid-base Asp290, and the putative “gating” residue Glu196

Martin Pfeiffer[1], Patricia Wildberger[1] and Bernd Nidetzky[1,2*]

[1]Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

[2] acib - Austrian Centre of Industrial Biotechnology, Petersgasse 14, A-8010 Graz, Austria

* Corresponding author; e-mail: bernd.nidetzky@tugraz.at

Running title: Substrate-selective sugar phosphate hydrolysis by HAD

3.1 Materials and methods.

3.1.1 Chemicals and reagents.

Unless stated otherwise, all chemicals were of highest purity available from Sigma-Aldrich (Vienna, Austria) or Roth (Karlsruhe, Germany). Oligonucleotide synthesis was performed by Sigma-Aldrich (Vienna, Austria). DNA sequencing was performed at LGC Genomics (Berlin, Germany).

3.1.2 Molecular cloning.

The desired mutations, H18→A, D290→A and E196→A, were introduced to the *Strep*-tagged wild-type (WT) α-D-glucose 1-phosphatase (αG1Pase) fusion protein by using a previously described two-stage PCR protocol [1]. pMS470_dsbC_ αG1Pase was used as template; the following pairs of oligonucleotide primers were used for introducing each of the desired substitutions (mutant triplet codon is underlined).

H18A fwd: 5´-CATGATGAGCCGCGCCAACTTACGTGCGCCGCTG-3´

H18A bwd: 5´-CAGCGGCGCACGTAAGTTGGCGCGGCTCATCATG-3´

D280A fwd: 5´-GTTGGGCACGCCTCCAACATTG-3´

D290A bwd: 5´-GAACCCGTGCGGAGGTTGTAAC-3´

E196A fwd: 5´-GTATCAACAAGCCCCAGGTG-3´

E196A bwd: 5´-CATAGTTGTTCGGGGTCCAC-3´

In the first step of the two-step PCR protocol, two separate PCR reactions, using either the forward or the backward primer, were performed for each desired mutation. The reaction profile consisted of a preheating step at 98°C for 30 sec followed by 6 cycles (98°C, 15 sec/60-70°C, 20 sec/72°C, 180 sec/72°C, 300 sec). In the second PCR reaction, the first PCR reaction products were combined 1:1 (v:v) and amplification was continued for 30 cycles (98°C, 15 sec/60-70°C, 20 sec/72°C, 180 sec/72°C, 420 sec). The final PCR product was subjected to parental template digest by *Dpn*I and transformed into chemically competent *E. coli* Top 10 cells and single-colony transformants were selected on agar plates containing 0.110 mg/mL ampicillin. Plasmids were amplified in *E. coli* Top 10 cells and subsequently the plasmids were isolated and sent for sequencing. Sequenced plasmids harboring the desired mutation were transformed into electro-competent *E. coli* Origami B cells, whereas single-colony transformants were selected as mentioned above.

3.1.3 Expression and purification.

E. coli Origami B cells harboring one of the four variants of pMS470_dsbC_αG1Pase (WT/H18A/D290A/E196A) were cultivated and harvested as described previously [2]. αG1Pase variants were purified from the crude extract using *Strep*-Tactin Sepharose column (IBA, Göttingen, Germany), like described previously [3].

3.1.4 Assays.

Phosphatase activities were quantified by using a continuous *para*-nitrophenylphosphate hydrolysis assay based on a spectrophotometric measurement at 405 nm on a Beckman

Coulter DU 800 UV/VIS spectrophotometer (Krefeld, Germany). Reaction mixtures used for assessing the activity of αG1Pase contained 4 mM *para*-nitrophenylphosphate in 50 mM MES, pH 7.0 at 30°C. The total protein concentration was determined as described by Bradford [4] employing Roti-Quant reagent (Roth, Karlsruhe, Germany) referenced against BSA.

Inorganic phosphate was measured colorimetrically at 850 nm [5]. α-Glucose 1-phosphate (αGlc 1-*P*) and glucose 6-phosphate (Glc 6-*P*) were assayed as described in chapter 2 of this thesis [6] .

3.1.5 Hydrolysis study.

Two different substrates, *p*NPP and αGlc 1-*P*, were subject to hydrolysis by each of the αG1Pase variants (WT/H18A/D290A/E196A). Reactions were performed in 50 mM MES pH 7.0, containing 20 mM substrate. Incubations were done at 37°C using agitation at 650 rpm on a Thermomixer comfort (Eppendorf, Hamburg, Germany). To account for the varying activities of the different αG1Pase variants, the *p*NPP conversions were started with the addition of 66 nM, 5.6 µM, 66 nM, and 0.22 µM for WT, H18A, D290A and E196A, respectively. The αGlc 1-*P* conversions were started with the addition of 66 nM, 6.6 µM, 6.6 µM, and 66 nM for WT, H18A, D290A and E196A, respectively. Substrate hydrolysis was monitored discontinuously. Therefore samples were taken every 15 min over a period of 75 min. The reaction was stopped by heating (99°C, 5 min). Protein was removed by centrifugation at 20,000 g for 5 min and free phosphate concentration was measured in the supernatant. The apparent reaction rate k_{cat_app} was calculated using the relationship $k_{cat\_app} = V_{max}/[E]$, where *[E]* is the molar enzyme concentration based on protein

concentration measurements according to Bradford. V_{max} is the maximum velocity based on the phosphate release rate.

3.1.6 Analysis performed with H18A.

Chemical rescue studies. Reactions were performed in 50 mM MES buffer, pH 7.0 containing 15 mM αGlc 1-*P* as substrate. Incubations were performed at 37°C using agitation at 650 rpm. Imidazole (20 mM), sodium azide (20 mM, 50 mM) or 20 mM sodium acetate were used as rescue agent. The reaction was started by addition of 6.6 µM H18A. αGlc 1-*P* conversion was monitored discontinuously. Samples were taken every 15 min over a period of 75 min. Samples were treated as mentioned above. Free phosphate, αGlc 1-*P* and Glc 6-*P* concentrations were measured. To exclude the possibility of phosphoryl group transfer to acceptors other than water (e.g. glucose produced in the reaction), we performed a conversion in 50 mM MES, pH 7.0, containing 50 mM sodium azide and 200 mM glucose as alternative phosphate acceptor. Protein concentration and sample treatment were the same as above.

3.1.7 Analysis performed with D290A.

Brønsted analysis. Reactions were performed in 50 mM MES, pH 7.0, at 37°C and 650 rpm agitation. The initial substrate concentration was 20 mM. Phenylphosphate, *para*-nitrophenyl phosphate (*p*NPP), pyrophosphate (PPi), sodium tripolyphosphate (PPP), napthylphosphate, phytate and αGlc 1-*P* were used as substrates. The reactions were started by addition of 6.6 µM D290A, except for the conversion with *p*NPP, in which 0.66 µM D290A were used. As control, conversions under the same conditions using 0.66 µM αG1Pase were performed. The reaction was allowed to proceed for 75 min and samples were taken every 15 min. Samples

were treated as mentioned previously and the free phosphate concentration was measured in all samples.

Phosphoryl transfer studies. Reactions were performed in 50 mM MES pH 7.0 containing 20 mM αGlc 1-*P* and 200 mM glucose at 37°C and 650 rpm. Incubation, sampling and sample preparation was performed exactly as mentioned above. Free phosphate, αGlc 1-*P* and Glc 6-*P* concentrations were measured. Phosphorylated products were measured indirectly via the phosphate balance (equation 1), whereas a non-closed balance indicated the production of a newly phosphorylated compound. The first section of the term Δ[αGlc 1-*P*] represents the phosphate donor substrate consumed. The second side of the term stands for the phosphorylated product produced, whereas [X-*P*] stands for the newly phosphorylated product.

(1) $\Delta[\alpha\text{Glc 1-}P] = [\text{X-}P] + [\text{Glc 6-}P] + [\text{free phosphate}]$

3.1.8 Analysis performed with E196A.

Phosphoryl transfer studies. Reactions were performed in 50 mM MES pH 7.0 containing 20 mM of αGlc 1-*P* and 200 mM L-ascorbic acid or guanosine as acceptor substrate at 37°C and 650 rpm. The reaction was started by addition of 0.1 µM E196A. Incubation, sampling and sample preparation was performed exactly as mentioned above. Free phosphate, αGlc 1-*P* and Glc 6-*P* concentrations were measured. Phosphorylated products were measured indirectly via the phosphate balance (equation 1).

3.1.9 Analysis of the substrate binding tunnel

Analysis was performed using the free software package MOLE2.0 [7]. PyMOL (http://pymol.sourceforge.net) was used for visualization.

3.2 Results

3.2.1 Selection of target amino acids.

His^{18} and Asp^{290} were chosen as mutation targets, because they are considered as the nucleophile and the general acid-base catalyst, respectively. In the available crystal structure of αG1Pase (PDB code: 1NT4) His^{18} is replaced by Ala. Therefore, we compared the active-site of αG1Pase with the active site of the closely related *E. coli* phytase *appA* (PDB code: 1DKO). This analysis showed that the phytase His^{17}, the homologous residue to αG1Pase His^{18}, is ideally positioned to act as the nucleophile [8]. Docking of αGlc 1-*P* into the active site of the G1Pase crystal structure showed that Asp^{290} is positioned in close proximity to the O1-*P* bond of αGlc 1-*P* [2]. This enables Asp^{290} to protonate the glucose leaving group in the first reaction step (Figure 1A), a catalytic reaction usually supported by the action of a general acid-base catalyst. To eliminate the functional side chain groups involved in catalysis, we created a H18A and a D290A mutant.

In addition to the nucleophile and the acid-base catalyst, we were interested in factors influencing the substrate specificity. Therefore we performed a detailed analysis of the αG1Pase substrate binding tunnel using the software MOLE 2.0 [7]. We could identify two tunnels originating from the active site (Figure 1B).

Tunnel 1 is approximately 16 Å long and involves Lys^{244}, Asn^{245}, Gln^{248}, Asp^{290}, Lys^{207}, Tyr^{247}, Leu^{24}, Ser^{291}, Glu^{196}, Asn^{26} and Asn^{27}. Around 14 Å from the inner tunnel origin a bottleneck can be found, in which the tunnel radius is decreased from 2.7 Å to 2.4 Å (Figure 1C). We identified Glu^{196} as an important participant in the formation of this bottleneck. Because of its negative charge at physiological pH, it has the potential to repel negative charged ligands. Therefore, we hypothesize that it might be a substrate determining factor, which prompted us to create an E196A mutant.

Tunnel 2 is approximately 22 Å long and involves Lys^{244}, Asn^{245}, Gln^{248}, Asp^{290}, Asn^{207}, Asp^{211}, Leu^{24}, Arg^{21}, Asn^{27}, Leu^{24}, Ser^{208}, Pro^{23}, Leu^{31} and Gly^{28}. Alongside tunnel 2, two bottlenecks can be identified (Figure 1D). The first bottleneck (at approximately 10 Å) is characterized by a narrowing of the tunnel radius from 2.4 Å to 1.5 Å. The second bottleneck (at around 18 Å) is characterized by a decrease in radius from 2.0 Å to 1.5 Å. We do not consider mutations alongside tunnel 2, because the two bottlenecks, with a minimal diameter of 3.0 Å, prohibit transport of αGlc 1-*P*, which has a diameter equal or greater 4.0 Å (measured in PyMOL). Due to the narrowness of tunnel 2 we conclude, that only very small molecules can travel alongside this tunnel. Thus, we speculate that tunnel 2 might be involved in the transport of a water molecule, which leads to hydrolysis of the covalent phospho-histidine intermediate in the second half reaction of the proposed reaction mechanism.

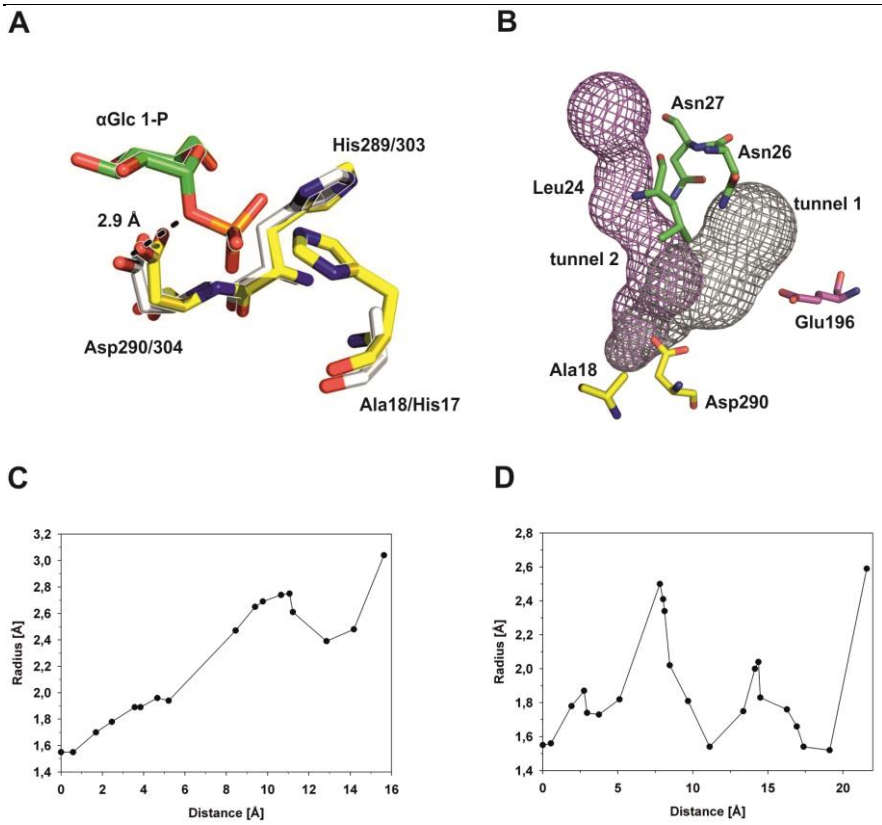


Figure 1. Structural analysis of the active-site and the binding pocket of αG1Pase. (A) Overlay of the active-site of αG1Pase (white) and *E. coli* phytase (yellow). αGlc 1-*P* was docked into the active site as described in chapter 1 of this thesis. (B) Two tunnels originate in the active site. Yellow residues represent residues of the catalytic-core; green residues are involved in building the bottleneck in tunnel 1. Glu^{196} (represented in magenta) is involved in building the bottleneck in tunnel 1 and represents a target for mutation to Ala. (C) Radius distance plot for tunnel 1 elucidates the presence of a bottleneck. (D) Radius distance plot for tunnel 2 elucidates the presence of two bottlenecks.

3.2.2 Molecular cloning, expression and purification of αG1Pase and mutants.

αG1Pase and mutants were expressed as *N*-terminal *Strep*-tag fusion proteins in *E. coli* Origami B (DE3) [2]. Purification was achieved via a *Strep*-tag [3] affinity chromatography and resulted in about 5.5 mg/L, 5.5 mg/L, 1.8 mg/L and 5.1 mg/L of pure H18A, D290A, E196A and αG1Pase, respectively. Purity of the enzyme preparation was verified via SDS-PAGE and silver staining (Figure 2). Protein preparations were stable and could be stored frozen (>3 mg/mL in 50 mM MES buffer pH 7.0) for several weeks without substantial loss of activity.

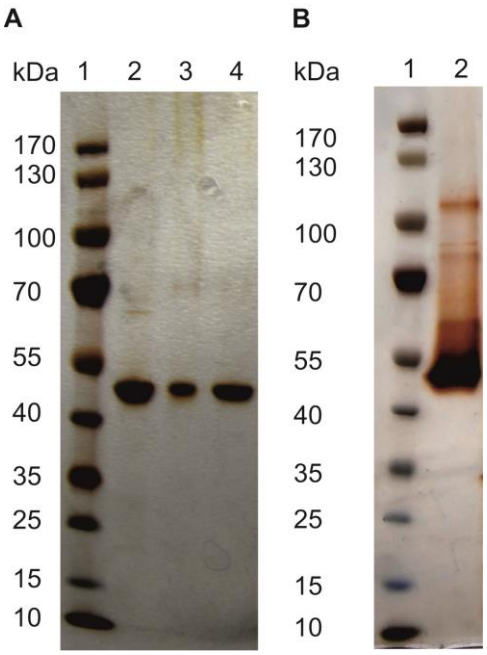


Figure 2. Purification of αG1Pase and mutants was monitored by SDS-PAGE. (A) 1: PAGE-ruler protein ladder; 2: purified H18A (2 µg); 3: purified E196A (1 µg); purified D290A (2 µg). (B) 1: PAGE-ruler protein ladder; 2: purified αG1Pase (5 µg).

3.2.3 Hydrolytic activity of αG1Pase.

Two different substrates, αGlc 1-*P* (20 mM) and *p*NPP (5 mM), were used to determine the enzymatic activity of αG1Pase. The results are summarized in (Table 1). The apparent phosphate release rates of αG1Pase are 44 s^{-1} for αGlc 1-*P* and 6.7 s^{-1} for *p*NPP. Both activities are at least 3-fold lower compared to the activity reported by Cotrill et al [9]. The differences might be explained by the fact that Cotrill et al performed hydrolysis studies at pH 4.5, whereas we used pH 7.0. Interestingly, the determination of *p*NPP hydrolysis rates varied up to a magnitude depending on the analytical method used. In general the discontinuous (Saheki) phosphate release measurement resulted in lower activities as compared to the continuous *para*-nitrophenyl (*p*NP) release measurement (405 nm).

Table 1. Apparent rate constants of αG1Pase and mutants (H18A, D290A, E196A) based on phosphate released.

Protein	k_{cat_app} [s^{-1}]		residual activity mutant/WT	
	αGlc 1-*P*	*p*NPP	αGlc 1-*P*	*p*NPP
αG1Pase	44	6.8	100%	100%
H18A	0.4	0.004*	1%	1%*
D290A	0.1	10	0.3%	150%
E196A	31	2.5	68%	36%

*k_{cat_app} was determined following release of *p*NP at 405 nm.

3.2.4 Characterization of H18A.

Hydrolytic activity of H18A. Apparent turnover frequencies (k_{cat_app}) for αGlc 1-*P* (20 mM) and *p*NPP (5 mM) were determined from phosphate release rates, or from release of *p*NP, respectively. The results are summarized in (Table 1).

H18A showed hydrolytic activities towards αGlc 1-*P* with a k_{cat_app} of 0.4 s^{-1}. The hydrolysis rate for *p*NPP was 0.004 s^{-1}. Compared to αG1Pase, H18A showed only 1% residual activity for both substrates. This drastic decrease of hydrolytic activity is consistent with the proposed role of His^{18} as the nucleophile, however, detection of residual activity is in contrast to the findings of Lee et al and Cortrill et al, which report complete loss of activity for their H18A mutant [9,10]. Residual activity might be explained by a minor contamination of αG1Pase resulting from co-purification of the WT-enzyme which might be expressed at a basal level by the expression host. An alternative explanation would be that His^{289} takes over the role as the nucleophile and forms a covalent phospho-histidine intermediate.

Restoring hydrolytic activity of H18A applying chemical rescue. In an attempt to partially restore the activity of H18A, we performed chemical rescue experiments introducing either azide (20 – 50 mM), imidazole (20 mM) or acetate (20 mM) as alternative nucleophiles. Apparent reaction rates were determined discontinuously measuring the decrease in αGlc 1-*P* concentration. As shown in Figure 3A, chemical rescue with 20 mM of any rescue agent did not restore the activity of H18A. A minor increase of activity, 15% compared to H18A without rescue agent, was observed when 50 mM azide was used as rescue agent. This

increase, however, may lie in the range of experimental failure, thus repetition of the experiment is recommended.

Phosphoryl transfer activity of H18A. To investigate possible phosphoryl transfer activity, we performed a conversion in 50 mM MES pH 7.0 supplemented with 50 mM azide and 200 mM glucose as phosphate acceptors. Interestingly, we observed phosphoryl transfer activity, resulting in the production of Glc 6-*P*, independent on the presence of rescue agent. The apparent reaction rate for the production of Glc 6-*P* and the consumption of αGlc 1-*P* were identical for both performed reactions (Figure 3B). In both conversions the phosphate balance was closed, thus indicating that no phosphate was transferred to azide.

The results can be interpreted in two ways: On the one hand the residual phosphoryl transfer to glucose as acceptor can be catalyzed by impurities of αG1Pase, originating from the expression host. If we assume that the whole residual activity is caused by this contamination, then the transfer activity of H18A should be roughly 1% of the αG1Pase transfer activity (Table 1). The k_{cat_app} for Glc 6-*P* production using αG1Pase is 30 s^{-1} [2], whereas for H18A it is 0.15 s^{-1}. Although 1% contamination with αG1Pase seems high, optimization of the purification protocol or production of an *E. coli* OrigamiB Δ*agp* expression strain is recommended to eliminate the possibility of contamination after all.

On the other hand, the observed transfer activity could also be explained by His^{289}, located in the active core, taking over the role as an alternative nucleophile. This would allow phosphoryl transfer without the presence of the original nucleophile His^{18} and independent of rescue agents. In previous studies with the closely related *E. coli* phytase *appA* (30% sequence

identity to αG1Pase) Ostanin et al could show that a phytase-H303A mutant exhibits a tremendous decrease of hydrolytic activity (≤ 0.4% residual activity). Unfortunately, they could not pin down the exact catalytic role of His^{303}, the homologues residue to His^{289} in αG1Pase [11].

The production of a H18A-H289A double mutant could elucidate the exact role of both residues during catalysis. Although very interesting, further investigations of these questions is beyond the scope of this master thesis.

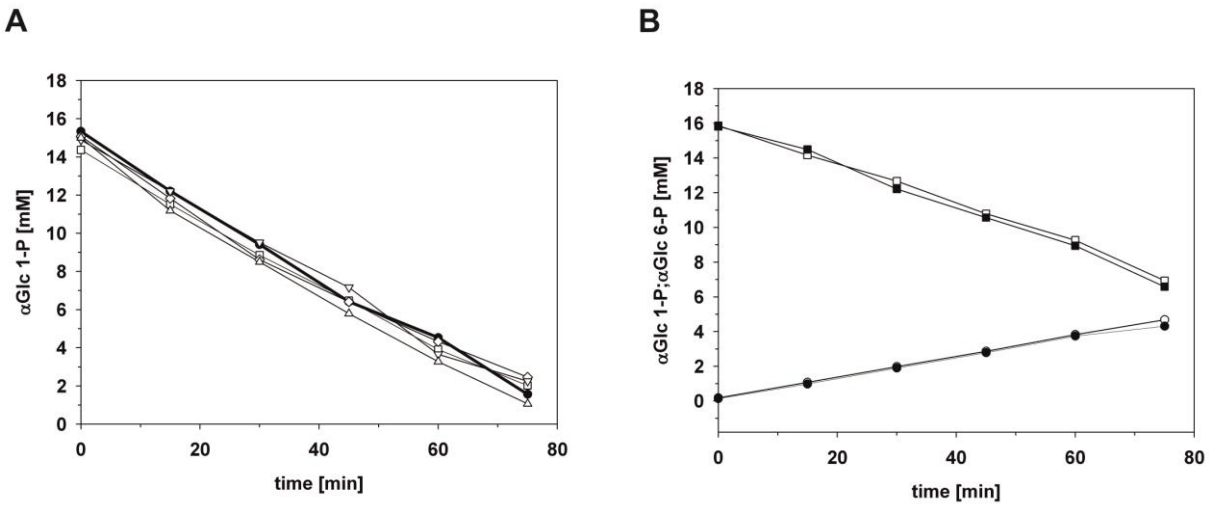


Figure 3. Chemical rescue of H18A. (A) Chemical rescue using different nucleophilic reagents. Symbols indicate: (●) without rescue agent, (△) 50 mM azide, (◇) 20 mM acetate, (□) 20 mM azide and (▽) 20 mM imidazole. (B) Non-hydrolytic phosphoryl transfer in H18A is independent on the presence of rescue agent (50 mM azide). Symbols indicate: (●) Glc 6-*P* concentration in the absence of azide, (○) Glc 6-*P* concentration in the presence of 50 mM

azide, (■) αGlc 1-*P* concentration in the absence of azide and (□) αGlc 1-*P* concentration in the presence of 50 mM azide.

3.2.5 Characterization of D290A.

Hydrolytic activity of D290A. D290A hydrolyzed αGlc 1-*P* with a k_{cat_app} of 0.1 $s^{-1}$and *p*NPP with a k_{cat_app} of 10 s^{-1} (Table 1). As expected, D290A showed only very low residual activity (0.3% compared to αG1Pase) for the hydrolysis of αGlc 1-*P*, however, activity towards *p*NPP is increased by a factor of 1.5 compared to αG1Pase. In contrast to the other produced mutants, the replacement of Asp^{290} by Ala leads to an increase of activity towards *p*NPP. As the catalytic efficiency of D290A is highly dependent on the substrate used, we suggest that the reason for this interesting catalytic behavior is mediated by the substrate.

Compared to αGlc 1-*P*, *p*NPP has a much better leaving group, thus protonation of the leaving group is less important for efficient catalysis. In consistency, the replacement of the acid-base catalyst, Asp^{290} by Ala, leads to a strong decrease in hydrolytic activity towards αGlc 1-*P*, but not towards *p*NPP. This finding leads to the conclusion that Asp^{290} might be involved in leaving group protonation.

Activity of αG1Pase and D290A is dependent on the substrate leaving group pK_a. A number of phosphate esters were used as substrates for αG1Pase and D290A, in order to study dependency of the apparent hydrolysis rate onto changes in the leaving group pK_a (Table 2). Both enzymes showed dependency of apparent phosphate release rates on changing leaving group pK_a. The correlation was clearly negative for D290A, whereas increase of the leaving

group's pK_a resulted in a decrease of hydrolytic activity (i.e. phenylphosphate ≤ 1 % activity relative to *p*NPP). αG1Pase shows moderate sensitivity to increased leaving group pK_a, as shown with phenylphosphate (37% activity relative to *p*NPP). Contrary to the trend of D290A, αG1Pase showed the highest activity with αGlc 1-*P*, which possesses a high leaving group pK_a of 12.46 (Figure 4). An exception of this trend represents the hydrolysis of PPi (Table 2). Both enzymes show low apparent turnover rates for this substrate (0.21 s^{-1} D290A, 0.11 s^{-1} αG1Pase) although the pK_a of the leaving group is the highest of the tested substrates. This indicates that PPi is a poor substrate. Interestingly, D290A shows a higher turnover compared to αG1Pase, as observed with *p*NPP. PPi and *p*NPP have a good leaving group and thus leaving group protonation is less important for efficient hydrolysis. However, this does not fully explain why substrates with a good leaving group are more efficiently hydrolyzed by D290A. A tentative explanation of enhanced turnover by D290A would be, that elimination of a negatively charged amino acid in the active center would facilitate entry of a negatively charged substrate and thus enhance the apparent turnover compared to αG1Pase.

Altogether the results support the hypothesis that Asp^{290}, the homologue residue to phytase Asp^{304}, is the acid-base catalyst in αG1Pase (Figure 1A). Furthermore, our results suggest that the formation of the phospho-enzyme intermediate and release of the leaving group is rate limiting in D290A. However, in αG1Pase, breakdown of the phospho-histidine intermediate is facilitated by protonation of the leaving group, which is mediated by the general acid-base catalyst. Thus, turnover in αG1Pase is less dependent on the pK_a of the leaving group, as compared to D290A.

Table 2. Kinetic parameters of αG1Pase and D290A towards various phosphate esters.

Substrate	Leaving group	k_{cat_app} [s^{-1}]	
	pK_a	D290A	αG1Pase
Pyrophosphate	6.82 [12]	0.21	0.11
*p*NPP	7.14 [13]	10	6.8
2-Naphtylphosphate	9.63 [13]	0.33	n.d.
*p*Phenylphosphate	9.99 [13,14]	0.08	1.7
αGlc 1-*P*	12.46 [13]	0.09	44
Phytate		0.94	2.07

[a] n.d. not determined

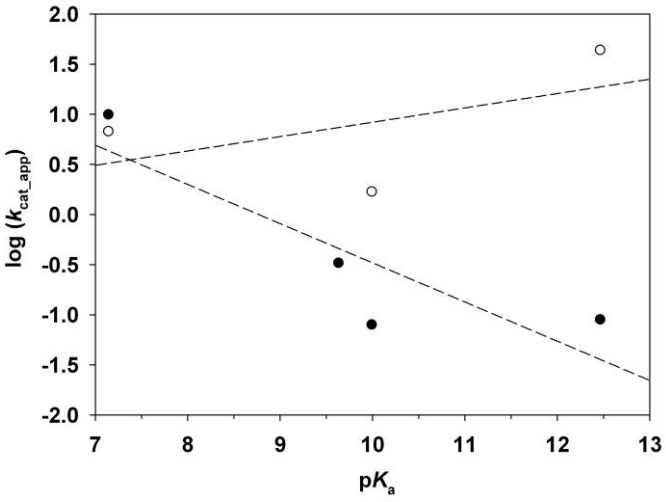


Figure 4. Brønsted plots for αG1Pase and D290A. Symbols indicate: (○) αG1Pase and (●) D290A.

***Phosphoryl transfer activity of D290A*.** To study possible non-hydrolytic phosphoryl transfer, catalyzed by D290A, a standard conversion was performed adding 15 mM αGlc 1-*P* as phosphoryl donor and 200 mM glucose as phosphoryl acceptor.

As shown in Figure 5, D290A was able to perform non-hydrolytic phosphoryl transfer, resulting in the production of 2.1 mM of Glc 6-*P* after 75 min of incubation. The αGlc 1-*P* consumption rate was reduced by a factor of 2 compared to the same reaction performed without addition of glucose. This apparent rate retardation might be caused by an enhanced back reaction, producing αGlc 1-*P*, instead of Glc 6-*P* [2]. The phosphate release rate is decreased by a factor of 10, which is also observed for αG1Pase and represents the suppressed substrate hydrolysis by the presence of an alternative phosphate acceptor substrate [2]. The overall reduced reaction rates, compared to αG1Pase, can be explained mechanistically due to the fact, that in D290A, catalysis rates are limited by the formation of the phospho-enzyme intermediate. The second half reaction, however, is not limiting and therefore proceeds with the same rate as in αG1Pase, thus, the transfer over hydrolysis ratio is identical to that observed in αG1Pase. Unfortunately, the observed transfer activity could be an artifact produced by minor contaminations of αG1Pase.

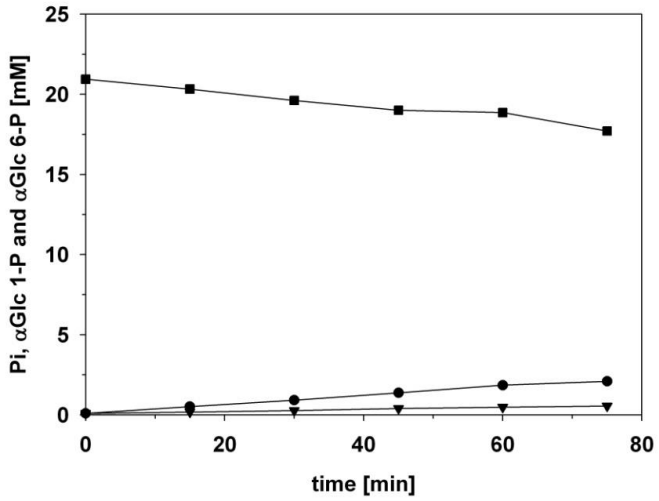


Figure 5. Time-course of phosphoryl transfer reactions catalyzed by D290A in the presence of 200 mM glucose as acceptor substrate at pH 7.0. Symbols indicate: **(■)** αGlc 1-*P*, (●) Glc 6-*P* and (▼) free phosphate.

In summary, we were able to show that Asp290 is involved in leaving group protonation and thus, can be considered as the general acid-base catalyst in αG1Pase. Observed phosphoryl transfer activity, however, needs further investigation. It is not clear if it results from αG1Pase contamination, or if it is an intrinsic property of D290A. To clarify this open question re-expression of the D290A variant in an *E. coli* OrigamiB Δ*agp* expression strain and subsequent characterization would be necessary. An alternative would be optimization of the purification process, where elimination of periplasmic αG1Pase would be the main goal. This can be archived by introducing a pre-purification step which specifically releases periplasmatic protein by cold osmotic shock [15], prior to cell disruption.

3.2.6 Characterization of E196A.

Hydrolytic activity of E196A. The hydrolysis rate of the gating residue, E196A, was determined with 31 s^{-1} for αGlc 1-*P* and 2.5 s^{-1} for *p*NPP (Table 1). Both activities are in the same order of magnitude compared to αG1Pase activities. Lee et al proposed, based on docking studies, that Glu^{196} might hinder the entrance of *p*NPP into the active site at elevated pH. They postulate that entry into the active site is denied by the negatively charged Glu^{196} carboxyl side chain, which repels *p*NPP. Exchange of Glu^{196} to Ala, however, did not promote hydrolysis of *p*NPP at pH 7.0 compared to αG1Pase. Contrary we found decreased hydrolytic activities towards both substrates (αGlc 1-*P* and *p*NPP), which might be explained due to slight alterations in the binding tunnel and thus less optimal substrate binding. An alternative hypothesis would be increased phosphate inhibition resulting from the lack of negative charge repulsion in the binding pocket, which is usually mediated by Glu^{196}. Further investigation of both hypotheses was beyond the scope of this master thesis.

Analysis of the substrate spectrum of E196A. As a result of Glu^{196} exchange to Ala we expected an enlargement of the binding pocket and thus an enhanced acceptance for bulkier phosphoryl acceptor substrates. To test this hypothesis, we performed phosphoryl transfer reactions at pH 7.0 using either 200 mM guanosine, or 200 mM L-ascorbic acid as acceptor substrate and 20 mM αGlc 1-*P* as donor substrate. Incorporation of phosphate into the acceptor substrate was accessed towards the phosphate balance, whereas a not closed balance was interpreted as phosphoryl transfer activity towards the offered acceptor substrate.

Both performed reactions showed a closed phosphate balance and thus we can conclude that neither guanosine nor L-ascorbic acid are suitable acceptor substrates.

Structural analysis of E196A. Lee et al proposed that Glu^{196} interacts with the C4-OH of αGlc 1-*P*. The crystal structure analyzed by Lee et al contained a βGlc 1-*P* instead of αGlc 1-*P*. The distance between Glu^{196} and the C4-OH of the βGlc 1-*P* is indeed very short (2.3 Å), representing a strong hydrogen bond. Analysis of the αG1Pase crystal structure docked with αGlc 1-*P* shows, that the distance between Glu^{196} and the C4-OH is roughly doubled (4.5 Å) and thus hydrogen bonding is unlikely to occur. Instead, analysis of the polar interactions of C4-OH and the C5-OH of the glucose ring revealed Ser^{291} as a possible interaction partner. The distance between the Ser^{291} hydroxyl group and the C4-OH and the C5-OH is ≤ 3 Å, representing possible hydrogen bonds that orientate the sugar ring in the active site.

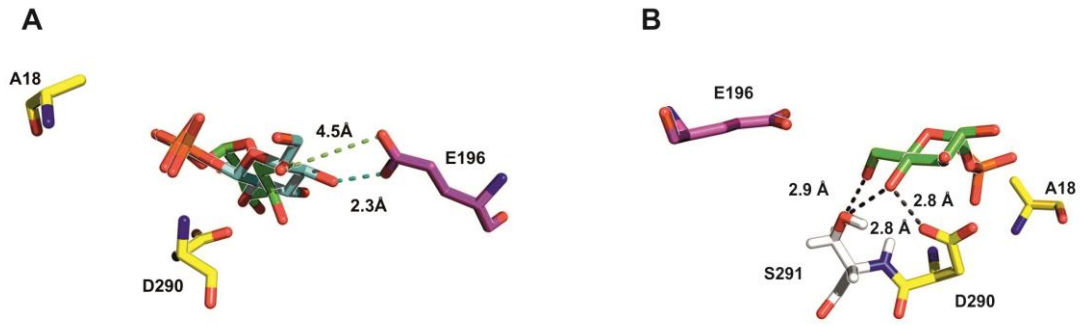


Figure 5. Interactions of gating residues with the glucose ring. (A) Interactions of the gating residue Glu^{196} with the C4-OH of αGlc 1-*P* (green) or the C4-OH of βGlc 1-*P* (cyan). (B) Interactions of Ser^{291} with the C4-OH and the C6-OH of αGlc 1-*P*.

In summary, we were able to show that the exchange of Glu^{196} to Ala did not result in enhanced activity towards *p*NPP at pH 7.0 and that it did not enhance the promiscuity towards alternative phosphoryl acceptor substrates such as L-ascorbic acid and guanosine. Furthermore, structural analysis revealed that Glu^{196} is unlikely to interact with the glucose ring of αGlc 1-*P*, however Ser^{291} seems to be involved in substrate binding. A more detailed characterization of E196A was beyond the scope of this master thesis.

3.3 Reference

1. Wang W, Malcolm BA (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. Biotechniques 26: 680-682.

2. Pfeiffer M, Nidetzky B, Wildberger P, Brecker L (2014) Chapter 1 of this thesis. In preparation for Appl. Env. Microbiol.

3. Wildberger P, Todea A, Nidetzky B (2012) Probing enzyme-substrate interactions at the catalytic subsite of Leuconostoc mesenteroides sucrose phosphorylase with site-directed mutagenesis: the roles of Asp(49) and Arg(395). Biocatal. Biotransform. 30: 326-337.

4. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248-254.

5. Saheki S, Takeda A, Shimazu T (1985) Assay of inorganic phosphate in the mild pH range, suitable for measurement of glycogen phosphorylase activity. Anal. Biochem. 148: 277-281.

6. Chapter 2 of this thesis, in preperation for the J. Mol. Catal. B: Enzym.

7. Berka K, Hanak O, Sehnal D, Banas P, Navratilova V, et al. (2012) MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. Nucleic Acids Res. 40: 222-227.

8. Lim D, Golovan S, Forsberg CW, Jia Z (2000) Crystal structures of Escherichia coli phytase and its complex with phytate. Nat. Struct. Biol. 7: 108-113.

9. Cottrill MA, Golovan SP, Phillips JP, Forsberg CW (2002) Inositol phosphatase activity of the Escherichia coli agp-encoded acid glucose-1-phosphatase. Can. J. Microbiol. 48: 801-809.

10. Lee DC, Cottrill MA, Forsberg CW, Jia Z (2003) Functional insights revealed by the crystal structures of Escherichia coli glucose-1-phosphatase. J. Biol. Chem. 278: 31412-31418.

11. Ostanin K, Harms EH, Stevis PE, Kuciel R, Zhou MM, et al. (1992) Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase. J. Biol. Chem. 267: 22830-22836.

12. Kumler WD, Eiler JJ (1943) The acid strength of mono and diesters of phosphoric acid. The n-alkyl esters from methyl to butyl, the esters of biological importance, and the natural guanidine phosphoric acids. Journal of the American Chemical Society 65: 2355-2361.

13. Lide DR (2006) CRC Handbook of Chemistry and Physics: A Ready-reference Book of Chemical and Physical Data: 2006-2007: CRC Press Incorporated.

14. Gawron O, Duggan M, Grelecki CJ (1952) Manometric Determination of Dissociation Constants of Phenols. Anal. Chem. 24: 969-970.

15. Quan S, Hiniker A, Collet J-F, Bardwell JA (2013) Isolation of Bacteria Envelope Proteins. In: Delcour AH, editor. Bacterial Cell Surfaces: Humana Press. pp. 359-366.