Characterization of Enzymes Involved in the Synthesis and Processing of Ether Lipids in Archaea and Bacteria

DISSERTATION ZUR ERLANGUNG

DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)

DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER

UNIVERSITÄT REGENSBURG

vorgelegt von

Mona Linde

aus

Fürth

im Jahr 2017

Das Promotionsgesuch wurde eingereicht am:

05.04.2017

Die Arbeit wurde angeleitet von: **Dr. Patrick Babinger**

Prof. Dr. Reinhard Sterner

Unterschrift:

This work was done in the period from July 2013 to April 2017 in the group of Dr. Patrick Babinger - Prof. Dr. Reinhard Sterner (Biochemistry II, Institute of Biophysics and Physical Biochemistry, University of Regensburg).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

Amino acids are abbreviated by the 1- or 3-letter code.

$^{0.1\,\%}A_x$	specific extinction coefficient at x nm
A	absorbance, adenosine
Å	Ångström (10^{-10} m)
Ac	acetate
afGGGPS	GGGPS from *Archaeoglobus fulgidus*
Amp	ampicillin
APS	ammonium persulfate
AraM	glycerol 1-phosphate dehydrogenase
ASP	ammonium sulfate precipitation
baYvoF	YvoF from *Bacillus anthracis*
bp	base pair
bsPcrB	PcrB from *Bacillus subtilis*
bsYvoF	YvoF from *Bacillus subtilis*
B. subtilis	*Bacillus subtilis*
c	concentration
C	cytosine
°C	degree Celsius
cAMP	cyclic adenosine monophosphate
Cam	chloramphenicol
CD	circular dichroism
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-Propanesulfonate, a zwitterionic detergent
CIP	calf intestinal phosphatase
CK	creatine kinase
CoA	coenzyme A
cpGGGPS	GGGPS from *Chitinophaga pinensis*
cpm	counts per minute
C-terminal	carboxy-terminal end of a polypeptide chain
CV	column volume
d	pathlength [cm]
D	dimer
Da	dalton [g/mol]
DGGGPS	digeranylgeranylglyceryl phosphate synthase
DGGGPR	digeranylgeranylglyceryl phosphate reductase
DHAP	dihydroxyacetone phosphate
DMSO	dimethyl sulfoxide
dNTP	deoxyribonucleotide triphosphate (N = A, C, G or T)
dsDNA	double-stranded DNA
DSC	differential scanning calorimetry
DTNB	5,5´-dithiobis-(2-nitrobenzoic acid); Ellman´s reagent
DTT	1,4-dithiothreitol

E	enzyme
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
Ery	erythromycine
et al.	and other authors (*et alii*)
EtBr	ethidium bromide
EtOH	ethanol
f	dilution factor
FBA	fructose 1,6-bisphosphate aldolase
fjGGGPS	GGGPS from *Flavobacterium johnsoniae*
FPP	farnesyl pyrophosphate
G	guanosine, Gibbs free energy
GDH	glutamate dehydrogenase
GdmCl	guanidinium chloride
GGG	geranylgeranylglycerol
GGGP	geranylgeranylglyceryl phosphate
GGGPS	geranylgeranylglyceryl phosphate synthase
G1P	glycerol 1-phosphate
G3P	glycerol 3-phosphate
GGPP	geranylgeranyl pyrophosphate
GPP	geranyl pyrophosphate
h	hour
H	enthalpy, hexamer
HepG	heptaprenylglycerol
HepGP	heptaprenylglyceryl phosphate
HepGPS	heptaprenylglyceryl phosphate synthase, PcrB
HepPP	heptaprenyl pyrophosphate
$(His)_6$ tag	hexahistidine tag
HK	hexokinase
HPLC	high pressure liquid chromatography
IPP	isopentenyl pyrophosphate
IPTG	isopropyl ß-D-thiogalactopyranoside
kan	kanamycine
kb	kilobase pair
k_{cat}	turnover number
k_{cat}/K_M	catalytic efficiency parameter
K_d	dissociation constant
k_M	Michaelis-Menten constant
KP	potassium phosphate (buffer)
lacZ	gene coding for the enzyme β-galactosidase
LB	Luria-Bertani (medium)
LIC	ligation independent cloning
LUCA	Last universal common ancestor
M	molar [mol/l], monomer
MAT	maltose O-acetyltransferase
MCS	multiple cloning site
MD	molecular dynamics

MSA	multiple sequence alignment
mtGGGPS	GGGPS from *Methanothermobacter thermautotrophicus*
MW	molecular weight
MWCO	molecular weight cut off
n	nano (10^{-9}), number of nucleotides
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
nanoDSF	nano differential scanning fluorimetry
NBRP	National BioResource Project
N_A	number of amino acids
N-terminal	amino-terminal end of a polypeptide chain
OD_x	optical density at x nm
ON	over night
p	pico (10^{-12})
P	pellet (insoluble cell fraction)
PAGE	polyacrylamide gel electrophoresis
PC	paper chromatography
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PFK	phosphofructokinase
PGI	phosphogluco isomerase
pH	negative decadic logarithm of the proton concentration
PhoB	alkaline phosphatase
PNPase	Purine nucleoside phosphorylase from microorganism
PP_i	pyrophosphate
P_i	orthophosphate
PPase	Pyrophosphatase from *Escherichia coli*
QCM	QuikChange mutagenesis
r.m.s.d.	root mean square deviation
rpm	revolutions per minute
R	resistance
RT	room temperature
S	supernatant (soluble cell fraction), substrate concentration, entropy
SDS	sodium dodecyl sulfate
slGGGPS	GGGPS from *Sphingosoma linguale*
ssDNA	single-stranded DNA
$t^{app1}/_2$	apparent half-life
T_x	transition temperature with x being 1, 1* or 2
T	temperature, thymidine
T_A	annealing temperature
taGGGPS	GGGPS from *Thermoplasma acidophilum*
TBE	Tris-Borat-EDTA buffer

TEMED	N,N,N',N'-tetramethylethylenediamine
TIM	triosephosphate isomerase
tkGGGPS	GGGPS from *Thermokokkus kodakarensis*
TLC	thin layer chromatography
T_M	melting temperature of primers; temperature at which 50% of the protein is in a non-native state
Tris	Tris(hydroxymethyl)aminomethane
TrxA	N-terminal thioredoxin tag
U	Unit, 1U is equivalent to the amount of enzyme that converts 1 µmol substrate per minute at standard conditions
UV	ultraviolet
V	volt
v_i	initial velocity
v_{max}	maximum velocity
(v/v)	volume per volume
wt	wild type
(w/v)	weight per volume
XOD	microbial xanthine oxidase
YvoF	polyprenylglycerol *O*-acetyltransferase
zpGGGPS	GGGPS from *Zunonwangia profunda*
Δ	difference
Δxxx	*B. subtilis* strain with deleted xxx gene
ε_x	molar extinction coefficient at x nm
λ	wavelength
µ	micro (10^{-6})
3D	three dimensional
Θ	ellipticity
Θ_{MRW}	average ellipticity per amino acid
Θ_{obs}	meausered ellipticity
Ω	ohm

1 ABSTRACT

In Archaea, ether lipids play an essential role as the main building blocks of the cellular membrane. Recently, ether lipids have also been discovered in the domain of Bacteria, and the key enzymes that catalyze their synthesis, glycerol 1-phosphate dehydrogenase (AraM) and heptaprenylglyceryl phosphate synthase (HepGPS), have been described. In Bacillales, heptaprenylglyceryl phosphate (HepGP) does not become linked to a second polyprenyl moiety like for ether lipids in Archaea, but is dephosphorylated and acetylated. The enzymes that catalyze these reactions have been identified and characterized whithin this work. The phosphatase PhoB acting on HepGP was enriched from a *Bacillus subtilis* (*B. subtilis*) cell extract and identified by mass spectrometry. The *B. subtilis* gene *phoB* was amplified and heterologously expressed. The dephosphorylation activity of PhoB could be verified *in vitro*. Nevertheless, the results indicate that any other phosphatase might catalyze dephosphorylation of HepGP as well, because those enzymes only exhibit a low subtrate specificity. By screening a *B. subtilis* knockout library for deficiency in acetylation, the *yvoF* gene product was identified to be the acetyltransferase. The *yvoF* gene was heterologously expressed and YvoF was characterized biochemically. Its acetyl-CoA dependent activity was verified *in vitro* and catalytic parameters were obtained by a 5,5´-dithiobis-(2-nitrobenzoic acid)-coupled activity assay and a discontinuous radiometric assay. A membrane assoziation of YvoF was inferred from ultracentrifugation experiments. The *in vitro* acetyltransferase activity was also verified for another YvoF representative, namely from *Bacillus anthracis*. Because YvoF is a close relative to maltose *O*-acetyltransferase (MAT), the activity of YvoF was compared to MAT. YvoF and MAT partially overlap in substrate and product range *in vitro*, but MAT is not able to complement the *yvoF* knockout *in vivo*. It remains to be clarified how this *in vivo* specialization is achieved, and one answer might be that the cellular localization of the two enzymes is different, as the YvoF substrate HepG is associated to the cellular membrane. The biological function of acetylated ether lipids in *B. subtilis* is still unknown. However, the identification of two enzymes involved in their biosynthesis pathway brings us closer to answer this intriguing question.

The HepGPS belongs to the family of geranylgeranylglyceryl phosphate synthases (GGGPS) and previous studies showed that the members of this family can be dimers or hexamers. Hexameric GGGPS complexes are built from three dimeric modules in the configuration

of the native dimers. Besides the dimeric interface (symmetric interface 1), two additional interfaces exist in the hexamer, the symmetric interface 2 and the asymmetric interface 3a/3b. To elucidate the impact of hexamerization on stability and activity, the GGGPS from *Methanothermobacter thermautotrophicus* (mtGGGPS) was investigated by mutational studies. The symmetric interface 2 and the asymmetric interface 3a/3b were shown to contribute equally to the stabilization of the hexameric structure, because both led to dimerization if disturbed. A mutation that disrupts the dimeric interface led to a further collapse of the dimers into monomers. The thermal stability of the resulting monomeric, dimeric and hexameric mtGGGPS mutants was determined via various biophysical techniques (circular dichroism-, differential scanning calorimetry-, nano differential scanning fluorimetry-measurements and irreversible heat inactivation). In almost all cases, a higher oligomerization state ensured an elevated thermostability and contributed to maintaining activity at higher temperatures. Based on these results, a putative unfolding pathway for mtGGGPS was proposed. First, a transition of the native structure to a stable partially folded intermediate occurs, which is inactive but maintains almost 80-90 % of the overall secondary structure. Depending on the oligomerization state, this 1st transition happens at a temperature between 50 °C (monomeric mutants), 60 °C (dimeric mutants) and 100 °C (hexameric wild type and mutants). Only at elevated temperatures of around 120 °C a second transition presumably leads to a complete denaturation of the protein, irrespective of the oligomerization state. Data obtained with two other couples of hexameric GGGPS wild type and dimeric mutant as well as the analysis of native hexameric and dimeric GGGPS enzymes from different species confirmed the results obtained with mtGGGPS. A steady-state kinetic experiment revealed that the dimeric mutants showed decreased catalytic efficiencies especially due to an increase of the K_M for G1P. In summary, it could be shown for almost all hexameric enzymes that when compared to dimeric or monomeric enzymes, hexamerization ensured structural integrity as a prerequisite for thermal stability and activity.

2 ZUSAMMENFASSUNG

Etherlipide spielen in Archaeen eine essentielle Rolle als Hauptbausteine der Zellmembran. Kürzlich wurden auch in Bakterien Etherlipide entdeckt und die Schlüsselenzyme, die deren Synthese katalysieren identifiziert: die Glycerin-1-phosphatdehydrogenase (AraM) und die Heptaprenylglycerinphosphatsynthase (HepGPS). In Bacillales wird das entstandene Heptaprenylglycerinphosphat (HepGP) jedoch nicht mit einer weiteren Polprenylkette verknüpft, wie es in Archaeen geschieht, sondern nachfolgend dephosphoryliert und acetyliert. Die Enzyme, die für diese Reaktionen verantwortlich sind, wurden im Rahmen dieser Arbeit identifiziert und charakterisiert. Die Phosphatase PhoB, die HepGP dephosphoryliert wurde aus einem *B. subtilis* Rohextrakt angereichert und mittels Massenspektrometrie identifiziert. Das Gen *phoB* von *B. subtilis* wurde amplifiziert und heterolog exprimiert. Die Dephosphorylierungsaktivität von PhoB konnte *in vitro* nachgewiesen werden. Nichtsdestotrotz kann angenommen werden, dass jede andere Phosphatase auf Grund geringer Substratspezifität diese Dephosphorylierung katalysieren kann. Bei dem *Screening* einer *B. subtilis knockout*-Bibliothek auf den Verlust der Acetylierung von HepG wurde das Gen *yvoF* identifiziert, welches für die Acetyltransferase YvoF kodiert. Das Gen wurde heterolog exprimierte und YvoF biophysikalisch charakterisiert. Die acetyl-CoA-abhängige Aktivität wurde *in vitro* bestätigt und katalytische Parameter wurden mittels eines 5,5´-Dithiobis-2-nitrobenzoesäure-gekoppelten *Assays* und eines diskontinuierlichen radioaktiven *Assays* ermittelt. Auf eine Membranassoziation von YvoF wurde auf Grund von Ultrazentrifugationsergebnissen geschlossen. Die *in vitro*-Acetyltransferaseaktivität konnte zusätzlich für einen weiteren YvoF-Vertreter aus *Bacillus anthracis* bestätigt werden. Da YvoF nahe mit der Maltose-*O*-Acetyltransferase (MAT) verwandt ist, wurden die Aktivitäten von YvoF und MAT untereinander verglichen. YvoF und MAT überlappen teilweise in ihrem Substrat- und Produktspektrum *in vitro*, aber MAT kann den *knockout* von YvoF *in vivo* nicht komplementieren. In der Zukunft muss diese *in vivo* Spezialisierung noch aufgeklärt werden, eine mögliche Antwort könnte jedoch sein, dass die zelluläre Lokalisierung beider Enzyme unterschiedlich ist, da das Substrat HepG von YvoF membrangebunden vorliegt. Die biologische Funktion von acetylierten Etherlipiden in *B subtilis* ist bislang unbekannt. Die Identifizierung zweier

Enzymen, die an deren Biosyntheseweg beteiligt sind, bringt uns jedoch näher daran diese Frage zu beantworten.

Die HepGPS gehört zu der Familie der Geranylgeranylglycerinphophatsynthasen (GGGPS) und vorangegangene Studien zeigten, dass Mitglieder dieser Familie als Hexamere oder Dimere vorkommen. Hexamere GGGPS sind aus drei dimeren Modulen aufgebaut, welche die Konfiguration der nativen Dimere aufweisen. Neben der Dimerkontaktfläche (symmetrische Kontaktfläche 1) existieren zwei weitere Kontaktflächen im Hexamer, die symmetrische Kontaktfläche 2 und die asymmetrische Kontaktfläche 3a/3b. Um die Auswirkung der Hexamerisierung auf Stabilität und Aktivität zu untersuchen, wurde die GGGPS aus *Methanothermobacter thermautotrophicus* in Mutagenesestudien untersucht. Dabei stellte sich heraus, dass die asymmetrische Kontaktfläche 3a/3b und die symmetrische Kontaktfläche 2 gleichermaßen wichtig sind für die Stabilisierung des Hexamers, da beide nach Deletion zu einer Dimerisierung führten. Eine Mutation, die die Dimerkontaktfläche zerstört, führte zu einer Monomerisierung der Dimere. Die thermische Stabilität der hieraus resultierenden Monomere, Dimere und Hexamere wurde mit verschiedenen biophysikalischen Methoden untersucht (Circulardichroismus, Differentielle *Scanning*-Kalorimetrie, Nano-differentielle *Scanning*-Fluorimetrie und irreversibler Hitzeinaktivierung). Eine höhere Oligomerisierung sicherte in fast allen untersuchten Fällen eine höhere Thermostabilität und trug zum Erhalt der Aktivität bei hohen Temperaturen bei. Basierend auf diesen Ergebnissen wurde ein möglicher Auffaltungsweg für mtGGGPS postuliert. Zuerst geschieht ein Übergang der nativen Struktur zu einem stabilen teilweise gefalteten Intermediat, welches inaktiv ist, jedoch noch ca. 80-90 % der gesamten Sekundärstruktur des nativen Zustands aufweist. Je nach Oligomerisierungszustand geschieht dieser erste Übergang bei Temperaturen von 50 °C (monomere Mutanten), 60 °C (dimere Mutanten) und 100 °C (hexamerer Wildtyp und Mutanten). Erst bei höheren Temperaturen um 120 °C führt ein zweiter Übergang unabhängig vom Oligomerisierungszustand vermutlich zu einer kompletten Entfaltung des Proteins. Daten von zwei weiteren hexameren wildtypischen GGGPS und dimeren Mutanten sowie die Analyse von nativen hexameren und dimeren GGGPS verschiedener Spezies bestätigten die Ergebnisse, die mit mtGGGPS erhalten wurden. *Steady-state*-Kinetiken zeigten, dass die dimere Mutanten eine drastische Verschlechterung in der katalytischen Effizienz aufwiesen, hauptsächlich aufgrund einer verminderten Affinität für

G1P. Zusammenfassend konnte für fast alle hexameren Enzyme im Vergleich zu dimeren und monomeren Enzymen gezeigt werden, dass Hexamerisierung eine strukturelle Integrität als Voraussetzung für thermische Stabilität und Aktivität gewährleistet.

3 Introduction

3.1 Chemical composition of lipids in Bacteria and Archaea

The domain of Archaea forms a unique phylogenetic group if compared to the other two domains of life, Bacteria and Eukarya (Woese and Gupta, 1981). This domain has been of an early interest in science as Archaea exhibit the fascinating property to live in extreme environments, as for example high temperatures, high salt conditions, or acidic pH, which is fatal for many other organisms. Archaea are sufficiently different regarding biochemical features like highly diverged ribosomal RNAs, exceptional metabolic cofactors and most strikingly the chemical composition of the lipids forming their cellular membranes (Woese *et al.*, 1978). Whereas the membrane lipids of Bacteria and Eukarya are composed of *sn*-glycerol 3-phosphate (G3P) being esterified with two fatty acids, the membranes of Archaea exhibit an architecture in which isoprene units (most commonly two C_{20} geranylgeranyl groups) are connected to *sn*-glycerol 1-phosphate (G1P) as backbone via ether linkages (Kates, 1993; Koga *et al.*, 1993; Matsumi *et al.*, 2011; Wächtershäuser, 2003; Figure 1).

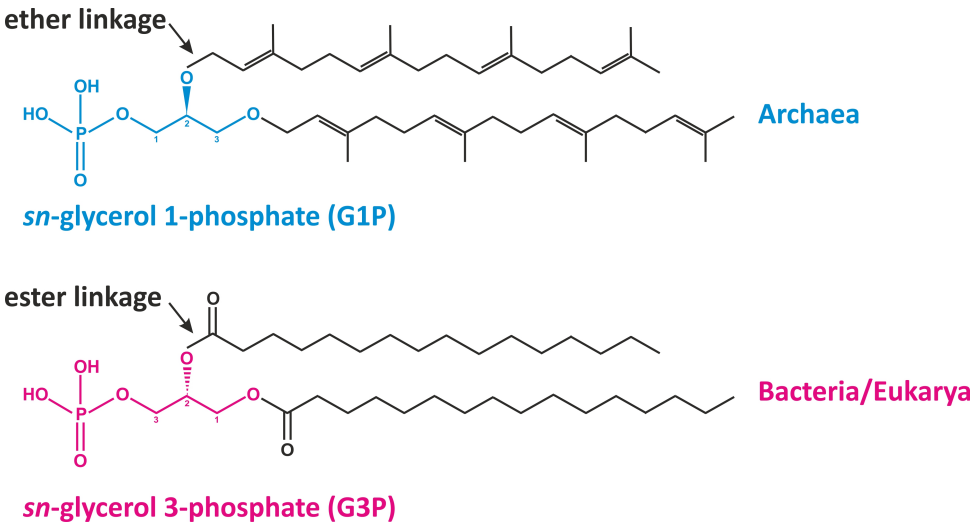


Figure 1 - Chemical composition of core phospholipids in Bacteria/Eukarya and Archaea
In Archaea G1P is bound to polyprenyl derivatives by ether linkages (blue), while in Bacteria and Eukarya G3P is bound to fatty acids by ester linkages (magenta).

This difference in the chemical composition of the membranes between Archaea and Bacteria/Eukarya has been considered to build a branch point in the early evolution of the three kingdoms (Boucher, 2007; Boucher *et al.*, 2004; Koga and Morii, 2007; Payandeh and Pai, 2007). Therefore huge efforts have been made identifying the enzymes involved

in the two different biosynthesis pathways for membrane lipids in Bacteria/Eukarya compared to Archaea (Figure 2).

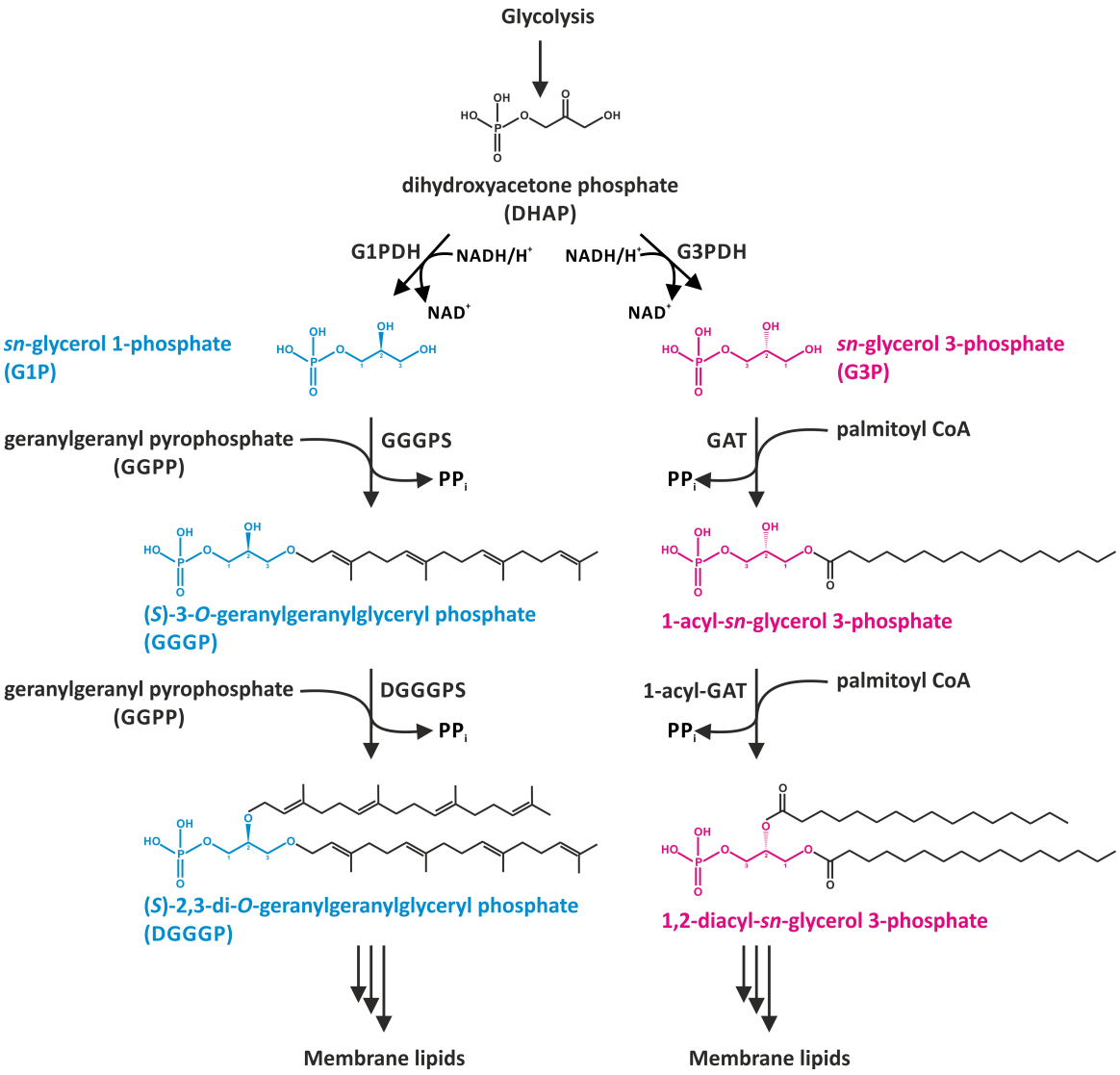


Figure 2 - Biosynthesis pathway of core phospholipids in Bacteria/Eukarya and Archaea
In Archaea (blue) G1P is provided by glycerol 1-phosphate dehydrogenase (G1PDH), while in Bacteria and Eukarya (magenta) G3P is provided by glycerol 3-phosphate dehydrogenase (G3PDH). Both enzymes exhibit no similarity among each other. In Archaea the GGPP is connected to G1P by the geranylgeranylglyceryl phosphate synthase (GGGPS). The second prenylation is catalyzed by the digeranylgeranylglyceryl phosphate synthase (DGGGPS). In Bacteria and Eukarya a glycerol 3-phosphate acyltransferase (GAT) and subsequently a 1-acyl-glycerol 3-phosphate acyltransferase (1-acyl-GAT) transfer the acyl moiety of acyl-CoA to G3P. PP_i: pyrophosphate, $NADH/H^+$: nicotinamide adenine dinucleotide (reduced form), NAD^+: nicotinamide adenine dinucleotide (oxidized). In this figure the acyl moiety is palmitoyl.

The committed step in the biosynthesis of ether lipids in Archaea, the synthesis of G1P through reduction of dihydroxyacetone phosphate (DHAP) and the transfer of a polyprenyl moiety to the C_3-hydroxyl group of G1P, is catalyzed by the glycerol 1-phosphate dehydrogenase (G1PDH) and the geranylgeranylglyceryl phosphate synthase (GGGPS). The second prenyltransfer to the C_2-hydroxyl group of G1P is catalyzed by the

digeranylgeranylglyceryl phosphate synthase (DGGGPS). Following reduction of the isoprenoid double bonds by the digeranylgeranylglyceryl phosphate reductase (DGGGPR) various head group modifications occur before the phosphoslipid gets incorporated into the cellular membrane (Koga and Morii, 2007). In Bacteria and Eukarya the addition of fatty acids is catalyzed by acyltransferase enzymes. Equally to Archaea, the core lipids get subsequently modified with polar head groups. The main polar head groups (L-serine, *myo*-inositol, and glycerol) are commonly found in both domains, Archaea and Bacteria, and the enzymes involved in the modification are homologous in all organisms (Koga, 2011). The GGGPS enzyme as well as the G1PDH, however, were considered to belong exlusively to the domain of Archaea and therefore highlight a key event in the evolutionary separation between Archaea and Bacteria (Glansdorff *et al.*, 2008; Koga, 2011; Lombard *et al.*, 2012b; Payandeh and Pai, 2007; Pereto *et al.*, 2004). However, prenylated compounds are not restricted to the domain of Archaea but are the most abundant and structurally most diverse natural products ocurring in all three domains of life (Vandermoten *et al.*, 2009; Winkelblech *et al.*, 2015). They are part of numerous biochemical pathways: next to being part of membranes (prenylated lipids in Archaea, sterols in Eukarya and Bacteria), isoprenylated compounds are also present as quinones in electron transport chains, as photosynthetic pigments (carotinoids), as hormons, etc. (Vandermoten *et al.*, 2009). Protein prenyltransferases for example play an important role in posttranslational modifications of proteins in eukaryotes inducing correct cellular localization or activity (Palsuledesai and Distefano, 2015). Additionally, prenylated tRNAs have been detected (Dumelin *et al.*, 2012; Xie *et al.*, 2007). The ComX pheromone from *B. subtilis* is an isoprenoide oligopeptide containing a regularly geranylated tryptophan residue, which stimulates natural genetic competence (Tsuji *et al.*, 2012). Other polyprenyl derivatives are used as lipid carrier in the synthesis of peptidoglycan (Kobayashi *et al.*, 2003; Leaver *et al.*, 2009; Thorne and Kodicek, 1966). Also bacteria possess ether lipids in form of alkyl ethers, which occur in small extents in anaerobes usually existing as mono-alkyl glycerol ethers with an ester-linked fatty acid (Goldfine and Langworthy, 1988). Next to prenylation, also G1P is known to be present in Bacteria as for example as compound of phosphoglycolipids and lipoteichoic acid polymers but gets produced by an alternate biosynthetic mechanism (Fischer and Arneth-Seifert, 1998; Percy and Grundling, 2014). The search for G1P-producing enzymes in bacteria revealed

that also Gram-positive Bacteria like *Bacillus subtilis* (*B. subtilis*) possess G1P-based ether lipids (Guldan *et al.*, 2011; Guldan *et al.*, 2008).

3.2 G1PDH and GGGPS as branch point in the evolutionary separation between the three domains of life

According to Woese and co-workers (1990; 1978) all living organisms stem from the early divergence of a hypothetical progenitor into three primary domains of life. As the origin of life on Earth remains an open question, elucidating detailed characteristics about the "last universal common ancestor" (LUCA) are still a big issue (Payandeh and Pai, 2007). One debate regarding early life concerns the nature of the membrane lipids of LUCA. Did LUCA exhibit mixed cell membranes or are the membranes strictly archaea- or bacteria-like in their composition or, most strikingly, did LUCA possess lipid membranes at all? If yes, did the different biosynthesis pathways evolve independently? Major functions of membrane lipids are to form a barrier between the cell and the environment, to generate an appropriate membrane fluidity at physiological temperatures, and to ensure proper functioning of the cells metabolism during other external influences (Edidin, 2003; Goldfine and Langworthy, 1988). Different controverse hypotheses and discussions have been put forward, which will be described shortly, based on a review of Lombard *et al.* (2012b).

Koga *et al.* (1998) proposed that LUCA was acellular, thereby indicating a late and independent emergence of lipid biosynthesis pathways. This, however, is difficult to integrate into the increasing evidence that LUCA was a complex organism containing several hundred genes. Martin and Russell (2003) envisaged that LUCA contained mineral membranes instead of lipid membranes, which would imply that the phospholipid biosynthesis pathways would have evolved independently during evolution of the bacterial and archaeal lineages in one hydrothermal chimney surprisingly by high speed (Lombard *et al.*, 2012b). Additionally to this somewhat particular assumption, the hypothesis fails to postulate a mechanism coupling the formation of these mineral compounds with replication (Lombard *et al.*, 2012b).

Wächtershäuser (2003) speculated in a very intriguing paper that the early cellularization occurred via membranes composed of simple lipids that were synthesized non-

enzymatically by inorganic transition metal catalysts or by primitive non-stereospecific enzymes. This hypothesis is mainly based on the pre-cell theory of Otto Kandler (1994a, b; Kandler and Konig, 1998). Kandler´s pre-cell theory imagines a trunk evolution of pre-cells, which have been emerged at some point in the first phase of evolution. The origin of these pre-cells is based on primordial chemo-autotrophic anabolism on surfaces of collodial or microcrystalline particles with transition metal/sulphur structure (Wächtershäuser, 1988, 1992). These pre-cells are metabolizing self-reproducing entities exhibiting most of the basic properties of a cell and display an overall population, which is distributed over a variety of habitats. Thus, pre-cells are multiphenotypical containing a stable racemate of chiral lipids (Kandler, 1994a, b; Kandler and Konig, 1998). In the second phase of evolution, the three domains of life diverged from this pre-cell population (Wächtershäuser, 2003). The heterochiral membranes would have spontaneously segregated into more stable homochiral membranes concomitant with the appearance of stereospecific enzymes, G1PDH and G3PDH (Figure 3).

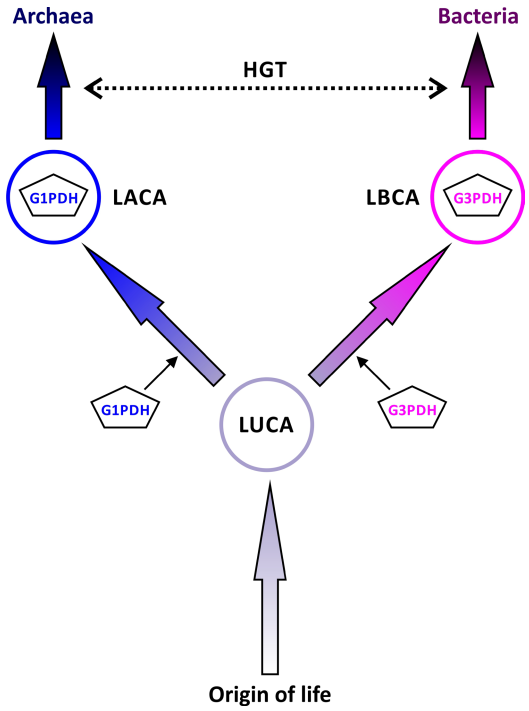


Figure 3 – Wächtershäusers hypothesis for the early evolution of cell membrane lipid composition
According to Wächtershäuser (2003), the last universal common ancestor (LUCA) had a heterochiral lipid membrane. Both G1P and G3P were used, but were synthesized non-enzymatically by inorganic transition metal catalysts or by primitive non-stereospecific enzymes. During evolution, the last bacterial common ancestor (LBCA) and the last archaeal common ancestor (LACA) diverged due to the emergence of specialized G3PDH and G1PDH, each containing homochiral lipid membranes. Certain bacterial lineages as Firmicutes/Bacillales/Bacteroidetes presumably received G1PDH from archaeal species through horizontal gene transfer (HGT). *Vice versa*, certain archaeal lineages may have acquired G3PDH from bacterial species via HGT. The figure was modified from Yokobori *et al.* (2016).

At each time point of divergence a founder population was supposed to branch off from the overall population. Kandler assumed the bacterial lineage to diverge first due to the relative simple information processing machinery at the so called pre-cell stage 1 (PC-1), thereby opening its own separate evolutionary path (Wächtershäuser, 2003). At a later more-evolved stage of the pre-cells (PC-2), the lineage of Archaea branched off and at last the eukaryal lineage evolved at PC-3.

Boucher and co-workers (2004) put forward yet another theory based on phylogenetic analysis of all enzymes involved in the biosynthesis of isoprenoid side-chains and of the G1P backbone. The results indicated that the isoprenoid biosynthesis of Archaea evolved through a combination of evolutionary processes including uptake of ancestral enzymes occurring prior to the diversification of the domain of Archaea, modification of enzyme specificity, orthologous and non-orthologous gene displacement and horizontal gene transfer (HGT). The biosynthesis of two precursors, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), is the foundation for the biosynthesis of archaeal membranes, because the polyprenyl chain, which gets attached to the G1P backbone by GGGPS enzymes, is assembled from these two isoprene units. The pathway synthesizing these precursors in most eukaryotes and some bacteria is the so called mevalonate pathway (MVA; Figure 4; Boucher and Doolittle, 2000; Boucher *et al.*, 2004; Lange *et al.*, 2000).

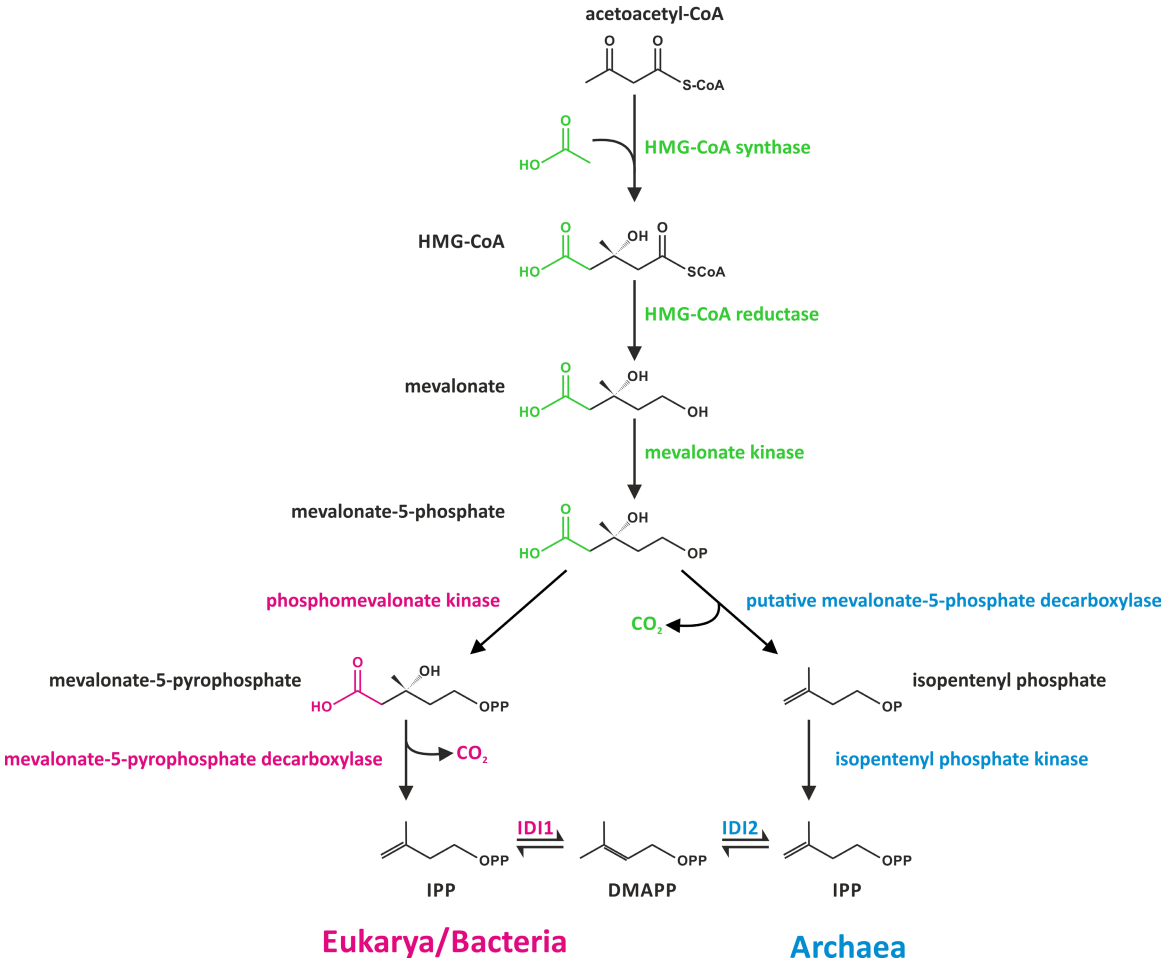


Figure 4 - Mevalonate pathway for synthesis of isoprenoid precursors
The structure of the acetoacetyl-CoA core is depicted in black, acetyl-CoA in green. Homologues enzymes are colored green. Starting from mevalonate-5-phosphate: the classical mevalonate pathway enzymes occurring mainly in Eukarya and Bacteria are colored magenta and the alternative mevalonate pathway enzymes occurring in most Archaea are colored blue. HMG-CoA: 3-hydroxy-3-methylglutaryl-CoA, IPP: isopentenyl pyrophosphate, DMAPP: dimethylallyl pyrophosphate, IDI1 and IDI2: isopentenyl pyrophosphate isomerase.

In all studied Archaea besides *Sulfolobus* the biosynthesis pathway of the precursors is somewhat different from the classical MVA pathway (Figure 4; Boucher *et al.*, 2004). While the first three enzymes are homologues to the bacterial and eukaryotic enzymes (Smit and Mushegian, 2000), the last two enzymes, the phosphomevalonate kinase and the mevalonate-5-pyrophosphate decarboxylase, are missing in most archaeal species (Boucher *et al.*, 2004). Lombard and co-workers (2012b) postulated an alternative archaeal MVA pathway starting from mevalonate-5-phosphate. In this alternative pathway instead of a second phosphorylation event followed by decarboxylation, a decarboxylation takes place first catalyzed by a putative decarboxylase, resulting in isopentenyl phosphate, which is subsequently phosphorylated to IPP by an isopentenyl

phosphate kinase. The enzyme catalyzing the conversion of IPP to DMAPP, the isopentenyl pyrophosphate isomerase (IDI1), is generally absent in Archaea with few exceptions, but the analogous non-homologous enzyme IDI2 is catalyzing the isomerization (Figure 4; Kaneda *et al.*, 2001). Following elongation of the isoprenoid chain by short-chain prenyltransferases, the next step in the biosynthesis of archaeal core phospholipids is the addition of geranylgeranyl isoprene chains to the G1P backbone. Most of the enzymes involved in lipid biosynthesis are assumed to have been present before the divergence of the domain of Archaea due to their relative widespread occurrence in all three domains of life (Boucher *et al.*, 2004; Lombard *et al.*, 2012b). Due to the fact that also Archaea exhibit variable concentrations of fatty acids or derivatives (Carballeira *et al.*, 1997; Nishihara *et al.*, 2000; Tornabene *et al.*, 1978), Lombard *et al*., (2012a) postulated an archaeal acyl carrier protein-independent fatty acid biosynthesis pathway. Archaeal fatty acids can be part of protein structures (Kolbe *et al.*, 2000) or have also been found in the membrane in very diverse euryarchaeotes (Gattinger *et al.*, 2002). Therefore it can be assumed that Bacteria and Archaea appear to have specialized their cell membranes by fine-tuning the relative importance of different components (Lombard *et al.*, 2012b). The impressive facts that isoprenoids are widely distributed and that the key steps in the MVA pathway as well as the fatty acid biosynthesis are conserved in all three domains of life lets assume that the LUCA possessed the complete toolkit for biosynthesis of isoprenoid and fatty acid-based phospholipids and therefore contained a heterochiral, complex modern-like phospholipid membrane (Lombard *et al.*, 2012b; Wächtershäuser, 2003). The driving force for the lipid divide during evolution could have been a certain instability of the mixed membranes, an adaption event of Archaea to high temperatures thereby preferring the ether-linked isoprenylated membranes and the evolution of the acyl carrier protein in Bacteria, which increased the efficiency of the fatty acid synthesis accompanied by a relegation of the isoprenoide biosynthesis (Lombard *et al.*, 2012b). Nevertheless, two enzymes seem to be a uniquely archaeal invention, which mirror the emergence of Archaea from LUCA and therefore are also called key enzymes in the evolutionary separation of the domain of Archaea: GGGPS and G1PDH (Boucher *et al.*, 2004; Koga *et al.*, 1998; Payandeh *et al.*, 2006; Payandeh and Pai, 2007; Pereto *et al.*, 2004). As also mesophilic archaea possess ether lipids, the GGGPS enzyme is proposed to have been present already in the last archaeal common ancestor (LACA).

3.3 Overall structural fold of GGGPS enzymes

GGGPS enzymes are of great interest as for one they lay at a branch point to biosynthesis of archaeal membrane lipids as already discussed above and on the other hand they are part of a large diverse family of prenyltransferases. The enzymatic characteristics of archaeal GGGPS have been investigated intensely (Chen *et al.*, 1993; Nemoto *et al.*, 2003; Payandeh and Pai, 2007; Soderberg *et al.*, 2001; Zhang and Poulter, 1993b) and a comprehensive analysis of the GGGPS enzyme family was performed by Peterhoff *et al.* (2014). Moreover, the crystal structures of several GGGPS enzymes have been solved (Payandeh *et al.*, 2006; Peterhoff *et al.*, 2014; Ren *et al.*, 2013; Ren *et al.*, 2012). GGGPS enzymes belong to a class of prenyltransferases that catalyze the transfer of isoprenoid groups onto non-isoprenoid acceptors. This class consists of a large amount of different prenyltransferases: *trans*- and *cis*-prenyltransferases, peptide-/protein prenyltransferases, tRNA prenyltransferases, and aromatic prenyltransferases; among the latter is the experimentally best investigated subgroup, the dimethylallyltryptophan synthase (DMATS; Winkelblech *et al.*, 2015). Although sharing a common reaction mechanism in terms of the structure barely any similarity can be found between GGGPS enzymes and other prenyltransferases. Only two types of prenyltransferases feature the well-known $(\beta\alpha)_8$-barrel fold (Oldfield and Lin, 2012; Payandeh *et al.*, 2006): GGGPS enzymes and MoeO5, which participates in the biosynthesis of the antibiotic moenomycin by catalyzing a *trans*-to-*cis* isomerization of farnesyl pyrophosphate (Doud *et al.*, 2011; Ren *et al.*, 2012). With about 10 % of all proteins with known structure exhibiting at least one $(\beta\alpha)_8$-barrel domain, this fold can be referred to as the most common enzyme fold (Sterner and Höcker, 2005). The $(\beta\alpha)_8$-barrel fold was first encountered in the triosephosphate isomerase (TIM) from chicken muscle (*Gallus gallus*; Banner *et al.*, 1975) and is therefore also called “TIM-barrel” fold (Figure 5).

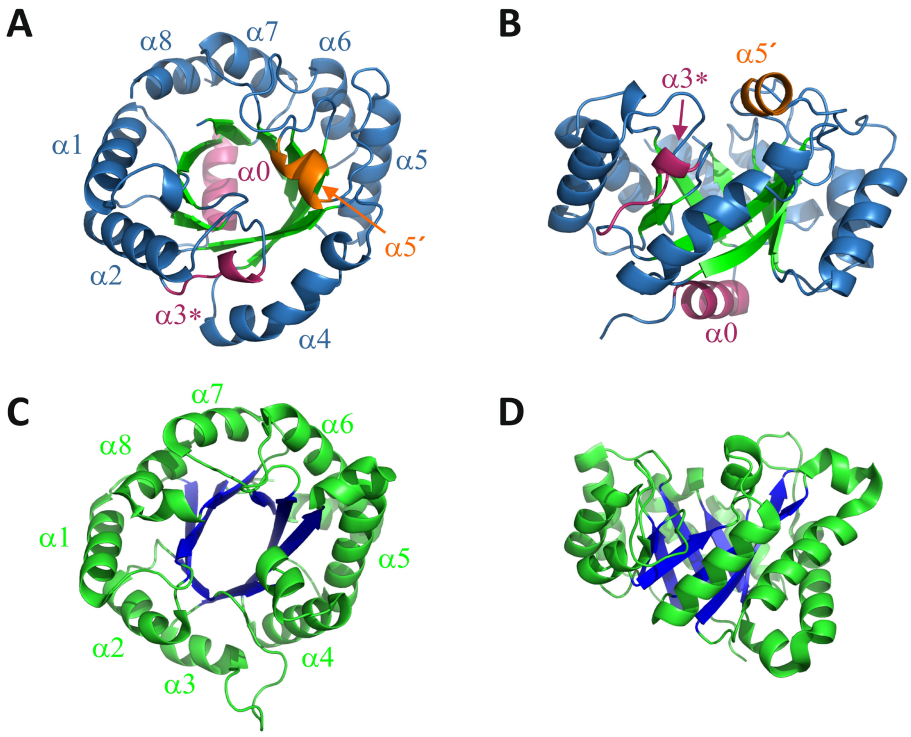


Figure 5 - Structural features of the $(\beta\alpha)_8$-barrel enzymes mtGGGPS and ggTIM
Secondary structure elements of the crystal structure of GGGPS from *Methanothermobacter thermautotrophicus* (mtGGGPS; A and B; pdb-code: 4mm1) and triosephosphate isomerase from *Gallus gallus* (ggTIM; C and D; pdb-code: 1tim; Banner *et al.*, 1975) are shown as ribbon diagrams. Top-view (**A**) and side-view (**B**) of mtGGGPS: central β-sheets and the surface-exposed α-helices are depicted in green and blue, respectively; the additional α-helix α0 and the flexible "swinging-door" helix α3* are in magenta; the additional small α-helix α5´ in orange. Only one protomer of the homohexamer is shown for clarity. Top-view (**C**) and side-view (**D**) of ggTIM: central β-sheets and the surface-exposed α-helices are depicted in blue and green, respectively. Only one protomer of the homodimer is shown for clarity.

The canonical $(\beta\alpha)_8$-barrel consists of eigth units each containing one β-strand and one α-helix. In each unit the C-terminus of the β-strand is linked via a βα-loop to the N-terminus of the corresponding α-helix. Eight units forming one $(\beta\alpha)_8$-barrel domain are connected by an αβ-loop linking the C-terminal end of α-helix n with the N-terminus of β-strand n+1 (Sterner and Höcker, 2005). The $(\beta\alpha)_8$-barrel commonly features a "catalytic face" localized at the C-terminal ends of the β-strands as well as in the βα-loops and a "stability face" being formed by the core structure and the opposite end of the barrel including the αβ-loops. About 50 % of known enzymes exhibiting a $(\beta\alpha)_8$-barrel fold among them the GGGPS enzymes need divalent metal ions for activity, which in prenyltransferases are commonly coordinated by conserved aspartates (Guo *et al.*, 2005; Guo *et al.*, 2004; Tarshis *et al.*, 1996; Vandermoten *et al.*, 2009). Comparison of the structural features of GGGPS enzymes with the first discovered $(\beta\alpha)_8$-barrel enzyme ggTIM revealed some

differences (Figure 5). GGGPS enzymes exhibit an additional helix (α0) at the N-terminus before the emergence of the first β-strand (Figure 5*A* and *B*). A sequence similarity network (SSN; Peterhoff *et al.*, 2014) and multiple sequence alignments (MSAs; Payandeh *et al.*, 2006) manifested that this helix is mainly built of positively charged and hydrophobic residues and therefore a contribution to the interaction of the enzyme with the membrane via this helix is presumed. Another specific feature is the substitution of α-helix 3 through a highly flexible loop (α3*). In GGGPS enzymes the average diameter of the barrel is a bit wider than the typical $(\beta\alpha)_8$-barrel fold mainly due to α-helix 4 and 5, which are kind of distorted to the outside of the barrel. Between α-helix 4 and 5 a small additional α-helix 5´ is inserted. The hydrophobic cavity essential for binding of the hydrophobic substrate is formed by the α-helices 3*, 4 and 5´ and the β-sheets 4 and partial 5 (Payandeh *et al.*, 2006; Ren *et al.*, 2013). α-helices 3* was assigned the function of a "swinging door", which allows for the access of the polyprenyl pyrophosphate. Moreover, α-helix 4 and 5´ are part of interfaces to other protomers of the complex (Payandeh *et al.*, 2006; Peterhoff *et al.*, 2014; Ren *et al.*, 2013). Based on structural similarity, the GGGPS is closely related to other $(\beta\alpha)_8$-barrel enzymes like the phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (HisA) and the cyclase subunit of the imidazole glycerol phosphate synthase (HisF; Holm and Sander, 1996), which exhibit a remarkable degree of internal two-fold symmetry and are both proposed to have evolved by duplication and fusion of $(\beta\alpha)_4$-half-barrels (Lang *et al.*, 2000; Payandeh and Pai, 2007). Consequently, Payandeh and Pai (2007) postulated that GGGPS evolved by gene duplication and subsequent fusion of a $(\beta\alpha)_4$-barrel ancestor protein or by duplication of an ancient *hisF*-like gene.

3.4 Bacterial homologues to the archaeal key enzymes G1PDH and GGGPS

Homology searches for proteins with significant sequence simiarities to the archaeal G1PDH and GGGPS facilitated the identification of both enzymes and revealed that also certain bacterial species as Firmicutes and Bacteroidetes possess G1P-based ether lipids (Guldan *et al.*, 2011; Guldan *et al.*, 2008; Payandeh *et al.*, 2006; Pereto *et al.*, 2004). The G1P producing homologue to the archaeal G1PDH in *B. subtilis* is AraM (Guldan, 2007;

Guldan *et al.*, 2008). AraM exhibits 21-24 % sequence identity to the archaeal G1PDH enzymes. G1PDH enzymes in common do not show any structural similarity to G3PDH, but are a close homologue of the glycerol dehydrogenase (GDH; Daiyasu *et al.*, 2002; Daiyasu *et al.*, 2005; Han and Ishikawa, 2005; Koga *et al.*, 1998; Pereto *et al.*, 2004; Yokobori *et al.*, 2016). AraM catalyzes the same $NADH^+$-dependent reduction of DHAP as G1PDH enzymes with a difference in the used metal cofactors. Whereas AraM needs Ni^{2+} to be active, archaeal G1PDH uses Zn^{2+}. Homology searches for GGGPS enzymes in Bacteria revealed that also Gram-positive Bacteria like *B. subtilis* contain polyprenylglyceryl phosphate synthases, namely the heptaprenylglyceryl phosphate synthase PcrB (HepGPS; Guldan, 2010; Guldan *et al.*, 2011). PcrB shares with archaeal GGGPS the well known $(\beta\alpha)_8$-barrel fold. Moreover, both enzymes catalyze almost identical reactions. For the reaction mechanism of prenylation either an electrophilic alkylation or a nucleophilic substitution has been proposed (Chen *et al.*, 1993; Jost *et al.*, 2010; Long *et al.*, 2002; Ren *et al.*, 2013; Ren *et al.*, 2012; Soderberg *et al.*, 2001; Zhang and Poulter, 1993b). In case of an electrophilic alkylation a highly electrophilic allylic carbocation is built, which subsequently alkylates the prenyl acceptor. Whereas in a nucleophilic substitution the C_3-OH group of G1P has to be activated to react as nucleophile on the C_1 of polyprenyl pyrophosphate with pyrophosphate acting as leaving group. With respect to the polyprenyl moiety, archaeal GGGPS enzymes favor the geranylgeranyl pyrophosphate (GGPP, C_{20}), whereas PcrB prefers the longer heptaprenyl pyrophosphate (HepPP, C_{35}; Badger *et al*., 2005; Doud *et al*., 2011; Peterhoff *et al*., 2012; Peterhoff *et al*., 2014; Figure 6).

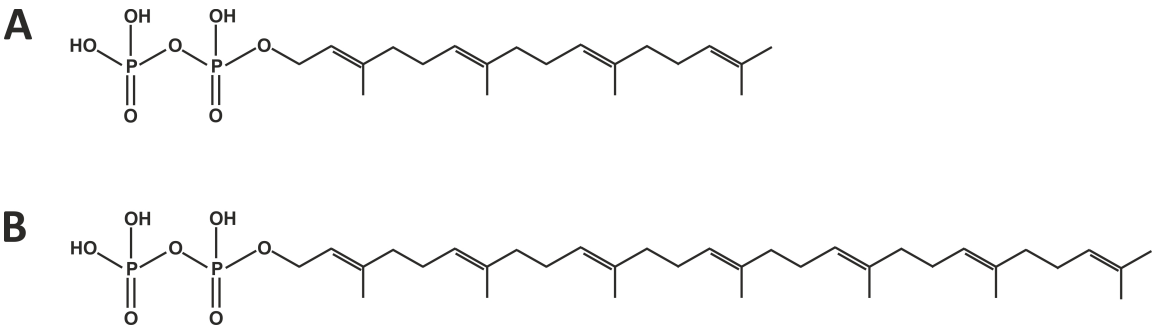


Figure 6 - Polyprenyl pyrophosphate substrates GGPP and HepPP
The favored substrate of archaeal GGGPS enzymes is GGPP (**A**), whereas the preferred substrate of PcrB is HepPP (**B**).

Comparison of the archaeal GGGPS from *Archaeoglobus fulgidus* (afGGGPS) with PcrB displayed that both enzymes share a dimeric oligomerization generated by an analogous dimerization interface (Peterhoff *et al.*, 2012). Structural superposition of crystal structures of afGGGPS (pdb-code: 2f6x) with the bacterial PcrB from *B. subtilis* (bsPcrB, pdb-code: 1viz) and MSAs revealed that the residues essential for binding of G1P are highly conserved and identified a large hydrophobic groove for binding of the long hydrophobic polyprenyl chain (Guldan *et al.*, 2011; Payandeh *et al.*, 2006; Figure 7).

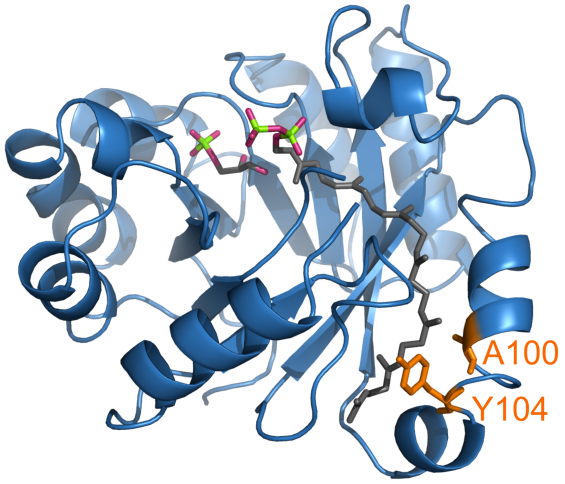


Figure 7 – Crystal structure of bsPcrB with modeled ligands
The crystal structure of bsPcrB (pdb-code: 1viz) is depicted as ribbon diagram. Important residues are displayed as sticks in orange. G1P and HepPP are shown in stick representation with carbon atoms in grey, oxygen atoms in magenta and phosphorus atoms in green. The figure was modified from Guldan *et al.* (2011).

In GGGPS enzymes, G1P is bound near the top inner rim of the β-barrel (Figure 7) and its phosphate moiety is coordinated by the very common standard phosphate-binding motif (Nagano *et al.*, 2002; Payandeh *et al.*, 2006; Vega *et al.*, 2003), which is constituted by side chains and backbone amino groups of the loops βα6, βα7 and βα8 (Peterhoff *et al.*, 2014). The nature of the hydrophobic tunnel important for binding of the polyprenyl moiety, especially substrate length determination, was intensely examined. Whereas afGGGPS exhibits a highly conserved tryptophan at position 99, annotated HepGPS enzymes show small aliphatic residues mostly alanine at the corresponding position (A100 in bsPcrB; Figure 7; Peterhoff *et al.*, 2014; Peterhoff *et al.*, 2012). Mutagenesis experiments elucidated that the conserved tryptophan is responsible for substrate length limitation (Guldan *et al.*, 2011) and is therefore also called “hydrocarbon ruler” or “limiter residue”

as already found in other prenyl- and acyltransferases (Ahn *et al.*, 2004; Liang *et al.*, 2002; Ohnuma *et al.*, 1996; Payandeh *et al.*, 2006; Tarshis *et al.*, 1996; Wyckoff *et al.*, 1998).

PcrB can bind and react with GGPP and shorter polyprenyl pyrophosphates besides its preferred substrate HepPP, albeit with lower efficiency (Guldan *et al.*, 2011). Such a catalytic promiscuity with respect to shorter substrates compared to the native one has been shown already for short-chain prenyltransferases, which catalyze condensation reactions up to C_{25} (Copley, 2003; Vandermoten *et al.*, 2009). Ren *et al.* (2013) proposed a limiter residue Y104 in bsPcrB, which constricts the substrate acceptance to a maximum of seven isoprene units (Figure 7). This position is randomly occupied in GGGPS sequences, whereas in almost all HepGPS sequences either a tyrosine or a phenylalanine can be found. The mutation of this position to alanine resulted in acceptance of even longer polyprenyl pyrophosphates (C_{40}) than the natural substrate C_{35} (Ren *et al*., 2013).

3.5 A new biosynthesis pathway of ether lipids in *Bacillus subtilis*

The discovery that also certain bacterial species possess G1P-based ether lipids implied to analyze if the same biosynthesis pathway is realized in these bacteria. Therefore *in vivo* radiolabeling experiments were performed (Guldan *et al.*, 2011). *B. subtilis* absorbes substantial amounts of radioactively labeled G1P present in the growth medium (Guldan *et al.*, 2011). Lipid extraction (Kates, 1986) followed by thin layer chromatography enabled the analysis of the *in vivo* products of bsPcrB. Reversed phase high pressure liquid chromatography followed by mass spectrometry and NMR studies revealed that radiolabeled G1P gets connected to HepPP producing HepGP, which gets subsequently dephosphorylated to HepG and than acetylated *in vivo* (Guldan *et al.*, 2011). After identification of both bacterial homologous enzymes to archaeal G1PDH and GGGPS (AraM and PcrB) and identification of the subsequently synthesized products a new biosynthesis pathway for archaea-like ether lipids in Bacteria was proposed (Figure 8).

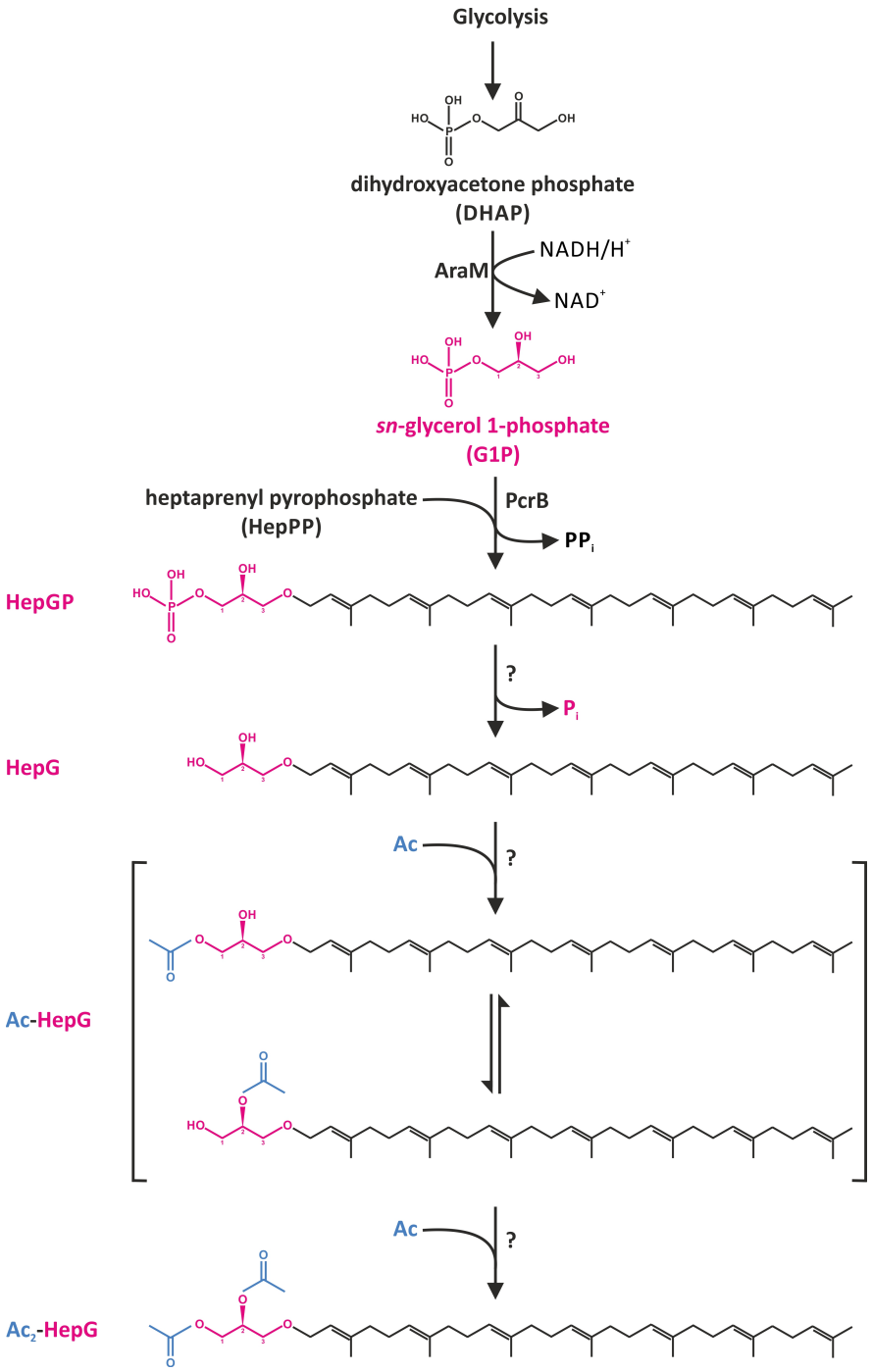


Figure 8 - Biosynthesis of G1P-based ether lipids in *B. subtilis*
A HepPP is attached to G1P (magenta) by PcrB resulting in heptaprenylglyceryl phosphate (HepGP). It could be shown that HepGP gets dephosphorylated to heptaprenyl glycerol (HepG) and subsequently acetylated at the two hydroxyl groups of G1P (acetyl group: blue) resulting in Ac-HepG and Ac_2-HepG. As discussed in the text, the acetyl group of Ac-HepG most likely swaps spontaneously between the two hydroxyl groups by acyl migration. P_i, orthophosphate; PP_i, pyrophosphate; Ac, acetyl group, $NADH/H^+$: nicotinamide adenine dinucleotide (reduced form), NAD^+: nicotinamide adenine dinucleotide (oxidized).

AraM produces the first substrate of PcrB (G1P) through reduction of DHAP. The second substrate HepPP is provided by the heptaprenyl pyrophosphate synthase (HepPPS) in *B. subtilis*, which catalyzes the consecutive condensation of four IPP molecules to farnesyl

pyrophosphate (FPP; Kobayashi *et al.*, 2003; Ogura and Koyama, 1998; Zhang *et al.*, 1998). HepPPS can not catalyze the formation of FPP itself, but is supplemented by the farnesyl pyrophosphate synthase, which occurs in all bacteria (Ogura and Koyama, 1998). The identification of the enzymes catalyzing the dephosphorylation of HepGP and subsequent acetylation of HepG was part of this doctoral thesis.

3.6 Comprehensive analysis of the GGGPS enzyme family

With the discovery of GGGPS enzymes also in bacterial species, the request for a comprehensive study of the diverse GGGPS enzyme family members emerged. Peterhoff *et al.* (2014) visualized the variability within the GGGPS family by applying the sequence similarity network (SSN) technique and calculating a phylogenetic tree (Figure 9).

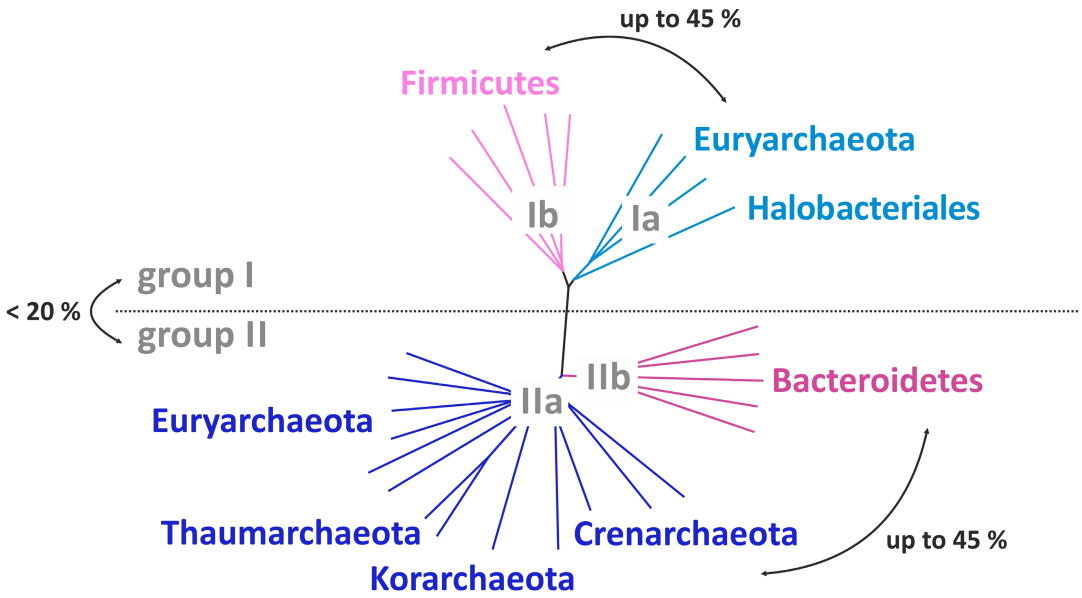


Figure 9 - Phylogenetic tree of the GGGPS enzyme family
The phylogenetic tree was generated on the basis of a multiple sequence alignment performed by Peterhoff *et al.* (2014). Latin number: group I and II, archaeal species (a) and bacterial species (b). Archaeal and bacterial subclusters are depicted in blue and magenta, respectively. Group I and group II share less than 20 % sequence identity, whereas the sequences in between the same group can be up to 45 %. The figure was modified from Peterhoff *et al.* (2014).

The combined analysis revealed that all Archaea except the Nanoarchaeota possess a gene coding for GGGPS as expected for a key enzyme in the evolutionary separation of the domain of Archaea (Payandeh and Pai, 2007). *Nanoarchaeum equitans* obtains its lipid membranes from its host *Ignicoccus hospitalis* (Jahn *et al.*, 2004; Peterhoff *et al.*, 2014) and therefore does not need an own lipid biosynthesis pathway. The splitting of the

GGGPS family enzymes into two groups, both composed of bacterial and archaeal sequences has already been known from earlier sequence comparisons. Group I and group II enzymes exhibit less than 20 % sequence identity among each other (Boucher *et al.*, 2004; Guldan *et al.*, 2011; Nemoto *et al.*, 2003; Peterhoff *et al.*, 2014). Group I assembles sequences from some Euryarchaeota such as *Archaeoglobus fulgidus* and all Halobacteria (Ia) sharing sequence identities of up to 45 % with the bacterial HepGPS deriving from Firmicutes (Ib equivalent to PcrB; Peterhoff *et al.*, 2014), whereas group II consists of representatives from all archaeal species (IIa) and some bacterial species, mainly Bacteroidetes (IIb). Much is known about the structure and function of group I GGGPS enzymes (Chen *et al.*, 1993; Guldan *et al.*, 2011; Nemoto *et al.*, 2003; Payandeh *et al.*, 2006; Payandeh and Pai, 2007; Peterhoff *et al.*, 2012; Soderberg *et al.*, 2001; Zhang and Poulter, 1993b). However, until recently, less information was available for the group II representatives. To elucidate structural features and substrate specificities of the GGGPS family in total, in a comprehensive study Peterhoff *et al.* (2014) characterized representative proteins from all branches of the phylogenetic tree. *In vitro* and *in vivo* activity tests proofed that bacterial group II enzymes exhibit a polyprenylglyceryl phosphate synthase function similar to the archaeal GGGPS with GGPP being the preferred substrate, whereas bacterial group I enzymes strongly favor HepGP. Two crystal structures of group II GGGPS enzymes have been resolved, the structure of GGGPS from *Methanothermobacter thermautotrophicus* (mtGGGPS, group IIa, pdb-code: 4mm1) and from *Flavobacterium johnsoniae* (fjGGGPS, group IIb, pdb-code: 4jej). Structural comparisons revealed that the aspartate, which is important for catalytic activity of group I enzymes by coordinating the essential Mg^{2+} (Ren *et al*., 2013) is also conserved in group II enzymes (Peterhoff *et al.*, 2014). During analysis of the hydrophobic tunnel with respect to limiter residues, it was astonishing to find that group II enzymes, although preferring C_{20} as substrate, do not contain an aromatic limiter residue at position 99 as present in archaeal group I enzymes (W99 in afGGGPS; Payandeh *et al.*, 2006; Peterhoff *et al.*, 2014; Ren *et al.*, 2013). A search for an alternative limiter residue in fjGGGPS and mtGGGPS revealed an isoleucine at position 90 (fjGGGPS) and a valine at position 86 (mtGGGPS) responsible for substrate length determination. The mutagenesis of these positions to alanine or glycine yielded in increased activity with longer polyprenyl substrates (Peterhoff *et al.*, 2014).

3.7 Oligomerization state of the GGGPS family enzymes

The enzymes were also analyzed regarding their oligomerization state by size exclusion chromatography (SEC) combined with static light scattering (SLS). All analyzed proteins belonging to group I showed a dimeric oligomerization state, whereas among the group II enzymes both dimeric and hexameric oligomerizations states could be found (Figure 10; Peterhoff *et al*., 2014).

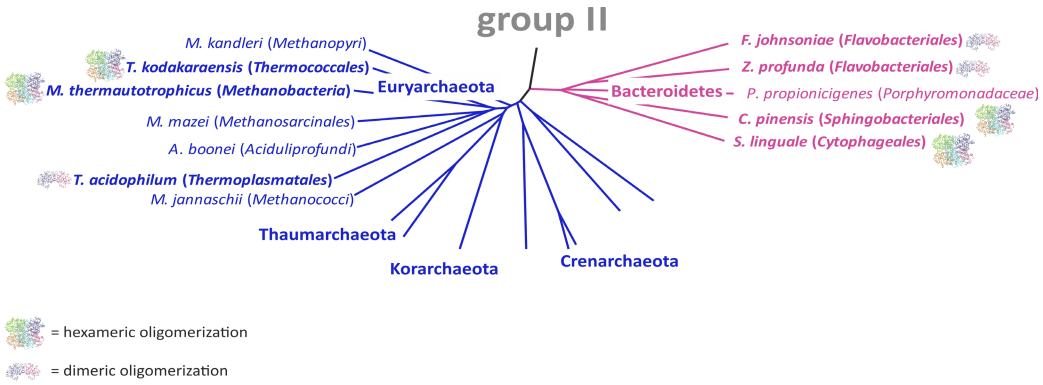


Figure 10 – Oligomerization state of group II GGGPS enzymes
The phylogenetic tree was generated on the basis of a multiple sequence alignment performed by Peterhoff *et al.* (2014). Representative sequences were selected from every order. For clarity, only group II is illustrated as rooted tree with group I serving as outgroup. The oligomerization state of some proteins was determined by size exclusion chromatography followed by static light scattering experiments. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit. Species whose GGGPS mutants were characterized in Peterhoff *et al.* (2014) are given in bold letters. Archaeal and bacterial organisms are depicted blue and magenta, respectively. The full species names are: *M. kandleri*: *Methanopyrus*, *T. kodakarensis*: *Thermococcus*, *M. thermautotrophicus*: *Methanothermobacter*, *M. mazei*: *Methanosarcina*, *A. boonei*: *Aciduliprofundum*, *T. acidophilum*: *Thermoplasma*, *M. jannaschii*: *Methanococcus*, *F. johnsoniae*: *Flavobacterium*, *Z profunda*.: *Zunongwangia*, *P. propionicigenes*: *Paludibacter*, *C. pinensis*: *Chitinophaga*, *S. linguale*: *Spirosoma*. The figure was taken and modified from Peterhoff *et al.* (2014).

In nature, many proteins fold into oligomeric structures and it has been estimated that about 75 % of proteins in *Escherichia coli* are forming oligomers with about 80 % of them being homooligomers (Doyle *et al.*, 2013; Goodsell and Olson, 2000). Also many $(\beta\alpha)_8$-barrels are known to assemble to homooligomers (Sterner and Höcker, 2005). Higher oligomerization can influence the thermal stability, the catalytic efficiency, allow for substrate channeling and provide other advantages. An increasing number of hyperthermophilic and thermophilic proteins have been found that exhibit a higher oligomerization state than their mesophilic counterparts (Vieille and Zeikus, 2001). Well-studied examples are the phosphoribosyl anthranilate synthase (TrpF) and TIM. In most mesophiles, TrpF forms monomers or is present as the bifunctional fusion protein TrpC-

TrpF (Luger *et al.*, 1989; Wilmanns *et al.*, 1992), whereas in thermophilic organisms, as for example *Thermotoga maritima,* TrpF is a highly thermostable homodimer (Sterner *et al.*, 1996). Analogously, eukaryotic, bacterial and mesophilic archaeal TIM enzymes build homodimers whereas in hyperthermophilic archaea higher oligomers as for example homotetramers are found (Kohlhoff *et al.*, 1996; Schramm *et al.*, 2001; Walden *et al.*, 2001). With respect to the abundance of homooligomeric proteins in nature it was proposed by various scientists that this form of protein assembly provides an evolutionary advantage based on greater adaptability, greater folding efficiency, a smaller genome size or greater stability (Doyle *et al.*, 2013; Goodsell and Olson, 2000; Liu *et al.*, 2002; Perica *et al.*, 2012). In case of the GGGPS enzymes of group II, consisting of either homodimers or homohexamers, it was part of this doctoral thesis to elucidate the function of higher homooligomerization.

Aim of this work

Compared to Gram-negative species like Bacteroidetes, which exhibit GGGPS with archaea-like activity, the polyprenylglyceryl phosphate synthases from Gram-positive species like *B. subtilis* possess enzymes with altered substrate specificity. These enzymes are called heptaprenylglyceryl phosphate synthases (PcrB in *B. subtilis*) and prefer C_{35} polyprenyl moieties (HepPP) instead of the shorter C_{20} polyprenyl moieties (GGPP). In contrast to archaeal G1P-based ether lipids, which get isoprenylated another time and further modified, the bacterial G1P-based ether lipids get subsequently dephosphorylated and acetylated *in vivo*. As such modifications have not been described so far, the invovled enzymes have been identified and characterized within this work. The results are presented and discussed in **chapter 6**.

Combining biochemical and computational methods, a comprehensive analysis for a systematic and comparative characterization of the GGGPS enzyme family revealed two separate groups (group I and group II) with less than 20 % sequence identity. During biochemical characterization of these two groups it became clear that all group I enzymes exhibit a dimeric oligomerization state, whereas in group II dimeric as well as hexameric oligomerization can be found. In the second part of this thesis, hexameric group II enzymes and dimeric mutants thereof were compared with respect to thermal stability and catalytic activity. The data were complemented by the comparative analysis of native dimers and hexamers from different archaeal and bacterial species. The results are presented and discussed in **chapter 7**.

Subsequent to the main part, **References** and **Supplement** can be found. The work is concluded by the **Acknowledgements**.

4 MATERIALS

4.1 Instrumentation

Autoclave:	
Series EC Stream Sterilizers	WEBECO, Selmsdorf
Balances:	
MC1	SARTORIUS, Göttingen
PL3000	METTLER TOLEDO, Gießen
SI-114	DENVER INSTRUMENT, Göttingen
CD spectro-polarimeter J-815	JASCO GmbH, Groß-Umstadt
Cell Density Meter Ultrospec 10	GE HEALTHCARE, München
Centrifuges:	
Centrifuge 5810R	EPPENDORF, Hamburg
Centrifuge 5415R	EPPENDORF, Hamburg
Sorvall RC 2B, 5C plus	DU PONT Instruments, Bad Homburg
Avanti J-26 XP	BECKMAN COULTER, Krefeld
Ultracentrifuge OPTIMA MAX-XP	BECKMAN COULTER, Krefeld
Centrifuge UNIVERSAL	HETTICH GmbH & Co KG, Tuttlingen
Centrifuge MIKRO 200R	HETTICH GmbH & Co KG, Tuttlingen
Chromatographic devices:	
ÄKTA basic better	GE HEALTHCARE, München
ÄKTA prime	GE HEALTHCARE, München
ÄKTA purifier 10	GE HEALTHCARE, München
Columns:	
HisTrap FF 5 ml	GE HEALTHCARE, München
Resource S	GE HEALTHCARE, München
Superdex 75 10/300 GL	GE HEALTHCARE, München
Superdex 200 10/300 GL	GE HEALTHCARE, München
Computer:	
Dell Optiplex Systems	DELL Inc., Round Rock, USA
Cyclone phosphorimager	PACKARD BioScience, Meriden, USA
Fluorescence spectrometer FP-6500	JASCO GmbH, Groß-Umstadt

Freezer -80° C	MDF-U72V, SANYO, Tokyo, Japan
Freezer -20° C	LIEBHERR, Nussbaumen
Gas burner, Gasprofi 2SCS	WLD-TEC GmbH, Göttingen
Gel electrophoresis system:	
Agarose gel electrophoresis chamber	
Agarose electrophoresis unit	HOEFER Pharmacia Biotech, USA
SDS electrophoresis chamber	
Mighty Small II	HOEFER Pharmacia Biotech, USA
Multi Gel Caster Assembling gel apparatus	GE HEALTHCARE, München
Glass pipettes and glassware	FISCHER SCIENTIFIC, Schwerte NOVOGLAS, Bern, Swiss SCHOTT, Mainz
Heating block-Thermostat	HBT-2 131 HLC, Bovenden
Incubator	BINDER GmbH, Tuttlingen
Liquid Szintillation Analyzer TRI-CARB 2900RT	PACKARD BioScience, Meriden, USA
Lyophilizer:	
Lyovac GT2	LEYBOLD-HERAEUS, Hürth
Concentrator plus	EPPENDORF, Hamburg
MicroCal VP-DSC	MALVERN Instruments, Malvern, UK
Magnetic stirrer:	
MR0, MR2000	HEIDOLPH, Kehlheim
MR1, MR3001 (heatable)	HEIDOLPH, Kehlheim
Microliter pipettes Research	EPPENDORF, Hamburg
Microwave	HMT 842C BOSCH, Nürnberg
Multichannel pipettes electronic:	
Research pro, 8-chanel, 5-100 µl, 20-300 µl, 50-1200 µl	EPPENDORF, Hamburg
Mosquito LCP roboter	TTP LABTECH, Melbourn, UK

Multi-Doc-It Digital Imaging System	UVP Inc., USA
PCR-cycler:	
Mastercycler personal	EPPENDORF, Hamburg
Mastercycler gradients	EPPENDORF, Hamburg
Peristaltic pump, Miniplus 2	GILSON Medical Electronics, France
pH-Meter Level1	INOLAB, Weilheim
PhosphorImager Film FLA-3000	FUJIFILM, Tokyo, Japan
Plate shaker Rocking Platform	BIOMETRA, Göttingen
Power supply unit:	
Power Pack P25	BIOMETRA, Göttingen
Power Supply EPS 301	GE HEALTHCARE, München
Quartz cuvettes:	
101-QS (layer thickness 10 mm)	HELLMA GmbH & Co. KG, Müllheim
105-QS (layer thickness 10 mm)	HELLMA GmbH & Co. KG, Müllheim
Shaking incubator:	
Certomat H	BRAUN Biotech, Melsungen
Certomat BS-1	BRAUN Biotech, Melsungen
Multitron	INFORS HT, Bottmingen, Swiss
Ultrafree-20 nanopore water system	MILLIPORE, Eschborn
Ultrasoni system:	
Branson Sonifier 250 D	HEINEMANN, Schwäbisch Gmünd
UV-Vis spectral photometer V650	JASCO GmbH, Groß-Umstadt
UV-Vis Biophotometer	EPPENDORF, Hamburg
Vakuum pump ME 2C	VACUUMBRAND, Wertheim
Vortex Genie 2	SCIENTIFIC IND., Bohemia, USA

4.2 Consumables

Easy-Xtal plates	QIAGEN, Hilden
Centrifugal Filter Device	
Amicon Ultra-15 (mwco: 10 kDa)	MILLIPORE, Bedford, USA

Deep-well plates, 96 well/2000 µl	EPPENDORF, Hamburg CORNING Inc., USA
Dialysis tubing	
Visking, 27/32, 14 kDa	ROTH GmbH &Co, Karlsruhe
Disposable syringes, Omnifix® 60 ml	BRAUN Biotech, Melsungen
Filter paper	WHATMAN, Maidstone, England
Membrane filter ME24 Ø47 mm; 0.2 µm	SCHLEICHER&SCHUELL, Dassel
Microtiter plates, 96 well/200 µl, flat bottom	
Vis, 96 well, sterile with lid	SARSTEDT, Nümbrecht
Nitrocellulose filter (Ø13 mm)	MILLIPORE, Eschborn
Parafilm „M“ Laboratory Film	PECHINEY, Menasha, USA
Pasteur pipettes	HIRSCHMANN, Ebermannstadt
PCR-tubes 0.2 ml	PEQLAB, Erlangen
Petri dish 94/16	GREINER bio-one, Nürtingen
Pipette tips	SARSTEDT, Nümbrecht
Plastic cuvettes:	
½ microcuvettes, UV-transparent	SARSTEDT, Nümbrecht
1ml cuvettes	SARSTEDT, Nümbrecht
Plastic tubes:15 ml, 50 ml	SARSTEDT, Nümbrecht
Reaction vessels 1.5 ml, 2 ml	ROTH, Karlsruhe EPPENDORF, Hamburg
Reaction vessel with screw-cap, 2 ml	SARSTEDT, Nümbrecht
Silica60 plates	MERCK, Darmstadt
Syringe filter, pore size 0.2 µm, 0.45 µm	RENNER GmbH, Daunstadt

4.3 Chemicals

All chemicals used were p.a. graded and purchased from the companies listed below.

ALFA AESAR	Karlsruhe
APPLICHEM GmbH	Darmstadt

BIO101 Inc.	Carlsbad, USA
BIO-RAD LABORATORIES	Hercules, USA
BIOZYM	Hessisch Oldendorf
BODE CHEMIE	Hamburg
BOEHRINGER MANNHEIM	Mannheim
CARL ROTH GMBH & Co. KG	Karlsruhe
DIFCO	Dreieich
FLUKA	Neu-Ulm
GE HEALTHCARE	München
GERBU Biotechnik GmbH	Gailberg
GIBCO/BRL	Eggestein
HARTMANN ANALYTIC	Braunschweig
MERCK	Darmstadt
MP BIOCHEMICALS	Illkirch, France
NATIONAL DIAGNOSTICS	Simerville, USA
OXOID	Wesel
RIEDEL-DE HAEN	Seelze
ROCHE DIAGNOSTICS	Mannheim
SERVA	Heidelberg
SIGMA-ALDRICH	Deisenhofen
VWR	Leuven, Belgium

4.4 Kits

4.4.1 Kits for molecular biology

GeneJET™ Plasmid Miniprep Kit	MBI FERMENTAS, St. Leon-Rot
GeneJET™ Gel Extraction Kit	MBI FERMENTAS, St. Leon-Rot
SuperSignal West Pico Kit	PIERCE, Rockford, USA
His SpinTrap™ Kit	GE HEALTHCARE, München
ROTI®-Spin MINI 10 Kit	CARL ROTH, Karlsruhe™

4.4.2 Kits for protein crystallization

Morpheus®	MOLECULAR DIMENSIONS, Suffolk, UK
MIDAS™	MOLECULAR DIMENSIONS,

	Suffolk, UK
PGA Screen™	MOLECULAR DIMENSIONS, Suffolk, UK
PEGRx™1+2	MOLECULAR DIMENSIONS, Suffolk, UK
The structure screen combination	MOLECULAR DIMENSIONS, Suffolk, UK

4.5 Enzymes

4.5.1 Molecular biology

Alkaline phosphatase (CIP)	NEW ENGLAND BIOLABS, Frankfurt am Main
DNA polymerases:	
GoTaq® Flexi	PROMEGA, Mannheim
Pwo	ROCHE DIAGNOSTICS, Mannheim
Pfu	PEQLAB, Erlangen
Lysozyme	SIGMA-ALDRICH, Deisenhofen
Restriction endonucleases	MBI FERMENTAS, St-Leon-Rot NEW ENGLAND BIOLABS, Frankfurt am Main
T4-DNA ligase	MBI FERMENTAS, St-Leon-Rot
Thrombin	SIGMA-ALDRICH, Deisendorf

4.5.2 G1P synthesis

Aldolase	EC 4.1.2.13, from *Oryctolagus cuniculus*
Phosphogluco isomerase type III	EC 5.6.1.9, from *Saccharomyces cerevisiae*
Hexokinase type F-300	EC 2.7.1.1., from *Saccharomyces cerevisiae*
Phosphofructokinase type VII	EC 2.7.1.11, from *Bacillus stearothermophilus*

Triosephosphate isomease type VII-S	EC 5.3.1.1, from *Oryctolagus cuniculus* SIGMA-ALDRICH, Deisendorf
Creatine kinase	EC 2.7.3.2, from *Oryctolagus cuniculus* ROCHE DIAGNOSTICS, Mannheim

4.5.3 Colorimetric assay for phosphate detection

PNPase	Purine nucleoside phosphorylase from microorganisms
PPase	Pyrophosphatase from *Escherichia coli*
XOD	Mircrobial xanthine oxidase SIGMA-ALDRICH, Deisendorf

4.6 Antibodies

Anti-$(His)_6$ tag antibody, peroxidase conjugated	ROCHE DIAGNOSTICS, Mannheim

4.7 Bacterial strains

4.7.1 Overview about used *E. coli* strains

4.7.1.1 ***E. coli* Turbo** (NEW ENGLAND BIOLABS, Frankfurt am Main)

F' *proA*$^{+}$*B*$^{+}$ *lacI*^{q} Δ*(lacZ)*M15*/fhuA2* Δ*(lac-proAB) glnV gal R(zgb-210::Tn10)* Tet^{S} *end*A1 *thi-1*Δ*(hsdS-mcrB)5*

E. coli Turbo cells are T1-phage resistant due to a deletion of *fhu*A2. As the *rec*A function of the strain is intact, *E. coli* Turbo cells grow fast and form visible colonies after 8 h incubation at 37° C.

4.7.1.2 ***E. coli* BL21-CodonPlus (DE3) RIPL** (STRATAGENE, La Jolla, USA)

B F´ *omp*T *hsd*S($r_B^-m_B^-$) *dcm*$^{+}$ Tet^{R} *gal* λ(DE3) *end*A Hte [*arg*U *pro*L Cam^{R}] [*arg*U *ile*Y *leu*W Strep^{R}/Spec^{R}]

E. coli T7 BL21-CodonPlus (DE3) RIPL cells contain extra copies of rare-codon tRNAs. They enable the efficient heterologous expression of proteins from genes, which lack the codon

usage of *E. coli*. The cells have a pACYC plasmid with extra copies of the *arg*U and *pro*L tRNA genes, and a pSC101 plasmid with extra copies of the *arg*U, *ile*Y and *leu*W tRNA genes. The pACYC plasmid confers resistance to chlorampehnicol and the pSC101 plasmid confers resistance to streptomycin and spectinomycin.

4.7.1.3 ***E. coli* T7 Express** (NEW ENGLAND BIOLABS, Frankfurt am Main)

fhuA2 lacZ::T7 gene1 [*lon] ompT gal sulA11 R(mcr-73::miniTn10*-Tet^S*)*2 [*dcm*] *R(zgb-210::Tn10*--Tet^S*) endA1 Δ(mcrC-mrr) 114::IS10*

E. coli T7 Express is a derivative of *E. coli* BL21(DE3), which has the gene for T7 polymerase in the *lac* operon. They are T1-phage resistent due to a deletion of *fhu*A2. The cells exhibit tetracycline resistance.

4.7.2 Overview about used *B. subtilis* strains

Table 1 - *B. subtilis* strains used in this thesis

strain	features[1]	origin reference	comment
Bacillus subtilis subsp. subtilis str. 168	Marburg; ATCC 6051, tryptophan auxotrophic	Zeigler et al. (1958)	In this thesis called *B. subtilis* wild type (wt).
***Bacillus subtilis subsp. subtilis str. 168* ΔpcrB**	Δ*pcrB* with pMUTIN4, Ery^R	Kobayashi *et al.* (2003)	In this thesis called *B. subtilis* Δ*pcrB*
***Bacillus subtilis subsp. subtilis str. 168* ΔA332**	Δ*yvoF* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*yvoF*
***Bacillus subtilis subsp. subtilis str. 168* ΔA391**	Δ*yvoF* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*yvoF*
***Bacillus subtilis subsp. subtilis str. 168* ΔB865**	Δ*maa* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*maa*

***Bacillus subtilis subsp. subtilis str. 168* ΔA331**	Δ*ppaX* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*ppaX*
***Bacillus subtilis subsp. subtilis str. 168* ΔA330**	Δ*yvoD* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*yvoD*
***Bacillus subtilis subsp. subtilis str. 168* ΔA388**	Δ*lgt* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*lgt*
***Bacillus subtilis subsp. subtilis str. 168* ΔA328**	Δ*hprK* with pMUTIN4, Ery^R	National BioResource Project[2]	In this thesis called *B. subtilis* Δ*hprK*

[1] Genetic nomenclature of Bachmann (1990)
[2] National BioResource Project from National Institute of Genetics, Microbial Genetics Laboratory in Japan

4.7.3 pET vectors

Genes inserted into the multiple cloning site (MCS) of pET vectors (plasmids for expression by T7 RNA Polymerase) are transcribed by the RNA-polymerase of the phage T7 (Studier et al., 1990). The expression of genes takes place in special *E. coli* strains, which carry a chromosomal copy of the T7 RNA polymerase. The expression of the T7 RNA polymerase gene proceeds under the control of the *lac*UV5 promotor operator and is induced by the addition of IPTG. The gene for the *lac* repressor (*lac*I), which is required for suppression of gene expression in the absence of induction, is located on the plasmid and is constitutively expressed.

The pET21a vector encodes an optional C-terminal $(His)_6$ tag sequence and confers ampicillin resistance. The pET28a vector encodes an optional C-terminal $(His)_6$ tag sequence and confers kanamycin resistance. The pET28atrxN vector is a pET28a derivative and provides an optional C-terminal $(His)_6$ tag sequence plus a N-terminally fused thioredoxin (TrxA). TrxA is removable by thrombin digestion leaving a N-terminal Gly-Ser extension to the protein. The construct pET28atrxN was generated by amplifying the thioredoxin gene *trx*A from genomic *E. coli* DNA and inserting it between the *Nco*I and *Nde*I restriction sites in pET28a. pET28atrxN confers kanamycin resistance (Peterhoff *et al.*, 2014).

4.7.4 Plasmids used in this thesis

Table 2 - Overview about plasmids used in this thesis[4]

vector	gene	cloning
pET21a	bs*yvoF*[1]	cloned with *Nde*I/*Xho*I, all genes include a C-terminal $(His)_6$ tag
	ba*yvoF*	
	bs*maa*	
	bs*phoB*	
	cp*gggps*[2]	
	cp*gggps*_Y143A[2]	
	sl*gggps*[2]	
	fj*gggps*[2]	
	zp*gggps*[2]	
	tk*gggps*[2]	
	tk*gggps*_W143A[2]	
	ta*gggps*[2]	
	mt*gggps*[2]	
	mt*gggps*_W141A[2]	
	mt*gggps*_K146A[3]	
	mt*gggps*_I155A[3]	
	mt*gggps*_D144A[3]	
	mt*gggps*_M168E[3]	
	mt*gggps*_A162E[3]	
	mt*gggps*_I107E[3]	
	mt*gggps*_Y105A	
	mt*gggps*_D57R	
	mt*gggps*_R88E	

	mt*gggps*_A162E_W141A	
	mt*gggps*_A162E_Y105A	
	mt*gggps*_A162E_D57R	
	mt*gggps*_A162E_R88E	
pET28atrxN	bs*yvoF*	cloned with *Nde*I/*Xho*I, all genes include a C-terminal $(His)_6$ tag and an N-terminal Trx tag
	ba*yvoF*	
	bs*maa*	
	bs*phoB*	
pET28a	ba*yvoF*	cloned with *Nde*I/*Xho*I, the gene includes a C-terminal $(His)_6$ tag

[1]This gene has already been cloned previously (Linde, 2013), [2]The genes of these enzymes have already been cloned previously (Peterhoff *et al.*, 2014), [3]The genes of these mutants have already been cloned previously (Kilu, 2014), [4]For clarity, all genes of polyprenylglyceryl phosphate synthases are called *gggps*.

4.8 Oligonucleotides

In all oligonucleotides the newly introduced restriction sites or adapter sequences are written in **bold** and the codon for a newly introduced amino acid is underlined. The oligonucleotides were ordered from METABION or BIOMERS.

4.8.1 Vector specific amplification and sequencing primers

Table 3 – Sequencing primers used in this thesis.

oligonucleotide	**sequence, 5'→ 3'**	**usage**
5´-T7Promotor	TAATACGACTCACTATAGGG	sequencing of pET vectors
3´-T7Terminator	GCTAGTTATTGCTCAGCGG	sequencing of pET vectors
5´-pMUTIN_up	CATCCAGAACAACCTCTGCT	sequencing of pDG148_*Stu* vector
3´-pMUTIN_down	CCCAGTCGACCCGTAATCT	sequencing of pDG148_*Stu* vector
3´-pDG148_down	CTGCGACATCGTATAACGTTACT	sequencing of pDG148_*Stu* vector

4.8.2 Amplification and mutagenic primers

Table 4 - Amplification and mutagenic primers for genes used in this thesis

oligonucleotide	sequence 5´→ 3´	usage
5´bs*yvoF Nde*I	CTC**CATATG**AGAAAAACAGATCGTCATCC	amplification of bs*yvoF* for cloning in pET21a[1] or pET28atrxN
3´bs*yvoF Xho*I	TGCG**CTCGAG**TTCAGCGGACTTTTTCAATC	
5´ba*yvoF Nde*I	CTC**CATATG**CGACGGACAACGCGCTATCC	amplification of ba*yvoF* for cloning in pET21a, pET28a or pET28atrxN
3´ba*yvoF Xho*I	TGCG**CTCGAG**GTATGAACCTTCTCTAGCC	
5´bs*maa Nde*I	CTC**CATATG**CTGAGGACAGAAAAAGAAAAAATGGC	amplification of bs*maa* for cloning in pET21a or pET28atrxN
3´bs*maa Xho*I	TGCG**CTCGAG**CAATTGCTTCAAAATCCTTGCGGGG	
5´bs*phoB Nde*I	CTC**CATATG**AAAAAATTCCCGAAGAAATTAC	amplification of bs*phoB* for cloning in pET21a or pET28atrxN
3´bs*phoB Xho*I	TGCG**CTCGAG**CTTATCGTTAATCTTAATGTTG	
5´cp*gggps Nde*I	ATCG**CATATG**CACAATAAAATATACAATTC	amplification of cp*gggps* for cloning in pET21a[2]
3´cp*gggps Xho*I	ATCG**CTCGAG**TTTTAACACAGGTGCAGCT	
5´cp*gggps*_Y143A	GCGCCGACTACTGTATCTGCGATCAGTAATACC	QuikChange mutagenesis (QCM)[3] of cp*gggps*[2]
3´cp*gggps*_Y143A	GGTATTACTGATCGCAGATACAGTAGTCGGCGC	
5´sl*gggps Nde*I	ATCG**CATATG**ACTATACTCCGGGATTAT	amplification of sl*gggps* for cloning in pET21a[2]
3´sl*gggps Xho*I	ATCG**CTCGAG**CGCTTGAACAACGGATTG	
5´fj*gggps Nde*I	ATCG**CATATG**GAGCAAAAAATACTTACC	amplification of fj*gggps* for cloning in pET21a[2]
3´fj*gggps Xho*I	ATCG**CTCGAG**CGCTTGAACAACGGATTG	
5´zp*gggps Nde*I	ATCG**CATATG**CCTAAAATTTTAGATGCC	amplification of zp*gggps* for cloning in pET21a[2]
3´zp*gggps Xho*I	ATCG**CTCGAG**TTTAAATTCTCCATTTTCAAAGC	

5´tk*gggps Nde*I	CTAG**CATATG**CTCAAGCTTGGAAAAG	amplification of tk*gggps* for cloning in pET21a[2]
3´tk*gggps Xho*I	CTAG**CTCGAG**TCCCTTAACGCCCCTGT	
5´tk*gggps*_W143A	GGAGAAACCGTGGGCGCGGTCGGGGACGCAAAG	QCM of tk*gggps*[2]
3´tk*gggps*_W143A	CTTTGCGTCCCCGACCGCGCCCACGGTTTCTCC	
5´ta*gggps Nde*I	CAGGG**CATATG**ATGACCGTCCTCGAAGACATG	amplification of ta*gggps* for cloning in pET21a[2]
3´ta*gggps Xho*I	CCGGTA**CTCGAG**TTGTATTTTTGATATTCCTAC	
5´mt*gggps Nde*I	ATCG**CATATG**TTCAAAATGAAGGTTGAAGATT	amplification of mt*gggps* for cloning in pET21a[2]
3´mt*gggps Xho*I	ATCG**CTCGAG**GACGGACCCCATACCCTCAACGATCTCC	
5´mt*gggps*_W141A	GGGGGTACGGTTGGAGCGGTCGGTGACACCAAG	QCM of mt*gggps*[2]
3´mt*gggps*_W141A	CTTGGTGTCACCGACCGCTCCAACCGTACCCCC	
5´mt*gggps*_K146A	GGATGGGTCGGTGACACCGCGCCGGTCCCCAGGAACAAAC	QCM of mt*gggps*[4]
3´mt*gggps*_K146A	GTTTGTTCCTGGGGACCGGCGCGGTGTCACCGACCCATCC	
5´mt*gggps*_I155A	AGGAACAAACCCGACGCGGCAGCCGCCTATGCC	QCM of mt*gggps*[4]
3´mt*gggps*_I155A	GGCATAGGCGGCTGCCGCGTCGGGTTTGTTCCT	
5´mt*gggps*_D144A	GTTGGATGGGTCGGTGCGACCAAGCCGGTCCCC	QCM of mt*gggps*[4]
3´mt*gggps*_D144A	GGGGACCGGCTTGGTCGCACCGACCCATCCAAC	
5´mt*gggps*_M168E	GCTGAATTCCTTGGCGAAAGGCTCTTCTACCTT	QCM of mt*gggps*[4]
3´mt*gggps*_M168E	AAGGTAGAAGAGCCTTTCGCCAAGGAATTCAGC	
5´mt*gggps*_A162E	GCCGCCTATGCCATGGAAGCTGAATTCCTTGGC	QCM of mt*gggps*[4]
3´mt*gggps*_A162E	GCCAAGGAATTCAGCTTCCATGGCATAGGCGGC	
5´mt*gggps*_I107E	ACAAACCCCTACTGGGAAATAGGTGCCCAGGCC	QCM of mt*gggps*[4]
3´mt*gggps*_I107E	GGCCTGGGCACCTATTTCCCAGTAGGGGTTTGT	

5´mt*gggps*_Y105A	AACTCCACAAACCCCGCCTGGATCATAG GTGCCC	QCM of mt*gggps*
3´mt*gggps*_Y105A	GGGCACCTATGATCCAGGCGGGGTTTG TGGAGTT	
5´mt*gggps*_D57R	GGAGGGTCAACCACACGCTCCAGTGAA CTTGATAACAC	QCM of mt*gggps*
3´mt*gggps*_D57R	GTGTTATCAAGTTCACTGGAGCGTGTG GTTGACCCTCC	
5´mt*gggps*_R88E	CCACGGGTGTCAGCGAATACGCCGATG CAATATTC	QCM of mt*gggps*
3´mt*gggps*_R88E	GAATATTGCATCGGCGTATTCGCTGACA CCCGTGG	
mt*gggps*_A162E_W141A	primer of mt*gggps*_W141A used on pET21a_mt*gggps*_A162E template	QCM of mt*gggps*_A162E
mt*gggps*_A162E_I107E	primer of mt*gggps*_I107E used on pET21a_mt*gggps*_A162E template	
mt*gggps*_A162E_Y105A	primer of mt*gggps*_Y105A used on pET21a_mt*gggps*_A162E template	
mt*gggps*_A162E_D57R	primer of mt*gggps*_D57R used on pET21a_mt*gggps*_A162E template	
mt*gggps*_A162E_R88E	primer of mt*gggps*_R88E used on pET21a_mt*gggps*_A162E template	

[1] This gene has already been cloned previously (Linde, 2013), [2]The genes of these enzymes have already been cloned previously (Peterhoff *et al.*, 2014), [3]QCM is described in chapter 5.3.4, [4]The genes of these mutants have already been cloned previously (Kilu, 2014).

4.9 Ladders and markers

The size of DNA fragments in agarose gels was determined with the help of the GeneRuler™ 1 kb Plus DNA Ladder (MBI FERMENTAS, St. Leon-Rot). The size of proteins in SDS-PAGE gels for Coomassie staining was determined with the help of the Unstained Protein Molecular Weight Marker (MBI FERMENTAS, St. Leon-Rot), for Western Blotting with the help of the PageRuler™ Prestained Protein Ladder (MBI FERMENTAS, St. Leon-Rot).

4.10 Buffers and solutions

Unless otherwise specified, buffers were filter-sterilized and stored at room temperature (RT). For all preparations, Milli-Q grade water (MILLIPORE) or double distilled water (ddH_2O) was used.

4.10.1 Buffers and solutions for working with *E. coli* and *B. subtilis*

Antibiotics

Dissolved and filter-sterilized antibiotics were placed at -20 °C for long-term storage.

Ampicillin (1000 x)	150 mg/ml ampicillin (sodium salt) dissolved in water
Chloramphenicol (1000 x)	30 mg/ml chloramphenicol dissolved in 100% EtOH
Erythromycin (1000x)	0.8 mg/ml erythromycin dissoveld in 70% EtOH
Kanamycin (1000 x)	75 mg/ml kanamycin dissolved in water

Further solutions

Glucose	20 % (w/v) glucose dissolved in water, filter-sterilized, and stored at RT
Glycerol (87%)	autoclaved and stored at RT
IPTG stock solution	0.5 M IPTG dissolved in water, filter-sterilized, and stored at -20 °C
$MgCl_2$	1 M $MgCl_2$, dissolved in water, filter-sterilized, and stored at RT
$MgSO_4$	1 M $MgSO_4$ dissolved in water, filter-sterilized, and stored at RT
TFB I buffer	100 mM KCl, 50 mM $MnCl_2$, 30 mM KAc, 10 mM $CaCl_2$, 15 % glycerol; stock solution for the single components were stored at 4 °C. A volume of 100 ml TFB I-buffer was prepared prior to use.

TFB II buffer	100 mM Tris/HCl, pH 7.0, 10 mM KCl, 75 mM $CaCl_2$, 15% glycerol; A volume of 100 ml TFB II-buffer was prepared prior to use.

4.10.2 Buffers and solutions for molecular biology

PCR dNTP solution (2 mM)	A solution of dNTP`s (2mM of each A, C, G, and T) was prepared and stored at -20 °C.
Agarose (1%)	5 g agarose dissolved in 500 ml 0.5 x TBE, boiled and stored at 60 °C
Ethidium bromide solution	10 mg/ml ethidium bromide (EtBr)
Sucrose color marker	60% (w/v) sucrose, 0.1% (w/v) bromphenol blue, 0.1% (w/v) xylencyanole FF dissolved in 0.5 x TBE
TBE (5x)	445 mM boric acid, 12.5 mM EDTA, 445 mM Tris (resulting pH-value: 8.15)

4.10.3 Buffers and solutions for working with proteins

Potassium phosphate (KP) buffer

1 M K_2HPO_4 (base) and 1 M KH_2PO_4 (acid), pH was adjusted by mixing

Tris/HCl buffer

1 M Tris, pH was adjusted by addition of fuming hydrochloric acid

HEPES buffer

1 M HEPES, pH 7.5

PIPES buffer

0.5 mM PIPES, pH 7.5

Resuspension buffer for YvoF

50 mM KP pH 7.5, 300 mM KCl, 10 mM imidazole, 10 mM $MgCl_2$, 9 mM CHAPS for DTNB-coupled assay

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT, 9 mM CHAPS for all other assays

Resuspension buffer for enhancing soluble YvoF

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT, 9 mM CHAPS

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT, 0.4 mM Triton X100

50 mM Tris/HCl pH 8.0, 1 M NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT

Resuspension buffer for MAT and PhoB

50 mM KP pH 7.5, 300 mM KCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 1 mM DTT

Resuspension buffer for all other proteins

50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole

<u>for irreversible heat inactivation studies</u>

50 mM KP pH 7.5, 300 mM KCl, 10 mM imidazole

<u>for all other assays</u>

Buffer for ion exchange chromatography

Equilibration buffer: 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

Elution buffer: 50 mM Tris/HCl pH 8.0, 3 M NaCl, 10 mM $MgCl_2$, 1 mM DTT

Buffer for size exclusion chromatography

For PhoB: 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

For YvoF: 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 5 mM $MgCl_2$

For all other proteins: 50 mM KP pH 7.5, 300 mM KCl

Buffer for dialysis and storage

For YvoF: 50 mM KP pH 7.5, 300 mM KCl, 10 mM $MgCl_2$

<u>for DTNB coupled assay</u>

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

<u>for all other assays</u>

For MAT: 50 mM KP pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT

	50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
For PhoB:	50 mM KP pH 7.5, 300 mM KCl, 10 mM $MgCl_2$, 1 mM DTT
	50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
	50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
For all other proteins:	50 mM KP pH 7.5 or 50 mM Tris/HCl pH 8.0

4.10.4 Buffers and solutions for SDS-PAGE

Ammonium persulfate (APS) solution	10 % (w/v) APS solution; filter-sterilized, and stored at -20 °C
Coomassie staining solution	0.2 % (w/v) Coomassie Brilliant Blue G-250 andR-250, 50 % (v/v) ethanol, 10 % (v/v) HAc; filtered and protected from light
Buffer for SDS-PAGE resolving gel	0.4 % (w/v) SDS, 1.5 M Tris/HCl pH 8.8
Buffer for SDS-PAGE stacking gel	0.4 % (w/v) SDS, 0.5 M Tris/HCl pH 6.8
SDS-PAGE electrophoresis buffer	0.1 % (w/v) SDS, 0.025 M Tris, 0.2 M glycine (resulting pH-value: 8.5)
SDS-PAGE sample buffer (5 x)	5 % (w/v) SDS, 25 % (w/v) glycerol, 12.5 % (v/v) β-mercaptoethanol, 0.025 % (w/v) bromphenol blue, 1.25 M Tris/HCl pH 6.8

4.10.5 Buffers and solutions for Western Blotting

Transfer buffer 1x	20 mM Tris/HCl pH 8.5, 150 mM glycine, 20 % (v/v) methanol, 0.02 % (w/v) SDS
10x phosphate buffered saline (PBS)	91 mM Na_2HPO_4, 17 mM NaH_2PO_4, 1.5 mM NaCl, pH 7.4 pH was adjusted by addition of NaOH

Wash solution PBS-T	1x PBS buffer with 0.1 % (v/v) Tween-20
Blocking buffer PBST-M	1x PBS-T wash solution with 5 % (w/v) milk powder

4.10.6 Solutions for paper - and thin layer chromatography

Preperative paper chromatography	isopropyl alcohol/ammonia/water, 6:3:1 (v/v/v)
Thin layer chromatography	ethyl acetate/hexane, 1:1 (v/v) acetone/chloroform/methanol/water, 15:2:2:1 (v/v/v/v)

4.11 Bacterial growth media for *E. coli* and *B. subtilis*

For sterilization, the medium was autoclaved for 20 min at 121 °C and 2 bar. For selective media, the corresponding antibiotics were added after cooling down the medium to 60 °C in terms of a filter-sterilized, 1000-fold concentrated stock solution.

Luria-Bertani (LB) medium

0.5 % (w/v) yeast extract, 1.0 % (w/v) NaCl, 1.0 % (w/v) tryptone

LB agar

LB medium plus 1.5 % (w/v) Bacto-Agar

SOB medium

0.5 % (w/v) yeast extract, 0.05 % (w/v) NaCl, 2.0 % (w/v) tryptone. After autoclaving, 10 mM $MgSO_4$, 10 mM $MgCl_2$, and 2.5 mM KCl (each filter-sterilized) were added prior to use.

4.12 Software

4.12.1 Local applications

ÄKTA Unicorn version 5.01 (318)	© GE HEALTHCARE
CLC main workbench 7.0.2	© CLC BIO
Corel Draw X6	© COREL Corp.
Endnote Version X7	© WINTERTREE Software Inc.

MS Office 2010	© 2010 MICROSOFT CORPORATION
Origin 7	© ORIGINLAB
OptiQuant 3.0	© PACKARD Instruments Co.
PyMOL v0.99	© DELANO SCIENTIFIC LCC.
SigmaPlot 12.5	© SPSS Inc.
Spectra Manager 1 and 2	© JASCO

4.12.2 Server supported applications

National BioResource Project	https://shigen.nig.ac.jp/bsub/resource
ProtParam	http://web.expasy.org/protparam/
BsubCyc (*B. subtilis* Datenbank)	https://bsubcyc.org
Basic Local Alignment Search Tool	https://blast.ncbi.nlm.nih.gov/Blast.cgi
EMBOSS Matcher	http://www.ebi.ac.uk/Tools/psa/emboss_matcher/

5 METHODS

5.1 Preparation of instrumentation and solutions

All thermostable solutions and media were autoclaved for 20 min at 121 °C and 2 bar prior to use. Glassware and consumables were autoclaved and subsequently dried at 50 °C in a compartment drier. Additionally, glassware was incubated at 200 °C for 4 h for sterilization. Heat-labile solutions were prepared in stock solutions and filtered, either via a membrane filter with a pore size of 0.2 µm by use of a vacuum pump, or by using a syringe filter with a pore size of 0.2 µm or 0.45 µm. Solutions for chromatographic systems were degassed for at least 30 min in a desiccator prior to use.

5.2 Microbiological methods

5.2.1 Experiments with *E. coli*

5.2.1.1 **Cultivation and storage of *E. coli* strains**

E. coli strains were cultivated at 37 °C by shaking at 140 rpm (1 liter cultures) or 170 rpm (5, 50, 100, 250 and 500 ml cultures), respectively. For cultivation, LB medium was used, unless otherwise stated. For plasmid-harboring strains, the medium was supplemented with the corresponding antibiotics (150 µg/ml ampicillin, 30 µg/ml chloramphenicol, 75 µg/ml kanamycin) using a filter-sterilized, 1000-fold concentrated stock solution. To obtain single colonies, the cell suspension was plated on agar plates containing the adequate antibiotics, and incubated over night at 37 °C. For temporary storage, plates and suspensions were sealed and stored at 4 °C. For long-term storage, glycerol cultures were prepared. For this purpose an aliquot of an overnight culture was mixed in a 1:1 ratio with 87 % glycerol, and stored in a sterile screw cap reaction vessel at -80 °C.

5.2.1.2 **Preparation of chemically competent *E. coli* cells**

For preparation of chemically competent *E. coli* cells according to Inoue *et al.* (1990), 500 ml SOB medium was inoculated with the respective overnight culture to an OD_{600} of 0.1, and cultured at 37 °C and 170 rpm until an OD_{600} of 0.6 was reached. The culture was incubated on ice for 15 min, transferred into 50 ml tubes, and cells were harvested by centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 10 min, 4 °C). The cell pellet was resuspended in 100 ml ice-cold TFB I buffer and centrifuged a second time under the same conditions as stated above. The resulting pellet was resuspended in 10 ml ice-cold TFB II

buffer. Immediately after resuspension, 100 µl aliquots of the cell suspension were transferred to EPPENDORF reaction vessels on ice and stored at -80 °C.

5.2.1.3 **Transformation of chemically competent *E. coli* cells**

For transformation of chemically competent *E. coli* cells, a 100 µl aliquot was thawed on ice, and about 100 ng plasmid DNA with a maximum volume of 20 µl were added. Following incubation on ice for 5 min, cells were heat-shocked for 45 s at 42 °C, and subsequently chilled for further 5 min on ice. A volume of 900 µl LB medium was added, and cells were incubated for 1 h at 37 °C in a shaker at 170 rpm to develop antibiotic resistance. Finally, dilutions of the cell suspension were plated on LB agar plates containing the appropriate antibiotics for selection.

5.2.2 Cultivation and storage of *B. subtilis* strains

B. subilis strains were cultivated at 40 °C while shaking at 140 rpm (1 liter cultures in 3 liter flasks) or 170 rpm (5 ml cultures in 50 ml plastic tubes (SARSTEDT) and 100 ml cultures in 500 ml flasks), respectively. For cultivation, LB-medium was used, unless otherwise stated. For plasmid-harboring strains, the medium was supplemented with 0.8 µg/ml erythromycin using a filter-sterilized, 1000-fold concentrated stock solution. To obtain single colonies, the cell suspension was plated on agar plates containing the adequate antibiotics, and incubated over night at 37 °C. For temporary storage, plates and suspensions were sealed and stored at 4 °C. For long-term storage, glycerol cultures were prepared. For this purpose an aliquot of an overnight culture was mixed in a 1:1 ratio with 87 % glycerol, and stored in a sterile screw cap reaction vessel, deep-well plates or microtiter plates at -80 °C.

5.3 Molecular biology methods

5.3.1 Determination of DNA concentration

The DNA concentration was determined spectrophotometrically at a wavelength of 260 nm. According to Lambert-Beer's law, an OD_{260} value of 1 (with ${}^{0.1\,\%}A_{260}$ = 20 cm^2 mg^{-1} and a pathlength of 1 cm) corresponds to a DNA concentration of 50 µg/ml double-stranded (ds) DNA (35 µg/ml RNA and 33 µg/ml ssDNA, respectively). Thus, the DNA concentration can be calculated as follows:

$$c_{dsDNA} = \frac{A_{260} \cdot 50 \cdot f}{1000}$$

Equation 1 - Determination of DNA concentration

c_{dsDNA}	concentration of double stranded DNA [µg/µl]
A_{260}	absorbance at 260 nm
f	dilution factor

A pure DNA solution should not show measurable absorption above 300 nm, and its OD_{260}/OD_{280} quotient should be at least 1.8.

5.3.2 Amplification of DNA fragments by standard polymerase chain reaction

The polymerase chain reaction (PCR; Mullis and Faloona, 1987; Saiki *et al.*, 1985) is used to amplify a specific DNA fragment *in vitro*. This is achieved by cyclic repetition of the denaturation of the dsDNA, followed by the hybridization (annealing) of primers (synthetic oligonucleotides that flank the DNA sequence of interest) and enzymatic DNA synthesis (extension). The DNA fragment is exponentially amplified. The reaction was performed in a total volume of 50 µl in a thermal cycler (lid temperature 110 °C). The standard reaction mixture contained 5-100 ng of template DNA, 2.5 U GoTaq DNA polymerase, 5 x Green GoTaq reaction buffer [contains 7.5 mM $MgCl_2$ (final concentration: 1.5 mM $MgCl_2$) and loading buffer], 0.2 mM dNTP mix, and 1 µM of each primer. For a high accuracy during the amplification, 2.5 U Pwo polymerase, which has a 3'→5' proofreading activity, was added to the reaction mixture. The standard PCR program was as follows:

step	temperature (°C)	duration
Initialization step	95	3 min
Denaturation	95	45 s
Annealing	T_A	45 s
Extension	72	1 min/kb
Final elongation	72	10 min
Final hold	16	∞

Steps 2 to 4 were repeated 30-35 times.

The optimum melting temperature T_M and annealing temperature T_A of the primers were calculated according to Equation 2 and Equation 3 (Chester and Marshak, 1993).

$$T_M = 69.3 + 0.41 \cdot (\%GC) - \frac{650}{n}$$

Equation 2 - Calculation of the melting temperature of an oligonucleotids

T_M	melting temperature of primer (°C)
% GC	GC content of primer
n	number of nucleotides in primers

$$T_A = \left(\frac{T_{M1} + T_{M2}}{2}\right) - 3\,°C$$

Equation 3 - Calculation of the optimum annealing temperature for a pair of primers

T_A	annealing temperature (°C)
T_{M1}, T_{M2}	melting temperatures of primers 1 and 2 (°C)

The optimum annealing temperature was also experimentally determined using a thermal cycler with gradient function (EPPENDORF Mastercycler gradient). To this end, different PCRs with annealing temperatures between 50 °C and 70 °C were set up in parallel, and the yields of the amplification products were determined.

5.3.3 Colony PCR

To verify the success of cloning, an insert screening was performed by colony PCR. Transformants grown on selective LB agar were added to the PCR mixture by picking a single colony with a pipette tip and transferring small quantities of cells to the reaction mixture. The cells were subsequently disrupted in the initial denaturation step at 95 °C, and the released DNA was used as template in the following amplification cycles. Vector-specific amplification primers were used (4.8.1). The standard reaction mixture included 1 U GoTaq DNA polymerase, 5 x Green GoTaq reaction buffer, 0.2 mM dNTP-mix, 1 µM of each primer, and water to a final volume of 20 µl. The amplification was performed according to the standard PCR protocol described in chapter 5.3.2. As *B. subtilis* are gram positive and therefore possess a more stable cellwall than *E. coli*, transformants grown on selective LB agar were picked, resuspended in 100 µl water and incubated at 95 °C for 15 min. After centrifugation (EPPENDORF Centrifuge 5415R, 13,200 rpm, 1 min, RT) 1 µl of supernatant was used as template for the PCR.

5.3.4 QuikChange site-directed mutagenesis (QCM)

The QuikChange site-directed mutagenesis (QCM) method allows for the efficient introduction of point mutations, insertions and deletions in any type of dsDNA plasmid. It is performed using Pfu DNA polymerase, which replicates both plasmid strains with high

fidelity due to its 3'→5' proofreading activity. The technique originally developed by STRATAGENE (La Jolla, USA) was modified according to a protocol published by Wang and Malcolm (1999) using a two-stage mutagenesis protocol. For introduction of point mutations, complementary mutagenic primers with up to 35 bases in length and 12-15 bases of template complementary sequence on both sides of the mismatch were used. Pfu DNA polymerase extends and incorporates mutagenic primers during temperature cycling and generates a mutated plasmid containing staggered nicks. To avoid unproductive primer dimer formation, two separate primer extension reactions were performed initially. In the second step both primer extension reactions are combined and linearly amplified. Methylated DNA template is digested by treatment of the product with *Dpn*I. *E. coli* cells are transformed with the nicked vector DNA carrying the desired mutations, and the nicks are sealed by the DNA repair apparatus of the cells. See Figure 11 for illustration.

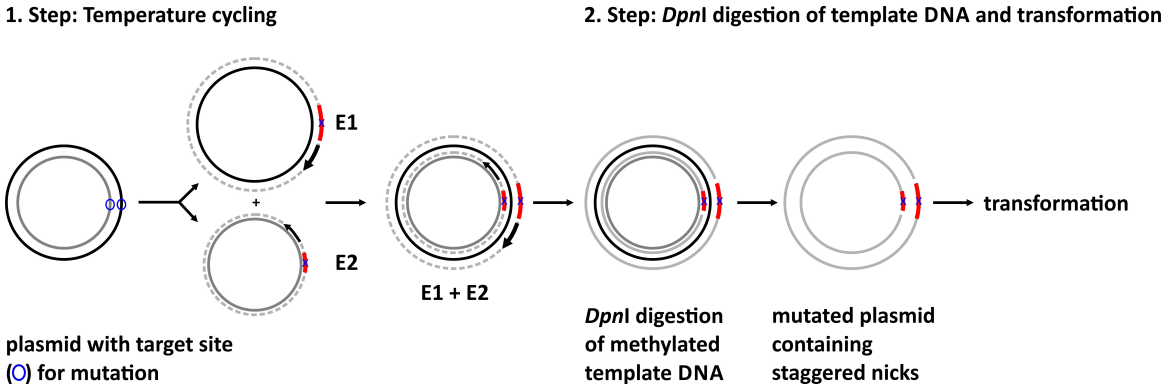


Figure 11 – Overview of the QuikChange site-directed mutagenesis method
In the first step the mutated plasmids containing staggered nicks are generated in two parallel primer extension reactions (E1 and E2). The reactions E1 and E2 are combined and amplified. In the second step, the DNA template is digested by treatment with *Dpn*I, and *E. coli* cells are transformed with the nicked vector DNA containing the desired mutations. The figure was modified according to the QuikChange Site-Directed Mutagenesis Kit manual (STRATAGENE).

The two separate primer extension reactions (E1 and E2) were performed each in a total volume of 50 µl containing 50 ng of template plasmid comprising the gene of interest, 2.5 U Pfu DNA polymerase, 10 x Pfu DNA polymerase reaction buffer [contains 20 mM $MgCl_2$ (final concentration: 2 mM $MgCl_2$)], 0.2 mM dNTP mix, and 1.6 µM of either the 5' or 3' primer. The standard primer extension program using Pfu DNA polymerase was as follows:

step	temperature (°C)	duration
Initialization step	95	3 min
Denaturation	95	45 s
Annealing	T_A	1 min
Extension	72	2 min/kb
Final elongation	72	10 min
Final hold	16	∞

Steps 2 to 4 were repeated four times.

Subsequently, 25 µl of each primer extension reaction (E1 and E2) were mixed, supplemented with additional 2.5 U Pfu DNA polymerase and subjected to 18 cycles of QCM as described above. The methylated template DNA was digested by directly adding 20 U DpnI to 25 µl of the QCM reaction mix and incubated at 37 °C for 1.5 h. Then, chemically competent *E. coli* Turbo cells were transformed using 20 µl of *Dpn*I-digested QCM mix.

5.3.5 Isolation and purification of plasmid DNA from *E. coli*

Purification of plasmid DNA was performed according to the principle of alkaline cell lysis (Le Gouill *et al.*, 1994). Bacterial cell cultures were lysed by adding SDS and sodium hydroxide. The suspension was subsequently neutralized by addition of ammonium acetate; proteins and genomic DNA were precipitated, whereas circular plasmid DNA remained in solution.

For analytical isolation of plasmid DNA from *E. coli* the mini preparation kit from FERMENTAS (GeneJET™ Plasmid Miniprep Kit) was used. For this purpose, 5 ml of overnight cell cultures were harvested by centrifugation (EPPENDORF 5415R, 13200 rpm, 1 min, RT). The isolation of plasmid DNA was performed according to the protocol supplied by the manufacturer. Bound plasmid DNA was eluted from the silica column with 50 µl sterile water. The recovered plasmid DNA was stored at -20 °C.

5.3.6 Agarose gel electrophoresis

DNA fragments are separated by agarose gel electrophorese according to their length. The bands in the gel become visible under UV-light by adding ethidium bromide, a DNA intercalating fluorescent dye (Sharp *et al.*, 1973). For preparation of agarose gels, 1 % (w/v) agarose was dissolved in 0.5 % TBE buffer by boiling in the microwave. Following

cooling down to 50-60 °C, 0.2 µl of an ethidium bromide stock solution (10 mg/ml) was added per ml agarose. The solution was then cast into a gel chamber, and a comb was inserted. Following solidification, the gel was covered with 0.5 % TBE buffer, and the comb was removed. The DNA samples were supplemented with DNA loading dye and pipetted into the gel pockets. The electrophoresis was performed using a voltage of 190 V for about 20 min. The negatively charged DNA migrates to the anode, whereby DNA fragments are retarded to a different extent by the agarose matrix due to differences in size. The fragments were detected under UV-light (λ = 302 nm) and documented using the Imager Multi-Doc-It Digital Imaging system. To estimate the size of the fragments, a volume of 5 µl of the GeneRuler™ 1 kb Plus DNA Ladder (FERMENTAS) was applied to a separate pocket of the gel.

5.3.7 Isolation of DNA fragments from agarose gels

The desired fragments were excised from the gel under UV-light (λ = 302 nm) with a scalpel and transferred to a reaction vessel. DNA was extracted using the GeneJET™ Gel Extraction Kit (FERMENTAS) according to the protocol supplied by the manufacturer. The isolated DNA was eluted with 30 µl sterile water and stored at -20 °C.

5.3.8 Enzymatic manipulation of dsDNA

5.3.8.1 **Cleavage of dsDNA by restriction endonucleases**

Type II endonucleases, which bind to a palindromic recognition sequence (recognition site) were applied for specific cleavage of dsDNA (Sambrook, 1989; Wilson and Murray, 1991). These restriction enzymes generate single stranded overhangs (sticky ends), which contain 3'-hydroxyl and 5'-phosphate ends. For analytical cleavage, a maximum of 1 µg DNA was incubated at 37 °C in the appropriate buffer with 20 U of each restriction enzyme for about 2 h in a volume of 50 µl. For preparative cleavage, with the objective of subsequent ligation, 2 µg or the entire amount of PCR product and 2 µg of vector DNA were digested with 20 U of each restriction enzyme in a volume of 50 µL at 37 °C for 2 h. The volume of added restriction endonuclease in the reaction mixture should not exceed 10 % of the total volume, as the activity of enzymes is influenced by glycerol, a component of the enzyme storage solution. The fragments were purified by agarose gel electrophoresis (chapter 5.3.6) for subsequent ligation (chapter 5.3.8.2).

5.3.8.2 **Ligation of DNA fragments**

For ligation, digested vector and insert were mixed at an estimated molar ratio of 1:3 and ligated in a total volume of 20 µl with 1 U T4-DNA ligase (FERMENTAS) in the buffer supplied by the manufacturer, either overnight in a thermal cycler at 16 °C or for 1 h at RT. Subsequently, competent *E. coli* cells were chemically transformed (chapter 5.2.1.3) with the ligation mixture.

5.3.9 DNA sequencing

All constructs generated and used in this work were sequenced to validate their integrity. The determination of nucleotide sequences was performed either by the company Geneart (part of Life Technologies, Regensburg) or by the company Seqlab (Göttingen). Samples for sequencing orders at Geneart contained about 100-300 ng DNA and 1.25 µM of sequencing primer in a total volume of 8 µl. Samples for sequencing orders at Seqlab contained about 600 ng DNA and 2 µM of sequencing primer in a total volume of 15 µl. The generated *.ABI files were analyzed by the program CLC main workbench.

5.4 Protein biochemistry methods

5.4.1 Gene expression

5.4.1.1 **Gene expression at analytical scale**

To test whether cloned genes are overexpressed and to analyze whether the recombinant proteins were produced in soluble form, expression was performed at analytical scale. To this end, 50 ml LB medium were supplemented with the appropriate antibiotics, inoculated with an overnight culture of a single colony to an OD_{600} of 0.1, and incubated at 37 °C and shaking (170 rpm). At an OD_{600} of 0.4-0.6, the cultures were induced by adding IPTG to a final concentration of 0.1 mM, 0.5 mM or 2 mM as stated in the text. Following incubation at 37 °C or 30 °C for 3 h to overnight, the cells were centrifuged (EPPENDORF Centrifuge 5415R, 4000 rpm, 10 min, 4 °C). The pellet was resuspended in 750 µl of 50 mM potassium phosphate pH 7.5. The cells were disrupted by sonication (HEINEMANN branson sonifier W-250D, amplitude 20 %, 2 min, 2 s pulse, 2 s pause, 4 °C) or the cells were incubated for 1 h on ice while vortexing all 10 min instead of sonication. Following centrifugation (EPPENDORF Centrifuge 5415R, 13,200 rpm, 5 min, 4 °C) the supernatant (containing soluble protein) was taken off, and the pellet (containing

insoluble protein) was resuspended in 750 µl of the resuspension buffer. Aliquots of the supernatant was taken off and subjected to HisSpin Trap columns (GE HELATHCARE, 5.4.2.1) for further purification. The eluate (E), supernatant (S) and the resuspended pellet (P) were supplemented with SDS sample buffer (5x), and incubated for 5 min at 95 °C. Subsequently, a volume of 5-10 µl of S and P was analyzed by SDS-PAGE (chapter 5.6.3). The expression conditions were optimized if necessary by using different *E. coli* host strains, by varying the resuspension buffers, IPTG concentrations or by utilizing Thioredoxin (TrxA) as fusion tag.

5.4.1.2 **Gene expression at preparative scale**

For purification of proteins at preparative scale, 1-14 x 1 liter LB medium (in standard flasks in case of *yvoF* expression and in baffled flasks for all other gene expressions) was supplemented with the corresponding antibiotics. The medium was inoculated with a freshly prepared overnight *E. coli* culture to an OD_{600} value of 0.1. The suspension was incubated at 37 °C and 150 rpm. At an OD_{600} of 0.4-0.6 the suspension was induced with 0.5 mM IPTG and cells were grown at 37 °C for 3 h to overnight. The cells were harvested by centrifugation (BECKMANN COULTER Avanti J-26SXP, JLA-8.1000, 4000 rpm, 20 min, 4 °C), and the pellet was resuspended in resuspension buffer (chapter 4.10.3). Cells were disrupted by sonication on ice (HEINEMANN branson sonifier W-250D, amplitude 40 %, 2 min, 2 s pulse, 2 s pause, 4 °C). In case of YvoF instead of sonication the resuspended cells were incubated for 1 h on ice while vortexing all 10 min. The cell debris was removed by centrifugation (BECKMAN COULTER Avanti J-26SXP, JA-25.50, 14,000 rpm, 30 min, 4 °C) and the supernatant was further processed.

5.4.2 Protein purification

Ammonium sulfate precipitation, ion exchange chromatography and analytical size exclusion chromatography were used for homologous PhoB enrichment in my master thesis (Linde, 2013). For all heterologous expressed proteins metal affinity chromatography resulted in pure protein samples. Subsequently, the buffer was exchanged by dialysis. The proteins were concentrated if necessary by ultrafiltration (MILLIPORE, Amicon 10 kDa cut-off) and stored at -80 °C.

5.4.2.1 **Immobilized metal ion affinity chromatography (IMAC) at analytical scale**

Metal ion affinity chromatography is based on specific, reversible interactions between proteins and immobilized metal ions. The columns consist of sepharose and a covalently linked iminodiacetic acid (IDA) type chelator. IDA coordinates metal ions such as Ni^{2+}, Cu^{2+}, Fe^{3+}, and Co^{2+}. Nitrogen and sulfur atoms can bind to the free binding sites at the metal ions.

In the framework of this thesis, bsPhoB, baYvoF and bsYvoF with a C-terminal hexahistidine tag ($(His)_6$ tag) were purified at analytical scale on His SpinTrap™ columns (GE HEALTHCARE). Bound proteins were eluted with 500 mM of imidazole, which competes with the histidines for binding to the immobilized Ni^{2+} ions. The purification was performed as follows (Table 5):

Table 5 - Protocol for protein purification with HisTrap column at analytical scale

equilibration	600µl binding puffer[1]
sample application	600 µl steppwise of protein solution in resuspension buffer[1]
wash	600µl binding puffer[1]
elution	2x 200 µl elution buffer[2]

[1]Resuspension buffer and binding buffer were identical

[2]The elution buffer was identical with the binding buffer except that it contained 500 mM imidazole.

After each step, the column was centrifuged for 30 s at 100 g (EPPENDORF Centrifuge 5417R, 4 °C). The eluate was further analyzed by SDS-PAGE (5.6.3), Western Blot (5.6.4) and activity assays.

5.4.2.2 **Immobilized metal ion affinity chromatography (IMAC) at preparative scale**

IMAC was developed by Porath *et al.* (1975). In this thesis, proteins with a C-terminal hexahistidine tag ($(His)_6$ tag) were purified on a HisTrap FF crude Ni^{2+}-IDA column (GE HEALTHCARE, CV: 5 ml, pressure limit: 0.5 MPa). Bound proteins were eluted by increasing concentrations of imidazole, which competes with the histidines for binding to the immobilized Ni^{2+} ions. Protein samples were sterile filtered (0.45 µm) and applied on the HisTrap FF crude column in an Äkta chromatographic system.

The purification with the HisTrap FF was performed using the chromatographic system ÄKTA Purifier 10 (GE HEALTHCARE) as follows (Table 6).

Table 6 - Protocol for protein purification with HisTrap column at preparative scale

flow rate	5 ml/min
equilibration	5 column volumes (CV) binding puffer[1]
sample application	40-200 ml protein solution in resuspension buffer[1]
wash	10 CV
elution	15 CV gradient from 0-75% elution buffer[2]; 2 ml fractions were collected
final wash-out	5 CV elution buffer
reequilibration	5 CV binding buffer
purging and storage	5 CV H_2O, 5 CV 20 % ethanol

[1]Resuspension buffer and binding buffer were identical

[2]The elution buffer was identical with the binding buffer except that it contained 500 mM imidazole.

The elution of protein was monitored by following the absorbance at 260 and 280 nm. The eluted fractions were analyzed by SDS-PAGE (5.6.3).

5.4.2.3 **Ammonium sulfate precipitation**

Ammonium sulfate precipitation (ASP) is an efficient method used for initial sample concentration and cleanup. ASP was used in my master thesis (Linde, 2013) to enrich the phosphatase bsPhoB from *B. subtilis* wild type cells.

A 50 ml overnight culture of *B. subtilis* was harvested by centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 15 min, 4 °C), the pellet was resuspended in 1/10 of culture volume 1x PBS (4.10.5) and the cells were digested by lysozyme (300 µg/ml) for 1 h at 37 °C. Following centrifugation (EPPENDORF Centrifuge 5415R, 13,200 rpm, 15 min, 4 °C) the supernatant was sonicated (HEINEMANN branson sonifier W-250D, amplitude 20 %, 2 min, 2 s pulse, 2 s pause, 4 °C). Subsequently, $(NH_4)_2SO_4$ was added to a final concentration of 60 % saturated solution, portion-wise as crystalline salt using a spatula, while stirring vigorously on ice. The amount of $(NH_4)_2SO_4$ was calculated using an online calculator by EnCor Biotechnology Inc (Gainesville, USA, http://www.encorbio.com/protocols/AM-SO4.htm) The suspension was stirred for further 30 min at 4 °C and subsequently centrifuged (SORVALL, SS34-rotor, 14,000 rpm, 15 min, 4 °C). The remaining soluble proteins containing the phosphatase activity were dialyzed against 50 mM

Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$ and 1 mM DTT followed by the second purification step, which was cation exchange chromatography (5.4.2.4).

5.4.2.4 **Ion exchange chromatography**

Ion exchange chromatography is based on the competition of charged molecules (proteins and salt ions) for interaction with immobilized ion exchange groups of opposite charge. In the first stage, charged molecules adsorb reversibly to the immobilized support material. Subsequently, bound molecules are eluted by a gradient of steadily increasing ionic strength, or alternatively a pH-gradient. The charge of the protein is mainly dependent on the amino acids with charged side chains. In the acidic or neutral pH-range the amino groups, mainly those of Lys, Arg, and His, are protonated and the protein exposes cationic behavior, whereas in a neutral or basic pH-range, the carboxyl groups of Asp and Glu residues are negatively charged and consequently the protein is anionic. The total charge of proteins is mainly dependent on the amino acid composition and the pH-value of the buffer. Ion exchange chromatography can be subdivided into cation exchange chromatography, in which positively charged ions bind to a negatively charged resin, and anion exchange chromatography, where negatively charged ions bind to a positively charged resin.

PhoB was purified using a cation exchange column (Resource S, GE HEALTHCARE, 1 ml). The protein sample was sterile filtered (0.45 µm) and applied on the column using the chromatographic system ÄKTA Purifier 10 (GE HEALTHCARE). A linear gradient of NaCl (0-3 M) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$ and 1 mM DTT was applied to elute the protein. Interfering salt was removed from the purified protein by dialysis against 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$ and 1 mM DTT. After pooling the active fractions, size exclusion chromatography was carried out as last purification step (5.6.5).

5.4.3 Buffer exchange by dialysis

To exchange the buffer of a protein solution or to remove salt, a dialysis was performed two times for at leat 4 hours against a 100-fold volume of buffer at 4 °C. A dialysis tubing (Visking) with a molecular cut-off of 14 kDa was used, which retains the protein while low molecular substances can pass through the membrane. All proteins were dialyzed against the buffer for dialysis as stated in 4.10.3.

5.4.4 Concentrating protein solutions

Protein solutions were concentrated using Amicon Ultra-15 centrifugal filter devices (MILLIPORE; molecular cut-off: 10 kDa; for complexes 30 kDa) by centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 4° C) according to the instructions of the manufacturer.

5.4.5 Storage of purified proteins

Purified and concentrated proteins were frozen in liquid nitrogen either dropwise or as 100 µl aliquots in PCR reaction vessels and stored at -80° C.

5.5 Preparative synthesis of glycerol 1-phosphate

5.5.1 Synthesis of glycerol 1–phosphate from glucose

As described earlier (Guldan, 2010), ^{14}C-labeled glycerol 1-phosphate (G1P) was synthesized *in vitro* from ^{14}C-labeled glucose (HARTMANN ANALYTIC) by using enzymes of the glykolysis (Figure 12). For the stereospecific reduction of ^{14}C-dihydroxyacetone phosphate (^{14}C-DHAP) to ^{14}C-G1P a glycerol 1-phosphate dehydrogenase (AraM) was used (Guldan *et al.*, 2008).

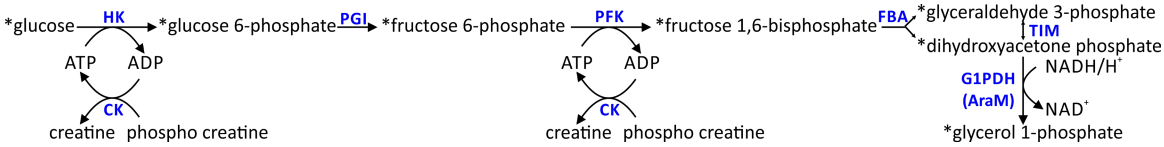


Figure 12 - Synthesis of ^{14}C-G1P from ^{14}C-glucose via the first steps of the glycolysis and AraM
HK: hexokinase, PGI: phosphogluco isomerase, PFK: phosphofructokinase, FBA: fructose 1,6-bisphosphate aldolase, TIM: triosephosphate isomerase, G1PDH: glycerol 1-phosphate dehydrogenase, CK: creatine kinase, ATP: adenosine triphosphate, ADP: adenosine diphosphate, NADH/H$^+$: nicotinamide adenine dinucleotide (reduced form), NAD$^+$: nicotinamide adenine dinucleotide (oxidized), *: ^{14}C-labeled.

To obtain maximal yields, the enzymes have been applied in a ratio for optimal efficiency (Fiévet *et al.*, 2006; Table 7). To prevent inhibition of PFK by increasing ATP concentrations, an ATP regeneration system was used consisting of phospho creatine and creatine kinase.

Table 7 - Concentration of enzymes for G1P synthesis

enzyme	concentration [µM]
hexokinase	0.1
phosphogluco isomerase	0.26
phosphofructokinase	0.72
fructos 1,6-bisphosphate aldolase	1.53
triosephosphate isomerase	0.06
creatine kinase	1
glycerole 1-phosphate dehydrogenase (AraM)	17.2

The used buffer conditions and substrate concentrations have been as follows (Table 8).

Table 8 – Substrate- and bufferconcentrations for G1P synthesis

substrate	concentration [mM]	buffer	concentration [mM]
^{14}C-glucose	0.214	**PIPES, pH 7.5**	50
NADH	3	**KCl**	100
ATP	1	**MgAc**	5
creatine phosphate	20		

The reaction volume was 200 µl and incubation was followed at RT. As AraM exhibits low stability in PIPES buffer this enzyme was added twice after a incubation time of 1 h at 40 °C. The reaction was followed over night and stopped the next day by applying to preparative paper chromatography.

5.5.2 Preparative paper chromatography

The separation of ^{14}C-labeled G1P from the other assay components was done by paper chromatography (Benson *et al.*, 1950). Paper chromatography is a form of partition chromatography (Consden *et al.*, 1944). The supporting material is a homogenous filter paper (WHATMAN No. 1) and the stationary phase is built of water, which is stored in the intercellular space of the cellulose fibres (content of about 30 %). The used mobile phase is a mix of isopropyl alcohol/ammonia/water, 6:3:1 (v/v/v). Compounds comigrate with the mobile phase in dependency of their solubility in the mobile phase. The chromatogram was developed descending (Guldan, 2010). 400 µl of reaction solution

from G1P synthesis was applied to the starting front of a paper with dimensions of 16x60 cm. The paper was folded over a clamp and hanged in a tray of the chromatography chamber containing the solvent of the mobile phase. At the bottom of the chamber a vessel with about 10 ml of solvent was located to ensure saturation of the air in the chamber with the solvent. The chromatography was performed over night, dried afterwards on air and the ^{14}C-labeled G1P was detected via autoradiography with the Cyclone phosphoimager system (PACKARD BioScience). The detected segment was cut out and eluted with water. Therefore the paper was put in a 50 ml plastic tube (SARSTEDT) and incubated with 3 ml water for 1 h while rotating permanently at RT. Following removal of the residual water this process was repeated twice. The eluated fractions were pooled and the concentration of ^{14}C-labeled G1P was determined through measuring of the activity in the scintillation counter (PACKARD BioScience Tri-Carb 2900RT).

5.5.3 Qualitative verification of glycerol 1-phosphate synthesis

To verify successful G1P synthesis, an enzymatic assay was performed at 40 °C for 2 h in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$ and 0.2 % Tween80. 1 µM GGGPS from *Archaeoglobus fulgidus* (afGGGPS), 50 µM GGPP (SIGMA-ALDRICH) and 0.15 µCi ^{14}C-labeled G1P were combined in a volume of 200 µl. After 1 h the reaction sample was splitted into 2x 100 µl. To one sample, 10 U calf intestinal alkaline phosphatase (NEW ENGLAND BIOLABS) was added and the reaction was followed further for 1 h at 40 °C. Products were extracted according to the method of lipid extraction (5.6.7), analyzed by thin layer chromatography (5.6.8) and visualized with a phosphoimager system (5.6.9).

5.6 Analytical methods

5.6.1 Protein concentration determination via absorption spectroscopy

The aromatic amino acids Trp, Tyr and Phe as well as disulfide bonds (cystine) absorb UV-light in a wavelength interval between 250 and 300 nm. The molar extinction coefficient at 280 nm (ε_{280}) can be determined from the amino acid composition (Pace *et al.*, 1995) (Equation 4).

$$\boldsymbol{\varepsilon_{280} = \sum Trp \cdot 5500 + \sum Tyr \cdot 1490 + \sum Cystine \cdot 125}$$

Equation 4 - Determination of the molar extinction coefficient ε_{280}

ε_{280} molar extinction coefficient at 280 nm [$M^{-1}cm^{-1}$]

The specific extinction coefficient ($^{0.1\,\%}A_{280}$) can be calculated according to Equation 5 taking the molecular weight into account.

$$^{0.1\%}\boldsymbol{A_{280}} = \frac{\boldsymbol{\varepsilon_{280}}}{\boldsymbol{MW}}$$

Equation 5 - Determination of the specific extinction coefficient $^{0.1\%}A_{280}$

$^{0.1\,\%}A_{280}$	specific extinction coefficient at 280 nm [cm^2/mg]
MW	molecular weight [g/mol]

Using Lambert-Beer's law, the protein concentration can be determined by measuring the absorbance at 280 nm (Equation 6).

$$\boldsymbol{A_{280}} = {}^{0.1\%}\boldsymbol{A_{280}} \cdot \boldsymbol{c} \cdot \boldsymbol{d}$$

$$\boldsymbol{c} = \frac{\boldsymbol{A_{280}}}{^{0.1\%}\boldsymbol{A_{280}} \cdot \boldsymbol{d}}$$

Equation 6 - Determination of the protein concentration by using the specific extinction coefficient $^{0.1\,\%}A_{280}$

A_{280}	absorbance at 280 nm
c	concentration [mg/ml]
d	pathlength [cm]
$^{0.1\,\%}A_{280}$	specific extinction coefficient at 280 nm [cm^2/mg]

Absorbance spectra were recorded between 220 and 350 nm. The absorbance maximum should be at 280 nm and the A_{280}/A_{260} ratio should be at least 1.8 for a pure protein solution. No absorbance above 300 nm should be detectable to exclude any distortion of the results caused by light scattering resulting from aggregation.

5.6.2 Bradford assay

The Bradford assay is also used for the quantification of proteins in aqueous solutions. The assay contains Coomassie Brilliant Blue G250, which predominantely binds to cationic, nonpolar and hydrophobic regions of the protein. The absorption maximum is shifted from 465 nm to 595 nm when the dye binds to the protein. The change in absorbance at 595 nm is linear in a protein range of 1.2 to 10.0 µg/ml. For determination of protein concentrations, 200 µl of Bradford reagent (BIORAD) were added to 800 µl of diluted protein solution. After 5 min incubation in the dark at RT, the absorbance was measured at 595 nm using diluted Bradford reagent as blank. The protein concentration was calculated by a linear regression standard curve of BSA according to Equation 7.

$$c = \frac{A_{595}}{m}$$

Equation 7 - Determination of protein concentration by the Bradford assay

A_{595}	absorbance at 595 nm
c	concentration [mg/ml]
m	slope of standard curve

5.6.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins are denatured by the detergent sodium dodecyl sulfate (SDS) and negatively charged proportional to their molecular weights. SDS binds to the protein in a ratio of approximately 1.4 g SDS per 1 g protein. The net charge of proteins can be neglected compared to the negative charge of the protein complexed with SDS, which yields an approximately uniform mass to charge ratio. As a consequence, electrophoretic mobility depends only on the sieve effect of the gel: the migration speed is inversely proportional to the logarithm of mass (Laemmli, 1970). Table 9 shows the composition of the 12.5 % SDS gels used in this thesis.

Table 9 - Composition of 12.5 % SDS-PAGE gel

Amount specification applies to 13 gels.

	resolving gel (12.5 %)	stacking gel (6 %)
resolving/sacking gel buffer	19.5 ml	7.38 ml
acrylamide-SL (30%)	26.2 ml	5.9 ml
H_2O	31.58 ml	15.95 ml
TEMED	0.089 ml	0.029 ml
APS (10%)	0.195 ml	0.089 ml

Samples were supplemented 1:4 with 5 x SDS-PAGE sample buffer and incubated for 5 min at 95 °C. Gel pockets were loaded with 5-20 µl of sample, and gels were run at 50 mA and 300 V for about 30 min. Subsequently, gels were stained with SDS-PAGE staining solution, whereby the detection limit of the Coomassie Brilliant Blue dye G-250 is between 200-500 ng protein/mm^2. Gels were swayed for 10 min in the staining solution and excess dye was removed by repeatedly boiling in water (microwave 900 W).

5.6.4 Semi dry western blot

Western blots are used for the specific identification and quantification of single proteins in complex protein mixtures as for example crude extracts. The detection of the protein

of interest is done with an anti-$(His)_6$ tag antibody (Towbin *et al.*, 1979). Following separation of proteins on SDS-PAGE (5.6.3) the gel is blotted on a polyvinylidene difluoride (PVDF) membrane, which enables unspecific binding of proteins via hydrophobic interactions with the polymer. In this thesis a Immobilon-P membrane from MILLIPORE with a pore size of 0.45 µm was used. The membrane was activated by shaking in 100 % methanol for 1 min. Afterwards membrane, gel and filter papers were incubated for 15 min in transfer buffer (0). As the blotting direction in the used blotting apparatus was downward, first three filter papers, then the PVDF membrane, the gel and on top again three filter papers were arranged. The transfer was carried out for 1 h at RT with 45 mA per gel. Following incubation in blocking buffer for 1 h at RT or over night at 4 °C to prevent unspecific binding, the Anti-$(His)_6$ tag antibody (1:500, ROCHE DIAGNOSTICS) was added for further 30 min incubation. Afterwards, the membrane was washed three times in wash solution. The detection of bound antibody was performed with the SuperSignal West Pico Kit (PIERCE) according to the protocol supplied by the manufacturer.

5.6.5 Analytical size exclusion chromatography (SEC)

Apparent molecular weights (MW_{app}) of proteins were determined by analytical size exclusion chromatography. Proteins are separated according to their size. The retention volume decreases approximately linear with the logarithm of the molecular weight. Globular calibration proteins were used for a standard curve in order to determine the apparent molecular weights. In this work, a Superdex 75 10/300 GL (GE HEALTHCARE, 23.56 ml) or a Superdex 200 10/300 GL (GE HEALTHCARE, 23.56 ml) was used with an Äkta Basic 10 better (GE HEALTHCARE) chromatographic device. Protein samples were centrifuged (EPPENDORF 5415R, 13,200 rpm, 30 min, 4° C) and the supernatant was applied to the pre-equilibrated column, which was run in various buffers depending on the protein to be analyzed (4.10.3). The elution was followed by absorbance measurement at 280 nm.

5.6.6 Membrane fractionation

The localization of YvoF in *B. subtilis* wild type was determined by ultracentrifugation. A 50 ml overnight culture of *B. subtilis* was harvested by centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 20 min, 4 °C), the pellet was resuspended in 1/10 culture volume of 1x PBS (4.10.5) and the cells were digested by lysozyme (300 µg/ml) for 1 h at

37 °C. Following centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 20 min, 4 °C), the pellet was dissolved in 500 µl 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 9 mM CHAPS. 100 µl of the supernatant were taken off for further analysis and the residual supernatant was subjected to ultracentrifugation for 1 h at 4 °C and 100,000 g in a fixed angle rotor (BECKMAN COULTER Optima^{TM} MAX-XP Centrifuge, MLA-130 fixed angle) for membrane fractionation. The pellet after ultracentrifugation was diluted in 500 µl 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 9 mM CHAPS. Supernatant and pellet of the crude extract as well as supernatant and pellet after ultracentrifugation were analyzed for activity by a radiometric *in vitro* activity assay for acetyltransferase activity (5.6.14.4).

5.6.7 Extraction of lipids

Extraction of lipids from cells or of hydrophobic products from reaction samples was done according the method of Bligh and Dyer (1959) modified by Kates (1986). Methanol and afterwards chloroform was added to the aqueous solution (lysed cells or reaction samples) to a ratio methanol/chloroform/water of 2:1:1 (v/v/v) and the sample was vortexed. The resulting mixture was monophasic. Following addition of chloroform and water to a ratio of 2:2:2, a separation of organic and aqueous phase could be observed. For complete phase separation, centrifugation was performed for 5 min at RT (HETTICH Centrifuge MIKRO 200R, 11,000 rpm). The organic phase was taken off, evaporated to dryness (EPPENDORF Concentrator plus) and the remaining sample was diluted in a smaller volume of chloroform.

5.6.8 Thin layer chromatography (TLC)

Thin layer chromatography is a form of adsorption chromatography and a convenient technique for the separation of multiple samples in parallel (Stahl, 1967). The stationary phase consists of adsorbent material, which is immobilized on a glas or an aluminium plate. Samples are spotted on the plate, which is subsequently placed in a chamber with liquid mobile phase. Initially, the spots are not in direct contact with the mobile phase. During the chromatographic process, the mobile phase is drawn up the plate by capillary force. Compounds comigrate with the mobile phase in dependency of their interaction with the stationary phase and in dependency of their solubility in the mobile phase.Here, Silica60 plates (MERCK, Darmstadt) were used as stationary phase and ethyl acetate/hexane at a ratio of 1:1 (v/v) as mobile phase. The size of the Silica60 plates varied

with the number of samples. The separation was terminated before the solvent front reaches the top of the stationary phase. Visualization of the separated spots was done with a phosphoimager system (5.6.9).

5.6.9 Phosphorimaging

Radiometricly labeled analytes on a TLC plate can be visualized by their ability to excite phosphor material. The stored energy is released as blue-light luminescence after excitation with visible light. In this work, a phosphorimager Film FLA-3000 (FUJIFILM) was exposed over night and the luminescence was read in a Cyclone phosphorimager system (PACKARD BioScience).

5.6.10 *In vivo* radiolabeling experiments

^{14}C-labeled G1P was synthesized as described in chapter 5.5.1. To test *B. subtilis* wild type and knockout strains from NBRP (4.7.2) for a loss of function, the cells were grown overnight in 5 ml of LB medium containing 1 µCi of ^{14}C-labeled G1P (1.34 µM). To stabilize the knockout strains, erythromycin (0.8 µg/ml) was added for their cultivation. Cells were grown over night at 40 °C and harvested by centrifugation (HETTICH Centrifuge UNIVERSAL, 4000 rpm, 15 min, RT). The pellet was resuspended in 1/10 culture volume of 1x PBS (chapter 4.10.5) and the cells were digested by lysozyme (300 µg/ml) for 1 h at 37 °C. Following centrifugation (HETTICH Centrifuge MIKRO 200R, 13.200 rpm, 15 min, 4 °C) the supernatant was used for further lipid extraction (5.6.7), analysis by thin layer chromatography (5.6.8) and visualization with a phosphoimager system (5.6.9).

5.6.11 Differential scanning calorimetry (DSC)

Differential scanning calorimetry is a classical thermally induced unfolding technique to determine the thermal stability of proteins in solution, based on monitoring heat effects that accompany conformational changes during protein unfolding (Freire, 1995). The calorimeter contains the sample cell for a solution of protein in buffer and the reference cell for a solution of buffer. The temperature in both cells is equally increased during the measurement. Protein conformation changes and protein unfolding result in a temperature difference between the sample cell and the reference cell, which is compensated by electric heating power. The melting temperature T_M of a protein is deduced from the changes in heat capacity within a temperature range.

Dependent on the aggregation tendency of the sample, 20 µM protein (subunit concentration) was heated in degassed 50 mM potassium phosphate pH 7.5 in an appropriate temperature range (30 °C - 130 °C, 30 °C - 70 °C or 20 °C - 70 °C) at a ramp rate of 1°C min^{-1} in a VP-DSC differential scanning microcalorimeter (MicroCal, MALVERN) with fixed reference and sample cell (0.51141 ml each). The change in heat capacity with raising temperature was followed. The proper equilibration of the calorimeter was ascertained by performing several buffer-buffer baselines. Overpressure was applied to prevent boiling above 100 °C. DSC transitions were fitted with a non two state model using the MicroCal Analysis Launcher – Origin 7 software as supplied by the manufacturer. The apparent midpoint temperature (T_{Mapp}) of the irreversible unfolding transition was determined as an operational measure of protein stability. Mean values of duplicates were calculated.

5.6.12 Circular dichroism spectroscopy (CD)

Circular dichroism (CD) spectroscopy in the far-UV range is a method for quanitative analysis of protein secondary structure and its unfolding. Right- and left-circular polarized light are absorbed to a different extent when passing through a solution of chiral molecules. As a result, the circular dichroism signal over the corresponding wavelengths can adopt positive and negative values. Different secondary structure elements like α-helix, β-sheet and random coil have characteristic CD spectra in the range of 170-250 nm. In this work, far-UV circular dichroism spectra of all proteins were recorded in a JASCO J-815 spectrometer. The spectra were measured with 10 µM protein (3.5 µM in case of bsYvoF) in degassed 50 mM potassium phosphate, pH 7.5 (in the storage buffer after dialysis in case of bsYvoF), in a 0.1 cm cuvette at a scan rate of 50 nm min^{-1} with a response time of 0.5 sec from 185 nm to 260 nm (195 nm to 260 nm in case of background noise). Spectra were recorded at 25 °C before and at 95 °C after thermal denaturation of the protein sample. Each spectrum was the average of three accumulations. Each protein was meausured in duplicates. Data was normalized to obtain the molar ellipticity per amino acid according to Equation 8 (Schmid, 1997):

$$\boldsymbol{\theta_{MRW}} = \frac{\boldsymbol{\theta_{obs} \cdot 100 \cdot MRW}}{\boldsymbol{\beta \cdot d}} = \frac{\boldsymbol{\theta_{obs} \cdot 100 \cdot MW}}{\boldsymbol{\beta \cdot d \cdot N_A}} = \frac{\boldsymbol{\theta_{obs} \cdot 10^5}}{\boldsymbol{c \cdot d \cdot N_A}}$$

Equation 8 - Calculation of normalized ellipticity per amino acid residue

Θ_{MRW}	mean molar ellipticity per amino acid (deg cm^2 $dmol^{-1}$)
Θ_{obs}	measured ellipticity (mdeg)
MRW	mean residue weight (kDa)
b	protein concentration (mg/ml)
c	protein concentration (µM)
d	pathlength (cm)
MW	molecular weight of protein (kDa)
N_A	number of amino acids

The protein sample was heated from 25 °C to 95 °C at a ramp rate of 1°C min^{-1}.The loss of ellipticity with raising temperature was followed at 222 nm or 205 nm. The apparent midpoint temperature ($T_{1/1*/2}$) of the irreversible unfolding transition was determined as an operational measure of protein stability. Data were fitted with a two state model. The apparent midpoint temperature ($T_{1/1*/2}$) of the irreversible unfolding transition was determined as an operational measure of protein stability. Average values of duplicates were calculated. The melting curves were smoothed with the LOESS algorithm (locally weighted regression) provided by SigmaPlot (https://systatsoftware.com/products/sigmaplot/sigmaplot-smoothing-routines/). Each datapoint along the smooth curve is obtained from a regression of data points close to the curve point with the closest points more heavily weighted. The LOESS smoothing parameters were varied to achieve the best visualization. A polynomial degree of one and a sampling proportion of 0.4 was used in all cases. The default number of intervals (100) for generation of the smooth curve was found to be the best.

5.6.13 Nano differential scanning fluorimetry (nanoDSF)

NanoDSF is an advanced differential scanning fluorimetry method for measuring protein folding and stability in a label-free fluorimetric analysis.The Prometheus NT.48 instrument by NanoTemper Technologies (München, Germany) uses a capillary-based system in combination with an UV detection system designed specifically to monitor shifts in the intrinsic fluorescence emission of tryptophan or tyrosine. As no secondary reporter fluorophores are required, protein solutions can be analyzed independent of buffer compositions, with a wide concentration range of 5 µg to 250 mg/ml. A further advantage of this method is that only very small protein quantities from about 10 µl are required and 48 samples can be analyzed in one run. For nanoDSF measurements, 30 µM protein (subunit concentration) was heated in degassed 50 mM potassium phosphate pH 7.5 from 20 °C to 95 °C at a ramp rate of 1°C min^{-1}.The excitation power was 10 % at 280 nm and

the shift of intrinsic tryptophan fluorescence of proteins upon temperature-induced unfolding was monitored by detecting the emission fluorescence at 330 nm and 350 nm. To exclude buffer effects, one buffer reference was heated, too. Protein melting points ($T_{1/1*}$) were determined as the temperature where the first derivative of the temperature-dependent course of the ratio of tryptophan emission intensities at 330 nm and 350 nm reaches a maximum. The apparent midpoint temperature ($T_{1/1*}$) of the irreversible unfolding transition was determined as an operational measure of protein stability. Mean values of duplicates were calculated.

5.6.14 Radiometric *in vitro* activity assays

The use of the natural substrate HepGP or HepG was not feasible in the *in vitro* assays of PhoB, YvoF and MAT due to its hydrophobicity and the associated insolubility under physiological conditions. Instead, the shorter substrates GGGP or GGG were used, which are less hydrophobic and more soluble in aqueous solutions. Due to the fact that PcrB also accepts GGPP as substrate, although with less efficacy than HepPP (Guldan *et al.*, 2011), we assumed that YvoF, MAT as well as PhoB do not exhibit a strict selectivity with respect to the number of polyprenyl moieties of the substrate, too.

5.6.14.1 **Assay for GGGPS activity**

^{14}C-G1P was synthesized as described in chapter 5.5.1. To test the activity of purified GGGPS enzymes, 1 µM of protein (subunit concentration) was incubated with 20 µM GGPP or 20 µM GPP and 20 µM ^{14}C-G1P (302.72 nCi) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a total volume of 100 µl for 2 h at 40°C. The products were dephosphorylated by adding 1 U calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) for 1 h at 40 °C and extracted according to the method of lipid extraction (5.6.7), analyzed by TLC (5.6.8) and visualized with a phosphorimager system (5.6.9).

5.6.14.2 **Synthesis of ^{14}C-GGG(P)**

To test the activity of PhoB, YvoF or MAT under different conditions, ^{14}C-GGGP, respectively ^{14}C-GGG were synthesized, and radiolabeling assays were performed. ^{14}C-G1P was synthesized as described in chapter 5.5.1. GGPP (150 µM; SIGMA-ALDRICH) was incubated with ^{14}C-G1P (1 µCi, 22.3 µM) and GGGPS from *Archaeoglobus fulgidus* (afGGGPS; 5 µM) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a volume of 300 µl for 2 h at 40°C to produce radiolabeled ^{14}C-GGGP. If ^{14}C-GGG was needed, it was

produced through dephosphorylation by calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) for 1 h at 40 °C. The product was extracted according to the method of lipid extraction (5.6.7). The solvent was evaporated to dryness in a rotary evaporator and the remaining ^{14}C-GGGP or ^{14}C-GGG was dissolved in 40 µl chloroform to a concentration of 167.5 µM under the assumption of 100 % turnover.

5.6.14.3 **Assay for phosphatase activity on GGGP**

To test *B. subtilis* wild type crude extract or samples after chromatographic separation for phosphatase activity the samples were incubated with ^{14}C-GGGP (5.6.14.2, 37.5 nCi, 2.5 µM) in a volume of 100 µl 50 mM Tris/HCl, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT for 2 h at 40 °C (equivalent to pH 8.0 at RT). The products were extracted according to the method of lipid extraction (5.6.7), analyzed by TLC (5.6.8) and visualized with a phosphorimager system (5.6.9).

5.6.14.4 **Assay for acetyltransferase activity on GGG**

In case of the acetylation reaction of YvoF and MAT, conditions were used as follows. <u>For initial activity assays:</u> ^{14}C-GGG (5.6.14.2, 37.5 nCi, 2.5 µM) was incubated with crude extract or 1 µM purified enzyme in a volume of 100 µl 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT for 2h at 40 °C. <u>For acetyl-CoA dependent assays:</u> ^{14}C-GGG (37.5 nCi, 2.5 µM) was incubated with 1 µM purified enzyme and 0-250 µM acetyl-CoA (SIGMA-ALDRICH) in a volume of 100 µl 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT for 2h at 40 °C. <u>For time dependent assays:</u> ^{14}C-GGG (37.5 nCi, 2.5 µM) was incubated with 1 µM purified enzyme and 250 µM acetyl-CoA in a volume of 100 µl 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT for 0-120 min at 40 °C. <u>For comparison of YvoF and MAT:</u> ^{14}C-GGG (37.5 nCi, 2.5 µM) was incubated with 5 µM purified enzyme and 250 µM acetyl-CoA in a volume of 100 µl 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT for 0-3 h at 40 °C.

5.6.14.5 **Discontinuous radiometric assay for YvoF**

<u>For acetyl-CoA dependent kinetics:</u> GGPP (30.8 µM; SIGMA-ALDRICH) was incubated with ^{14}C-G1P (1.2 µCi, 16.08 µM) and afGGGPS (1 µM) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a volume of 500 µl for 3 h at 40 °C to produce ^{14}C-GGGP. ^{14}C-GGG was produced by dephosphorylation with calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) for 1 h at 40 °C. To initiate the acetylation reaction, YvoF (20 nM) in

50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT containing 0-500 µM of acetyl-CoA (SIGMA-ALDRICH) in a volume of 300 µl was added. After 2, 4, 6 and 9 min (80 and 120 µM acetyl-CoA) or 1, 3, 4 and 6 min (200-500 µM acetyl-CoA), 200 µl of the sample were removed and the reaction was stopped immediately by lipid extraction (5.6.7). For GGG-dependent kinetics: 0-410 µM GGPP were incubated with half-stoichiometric amounts of ^{14}C-G1P and afGGGPS (1 µM) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a volume of 900 µl for 3 h at 40 °C to produce ^{14}C-GGGP. Following dephosphorylation by CIP for 1 h at 40°C, ^{14}C-GGG was produced. The GGG concentration (0-8.6 µM in the final setup) was calculated in retrospect to account for losses during synthesis (see below). To initiate the reaction, YvoF (5 nM) in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT containing acetyl-CoA (1 mM) in a volume of 100 µl was added. After 1, 2, 3 and 5 min, 200 µl of the sample were removed and the reaction was stopped immediately by extraction. The products were extracted according to the method of lipid extraction (5.6.7), analyzed by thin layer chromatography (5.6.8) and visualized with a phosphorimager system (5.6.9). Intensities of the spots were quantified using the software OptiQuant 3.0 (PACKARD BioScience), reference spots with known amounts of radioactivity served for calibration. The product concentrations could be calculated from the reaction stoichiometry and the known specific activity of the initial substrate [U-14C]-Glucose. The starting GGG concentrations were calculated in retrospect from the total radioactivity per lane. Finally, reaction velocities were calculated from the protein and product concentrations and incubation times. Kinetic constants were obtained by fitting the Michaelis-Menten equation to the data, using SigmaPlot 12.0.

5.6.14.6 **Maltose dependent activity assay**

To test the activity of MAT or YvoF on the substrate maltose, ^{14}C-maltose was used in radiolabeling assays. [U-14C]-maltose (0.15 µCi, 1.25 µM or 0.3 µCi, 2.5 µM; HARTMANN ANALYTIC) was incubated with MAT or YvoF (4 µM) in 50 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT containing 0.25 mM acetyl-CoA (SIGMA-ALDRICH) in a volume of 200 µl overnight at 40 °C to produce radiolabeled acetyl-maltose. As a control that the acetyltransferase activity is solely due to the enzymes, ^{14}C-maltose was incubated in 50 mM potassium phosphate, pH 7.5 (0.15 µCi, 1.25 µM) with or without acetyl-CoA as described above. The resulting products were dried in a rotary evaporator (EPPENDORF

Centrifuge plus), dissolved in 30 µl water, analyzed by thin layer chromatography on Silica60 plates (MERCK) developed in acetone/chloroform/methanol/water 15:2:2:1 (v/v/v/v), and visualized with a phosphorimager system (PACKARD BioScience, 5.6.9). As standards, ^{14}C-maltose in water (2.5 nCi, 8.3 µM) and acetylated ^{14}C-maltose (1.5 nCi, 1.25 µM) were analyzed. Maltose was acetylated *in vitro* by incubation in 50 µl acetic anhydride overnight at room temperature.

5.6.15 Test for isomerization of single acetylated GGG

To test the single acetylation of GGG for isomerization (Fischer, 1920), acetylated ^{14}C-GGG was synthesized as follows. ^{14}C-G1P was synthesized as described in chapter 5.5.1. GGPP (150 µM; SIGMA-ALDRICH) was incubated with ^{14}C-G1P (10 µCi, 132 µM) and GGGPS from *Chitinophaga pinensis* (cpGGGPS; 20 µM) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a volume of 650 µl for over night at 40 °C to produce radiolabeled ^{14}C-GGGP. To produce ^{14}C-GGG, 5 U calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) was added for 1 h at 40 °C. For production of mainly monoacetylated ^{14}C-GGG bsYvoF (2 µM) and acetyl-CoA (0.2 µM, SIGMA-ALDRICH) were added and incubation was continued for another 2 h. The product was extracted according to the method of lipid extraction (5.6.7), analyzed by thin layer chromatography (5.6.8) and visualized with a phosphorimager system (5.6.9). The upper and lower band of the monoacetylated ^{14}C-GGG was scraped off the Silica60 plate and dissolved in 500 µl chloroform by vortexing. The silica particles were removed by centrifugation for 5 min at 4 °C (HETTICH Centrifuge MIKRO 200R, 13,200 rpm). The supernatant was evaporated to dryness in a rotary evaporator (EPPENDORF Concentrator plus) and the remaining monoacetylated ^{14}C-GGG was dissolved in 100 µl 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 by vortexing. As blank value 25 µl were taken immediately and stored at -20 °C to slow down further acyl migration and the remaining solution was incubated at 40 °C. Other probes were taken after 19 h, 25 h and 43 h and stored at -20 C. All probes were evaporated to dryness in a rotary evaporator and the remnants were dissolved in 20 µl chloroform, analyzed by thin layer chromatography (5.6.8) and visualized with a phosphorimager system (5.6.9).

5.6.16 Steady-state enzyme kinetics

For functional characterization of YvoF and various GGGPS steady-state enzyme kinetics were measured at 40 °C. The GGG- and acetyl-CoA dependent YvoF reaction and the GGPP- and G1P dependent GGGPS reaction were recorded.

5.6.16.1 **DTNB-coupled assay for YvoF**

The YvoF activity for the acetylation of GGG was measured in a coupled colorimetric assay (adapted from Yang *et al.* (2003)). Acetylation of GGG with acetyl-CoA as acetyl group donor resulted in acetylated GGG and CoA-SH as free thiol, which reacts with DTNB (Ellman's reagent, 5,5'-dithiobis-(2-nitrobenzoic acid)). The disulfide bond of DTNB is cleaved yielding 2-nitro-5-thiobenzoate (TNB^-), which ionizes to a TNB^{2-} dianion (yellow color) under basic conditions.

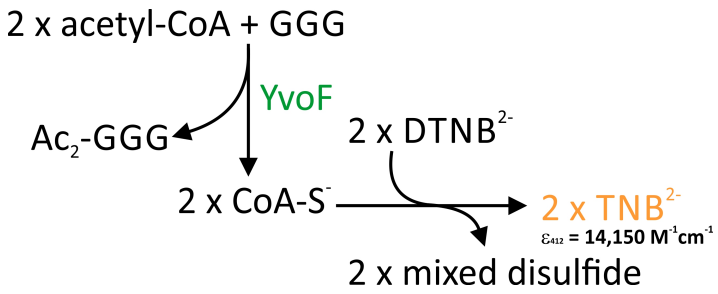


Figure 13 - DTNB-coupled assay of YvoF reaction
GGG: geranylgeranylglycerol, Ac_2-GGG: double acetylated geranylgeranylglycerol, DTNB: 5,5'-dithiobis-(2-nitrobenzoic acid), TNB^{2-}: 2-nitro-5-thiobenzoate dianion, YvoF: polyprenylglycerol *O*-acetyltransferase, ε_{412}: molar extinction coefficient at 412 nm.

The reaction is followed at 412 nm (ε_{412} = 14,150 $M^{-1}cm^{-1}$) and 40 °C using a V650 spectrophotometer (Jasco; d = 1 cm). To determine the concentration of the substrate GGG, it was traced with ^{14}C and synthesized similar as described in chapter 5.6.14.4. Briefly, ^{14}C-G1P was produced from [U-14C]-glucose as described in chapter 5.5.1. GGPP (0.6 mM; SIGMA-ALDRICH) was incubated with ^{14}C-G1P (1.4 µCi, 8.4 µM), cold G1P (1.2 mM, SIGMA-ALDRICH) and GGGPS from *Chitinophaga pinensis* (cpGGGPS; 20 µM) in 1100 µl of 50 mM Tris/HCl, pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 overnight at 40°C to produce ^{14}C-GGGP. ^{14}C-GGG was produced through dephosphorylation by adding calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) and incubation for 2 h at 40 °C. To inactivate and remove the proteins, the solution was heated to 95 °C for 5 min and centrifuged for 5 min at 13,200 g and 4 °C (HETTICH Centrifuge MIKRO 200R). Five GGG preparations of this scale were pooled and stored at 4 °C until kinetic analysis. To

determine the GGG concentration, GGG was analyzed by thin layer chromatography (5.6.8). The intensity of the GGG spot was quantified using the software OptiQuant 3.0 (PACKARD BioScience), reference spots with known amounts of radioactivity served for calibration. The GGG concentration could be calculated from the known specific activity of the initial substrate [U-14C]-Glucose. DTNB was obtained from SIGMA-ALDRICH and used without further purification. The DTNB stock solution (10 mM in 50 mM Tris/HCl pH 8.0) was prepared immediately before use and kept on ice in aluminium foil before pipetting into the assay medium. Acetyl-CoA (SIGMA-ALDRICH) was dissolved in water to a concentration of 10 mM and stored at -80 °C. For acetyl-CoA dependent kinetics, GGG (290 µM), acetyl-CoA (0-870 µM) and DTNB (2 mM) were mixed with YvoF (350 nM) in a total volume of 220 µl 50 mM Tris/HCl, pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80. The reaction was started by adding the enzyme. For GGG-dependent kinetics, GGG (0-258 µM), acetyl-CoA (909 µM) and DTNB (2 mM) were mixed with YvoF (350 nM) in a total volume of 220 µl 50 mM Tris/HCl, pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80. Reaction velocities were calculated from the protein concentration and the initial slopes. Kinetic constants were obtained by fitting the Michaelis-Menten equation to the data, using SigmaPlot 12.0.

5.6.16.2 **Photometric assay for GGGPS enzymes**

For determining kinetic constants for various GGGPS enzymes (cpGGGPS_wt, cpGGGPS_Y143A, tkGGGPS_wt, tkGGGPS_W143A, mtGGGPS_wt, mtGGGPS_W141A and mtGGGPS_A162E) a continuous enzyme-coupled assay for phosphate detection was used (Suarez *et al.*, 2012).

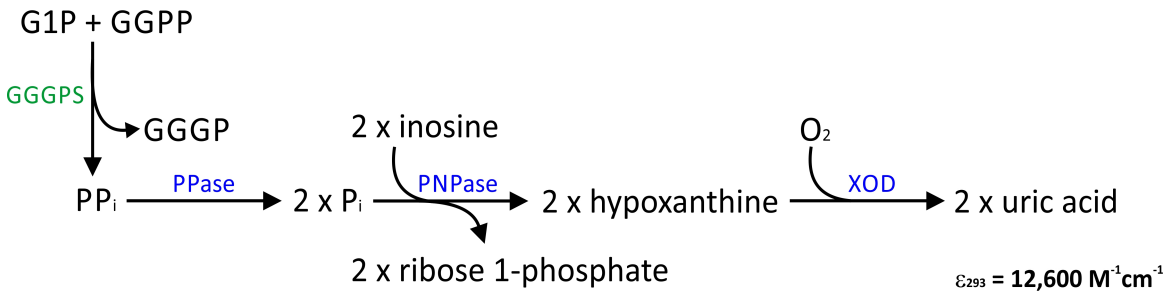


Figure 14 - Enzyme-coupled assay for phosphate detection
G1P: glycerol 1-phosphate, GGPP: geranylgeranyl pyrophosphate, GGGP: geranylgeranylglyceryl phosphate, GGGPS: geranylgeranylglyceryl phosphate synthase, PP_i: pyrophosphate, P_i: orthophosphate, PPase: pyrophosphatase, PNPase: purine nucleoside phosphorylase, XOD: xanthine oxidase, ε_{293}: molar extinction coefficient at 293 nm.

The activity of GGGPS enzymes was assayed in the presence of GGPP and G1P. The assay mixture contained 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80, 1.25 mM inosine (SIGMA-ALDRICH), and 0.0027, 0.25 and 1.25 U/ml of PPase, PNPase (in 50 mM Tris/HCl pH 7.5, 100 mM NaCl), and XOD (in 50 mM Tris pH 8.0), respectively (4.5.3). For GGPP-dependent kinetics, GGPP (0.1-10 µM, SIGMA-ALDRICH) and G1P (SIGMA-ALDRICH, 10 µM for tkGGGPS_wt and tkGGGPS_W143A, 350 µM for cpGGGPS_wt and mtGGGPS_wt, 1 mM mtGGGPS_A162E or 3 mM for mtGGGPS_W141A and cpGGGPS_Y143A) were mixed with the assay mixture in a total volume of 200 µl and incubated at 40 °C. For G1P-dependent kinetics, G1P (0.3-5000 µM) and GGPP (5.55 µM for tkGGGPS_wt and tkGGGPS_W143A, 11 µM for cpGGGPS_wt, mtGGGPS_wt and mtGGGPS_A162E or 7 µM for mtGGGPS_W141A and cpGGGPS_Y143A) were mixed with the assay mixture in a total volume of 200 µl and incubated at 40 °C. The reaction was started by addition of enzyme (30 nM for tkGGGPS_wt and tkGGGPS_W143A, 100 nM for mtGGGPS_wt, cpGGGPS_wt, mtGGGPS_A162E and 500 nM for mtGGGPS_W141A and cpGGGPS_Y143A), and monitored at 293 nm using a V650 spectrophotometer (Jasco; d = 1 cm). At this wavelength, the ε of uric acid was assumed equal to 12,600 $M^{-1}cm^{-1}$ (Suarez *et al.*, 2012). Reaction velocities were calculated from the protein concentration and the initial slopes. Kinetic constants were obtained by fitting the Michaelis-Menten equation to the data, using SigmaPlot 12.0. Average values of triplicates were calculated.

5.6.16.3 **Irreversible heat inactivation**

For determining kinetics of irreversible heat inactivation for GGGPS enzymes, a continuous enzyme-coupled assay for phosphate detection was used as described in chapter 5.6.16.2. All enzymes were diluted or concentrated to a concentration of 40 µM in 50 mM Tris/HCl pH 8.0. The pH values reported here are nominal pHs at room temperature. Actually, the pK_a value of Tris buffer decreases 0.031 units per 1 °C increase and therefore pH values at the denaturating temperature of highly stable proteins studied here could be about 2.3 units smaller than the values measured at lower temperatures (at 100 °C: pH 5.675, at 85 °C: pH 6.2, at 75 °C: pH 6.5, at 65 °C: pH 6.8 and at 50 °C: pH 7.25; Good *et al*., 1966). Tris/HCl buffer had to be used as in this assay potassium phosphate buffer was not usable due to high background effects. 100 µl aliquots were incubated at given temperature for different amounts of time. The samples were chilled

on ice and centrifuged at 13,200 rpm for 5 min at 4 °C (EPPENDORF Centrifuge 5415R). The remaining maximum velocity of the supernatant was assayed in the presence of GGPP and G1P. The assay mixture contained 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80, 1.25 mM inosine, 300 µM G1P, 10 µM GGPP and 0.027, 0.25, and 2.5 U/ml of PPase, PNPase, and XOD, respectively, in a total volume of 200 µl and was incubated at 40 °C. The reaction was started by addition of 1 µl enzyme (200 nM resulting from a 40 µM enzyme sample) or 6 µl enzyme (1.2 µM resulting from 40 µM enzyme sample for mtGGGPS_W141A and cpGGGPS_Y143A), and monitored at 293 nm using a V650 spectrophotometer (Jasco; d = 1 cm). At this wavelength, the ε of uric acid was assumed equal to 12,600 $M^{-1}cm^{-1}$ (Suarez *et al.*, 2012). Reaction velocities were calculated from the initial slopes. A half-life time $t^{app}_{1/2}$ was obtained by fitting a single exponential decay equation to the data, using SigmaPlot 12.0. Average values of triplicates were calculated.

For heat inactivation studies of the monomeric mutant mtGGGPS_A162E_W141A, radiometric assays had to be performed as the activity of the monomeric mutant is so low, that no activity could be observed in the spectrophotometric assay. The protein was diluted to a concentration of 40 µM in 50 mM potassium phosphate pH 7.5. 100 µl aliquots were incubated at given temperatures (45 °C, 50 °C, 55 °C, 60 °C) for different periods of time. The samples were chilled on ice and centrifuged at 13,200 rpm for 5 min at 4 °C (HETTICH Centrifuge MIKRO 200R). The remaining activity of the supernatant was assayed in the presence of GGPP and ^{14}C-G1P. The assay mixture contained 20 µM GGPP and 20 µM ^{14}C-G1P (17.84 nCi) in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 in a total volume of 100 µl. The reaction was started by addition of 25 µl enzyme (10 µM of 40 µM enzyme sample) and incubated for 2 h at 40 °C. The products were dephosphorylated by adding 1 U calf intestinal alkaline phosphatase (CIP, NEW ENGLAND BIOLABS) for 1 h at 40 °C and extracted according to the method of lipid extraction (5.6.7), analyzed by thin layer chromatography (5.6.8) and visualized with a phosphorimager system (5.6.9).

5.7 Bioinformatic analysis – MD simulations

The bioinformatic analysis of mtGGGPS was performed by Kristina Heyn (group of R. Merkl, University of Regensburg). MD simulations were conducted with Yasara (version 16.9., force field AMBER03) based on one monomer deduced from the crystal structure

of mtGGGPS (pdb-id 4mm1). A simulation cells was created, which was 5 Å larger than the protein along each axis and filled with water to a density of 0.997 g/ml, and counterions were added to a final concentration of 0.9% NaCl. The simulation was conducted at three different temperatures, namely 25 °C, 45 °C, and 80 °C, adjusted by the Berendsen thermostat. For each temperature, three independent simulation runs of length 100 ns were performed with slightly different temperatures (± 0.1°C) to alter starting conditions. Each minimization consisted of an initial equilibration step of length 14 ps followed by a heating step to reach the end temperature within 4 ps. During the next 100 ns of simulation, a snapshot was recorded every 20 ps. The snapshot-specific distribution of secondary structure elements of the full protein was determined by means of the function *Yasara:SecStrAll*. Corresponding values originating from the three runs conducted in parallel for each temperature were averaged.

The root mean square fluctuation (RMSF) value gives for a set of atoms their deviation from the mean positions deduced from a series of snapshots. For their calculation, the function *R:rmsf* was used; corresponding values resulting from the three temperature specific runs were averaged. For visualization in PyMol, these values were read as b-factors and a color gradient ranging from blue over yellow to red was used to indicate deviations from 0 Å to 3 Å.

5.8 Protein crystallization and X-ray structure determination

Protein crystallization was performed by hanging drop vapor diffusion method using protein concentrations of 28-40 mg/ml. To explore various conditions, an initial crystal screening was done using commercially available crystallization kits containing 96 unique conditions (4.4.2) using an automatic robotic system. The mutants mtGGGPS_I107E, mtGGGPS_W141A and mtGGGPS_A162E were crystallized with $MgCl_2$ (10 mM end concentration) and with or without glycerol 1-phosphate (1 mM) using the screens Morpheus® (MD1-46), MIDAS™ (MD1-59), PGA Screen™ (MD1-50), PEGRx™1+2 and The structure screen combination (MD1-03) from MOLECULAR DIMENSIONS. Single crystals were transferred to a cryo protectant solution (ethylene glycol) and flash frozen in liquid nitrogen. Data collection and X-ray structure determination was performed in collaboration with Dr. Chitra Rajendran (group of Prof. Dr. Christine Ziegler, University of Regensburg). Data of single crystals were collected at synchrotron beamlines PXIII and PXI

(SLS, Swiss Light Source, Switzerland) at cryogenic temperature. The data processing was done using XDS (Kabsch, 1993) and the data quality assessment was done using phenix.xtriage (Adams *et al.*, 2002). Molecular replacement was performed with MOLREP within the CCP4i suite (Potterton *et al.*, 2004) using wild type structure with PDB ID 4mm1 (Peterhoff *et al.*, 2014) as search model. Initial refinement was performed using REFMAC (Murshudov *et al.*, 1997). The model was further improved in several refinement rounds using automated restrained refinement with the program PHENIX (Adams *et al.*, 2002) and interactive modelling with Coot (Emsley and Cowtan, 2004). The final model was analyzed using the program MolProbity (Davis *et al.*, 2007). The refinement statistics are listed in Table S2 - Table S4.

Diffraction data and model coordinates have been deposited at the PDB with entry codes 5nez (mtGGGPS_I107E), 5nf1 (mtGGGPS_W141A) and 5ndy (mtGGGPS_A162E).

6 IDENTIFICATION AND CHARACTERIZATION OF HEPTAPRENYLGLYCERYL PHOSPHATE PROCESSING ENZYMES IN *Bacillus subtilis*

6.1 Preface

Eduard Hochmuth (Lehrstuhl für Biochemie, Biochemie I) performed ESI-MS with bsPhoB samples. Parts of this chapter (including the methods section) as well as figures have been published equally worded in Linde *et al.* (2016). The publication was written by Prof. Dr. Reinhard Sterner, Dr. David Peterhoff, Dr. Patrick Babinger and myself. This chapter also contains data from my master thesis (Linde, 2013). Details are indicated in the following text. This data was included in this work to provide the complete story of the identification of polyprenylglyceryl phosphate processing enzymes.

6.2 Objective of this thesis part

As outlined in the introduction, the first committed step in the biosynthesis of archaea-like ether lipids in Archaea catalyzed by GGGPS is followed by an addition of a second polyprenyl moiety by the enzyme DGGGPS. The resulting core phospholipid gets subsequently processed and modified before it is incorporated into the membrane (Figure 2; Koga and Morii, 2007). To analyze if the same biosynthesis pathway is used in Bacteria containing *araM* and *pcrB* genes, *in vivo* radiolabeling experiments were performed. *B subtilis* absorbes substantial amounts of radiometricly labeled G1P present in the growth medium (Guldan *et al.*, 2011). Lipid extraction (Kates, 1986) followed by thin layer chromatography enabled the analysis of the *in vivo* products of bsPcrB. Reversed phase high pressure liquid chromatography followed by mass spectrometry and NMR studies revealed that radiolabeled G1P is connected to HepPP producing HepGP, which gets subsequently dephosphorylated to HepG and than acetylated *in vivo* (Figure 8; Guldan *et al.*, 2011). Because this displays a main difference to archaeal G1P-based ether lipids, it was attempted to identify the enzymes involved in these modifications in *B*. *subtilis*.

6.3 Results and Discussion

In a previous study, Harald Guldan has developed a method to visualize the formation of polyprenyl ether derivatives *in vivo*, based on feeding *B. subtilis* cells with radiolabelled ^{14}C-G1P (Guldan *et al.*, 2011). The *B. subtilis* cells thereby provide the polyprenyl pyrophosphate substrate, particularly GGPP and HepPP in about equal amounts. Because a large collection of individual *B. subtilis* knockout strains is available from the National BioResource Project (NBRP) in Japan (Kobayashi *et al.*, 2003), it was obvious to use this methodology and screen this library to identify the phosphatase and acetyltransferase that act in the bacterial ether lipid synthesis pathway (Figure 8).

6.3.1 Identification and Characterization of a heptaprenylglyceryl phosphate processing phosphatase from *B. subtilis*

6.3.1.1 Purification of a heptaprenylglyceryl phosphate phosphatase from *B. subtilis* wild type

A search for all *B. subtilis* genes that have been annotated as a proved or putative phosphatase/pyrophosphatase yielded 66 hits, and all 26 available knockout strains for these genes were ordered from NBRP. The strains were grown in presence of ^{14}C-labelled G1P, and lipids were extracted and analyzed by thin layer chromatography (5.6.10). No strain showed an altered ether lipid composition that would indicate phosphatase deficiency. The reason could be that a knockout of the distinct phosphatase acting on HepGP was not available or alternatively, that many phosphatases can complement each other, as discussed later.

As an alternative approach, the HepGP phosphatase activity was enriched from *B. subtilis* wild type cell extracts using various conventional biochemical techniques, namely a combination of ammonium sulfate precipitation (5.4.2.3), followed by cation exchange chromatography (5.4.2.4) and size exclusion chromatography (5.6.5). After each purification step, fractions were screened for dephosphorylation activity (5.6.14.3), and the most active fractions were used for the next enrichment step. The use of the natural substrate HepGP was not feasible in the *in vitro* assays due to its hydrophobicity and the associated insolubility under physiological conditions. Instead, the shorter substrate GGGP was used, which is less hydrophobic and more soluble in aqueous solutions. Due to

the fact that PcrB also accepts GGPP as substrate, although with less efficacy than HepPP (Guldan *et al.*, 2011; Peterhoff *et al.*, 2014), it was assumed that the enzymes involved in subsequent modifications do not exhibit a strict selectivity with respect to the number of polyprenyl moieties of the substrate, too. The enriched proteins were identified by HPLC-coupled electrospray ionisation - mass spectrometry. Among eleven proteins one phosphatase was found, the alkaline phosphatase PhoB. A knockout of *phoB* was not available from NBRP. All experiments up to this point have already been performed in my master thesis (Linde, 2013).

6.3.1.2 **Heterologous expression and purification of bs*phoB***

To express bs*phoB* heterologously in *E. coli* cells, the gene was amplified from genomic DNA of *B. subtilis* (5.3.2) and cloned into the pET21a vector via *Nde*I/*Xho*I restriction sites to add a C-terminal $(His)_6$ tag to the recombinant protein (5.3.8). The cloned gene was verified by sequencing (5.3.9). *E. coli* BL21-CodonPlus (DE3) RIPL cells were transformed with pET21a_bs*phoB* and the gene expression was induced by IPTG for 3 h at 37 °C (5.4.1.1). The protein was enriched by metal chelate affinity chromatography (5.4.2.1) and its purity was estimated to be about 90 % by SDS-PAGE (5.6.3, Figure 15). Monomeric bsPhoB consists of 470 residues and has a molecular weight of 50 kDa, which is consistent with the size determined by SDS-PAGE. Protein yields of about 2 mg/l culture could be obtained for bsPhoB.

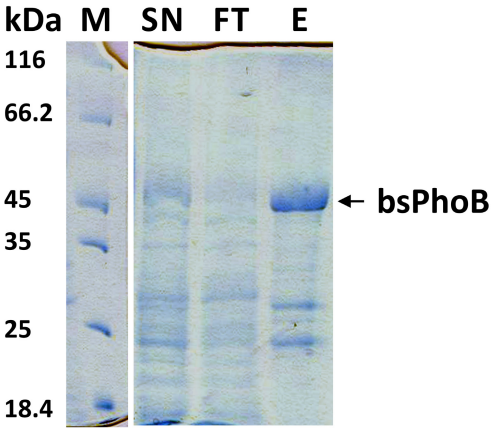


Figure 15 - Purity of bsPhoB
SDS-PAGE analysis (12.5 % polyacrylamide) of purified bsPhoB. Size marker (M), molecular weight in kDa. SN: supernatant, FT: flowthrough, E: eluate. bsPhoB is marked by an arrow.

6.3.1.3 *In vitro* activity of bsPhoB

The heterologously expressed bsPhoB was tested for *in vitro* dephosphorylation activity on ^{14}C-GGGP (5.6.14.3, Figure 16).

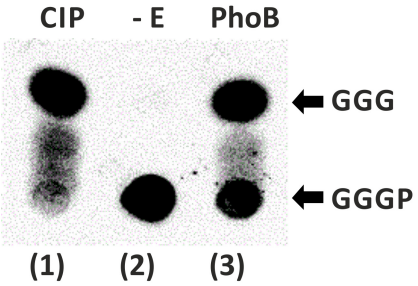


Figure 16 - Activity assay of purified bsPhoB
The substrate ^{14}C-GGGP was incubated with different protein samples for 4 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: calf intestinal alkaline phosphate (CIP, positive control), lane 2: no enzyme added (- E, negative control), lane 3: purified bsPhoB (PhoB). The origin of the chromatography (marked by GGGP from GGGP production) and the product spot (GGG) are marked by arrows. The figure was taken from Linde *et al.* (2016).

In fact, PhoB could dephosphorylate GGGP, as it is the case for calf intestinal phosphatase (CIP), which was used as a positive control. PhoB is a member of an alkaline phosphatase multigene family consisting of at least four members in *B. subtilis*. Which of these phosphatases gets expressed is dependent on the physiological conditions. Their cellular localization varies from secreted to membrane associated, and it has been shown that the main members of this family, PhoA and PhoB, can mutually complement each other (Abdel-Fattah *et al.*, 2005; Hulett *et al.*, 1991). Although it has been proposed that the main biological function of the Pho enzymes is the provision of phosphate during phosphate starvation and that *phoB* expression is accordingly controlled, it was possible to purify substantial amounts of PhoB from *B. subtilis* cells grown in rich medium. Like most other alkaline phosphatases, PhoB can hydrolyze a large number of phosphorylated components (Hulett *et al.*, 1991). The same is the case for CIP, which has already been used by others to dephosphorylate ether lipids (Zhang and Poulter, 1993a) and therefore has served as a positive control here. In consequence, it is very likely that HepGP gets dephosphorylated in an unspecific way, either by PhoB or by any other alkaline phosphatase existing in the cell. For that reason investigating the dephosphorylation reaction in further detail was abandoned.

6.3.2 Identification and Characterization of a heptaprenylglycerol processing acetyltransferase from *B. subtilis*

6.3.2.1 *In vivo* radiolableling experiments using a knockout library

Like for phosphatase activity, all *B. subtilis* genes with an annotation as a (putative) acetyltransferase, deacetylase, acetylase or phosphotransacetylase were identified, resulting in 63 candidate genes. For 41 of them, knockout strains were available from NBRP. All those strains were tested in the *in vivo* radiolabeling assay for their ability to acetylate HepG (5.6.10). These radiolabeling assays were already performed in my master thesis (Linde, 2013). Two of the knockout strains showed an altered thin layer chromatography profile compared to the wild type, with missing spots for single acetylated Ac-HepG and double acetylated Ac_2-HepG (Figure 17).

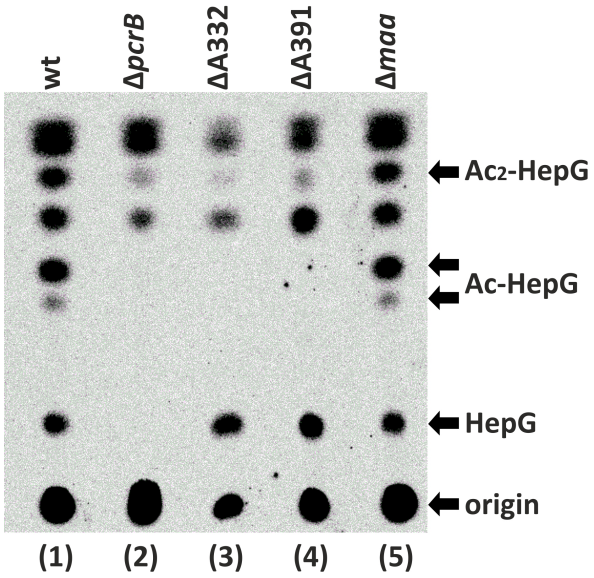


Figure 17 - Identification of acetyltransferase-deficient strains
Different *B. subtilis* strains were grown overnight in the presence of radiolabeled G1P. Lipids were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: *B. subtilis* wild type (wt, reference), lane 2: *B. subtilis ΔpcrB* (background reference, produces no HepGP and thus no HepG), lane 3: *B. subtilis ΔyvoF* (NBRP code: MGNA-A332), lane 4: *B. subtilis ΔyvoF* (NBRP code: MGNA-A391), lane 5: *B. subtilis Δmaa* (NBRP code: MGNA-B865). The origin of chromatography as well as spots of HepG, single acetylated Ac-HepG and double acetylated Ac_2-HepG are marked with arrows. Two separate spots occur for Ac-HepG, because there are two alternative positions for the first acetylation on the glycerol backbone. The acetyl group can most likely swap spontaneously between them, as discussed in the text (cf. Figure 8). Please note that all strains, including the background reference Δ*pcrB*, produce some background spots at low but comparable intensities, because ^{14}C-G1P can go into glycolysis and by this way into many cellular compounds (Guldan *et al.*, 2008). One of those faint spots superimposes with the Ac_2-HepG spot (Guldan *et al.*, 2011). The figure was taken from Linde *et al.* (2016).

In both strains (NBRP-codes MGNA-A332 and MGNA-A391) the same uncharacterized gene, *yvoF,* is inactivated, which codes for a putative *O*-acetyltransferase. A Δ*pcrB* knockout strain served as a background reference. It does not produce HepGP, and therefore no acetylated HepG as well (Guldan *et al.*, 2011). A knockout strain of the *maa* gene encoding maltose *O*-acetyltransferase (MAT) served as a positive control for the *B. subtilis* strains used from NBRP. As discussed later, MAT is the closest homologue to YvoF in *B. subtilis*.

In the *B. subtilis* genome, the *yvoF* gene is located in one operon together with the genes of a serine/threonine protein kinase/phosphorylase (*hprK*), prelipoprotein diacylglycerol transferase (*lgt*), a putative integral inner membrane protein (*yvoD*), and a pyrophosphatase/ glycerol 3-phosphatase (*ppaX*, Galinier *et al*., 1998). To enlighten a possible involvement of the products of these genes in the processing of HepGP, the corresponding knockout strains were ordered from NBRP and subjected to *in vivo* radiolabeling experiments as described before (Figure 18).

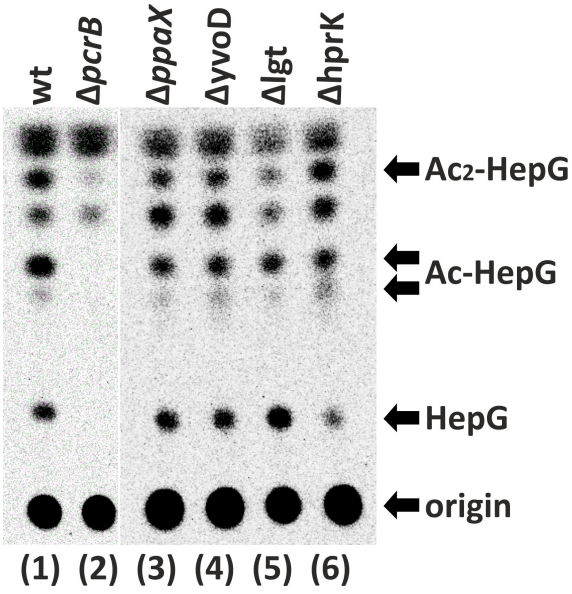


Figure 18 - Testing of *yvoF* operon-associated genes for deficiencies in HepGP processing
Different *B. subtilis* strains were grown overnight in the presence of radiolabeled G1P. Lipids were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: *B subtilis* wild type (wt, reference), lane 2: *B. subtilis* Δ*pcrB* (background reference), lane 3: *B. subtilis* Δ*ppaX* (pyrophosphatase, NBRP code: MGNA-A331), lane 4: *B. subtilis* Δ*yvoD* (uncharacterized membrane protein, NBRP code: MGNA-A330), lane 5: *B. subtilis* Δ*lgt* (prolipoprotein diacylglyceryl transferase, NBRP code: MGNA-A388), lane 6: *B. subtilis* Δ*hprK* (HPr kinase/phosphorylase, NBRP code: MGNA-A328). The origin of chromatography as well as spots of HepG, single acetylated Ac-HepG and double acetylated Ac_2-HepG are marked with arrows. The samples were run on two separate TLC plates. The figure was taken from Linde *et al.* (2016).

No altered thin layer chromatography profile could be observed. Therefore, an essential involvement of these gene products in the processing of HepGP is rather unlikely.

6.3.2.2 **Cellular localization of bsYvoF**

To identify the localization of the bsYvoF protein in the cell, membrane fractionation experiments were performed (5.6.6). A *B. subtilis* crude extract was separated by centrifugation into a pellet fraction containing cell debris and insoluble proteins, and a supernatant fraction containing soluble proteins including membranes. The supernatant was further fractionated by ultracentrifugation, resulting in a supernatant fraction containing soluble proteins only and a pellet fraction containing membranes and membrane proteins only. All fractions were tested for acetyltransferase activity by thin layer chromatography (Figure 19).

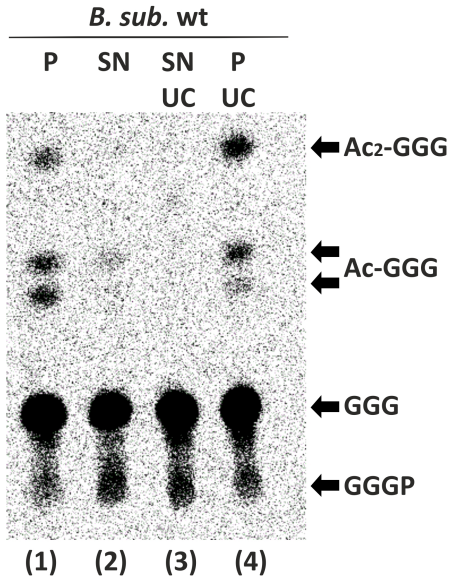


Figure 19 - Cellular localization of bsYvoF
Supernatant (SN) and pellet (P) fractions of a *B. subtilis* wild type crude extract were tested for acetyltransferase activity. The supernatant was further fractionated by ultracentrifugation (UC). The extracts were incubated with the substrate ^{14}C-GGG for 2 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: Pellet crude extract, lane 2: supernatant crude extract, the supernatant was re-centrifuged: lane 3: supernatant after ultracentrifugation (UC), lane 4: pellet after ultracentrifugation (UC). The origin of chromatography (marked by residual GGGP from GGG production) and the product spots are marked by arrows. The figure was taken from Linde *et al.* (2016).

Most of the activity was present in the pellet fraction of the crude extract. Only little activity was present in the supernatant, which could be concentrated by ultracentrifugation, resulting in a significant activity in the membrane fraction. No activity

was left in the fraction of soluble proteins. This indicates that bsYvoF is in some way associated to the membrane and is not present in the cytosol.

6.3.2.3 Heterologous expression and purification of bs*yvoF*

To express bs*yvoF* heterologously in *E. coli* cells, the gene was amplified from genomic DNA of *B. subtilis* (5.3.2) and cloned into the pET21a vector via *Nde*I/*Xho*I restriction sites to add a C-terminal $(His)_6$ tag to the recombinant protein (5.3.8). The cloned gene was verified by sequencing (5.3.9). *E. coli* BL21-CodonPlus (DE3) RIPL cells were transformed with pET21a_bs*yvoF* and gene expression was induced by IPTG as described in chapter 5.4.1.1. The protein was enriched with a His SpinTrap™ column (GE HEALTHCARE, 5.4.2.1) and the expression level was analyzed by SDS-PAGE and Western Blotting (5.6.3, 5.6.4, data not shown). As the initial test expressions afforded poor protein yields, experiments to obtain a larger amount of soluble protein were performed. For this purpose, different resuspension buffers were used for purification of YvoF after expression (4.10.3), and the amount of eluted protein was again analyzed by Western Blotting (Figure 20). The amplification and cloning of bs*yvoF* into pET21a as well as initial test expressions and resuspension buffer optimization were already performed in my master thesis (Linde, 2013).

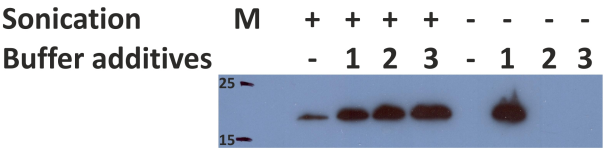


Figure 20 - Western blot for analysis of the best purification conditions of heterologously expressed YvoF Equal portions of bs*yvoF* expressing cells were harvested by centrifugation and resuspended in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, containing either 9 mM CHAPS (1), 1 M NaCl (2) or 0.4 mM TritonX100 (3). One single culture was used to ensure an identical number of cells in each sample. After 1 h, cells were either disrupted by sonication (+) or directly used (-) for SDS-PAGE and Western Blot analysis. Detection was done with an anti-$(His)_6$ tag antibody. The lines on the left represent the size marker (M), molecular weights in kDa. The figure was taken from Linde *et al.* (2016).

Monomeric bsYvoF consists of 180 residues and has a molecular weight of 19.3 kDa, which is consistent with the size determined by SDS-PAGE. The fact that high salt concentrations were sufficient to increase purification yield suggests that bsYvoF is not an integral membrane protein. It has been shown that a solubilization by elevated salt concentrations is only possible for peripherally attached membrane proteins or proteins with lipid anchor, but not for transmembrane proteins (Steck and Yu, 1973). Sequence analysis and

comparisons with the related MAT protein, of which several structures have been solved (e.g. pdb codes: 3hjj, *B. anthracis*; 1ocx, *E. coli*, (Lo Leggio *et al.*, 2003)), also revealed no indications for YvoF being an integral membrane protein. How YvoF is associated to the membrane remains unknown to date. The highest yield of soluble protein could be obtained by using the resuspension buffer containing 9 mM CHAPS and without sonication of the crude extract. Therefore, this purification method was used for all following expressions of bs*yvoF*. As expression of pET21a_bs*yvoF* resulted in pretty low protein yields that were only detectable by western blot even at preparative scale (data not shown), another expression vector, pET28atrxN (4.7.3), was tested for expression at analytical scale. This vector provides an optional C-terminal $(His)_6$ tag sequence plus a N-terminally fused thioredoxin (TrxA). Thioredoxin is one out of many commonly used fusion partners to mediate production of soluble protein (LaVallie *et al.*, 1993). To further improve expression of bs*yvoF*, different IPTG concentrations were used for induction of the gene expression, and to analyze possible temperature affects, the expression was carried out at 30 °C and 37 °C. The proteins were purified with His SpinTrap™ columns (GE HEALTHCARE, 5.4.2.1) and the expression level was analyzed by SDS-PAGE (5.6.3, Figure 21).

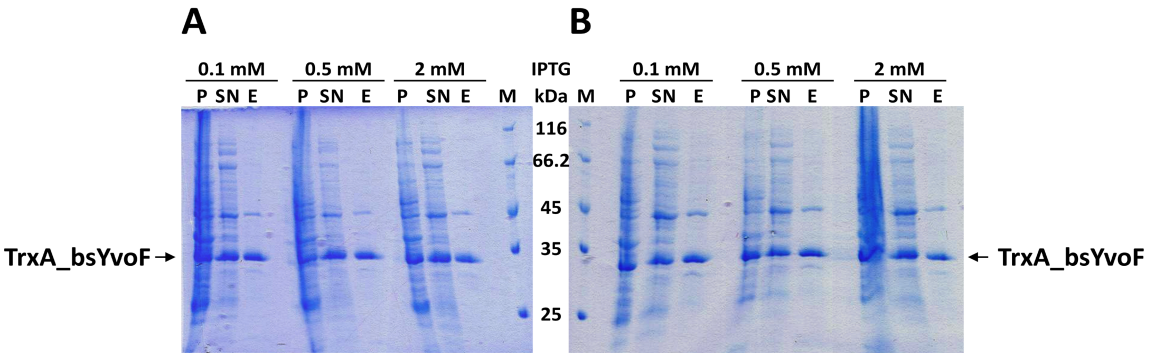


Figure 21 - *TrxA*_bs*yvoF* expression as induced by different IPTG concentrations and temperatures
SDS-PAGE analysis (12.5 % polyacrylamide) of TrxA_bsYvoF after expression at 37 °C (A) and 30 °C (B). Size marker (M), molecular weights in kDa. P: pellet, SN: supernatant, E: eluate. TrxA_bsYvoF is marked by arrows.

Obviously the fusion of TrxA to bsYvoF led to improved solubility of the protein, as no western blot is needed for protein detection. Monomeric bsYvoF fused to TrxA consists of 297 residues and has a calculated molecular weight of 32.6 kDa, which is consistent with the size determined by SDS-PAGE. No difference between various IPTG concentrations or

incubation temperatures could be detected, therefore preparative scale expressions were performed at 37 °C with 0.5 mM IPTG (Figure 22). The protein was enriched by metal chelate affinity chromatography (5.4.2.2) and the purity was estimated to be > 95 % by SDS-PAGE (5.6.3). The fusion of TrxA to bsYvoF enabled protein yields of about 0.4 mg/l culture.

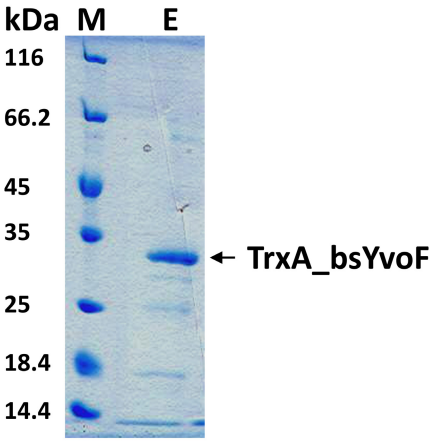


Figure 22 - Purity of TrxA_bsYvoF
SDS-PAGE analysis (12.5 % polyacrylamide) of the purified TrxA_bsYvoF. Size marker (M), molecular weights in kDa, E: eluate. TrxA_bsYvoF is marked by an arrow.

The fusion tag TrxA is removable by thrombin digestion, leaving an N-terminal Gly-Ser extension to the protein (Waugh, 2011). Different thrombin concentrations were added to the purified protein and incubated at different temperatures over night, and after centrifugation cleavage was analyzed by SDS-PAGE (5.6.3, Figure 23).

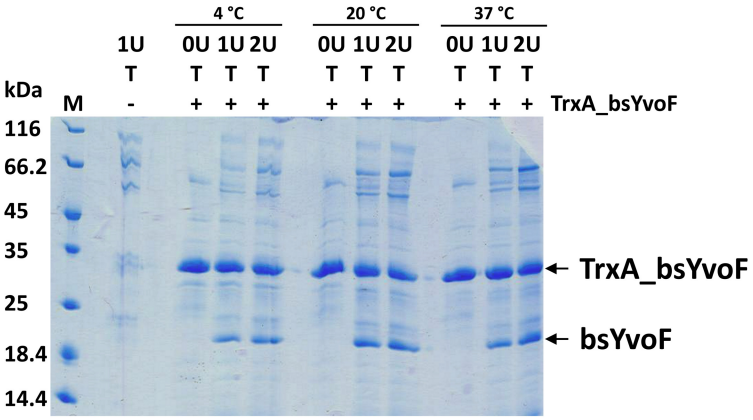


Figure 23 - Cleavage of the TrxA_bsYvoF fusion protein by thrombin
SDS-PAGE analysis (12.5 % polyacrylamide) of thrombin cleavage of TrxA_bsYvoF. Size marker (M), molecular weights in kDa. 1 unit thrombin per mg protein (1U T) or 2 units thrombin per mg protein (2U T) were added to TrxA_bsYvoF (+) at different temperatures (4 °C, 20 °C, 37 °C), as reference control no thrombin was added to the protein (0U T +), as background reference only thrombin without protein was analyzed (1U T -). TrxA_bsYvoF and cleaved bsYvoF are marked by arrows.

In all cases, no more than 30 % cleavage could be obtained, independent of temperature or thrombin concentration. Similarly, subsequent thrombin cleavage experiments with higher thrombin concentrations of 5 U, 10 U, 20 U, 40 U or 80 U thrombin per mg protein did not show an altered cleavage profile (data not shown). Therefore, all following experiments were performed with the fusion protein TrxA_bsYvoF, in the text referred to as bsYvoF.

6.3.2.4 *In vitro* activity of bsYvoF

Different samples of the heterologously expressed bsYvoF were tested for acetyltransferase activity by incubation of protein samples with the *in vitro* produced substrate ^{14}C-GGG (5.6.14.2). The use of the natural substrate HepG was not feasible in the *in vitro* assays due to its hydrophobicity and the associated insolubility under physiological conditions. Instead, the shorter substrate GGG was used, which is less hydrophobic and more soluble in aqueous solutions as already discussed for bsPhoB. The products were extracted, separated by thin layer chromatography, and visualized by autoradiography (5.6.14.4). First, crude extracts of *E. coli* expression cultures were tested (Figure 24).

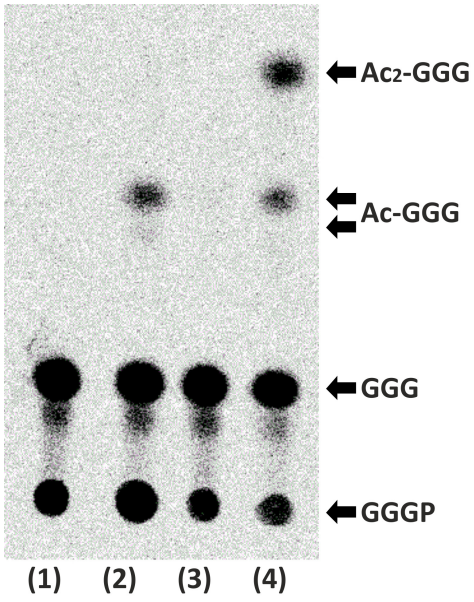


Figure 24 - Activity of bsYvoF in *E. coli* crude extract and purified protein samples
Different protein samples were incubated with ^{14}C-GGG for 2 h at 40 °C. The products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: *E.coli* crude extract (negative control); lane 2: *E.coli* crude extract from cells overexpressing wild type *yvoF*; lane 3: purified bsYvoF; lane 4: purified bsYvoF plus 250 µM acetyl-CoA. The origin of chromatography (marked by residual GGGP from GGG production) and the product spots are marked by arrows. The figure was taken from Linde *et al.* (2016).

No acetyltransferase activity could be observed in the *E. coli* crude extract without over-expressed bsYvoF (background control). This excludes a disturbing acetyltransferase cross-contamination originating from *E. coli*. *E. coli* crude extracts with over-expressed bsYvoF show significant acetyltransferase activity. However, acetyltransferase activity is lost when testing a purified bsYvoF sample. This indicates that the second substrate providing the acetyl moiety is present in the *E. coli* cell extract, but not in the purified protein solution. After the addition of 250 µM acetyl-CoA to the assay with purified bsYvoF, acetyltransferase activity is regained. This demonstrates that acetyl-CoA is the cosubstrate of the YvoF reaction. Interestingly, the thin layer chromatography profile differs between *lane 2* and *lane 4*. In *lane 2* only the monoacetylated Ac-GGG can be found, whereas in *lane 4* both products are visible, the mono- as well as the diacetylated GGG. Because we assumed that the generation of mono- and diacetylated product might be dependent on acetyl-CoA concentration, an activity assay with varying acetyl-CoA concentrations was performed (Figure 25*A*).

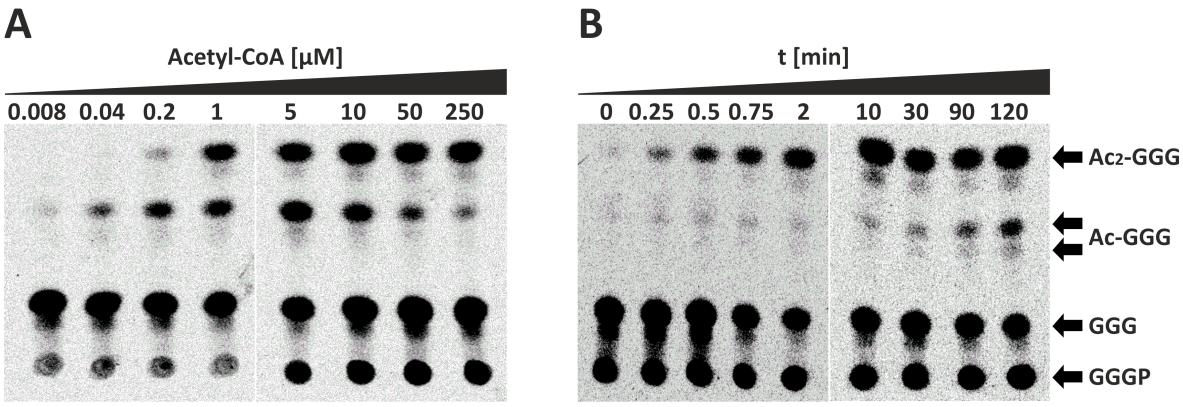


Figure 25 - Dependence of product formation by purified bsYvoF on substrate concentration and time
(**A**) bsYvoF was incubated with ^{14}C-GGG and increasing concentrations of acetyl-CoA for 2 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. (**B**) bsYvoF was incubated with ^{14}C-GGG and 250 µM acetyl-CoA for different times at 40°C. The generated products were extracted, separated by thin layer chromatography and visualized by autoradiography. The origin of chromatography (marked by residual GGGP from GGG production) and the product spots are marked by arrows. The samples were run on two separate TLC plates each. The figure was taken from Linde *et al.* (2016).

With increasing acetyl-CoA concentration, the diacetylated Ac_2-GGG band gets more prominent, but interestingly, the mono-acetylated Ac-GGG band diminishes again after reaching a maximum at concentrations of about 5 µM. To get more insight into the kinetics of mono- and diacetylation, the accumulation of the products over time was

followed at a constant acetyl-CoA concentration of 250 µM (Figure 25*B*). Under those conditions, the formation of the diacetylated product is favored at the beginning. During the first 10 min after starting the reaction, only a tiny, but constant fraction of Ac-GGG can be detected, while the amount of Ac_2-GGG increases with time. Later than 10 minutes after starting the reaction, the concentration of monoacetylated product is increasing, while the amount of the diacetylated product almost remains constant. It can be concluded from the two latter experiments that when acetyl-CoA is redundantly available at high concentration (i.e. at the beginning of the reaction shown in Figure 25*B*), the second acetylation takes place before the mono-acetylated product can diffuse away from the active site of the enzyme. With decreasing concentration of acetyl-CoA, more monoacetylated product is released before the second acetylation can take place.

As mentioned in the legend to Figure 8, there are two alternative hydroxyl groups on the glycerol backbone for the first acetylation to take place, giving two spots on the chromatograms in varying intensities. We speculated that the acetyl group swaps spontaneously between the two hydroxyl groups by acyl migration. To verify this hypothesis, isomerization experiments were performed (5.6.15). Briefly, monoacetylated ^{14}C-GGG was synthesized *in vitro* under conditions where monoacetylation is favored. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. The upper and lower band of the monoacetylated ^{14}C-GGG were cut out from the Silica60 plate and extracted according to 5.6.15. ^{14}C-GGG was dissolved in 100 µl 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80 by vortexing. As a blank value, 25 µl were taken immediately and stored at -20 °C to slow down further acyl migration and the remaining solution was incubated at 40 °C. Other probes were taken after 19 h, 25 h and 43 h and stored at -20 C. All probes were evaporated to dryness in a rotary evaporator and the remnants were dissolved in 20 µl chloroform, analyzed by thin layer chromatography (5.6.8), visualized with a phosphorimager system (5.6.9, Figure 26), and quantified using the OptiQuant Software.

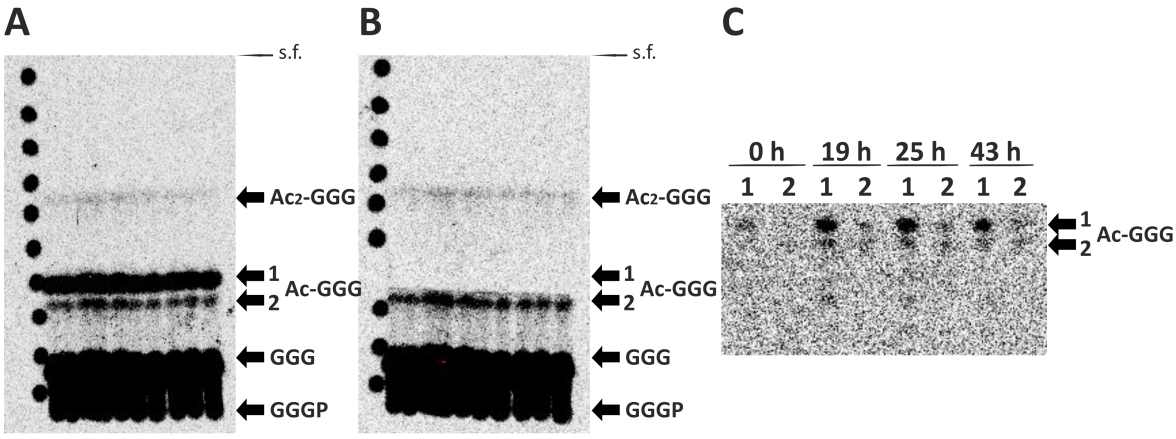


Figure 26 - Isomerization experiment of monoacetylated ^{14}C-GGG
(**A**) The YvoF reaction was carried out at conditions where monoacetylation of GGG is favored (0.2 µM acetyl-CoA, cf. Figure 25). The generated products were extracted, separated by thin layer chromatography and visualized by autoradiography. The two isomers (1, 2) of monoacetylated GGG are marked by arrows. (**B**) Isomer 1 of monoacetylated GGG was cut out from the TLC plate and the plate was re-autoradiographed to verify a proper separation of isomer 1 and 2. Afterwards, isomer 2 was cut out as well. Both samples were eluted from the silica60 material by $CHCl_3$, the solvent was evaporated, and the samples were redissolved in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2 % Tween80, because acyl migration is known to be faster in aqueous buffers. (**C**) The isomers were incubated at 40 °C, samples were taken after 0 h, 19 h, 25 h, 43 h and frozen to -20 °C to slow down further acyl migration. The samples were again analyzed by TLC and autoradiography. The intensities of the spots representing isomer 1 and 2 were quantified by OptiQuant Software.

The ratio isomer 1 : isomer 2 was 6 : 1 - 8 : 1 after the first thin layer chromatography (Figure 26*A*). Isomer 1 of monoacetylated GGG was cut out from the TLC plate and the plate was re-autoradiographed to verify a proper separation of isomer 1 and 2 (Figure 26*B*). After extraction of isomer 1 and incubation for 19 h, the ratio was 3 : 1 – 4 : 1 (Figure 26C, *lane 19 h*). This indicates that within 19 h isomer 1 of Ac-GGG has re-isomerized to at least the same ratio as it was at the beginning. It must be noted that the blank value unfortunately but reproducibly shows only faint spots. This could be due to the poor solubility of Ac-GGG in the aqueous buffer, making it necessary to incubate for some hours for redissolving. This isomerization makes a putative regiospecificity of YvoF irrelevant under physiological conditions, and a detailed kinetic study very difficult. Acyl migration has first been described on acylated esters of glycerol (Fischer, 1920) and is a well known hindering factor in the isolation, characterization and synthesis of lipids (Kodali *et al.*, 1990; Lyubachevskaya and Boyle-Roden, 2000; Turon *et al.*, 2003).

To determine the catalytic parameters of bsYvoF both a DTNB-coupled assay (5.6.16.1, Figure 27) and a discontinuous radiometric assay (5.6.14.5) were used.

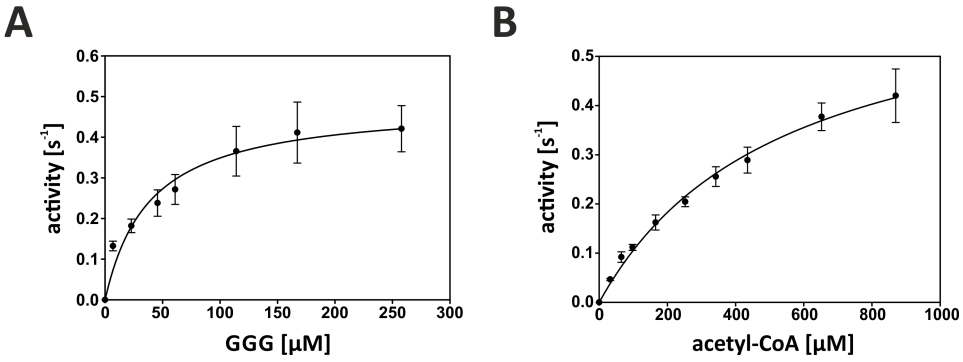


Figure 27 – Steady-state kinetic measurements of acetyltransferase activity of bsYvoF
(**A**) 350 nM bsYvoF was incubated with 0-258 µM of the substrate ^{14}C-GGG and acetyl-CoA (909 µM).
(**B**) 350 nM bsYvoF was incubated with the substrate ^{14}C-GGG (290 µM) and 0-870 µM of acetyl-CoA. Reactions were performed in 50 mM Tris/HCl pH 8.0, 10 mM $MgCl_2$, 0.2% Tween80 plus 2 mM DTNB at 40°C and started by adding the enzyme. The increase of absorbance at 412 nm was followed and reaction velocities were calculated from the protein concentration and the initial slopes. The experiment was done in triplicates, the error bars show standard deviations. Kinetic constants were obtained by fitting the Michaelis-Menten equation to the data, using SigmaPlot 12.0 (Table 10). The figure was taken from Linde *et al.* (2016).

The results of both assays are in good accordance (Table 10). The high K_M for acetyl-CoA supports that when this substrate is present at low, non-saturating concentrations (<< K_M), Ac-GGG may be released before a second acetylation of the same acceptor molecule can take place (Figure 25).

Table 10 - Catalytic parameters of bsYvoF at 40 °C
The parameters were determined independently with a DTNB-coupled assay in triplicate (given with standard deviations) and a discontinuous radiometric assay without replicates (values in parentheses). Saturation curves of the DTNB assay are shown in Figure 27.

substrate	GGG	acetyl-CoA
K_M [µM]	39 ± 9 (> 8)[a]	538 ± 70 (162)
k_{cat} [s^{-1}]	0.48 ± 0.03 (> 0.22)[a]	0.67 ± 0.05 (> 0.22)[a]
k_{cat}/K_M [$M^{-1}s^{-1}$]	$1.25 \times 10^4 \pm 0.38 \times 10^4$ (2.8×10^4)	$0.13 \times 10^4 \pm 0.07 \times 10^4$ (> 0.14×10^4)[b]

[a] In the discontinuous radiometric assay, saturation could not be reached with the substrate GGG. Therefore, only $k_{cat}/K_{M(GGG)}$ could be derived experimentally from the initial slope of the saturation curve. Because the saturation curve is linear up to the highest GGG concentration that could be used (8 µM), $K_{M\,(GGG)}$ must be > 8 µM. A minimal k_{cat} was calculated from $k_{cat}/K_{M(GGG)}$ with the assumption $K_{M(GGG)}$=8 µM.
[b] Because the acetyl-CoA dependent reaction was assayed under non-saturating concentrations for GGG, $k_{cat}/K_{M(acetyl\text{-}CoA)}$ could only be estimated from the calculated k_{cat} and the experimentally determined $K_{M(AcCoA)}$. The table was taken from Linde *et al.* (2016).

6.3.3 Evolutionary and functional relationship between bsYvoF and bsMAT

The closest homologue of YvoF is the maltose *O*-acetyltransferase (MAT), the *maa* gene product (Boos *et al.*, 1981; Brand and Boos, 1991), which is in turn a close relative of the galactoside *O*-acetyltransferase (GAT), the *lacA* gene product (Zabin *et al.*, 1962). MAT and GAT acetylate a large variety of hexoses with different efficiencies, with a focus of MAT on glucosides and GAT on galactosides (Lo Leggio *et al.*, 2003). The structures of MAT from *E. coli* (pdb-code: 1ocx) and *B. anthracis* (pdb-code: 3hjj, 3igj in complex with acetyl-CoA) have been solved, but the biological functions of MAT and GAT have not yet been verified experimentally. It is supposed that they act as detoxifying enzymes at excessive intracellular sugar concentrations, because the acetylated sugars can easily diffuse out of the cell, but are not actively transported back (Boos *et al.*, 1981; Roderick, 2005). Both enzymes belong to the family of hexapeptide repeat proteins (Vuorio *et al.*, 1994). This family comprises many acetyltransferases, and they all share a characteristic structural architecture, a large left-handed parallel β-helix that is built from repetitive hexapeptide motifs (Figure 28).

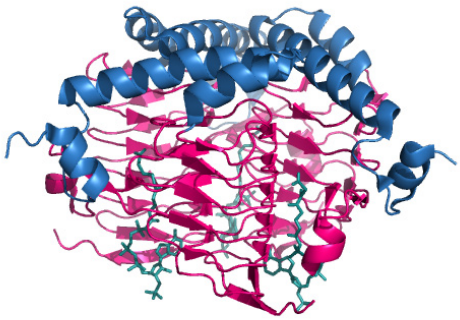


Figure 28 - Maltose *O*-acetyltransferase from *B. anthracis* in complex with acetyl-CoA (pdb-code: 3igj)
Maltose *O*-acetyltransferase (MAT) is shown as a ribbon diagram. The β-helix domain is depicted in magenta, acetyl-CoA in cyan, the N-terminal section in blue.

The β-helix domain mediates the formation of a trimer, and three molecules of the substrate acetyl-CoA can be bound at the contact interfaces between the protomers. The sequence identity between YvoF and MAT is 56.9 % within a C-terminal sequence of approx. 50 amino acids (total sequence length of YvoF: 172) that forms the central part of the β-helix domain, including the (putative) acetyl-CoA binding site. This is why it can be supposed that YvoF has the same basic structural features like MAT. However, the N-terminal section of YvoF differs significantly in sequence from known acetyltransferases

and might provide the specificity for polyprenylglyceryl substrates, perhaps by controlling the cellular localization of the enzyme.

MAT cannot complement the YvoF function *in vivo*. The *ΔyvoF* strains have intact *maa* genes, but nevertheless they are deficient in HepG acetylation (Figure 17). However, this might be rather due to a different cellular localization of YvoF and MAT than due to different substrate ranges of the two enzymes. It was shown that YvoF is somehow membrane-associated (Figure 19) like its substrate HepG, whereas MAT is likely to be not.

6.3.3.1 **Heterologous expression and purification of bs*maa***

To express bs*maa* heterologously in *E. coli* cells, the gene was amplified from genomic DNA of *B. subtilis* (5.3.2) and cloned into the pET21a vector and into the pET28atrxN vector via *Nde*I/*Xho*I restriction sites. This allows for the expression of bsMAT with a C-terminal $(His)_6$ tag and additionally a N-terminal thioredoxin in case of pET28atrxN (5.3.8). *E. coli* T7 Express cells were transformed with pET21a_bs*maa* and pET28atrxN_bs*maa* and gene expression was induced by IPTG as described in chapter 5.4.1.1. The protein was enriched with His SpinTrap™ columns (GE HEALTHCARE, 5.4.2.1), and its purity was analyzed by SDS-PAGE (5.6.3). The fusion of TrxA to bsMAT led to a higher amount of soluble protein compared to expression of pET21a_bs*maa* (data not shown). Therefore, preparative scale expression of pET28atrxN_bs*maa* was performed at 37 °C with 0.5 mM IPTG. The fusion protein was enriched by metal chelate affinity chromatography (5.4.2.2), and the purity was estimated to be > 95 % by SDS-PAGE (5.6.3, Figure 29).

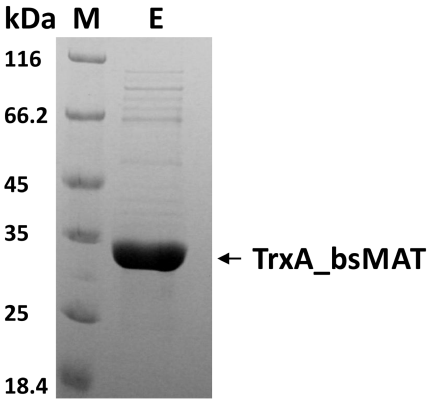


Figure 29 - Purity of TrxA_bsMAT
SDS-PAGE analysis (12.5 % polyacrylamide) of the purified TrxA_bsMAT. Size marker (M), molecular weight in kDa, E: eluate. TrxA_bsMAT is marked by an arrow.

Monomeric bsMAT fused to TrxA consists of 378 residues and has a molecular weight of 33.5 kDa, which is consistent with the size determined by SDS-PAGE. The fusion of TrxA to bsMAT afforded protein yields of about 9 mg/l culture.

To test whether TrxA can be removed by thrombin digestion and whether the free bsMAT is still soluble after cleavage, 20 U thrombin per mg protein were added to the purified protein and incubated over night at 4 °C. After centrifugation the cleavage was analyzed by SDS-PAGE (5.6.3, Figure 30).

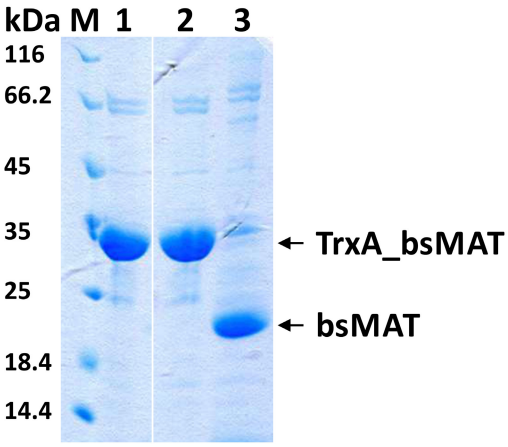


Figure 30 - Cleavage of the TrxA_bsMAT fusion protein by thrombin
SDS-PAGE analysis (12.5 % polyacrylamide) of thrombin cleavage of TrxA_bsMAT. Size marker (M), molecular weights in kDa. TrxA-bsMAT before thrombin cleavage as reference (1); as background reference no thrombin was added to TrxA_bsMAT at 4 °C over night (2); 20 units thrombin per mg protein were added to TrxA_bsMAT at 4 °C over night (3). TrxA_bsMAT and cleaved bsMAT are marked by arrows.

About 90 % cleavage of could be obtained. As bsYvoF could not be quantitatively cleaved from TrxA (Figure 23), like for bsYvoF the fusion protein TrxA-bsMAT was used in subsequent experiments. TrxA-bsMAT will be referred to as bsMAT in the following text.

6.3.3.2 *In vitro* activity of bsMAT

The overlap of substrate acceptance of bsYvoF and bsMAT was tested qualitatively *in vitro*, using GGG and maltose as acetyl acceptors (Figure 31).

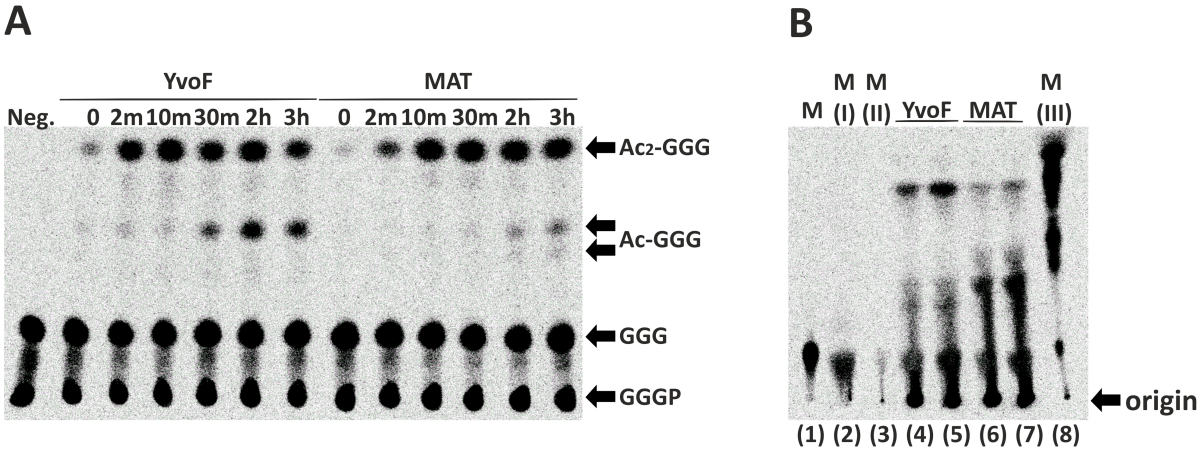


Figure 31 - Overlap in substrate acceptance of bsYvoF and bsMAT
(**A**) GGG acetyltransferase activity of bsYvoF and bsMAT. Same amounts of protein (5 µM) were incubated with the substrates ^{14}C-GGG and 250 µM acetyl-CoA for different times at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. As negative control (Neg.), the substrate ^{14}C-GGG was incubated solely with acetyl-CoA without enzyme. The origin of chromatography (marked by residual GGGP from GGG production) and the product spots are marked by arrows. (**B**) Maltose acetyltransferase activity of bsYvoF and bsMAT. Same amounts of protein (4 µM) were incubated with two different concentrations of the substrate ^{14}C-maltose (1.25 µM, lanes 4, 6; and 2.5 µM, lanes 5, 7) overnight at 40°C. The generated products were separated by thin layer chromatography, and visualized by autoradiography. M, ^{14}C-maltose without added enzyme and acetyl-CoA; M(I), ^{14}C-maltose in 50 mM potassium phosphate, no enzyme; M(II), ^{14}C-maltose in 50 mM potassium phosphate, 250 µM acetyl-CoA, no enzyme; M(III), ^{14}C-maltose that has been completely acetylated by incubation with acetic anhydride. The origin of chromatography is marked by an arrow. The figure was taken from Linde *et al.* (2016).

It is obvious that each enzyme accepts both substrates, but with different efficiencies and product specificities. The polyprenylglyceryl acetyltransferase activity of bsMAT is somewhat lower than that of bsYvoF under the applied specific conditions (Figure 31*A*). In contrast, the maltose acetyltransferase activity of bsYvoF is about the same as that of bsMAT, but the product spots obviously differ (Figure 31*B*). Without reference substances or detailed analysis, it is impossible to identify them specifically, but it can be assumed that they represent maltose derivates that are (multiply) acetylated at different hydroxyl groups. MAT is reported to acetylate maltose exclusively at the C_6 position of the nonreducing glucose moiety (Lo Leggio *et al.*, 2003). Although these experiments only provide qualitative results, they allow two conclusions: First, YvoF and MAT have only partial overlapping activities and therefore certainly serve different cellular functions. Second, because MAT has significant polyprenylglyceryl acetyltransferase activity *in vitro*, but cannot complement the *yvoF* knockout *in vivo*, it is very likely that YvoF acts in a membrane associated manner, as indicated by the previous experiments.

6.3.3.3 **Oligomerization state of bsYvoF and bsMAT**

As already discussed above, maltose *O*-acetyltransferases commonly forms trimers, which allows for the binding of three molecules of the substrate acetyl-CoA at the contact interfaces between the protomers (Figure 28). To verify this oligomerization state also for bsYvoF, analytical size exclusion chromatography was performed using a S75 column (GE HEALTHCARE) for both enzymes (5.6.5). Calibration curves used for the analysis of the results are shown in chapter 9.1.

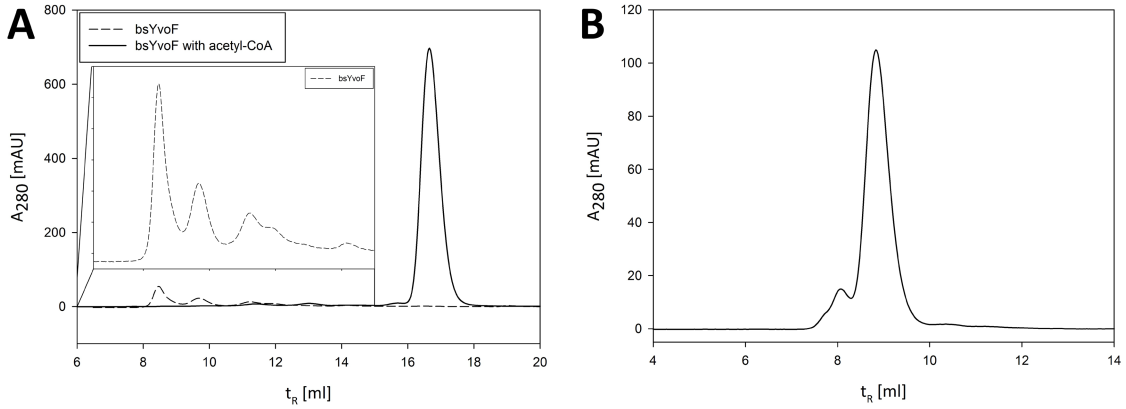


Figure 32 – Analytical size exclusion chromatography of bsYvoF and bsMAT
The proteins were applied to a S75 analytical column equilibrated with 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 5 mM $MgCl_2$ for bsYvoF and with 50 mM KP pH 7.5, 300 mM KCl for bsMAT. Elution was performed at a flow rate of 0.5 ml/min at room temperature, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). (**A**) Analytical SEC of bsYvoF (17 µM, subunit concentration) without acetyl-CoA (dashed line and inlet (Y-axis about 10x zoomed)) and with acetyl-CoA (44 µM bsYvoF, subunit concentration and 0.25 mM acetyl-CoA, solid line). (**B**) Analytical SEC of bsMAT (35 µM, subunit concentration).

BsYvoF without acetyl-CoA could not be concentrated higher than 17 µM, resulting in very low intensities of absorbance at 280 nm (Figure 32*A*). Two main distinct peaks were obtained, one peak at the exclusion volume, maybe referring to soluble aggregates, the other peak at an apparent molecular weight of about 90 kDa, indicating a trimeric oligomerization state (Figure 32*A,* dashed line and inlet). As the second peak had only about 30 % intensity compared to the first peak and the protein sample showed other impurity bands after concentrating, a clear determination of the oligomerization state of bsYvoF was not possible. To gain higher protein concentrations, acetyl-CoA was added to a final concentration of 0.25 mM. Acetyl-CoA stabilized bsYvoF during concentrating, resulting in 44 µM bsYvoF. The SEC profile of concentrated bsYvoF with acetyl-CoA resulted in a homogenous symmetric peak, but surprisingly at a very late retention

volume, corresponding to an apparent molecular weight of only 7 kDa (Figure 32*A*, solid line), which is difficult to explain. The high instability of bsYvoF and the fact that addition of acetyl-CoA somehow alters the overall protein structure resulting in abnormal behaviour on the SEC makes a clear determination of the oligomerization state of bsYvoF impossible. BsMAT eluted as homogenous symmetric peak from the analytical size exclusion column (Figure 32*B*), corresponding to an apparent molecular weight of about 80 kDa. This indicates a trimeric oligomerization state, as expected

6.3.3.4 **Structural integrity and thermal stability of bsYvoF and bsMAT**

To ensure the structural integrity and to determine the thermal stability of bsYvoF and bsMAT, far-UV circular dichroism (CD) was used (5.6.12, Figure 33).

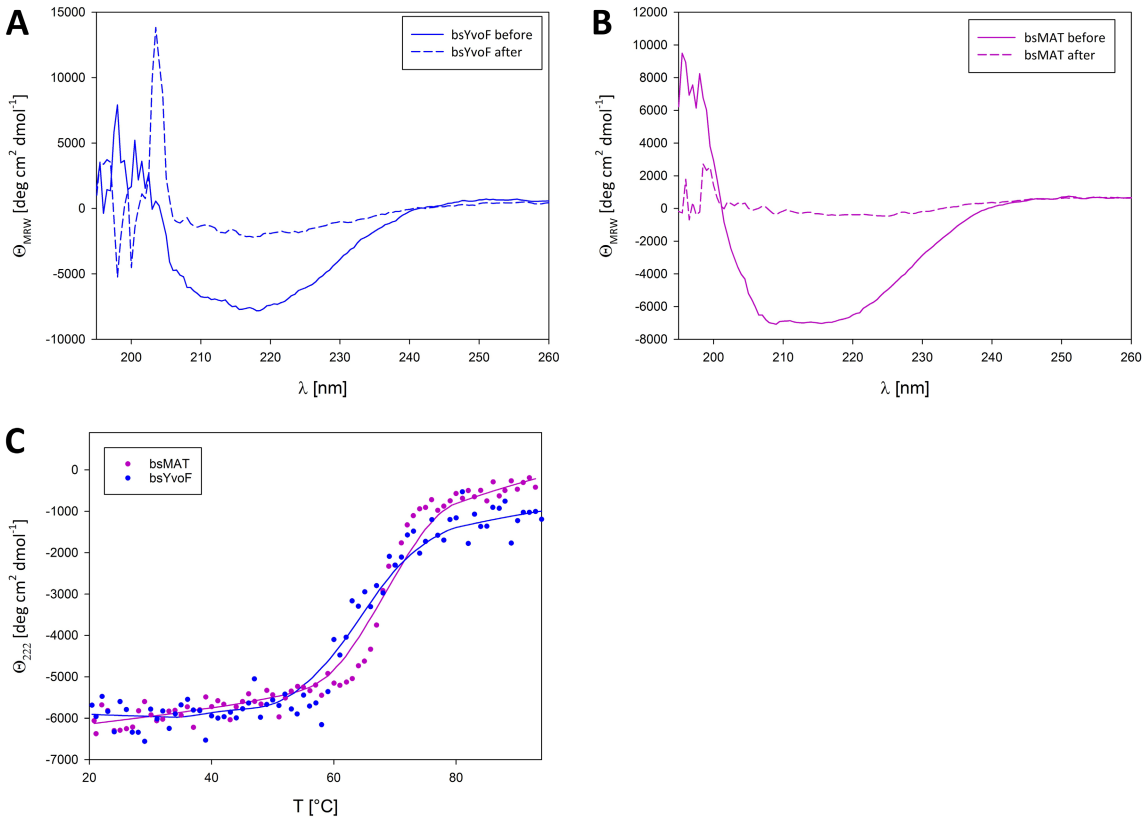


Figure 33 - Structural integrity and thermal stability of bsYvoF and bsMAT
Far-UV CD spectra were recorded from 195 nm to 260 nm (d = 1 mm) at room temperature before thermal heating (before, solid line) and immediately after heating to 95 °C (after, dashed line) for (**A**) 3.5 µM bsYvoF in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT and (**B**) 10 µM bsMAT in 50 mM KP pH 7.5 (**C**) Thermal denaturation was followed by CD spectroscopy. The loss of ellipticity at 222 nm of 3.5 µM bsYvoF in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT (data points: blue dots, smoothed: solid blue line) and of 10 µM bsMAT in 50 mM KP pH 7.5 (data points: magenta dots, smoothed: solid magenta line) was monitored from 25 °C to 95 °C at a scan rate of 1 °C/min (d = 1 mm). The LOESS algorithm was used to smooth the data sets.

Due to the high instability of bsYvoF and to avoid increasing the amount of background protein (cf. 6.3.3.3) it was refrained from concentrating the protein or removing the salts before CD meausurements, and only 3.5 µM protein were used. BsYvoF (Figure 33*A*) and bsMAT (Figure 33*B*) show CD spectra dominated by α-helical structures with the difference that bsYvoF lacks the typical second minimum of α-helices at 208 nm. This lets assume that bsYvoF exhibits less α-helices than bsMAT. Recalling the structure of bsMAT (Figure 28) and the assumption that bsYvoF differs significantly in the N-terminal section which is built of mainly α-helices in MAT (6.3.3), this spectoscopic result would support a different structural composition of bsYvoF in this area. The CD spectra of bsYvoF showed strong background noise below 200 nm. This could be due to the high amount of chloride ions coming from NaCl and $MgCl_2$ dissolved in the protein buffer as well as additional chloride ions due to pH adjustment of Tris buffer with HCl. Chloride ions absorb strongly below 200 nm (Kelly *et al.*, 2005). As the chloride ions did not disturb the CD signal above 200 nm, the thermal denaturation could be followed by CD at 222 nm. BsYvoF showed a cooperative thermal unfolding with a melting temperature (T_M) of 63.1 ± 0.02 °C and bsMAT showed a highly cooperative thermal unfolding with a melting temperature (T_M) of 68.39 ± 0.38 °C (Figure 33*C*).

6.3.4 Characterization of YvoF from *Bacillus anthracis*

To obtain higher amounts of soluble YvoF for further characterization, e.g. for size exclusion chromatography and crystallization, YvoF from *B. anthracis* was produced. *B. anthracis* exhibits 80.7 % protein sequence identity to bsYvoF (EMBOSS Matcher). The gen was amplified from genomic DNA (5.3.2) and cloned into the pET21a vector, into the pET28a vector, and into the pET28atrxN vector via *Nde*I/*Xho*I restriction sites to add a C-terminal $(His)_6$ tag and additionally an N-terminal thioredoxin in case of pET28atrxN to the recombinant protein (5.3.8). *E. coli* BL21-CodonPlus (DE3) RIPL cells were transformed with pET21a_ba*yvoF*, pET28_ba*yvoF*, or pET28atrxN_ba*yvoF*, and gene expression was induced by IPTG for 6 h or over night as described in chapter 5.4.1.1. The protein was either enriched from cells in resuspension buffer containing 9 mM CHAPS without sonication, identical to bsYvoF purification, or resuspended in 50 mM KP pH 7.5, 300 mM KCl, 10 mM imidazole and sonication. Purification was performed with His SpinTrap™ columns (GE HEALTHCARE, 5.4.2.1) and the expression level was analyzed by SDS-PAGE

and Western Blotting (5.6.3, 5.6.4, 9.2, Figure S1). The purification with sonication exclusively resulted in insoluble expression (pellet fraction) in all vector systems. The purification with 9 mM CHAPS and without sonication, however, led to soluble protein with all used vectors, except the expression of pET28a_ba*yvoF* over night, which did not afford any protein. As the test expressions did not show any improvement of *yvoF* expression compared to bsYvoF, no further expression at preparative scale was performed. To verify the acetyltransferase activity of baYvoF, *in vitro* activity assays were performed (5.6.14.4, Figure 34).

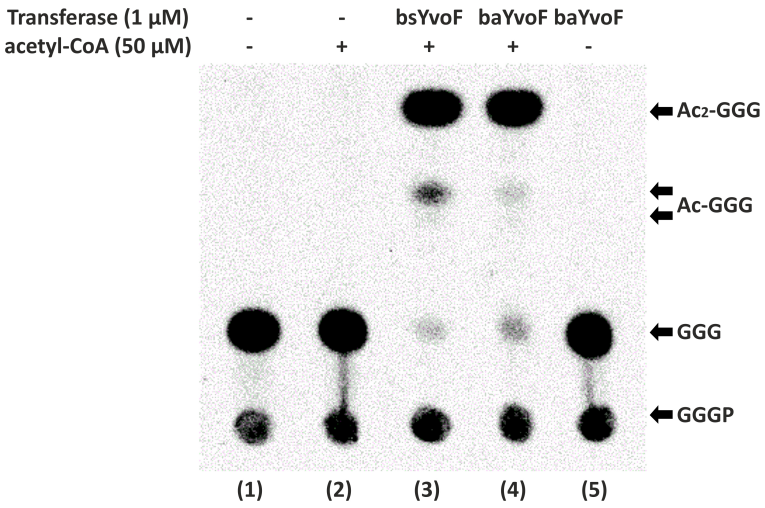


Figure 34 - Activity of baYvoF
BaYvoF and bsYvoF (positive control) were incubated with ^{14}C-GGG for 2 h at 40 °C. The products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Lane 1: ^{14}C-GGG without enzyme (negative control); lane 2: ^{14}C-GGG without enzyme plus 50 µM acetyl-CoA (negative control); lane 3: ^{14}C-GGG plus purified bsYvoF plus 50 µM acetyl-CoA; lane 4: ^{14}C-GGG plus purified baYvoF plus 50 µM acetyl-CoA; lane 5: ^{14}C-GGG plus purified baYvoF without acetyl-CoA. The origin of chromatography (marked by residual GGGP from GGG production) and the product spots are marked by arrows.

baYvoF showed the same activity pattern as bsYvoF. Only after addition of acetyl-CoA, an acetylation reaction could be detected. Even if the expression of a *yvoF* gene from another organism did not result in a higher yield of soluble protein, the acetylation activity of YvoF could be confirmed.

6.3.5 Occurrence of YvoF-like enzymes

Last, the species distribution of YvoF enzymes was analyzed. The NCBI refseq database was searched using BLAST and *B. subtilis* YvoF as query and the species distribution was visualized. Only hits better than an E-value threshold that was set by the best hit with a "maltose acetyltransferase" annotation were included (754 hits). Just below the threshold, a rapid drop of E-values occurred, and the hits did not cover the whole sequence length of YvoF anymore. This proves that the selected sequences represent almost all YvoF orthologues that are known today. Most hits (95 %) were among Bacillales (Bacillaceae, Listeriaceae and Staphylococcaceae) and Clostridiales. This means that the species distribution of YvoF exactly coincides with that of PcrB, the enzyme that produces HepGP (Peterhoff *et al.*, 2014) and supports that both enzymes in fact belong to a common, biologically relevant pathway. Interestingly, there was a significant number of additional hits among the archaeal Halobacteria, the best ones with 44 % sequence identity over the complete sequence length to *B. subtilis* YvoF, and it will be a challenging task to elucidate their function.

6.4 Conclusion

In Archaea, ether lipids play an essential role as the main building blocks of the cellular membrane. Recently, ether lipids have also been discovered in the domain of Bacteria, and the key enzymes that catalyze their synthesis, G1PDH and HepGPS, have been analyzed in detail (Guldan, 2010; Peterhoff *et al.*, 2012). In Bacillales, HepGP does not become linked to a second polyprenyl moiety like ether lipids in Archaea, but is dephosphorylated and acetylated (Guldan *et al.*, 2011). The enzymes that catalyze these reactions have been identified and characterized within this thesis. The phosphatase activity for HepGP was enriched from a *B. subtilis* cell extract and the phosphatase PhoB was identified by mass spectrometry. The gene of bs*phoB* was amplified and heterologously expressed. The dephosphorylation activity of bsPhoB on GGGP could be verified *in vitro*. Nevertheless it can be assumed that any other phosphatase might catalyze dephosphorylation of HepGP in an unspecific manner as well. By screening a *B. subtilis* knockout library for deficiency in acetylation, the bs*yvoF* gene product was identified to be the acetyltransferase. The gene of bs*yvoF* could be amplified and the heterologously expressed bsYvoF could be characterized biochemically. Its *in vitro* activity

was verified and catalytic parameters could be obtained by using a DTNB-coupled activity assay as well as a discontinuous radiometric assay. The membrane assoziation of bsYvoF was determined by ultracentrifugation. The acetyl-CoA dependent enzyme YvoF is a close relative of maltose *O*-acetyltransferase (MAT). The activity of YvoF was compared to MAT. YvoF and MAT partially overlap in substrate and product range *in vitro*, but MAT is not able to complement the *yvoF* knockout *in vivo*. It remains to be clarified how this *in vivo* specialization is achieved, and one answer might be that the cellular localization of the two enzymes is different, as the YvoF substrate HepG is associated to the cellular membrane. Acetylated lipids and their function are a rather unexplored field. They play a role in activation of thrombozytes (Lee *et al.*, 1990; Lee *et al.*, 1986) or in the modification of lipids in industry, but their function in Bacteria is still an enigma. However, the identification of two enzymes involved in their biosynthesis pathway brings us closer to answer this intriguing question.

6.5 Ongoing research and future work

Current investigations focus on solving the biological function of G1P-based ether lipids in Bacteria. Phentoype assays have been established (Seisenberger, 2017) and different conditions will be tested on *B. subtilis* wild type, *B. subtilis* Δ*pcrB*, and *B. subtilis* Δ*yvoF* to elucidate their role. Possible effects of these ether lipids on the membrane fluidity or membrane composition under different physiological conditions should be investigated. The localization and orientation of HepG or its acetylated products in the membrane of *B. subtilis* is still unkown. This could maybe be elucidated with immunological techniques or electron microscopy. To get a better insight into the molecular details underlying the binding of the long hydrophobic polyprenylglycerol as substrate instead of maltose or other sugars, the purification of stable and highly concentrated YvoF has to be optimized. Therefore different *yvoF* genes from different organsisms have already been tested in small scale expression in *E. coli*, but this did not result in better yields of soluble protein (Seisenberger, 2017). As the analysis of occurrence of YvoF-like enzymes in other organisms lead to the knowledge that also Halobacteria exhibit *yvoF* genes, it was obvious to test expression of *yvoF* in halobacterial strains. Therefore *Haloverax volcanii* (hv) was chosen as expression system for hv*yvoF*. The expression and purification was successful and resulted in high protein concentrations (Seisenberger, 2017). Initial screening yielded already crystals, which have now to be optimized. The crystallization of YvoF and the investigation of phenotype assays are a big step closer to answer the question which role G1P-based ether lipids play in Bacteria. It is also very interesting which role the *yvoF* genes play in Halobacteria. Therefore *yvoF* in *Haloferax volcanii* should get deleted and be subjected to a phenotype screening. Also *in vivo* labeling experiments should be performed in these organisms to analyze the function of YvoF in these species. Do these Halobacteria contain acetylated GGG in their membranes or does YvoF play another role in these species? Payandeh and Pai (2007) made an interesting statement in their paper by imagining the reconstruction of the whole archaeal lipid biosynthesis apparatus, in particular G1PDH, GGGPS, DGGGPS and DGGGPR in Bacteria containing an knockout of the endogenous lipid production, maybe resulting in a bacterium with mainly archaeal ether lipids in the membrane. The prospect of engineering *E. coli* with enzymes involved in the biosynthesis of archaeal lipids has been already initiated in several studies (Caforio

et al., 2015; Jain *et al.*, 2014). Taking up this idea it would be quite interesting to spread this onto *B. subtilis*. Here only the expression of DGGGPS and DGGGPR would be necessary as the G1PDH (AraM) and the GGGPS-like enzyme (HepGPS, PcrB) are already present in the organism and *B. subtilis* synthesizes C_{15}, C_{20}, C_{35} and C_{55} *in vivo* (Kabisch *et al.*, 2013; Peterhoff *et al.*, 2014; Takahashi and Ogura, 1981). As DGGGPS and DGGGPR prefer the C_{20} product like archaeal GGGPS enzymes, these enzymes would have to be engineered for also accepting the longer polyprenyl chain C_{35}. After successful over-expression of the enzymes, it would be interesting whether the main compound of the membrane is built of archaeal ether lipids with double isoprenylated G1P or whether the dephosphorylation and subsequent acetylation of the single prenylated G1P is favored. It would be equally interesting to elucidate the difference between membranes containing C_{20} diether lipids compared to membranes containing C_{35} diether lipids. For this purpose, the heptaprenyl pyrophosphate synthase (HepPPS) would need to be knocked out in *B. subtilis* and the native DGGGPS and DGGGPR with preferred substrate acceptance for C_{20} would have be to expressed such that mainly C_{20}-based compounds get synthesized leading to membranes containing mainly C_{20} diether lipids. Then these membranes should be analyzed regarding their membrane fluidity and phenotypes.

7 THE GGGPS ENZYME FAMILY: HEXAMERIZATION ENSURES STRUCTURAL INTEGRITY AS A PREREQUISITE FOR THERMAL STABILITY AND ACTIVITY

7.1 Preface

Dr. Chitra Rajendran collected the X-ray data and solved the structure of mtGGGPS_A162E, mtGGGPS_W141A and mtGGGPS_I107E. MD simulations on a mtGGGPS subunit have been performed by Kristina Heyn, AG Prof. R. Merkl, University of Regensburg. This chapter also contains data from my bachelor student Whitney Kilu (Kilu, 2014). Details are indicated in the following text. This data was included in this work to draw a complete picture on the hexamerization of GGGPS from *Methanothermobacter thermautotrophicus*.

7.2 Objective of this thesis part

Due to the special interest of our group in enzymes with a $(\beta\alpha)_8$-barrel fold, an ancient enzyme family, the geranylgeranylglyceryl phosphate synthases (GGGPS), came into the focus of our research. Combining biochemical and computational methods, a comprehensive analysis for a systematic and comparative characterization of the GGGPS enzyme family revealed two separate groups with less than 20 % sequence identity. During biochemical characterization of these two groups it became clear that all group I enzymes exhibit a dimeric oligomerization state, whereas in group II dimeric as well as hexameric oligomerization can be found. This work analyzes the impact of oligomerization for stability and activity by performing and analysing mutagenesis studies with the GGGPS from *Methanothermobacter thermautotrophicus*. Furthermore, a variety of native dimeric and hexameric GGGPS enzymes were characterized and compared regarding their thermal stability and activity.

7.3 Results and Discussion

7.3.1 Quaternary structure organization of the hexameric GGGPS from *Methanothermobacter thermautotrophicus*

To study the structural principles of the hexamerization of GGGPS enzymes, the GGGPS from *Methanothermobacter thermautotrophicus* (mtGGGPS) was selected. The hexamer of mtGGGPS consists of three dimeric modules (Figure 35*A*). Two subunits interact via the prototypic dimeric interface (in the following called symmetric interface 1) to build a dimeric module (Figure 35*A*). Via two other interfaces, two additional dimeric modules connect with the first dimeric module to build the hexamer (Figure 35*A*, green and cyan dimer). This means, the quaternary structure of mtGGGPS is tied together by three different interfaces as described previously by Peterhoff *et al*. (2014; Figure 35*B*). Next to the symmetric interface 1 (Figure 35*B*, red plane) there are two other interfaces presumably important for the stabilization of the hexameric oligomerization. Two interacting surfaces from adjacent protomers build the symmetric interface 2 (Figure 35*B*, dashed grey line) and a third interface is built by two asymmetric interacting surfaces from adjacent protomers (asymmetric interface 3a/3b, Figure 35*B*, solid black line).

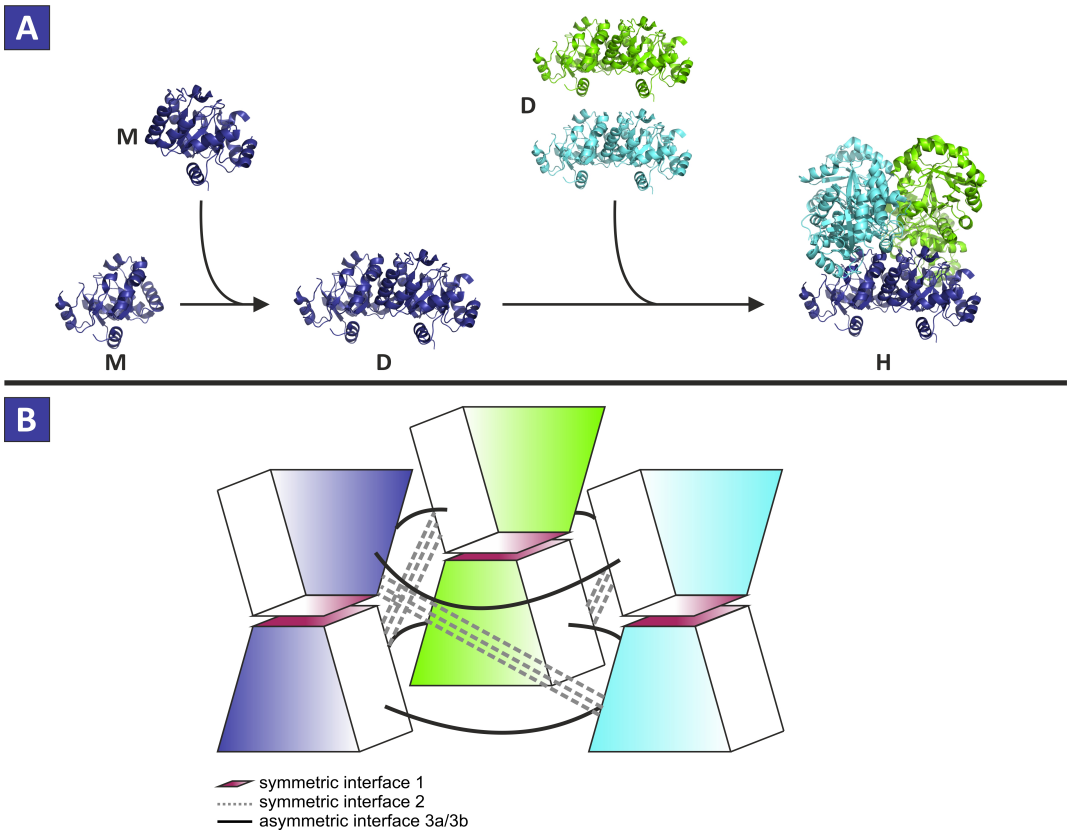


Figure 35 - Constitution and oligomerization interfaces of the hexameric mtGGGPS
(**A**) Constitution of mtGGGPS (pdb-code: 4mm1). Two subunits build a dimeric module in the prototypic configuration of the native dimers, depicted in deepblue. Three of these dimeric modules (deepblue, green, cyan) build one hexamer. M: monomer, D: dimer, H: hexamer (**B**) Schematic depiction of the constitution of hexameric mtGGGPS. The associated subunits of each dimeric module are depicted in the same color. The symmetric interface 1 is depicted in magenta. The associations of symmetric interface 2 are depicted as dashed grey lines and the associations of asymmetric interface 3a/3b as solid black line.

While three interfaces 3a/3b link three protomers (one from each dimeric module) to a ring-like structure, the interfaces 1 and 2 connect the two 3-mer rings with each other (Figure 35*B*).

The aim of this study was to intensely probe residues at the interfaces with respect to their impact for oligomerization, and to figure out what the hexameric structure is good for. For this purpose, mutational studies were performed and the oligomerization state was analyzed by analytical size exclusion chromatography (SEC). Afterwards, the presumably stabilizing effect of the hexamer compared to the mutants with a lower oligomerization state was studied with thermal denaturation methods (circular dichroism (CD), differential scanning calorimetry (DSC), nano differential scanning fluorimetry (nanoDSF)), and irreversible heat inactivation.

7.3.1.1 Heterologous expression and purification of GGGPS enzymes

The genes of ten GGGPS enzymes were already amplified by Peterhoff *et al*. (2014; 9.3, Table S1). Twelve mutants of *Methanothermobacter thermautotrophicus* GGGPS were constructed partly in the bachelor thesis of Whitney Kilu (2014) and partly in this thesis (Table S1). The genes were cloned into the expression vector pET21a via the *Nde*I/*Xho*I restriction sites, providing recombinant proteins with a C-terminal $(His)_6$ tag (5.3.8). Mutants were generated by QuickChange Mutagenesis (5.3.4). All cloned genes were verified by sequencing (5.3.9). *E. coli* BL21-CodonPlus (DE3) RIPL cells, or in case of zpGGGPS, *E. coli* T7 Express cells were transformed with the pET21a-based constructs and gene expression was IPTG-induced over night at 37 °C (5.4.1.2). The proteins were enriched by metal chelate affinity chromatography (5.4.2.2), and the purity was estimated to be > 95 % by SDS-PAGE (5.6.3, Figure 15).

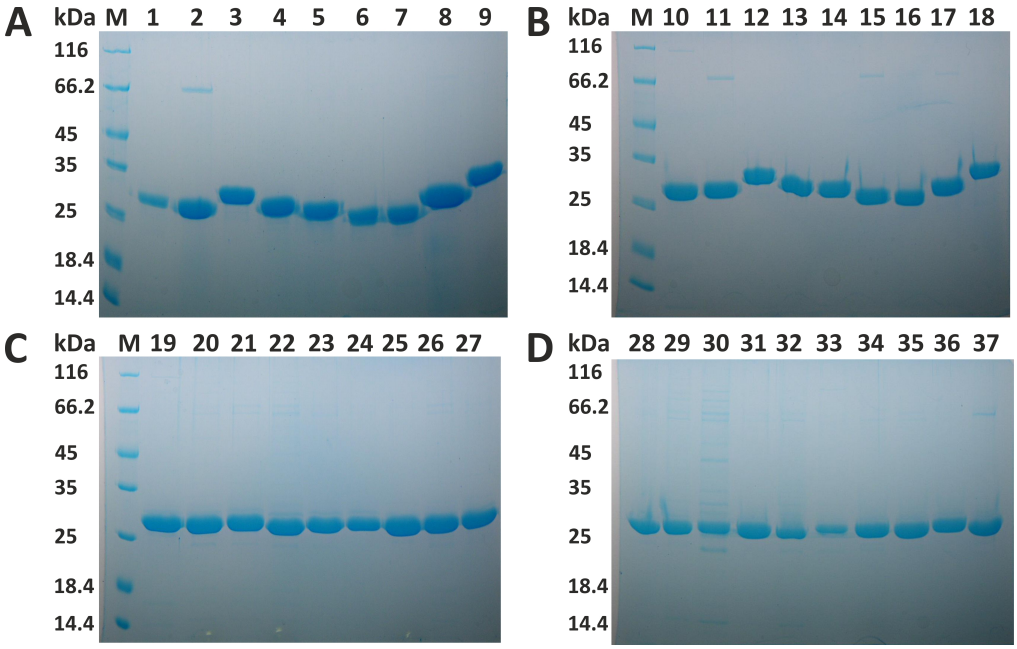


Figure 36 - Purity of all GGGPS wild type and mutant enzymes characterized in this thesis
SDS-PAGE analysis (12.5 % polyacrylamide) of the purified GGGPS enzymes (13.8 µg protein each). Size marker (M), molecular weight in kDa. (**A**) GGGPS enzymes purified in 50 mM Tris/HCl pH 8.0: (1) mtGGGPS_wt, (2) cpGGGPS_wt, (3) tkGGGPS_wt, (4) taGGGPS_wt, (5) fjGGGPS_wt, (6) slGGGPS_wt, (7) zpGGGPS_wt, (8) cpGGGPS_Y143A, (9) tkGGGPS_W143A. (**B**) GGGPS enzymes purified in 50 mM potassium phosphate pH 7.5: (10) mtGGGPS_wt, (11) cpGGGPS_wt, (12) tkGGGPS_wt, (13) taGGGPS_wt, (14) fjGGGPS_wt, (15) slGGGPS_wt, (16) zpGGGPS_wt, (17) cpGGGPS_Y143A, (18) tkGGGPS_W143A. (**C**) GGGPS enzymes purified in 50 mM potassium phosphate pH 7.5: (19) mtGGGPS_wt, (20) mtGGGPS_A162E, (21) mtGGGPS_R88E, (22) mtGGGPS_D57R, (23) mtGGGPS_I155A, (24) mtGGGPS_K146A, (25) mtGGGPS_W141A, (26) mtGGGPS_Y105A, (27) mtGGGPS_I107E. (**D:28-32**) GGGPS enzymes purified in 50 mM potassium phosphate pH 7.5: (28) mtGGGPS_A162E_W141A, (29) mtGGGPS_Y105A, (30) mtGGGPS_I107E, (31) mtGGGPS_R88E, (32) mtGGGPS_D57R, (**D:33-37**) GGGPS enzymes purified in 50 mM Tris/HCl pH 8.0: (33) mtGGGPS_wt, (34) mtGGGPS_A162E, (35) mtGGGPS_W141A, (36) mtGGGPS_I107E, (37) mtGGGPS_A162E_W141A.

The molecular weight of one subunit of GGGPS enzymes is about 27 kDa, which is consistent with their size determined by SDS-PAGE. Protein yields of about 3 – 90 mg/l culture could be obtained (Table S1).

7.3.1.2 Mutational analysis of the asymmetric interface 3a/3b

Residues at the asymmetric interface 3a/3b were examined first. The main part of the small interface 3a/3b is formed by the residues W141, D144, K146, and a hydrophobic cleft which is formed by small hydrophobic amino acids, mainly I155 (Figure 37). The interface residues from the interacting surface 3a are localized in α-helix 5´ (W141) and in the loop region α5´α5 (D144) and are in contact with residues from the interacting surface 3b of the opposite protomer, which is composed of K146 and I155 laying in the loop region α5´α5 and at the beginning of α-helix 5, respectively. During the study of the GGGPS enzyme family, our group has shown that the W141 localized in the asymmetric interface 3a/3b is essential for the stabilization of the hexameric enzymes. If it is mutated to an alanine (mtGGGPS_W141A), the hexamer collapses into its dimeric modules (Peterhoff *et al.*, 2014). A bioinformatic analysis (Patrick Löffler, AG Prof. R. Merkl, University of Regensburg) was done in the bachelor thesis of Whitney Kilu (Kilu, 2014) and supported the hypothesis of Peterhoff *et al.* (2014) that W141 forms a cation-π interaction with K146 of the opposite protomer (Figure 37).

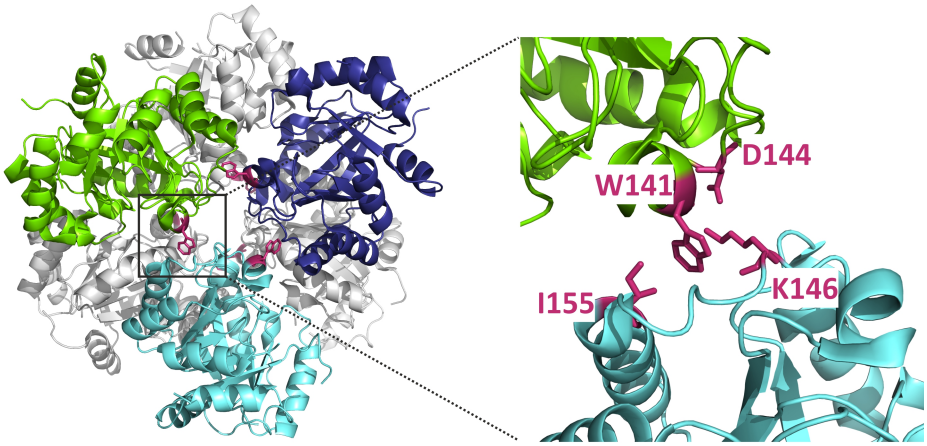


Figure 37 –The asymmetric interface 3a/3b of mtGGGPS
The interaction between the interacting surfaces 3a and 3b of two adjacent protomers (green and cyan) of mtGGGPS (pdb-code: 4mm1) is shown in survey (left side) and in detail (right side). To distinguish between the associated subunits of each dimeric module, one subunit is shown in color, the other in grey (left side). Important interface residues of interface 3a/3b are depicted as sticks in magenta (right side).

A cation-π interaction is an electrostatic attraction between a cation or rather a positively charged side-group (here lysine) and delocalized π-electrons (tryptophan). K146 additionally shows an ionic interaction to D144, whereas W141 extends into a hydrophobic cleft (Peterhoff *et al.*, 2014). The residues D144, K146 and I155 were exchanged to alanine in oder to interrupt possible interactions, and the oligomerization state of the mutants was analyzed by analytical SEC (5.6.5, Figure 38). These mutations have already been analyzed previously (Kilu, 2014).

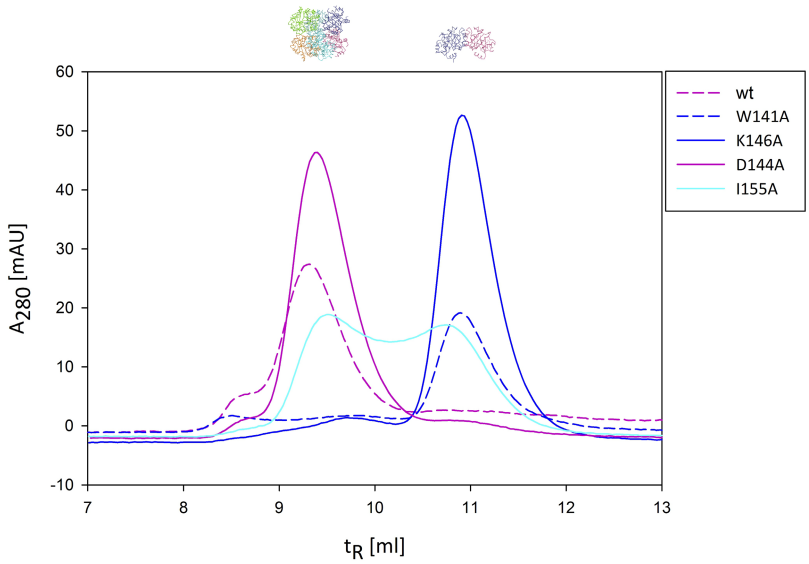


Figure 38 – Analytical size exclusion chromatography of mtGGGPS mutants with amino acid exchanges at the asymmetric interface 3a/3b
The proteins (40 µM, subunit concentration) were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta) and mtGGGPS_W141A (W141A, dashed lines, blue) are shown as references for the hexameric and dimeric oligomerization states. mtGGGPS_K146A (K146A, blue), mtGGGPS_D144A (D144A, magenta) and mtGGGPS_I155A (I155A, cyan) are drawn as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

MtGGGPS_wt and mtGGGPS_W141A served as references in the SEC experiments, because their oligomerization state has already been determined previously by SEC coupled with static light scattering (SLS; Peterhoff *et al.*, 2014). It was observable that the mutation K146A, which deletes the proposed cation-π interaction between W141 and K146, leads to dimeric proteins (Figure 38). This was not surprising as the mutation W141A leads to the same effect. To confirm that the dimerization is due to the loss of the cation-π-interaction and not to the loss of the ionic interaction between K146 and D144, D144 was separately mutated to alanine. This mutation did not lead to a collpase of the

hexameric conformation (Figure 38). This clearly shows that the interface 3a/3b is mainly stabilized by the interaction between W141 and K146. As W141 extends into a hydrophobic pocket of the opposite protomer, which is mainly formed by I155, this amino acid was also mutated to elucidate the importance of the hydrophobic pocket for hexamerization. Interestingly, the mtGGGPS_I155A mutant obviously falls into an association-dissociation equilibrium between the dimeric and hexameric state under the given experimental conditions (Figure 38). Analytical SEC of mtGGGPS_I155A with varying concentrations approved a concentration-dependent association-dissociation equilibrium (5.6.5, Figure 39).

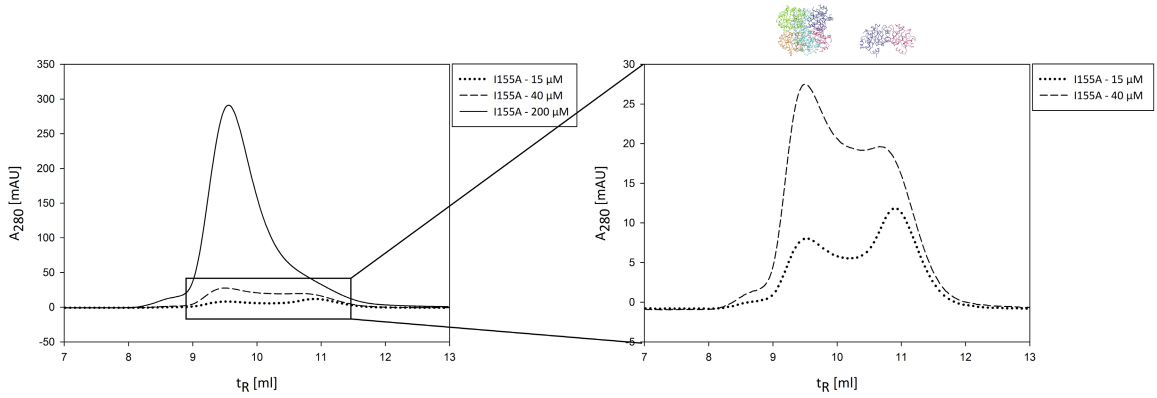


Figure 39 - Analytical size exclusion chromatography with varying concentrations of mtGGGPS_I155A
The mtGGGPS_I155A was applied at three different subunit concentrations to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). The lowest concentration of 15 µM is depicted as dotted lines, the medium concentration of 40 µM as dashed lines and the highest concentration of 200 µM as solid lines. The SEC profile of the two lower concentrations was enlarged to give a better insight and is shown on the right side of the figure. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

Whereas a low concentration led to an enrichment of the dimeric protein with a ratio of about 2:3 (hexamer:dimer, Figure 39, 15 µM), higher concentrations yielded higher fraction of the hexameric oligomerization state. Already an almost 3 fold increase in protein concentration (Figure 39, 40 µM) resulted in an inverted ratio of 3:2 (hexamer:dimer), whereas a concentration of 200 µM (Figure 39) afforded mainly hexameric proteins. Nevertheless, the peak at 200 µM concentration was not symmetric, but asymmetric with a shoulder due to small amount of dimers. It can be assumed that the hydrophobic interactions between I155 and the tryptophan anchor W141 are important for the exact localization of the tryptophan and consequently help to stabilize the hexamer, but are not as important as the cation-π-interaction with K146. Therefore,

the effect of the mutation I155A on the oligomerization state is not as drastic as that of K146A or W141A.

7.3.1.3 Mutational analysis of the symmetric interface 1 (dimeric interface)

Three interfaces 3a/3b connect one subunit of each of the dimeric modules resulting in two "rings" (Figure 35*B*). The rings are not superposable, but staggered above each other (Figure 37). This leads to the interesting question whether the destabilization of the symmetric interface 1 results in trimers, which would be expected if the symmetric interface 2 were loose. A mutational analysis of this interface will be presented later (7.3.1.4). Therefore amino acids had to be found, which are important for the symmetric interface 1. Such residues have already been identified previously for the dimeric bsPcrB (Peterhoff *et al.*, 2012). Therefore, the structure of mtGGGPS (pdb-code: 4mm1) was superimposed with bsPcrB (pdb-code: 1viz; Kilu 2013). Three equivalent amino acids could be found: bsPcrB_I93 superimposed with mtGGGPS_I107, bsPcrB_V148 with mtGGGPS_A162 and bsPcrB_L152 with mtGGGPS_M168 (Figure 40).

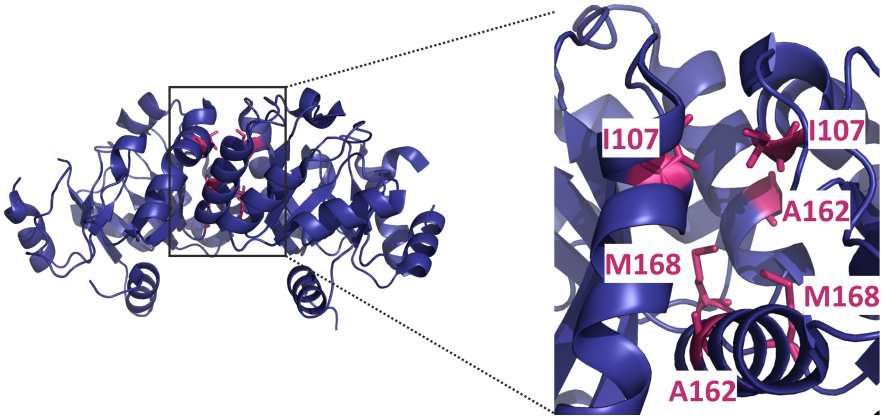


Figure 40 - The symmetric interface 1 of mtGGGPS
Superimposing of the structure of mtGGGPS (pdb-code: 4mm1) with bsPcrB (pdb-code: 1viz) in PyMol revealed equivalent interface residues within symmetric interface 1, depicted as sticks in magenta. The symmetric interface 1 is shown in survey (left side) and in detail (right side). For clarity only one dimeric module of the hexameric mtGGGPS is shown.

I107 is localized at the upper part of the symmetric interface 1 (at the beginning of α-helix 4) and near the catalytic face of the $(\beta\alpha)_8$-barrel in close proximity to the symmetric interface 2. A162 is localized in the center of the symmetric interface 1 (in the middle of α-helix 4) and M168 at the lower part of the interface at the loop region α5β6. The three amino acids were mutated independently to glutamate in the hexameric mtGGGPS_wt as well as in the dimeric mtGGGPS_W141A to elucidate their effect on the symmetric

interface 1. Glutamate was chosen to introduce an electrostatic repulsion between the two adjacent protomers of one dimeric module in order to disrupt the symmetric interface 1. The oligomerization state of the mutants was analyzed by analytical SEC (5.6.5). While M168E did not show any effect at all and was therefore not further characterized (data not shown), the mutation A162E was effective (Figure 41).

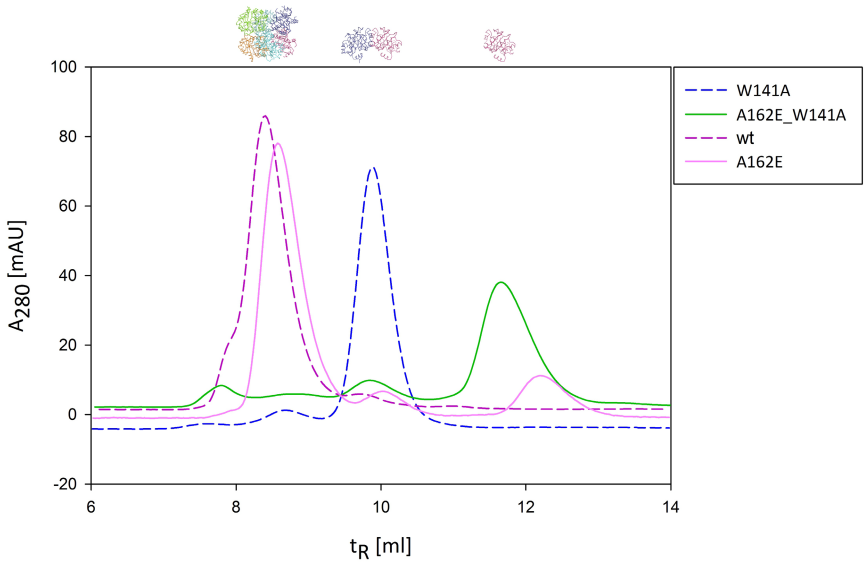


Figure 41 - Analytical size exclusion chromatography of mtGGGPS mutants with amino acid exchanges at the core of the symmetric interface 1
The proteins (40 µM, subunit concentration) were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta) and mtGGGPS_W141A (W141A, dashed lines, blue) are shown as reference peaks for hexameric and dimeric oligomerization state. mtGGGPS_A162E (A162E, lightpink) and mtGGGPS_A162E_W141A (A162E_W141A, green) are drawn as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

The incorporation of the mutation A162E into the dimer (resulting in mtGGGPS_A162E_W141A) lead to a monomeric oligomerization state (Figure 41). The incorporation of this monomerizing mutation into the hexameric mtGGGPS_wt (resulting in mtGGGPS_A162E), however, did not lead to the expected collapse of the hexameric structure into trimers (Figure 41). Only a slight amount of monomeric and dimeric peak could be detected in the SEC profile of mtGGGPS_A162E besides the main hexameric peak. To analyze whether this is associated with a concentration-dependent association-dissociation equilibrium as it was the case for mtGGGPS_I155A, analytical SEC with varying concentrations of mtGGGPS_A162E was performed (5.6.5, Figure 42).

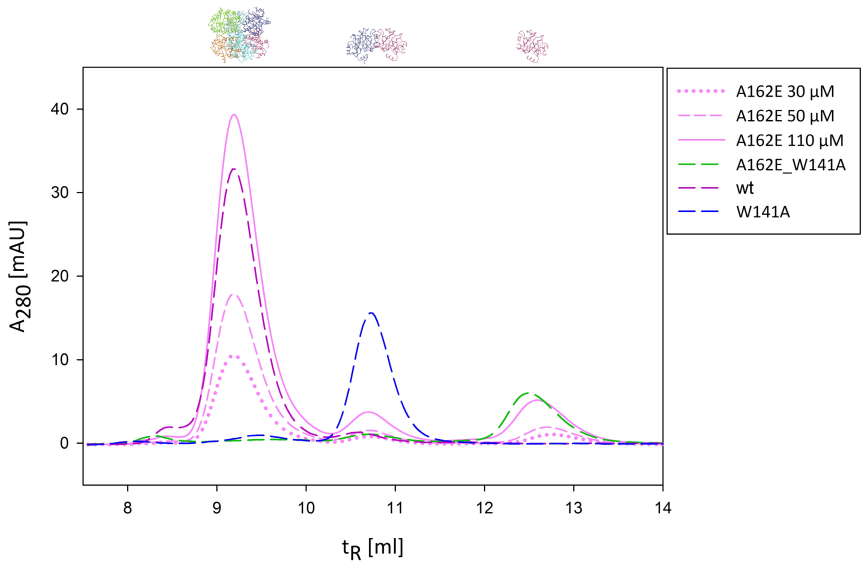


Figure 42 - Analytical size exclusion chromatography with varying concentrations of mtGGGPS_A162E
The proteins were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta, 40 µM, subunit concentration), mtGGGPS_W141A (W141A, dashed lines, blue, 40 µM, subunit concentration) and mtGGGPS_A162E_W141A (A162E_W141A, dashed lines, green, 33 µM) are shown as reference peaks for hexameric, dimeric and monomeric oligomerization state. Different concentrations of mtGGGPS_A162E (A162E) are colored lightpink. The lowest concentration of 30 µM (subunit concentration) is depicted as dotted lines, the middle concentration of 50 µM (subunit concentration) as dashed lines and the highest concentration of 110 µM (subunit concentration) as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

The ratio between the monomeric, dimeric and hexameric amount of protein in the different SEC profiles did not change, so it can be concluded that there is no rapid concentration-dependent association-dissociation equilibrium under the prevailing conditions.

The mutation I107E in the mtGGGPS_wt background resulted in a somewhat surprising effect as it did neither retain the hexameric state as seen for the A162E exchange nor cause its collapse into trimers. Instead, dimers are observed as it is the case for W141A and K146A (Figure 43). I107 definitely lies in the symmetric interface 1, but not in the center like A162, but more at the periphery and close to the symmetric interface 2. Therefore it can be speculated that the I107E exchange has a stronger effect on this interface than on the symmetric interface 1. It will be shown later that mutations at the symmetric interface 2 result in dimers retaining symmetric interface 1. In order to test whether this is also the case for the mtGGGPS_I107E mutant, the monomerizing mutation

A162E was introduced into mtGGGPS_I107E (resulting in mtGGGPS_A162E_I107E) and its oligomerization state was analyzed by analytical SEC (5.6.5, Figure 43).

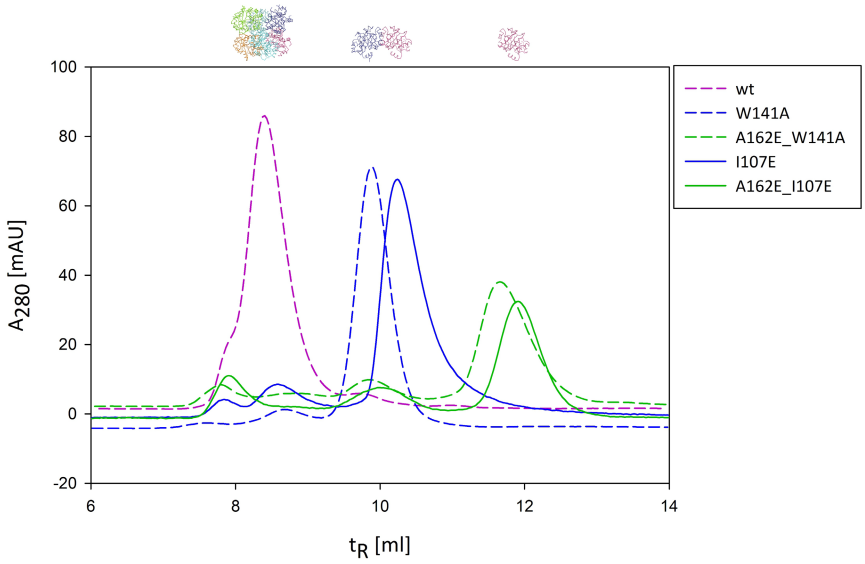


Figure 43 - Analytical size exclusion chromatography of mtGGGPS mutants with amino acid exchanges at the periphery of the symmetric interface 1 (dimeric interface)
The proteins were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta, 40 µM, subunit concentration), mtGGGPS_W141A (W141A, dashed lines, blue, 40 µM, subunit concentration) and mtGGGPS_A162E_W141A (A162E_W141A, dashed lines, green, 33 µM) are shown as reference peaks for hexameric, dimeric and monomeric oligomerization state. mtGGGPS_I107E (I107E, blue, 40 µM, subunit concentration) and mtGGGPS_A162E_I107E (A162E_I107E, green, 25 µM, subunit concentration) are drawn as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

The introduction of the monomerizing mutation A162E into the mtGGGPS_I107E mutant led to a collapse of the dimer into its monomeric subunits as it is the case for mtGGGPS_A162E_W141A. MtGGGPS_I107E as well as mtGGGPS_A162E_I107E showed somewhat shifted SEC profiles compared to the reference proteins. To analyze whether this shift is due to a concentration-dependent behavior, analytical SEC of mtGGGPS_I107E with varying concentrations was performed (5.6.5, Figure 44).

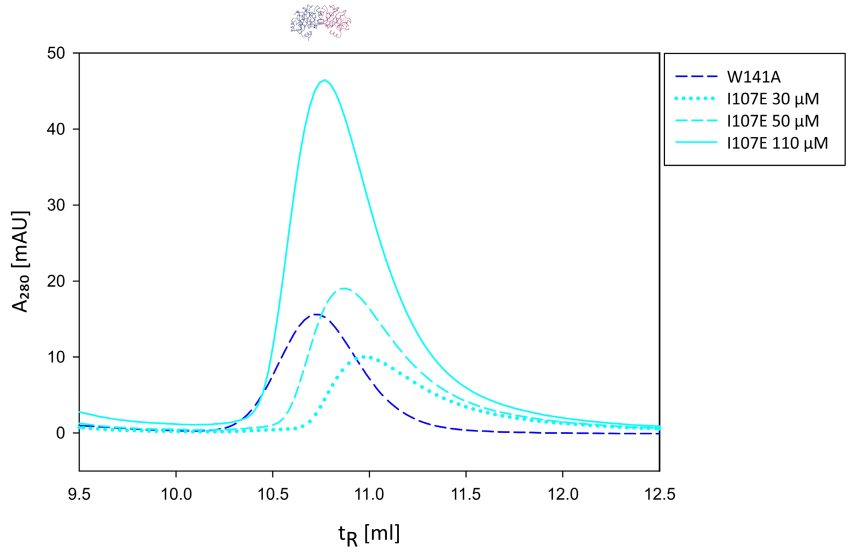


Figure 44 - Analytical size exclusion chromatography with varying concentrations of mtGGGPS_I107E
The proteins were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_W141A (W141A, dashed lines, blue, 40 µM, subunit concentration) is shown as reference peak for the dimeric oligomerization state. Different concentrations of mtGGGPS_I107E (I107E) are colored cyan. The lowest concentration of 30 µM (subunit concentration) is depicted as dotted lines, the middle concentration of 50 µM (subunit concentration) as dashed lines and the highest concentration of 110 µM (subunit concentration) as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

MtGGGPS_I107E clearly exhibits a concentration-dependent association-dissociation behavior. Whereas low concentrations shifted the equilibrium to lower molecular weights, compared to the dimeric reference protein mtGGGPS_W141A (Figure 44, 30 µM and 50 µM), very high concentrations of 110 µM showed a retention volume almost identical to that of mtGGGPS_W141A (Figure 44). In combination with the behavior of mtGGGPS_A162E_I107E (Figure 43), this implicates that the mutation I107E leads to a disruption of the hexamer into its dimeric modules. As already mentioned above the reason for this could be that the mutation I107E rather disturbs interface 2 than interface 1.

7.3.1.4 **Mutational analysis of the symmetric interface 2**

Interface 2 always connects one subunit from the upper ring with one subunit from the lower ring, which belong to different dimeric modules (Figure 35*B*). In combination with the dimeric interface 1, a kind of “zigzag” link is generated: lower subunit of dimeric module 1 (deepblue) with upper subunit of dimeric module 2 (green), lower subunit of dimeric module 2 (green) with upper subunit of dimeric module 3 (cyan) and lower subunit of dimeric module 3 (cyan) with upper subunit of dimeric module 1 (deepblue).

It was shown in the previous chapter that even if the symmetric interface 1 is disturbed through the monomerizing mutation A162E, the symmetric interface 2 is able to stabilize the hexameric conformation signalizing that this interface is more important for hexamerization than the symmetric interface 1. To confirm this hypothesis, mutational analysis was performed at the symmetric interface 2. The symmetric interface 2 is mainly built of four interface residues (D57, N82, R88, Y105). N82 is localized at the beginning of α-helix 3* building H-bonds with Y105 of the adjacent protomer, Y105 is localized at the beginning of α-helix 4 and forms a stacking interaction with Y105 of the adjacent protomer, R88 is localized at the end of α-helix 3*, and D57 in the βα2-loop. D57 and R88 build polar contacts with each other (Figure 45).

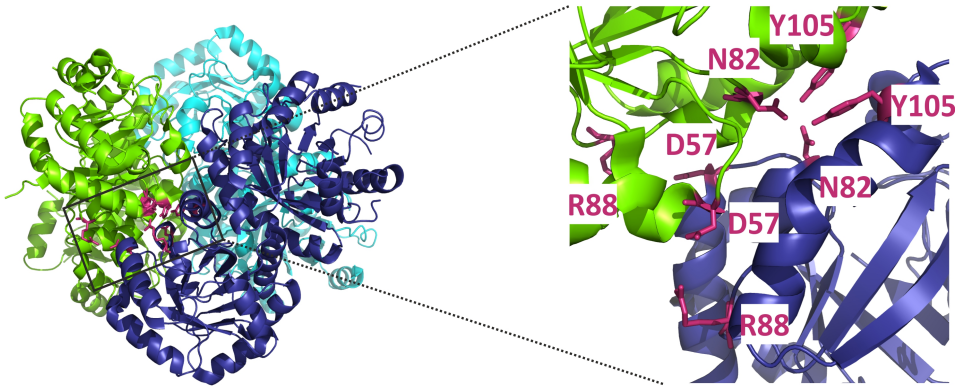


Figure 45 - The symmetric interface 2 of mtGGGPS
The symmetric interface 2 of mtGGGPS (pdb-code: 4mm1) is shown in survey (left side) and in detail (right side). The associated subunits of each dimeric module are depicted in the same color. Important interface residues of the symmetric interface 2 are depicted as sticks in magenta.

As Y105 is participating in two interactions, this residue was subjected to mutagenic analysis. Y105 was mutated to the small amino acid alanine in the background of mtGGGPS_wt (resulting in mtGGGPS_Y105A) and mtGGGPS_A162E (resulting in mtGGGPS_A162E_Y105A) to analyze its cross-connecting properties. Alanine was chosen to delete all possible stacking interactions or H-bonds. Additionally, the residues D57 and R88 were subjected to mutagenesis as well. As these two residues are involved in a polar contact between each other, D57 was mutated to arginine or R88 to glutamate to delete the polar contact and introduce repulsion. As for the mutation Y105A, both mutations were separately introduced into mtGGGPS_wt (resulting in mtGGGPS_D57R and mtGGGPS_R88E) and into mtGGGPS_A162E (resulting in mtGGGPS_A162E_D57R and mtGGGPS_A162E_R88E). The mutants were analyzed via analytical SEC (5.6.5, Figure 46).

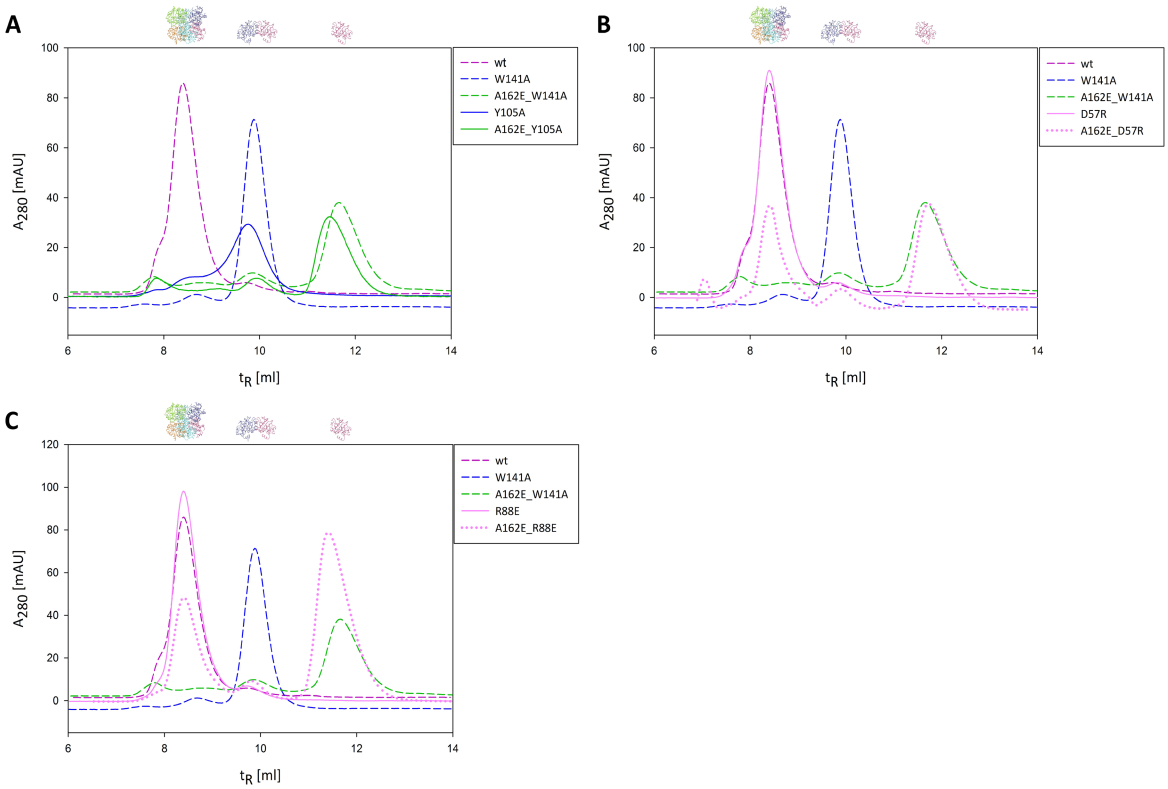


Figure 46 - Analytical size exclusion chromatography of mtGGGPS mutants with amino acid exchanges at the symmetric interface 2

The proteins were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta, 40 µM, subunit concentration), mtGGGPS_W141A (W141A, dashed lines, blue, 40 µM, subunit concentration) and mtGGGPS_A162E_W141A (A162E_W141A, dashed lines, green, 33 µM) are shown as reference peaks for hexameric, dimeric and monomeric oligomerization state. (**A**) mtGGGPS_Y105A (Y105A, blue, 25 µM, subunit concentration) and mtGGGPS_A162E_Y105A (A162E_Y105A, green, 40 µM, subunit concentration) are drawn as solid lines. (**B**) mtGGGPS_D57R (D57R, solid line, 40 µM, subunit concentration) and mtGGGPS_A162E_D57R (A162E_D57R, dotted line, 40 µM, subunit concentration) are colored lightpink. (**C**) mtGGGPS_R88E (R88E, solid line, 40 µM, subunit concentration) and mtGGGPS_A162E_ R88E (A162E_ R88E, dotted line, 40 µM, subunit concentration) are colored lightpink. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

Incorporation of Y105A into the mtGGGPS_A162E mutant resulted in monomers, whereas this mutation in mtGGGPS_wt lead to a collapse of the hexamer into dimeric modules (Figure 46*A*). This indicates that the symmetric interface 2 is as important as the asymmetric interface 3a/3b (W141, K146) for hexamerization. If either the symmetric interface 2 or the asymmetric interface 3a/3 is disturbed, the hexameric structure collapses into its dimeric modules. The interacting surfaces of both interfaces are not strong enough to keep the hexamer together once the other interface is destroyed. Asymmetric interface 3a/3b can not stabilize the hexamer over the “ring-system”, if the symmetric interface 2 is destroyed, and symmetric interface 2 can not crossconnect the

dimeric modules and stabilize the hexamer, if the asymmetric interface 3a/3b is deleted. The resulting dimeric modules after disturbance of one of these interfaces are of the same configuration as they can both be dissected into monomers with the monomerizing mutation A162E. mtGGGPS_Y105A shows a small shoulder at the corresponding retention time of the hexamer in the SEC profile (Figure 46*A*). To analyze whether this is associated with a concentration-dependent association-dissociation equilibrium, analytical SEC of mtGGGPS_Y105A with varying concentrations was performed (5.6.5). MtGGGPS_Y105A did not show a concentration-dependent association-dissociation equilibrium under the tested conditions as the percentual amount of the shoulder in the different SEC profiles did not change (Figure 47*A*).

Incorporation of D57R into the hexameric mtGGGPS_wt did not lead to dimerization, but to partial monomerization in the mtGGGPS_A162E mutant background (hexamer:monomer ratio of 1:1; Figure 46*B*). Obviously, the deletion of the polar contact does effect the oligomerization state and is partial disrupting interface 2, but not as severely as the mutation of Y105.

As for the D57R mutation, the incorporation of R88E into the hexameric mtGGGPS_wt did not lead to dimerization, but to partial monomerization in the mtGGGPS_A162E mutant background (hexamer:monomer ratio of about 2:3; Figure 46*C*). The deletion of the polar contact does effect the oligomerization state and is partial disrupting the interface 2, but again not as severely as the Y105A mutation. To analyze whether this is associated with a concentration-dependent association-dissociation equilibrium analytical SEC with varying concentrations of mtGGGPS_R88E was performed (5.6.5, Figure 47*B*).

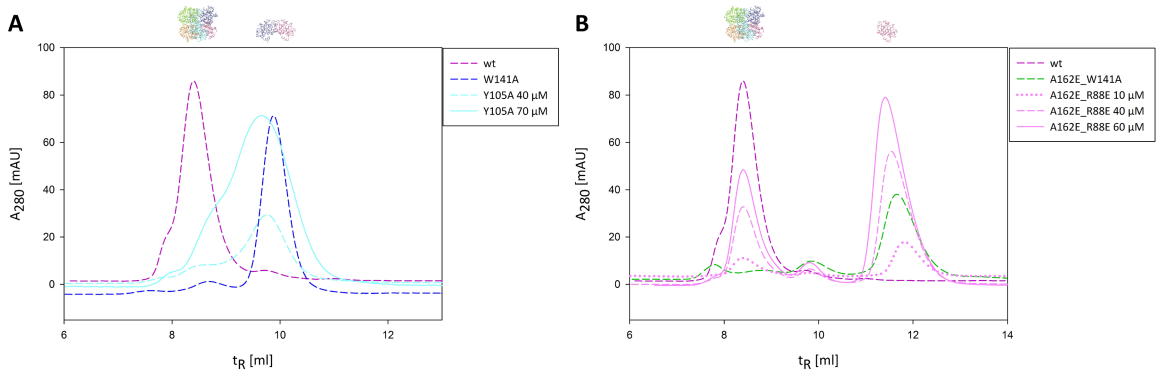


Figure 47 – Analytical size exclusion chromatography with varying concentrations of mtGGGPS_Y105A and mtGGGPS_R88E

The proteins were applied to a S75 analytical column equilibrated with 50 mM KP pH 7.5, 300 mM KCl. Elution was performed at a flow rate of 0.5 ml/min, followed by measuring the absorption at 280 nm (A_{280}) and plotted against the retention volume (t_R). mtGGGPS_wt (wt, dashed lines, magenta, 40 µM, subunit concentrations), mtGGGPS_W141A (W141A, dashed lines, blue, 40 µM, subunit concentrations) and mtGGGPS_A162E_W141A (A162E_W141A, dashed lines, green, 33 µM) are shown as reference peaks for hexameric, dimeric and monomeric oligomerization state. (**A**) Different concentrations of mtGGGPS_Y105A (Y105A) are colored cyan. 40 µM(subunit concentrations) is depicted as dashed lines and 110 µM (subunit concentrations) as solid lines. (**B**) Different concentrations of mtGGGPS_R88E (R88E) are colored lightpink. The lowest concentration of 10 µM (subunit concentrations) is depicted as dotted lines, the middle concentration of 40 µM (subunit concentrations) as dashed lines and the highest concentration of 60 µM (subunit concentrations) as solid lines. The oligomerization state is symbolized by a ribbon structure, different colors pointing out each subunit.

The ratio between the monomeric and hexameric amount of protein in mtGGGPS_R88E in the different SEC profiles did not change and was always about 2:3 (hexamer:monomer). Thus, it could be concluded that there is no concentration-dependent association-dissociation equilibrium under the tested conditions. The higher amount of monomeric protein compared to the 1:1 ratio of mtGGGPS_D57R indicates a slightly stronger effect of this mutation on the symmetric interface 2. It should be noted that Y105 is located near the center of the hexameric overall structure as it is also the case for W141, whereas R88 is localized a little bit further apart from the center of the structure and with D57 being even more further apart at the outside of the hexameric structure. One could assume that perturbation at the inner core of the hexamer does have a stronger effect on the oligomerization than disturbing at the outer shell. This would also explain why the oligomerization states of mtGGGPS_R88E and mtGGGPS_D57R, although natively interacting with each other, show a somewhat diverse behavior. Furthermore, this would also elucidate why neither the R88E nor the D57R mutation cause a dimerization effect if introduced into the hexameric mtGGGPS_wt, in contrast to the

Y105A mutation. Both mutations lie to far apart from the center to disturb the overall hexameric structural fold.

7.3.2 Biochemical characterization of mtGGGPS wild type and its mutants

Table 11 gives an overview about mtGGGPS_wt and its mutants in this work, which will be analyzed in the further part of the thesis. mtGGGPS_D144A and mtGGGPS_M168E did not show any effect on the hexameric oligomerization state and were therefore not further analyzed.

Table 11 - Overview about the oligomerization state of mtGGGPS_wt and its mutants biochemically characterized in this thesis[1]

Enzyme	Oligomerization state
mtGGGPS_wt	Hexamer
mtGGGPS_D57R	Hexamer
mtGGGPS_R88E	Hexamer
mtGGGPS_A162E	Hexamer
mtGGGPS_I155A	Hexamer - Dimer
mtGGGPS_Y105A	Dimer
mtGGGPS_I107E	Dimer
mtGGGPS_W141A	Dimer
mtGGGPS_K146A	Dimer
mtGGGPS_A162E_D57R	Hexamer - Monomer
mtGGGPS_A162E_R88E	Hexamer - Monomer
mtGGGPS_A162E_Y105A	Monomer
mtGGGPS_A162E_I107E	Monomer
mtGGGPS_A162E_W141A	Monomer

[1] For clarity, the order of the mutated amino acids in mutants containing two mutations is always beginning with the monomerizing mutation A162E

7.3.2.1 Circular dichroism spectroscopy

To ensure structural integrity of the mtGGGPS_wt and its mutants, far-UV circular dichroism (CD) spectroscopy was used. CD spectra were obtained between 185 nm and 260 nm (5.6.12, Figure 48).

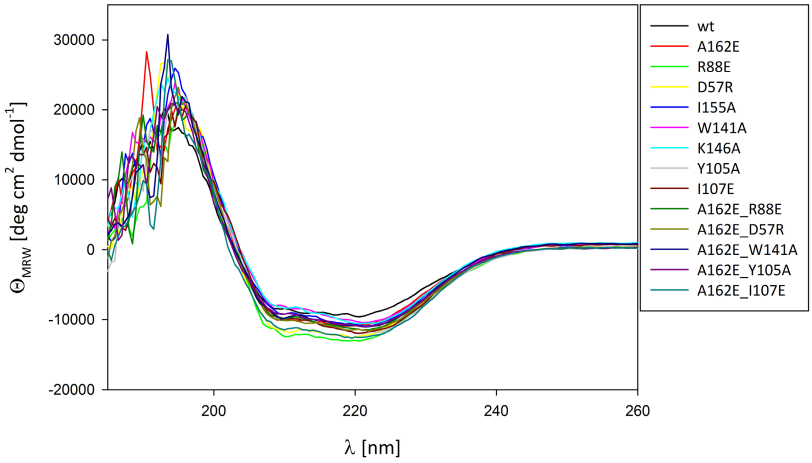


Figure 48 - Structural integrity of mtGGGPS_wt and its mutants
Far-UV CD spectra of mtGGGPS_wt and its mutants (10 µM, subunit concentration) were recorded in 50 mM potassium phosphate pH 7.5 from 185 nm to 260 nm (d = 1 mm) at room temperature.

MtGGGPS_wt and its mutants showed CD spectra indicative for a mixture of α-helices and β-sheets in agreement with the $(\beta\alpha)_8$-barrel fold, which confirmed a native configuration for all proteins. α-helices show two characteristic minima at about 208 nm and 222 nm and a maximum at around 190 nm. β-sheets exhibit a minimum around 218 nm and a maximum at 196 nm (Kelly *et al.*, 2005; Wallace *et al.*, 2003).

As higher oligomerization is known to contribute to higher thermostability (3.7), mtGGGPS_wt and its mutants were subjected to thermal heating followed by CD spectroscopy (4.6.12, Figure 49).

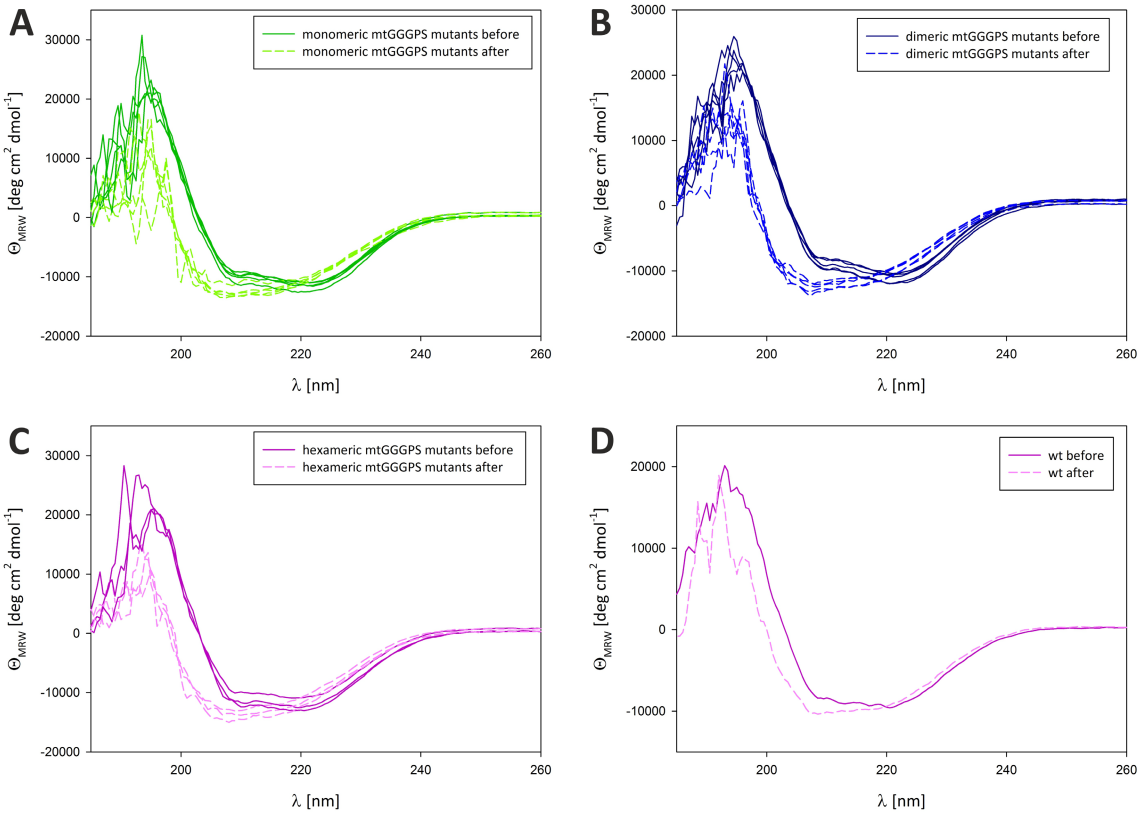


Figure 49 – CD spectra before and after heating of mtGGGPS_wt and its mutants
Far-UV CD spectra of mtGGGPS_wt and mutants (10 µM, subunit concentration). Spectra were recorded in 50 mM potassium phosphate pH 7.5 from 185 nm to 260 nm (d = 1 mm) at room temperature (before) and immediately after heating to 95 °C (after). (**A**) All monomeric mtGGGPS mutants (A162E_D57R, A162E_R88E, A162E_Y105A, A162E_I107E, A162E_W141A) before heating at room temperature (solid green lines) and after heating to 95 °C (dashed palegreen lines). (**B**) All dimeric mtGGGPS mutants (Y105A, I107E, W141A, K146A, I155A) before heating at room temperature (solid blue lines) and after heating to 95 °C (dashed lightblue lines). (**C**) All hexameric mtGGGPS mutants (D57R, R88E, A162E) before heating at room temperature (solid magenta lines) and after heating to 95 °C (dashed lightpink lines). (**D**) mtGGGPS_wt before heating at room temperature (solid magenta line) and after heating to 95 °C (dashed lightpink line). Each spectrum represents an individual mtGGGPS mutant, but for clarity an individual labelling was omitted.

From the CD spectra can be interpreted that the α-helical content (minimum at 222 nm) decreased slightly upon heating to 95 °C, whereas the random coil content increased, which is shown through the broadening of the spectrum towards 200 nm. Random coils typically exhibit a minimum around 195 nm (Kelly *et al.*, 2005; Wallace *et al.*, 2003). This indicates that mtGGGPS and its mutants are rather thermostable.

7.3.2.2 **Differential scanning calorimetry**

In order to quantify and compare the thermal stabilities of mtGGGPS_wt and its mutants, differential scanning calorimetry (DSC) was performed (5.6.11). DSC is a powerful technique for investigating thermal unfolding of proteins by measuring changes in heat

capacity thereby revealing changes in the tertiary structure of the protein. This technique enables the heating to a maximal temperature of 130 °C. The results are shown in Figure 50.

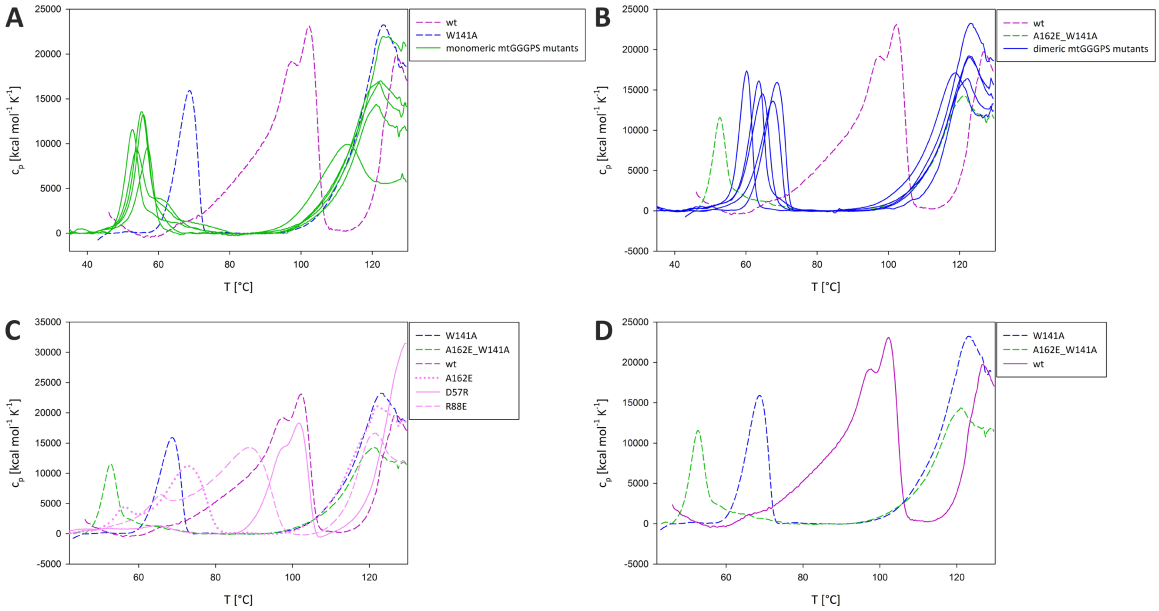


Figure 50 – Thermal stability of mtGGGPS_wt and its mutants
Thermal denaturation was followed by DSC. Changes in heat capacity of 20 µM protein (subunit concentration) in 50 mM potassium phosphate pH 7.5 were monitored from 30 °C to 130 °C at a scan rate of 1 °C/min. 50 mM potassium phosphate served as reference. The curves were baseline corrected. (**A**) All monomeric mtGGGPS mutants (A162E_D57R, A162E_R88E, A162E_Y105A, A162E_I107E, A162E_W141A) are depicted as solid green lines. mtGGGPS_wt (wt, dashed magenta line) and mtGGGPS_W141A (W141A, dashed blue line) serve as hexameric and dimeric reference. (**B**) All dimeric mtGGGPS mutants (Y105A, I107E, W141A, K146A, I155A) are depicted as solid blue lines. mtGGGPS_wt (wt, dashed magenta line) and mtGGGPS_A162E_W141A (A162E_W141A, dashed green line) serve as hexameric and monomeric reference. (**C**) All hexameric mtGGGPS mutants are colored lightpink. mtGGGPS_A162E (A162E, dotted line), mtGGGPS_D57R (D57R, solid line), mtGGGPS_R88E (R88E, dashed line). MtGGGPS_wt (wt, dashed magenta line), mtGGGPS_W141A (W141A, dashed blue line) and mtGGGPS_A162E_W141A (A162E_W141A, dashed green line) serve as hexameric, dimeric and monomeric reference. (**D**) mtGGGPS_wt (wt, solid magenta line), mtGGGPS_W141A (W141A, dashed blue line) and mtGGGPS_A162E_W141A (A162E_W141A, dashed green line) serve as hexameric, dimeric and monomeric reference. Each thermogram represents an individual mtGGGPS mutant, but for clarity an individual labelling was omitted.

All DSC thermograms did not show any signs of heavily aggregated protein after heat denaturation. MtGGGPS_wt and its mutants showed a characteristic behavior, which can be seen best by looking at the thermogram of the dimeric mutant mtGGGPS_W141A (Figure 50*D*, blue dashed curve). This mutant shows a distinct symmetric transition at approx. 68 °C (T_1) and a second, even higher transition peak in the DSC thermogram at approx. 124 °C (T_2). This high temperature transition is common for mtGGGPS_wt and all mutants and immediately explains why the loss of ellipticity in the CD spectroscopic

heating experiments performed up to 95 °C is only about 20 %. The transitions confirm the proposed thermostabilizing effect of higher oligomerization as the monomeric mutants showed the 1st transition at around 50 °C (Figure 50*A*), the dimeric mutants at around 60 °C (Figure 50*B*), and the hexameric mtGGGPS_wt and its mutants at around 95 °C (Figure 50*C* and *D*).

The DSC thermograms of mtGGGPS_A162E and mtGGGPS_R88E differed from the other thermograms (Figure 50*C*). In mtGGGPS_A162E three transitions occurred instead of two. As it was shown by SEC, mtGGGPS_A162E forms small amounts of monomeric protein (Figure 42), suggesting that the 1st transition at 55 °C is caused by this monomeric fraction. The 2nd transition at 73 °C on the other hand might be caused by the hexameric fraction. This change happens at an about 20 K lower temperature compared to the first transition in the wild type. This could be explained by the introduced mutation, which presumably disturbs the structure. As this transition is lying in between T_1 and T_2 it will be in the following referred to as T_{1*}.

Also mtGGGPS_R88E exhibits three transitions instead of two. As already discussed in the SEC experiments, the mutation R88E did not show an effect if incorporated into the mtGGGPS_wt, but it led to a mixture of monomeric and hexameric proteins if incorporated into mtGGGPS_A162E (Figure 47*B*). This indicates that the mutation somehow affects the structure of the hexamer leading to small dimeric amounts of protein not visible in the SEC experiments under the prevailing conditions. Therefore, it can be assumed here that the 1st transition at 66 °C is caused by the dimeric fraction and that the 2nd transition at 90 °C is caused by the hexameric fraction. As it is the case for mtGGGPS_A162E this transition happens at an about 15 K lower temperature compared to the wild type, which again could be explained by the introduced mutation leading to a perturbation in the native structure.

Repetitive thermal denaturation experiments revealed that the thermal transitions observed by DSC are mostly irreversible (data not shown).

7.3.2.3 **Nano differential scanning fluorimetry**

To further confirm the 1st transitions, nano differential scanning fluorimetry (nanoDSF) was conducted (5.6.13). nanoDSF is a powerful method for monitoring changes in the protein side chain environment during thermal heating in a label-free fluorimetric analysis. This technique takes into account the strong temperature dependence of the fluorescence signal by calculating the ratio of fluorescence emission intensities at 350 nm and 330 nm. The nanoDSF thermograms for mtGGGPS_wt and its mutants are shown in Figure 51.

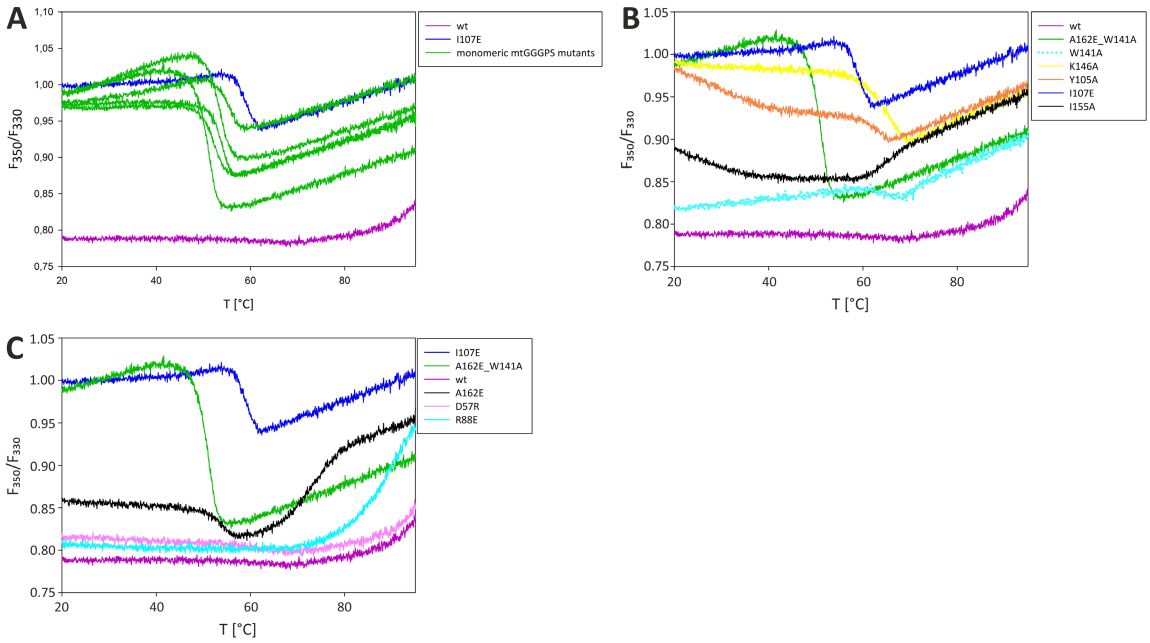


Figure 51 – Thermal stability of mtGGGPS_wt and its mutants
Thermal denaturation of mtGGGPS_wt and its mutants was followed by nanoDSF. The change in the ratio of the emission fluorescence at 350 nm to 330 nm of 30 µM protein (subunit concentration) in 50 mM potassium phosphate pH 7.5 was monitored from 20 °C to 95 °C at a scan rate of 1 °C/min. (**A**) All monomeric mtGGGPS mutants (A162E_D57R, A162E_R88E, A162E_Y105A, A162E_I107E, A162E_W141A) are depicted as green lines. mtGGGPS_wt (wt, magenta) and mtGGGPS_I107E (I107E, blue) serve as hexameric and dimeric reference. (**B**) Dimeric mtGGGPS mutants: mtGGGPS_I107E is depicted in deepblue, mtGGGPS_Y105A in orange, mtGGGPS_W141A as dotted cyan lines, mtGGGPS_K146A in yellow and mtGGGPS_I155A in black. mtGGGPS_wt (wt, magenta) and mtGGGPS_A162E_W141A (A162E_W141A, green) serve as hexameric and monomeric reference. (**C**) Hexameric mtGGGPS mutants: mtGGGPS_D57R is depicted in lightpink, mtGGGPS_R88E in cyan, mtGGGPS_A162E in black. mtGGGPS_wt (wt, magenta), mtGGGPS_I107E (I107E, blue) and mtGGGPS_A162E_W141A (A162E_W141A, green) serve as hexameric, dimeric and monomeric reference. Each thermogram represents an individual mtGGGPS mutant, but for clarity an individual labelling was omitted in case of monomeric enzymes.

All proteins showed a characteristic behavior, which can be seen best by looking at the thermograms of monomeric mutants (Figure 51*A*, green curves). These mutants show transitions at about 50 °C, whereas the dimeric proteins showed transitions at about 60 °C

(Figure 51*B*), in accordance with the T_1 values observed in the DSC measurements. For the hexameric mtGGGPS_wt and hexameric mtGGGPS mutants no or only the onset of a transition could be observed (Figure 51*C*), as the nanoDSF technique is limited to a maximal temperature of 95 °C. Astonishingly, the monomeric and dimeric proteins showed a decrease in the F_{350}/F_{330}-ratio indicating a blue-shift of the emission fluorescence of tryptophan, which indicates that the tryptophans are exposed to a more hydrophobic environment after thermal heating (Figure 51*A* and *B*). In contrast, the hexameric proteins showed the expected increase in the F_{350}/F_{330}-ratio upon thermal heating (Figure 51*C*).

Again, the mtGGGPS_A162E mutant showed a different nanoDSF thermogram compared to the hexameric mtGGGPS_wt (Figure 51*C,* black line). As already seen in the DSC thermograms, two transitions occurr. The first transition exhibiting a decrease in the ratio F_{350}/F_{330} and the second transition exhibiting an increase in the ratio F_{350}/F_{330}. It could be assumed that the 1st transition (T_1) belongs to the partial unfolding of the monomeric proteins therefore exhibiting a blue-shift and that the 2nd transition (T_{1*}) belongs to the unfolding of the hexameric proteins therefore exhibiting a red-shift.

To sum up the results from CD, DSC and nanoDSF experiments, a small signal shift in the CD measurement as a consequence of heating to 95 °C could be observed for all mtGGGPS mutants, which is presumably due to the loss of α-helical parts and only a small loss of secondary structure. Accordingly, a significant transition in the DSC measurements due to strong changes in the heat capacity and a transition in the nanoDSF measurements due to changes in tertiary structure could be observed. All these transitions are in total agreement (Table 12). Only in the DSC experiments, an independent second transition (T_2) at around 120 °C was observed, which most likely represents the complete denaturation of the proteins.

Table 12 - Overview of the data of DSC- and nanoDSF measurements[1,2]

Enzyme	T_1 [°C] DSC	T_1/T_{1*} [°C] nanoDSF	T_{1*} [°C] DSC	T_2 [°C] DSC
mtGGGPS_wt	97.5 ± 0.7	> 95	102.5 ± 0.6	127.0 ± 0.5
mtGGGPS_A162E	55.0 ± 0.0	53.8 ± 0.4 73.6 ± 0.1	73.0 ± 0.7	125.8 ± 1.8
mtGGGPS_R88E	65.7 ± 1.1	> 95	89.1 ± 1.3	122.2 ± 0.7
mtGGGPS_D57R	97.2 ± 0.1	> 95	101.7 ± 0.5	129.9 ± 0.1
mtGGGPS_I155A	64.8 ± 0.3	65.3 ± 0.5	-	123.3 ± 2.4
mtGGGPS_W141A	68.3 ± 0.5	72.1 ± 0.1	-	124.4 ± 0.1
mtGGGPS_K146A	67.4 ± 0.2	66.0 ± 0.1	-	123.5 ± 0.7
mtGGGPS_Y105A	64.2 ± 1.1	63.2 ± 0.5	-	124.3 ± 0.7
mtGGGPS_I107E	60.0 ± 0.2	59.0 ± 0.1	-	120.5 ± 0.7
mtGGGPS_A162E_R88E	55.8 ± 0.2	53.9 ± 0.1	-	114.6 ± 0.6
mtGGGPS_A162E_D57R	53.8 ± 0.1	52.2 ± 0.1	-	126.2 ± 0.8
mtGGGPS_A162E_W141A	52.7 ± 0.1	51.0 ± 0.1	-	120.7 ± 1.0
mtGGGPS_A162E_Y105A	55.5 ± 0.5	53.4 ± 0.1	-	122.7 ± 0.1
mtGGGPS_A162E_I107E	56.5 ± 0.1	55.3 ± 0.2	-	124.2 ± 0.4

[1] Hexameric oligomerization state is depicted in magenta, dimeric in blue and monomeric in green

[2] $T_1/T_{1*}/T_2$: apparent midpoint temperature, thermal transitions with T_1 referring to the first observable transition, T_2 referring to the last observable transition (complete unfoding of the protein) and T_{1*} referring to observable transitions in between, average values of duplicates were calculated (given with standard deviations)

7.3.2.4 Irreversible heat inactivation

To estimate whether this 1st transition causes in the enzyme a loss of activity, irreversible heat inactivation was performed. For confirmation that the mtGGGPS mutants show activity at all, radiometric activity assays were performed (5.6.14.1, Figure 52).

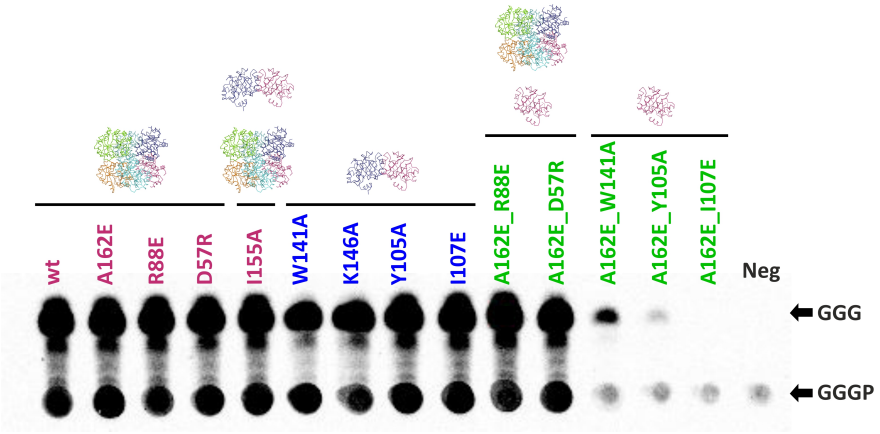


Figure 52 – Activity assay of mtGGGP_wt and its mutants
The substrate ^{14}C-G1P was incubated with GGPP and 1 µM of mtGGGPS enzymes (1 µM, subunit concentration) for 2 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Hexameric mtGGGPS mutants are shown in magenta, dimeric mtGGGPS mutants in blue and monomeric mtGGGPS mutants in green. As negative control (Neg), no enzyme was added. As positive control, mtGGGPS_wt (wt) was added. The origin of the chromatography (marked by the product spot GGGP from GGGPS activity) and the product of CIP dephosphorylation (GGG) are marked by arrows. The oligomerization state is symbolized by ribbon structure, different color pointing out each subunit.

Between hexamers and dimers, no obvious difference in activity after 2 h of incubation could be detected. The monomers, however, show no or only low (mtGGGPS_A162E_W141A) activity. Because Peterhoff *et al.* (2012) proposed that monomeric bsPcrB is more active with shorter substrates like geranyl pyrophosphate (GPP, C_{10}), this substrate was also tested for all mtGGGPS mutants (9.4, Figure S2). An overall reduced activity of all mtGGGPS mutants with GPP compared to the activity with the preferred substrate GGPP could be detected. Monomeric mtGGGPS mutants did not show any activity at all under the given conditions.

Next, irreversible heat inactivation was performed. Because the activity of the monomers was very low, only hexameric and dimeric enzymes could be analyzed with a spectroscopic assay (5.6.16.3). MtGGGPS_wt (hexamer), mtGGGPS_A162E (hexamer) and mtGGGPS_W141A (dimer) were chosen for analysis. mtGGGPS_wt and the two mutants were incubated at different temperatures. After various amounts of time, the remaining activity of the sample was measured and the apparent half-lifes ($t^{app}_{1/2}$) were calculated by fitting a single exponential decay equation to the data (9.5, Figure S3). The obtained $t^{app}_{1/2}$ values are summarized in Table 13 and compared to the transition temperatures obtained by DSC and nanoDSF measurements.

Table 13 - Results from heat inactivation and thermal denaturation studies of mtGGGPS_wt, mtGGGPS_W141A and mtGGGPS_A162E_W141A[2]

Enzyme	$t^{app}_{1/2}$ [min]	T_1/T_{1*} [°C][1] DSC	T_1/T_{1*} [°C][1] nanoDSF
mtGGGPS_wt	**100°C: 16.5****	97.5 ± 0.7 102.5 ± 0.6	> 95
mtGGGPS_A162E	**85°C: 3.3**	55.0 ± 0.0 73.0 ± 0.7	53.8 ± 0.4 73.6 ± 0.1
mtGGGPS_W141A	**65°C: 15.0**	68.3 ± 0.5	72.1 ± 0.0

[1] values are from Table 12
[2] **: 0.1 mM GdmCl was added before heat incubation

In case of mtGGGPS_wt, the fit of the data points observed under native buffer conditions at 100 °C with a single exponential decay was not feasible (Figure S3). Therefore, the heat inactivation was performed in the presence of 0.1 M GdmCl, which resulted in a $t^{app}_{1/2}$ of 16.5 min at 100 °C. In comparison, mtGGGPS_A162E showed a $t^{app}_{1/2}$ of 3.3 min at 85 °C and mtGGGPS_W141A showed a $t^{app}_{1/2}$ of 15 min at 65 °C. This clearly indicated that both mutants, mtGGGPS_A162E and mtGGGPS_W141A, are much less thermostable than mtGGGPS_wt. The temperatures at which the mutants rapidly loose their activity, 85 °C for mtGGGPS_A162E, and 65 °C for mtGGGPS_W141A, coincide nicely with the first transitions in the DSC and nanoDSF measurements (Table 13).

In order to estimate the stability against heat inactivation for a monomeric protein, the heat inactivation of mtGGGPS_A162E_W141A was followed by a radiometric activity assay (5.6.16.3). This assay is much more sensitive than the spectroscopic assay, which allows one to detect even very low activities (Figure 53).

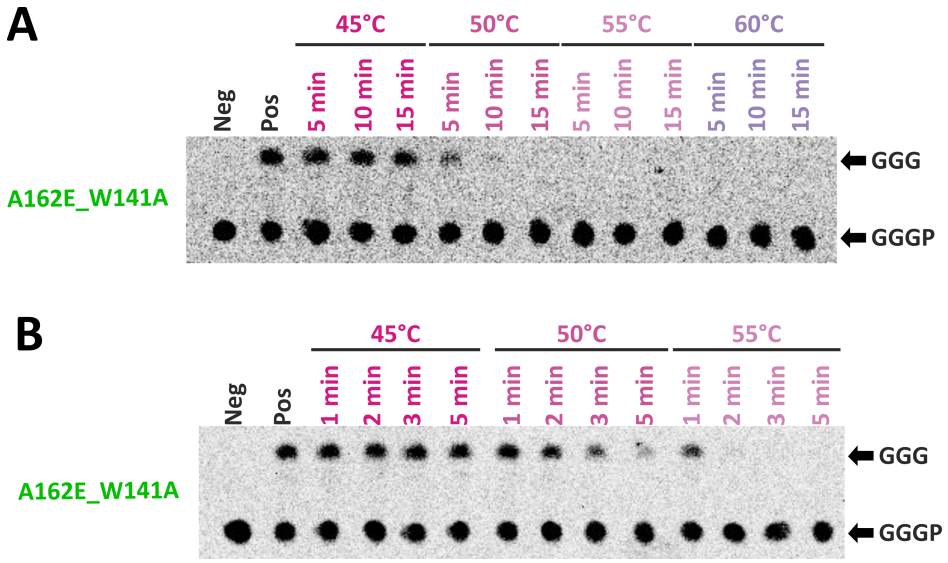


Figure 53 – Radiometric heat inactivation of monomeric mtGGGPS_A162E_W141A
40 µM mtGGGPS_A162E_W141A was incubated in a volume of 100 µl at given temperatures. After different periods of time, samples were withdrawn, chilled on ice and centrifuged. The remaining activity of the supernatant was assayed in the presence of GGPP and ^{14}C-G1P. The reaction was started by addition of 25 µl enzyme and incubated for 2 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. (**A**) Incubation at 45 °C, 50 °C, 55 °C and 60 °C for 5 min, 10 min and 15 min. (**B**) Incubation at 45 °C, 50 °C and 55 °C for 1 min, 2 min, 3 min and 5 min. As negative control (Neg), no enzyme was added. As positive control (Pos), the enzyme was added without heat incubation. The origin of the chromatography (marked by the product spot GGGP from GGGPS activity) and the product of CIP dephosphorylation (GGG) are marked by arrows.

Monomeric mtGGGPS_A162E_W141A showed no activity above 50 °C (Figure 53*A*). It could clearly be seen that the monomeric protein did not lose activity at 45 °C even after 15 min incubation, whereas at 50 °C the activity decreased within 5 min of incubation to almost zero (Figure 53*B*). The rapid loss of activity at 50 °C coincides with the T_1 values observed in DSC and nanoDSF measurements, which lie between 48 °C to 53 °C. This suggests that the first thermal transition leads to inactivation and raises the intriguing question: Which part of the protein collapses at the 1st transition temperature of the monomeric, dimeric and hexameric proteins and leads to inactivation of the enzymes that still exhibit almost 80-90 % of residual secondary structure? It can not be due to monomerization of the hexamer or dimer, as the monomer itself shows this transition as well, concomitant with a loss of activity.

7.3.2.5 **MD simulations**

The parts of the protein that unfold partially and thus lead to inactivation were identified using computational analysis. Studying the dynamics of the full mtGGGPS complex via Molecular Dynamics Simulation (MD simluations, 5.7) within an acceptable time interval was not feasible. Thus the analysis had to be restricted to one subunit of mtGGGPS_wt. Monomers have their 1st transition at around 50 °C, suggesting, that MD simulations performed at room temperature (RT) and at 45 °C have no effect on 2D structure. In contrast, for a MD simulation at 80 °C, which is above the 1st transition temperature, a change in the α-helical content was anticipated. Indeed, the MD simulations are in agreement with these expectations (Figure 54).

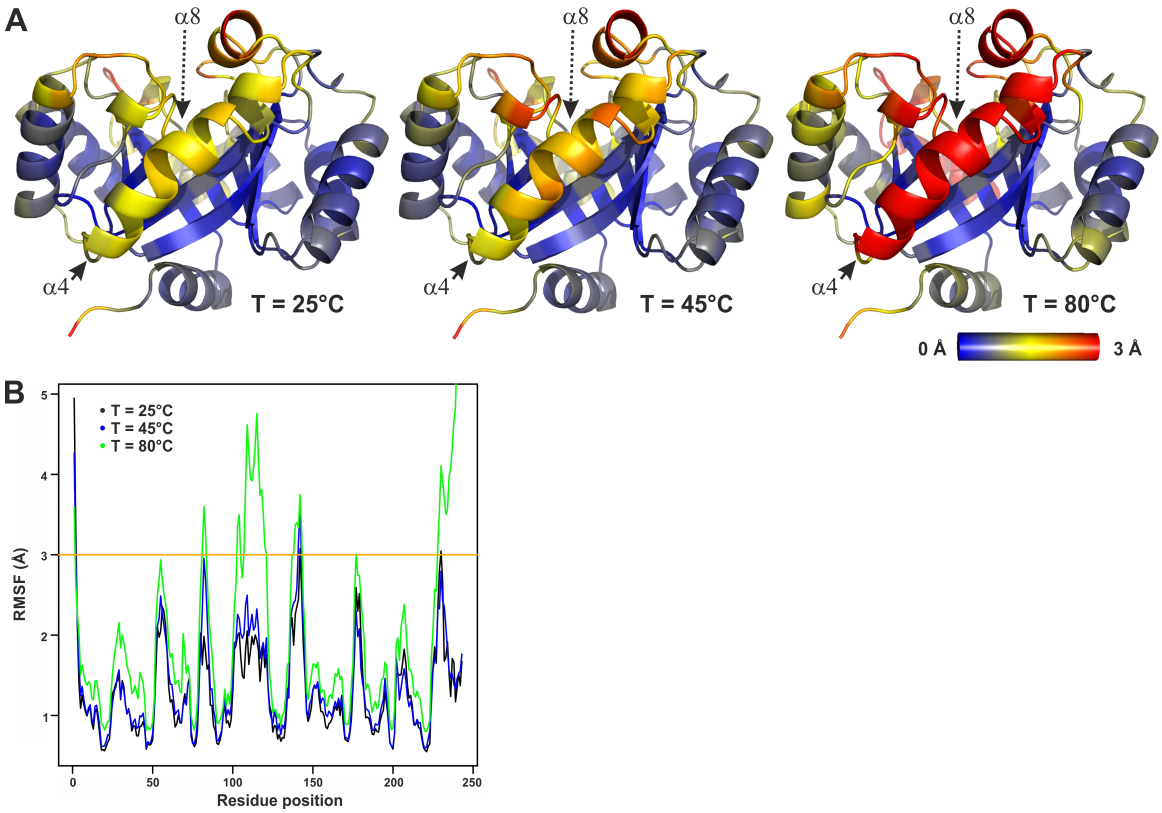


Figure 54 - MD Simulation analysis of one subunit of hexameric mtGGGPS_wt
All calculations were done for three different temperatures, namely T = 25 °C, 45 °C, and 80 °C. **(A)** Residue-specific RMSF values resulting from one series of snapshots. The RMSF values were mapped onto the initial structure and are color coded, see bar depicted in panel. (**A**) All values above 3 Å are shown in red. (**B**) Same results as shown in (**A**); however the residue-specific RMSF values are plotted residue-wise without an upper cut-off at 3 Å (orange line) applied in (**A**). The figure was provided by Kristina Heyn (AG Prof. R. Merkl, University of Regensburg). α-helix 4 (solid arrow, front) and α-helix 8 (dotted arrow. background) are marked.

Different parts of mtGGGPS showed variabilities over the MD stimulation. Whereas the MD simuation at 25 °C and 45 °C exhibited a very similar fluctuation pattern (Figure 54*A* and *B*, black and blue line), the MD simulation at 80 °C showed significant enhanced fluctuations (Figure 54*A* and *B*, green line). The α-helix 4 (Figure 54*B*, residues 101-121) as well as the C-terminal domain containing α-helix 8 (Figure 54*B*, residues 234-246) displayed the most significant increases in the RMSF value of more than 3 Å compared to the values after MD simulation at 25 °C and 40 °C. Also the "swinging door" α-helix 3* (Figure 54*B*, residues 82-88), the interacting surface 3a (α-helix 5´, Figure 54*B*, residues 138-143) and the loop βα6 (Figure 54*B*, residues 175-184) turned out to be more flexible with a difference in RMSF value of about 0.5-1 Å. Regarding the swinging door, building an access path for the polyprenyl pyrophosphate (Payandeh *et al.*, 2006), and the loop βα6, participating in G1P binding (Peterhoff *et al.*, 2014), the increases in the RMSF value occur in equivalent pace indicating a higher flexibility by nature as already proposed by average temperature factors before (Payandeh *et al.*, 2006). In contrast the difference in RMSF regarding the α-helix 5´ stayed almost the same at 25 °C and 40 °C but increased by about 1 Å at 80 °C.

7.3.2.6 **Proposed unfolding pathway of mtGGGPS**

Based on the results of the MD simulation and the results of the various biophysical studies (CD, DSC, nanoDSF, analytical SEC and heat inactivation) an unfolding pathway for the enzyme mtGGGPS_wt and mutants was proposed (Figure 55).

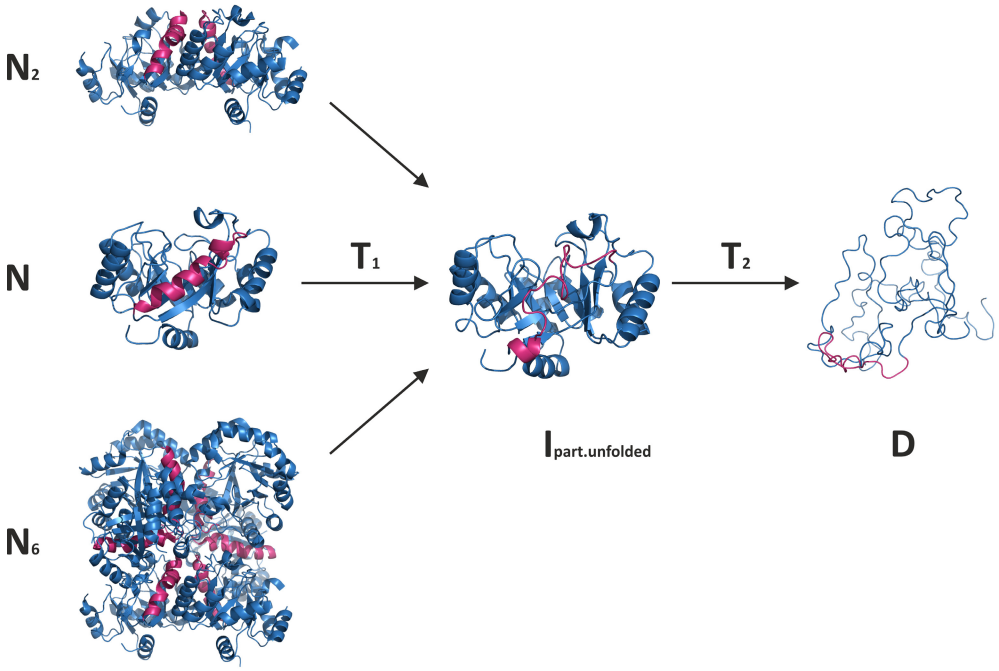


Figure 55 - Proposed unfolding pathway of mtGGGPS_wt and its mutants
The native states of monomeric (N), dimeric (N_2) and hexameric (N_6) mtGGGPS (pdb-code: 4mm1) are depicted as ribbon diagrams. α-helix 4 is depicted in magenta. Concomitant with the first thermal transition (T_1), α-helices - among them α-helix 4 - get partially unfolded resulting in a locally denatured intermediate ($I_{part.unfolded}$) presumably in form of higher-order soluble aggregates. For clarity only one subunit is shown. With rising temperature, the partially folded intermediate completely unfolds at the second thermal transition (T_2) to the denatured state (D).

The first thermal transition (T_1) leads to a perturbation of mainly α-helix 4 (Figure 55), α-helix 8 and α-helix 5´ resulting in a partially unfolded inactive intermediate ($I_{part.unfolded}$) exhibiting almost 80-90 % of residual secondary structure compared to the native structure (N) as confirmed by CD (Figure 49). The same mechanism happens in the dimeric (N_2) and hexameric (N_6) enzymes, but at higher temperatures due to their higher oligomerization state, which stabilizes α-helix 4 and α-helix 5´ through the interacting surfaces of the adjacent protomers. The partially unfolded intermediate forms inactive, soluble higher-order oligomers as detected by analytical SEC experiments. α-helix 4 is lying at the edge of the symmetric interface 1 (lower part of the helix) and the symmetric interface 2 (upper part of the helix) and α-helix 5´ is part of the asymmetric interface 3a/3b. It can be assumed that the perturbation of these helices in the hexamer due to increasing temperature, is followed first by a dissociation to partially unfolded dimers. As the α-helix 4 is also a central element of the symmetric interface 1 (dimeric interface) it can further be assumed that these partially unfolded dimers monomerize before building soluble aggregates. The second transition (T_2) at around 120 °C, that is common for

mtGGGPS_wt and all mutants is presumed to induce the full denaturation of the protein (D).

The proposed mechanism including a partially folded inactive, but stable intermediate with almost 80-90 % residual secondary structure is not inconceivable as there are several examples of similar unfolding mechanisms in literature. Tarafdar *et al.* (2014) proposed for a highly thermostable hairpin protein (HrpZ), that the polydisperse oligomeric structures at the beginning of thermal upheating first dissociate to dimers, then to partially unfolded dimers and lastly to completely unfolded monomers. They also observed multiple transitions in DSC experiments concomitant with no changes in CD thermograms until the onset of transition 2. Sedlak *et al.* (2014) investigated the kinetic stability of cytochrome c oxidase. In DSC thermograms, two irreversible transitions could be detected. The 1st transition is associated with conformational changes that are critical for activity, as they observed a complete loss of enzymatic activity. Only small changes in molar ellipticity were found at this transition that were interpreted as a significant perturbation near the enzyme active site and at least partial unfolding of the secondary structure. The 2nd transition was proposed to consist of several processes including subunit dissociation and unfolding of secondary and tertiary structure. For the triose phosphate isomerase (TIM) from *Trypanosoma cruzi* two existing intermediates were found, one of them exhibiting considerable secondary structure but lacking most of the tertiary interactions present in native yTIM (Mixcoha-Hernandez *et al.*, 2007). For glycosomal TIM, Guzman-Luna and Garza-Ramos (2012) revealed two inactive partially folded intermediates, one of them exhibiting 45 % residual secondary structure, but unable to bind the substrate. Zhu *et al.* (2001) engineered a leucine zipper mutant of the GCN4 leucine zipper with a three-state unfolding pathway. They proposed that the native dimer gets partially unfolded, thereby yielding a mixture of partially unfolded oligomeric intermediates of order greater than two, which change to unstructured monomers in a second transition. Ramstein *et al.* (2003) manifested a dimeric thermal unfolding intermediate for the *E. coli* histone-like HU proteins. They investigated the thermal stability by using CD and DSC techniques. The proposed pathway starts with the formation of a partially folded dimeric intermediate exhibiting about 50 % residual secondary structure and the inability to bind DNA anymore, which in a second transition denatures

into unfolded monomers. In DSC thermograms, two partial overlapping heat absorption peaks could be observed with at least 25 K difference. The CD thermograms at 220 nm did not show any transition, therefore CD meausurements were performed at 200 nm resulting in thermal transitions at the same temperature as observed in the DSC thermograms. Lastly, Groemping and Reinstein (2001) investigated the folding properties of the nucleotide exchange factor GrpE from *Thermus thermophilus,* thereby revealing two transitions. The 1st transition leads to stable, partially folded intermediates with reduced activity and the 2nd laeds to monomerization and complete denaturation. The first transtition is silent in CD-spectroscopy, while the second transition mainly constitutes the CD signal. All these examples support the above proposed unfolding mechanism for mtGGGPS.

7.3.2.7 **Crystallization of mtGGGPS_I107E, mtGGGPS_W141A and mtGGGPS_A162E**

To elucidate how the introduced mutations affect the overall native structure and to analyze possible differences in substrate binding, dimeric mtGGGPS_I107E, dimeric mtGGGPS_W141A, and hexameric mtGGGPS_A162E were subjected to crystallography. The mutants were expressed with a C-terminal $(His)_6$ tag and purified by metal chelate affinity chromatography (5.4.2.2). Subsequently, the proteins were dialyzed against 50 mM Tris/HCl pH 8.0 and concentrated to 29.6 mg/ml (mtGGGPS_I107E), 38.7 mg/ml (mtGGGPS_W141A), and 27.7 mg/ml (mtGGGPS_A162E). Crytsals were obtained by hanging drop vapour diffusion methods in 96 well plates at 18 °C. The mtGGGPS mutants were supplemented with 10 mM $MgCl_2$ and with or without 1 mM G1P. First crystals appeared after 2 weeks at 18°C without G1P in 0.1 M BIS-TRIS propane pH 9.0, 10 % (v/v) Jeffamine®ED-2001 pH 7.0 (I107E, PEGRx™1+2 (MOLECULAR DIMENSIONS)), in 0.2 M lithium citrate tribasic tetrahydrate, 35 % (v/v) glycerol ethoxylate (W141A, MIDAS™ (MOLECULAR DIMENSIONS)) and in 2.0 M sodium formate, 0.1 M sodium acetate pH 4.6 (A162E, The structure screen combination (MOLECULAR DIMENSIONS)) as reservoir solution (Figure 56).

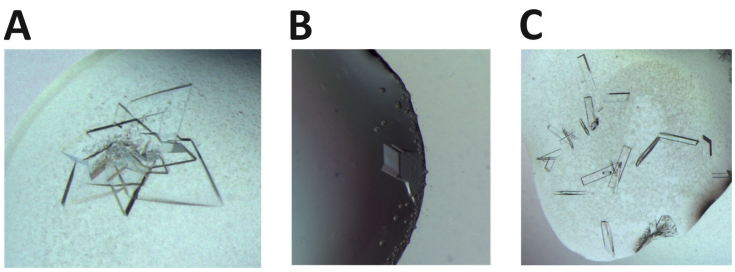


Figure 56 - Crystallization of mtGGGPS_A162E, mtGGGPS_W141A and mtGGGPS_I107E
(**A**) Crystals of mtGGGPS_A162E appeared after 2 weeks at 18 °C with 2.0 M sodium formate, 0.1 M sodium acetate pH 4.6 as reservoir solution. (**B**) Crystals of mtGGGPS_W141A appeared after 2 weeks at 18 °C with 0.2 M lithium citrate tribasic tetrahydrate, 35 % (v/v) glycerol ethoxylate as reservoir solution. (**C**) Crystals of mtGGGPS_I107E appeared after 2 weeks at 18 °C with 0.1 M BIS-TRIS propane pH 9.0, 10 % (v/v) Jeffamine®ED-2001 pH 7.0 as reservoir solution.

After 1 month at 18°C, single crystals were collected and flash frozen in liquid nitrogen. The data collection and refinement statistics are shown in the supplement 9.6 (Table S2 - Table S4). The obtained structures of all three mtGGGPS mutants showed a hexameric oligomerization state identical to the wild type (data not shown), although mtGGGPS_W141A and mtGGGPS_I107E were shown to be dimers in analytical size exclusion chromatography experiments. This is presumably due to the high concentration needed for crystallization that do not correspond to physiological conditions. Therefore, these results could not elucidate differences in the overall native structure or G1P-binding compared to the wild type and thus are not further discussed.

7.3.3 Characterization of hexameric archaeal tkGGGPS and bacterial cpGGGPS compared to their dimeric mutants

To confirm the differences for hexamers compared to dimers as observed for mtGGGPS, two other pairs of hexamers and dimeric mutants were subjected to thermal denaturation analysis and heat inactivation: the archaeal GGGPS from *Thermococcus kodakarensis* (tkGGGPS) and its dimeric mutant (tkGGGPS_W143A), and the bacterial GGGPS from *Chitinophaga pinensis* (cpGGGPS) and its dimeric mutant (cpGGGPS_Y143A). Peterhoff *et al.* (2014) verified the dimeric oligomerization state for both mutants by analytical size exclusion chromatography in combination with static light scattering. In both mutants, the aromatic amino acid of the interacting surface 3a that corresponds to W141 in mtGGGPS was mutated to alanine (tkGGGPS_W143A and cpGGGPS_Y143A). To analyze a stabilizing effect of hexamerization on thermal stability, CD and DSC measurements were performed as for the mtGGGPS mutants (5.6.12, 5.6.11, Figure 57).

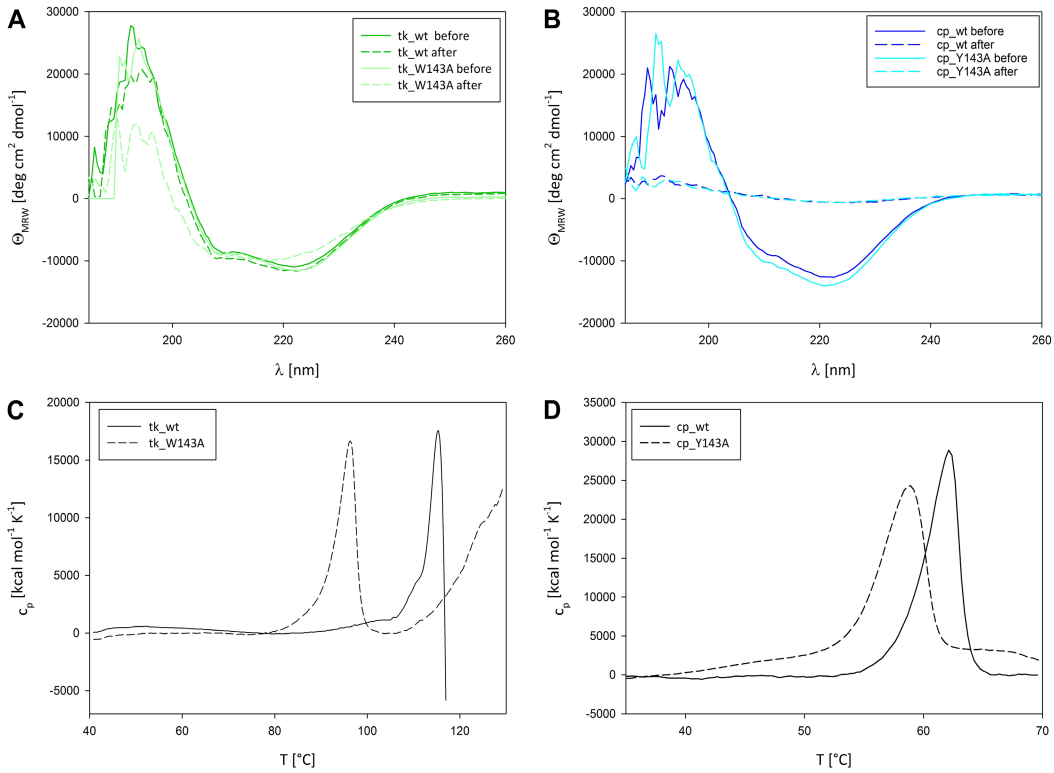


Figure 57 – Structural integrity and thermal stability of tkGGGPS_wt, cpGGGPS_wt, tkGGGPS_W143A and cpGGGPS_Y143A

(**A** and **B**) Far-UV CD spectra of tkGGGPS_wt, cpGGGPS_wt, tkGGGPS_W143A and cpGGGPS_Y143A (10 µM, subunit concentration). Spectra were recorded in 50 mM potassium phosphate pH 7.5 from 185 nm to 260 nm (d = 1 mm) at room temperature (before) and immediately after heating to 95 °C (after). (**A**) tkGGGPS_wt (tk_wt, deepgreen) and tkGGGPS_W143A (tk_W143A, palegreen) before heating at room temperature (solid lines) and after heating to 95 °C (dashed lines). (**B**) cpGGGPS_wt (cp_wt, deepblue) and cpGGGPS_Y143A (cp_Y143A, cyan) before heating at room temperature (solid lines) and after heating to 95 °C (dashed lines). (**C** and **D**) Thermal denaturation was followed by DSC. Changes in heat capacity of 20 µM protein in 50 mM potassium phosphate pH 7.5 were monitored from 30 °C to 130 °C at a scan rate of 1 °C/min. 50 mM potassium phosphate served as reference. The curves were baseline corrected. (**C**) tkGGGPS_wt (tk_wt, solid line) and tkGGGPS_W143A (tk_W143A, dashed line). (**D**) cpGGGPS_wt (cp_wt, solid line) and cpGGGPS_Y143A (cp_Y143A, dashed line).

At room temperature, all enzymes showed CD spectra indicative for a mixture of α-helices and β-sheets in agreement with the $(\beta\alpha)_8$-barrel fold, which confirmed a native configuration for all proteins.

Following heating to 95 °C, tkGGGPS_W143A showed a CD spectrum slightly shifted to lower wavelengths as it was the case for all mtGGGPS mutants (Figure 49), whereas tkGGGS_wt did not show any difference in CD spectra before and after heating (Figure 57*A*). Moreover, regarding the DSC thermograms an obvious difference between tkGGGPS_wt and tkGGGPS_W143A could be detected. As for the mtGGGPS mutants, tkGGGPS_W143A exhibited a first transition at 97 °C, followed by a second transition at

above 130 °C. tkGGGPS_wt in contrast did not exhibit two transitions but showed one transition at 115 °C, which resulted in immediate aggregation of the protein visible through a vertical decay of the DSC signal after the midpoint of the transition (Figure 57*C*).

In contrast, both proteins cpGGGPS_wt and cpGGGPS_Y143A showed a complete loss of ellipticity after heating (Figure 57*B*). Moreover, cpGGGPS_wt and cpGGGPS_Y143A denatured in a single transition at about 60 °C in DSC (Figure 57*D*). The results of the DSC experiments are summarized in Table 14.

Table 14 – Results from thermal denaturation and heat inactivation of tkGGGPS_wt, tkGGGPS_W143A, cpGGGPS_wt and cpGGGPS_Y143A[1]

Enzyme	Oligomerization state	T_1 [°C] DSC	T_2 [°C] DSC	remaining activity
tkGGGPS_wt	Hexamer	115.0 ± 0.0	-	100°C: 120 min: 86.5 ± 5.1 %
tkGGGPS_W143A	Dimer	96.6 ± 0.1	130.6 ± 1.3	100°C: 15 min: 42.3 ± 3.8 %
cpGGGPS_wt	Hexamer	-	61.3 ± 1.2	50°C: 60 min: 62.5 ± 4.0 %
cpGGGPS_Y143A	Dimer	-	58.4 ± 0.3	50°C: 30 min: 68.0 ± 6.7 %

[1] T_1/T_2: apparent midpoint temperature, thermal transitions with T_1 referring to the first observable transition and T_2 referring to the last observable transition (complete denaturation of the protein), average values of duplicates were calculated (given with standard deviations); remaining activity., average values of triplicates were calculated

To still verify that hexamerization ensures stabilization of activity at higher temperatures, tkGGGPS_wt, tkGGGPS_W143A, cpGGGPS_wt and cpGGGPS_Y143A were subjected to heat inactivation (5.6.16.3, Table 14). In all four cases a fitting of the obtained data points was not feasible (9.5, Figure S3). Therefore values of remaining activity after different incubation times were consulted for comparison. tkGGGPS_wt exhibited after 2 h incubation at 100 °C still about 86 % remaining activity, whereas tkGGGPS_W143A showed only about 42 % remaining activity after 15 min incubation at the same temperature. cpGGGPS_wt and cpGGGPS_Y143A exhibited almost the same remaining activity of about 65 % after incubation at 50 °C with the difference in incubation duration. Whereas cpGGGPS_wt was incubated for 1 h, cpGGGPS_Y143A was already inactivated to the same amount of remaining activity after half of the incubation time. So it can be concluded for these two cases as well, that hexamerization maintains activity at higher temperature (Table 14).

7.3.4 Comparative kinetic analysis of hexameric GGGPS wild type enzymes to their dimeric mutants

The radiometric activity assays of all mtGGGPS mutants confirmed the overall activity of the hexameric and dimeric mtGGGPS mutants, but gave no information on substrate binding or efficieny of the reaction (Figure 52). To determine catalytic parameters and elucidate a possible effect of dimerization, a photometric assay was used (5.6.16.2). As the activity of the monomers is pretty low (Figure 52) and the used photometric assay is not as sensitive as needed for these low activities, only hexamers and the corresponding dimeric mutants were analyzed; mtGGGPS_wt, mtGGGPS_A162E, and mtGGGPS_W141A were chosen for kinetic determination. To have additional pairs of hexamer and dimeric mutant, also the bacterial cpGGGPS_wt and its dimeric mutant cpGGGPS_Y143A, as well as the archaeal tkGGGS_wt and its dimeric mutant tkGGGPS_W143A were examined kinetically (9.7, Figure S4). The results are summarized in Table 15.

Table 15 – Steady-state kinetic parameters at 40 °C of hexameric and dimeric GGGPS variants[1]

Enzyme	Oligomerization state	K_M (G1P) [µM]	K_M (GGPP)[µM]	k_{cat} [s^{-1}]	k_{cat}/K_M (G1P) [$M^{-1}s^{-1}$]	k_{cat}/K_M (GGPP) [$M^{-1}s^{-1}$]
mtGGGPS_wt [a]	Hexamer	13.50 ± 1.00	0.51 ± 0.05	0.34 ± 0.03	2.52 (± 3.00) *10^4	6.67 (± 6.00) *10^5
mtGGGPS_wt	Hexamer	7.50 ± 0.65	1.68 ± 0.10	0.41 ± 0.01	5.29 (± 1.11) *10^4	2.46 (± 0.87) *10^5
mtGGGPS_A162E	Hexamer	11.70 ± 2.39	2.70 ± 0.27	0.92 ± 0.04	5.20 (± 1.54) *10^4	3.40 (± 1.40) *10^5
mtGGGPS_W141A	Dimer	573.00 ± 74.02	3.39 ± 0.56	0.10 ± 0.01	147.30 ± 55.00	0.29 (± 0.18) *10^5
cpGGGPS_wt [b]	Hexamer	22	n.d.	0.12	0.55 * 10^4	n.d.
cpGGGPS_wt	Hexamer	15.02 ± 1.56	0.47 ± 0.06	0.20 ± 0.01	1.05 (± 0.25) *10^4	4.30 (± 1.06) *10^5
cpGGGPS_Y143A	Dimer	276.60 ± 34.10	n.d.	0.08 ± 0.01	269.70 ± 79.00	1.13 (± 0.06) *10^5
tkGGGPS_wt	Hexamer	0.83 ± 0.19	0.31 ± 0.06	0.28 ± 0.02	3.40 (± 1.00) *10^5	0.89 (± 0.19) *10^5
tkGGGPS_W143A	Dimer	0.61 ± 0.05	1.12 ± 0.11	0.19 ± 0.00	3.10 (± 0.77) *10^5	1.60 (± 0.43) *10^5

[1] The parameters were determined with a colorimetric assay for phosphate detection in triplicates (given with standard deviations), n. d.: not determined, k_{cat}/k_M (GGPP) for cpGGGPS_Y143A was calculated from the initial linear slope of the curve as no saturation could be reached.
[a] Soderberg et al., 2001
[b] Beer, 2012

The obtained catalytic parameters of mtGGGPS_wt and cpGGGPS_wt are in good agreement with previous results (Beer, 2012; Soderberg *et al.*, 2001). Two of the three dimeric mutants showed an about 360 fold (mtGGGPS_W141A) or 40 fold (cpGGGPS_Y143A) reduced catalytic efficiency (k_{cat}/k_M (G1P)) compared to the native enzymes, which is in both cases mainly due to an increase in the K_M for G1P (about 76 fold for mtGGGPS_W132A and 18 fold for mtGGGPS_Y143A). Interestingly, tkGGGPS_wt and tkGGGPS_W143A showed virtually identical catalytic parameters.

In summary, mtGGGPS mutants and cpGGGPS mutants clearly showed that hexamerization maintains catalytic activity whereas dimerization leads to decreased binding affinity for the G1P substrate and a consequently reduced catalytic efficiency. G1P is bound via the common standard phosphate binding motif of $(\beta\alpha)_8$-barrel (Nagano *et al.*, 2002; Payandeh *et al.*, 2006; Vega *et al.*, 2003), which in GGGPS enzymes is constitued by side chains and backbone amino groups of the loops $\beta\alpha6$, $\beta\alpha7$ and $\beta\alpha8$ (Peterhoff *et al.*, 2014). Maybe the deletion of the aromatic residue in the interacting surface 3a (α-helix 5´, W141 for mtGGGPS, Y143A for cpGGGPS) somehow affects these loops or the already high flexibility of the loop $\beta\alpha6$ itself as shown in the MD simulation (7.3.2.5) is even higher in the dimeric enzymes due to the loss of stabilizing adjacent protomers thereby leading to a decrease in the binding affinity for G1P.

7.3.5 Comparative characterization of native dimeric and hexameric GGGPS enzymes

To show that the hyopthesis of thermal stabilization and maintainance of activity at high temperatures through higher oligomerization is also applicable to native enzymes, native dimeric and hexameric wild type proteins from group II of the GGGPS enzymes family were chosen for thermal denaturation and irreversible heat inactivation. As native hexameric enzymes, the archaeal mtGGGPS and tkGGGPS as well as two bacterial GGGPS, cpGGGPS and the GGGPS from *Sphingosoma linguale* (slGGGPS) were chosen. As native dimeric enzymes, the archaeal GGGPS from *Thermoplasma acidophilum* (taGGGPS) and two bacterial GGGPS from *Flavobacterium johnsoniae* (fjGGGPS) and *Zunongwangia profunda* (zpGGGPS) were selected. The oligomerization states of these enzymes were determined previously (Peterhoff *et al.*, 2014). To ensure structural integrity of all analyzed enzymes, CD spectra were recorded from 195 nm to 260 nm (5.6.12, Figure 58).

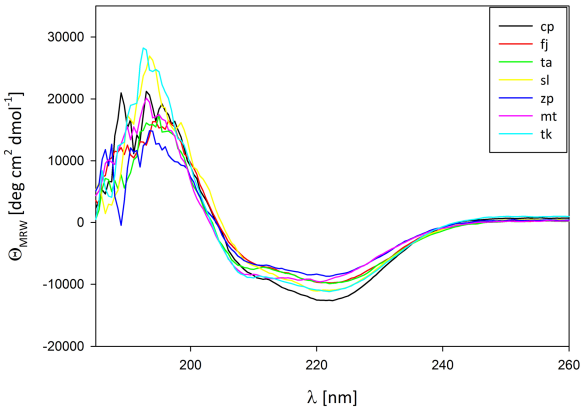


Figure 58 – Structural integrity of native GGGPS enzymes
Far-UV CD spectra of native GGGPS enzymes (10 µM, subunit concentration) were recorded in 50 mM potassium phosphate pH 7.5 from 185 nm to 260 nm (d = 1 mm) at room temperature.

All enzymes showed CD spectra indicative for a mixture of α-helices and β-sheets in agreement with the $(\beta\alpha)_8$-barrel fold, which confirmed a native configuration for all proteins. All enzymes were subjected to thermal heating followed by CD and DSC (5.6.12, 5.6.11, Figure 59). All GGGPS enzymes except mtGGGPS, tkGGGPS and fjGGGPS showed complete unfolding after thermal denaturation to 100 °C (Figure 59*A*). fjGGGPS exhibited the same shifted CD spectrum to lower wavelenghts after thermal heating, which is associated with a loss of activity as observed for the mtGGGPS mutants (Figure 59*B*, Table 16).

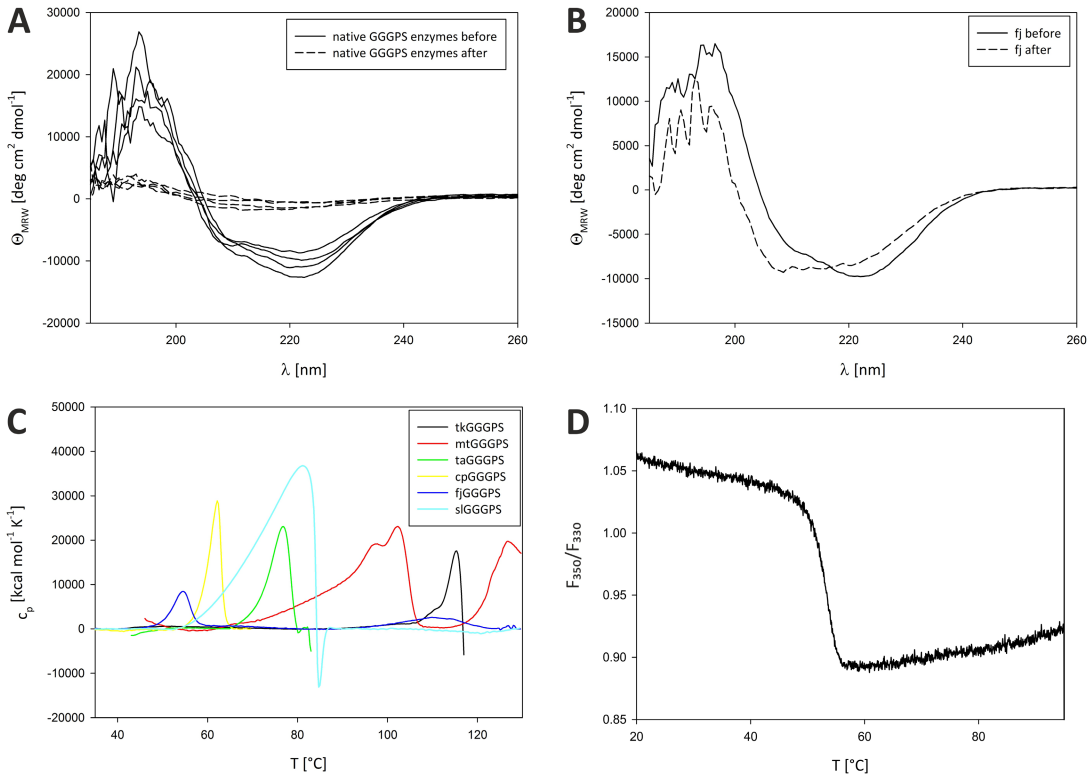


Figure 59 - Thermal stability of native GGGPS enzymes
(**A-B**): Far-UV CD spectra of native GGGPS enzymes (10 µM, subunit concentration) were recorded in 50 mM potassium phosphate pH 7.5 from 185 nm to 260 nm (d = 1 mm) at room temperature (before) and immediately after heating to 95 °C (after). (**B**) native GGGPS enzymes (cpGGGPS, taGGGPS, slGGGPS, zpGGGPS) except fjGGGPS, mtGGGPS and tkGGGS before heating at room temperature (solid lines) and after heating to 95 °C (dashed lines). Each spectrum represents an individual GGGPS enzyme, but for clarity an individual labelling was omitted. (**B**) fjGGGPS before heating at room temperature (solid line) and after heating to 95 °C (dashed line). (**C**) Thermal denaturation was followed by DSC. Changes in heat capacity of 20 µM protein in 50 mM potassium phosphate pH 7.5 were monitored from 30 °C to 130 °C at a scan rate of 1 °C/min. 50 mM potassium phosphate served as reference. The curves were baseline corrected. tkGGGPS is depicted in black, mtGGGPS in red, taGGGPS in green, cpGGGPS in yellow, fjGGGPS in blue and slGGGPS in cyan. (**D**) Thermal denaturation of fjGGGPS was followed by nanoDSF. The change in the ratio of the emission fluorescence at 350 nm to 330 nm of 30 µM protein (subunit concentration) in 50 mM potassium phosphate pH 7.5 was monitored from 20 °C to 95 °C at a scan rate of 1 °C/min.

Also the DSC and nanoDSF thermogram of fjGGGPS (Figure 59*C* and *D*) showed the same behavior as it could be observed for monomeric and dimeric mtGGGPS mutants (Figure 51). The DSC thermogram showed two transitions and the nanoDSF thermogram a decrease of the F_{350}/F_{330}-ratio indicating a blue-shift of the emission fluorescence of tryptophan. This indicates that such kind of unfolding behavior is not unique for mtGGGPS. fjGGGPS was subjected to repetitive thermal denaturation experiments, which revealed almost no reversibility of the observed DSC transitions (data not shown). tkGGGPS, taGGGPS, slGGGPS and cpGGGPS denatured in a single transition in DSC, which

resulted in immediate aggregation of the protein visible through a vertical decay of the DSC signal after the midpoint of the transition (Figure 59*C*).

Comparing the growth temperature optimum of the archaeal organisms, the dimeric taGGGPS is operating at 15 K to 25 K less than the two hexameric GGGPS enzymes (Table 16). The missing physiological pressure for high thermostability is reflected by the thermal denaturation analysis, as taGGGPS showed a complete unfolding at 75 °C, about 25 K lower than the 1st transition of mtGGGPS and about 40 K lower than the unfolding of tkGGGPS (Table 16). For the bacterial enzymes, an equal behavior could be found at least with respect to the oligomerization state. Even though the organisms do not differ in their growth optimum temperature (each around 20 °C), the hexameric enzymes slGGGPS and cpGGGPS showed about 10 K to 30 K higher thermal transitions in DSC measurements, compared to the dimeric enzymes (Figure 59*C*, Table 16). Table 16 summarizes all results.

Table 16 - Characteristica of native dimeric and native hexameric GGGPS enzymes[1,2,3]

Enzyme	Domain	Oligomerization state	Growth optimum [°C]	T_1 [°C] DSC	T_2 [°C] DSC	$t^{app}_{1/2}$ [min]/ remaining activity
tkGGGPS	Archaea	Hexamer	85	-	115.0 ± 0.0	100°C: 120 min: 86.5 ± 5.1 % rem.act.
mtGGGPS	Archaea	Hexamer	75	97.5 ± 0.7	127.0 ± 0.5	100°C: 16.5**
taGGGPS	Archaea	Dimer	59	-	76 ± 1.4	75°C: 7.0
slGGGPS	Bacteria	Hexamer	20	-	83.1 ± 0.89	75°C: 16.4
cpGGGPS	Bacteria	Hexamer	23	-	61.3 ± 1.2	50°C: 60 min: 62.5 ± 4.0 % rem.act.
zpGGGPS	Bacteria	Dimer	25-30	n.d.	n.d.	50°C: 4.5
fjGGGPS	Bacteria	Dimer	20	54.1 ± 0.1	110.0 ± 0.1	50°C: 5.3

[1] T_1/T_2: apparent midpoint temperature, thermal transitions with T_1 referring to the first observable transition and T_2 referring to the last observable transition (complete denaturation of the protein); $t^{app}_{1/2}$: half-life, determined by fitting data with a single exponential decay equation, average values of triplicates were calculated; rem.act.: remaining activity, average values of triplicates were calculated, n.d.: not determined

[2] For clarity T_{1*} of mtGGGPS is not shown

[3] **: 0.1 mM GdmCl was added before heat incubation

As already mentioned above, the fit of the data from the irreversible heat inactivation studies for tkGGGPS and cpGGGPS with a single exponential decay was not feasible, for all other enzymes, however, half-lifes ($t^{app}_{1/2}$) could be determined (9.5, Figure S3). Irreversible heat inactivation showed that the dimeric taGGGPS exhibited an apparent half-life ($t^{app}_{1/2}$) of 7 min at 75 °C, whereas mtGGGPS displayed a $t^{app}_{1/2}$ of 16.5 min at

100 °C. Additionally, tkGGGPS showed still 87 % remaining activity after 2 h incubation at 100 °C, clearly indicating a stabilizing effect through hexamerization. Compared to both dimeric enzymes (zpGGGPS and fjGGGPS), hexameric cpGGGPS still showed 63 % remaining activity after 1 h incubation at 50 °C, whereas the dimeric enzymes exhibited $t^{app}_{1/2}$ of about 5 min at the same temperature. Hexameric slGGGPS showed an about 3 fold increased $t^{app}_{1/2}$ of 16.4 min at a 25 K higher temperature, compared to the native dimers. Summarized, the native hexamers showed higher half-lifes at the same incubation temperature or showed about the same half-life at higher temperatures.

7.4 Conclusion

The members of GGGPS enzyme family exhibit dimeric and hexameric oligomerization states (Peterhoff *et al.*, 2014). To elucidate the role of hexamerization on stability and activity, the GGGPS from *Methanothermobacter thermautotrophicus* was investigated. The constitution of the hexamer and important interfaces for hexamerization were analyzed in mutagenic studies. The asymmetric interface 3a/3b and the symmetric interface 2 have been revealed to contribute equally to the stabilization of the hexameric structure with mutations of interface residues closer to the core of the hexamer (Y105A) being more deleterious for hexamerization than residues lying further apart (R88E and D57R). One amino acid could be found whose mutation disturbs the dimeric interface thereby leading to a collapse of the dimer into monomers (A162E). The thermal stability of the resulting monomeric, dimeric and hexameric mtGGGPS mutants as well as the mtGGGPS_wt was determined via CD, DSC, nanoDSF measurements and irreversible heat inactivation. A higher oligomerization state ensured a higher thermostability and contributed to maintaining activity after incubation at higher temperatures or after incubation for longer time periods at the same temperature. A possible unfolding pathway for mtGGGPS could be proposed. First, a transition of the native structure to a stable partially folded intermediate occurs, which is inactive but maintains almost 80-90 % of the overall secondary structure. Depending on the oligomerization state, this 1st transition happens at a temperature around 50 °C (monomeric mutants), 60 °C (dimeric mutants) and 100 °C (hexameric mtGGGPS_wt and mutants). Only at elevated temperatures of around 120 °C, which is in common for all mtGGGPS_wt and its mutants, a second transition presumably leads to a complete denaturation of the protein.

Two other couples of native hexamer and dimeric mutant (tkGGGPS and tkGGGPS_W143A as well as cpGGGPS_wt and cpGGGPS_Y143A) were analyzed for their behavior regarding thermal stability and activity. Again, the hexameric proteins showed higher thermal stability and maintained activity after incubation at higher temperatures, while the dimeric proteins lost their activity in about half the time compared to the hexameric enzymes. Catalytic parameters were obtained for mtGGGPS_wt, mtGGGPS_A162E, mtGGGPS_W141A, tkGGGPS_wt, tkGGGPS_W143A, cpGGGPS_wt and cpGGGPS_Y143A. Except for tkGGGPS_wt and tkGGGPS_W143A, the dimeric enzymes showed decreased catalytic efficiencies, especially due to an increase of the K_M for G1P. To verify the observed stabilizing effect of hexamerization also for native enzymes, native dimeric and hexameric GGGPS enzymes from the GGGPS enzyme family group II were analyzed biophysically regarding their thermal stability. It could be shown that native dimers are less thermostable and less active after incubation at high temperatures or after long incubation at temperatures near their melting temperature compared to native hexamers. In almost all examined cases, hexamerization ensured structural integrity as a prerequisite for thermal stability and activity.

7.5 Ongoing research and future work

GGGPS enzymes play an important role in the biosynthesis of archaea-like membrane lipids and are considered to play a key role in the early separation of the domain of Archaea (Boucher *et al.*, 2004; Koga *et al.*, 1998; Payandeh *et al.*, 2006; Payandeh and Pai, 2007; Pereto *et al.*, 2004). To the best of knowledge, no direct protein ancestor of GGGPS enzymes has been proposed up to date. Payandeh and Pai (2007) postulated based on sequence and structural similarities that GGGPS evolved by gene duplication and subsequent fusion of a $(\beta\alpha)_4$-barrel ancestor protein or by duplication of an ancient *hisF*-like gene. Kristina Heyn (AG Prof. R. Merkl, University of Regensburg) is currently working on an ancestral protein reconstruction of GGGPS enzymes, which could give an interesting aspect on how the GGGPS activity was established in the early steps of evolution. The reconstructed ancestor protein can then be analyzed on how the ancestral GGGPS enzymes differ from the recent GGGPS enzymes depending on oligomerization state, calatytic activity and substrate specificity. By expression of GGGPS sequences lying at

branch points of the reconstructed phylogenetic tree, the evolution of GGGPS enzymes to recent GGGPS enzymes especially on basis of their oligomerization state can be chased.

In the MD simulations α-helix 4, α-helix 5´and α-helix 8 have exhibited the most significant change in flexibility upon heating. Payandeh *et al.* (2006) illustrated using average temperature factors that a bound substrate induces stabilization of α-helix 2, α-helix 3* and α-helix 4 but increases mobility of α-helix 8. It would be interesting to repeat the MD simulation with a bound substrate and observe the changes in fluctuation behavior regarding the above announced α-helices.

As the crystallization of mtGGGPS_A162E, mtGGGPS_I107E and mtGGGPS_W141A did not show the physiological oligomerization state of the proteins, another structure determination experiment could be performed, transmission electron mycroscopy (TEM). Here, also low protein concentrations can be analyzed and therefore a structure of GGGPS enzymes in a physiological state could be obtained, albeit with lower resolution. With this technique, also more GGGPS family enzymes could be analyzed regarding their oligomerization state to enrich the knowledge of oligomerization state distribution in the GGGPS enzyme family.

Another work which is still in progress, is the construction of a hexamer starting from a native dimeric GGGPS enzyme. fjGGGPS was chosen as model. By sequence comparison with hexameric GGGPS enzymes and superposition with mtGGGPS, seven residues (E113G, A144T, A146S, R147Y, V148I, N155P and N158K) have been selected for mutagenesis (Figure 60).

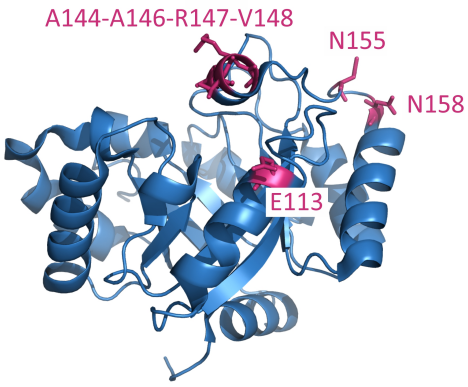


Figure 60 - Localization of residues chosen for mutagenesis in fjGGGPS
Constitution of one protomer of the dimeric fjGGGPS (pdb-code: 4jej) is shown as ribbon diagram. The seven amino acids chosen for mutagenesis are shown as sticks and colored magenta.

E113 is localized at the beginning of α-helix 4. Because this position is relevant for forming interface 2 and glutamate could cause steric hindrance to the opposite protomer during hexamerization, this amino acid is currently mutated to glycine. Glycine was chosen, as mtGGGPS also exhibits a glycine at this position. A144, A146, R147 and V148 are localized in the α-helix 5´ participating in the asymmetric interface 3a/3b and are therefore mutated to the same amino acids present in the hexameric enzymes. R147 is mutated to tyrosine as it is localized at the position of the aromatic amino acid needed for hexamerization (W141 in mtGGGPS, Y143 in cpGGGPS and W143 in tkGGGPS). N155 and N158 lie in the loop α5´α5 and are mutated to proline and lysine, respectively. N158 is localized at the position of the highly correlated K152 of hexameric enzymes, which will be discussed below and is therefore mutated to K. At the moment, the cloning of fjGGGPS coding for the seven exchanged amino acids (fjGGGPS_7Mut) as well as first test expressions were successful, therefore large scale expression and analysis of the oligomerization state as well as thermal stability is planned. If the oligomerization state indeed assembles to a hexameric structure, the mutated amino acids will be one by one mutated back to the native amino acid in fjGGGPS, thereby elucidating the most important residues for hexamerization.

Another interesting approach is elucidating the function of K152 in mtGGGPS. During analysis of group II enzymes, Peterhoff *et al.* (2014) ascertained that subgroups with a hexameric representative exhibited a strong correlation between the presence of the aromatic amino acid of interacting surface 3a (W141 in mtGGGPS) and the occurrence of a positively charged residue (K152 in mtGGGPS). A possible involvement of this lysine in the stabilization of the hexamer was assumed, but it was already shown by analytical size exclusion chromatography that the mutation of K152 to alanine in mtGGGPS did not affect the hexamerization (data not shown). If this in hexameric representatives highly conserved amino acid does not affect the oligomerization state it could possibly affect the catalytic activity of the enzymes at high temperature. Investigating the localization of K152 in mtGGGPS via PyMol, it can be seen that K152 is pointing in the direction of the loop βα6 (Figure 61).

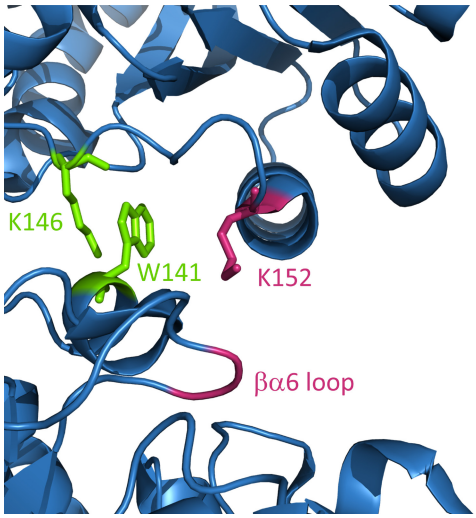


Figure 61 - Localization of K152 in mtGGGPS
Detail of the mtGGGPS structure (pdb-code: 4mm1). Interface residues from interacting surfaces 3a (W141) and 3b (K146) are shown as green sticks. The highly conserved K152 is shown as sticks in magenta. The two highly conserved residues S177 and G178 of loop βα6 participating in G1P binding are colored magenta.

As already mentioned above, this loop βα6, mainly constituted by S177 and G178 in mtGGGPS, is participating in G1P binding by forming a part of the common standard phosphate binding motif and therefore, a possible involvment of K152 in coordinating or stabilizing this loop and thereby influencing G1P binding could be assumed. To verify this assumption, the mtGGGPS_K152A mutant will be analyzed regarding its catalytic parameters. If the hexameric mtGGGPS_K152A also shows a decreased catalytic efficiency and a decreased substrate binding affinity for G1P like the dimeric mtGGGPS_W141A and cpGGGPS_Y143A, its role as stabilizer of the catalytic activity in GGGPS enzymes is verified.

In hexameric cpGGGPS, the tyrosine of the symmetric interface 2 (Y105 in mtGGGPS) is not conserved and therefore the question arose how interface 2 is stabilized in this hexamer. Therefore, a monomerizing mutation (A162E in mtGGGPS, A164E in cpGGGPS) has been introduced into cpGGGPS_wt and into the dimeric cpGGGPS_Y143A. Initial test expressions and analytical SEC let speculate that the cpGGGPS hexamer is only stabilized through interface 3a/3b, but is lacking the cross connecting symmetric interface 2, because cpGGGPS_A164E as well as cpGGGPS_A164E_Y143A showed monomeric peaks in the preliminary SEC experiment. This assumption has to be verified through large scale expressions and further SEC.

8 REFERENCES

Abdel-Fattah, W.R., Chen, Y., Eldakak, A., and Hulett, F.M. (2005). *Bacillus subtilis* phosphorylated PhoP: direct activation of the E(sigma)A- and repression of the E(sigma)E-responsive phoB-PS+V promoters during pho response. *Journal of bacteriology **187***, 5166-5178.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr **58***, 1948-1954.

Ahn, V.E., Lo, E.I., Engel, C.K., Chen, L., Hwang, P.M., Kay, L.E., Bishop, R.E., and Prive, G.G. (2004). A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J **23***, 2931-2941.

Bachmann, B.J. (1990). Linkage map of *Escherichia coli* K-12, edition 8. *Microbiol Rev **54***, 130-197.

Badger, J., Sauder, J.M., Adams, J.M., Antonysamy, S., Bain, K., Bergseid, M.G., Buchanan, S.G., Buchanan, M.D., Batiyenko, Y., Christopher, J.A.*, et al.* (2005). Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins **60***, 787-796.

Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E.*, et al.* (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature **255***, 609-614.

Beer, B. (2012). Archaeen-typische Lipide in Bakterien: Vergleichende Charakteristiken von Prenyltransferasen. In Institut für Biophysik und physikalische Biochemie, Lehrstuhl Biochemie II (Regensburg, Universität Regensburg), p. 96.

Benson, A.A., Bassham, J.A., Calvin, M., Goodale, T.C., Haas, V.A., and Stepka, W. (1950). The Path of Carbon in Photosynthesis. V. Paper Chromatography and Radioautography of the Products. *J. Am. Chem. Soc. **72***, 1710-1718.

Bligh, E.G., and Dyer, W.J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol. **37***, 911-917.

Boos, W., Ferenci, T., and Shuman, H.A. (1981). Formation and excretion of acetylmaltose after accumulation of maltose in *Escherichia coli*. *Journal of bacteriology* ***146***, 725-732.

Boucher, Y. (2007). Lipids: biosynthesis, function, and evolution. In Archaea, Molecular and Celluar Biology, R. Cavicchioli, ed. (Washington, D.C.: ASM Press).

Boucher, Y., and Doolittle, W.F. (2000). The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Molecular Microbiology* ***37***, 703-716.

Boucher, Y., Kamekura, M., and Doolittle, W.F. (2004). Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol. Microbiol.* ***52***, 515-527.

Brand, B., and Boos, W. (1991). Maltose transacetylase of *Escherichia coli*. Mapping and cloning of its structural, gene, *mac*, and characterization of the enzyme as a dimer of identical polypeptides with a molecular weight of 20,000. *J Biol Chem* ***266***, 14113-14118.

Caforio, A., Jain, S., Fodran, P., Siliakus, M., Minnaard, A.J., van der Oost, J., and Driessen, A.J. (2015). Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in *Escherichia coli*. *Biochem J* ***470***, 343-355.

Carballeira, N.M., Reyes, M., Sostre, A., Huang, H., Verhagen, M.F., and Adams, M.W. (1997). Unusual fatty acid compositions of the hyperthermophilic archaeon *Pyrococcus furiosus* and the bacterium *Thermotoga maritima*. *J Bacteriol* ***179***, 2766-2768.

Chen, A., Zhang, D., and Poulter, C.D. (1993). (S)-geranylgeranylglyceryl phosphate synthase. Purification and characterization of the first pathway-specific enzyme in archaebacterial membrane lipid biosynthesis. *J. Biol. Chem.* ***268***, 21701-21705.

Chester, N., and Marshak, D.R. (1993). Dimethyl sulfoxide-mediated primer Tm reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal Biochem* ***209***, 284-290.

Consden, R., Gordon, A.H., and Martin, A.J. (1944). Qualitative analysis of proteins: a partition chromatographic method using paper. *Biochem. J.* ***38***, 224-232.

Copley, S.D. (2003). Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr Opin Chem Biol* ***7***, 265-272.

Daiyasu, H., Hiroike, T., Koga, Y., and Toh, H. (2002). Analysis of membrane stereochemistry with homology modeling of *sn*-glycerol-1-phosphate dehydrogenase. *Protein Eng* ***15***, 987-995.

Daiyasu, H., Kuma, K., Yokoi, T., Morii, H., Koga, Y., and Toh, H. (2005). A study of archaeal enzymes involved in polar lipid synthesis linking amino acid sequence information, genomic contexts and lipid composition. *Archaea* ***1***, 399-410.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* ***35***, W375-383.

Doud, E.H., Perlstein, D.L., Wolpert, M., Cane, D.E., and Walker, S. (2011). Two Distinct Mechanisms for TIM Barrel Prenyltransferases in Bacteria. *J. Am. Chem. Soc.* ***133***, 1270-1273.

Doyle, C.M., Rumfeldt, J.A., Broom, H.R., Broom, A., Stathopulos, P.B., Vassall, K.A., Almey, J.J., and Meiering, E.M. (2013). Energetics of oligomeric protein folding and association. *Arch Biochem Biophys* ***531***, 44-64.

Dumelin, C.E., Chen, Y., Leconte, A.M., Chen, Y.G., and Liu, D.R. (2012). Discovery and biological characterization of geranylated RNA in bacteria. *Nat Chem Biol* ***8***, 913-919.

Edidin, M. (2003). Timeline - Lipids on the frontier: a century of cell-membrane bilayers. *Nat Rev Mol Cell Bio* ***4***, 414-418.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* ***60***, 2126-2132.

Fiévet, J.B., Dillmann, C., Curien, G., and de Vienne, D. (2006). Simplified modelling of metabolic pathways for flux prediction and optimization: lessons from an *in vitro* reconstruction of the upper part of glycolysis. *Biochem J* ***396***, 317-326.

Fischer, E. (1920). Wanderung von Acyl bei den Glyceriden. *Ber. d. deutsch. chem. Ges.* ***53***, 1621-1633.

Fischer, W., and Arneth-Seifert, D. (1998). D-Alanylcardiolipin, a major component of the unique lipid pattern of *Vagococcus fluvialis*. *J Bacteriol* ***180***, 2950-2957.

Freire, E. (1995). Differential Scanning Calorimetry. In Protein stability and folding theory and practice, B.A. Shirley, ed. (Totowa, NJ: Humana Press), pp. 191-218.

Galinier, A., Kravanja, M., Engelmann, R., Hengstenberg, W., Kilhoffer, M.C., Deutscher, J., and Haiech, J. (1998). New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proceedings of the National Academy of Sciences of the United States of America* ***95***, 1823-1828.

Gattinger, A., Schloter, M., and Munch, J.C. (2002). Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling. *FEMS Microbiol Lett* ***213***, 133-139.

Glansdorff, N., Xu, Y., and Labedan, B. (2008). The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct* ***3***, 29.

Goldfine, H., and Langworthy, T.A. (1988). A growing interest in bacterial ether lipids. *Trends Biochem Sci* ***13***, 217-221.

Good, N.E., Winget, G.D., Winter, W., Connolly, T.N., Izawa, S., and Singh, R.M. (1966). Hydrogen ion buffers for biological research. *Biochemistry* ***5***, 467-477.

Goodsell, D.S., and Olson, A.J. (2000). Structural symmetry and protein function. *Annu Rev Bioph Biom* ***29***, 105-153.

Groemping, Y., and Reinstein, J. (2001). Folding properties of the nucleotide exchange factor GrpE from *Thermus thermophilus*: GrpE is a thermosensor that mediates heat shock response. *J Mol Biol* ***314***, 167-178.

Guldan, H. (2007). Charakterisierung der *sn*-Glycerin-1-Phosphat-abhängigen Enzyme AraM und PcrB aus *Bacillus subtilis*. Diplomarbeit, Universität Regensburg.

Guldan, H. (2010). Nachweis Archaea-typischer Lipide in Bacteria über die Aufklärung der Funktion von AraM und PcrB aus *Bacillus Subtilis*. (Regensburg, Universität Regensburg).

Guldan, H., Matysik, F.M., Bocola, M., Sterner, R., and Babinger, P. (2011). Functional Assignment of an Enzyme that Catalyzes the Synthesis of an Archaea-Type Ether Lipid in Bacteria. *Angew Chem Int Ed Engl* ***50***, 8188-8191.

Guldan, H., Sterner, R., and Babinger, P. (2008). Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni(2+)-dependent AraM from *Bacillus subtilis*. *Biochemistry* ***47***, 7376-7384.

Guo, R.T., Ko, T.P., Chen, A.P., Kuo, C.J., Wang, A.H., and Liang, P.H. (2005). Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: roles of the metal ion and conserved residues in catalysis. *J Biol Chem* ***280***, 20762-20774.

Guo, R.T., Kuo, C.J., Chou, C.C., Ko, T.P., Shr, H.L., Liang, P.H., and Wang, A.H. (2004). Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic *Thermotoga maritima* and mechanism of product chain length determination. *J Biol Chem* ***279***, 4903-4912.

Guzman-Luna, V., and Garza-Ramos, G. (2012). The folding pathway of glycosomal triosephosphate isomerase: structural insights into equilibrium intermediates. *Proteins* ***80***, 1669-1682.

Han, J.S., and Ishikawa, K. (2005). Active site of Zn^{2+}-dependent sn-glycerol-1-phosphate dehydrogenase from *Aeropyrum pernix* K1. *Archaea* ***1***, 311-317.

Holm, L., and Sander, C. (1996). Mapping the protein universe. *Science* ***273***, 595-603.

Hulett, F.M., Kim, E.E., Bookstein, C., Kapp, N.V., Edwards, C.W., and Wyckoff, H.W. (1991). *Bacillus subtilis* alkaline phosphatases III and IV. Cloning, sequencing, and comparisons of deduced amino acid sequence with *Escherichia coli* alkaline phosphatase three-dimensional structure. *The Journal of biological chemistry* ***266***, 1077-1084.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* ***96***, 23-28.

Jahn, U., Summons, R., Sturt, H., Grosjean, E., and Huber, H. (2004). Composition of the lipids of *Nanoarchaeum equitans* and their origin from its host *Ignicoccus sp.* strain KIN4/I. *Arch Microbiol* ***182***, 404-413.

Jain, S., Caforio, A., and Driessen, A.J. (2014). Biosynthesis of archaeal membrane ether lipids. *Front Microbiol **5***, 641.

Jost, M., Zocher, G., Tarcz, S., Matuschek, M., Xie, X., Li, S.M., and Stehle, T. (2010). Structure-function analysis of an enzymatic prenyl transfer reaction identifies a reaction chamber with modifiable specificity. *J Am Chem Soc **132***, 17849-17858.

Kabisch, J., Thurmer, A., Hubel, T., Popper, L., Daniel, R., and Schweder, T. (2013). Characterization and optimization of *Bacillus subtilis* ATCC 6051 as an expression host. *J Biotechnol **163***, 97-104.

Kabsch, W. (1993). Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J Appl Crystallogr **26***, 795-800.

Kandler, O. (1994a). Cell-Wall Biochemistry and 3-Domain Concept of Life. *Systematic and Applied Microbiology **16***, 501-509.

Kandler, O. (1994b). Cell-Wall Biochemistry in Archaea and Its Phylogenetic Implications. *J Biol Phys **20***, 165-169.

Kandler, O., and Konig, H. (1998). Cell wall polymers in Archaea (Archaebacteria). *Cell Mol Life Sci **54***, 305-308.

Kaneda, K., Kuzuyama, T., Takagi, M., Hayakawa, Y., and Seto, H. (2001). An unusual isopentenyl diphosphate isomerase found in the mevalonate pathway gene cluster from *Streptomyces sp*. strain CL190. *Proc Natl Acad Sci U S A **98***, 932-937.

Kates, M. (1986). Techniques of lipidology: isolation, analysis, and identification of lipids. In Laboratory techniques in biochemistry and molecular biology vol. 3, part 2, R.H. Burdon, and P.H. van Knippenberg, eds. (Amsterdam: Elsevier), pp. 1-464.

Kates, M. (1993). Membrane lipids in Archaea. In The biochemistry of archaea, M. Kates, D.J. Kushner, and A.T. Matheson, eds. (Amsterdam: Elsevier), pp. 261-295.

Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. *Bba-Proteins Proteom **1751***, 119-139.

Kilu, W. (2014). Strukturelle Grundlagen der Hexamerisierung der GGGPS aus *Methanothermobacter thermautotrophicus*. Bachelorarbeit, Universität Regensburg. In Bachelorarbeit (Universität Regensburg).

Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P.*, et al.* (2003). Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. U. S. A.* ***100***, 4678-4683.

Kodali, D.R., Tercyak, A., Fahey, D.A., and Small, D.M. (1990). Acyl migration in 1,2-dipalmitoyl-*sn*-glycerol. *Chem Phys Lipids* ***52***, 163-170.

Koga, Y. (2011). Early Evolution of Membrane Lipids: How did the Lipid Divide Occur? *J Mol Evol* ***72***, 274-282.

Koga, Y., Kyuragi, T., Nishihara, M., and Sone, N. (1998). Did archaeal and bacterial cells arise independently from noncellular precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *J Mol Evol* ***46***, 54-63.

Koga, Y., and Morii, H. (2007). Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol Mol Biol Rev* ***71***, 97-120.

Koga, Y., Nishihara, M., Morii, H., and Akagawa-Matsushita, M. (1993). Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. *Microbiol. Rev.* ***57***, 164-182.

Kohlhoff, M., Dahm, A., and Hensel, R. (1996). Tetrameric triosephosphate isomerase from hyperthermophilic Archaea. *FEBS Lett* ***383***, 245-250.

Kolbe, M., Besir, H., Essen, L.O., and Oesterhelt, D. (2000). Structure of the light-driven chloride pump halorhodopsin at 1.8 A resolution. *Science* ***288***, 1390-1396.

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

Lang, D., Thoma, R., Henn-Sax, M., Sterner, R., and Wilmanns, M. (2000). Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. *Science* ***289***, 1546-1550.

Lange, B.M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc Natl Acad Sci U S A* ***97***, 13172-13177.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Biotechnology* ***11***, 187-193.

Le Gouill, C., Parent, J.L., Rola-Pleszczynski, M., and Stankova, J. (1994). Analysis of recombinant plasmids by a modified alkaline lysis method. *Anal Biochem* ***219***, 164.

Leaver, M., Dominguez-Cuevas, P., Coxhead, J.M., Daniel, R.A., and Errington, J. (2009). Life without a wall or division machine in *Bacillus subtilis*. *Nature* ***457***, 849-853.

Lee, T.C., Malone, B., Blank, M.L., Fitzgerald, V., and Snyder, F. (1990). Regulation of the synthesis of platelet-activating factor and its inactive storage precursor (1-alkyl-2-acyl-*sn*-glycero-3-phosphocholine) from 1-alkyl-2-acetyl-*sn*-glycerol by rabbit platelets. *J Biol Chem* ***265***, 9181-9187.

Lee, T.C., Malone, B., and Snyder, F. (1986). A new *de novo* pathway for the formation of 1-alkyl-2-acetyl-*sn*-glycerols, precursors of platelet activating factor. Biochemical characterization of 1-alkyl-2-lyso-*sn*-glycero-3-P:acetyl-CoA acetyltransferase in rat spleen. *J Biol Chem* ***261***, 5373-5377.

Liang, P.H., Ko, T.P., and Wang, A.H. (2002). Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* ***269***, 3339-3354.

Linde, M. (2013). Identifizierung und Charakterisierung der HepGP-prozessierenden Enzyme aus *Bacillus subtilis*. Masterarbeit, Universität Regensburg.

Linde, M., Peterhoff, D., Sterner, R., and Babinger, P. (2016). Identification and Characterization of Heptaprenylglyceryl Phosphate Processing Enzymes in *Bacillus subtilis*. *J Biol Chem* ***291***, 14861-14870.

Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2002). Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci* ***11***, 371-380.

Lo Leggio, L., Dal Degan, F., Poulsen, P., Andersen, S.M., and Larsen, S. (2003). The structure and specificity of *Escherichia coli* maltose acetyltransferase give new insight into the LacA family of acyltransferases. *Biochemistry* ***42***, 5225-5235.

Lombard, J., Lopez-Garcia, P., and Moreira, D. (2012a). An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids. *Mol Biol Evol* ***29***, 3261-3265.

Lombard, J., Lopez-Garcia, P., and Moreira, D. (2012b). Phylogenomic investigation of phospholipid synthesis in archaea. *Archaea* ***2012***, 630910.

Long, S.B., Casey, P.J., and Beese, L.S. (2002). Reaction path of protein farnesyltransferase at atomic resolution. *Nature* ***419***, 645-650.

Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1989). Correct folding of circularly permuted variants of a beta alpha barrel enzyme *in vivo*. *Science* ***243***, 206-210.

Lyubachevskaya, G., and Boyle-Roden, E. (2000). Kinetics of 2-monoacylglycerol acyl migration in model chylomicra. *Lipids* ***35***, 1353-1358.

Martin, W., and Russell, M.J. (2003). On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* ***358***, 59-85.

Matsumi, R., Atomi, H., Driessen, A.J., and van der Oost, J. (2011). Isoprenoid biosynthesis in Archaea - biochemical and evolutionary implications. *Res. Microbiol.* ***162***, 39-52.

Mixcoha-Hernandez, E., Moreno-Vargas, L.M., Rojo-Dominguez, A., and Benitez-Cardoza, C.G. (2007). Thermal-unfolding reaction of triosephosphate isomerase from *Trypanosoma cruzi*. *Protein J* ***26***, 491-498.

Mullis, K.B., and Faloona, F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol* ***155***, 335-350.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* ***53***, 240-255.

Nagano, N., Orengo, C.A., and Thornton, J.M. (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.

Nemoto, N., Oshima, T., and Yamagishi, A. (2003). Purification and characterization of geranylgeranylglyceryl phosphate synthase from a thermoacidophilic archaeon, *Thermoplasma acidophilum*. *J. Biochem. (Tokyo)* ***133***, 651-657.

Nishihara, M., Nagahama, S., Ohga, M., and Koga, Y. (2000). Straight-chain fatty alcohols in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Extremophiles* ***4***, 275-277.

Ogura, K., and Koyama, T. (1998). Enzymatic Aspects of Isoprenoid Chain Elongation. *Chem Rev* ***98***, 1263-1276.

Ohnuma, S., Hirooka, K., Hemmi, H., Ishida, C., Ohto, C., and Nishino, T. (1996). Conversion of product specificity of archaebacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. *J Biol Chem* ***271***, 18831-18837.

Oldfield, E., and Lin, F.Y. (2012). Terpene biosynthesis: modularity rules. *Angew Chem Int Ed Engl* ***51***, 1124-1137.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* ***4***, 2411-2423.

Palsuledesai, C.C., and Distefano, M.D. (2015). Protein prenylation: enzymes, therapeutics, and biotechnology applications. *ACS Chem Biol* ***10***, 51-62.

Payandeh, J., Fujihashi, M., Gillon, W., and Pai, E.F. (2006). The crystal structure of (S)-3-*O*-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme. *J. Biol. Chem.* ***281***, 6070-6078.

Payandeh, J., and Pai, E.F. (2007). Enzyme-driven speciation: crystallizing Archaea via lipid capture. *J Mol Evol* ***64***, 364-374.

Percy, M.G., and Grundling, A. (2014). Lipoteichoic acid synthesis and function in Gram-positive bacteria. *Annu Rev Microbiol* ***68***, 81-100.

Pereto, J., Lopez-Garcia, P., and Moreira, D. (2004). Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem. Sci.* ***29***, 469-477.

Perica, T., Chothia, C., and Teichmann, S.A. (2012). Evolution of oligomeric state through geometric coupling of protein interfaces. *Proc Natl Acad Sci U S A* ***109***, 8127-8132.

Peterhoff, D., Beer, B., Rajendran, C., Kumpula, E.P., Kapetaniou, E., Guldan, H., Wierenga, R.K., Sterner, R., and Babinger, P. (2014). A comprehensive analysis of the geranylgeranylglyceryl phosphate synthase enzyme family identifies novel members and reveals mechanisms of substrate specificity and quaternary structure organization. *Mol Microbiol* ***92***, 885-899.

Peterhoff, D., Zellner, H., Guldan, H., Merkl, R., Sterner, R., and Babinger, P. (2012). Dimerization determines substrate specificity of a bacterial prenyltransferase. *Chembiochem* ***13***, 1297-1303.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* ***258***, 598-599.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A., and Noble, M. (2004). Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* ***60***, 2288-2294.

Ramstein, J., Hervouet, N., Coste, F., Zelwer, C., Oberto, J., and Castaing, B. (2003). Evidence of a thermal unfolding dimeric intermediate for the *Escherichia coli* histone-like HU proteins: thermodynamics and structure. *J Mol Biol* ***331***, 101-121.

Ren, F., Feng, X., Ko, T.P., Huang, C.H., Hu, Y., Chan, H.C., Liu, Y.L., Wang, K., Chen, C.C., Pang, X.*, et al.* (2013). Insights into TIM-barrel prenyl transferase mechanisms: crystal structures of PcrB from *Bacillus subtilis* and *Staphylococcus aureus*. *Chembiochem* ***14***, 195-199.

Ren, F., Ko, T.P., Feng, X., Huang, C.H., Chan, H.C., Hu, Y., Wang, K., Ma, Y., Liang, P.H., Wang, A.H.*, et al.* (2012). Insights into the mechanism of the antibiotic-synthesizing enzyme MoeO5 from crystal structures of different complexes. *Angew Chem Int Ed Engl* ***51***, 4157-4160.

Roderick, S.L. (2005). The lac operon galactoside acetyltransferase. *C R Biol* ***328***, 568-575.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* ***230***, 1350-1354.

Sambrook, J., Frisch, E. E., Maniatis, T. (1989). Molecular cloning: a laboratory manual. (New York: Cold Spring Harbour).

Schmid, F.X. (1997). Optical spectroscopy to characterize protein conformation and conformational changes. In Protein structure : a practical approach, T.E. Creighton, ed. (Oxford {[u.a.]: IRL Press), pp. 259-295.

Schramm, A., Kohlhoff, M., and Hensel, R. (2001). Triose-phosphate isomerase from *Pyrococcus woesei* and *Methanothermus fervidus*. *Methods Enzymol* ***331***, 62-77.

Sedlak, E., Varhac, R., Musatov, A., and Robinson, N.C. (2014). The kinetic stability of cytochrome C oxidase: effect of bound phospholipid and dimerization. *Biophys J* ***107***, 2941-2949.

Seisenberger, C. (2017). Aufklärung der Struktur von YvoF und physiologischen Funktion von G1P-basierten Etherlipiden in *Bacillus subtilis.* Masterarbeit, Universität Regensburg.

Sharp, P.A., Sugden, B., and Sambrook, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose--ethidium bromide electrophoresis. *Biochemistry* ***12***, 3055-3063.

Smit, A., and Mushegian, A. (2000). Biosynthesis of isoprenoids via mevalonate in Archaea: the lost pathway. *Genome Res* ***10***, 1468-1484.

Soderberg, T., Chen, A., and Poulter, C.D. (2001). Geranylgeranylglyceryl phosphate synthase. Characterization of the recombinant enzyme from *Methanobacterium thermoautotrophicum*. *Biochemistry* ***40***, 14847-14854.

Spizizen, J. (1958). Transformation of Biochemically Deficient Strains of Bacillus Subtilis by Deoxyribonucleate. *Proc Natl Acad Sci U S A* ***44***, 1072-1078.

Stahl, E. (1967). Dünnschicht-Chromatographie ein Laboratoriumshandbuch, 2., gänzlich neubearb. und stark erw. Aufl. edn (Berlin: Springer-Verlag).

Steck, T.L., and Yu, J. (1973). Selective solubilization of proteins from red blood cell membranes by protein perturbants. *J Supramol Struct* ***1***, 220-232.

Sterner, R., and Höcker, B. (2005). Catalytic versatility, stability, and evolution of the (betaalpha)8-barrel enzyme fold. *Chem Rev* ***105***, 4038-4055.

Sterner, R., Kleemann, G.R., Szadkowski, H., Lustig, A., Hennig, M., and Kirschner, K. (1996). Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer. *Protein Sci* ***5***, 2000-2008.

Suarez, A.S.G., Stefan, A., Lemma, S., Conte, E., and Hochkoeppler, A. (2012). Continuous enzyme-coupled assay of phosphate- or pyrophosphate-releasing enzymes. *Biotechniques* ***53***, 99-103.

Takahashi, I., and Ogura, K. (1981). Farnesyl pyrophosphate synthetase from Bacillus subtilis. *J Biochem* ***89***, 1581-1587.

Tarafdar, P.K., Vedantam, L.V., Sankhala, R.S., Purushotham, P., Podile, A.R., and Swamy, M.J. (2014). Oligomerization, conformational stability and thermal unfolding of Harpin, HrpZPss and its hypersensitive response-inducing c-terminal fragment, C-214-HrpZPss. *PLoS One* ***9***, e109871.

Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C., and Poulter, C.D. (1996). Regulation of product chain length by isoprenyl diphosphate synthases. *Proc Natl Acad Sci U S A* ***93***, 15018-15023.

Thorne, K.J., and Kodicek, E. (1966). The structure of bactoprenol, a lipid formed by lactobacilli from mevalonic acid. *Biochem J* ***99***, 123-127.

Tornabene, T.G., Wolfe, R.S., Balch, W.E., Holzer, G., Fox, G.E., and Oro, J. (1978). Phytanyl-glycerol ethers and squalenes in the archaebacterium *Methanobacterium thermoautotrophicum*. *J Mol Evol* ***11***, 259-266.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* ***76***, 4350-4354.

Tsuji, F., Ishihara, A., Kurata, K., Nakagawa, A., Okada, M., Kitamura, S., Kanamaru, K., Masuda, Y., Murakami, K., Irie, K.*, et al.* (2012). Geranyl modification on the

tryptophan residue of ComXRO-E-2 pheromone by a cell-free system. *FEBS Lett* ***586***, 174-179.

Turon, F., Bonnot, F., Caro, Y., Pina, M., and Graille, J. (2003). Acyl migration incidence on accuracy of a triacylglycerol regioanalysis--a theoretical evaluation. *Chem Phys Lipids* ***125***, 41-48.

Vandermoten, S., Haubruge, E., and Cusson, M. (2009). New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. *Cell Mol Life Sci* ***66***, 3685-3695.

Vega, M.C., Lorentzen, E., Linden, A., and Wilmanns, M. (2003). Evolutionary markers in the $(b/a)_8$-barrel fold. *Curr. Opin. Chem. Biol.* ***7***, 694-701.

Vieille, C., and Zeikus, G.J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* ***65***, 1-43.

Vuorio, R., Harkonen, T., Tolvanen, M., and Vaara, M. (1994). The novel hexapeptide motif found in the acyltransferases LpxA and LpxD of lipid A biosynthesis is conserved in various bacteria. *FEBS Lett* ***337***, 289-292.

Wächtershäuser, G. (1988). Before enzymes and templates: theory of surface metabolism. *Microbiol Rev* ***52***, 452-484.

Wächtershäuser, G. (1992). Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog Biophys Mol Biol* ***58***, 85-201.

Wächtershäuser, G. (2003). From pre-cells to Eukarya--a tale of two lipids. *Mol. Microbiol.* ***47***, 13-22.

Walden, H., Bell, G.S., Russell, R.J., Siebers, B., Hensel, R., and Taylor, G.L. (2001). Tiny TIM: a small, tetrameric, hyperthermostable triosephosphate isomerase. *J Mol Biol* ***306***, 745-757.

Wallace, B.A., Lees, J.G., Orry, A.J., Lobley, A., and Janes, R.W. (2003). Analyses of circular dichroism spectra of membrane proteins. *Protein Sci* ***12***, 875-884.

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

Waugh, D.S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr Purif* ***80***, 283-293.

Wilmanns, M., Priestle, J.P., Niermann, T., and Jansonius, J.N. (1992). Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from *Escherichia coli* refined at 2.0 A resolution. *J Mol Biol* ***223***, 477-507.

Wilson, G.G., and Murray, N.E. (1991). Restriction and modification systems. *Annu Rev Genet* ***25***, 585-627.

Winkelblech, J., Fan, A., and Li, S.M. (2015). Prenyltransferases as key enzymes in primary and secondary metabolism. *Appl Microbiol Biotechnol* ***99***, 7379-7397.

Woese, C.R., and Gupta, R. (1981). Are archaebacteria merely derived 'prokaryotes'? *Nature* ***289***, 95-96.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U. S. A.* ***87***, 4576-4579.

Woese, C.R., Magrum, L.J., and Fox, G.E. (1978). Archaebacteria. *J Mol Evol* ***11***, 245-251.

Wyckoff, T.J., Lin, S., Cotter, R.J., Dotson, G.D., and Raetz, C.R. (1998). Hydrocarbon rulers in UDP-N-acetylglucosamine acyltransferases. *J Biol Chem* ***273***, 32369-32372.

Xie, W., Zhou, C., and Huang, R.H. (2007). Structure of tRNA dimethylallyltransferase: RNA modification through a channel. *J Mol Biol* ***367***, 872-881.

Yang, X., Qin, W., Lehotay, M., Toki, D., Dennis, P., Schutzbach, J.S., and Brockhausen, I. (2003). Soluble human core 2 beta6-N-acetylglucosaminyltransferase C2GnT1 requires its conserved cysteine residues for full activity. *Biochim Biophys Acta* ***1648***, 62-74.

Yokobori, S.I., Nakajima, Y., Akanuma, S., and Yamagishi, A. (2016). Birth of Archaeal Cells: Molecular Phylogenetic Analyses of G1P Dehydrogenase, G3P Dehydrogenases, and Glycerol Kinase Suggest Derived Features of Archaeal Membranes Having G1P Polar Lipids. *Archaea **2016***, 1802675.

Zabin, I., Kepes, A., and Monod, J. (1962). Thiogalactoside transacetylase. *The Journal of biological chemistry **237***, 253-257.

Zhang, D., and Poulter, C.D. (1993a). Analysis and purification of phosphorylated isoprenoids by reversed-phase HPLC. *Anal Biochem **213***, 356-361.

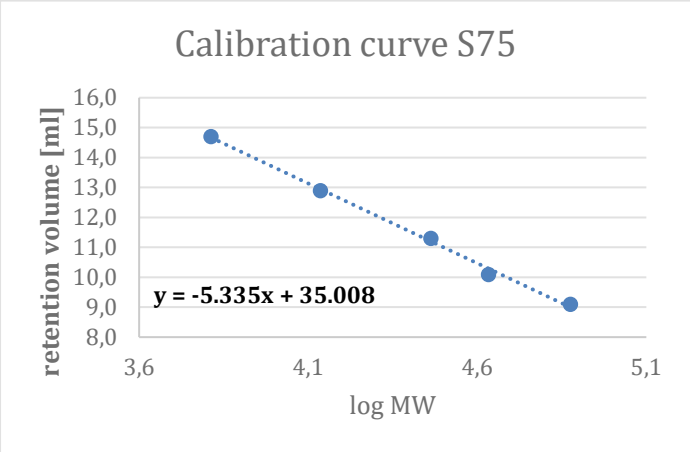
Zhang, D., and Poulter, C.D. (1993b). Biosynthesis of Archaebacterial Ether Lipids. Formation of Ether Linkages by Prenyltransferases. *J. Am. Chem. Soc. **115***, 1270 - 1277.

Zhang, Y.W., Koyama, T., Marecak, D.M., Prestwich, G.D., Maki, Y., and Ogura, K. (1998). Two subunits of heptaprenyl diphosphate synthase of *Bacillus subtilis* form a catalytically active complex. *Biochemistry **37***, 13411-13420.

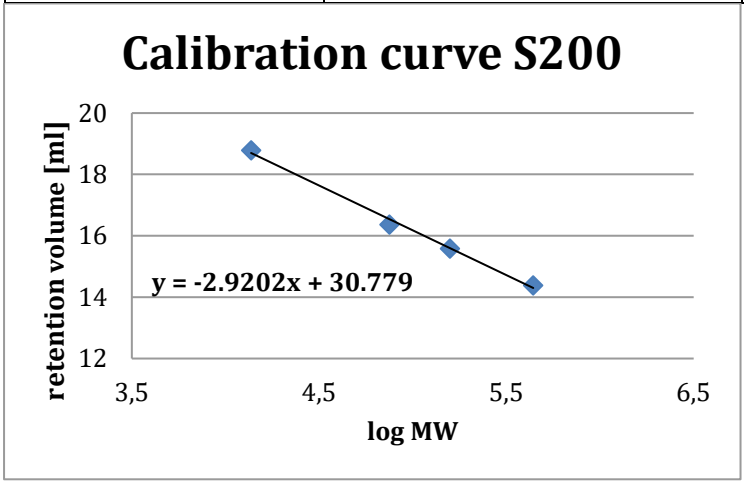
Zhu, H., Celinski, S.A., Scholtz, J.M., and Hu, J.C. (2001). An engineered leucine zipper a position mutant with an unusual three-state unfolding pathway. *Protein Sci **10***, 24-33.

9 SUPPLEMENT

9.1 Calibration curves



Protein	Retention volume [ml]	MW [Da]	log MW
Conalbumin	9.1	75000	4.8751
Ovalbumin	10.1	43000	4.6335
Carbonic anhydrase	11.3	29000	4.4624
Ribonuclease A	12.9	13700	4.1367
Aprotinin	14.7	6500	3.8129
Blue Dextran	7.43		



Protein	MW [Da]	log (MW)	Retention volume [ml]
Thyroglobulin	668000	5.82477646	12.62
Ferritin	440000	5.64345268	14.39
Aldolase	158000	5.19865709	15.59

Conalbumin	75000	4.87506126	16.37
Ribonuclease A	13700	4.13672057	18.79
Blue Dextran			8.16

9.2 Western Blots of baYvoF

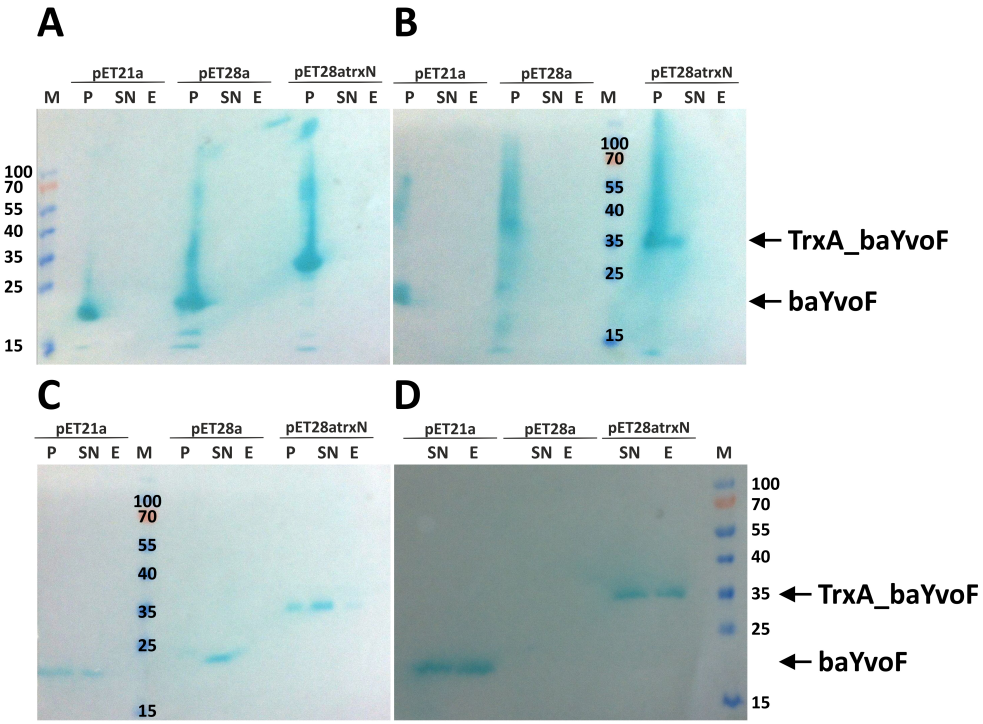


Figure S1 - Western blot for analysis of the best purification conditions for heterologously expressed baYvoF

E. coli BL21-CodonPlus (DE3) RIPL cells were transformed with pET21a_ba*yvoF*, pET28_ba*yvoF* or pET28atrxN_ba*yvoF,* and the gene was expressed for 6 h (A and C) or over night (B and D). The protein was either resuspended in 50 mM KP pH 7.5, 300 mM KCl, 10 mM imidazole and sonicated (A and B) or purified in the resuspension buffer containing 9 mM CHAPS without sonication (C and D). Purification was performed with His SpinTrap™ columns**.** Detection was done with an anti-$(His)_6$ tag antibody and visualized with 3,3´,5,5´-tetramethylbenzidine liquid substrate. The substrate system develops a blue reaction product after reaction with peroxidase. Size marker (M), molecular weights in kDa. P: pellet, SN: supernatant, E: eluate. TrxA_baYvoF and baYvoF are marked by arrows.

9.3 Overview about purification yields of biochemically characterized GGGPS enzymes

Table S1 – GGGPS enzymes and mutants characterized in this thesis with their purification yields following expression in *E. coli*

species	protein name	yield (mg/l culture volume)
Chitinophaga pinensis	cpGGGPS[1]	42.5[2] / 24.3[3]
Chitinophaga pinensis	cpGGGPS_Y143A[1]	31.7[2] / 21.2[3]
Spirosoma linguale	slGGGPS[1]	**3.7**[2] / 3.4[3]
Flavobacterium johnsoniae	fjGGGPS[1]	**19.2**[2] / 36.8[3]
Zunongwangia profunda	zpGGGPS[1]	**2.6**[2] / 7.7[3]
Thermococcus kodakaraensis	tkGGGPS[1]	**44**[2] / 20.6[3]
Thermococcus kodakaraensis	tkGGGPS_W143A[1]	**18.6**[2] / 16.7[3]
Thermoplasma acidophilum	taGGGPS[1]	**28.1**[2] / 18.5[3]
Methanothermobacter thermautotrophicus	mtGGGPS[1]	20.2[2] / 52.4[3]
Methanothermobacter thermautotrophicus	mtGGGPS_W141A[1]	**41**[2] / 42.5[3]
Methanothermobacter thermautotrophicus	mtGGGPS_K146A[4]	64.8[2]
Methanothermobacter thermautotrophicus	mtGGGPS_I155A[4]	45[2]
Methanothermobacter thermautotrophicus	mtGGGPS_D144A[4]	35[2]
Methanothermobacter thermautotrophicus	mtGGGPS_I107E[4]	78.4[2] / 92.3[3]
Methanothermobacter thermautotrophicus	mtGGGPS_Y105A	38.1[2]
Methanothermobacter thermautotrophicus	mtGGGPS_D57R	29.8[2]
Methanothermobacter thermautotrophicus	mtGGGPS_R88E	39.7[2]
Methanothermobacter thermautotrophicus	mtGGGPS_A162E[4]	31.5[2] / 44.7[3]
Methanothermobacter thermautotrophicus	mtGGGPS_A162E_W141A	16.6[2] / 34.2[3]
Methanothermobacter thermautotrophicus	mtGGGPS_A162E_I107E	13.8[2]
Methanothermobacter thermautotrophicus	mtGGGPS_A162E_Y105A	36.4[2]
Methanothermobacter thermautotrophicus	mtGGGPS_A162E_D57R	33.1[2]

Methanothermobacter thermautotrophicus	mtGGGPS_A162E_R88E	46.4[2]

[1] The genes of these enzymes have already been cloned previously (Peterhoff *et al.*, 2014) [2]Purification of enzyme in 50 mM potassium phosphate pH 7.5 [3]Purification of enzyme in 50 mM Tris/HCl pH 8.0 **Bold letters** → These enzymes have already been purified previously (Peterhoff *et al.*, 2014) [4]The genes of these mutants have already been cloned previously (Kilu, 2014)

9.4 Radiometric activity assays of mtGGGPS_wt and its mutants with GPP and GGPP

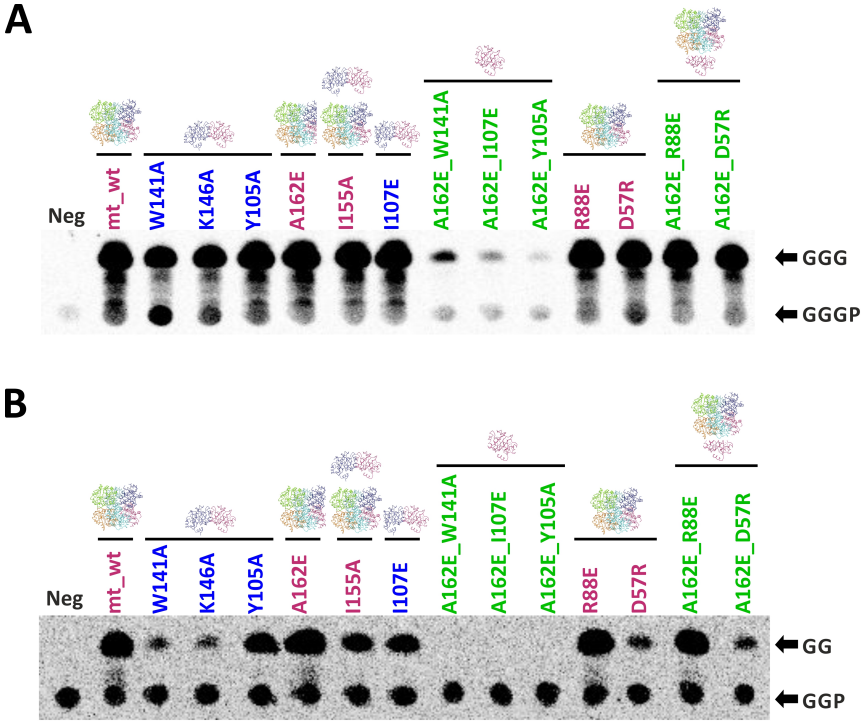
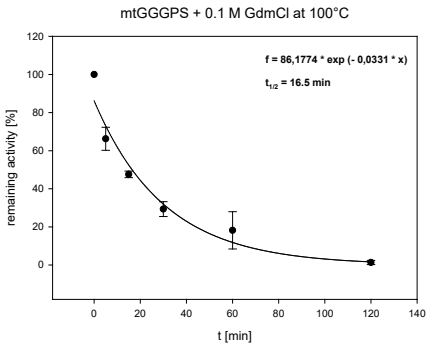
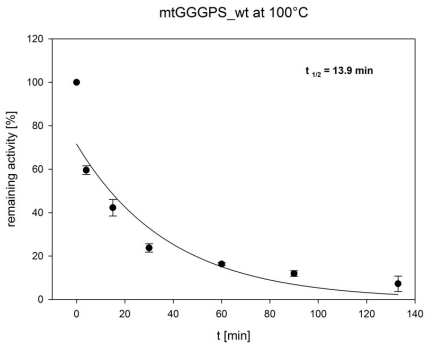
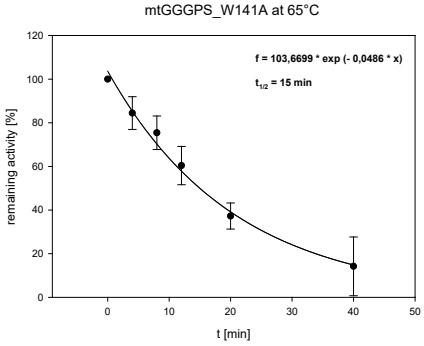


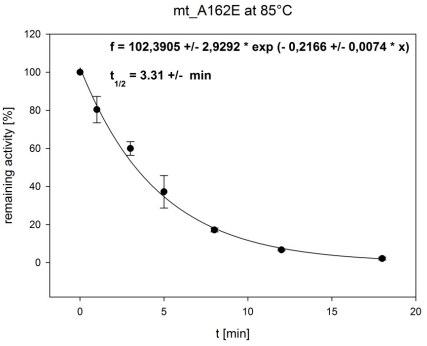
Figure S2 - Activity assay of mtGGGP_wt and its mutants with GGPP and GPP
(**A**) The substrate ^{14}C-G1P was incubated with GGPP and 1 µM of mtGGGPS_wt or mutant (subunit concentration) for 2 h at 40 °C. (**B**) The substrate ^{14}C-G1P was incubated with GPP and 1 µM of mtGGGPS_wt or mutant (subunit concentration) for 2 h at 40 °C. The generated products were extracted, separated by thin layer chromatography, and visualized by autoradiography. Hexameric mtGGGPS mutants are shown in magenta, dimeric mtGGGPS mutants in blue and monomeric mtGGGPS mutants in green. As negative control (Neg) no enzyme was added. The origin of the chromatography (marked by the product spot GGGP from GGGPS activity) and with CIP dephosphorylated GGG are marked by arrows. The oligomerization state is symbolized by ribbon structure, different color pointing out each subunit.

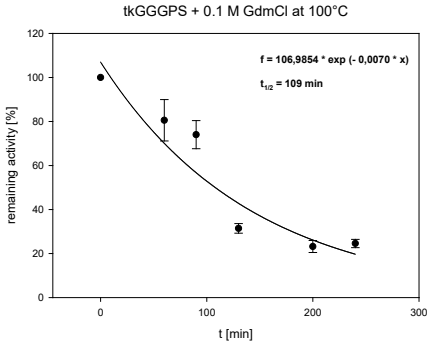
9.5 Irreversible heat inactivation of GGGPS enzymes

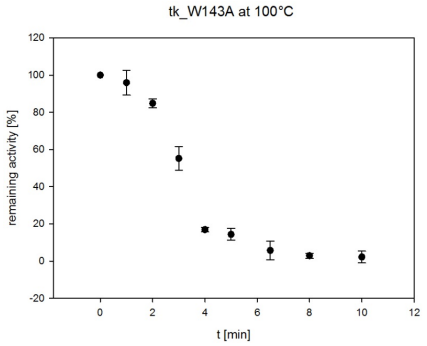


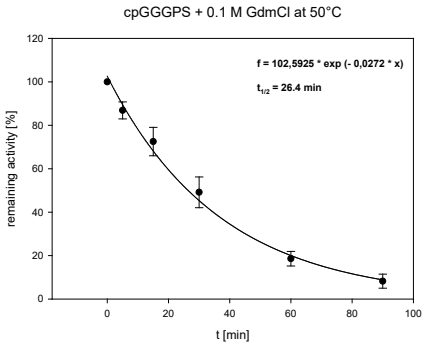


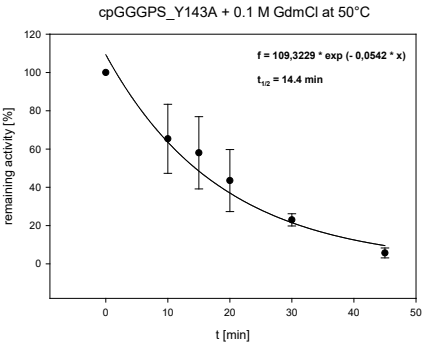












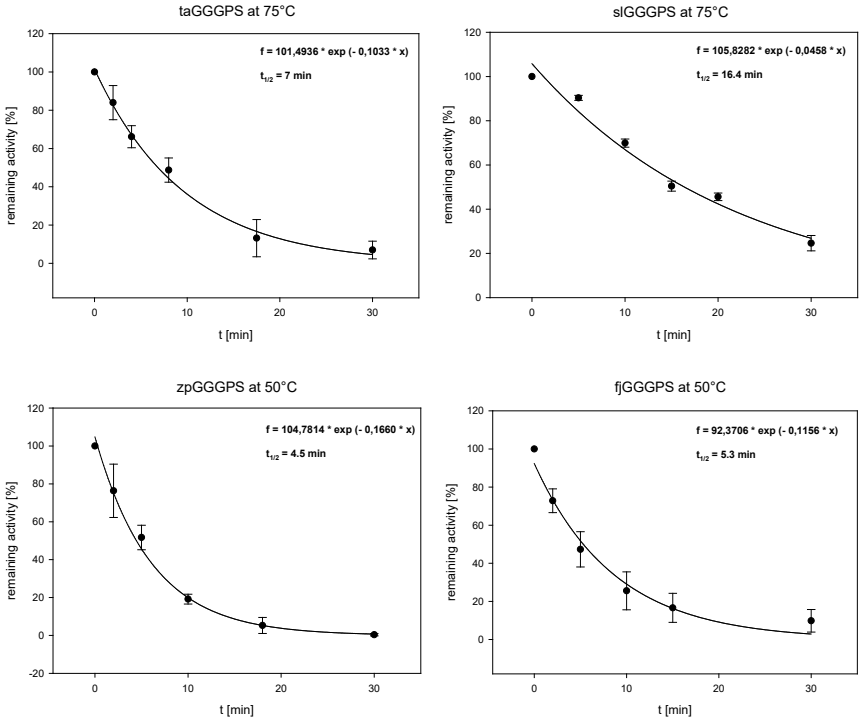


Figure S3- Irreversible heat inactivation of GGGPS enzymes

All GGGPS enzymes were irreversibly inactivated by heat and analyzed as described in chapter 5.6.16.3. The reaction was started by adding the enzyme, the increase of absorbance at 293 nm was followed and reaction velocities were calculated from the protein concentration and the initial slopes. The experiment was done in triplicates, the error bars show standard deviations. $t_{1/2}$ was obtained by fitting a single exponential decay equation to the data using SigmaPlot.

9.6 Data collection and refinement statistics of mtGGGPS_I107E, mtGGGPS_W141A and mtGGGPS_A162E

Table S2 - mtGGGPS_I107E: Data collection and refinement statistics[1]

Wavelength	1
Resolution range	47.63 - 2.394 (2.48 - 2.394)
Space group	P 21 21 21
Unit cell	76.978 136.409 152.503 90 90 90
Total reflections	863695 (75825)
Unique reflections	63701 (6038)
Multiplicity	13.6 (12.6)
Completeness (%)	0.99 (0.95)
Mean I/sigma(I)	26.06 (4.79)
Wilson B-factor	39.12
R-merge	0.07903 (0.5179)
R-meas	0.08213 (0.5393)
CC1/2	0.999 (0.915)
CC*	1 (0.977)
Reflections used in refinement	63700 (6039)
Reflections used for R-free	3187 (302)
R-work	0.1756 (0.2002)
R-free	0.2298 (0.2621)
CC(work)	0.950 (0.900)
CC(free)	0.902 (0.811)
Number of non-hydrogen atoms	11046
macromolecules	10678
Protein residues	1445
RMS(bonds)	0.008
RMS(angles)	0.97
Ramachandran favored (%)	96
Ramachandran allowed (%)	2.7
Ramachandran outliers (%)	1.4

[1] Statistics for the highest-resolution shell are shown in parentheses.

Table S3 - mtGGGPS_W141A: Data collection and refinement statistics[1]

Wavelength	1
Resolution range	47.96 - 2.697 (2.794 - 2.697)
Space group	P 1 21 1

Unit cell	83.198 62.783 136.259 90 96.352 90
Total reflections	113777 (1000)
Unique reflections	35033 (800)
Multiplicity	3.2 (1.2)
Completeness (%)	0.90 (0.21)
Mean I/sigma(I)	9.92 (0.95)
Wilson B-factor	58.53
R-merge	0.08797 (0.5434)
R-meas	0.105 (0.7473)
CC1/2	0.996 (0.513)
CC*	0.999 (0.824)
Reflections used in refinement	35020 (800)
Reflections used for R-free	1751 (40)
R-work	0.2344 (0.3516)
R-free	0.3197 (0.4467)
CC(work)	0.932 (0.646)
CC(free)	0.873 (0.492)
Number of non-hydrogen atoms	9955
macromolecules	9911
Protein residues	1369
RMS(bonds)	0.011
RMS(angles)	1.29
Ramachandran favored (%)	91
Ramachandran allowed (%)	7.1
Ramachandran outliers (%)	2.1

[1] Statistics for the highest-resolution shell are shown in parentheses.

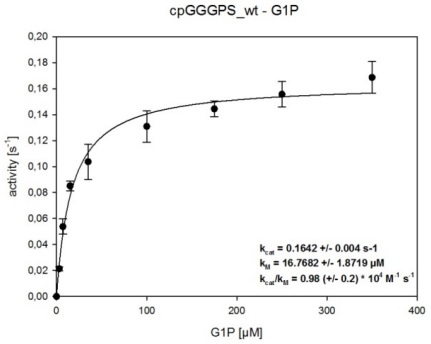
Table S4 - mtGGGPS_A162E: Data collection and refinement statistics[1]

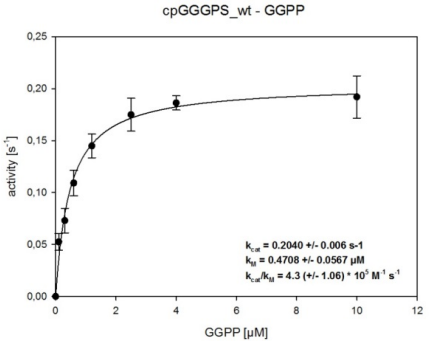
Wavelength	1
Resolution range	48.36 - 1.949 (2.019 - 1.949)
Space group	C 1 2 1
Unit cell	142.539 147.662 90.347 90 108.993 90
Total reflections	856285 (61391)
Unique reflections	125165 (9604)
Multiplicity	6.8 (6.4)
Completeness (%)	0.97 (0.75)
Mean I/sigma(I)	17.16 (1.79)
Wilson B-factor	37.34

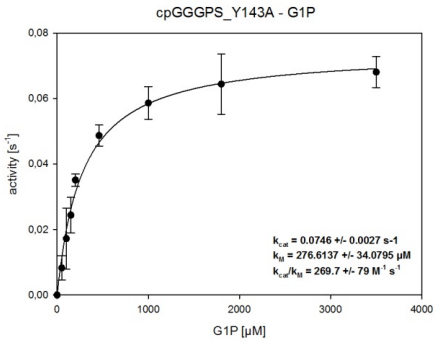
R-merge	0.06882 (1.068)
R-meas	0.07452 (1.16)
CC1/2	0.999 (0.746)
CC*	1 (0.925)
Reflections used in refinement	125165 (9566)
Reflections used for R-free	6257 (479)
R-work	0.2129 (0.3442)
R-free	0.2556 (0.3623)
CC(work)	0.958 (0.759)
CC(free)	0.935 (0.780)
Number of non-hydrogen atoms	11353
macromolecules	10776
Protein residues	1446
RMS(bonds)	0.008
RMS(angles)	0.98
Ramachandran favored (%)	97
Ramachandran allowed (%)	2.5
Ramachandran outliers (%)	0.56

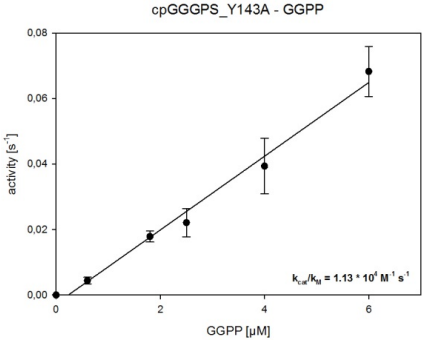
[1] Statistics for the highest-resolution shell are shown in parentheses.

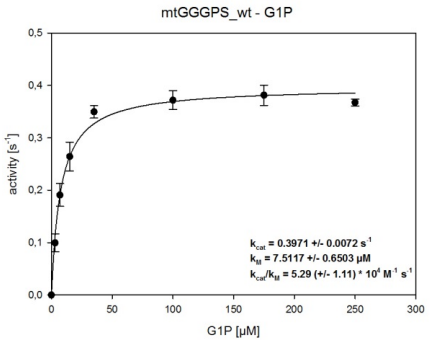
9.7 Steady-state kinetic measurements of GGGPS enzymes

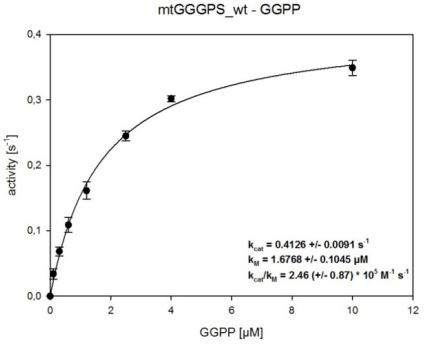


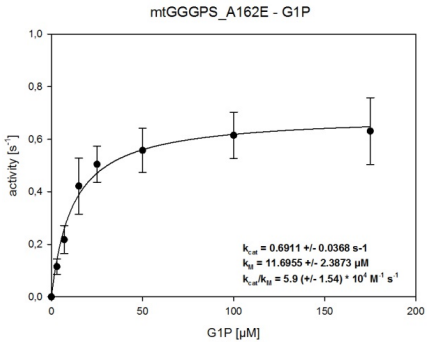


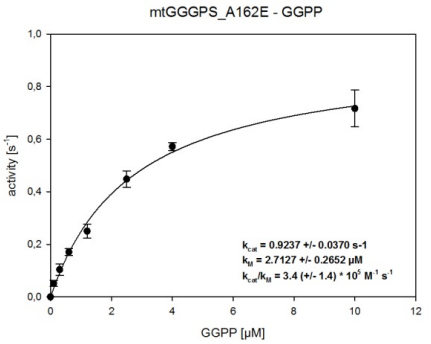












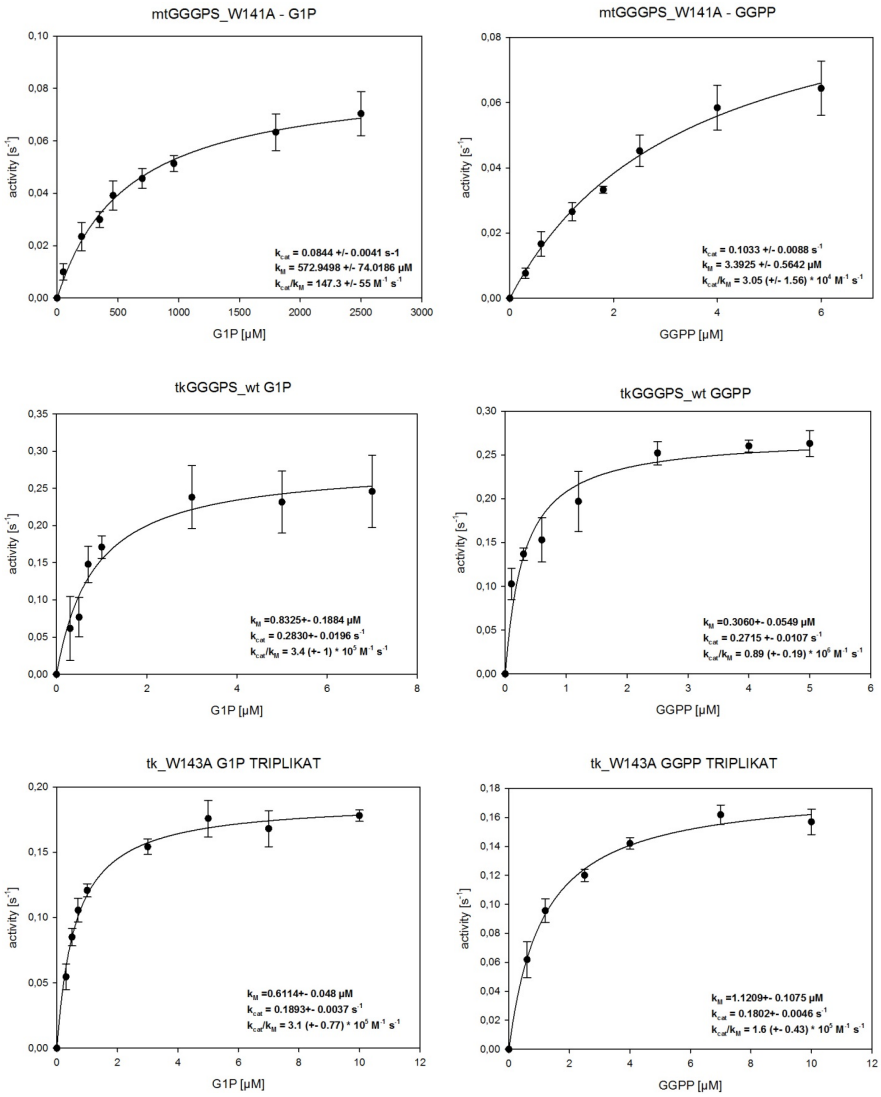


Figure S4 - Steady-state kinetic meausurements of GGGPS enzymes

All GGGPS enzymes were kinetically analyzed as described in chapter 5.6.16.2. The reaction was started by adding the enzyme, the increase of absorbance at 293 nm was followed and reaction velocities were calculated from the protein concentration and the initial slopes. The experiment was done in triplicates, the error bars show standard deviations. Kinetic constants were obtained by fitting the Michaelis-Menten equation to the data using SigmaPlot.

9.8 Protein sequences

The amino acids LE from restriction site *Xho*I and $(His)_6$ tag included depicted in **bold letters**, the thioredoxin tag and the amino acid H from restriction site *Nde*I is depicted in **blue bold letters**.

bsPcrB

MKKFPKKLLPIAVLSSIAFSSLASGSVPEASAQEKKKGNQDEIKNVIVLIGDGMGVSYTSAYRYLKDNKKTKVVE
PTAFDQYLVGQQTTYPDDPEQNVTDSAAAATAMSAGIKTYNNAIAVDNDGSEAKTVLEAAKEKGKATGLVA
TSEITHATPASFGSHDHSRKNMNSIADDYFDEMVNGKHKIDVLLGGGKSNFDRKDRNLIKEFKKAGYSYVDDR
KDMLKNKDSQVLGLFADGGLPKKIDRTKDIPSLKDMTNTAIKKLNKDKDGFFLMVEGSQIDWAGHDNDIVG
AMSEMEDFEQAYKAAIDFAKKDKHTLVVATADHSTGGYSIGADGIYNWFSEPIKAAKRTPDFMAEKIADGAD
VEKTLKTYIDQKKLALTKAEIQSVEEAAKSKEVLDIDNAIENIFNKRSHTGWTTGGHTGEDVPVYAYGPSSETFA
GQIDNTEIAKNVFKALQYNIKINDK**LEHHHHHH**

bsYvoF

**MGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPK
YGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLALVPRGSH**MRKTDRHPVSGANSLWHVYQTVPFLK
VVKNFIVIQIARYTPFIGMKNWLYRTFLRMKVGKQTSFALMVMPDIMFPEKISVGTNTIIGYNTTILAHEYLIHE
YRIGKVLIGDEVMIGANTTILPGVKIGDGAVVSAGTLVHKDVPDGAFVGGNPMRIIYTKEEMQERLKKSAE**LE
HHHHHH**

baYvoF

**MGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPK
YGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLALVPRGSH**MRRTTRYPVSGESSLWNVYKTVSFWK
VMKNFIIIQIARYTPFLSVKNWLYRTFLRMKVGKKTSFALMVMPDIMFPEKITVGDNSIIGYNTTLLAHEYLIREY
RLGEIVIGNEVMIGANTTILPGVRIGDGAIVSAGTLVHKDVPSGAFVGGNPMRVIYTKEQMEAREGSY**LEHHH
HHH**

bsMAT

**MGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPK
YGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLALVPRGSH**MMTVLEDMLRKTRNGKVHMTLIDPG
AKPPQECARIAEEAEMAGTDFIMVGGSTDIDSRAMDEAISAIKAKTDLKVIIFPGSSLMISPKADAIFFMSLLNS
GSLEYVVGHQVKAAIPLSAMKIEKIPMAYLVFDPGMTVGRVGKAHLIPRDDEKTALSYALAAQYFGFRLVYFE
AGSGSPYHVGENVVRRVKQELDIPVIVGGGIRTPEAAKALAQAGADMIVTGTIAERSVNVYEALHPIVESIKEV
GISKIQ**LEHHHHHH**

taGGGPS

MMTVLEDMLRKTRNGKVHMTLIDPGAKPPQECARIAEEAEMAGTDFIMVGGSTDIDSRAMDEAISAIKAKT
DLKVIIFPGSSLMISPKADAIFFMSLLNSGSLEYVVGHQVKAAIPLSAMKIEKIPMAYLVFDPGMTVGRVGKAHL
IPRDDEKTALSYALAAQYFGFRLVYFEAGSGSPYHVGENVVRRVKQELDIPVIVGGGIRTPEAAKALAQAGAD
MIVTGTIAERSVNVYEALHPIVESIKEVGISKIQ**LEHHHHHH**

cpGGGPS

MHNKIYNSFIDRKAKGIKSFAVLIDPDKVNPADIADLAAKCTAAKVDYIFLGGSLVITNHLDECVQQFKTLCDIP
VVLFPGSPSQVSRYADALLYLSVISGRNPELLIGQHVLSAPAVKKSGLEVISTGYVLIDGGAPTTVSYISNTTPIPS
DKDDIAMCTAMAGEMLGMKVVFMDAGSGARKPITESMISRVASQVSAPIIVGGGIRDAEKAYLNCKAGADII
VVGNAIEKETSLIKEMADAVHAAAPVLK**LEHHHHHH**

mtGGGPS

MFKMKVEDYFHDILRERKIHLTLIDPEEQTPEEAVEIARAAIRGGTDGIMLGGSTTDSSELDNTARALRENIDVP
IILFPGNTTGVSRYADAIFFMSLLNSTNPYWIIGAQALGAATVKKMGIEALPMGYLVVEPGGTVGWVGDTKPV
PRNKPDIAAAYAMAAEFLGMRLFYLEAGSGAPEHVPEEMIALVKRCTDQILIVGGGIRSGEDAARVAGAGAD
VVVTGTVVENSDNVEDKIREIVEGMGSV**LEHHHHHH**

fjGGGPS

MEQKILTTIHQQILEAKKNGQKLLAILLDPDKIVWENLDHLLLKINQSPATHIFVGGSIVESTIIEDLIAQLKQKTRL
PVVIFPGDPSQISPKADAILFLSLLSGRNPDYLIEYQVQAAPILKKTNLEVISTGYILIESGNETAVARVSKTEPLNR
ENFDLALATAQAGEMLGSKLIYLEAGSGAKKPVPLEMISVISQNVEIPIIVGGGIVDLHGIKKAYNAGADLVVIG
TAFENDSHFFDS**LEHHHHHH**

slGGGPS

MTILRDYKLSGRKAFAVLLDPDKVEQDAFSTLLQRTADYPVDFFLVGGSLVTDYAHKEVIATIRRYSSTPVILFPG
NPLHIESSADAILLLSLISGRNADFLIGQHVIAAPLLKKSGLEILPTGYMVVDSGTQTTVSYISGTMPLPHDKPDV
AACTALAGEMLGLQLMYLDAGSGARRPVSAAMIAAVRKAVNVPIIVGGGITSGEKAYEALKAGADMIVVGN
GVEQDPDLLPQLATVVREFNQSVVQA**LEHHHHHH**

tkGGGPS

MLKLGKVETYIHEKLEREKLHFVLLDPDDVSPELAGELASMSEEVGVDAIMVGGSTGAEGEVLDSVVRAIKESS
NLPVILFPGSHGGISKYADAIFFMSLLNSRNPFFITGAQALGAFQVKRYGIEPIPMAYLIIEPGETVGWVGDAKPI
PRHKPKIAAAYALAGQYLGMRLVYLEAGSGAPQPVPPEMIGLVKRVIDVPLIVGGGIRTEEQARAAVKAGADII
VTGTAIEKAGSVEKAREKLEELNRGVKG**LEHHHHHH**

zpGGGPS

MPKILDAIIKASKINKKLLAVLIDPEKFATENYSYFIEKLPEAVTHIFVGGSTATTAQSEVCVDFIKTKTNLPVILFPG
DKEQITEKADGILLLSLISGRNPEYLIEQHIKAVPKLLNAGLEIIPTGYLLLDGGNQSAVARVSKTKPIQQDEIELIR
NTALAGAMLGKQLVYLEAGSGALIPVSEKVIAEVKRDLNIPLIVGGGIRNATQLKKAYKAGADLVVIGTAFENG
EFK**LEHHHHHH**

10 ACKNOWLEDGEMENTS

Foremost, I would like to thank my PhD supervisor Prof. Dr. Reinhard Sterner for his support and constant guidance throughout this thesis. I´m very grateful for your generous and kind support in all scientific projects.

Further on, I would like to thank Dr. Patrick Babinger most sincerely for his permanent support, encouragement, patience, practical advice and expertise. I deeply appreciated your mentorship on a scientific as well as personal level.

I would like to thank Prof. Dr. Christine Ziegler for enabling protein crystallization in Regensburg, and Dr. Chitra Rajendran for solving the structures of some of my proteins. In accordance, I´m very grateful to Prof. Dr. Thomas Brück for mentoring my thesis in the context of the Internationale Graduate School of Life Science.

Many thanks to Prof. Dr. Rainer Merkl and Kristina Heyn for their ideas and valuable contributions to this thesis.

I express my deepest gratitude to Hermine Reisner, Sonja Fuchs, Christiane Endres, Sabine Laberer and Klaus-Jürgen Tiefenbach for their reliable technical assistance and pleasant teamwork.

Also, I am very grateful to Dr. Harald Guldan and especially Dr. David Peterhoff for introducing me to the PcrB-project, answering all my questions and for their constant support.

Special thanks to my colleagues Alexandra Holinski, Wolfgang Kaiser and Michael Schupfner for your awesome personal support and many lovely hours at work as well as during leisure time. I am more than grateful for enjoyable “Opera” breakfast sessions with you two (Alexandra and Hermine).

I would also like to thank my interns Whitney Kilu and Christina Seißenberger for their valuable contribution to this work. Thanks to Claudia Pauer for administrative support.

Cordial thanks to all current and former members of the Sterner lab for an excellent atmosphere and lots of valuable discussions as well as outstanding excursions and christmas parties.

Finally, I am deeply grateful to Markus and my family, my friends and Leonardo for their unlimited encouragement and support.