

Technische Universität München

Lehrstuhl für Mikrobiologie

Entwicklung eines Systems zur
komplementationsbasierten Selektion
thermostabiler Glycosidhydrolasen
in *Thermus thermophilus*

Matthias Mörch, M.Sc. (TUM)

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Aphrodite Kapurniotu

Prüfer der Dissertation: 1. Prof. Dr. Wolfgang Liebl

2. Prof. Dr. Rudi F. Vogel

Die Dissertation wurde am 15.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 24.09.2019 angenommen.

INHALTSVERZEICHNIS

Abkürzungen & Akronyme	V
1.1 Allgemeine Abkürzungen	V
1.2 Chemische Verbindungen	VI
Zusammenfassung	VIII
Abstract	X
1 Einleitung	1
1.1 Glycosidhydrolasen	1
1.2 Mikroorganismen und ihr Temperaturspektrum	2
1.2.1 <i>Oleispira antarctica</i>	3
1.2.2 <i>Bacillus subtilis</i>	3
1.2.3 <i>Escherichia coli</i>	4
1.2.4 <i>Geobacillus stearothermophilus</i>	4
1.2.5 <i>Picrophilus torridus</i>	5
1.2.6 <i>Thermus thermophilus</i>	6
1.3 Auf Massenspektrometrie basierende Proteomik	8
1.4 Thermal proteome profiling (TPP)	10
1.5 Cellular Thermal Shift Assay (CETSA)	12
1.6 Ziel der Arbeit	13
2 Material & Methoden	15
2.1 Chemikalien	15
2.2 Nährmedien	17
2.3 Puffer & Lösungen	19
2.4 Kits, kommerzielle Lösungen & Reagenzien	20
2.5 Technische Gerätschaften	21
2.6 Verbrauchsmaterialien	23
2.7 Computersoftware	24
2.8 Mikroorganismen	25
2.8.1 Kultivierung von Mikroorganismen	25
2.8.2 Antibiotika	27
2.8.3 Kultivierung von <i>Picrophilus torridus</i>	28
2.9 Bestimmung der Zelldichte	29
2.10 Plasmide	29

Inhaltsverzeichnis

2.11	Isolierung genetischen Materials aus Mikroorganismen	30
2.11.1	Genomische DNS.....	30
2.11.2	Plasmide	30
2.12	Polymerase-Kettenreaktion	30
2.13	Reinigung von Plasmiden und PCR-Produkten	32
2.14	DNS-Quantifizierung	32
2.15	Restriktionsverdau	32
2.16	DNS-Fragmentassemblierung nach GIBSON	33
2.17	Agarose-Gelelektrophorese	33
2.18	DNS-Sequenzierung	34
2.19	Transformation von Genmaterial	34
2.19.1	Chemische Transformation von <i>E. coli</i>	34
2.19.2	Transformation von <i>E. coli</i> mittels Elektroporation	35
2.19.3	Transformation von <i>T. thermophilus</i>	35
2.20	Kolonie-Cracking	36
2.21	Herstellung eines GH-defizienten <i>T. thermophilus</i> Stammes	37
2.22	Herstellung rekombinanter Stämme mit neu konstruierten Plasmiden.....	39
2.23	GH-Aktivitätsbestimmung mit <i>pNP</i> -Substraten	39
2.24	Herstellung von Proteinextrakten für TPP Experimente	40
2.25	Proteinkonzentrationsbestimmung nach Bradford.....	41
2.26	Dynamische Differenzfluorimetrie	42
2.27	Thermische Inaktivierung von Bakterien anhand der Überlebensrate	43
2.28	Logistische Regression.....	44
3	Ergebnisse	45
3.1	Komplementierung eines GH-defizienten <i>Thermus thermophilus</i> -Stammes.....	45
3.1.1	Aminosäuresupplementierung von basalem Minimalmedium.....	45
3.1.2	Wachstum von <i>T. thermophilus</i> und GH-defizienten Stämmen in TMB	47
3.1.3	Aktivitäten der Enzyme Bgl, CglT und CglT KIR	49
3.2	Proteomweite Untersuchung der Proteininstabilität bei <i>T. thermophilus</i> und anderen Mikroorganismen	51
3.2.1	Übersicht aller gemessenen Proteome	52
3.2.2	Verteilung der proteomweiten Thermostabilität.....	54
3.2.3	Proteomweite Verteilung von <i>pI</i> -Werten.....	55
3.2.4	Vergleich thermostabiler und instabiler Proteine	57
3.2.5	Analyse von Paralogen und Orthologen.....	64
3.2.6	Gegenüberstellung von intakten Zellen und Zellextrakten	70
3.2.7	<i>T. thermophilus</i> : Vergleich von Chromosom und Megaplasmid	75
3.2.8	Thermische Inaktivierung von <i>E. coli</i> und <i>T. thermophilus</i>	77
3.2.9	DSF mit intakten Zellen und Zellextrakten.....	80

Inhaltsverzeichnis

4	Diskussion	83
4.1	Komplementierung eines GH-defizienten <i>Thermus thermophilus</i> -Stammes.....	83
4.1.1	Das Minimalmedium (TMB)	84
4.1.2	Herstellung einer für die komplementationsbasierte Selektion geeigneten GH- Deletionsmutante.....	86
4.1.3	Validierung des Multideletionsstamms Δ GH für die komplementationsbasierte Selektion.....	89
4.1.4	Über die Anwendung komplementationsbasierter Selektion	92
4.2	Untersuchung der Proteinstabilität unterschiedlicher Mikroorganismen	93
4.2.1	pI-Werte.....	94
4.2.2	Proteinlänge.....	95
4.2.3	Hydrophobizität.....	96
4.2.4	Aminosäurefrequenz	98
4.2.5	Betrachtung von Paralogen und Orthologen	99
4.2.6	Gegenüberstellung von intakten Zellen und Zellextrakten	102
4.2.7	Thermische Zellinaktivierungskinetiken und DSF.....	104
5	Ausblick	108
	Literaturverzeichnis	XII
	Anhang	XX
	Abbildungsverzeichnis	XXVIII
	Tabellenverzeichnis	XXX
	Eigenständigkeitserklärung	XXXII
	Danksagung	XXXIII

ABKÜRZUNGEN & AKRONYME

1.1 Allgemeine Abkürzungen

Tabelle 1: Abkürzungen und deren ausgeschriebene Form.

Abkürzung	Definition
×	ein Vielfaches von
× g	Vielfaches der Erdbeschleunigung
°C	Grad Celsius
μ-	mikro (10 ⁻⁶)
A	Ampere (elektrische Stromstärke)
AS	Aminosäure
<i>B. sub.</i>	<i>Bacillus subtilis</i>
<i>blm</i>	Bleomycinresistenzkassette (A2)
Bp	Basenpaar
d	Tag
Da	Dalton
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytidintriphosphat
dGTP	Desoxyguanosintriphosphat
DNS	Desoxyribonukleinsäure
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
dTTP	Desoxythymidintriphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
<i>engl.</i>	Englisch
EtOH	Ethanol
F	Farad (elektrische Kapazität)
g	Gramm
<i>G. stea.</i>	<i>Geobacillus stearothermophilus</i>
GH	Glycosidhydrolase
h	Stunde
k-	kilo (10 ³)
KBE	koloniebildende Einheit
L	Liter
Lsg.	Lösung

Abkürzungen & Akronyme

Abkürzung	Definition
m	Meter
M	Molar (mol L ⁻¹)
m-	milli (10 ⁻³)
M-	Mega (10 ⁶)
m/z	Masse-zu-Ladungsverhältnis
MW	Molekulargewicht
n	Anzahl
n-	nano (10 ⁻⁹)
<i>O. ant.</i>	<i>Oleispira antarctica</i>
<i>P. tor.</i>	<i>Picrophilus torridus</i>
Pa	Pascal
PFGE	Pulsed-Field-Gelelektrophorese
pH	negativ dekadischer Logarithmus der Hydroniumionenkonzentration
rpm	Umdrehungen pro Minute (<i>engl.</i> : revolutions per minute)
s	Sekunde
s.o.	siehe oben
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese
<i>slp</i>	S-Layer-Protein
sp.	Spezies
ssp.	Subspezies
t	Zeit
<i>T. th.</i>	<i>Thermus thermophilus</i>
V	Volt (elektrische Spannung)
VE (-Wasser)	vollentsalzt (-es Wasser)
Ω	Ohm (elektrischer Widerstand)

1.2 Chemische Verbindungen

Tabelle 2: Summenformeln chemischer Verbindungen und deren ausgeschriebene Form.

Summenformel	Name
(NH ₄)NO ₃	Ammoniumnitrat
CaCl ₂	Kalziumchlorid
CaSO ₄	Kalziumsulfat
CoCl ₂	Cobalt(II)-chlorid
CTAB	Cetyltrimethylammoniumbromid
CuSO ₄	Kupfersulfat
DMF	Dimethylformamid

Abkürzungen & Akronyme

Summenformel	Name
DMSO	Dimethylsulfoxid
EDTA	Ethylendiamintetraessigsäure
FeCl ₃	Eisen(III)-chlorid
H ₃ BO ₃	Borsäure
KBr	Kaliumbromid
KCl	Kaliumchlorid
KH ₂ HPO ₄	Kaliumdihydrogenphosphat
KNO ₃	Kaliumnitrat
MgCl ₂	Magnesiumchlorid
MgSO ₄	Magnesiumsulfat
MOPS	3-(N-Morpholino)propansulfonsäure
Na ₂ HPO ₄	Dinatriumhydrogenphosphat
Na ₂ MoO ₄	Natriummolybdat
Na ₂ SO ₄	Natriumsulfat
NaCl	Natriumchlorid
NaCO ₃	Natriumcarbonat
NaF	Natriumfluorid
NaH ₂ PO ₄	Natriumdihydrogenphosphat
NaHCO ₃	Natriumhydrogencarbonat
NaNO ₃	Natriumnitrat
NTA	Nitrilotriessigsäure
PIPES	Piperazin-1,4-bis-(2-ethansulfonsäure)
PVP	Polyvinylpyrrolidon
SrCl ₂	Strontiumchlorid
Tris-HCl	Tris-(hydroxymethyl)-aminomethanhydrochlorid
X-Gal	5-Brom-4-chlor-3-indoxyl-β-D-galactopyranosid
ZnSO ₄	Zinksulfat

ZUSAMMENFASSUNG

Eine erhöhte Thermostabilität von Enzymen wird mit längeren Halbwertszeiten ihrer Aktivität und oft unter Erhalt dieser Aktivität bei höheren Temperaturen in Verbindung gebracht. Aus diesem Grund besteht industrielles Interesse an neuartigen, thermostabilen Enzymen bzw. an Verfahren zur Thermostabilisierung bereits eingesetzter Enzyme. Mithilfe solcher Enzyme können neue und robuste biotechnologische Verfahren für die Umwandlung bisher noch nicht erschlossener Rohstoffquellen pflanzlicher Biomasse in Biotreibstoffe der nächsten Generation oder für die nachhaltige Herstellung von Ausgangsverbindungen für die pharmazeutische und chemische Industrie entwickelt werden. Für die Identifizierung von thermostabilen Glycosidhydrolasen (GHs) wurde im Rahmen dieser Arbeit ein System zur komplementationsbasierten Selektion im extrem thermophilen Bakterium *Thermus thermophilus* entwickelt. Mit Hilfe rekombinanter Expression von GHs in diesem Wirtsorganismus konnte der prinzipielle Funktionsnachweis dieses Systems erbracht werden. Modellhaft wurde hierfür die Wildtyp-Form von CglT (eine GH der Enzymfamilie GH1) von *Thermoanaerobacter brockii* und die thermisch stabilere Mutante CglT KIR verwendet. Neben der wirtseigenen β -Glucosidase konnten beide CglT-Varianten die Fähigkeit zu Wachstum einer GH-Deletionsmutante in Minimalmedium mit D-Cellobiose wiederherstellen.

Im zweiten Teil dieser Arbeit wurde proteomweit die Protein-Thermostabilität – das sogenannte Meltom – von *T. thermophilus* und fünf weiteren Mikroorganismen (*Oleispira antarctica*, *Bacillus subtilis*, *Escherichia coli*, *Geobacillus stearothermophilus* und *Picrophilus torridus*) analysiert. Diese Organismen weisen ein breites Spektrum unterschiedlicher Wachstumstemperatur-Präferenzen auf; von psychrophil bis extremthermophil. Es konnte gezeigt werden, dass allgemein die Länge eines Proteins negativ mit dessen Thermostabilität korreliert, jedoch Eigenschaften wie Hydrophobizität und Aminosäurefrequenz keine Indikatoren für Thermostabilität in aufgeschlossenen Zellen darstellen. Die Gegenüberstellung von Proteinen aus Zellextrakten und intakten Zellen von

Zusammenfassung

E. coli und *T. thermophilus* zeigte, dass überwiegend Membranproteine unter *in vivo*-Bedingungen Stabilisierung erfahren. Der Vergleich des Denaturierungsverhaltens von Paralogen und Orthologen ergab, dass sich die Schmelztemperaturen nachweislich mit steigender Sequenzidentität annähern. Dieser Zusammenhang weist darauf hin, dass Thermostabilität eine evolutionär konservierte Eigenschaft darstellt. Des Weiteren ließ sich hierbei ableiten, dass die thermische (De-)Stabilisierung von einigen wenigen Aminosäuren an gewissen Schlüsselstellen ausgeht. Zudem konnte anhand der vorliegenden Meltdaten, in Kombination mit Zellinaktivierungsexperimenten und dynamischer Differenzfluorimetrie für *E. coli* und *T. thermophilus*, die Frage um die Gründe thermischer Zellinaktivierung erhellt werden. Mit Hilfe einer Anreicherung von *Gene Ontologie*-Begriffen jener Proteine, welche im Bereich thermischer Zellinaktivierung denaturierten, ergaben sich Hinweise darauf, dass der Verlust DNS-/RNS-bindender Proteine und ribosomaler Struktur hierbei eine Rolle spielen und somit diese Proteingruppen für den thermischen Zelltod von wichtiger Bedeutung sein könnten.

ABSTRACT

Increased thermostability of enzymes is associated with longer half-lives of their activity and maintaining this activity at higher temperatures. For this reason, there is industrial interest in novel, thermostable enzymes or in processes for the thermal stabilisation of enzymes already used. Such enzymes can be used to develop new and robust biotechnological processes for transforming so far untapped raw material sources of plant biomass into next-generation biofuels or for the sustainable production of precursor compounds for the pharmaceutical and chemical industries. For the identification of thermostable glycoside hydrolases (GHs), a system for complementation-based selection in the extremely thermophilic bacterium *Thermus thermophilus* was developed. With the help of recombinant expression of GHs in this host organism, the principle function of this system could be demonstrated. The wild type form of CglT (a GH of the enzyme family GH1) of *Thermoanaerobacter brockii* and a thermally more stable mutant CglT KIR were used as models. In addition to the host's own β -glucosidase, both CglT variants were able to restore the ability of growth of a GH deletion mutant in minimal medium with D-cellobiose.

In the second part of this work, the protein thermal stability – the so-called meltome – of *T. thermophilus* and five other microorganisms (*Oleispira antarctica*, *Bacillus subtilis*, *Escherichia coli*, *Geobacillus stearothermophilus* und *Picrophilus torridus*) was analysed on a proteomic level. These organisms have a wide range of different growth temperature preferences; from psychrophilic to extremely thermophilic. It has been shown that in general the length of a protein correlates negatively with its thermostability, but properties such as hydrophobicity and amino acid frequency are not indicators of thermostability in solubilised cells. The comparison of proteins from cell extracts and intact cells of *E. coli* and *T. thermophilus* showed that predominantly membrane proteins undergo stabilisation under *in vivo* conditions.

Abstract

The comparison of the denaturation behaviour of paralogs and orthologs showed that the melting temperatures converged with increasing sequence identity. This correlation indicates that thermostability is an evolutionary conserved property. Furthermore, it could be deduced that thermal (de-)stabilisation is caused by a few amino acids at certain key positions. In addition, the available meltome data, in combination with cell inactivation experiments and dynamic differential fluorimetry for *E. coli* and *T. thermophilus*, could be used to shed light on the reasons for thermal cell inactivation. By enriching gene ontology terms of those proteins which denatured in the range of thermal cell inactivation, it was indicated that the loss of DNA/RNA-binding proteins and ribosomal structure play a role in this process and that these protein groups could therefore be of importance for thermal cell death.

1 EINLEITUNG

1.1 Glycosidhydrolasen

Pflanzliches Material – im Speziellen die pflanzliche Zellwand – stellt das größte Vorkommen organischen Kohlenstoffs unseres Planeten an Land dar. Einen Großteil davon repräsentiert Cellulose, ein Polysaccharid aus β -1,4-glycosidisch verknüpften D-Glucoseeinheiten. Um diese in kurzkettigere Oligosaccharide oder Monosaccharide abzubauen, um daraus durch den zellulären Stoffwechsel Energie zu gewinnen, sind Glycosidhydrolasen (GHs) nötig. Enzyme dieser Enzymklasse hydrolysieren die glycosidischen Bindungen zwischen zwei oder mehreren Zuckermolekülen bzw. zwischen einer Zuckereinheit und einem Nichtkohlenhydratbestandteil. Im Zusammenspiel mit anderen Enzymen wie Transglycosidasen, Polysaccharidlyasen und Kohlenhydratesterasen fungieren sie als Schlüsselsystem im Polymerabbau unterschiedlichster Ökosysteme¹.

GHs werden anhand der Ähnlichkeit ihrer Aminosäuresequenzen in GH-Familien und basierend auf Übereinstimmungen ihrer Tertiärstrukturen in Clans eingeteilt^{2,3}. Die Datenbank CAZy (Carbohydrate-Active enZYmes database) gibt hierbei einen Überblick und teilt zum aktuellen Zeitpunkt dieser Arbeit (Juli 2019) GHs in 165 unterschiedliche Familien bzw. 18 Clans ein⁴.

Einleitung

1.2 Mikroorganismen und ihr Temperaturspektrum

(Mikro-) Organismen lassen sich anhand der Temperaturbereiche/-optima ihres Wachstums in unterschiedliche Gruppen einteilen: man unterscheidet anhand des Temperaturoptimums (Temperatur maximaler Wachstumsrate) *Psychrophile* (niedriges Temperaturoptimum unter 20 °C), *Mesophile* (mittleres Temperaturoptimum – ähnlich der Körpertemperatur von Warmblütern; etwa zwischen 25 und 45 °C) und *Thermophile*. Letztere spalten sich wiederum auf in⁵:

- (1) (moderat) Thermophile (Temperaturoptimum zwischen 50 und 64 °C; Wachstum ist nur unterhalb von 70 °C möglich)
- (2) Extrem Thermophile (optimales Wachstum zwischen 65 und 85 °C; die Maximaltemperatur für Wachstum liegt oberhalb von 70 °C)
- (3) Hyperthermophile (wachsen optimal bei Temperaturen über 85 °C; Wachstum ist noch über 90 °C möglich)

Die genannten Temperaturbereiche sind empirisch vorgeschlagene Werte, manche Forscher setzen die Grenze für hyperthermophiles Wachstum auch bei einem Temperaturoptimum von 80 °C oder höher. Bis heute sind keine Eukaryoten bekannt, welche oberhalb 62 °C gutes Wachstum aufweisen. Somit sind die Bezeichnungen extrem Thermophile und Hyperthermophile ausschließlich Bakterien und Archaeen vorbehalten. In den obersten Temperaturebereichen sind jedoch nur Archaeen vertreten und unter ihnen ist *Methanopyrus kandleri* Rekordhalter mit maximalen 122 °C⁶.

Unter den extrem thermophilen Bakterien zählt *Thermus* sp. als eine der bekanntesten Gattungen⁷. Den wohl prominentesten Vertreter stellt *Thermus aquaticus* dar⁸. Isoliert und beschrieben vor beinahe einem halben Jahrhundert, fand dessen thermostabile DNS-Polymerase – die *Taq*-Polymerase – Anwendung in der Polymerase-Kettenreaktion (PCR) und ermöglichte seither die *in vitro* Vervielfältigung genetischen Materials⁹.

Einleitung

1.2.1 *Oleispira antarctica*

Oleispira antarctica wurde aus einer Umweltprobe eines Küstenabschnitts des Rossmeers in der Antarktis isoliert¹⁰. Es handelt sich um psychrophile, aerobe sowie gram-negative Spirillen mit polaren Flagellen. Das Genom umfasst 4,4 MBp mit einem GC-Gehalt von 41,6 % und kodiert für insgesamt 3.919 Proteine¹¹. *O. antarctica* ist durch seine Alkan-Monooxygenase und Desaturasen für die Forschung sowie Industrie von besonderer Bedeutung. Aktuelle Arbeiten orientieren sich daran, das Potential zum Kohlenwasserstoffabbau von *Oleispira* und verwandter Bakterien zu nutzen, um mögliche Umweltverschmutzungen durch Öl zu beseitigen^{12,13}.

Mit einer optimalen Wachstumstemperatur von 2-4 °C bildet dieses Bakterium das untere Ende im Temperaturspektrum aller in dieser Arbeit vertretenen Organismen.

1.2.2 *Bacillus subtilis*

Bereits 1835 wurde *Bacillus subtilis* erstmals durch CHRISTIAN G. EHRENBERG beschrieben. Allerdings ist man sich über den genauen Ursprung heute verwendeter Stämme im Unklaren¹⁴. Durch seine unter anderem bereits lange Bekanntheit und Verfügbarkeit, wird *B. subtilis* als das eingehendste untersuchte Gram-positive Bakterium betrachtet. Es handelt sich um ubiquitär vorkommende, strikt aerobe, stäbchenförmig Endosporenbildner mit peritricher Begeißelung^{15,16}. Insgesamt bilden 4,2 MBp das Genom, welches für 4.100 Proteine kodiert¹⁷. Der durchschnittliche GC-Gehalt liegt bei 43,5 %.

Die bisherigen Anwendungsbereiche sind und waren für *B. subtilis* sehr breit gefächert; vom Einsatz als biologische UV-Dosimeter in der Raumfahrt bis hin zur Fermentation von Lebensmitteln^{18,19}.

Einleitung

1.2.3 *Escherichia coli*

Das im Jahr 1884 von seinem Namensgeber THEODOR ESCHERICH isolierte, Gram-negative und fakultativ anaerobe Darmbakterium *Escherichia coli* ist der ausführlichste studierte und bestverstandene Organismus unseres Planeten²⁰. Mit seiner Hilfe wurden bereits vor mehr als einem halben Jahrhundert unter anderem fundamentale Erkenntnisse um den genetischen Code, Transkription, Translation und Replikation gewonnen²¹⁻²³. Das Genom wurde 1997 erstmals vollständig sequenziert und umfasst 4,6 MBp bei einem GC-Gehalt von 50,8 %^{24,25}. Insgesamt kodiert es für 4.288 Proteine (*E. coli* K-12 MG1655).

Seine einfache Handhabung und kurze Generationszeit machen dieses Bakterium zum optimalen Studienobjekt. Zudem verhalten die einfache Transformierbarkeit und das Potential, Proteine in großen Mengen rekombinant zu exprimieren, in akademischer Forschung und Industrie zu hoher Popularität^{26,27}. Auch im Rahmen dieser Arbeit wurden diverse *E. coli* Stämme kultiviert, um Klonierungsarbeiten oder Genexpressionsexperimente durchzuführen.

1.2.4 *Geobacillus stearothermophilus*

Erstmalig wurde *G. stearothermophilus* (früher *Bacillus stearothermophilus*) 1920 aus einer verdorbenen Konservendose mit Nahrungsmitteln isoliert²⁸. Dieses Bakterium zählt als erstes unter den bisher Aufgeführten zur Kategorie der Thermophilen; seine optimale Wachstumstemperatur liegt bei 55 °C. Sein Gramverhalten ist positiv. Es ist in der Lage Sporen zu bilden und sein Stoffwechsel verläuft fakultativ anaerob^{29,30}.

An der Gattung *Geobacillus* herrscht starkes biotechnologisches Interesse. Beispielhaft sind die Bereiche Biokraftstoffe, mikrobiologische Sanierung von Böden sowie die Produktion thermostabiler Enzyme zu nennen³¹⁻³³. Entsprechend seinem historischen Erbe findet dieser Organismus wegen seiner Fähigkeit zur Bildung hitzeresistenter Endosporen auch heute noch Anwendung in von Qualitätskontrolle von Sterilisationsprozessen der Lebensmittelindustrie³⁴.

Einleitung

Zum Zeitpunkt dieser Arbeit liegt für den hier verwendeten *G. stearothermophilus* Stamm ATCC 12980 noch kein offizielles Referenzgenom vor. Die aktuell verfügbare Genomsequenz (*draft genome*) umfasst 2,6 MBp mit insgesamt 2.815 Genen, wovon wiederum 2.552 für Proteine kodieren^{35,36}.

1.2.5 *Picrophilus torridus*

Aufgrund seiner optimalen Wachstumstemperatur von 60 °C in Kombination mit einem sehr niedrigen pH-Optimum von 0,7, zählt dieses Archaeon zu den Thermoacidophilen³⁷. *P. torridus* ist sogar in der Lage, bei einem pH-Wert von -0,06 (entspricht 1,2 M Schwefelsäure) zu wachsen, was einem Rekordwert entspricht³⁸. Dies wird durch eine für Protonen nur gering durchlässige Membran sowie weiterer Anpassungen an extrem saure Bedingungen möglich gemacht^{37,39}. Der pH-Wert des Cytoplasmas von *P. torridus* ist mit 4,6 selbst für einen thermoacidophilen Organismus außergewöhnlich niedrig, da andere Acidophile nahe am Neutralbereich liegen.

Die Sequenzierung des Genoms ergab eine Größe von knapp 1,6 MBp³⁷. Von den 1.614 Genen kodieren 1.563 für Proteine⁴⁰.

Einleitung

1.2.6 *Thermus thermophilus*

Isoliert wurde *Thermus thermophilus* von OSHIMA & IMAHORI^{41,42} im Jahre 1968 aus einer heißen Quelle in Mine (Präfektur Shizuoka, Japan) (Abbildung 1).

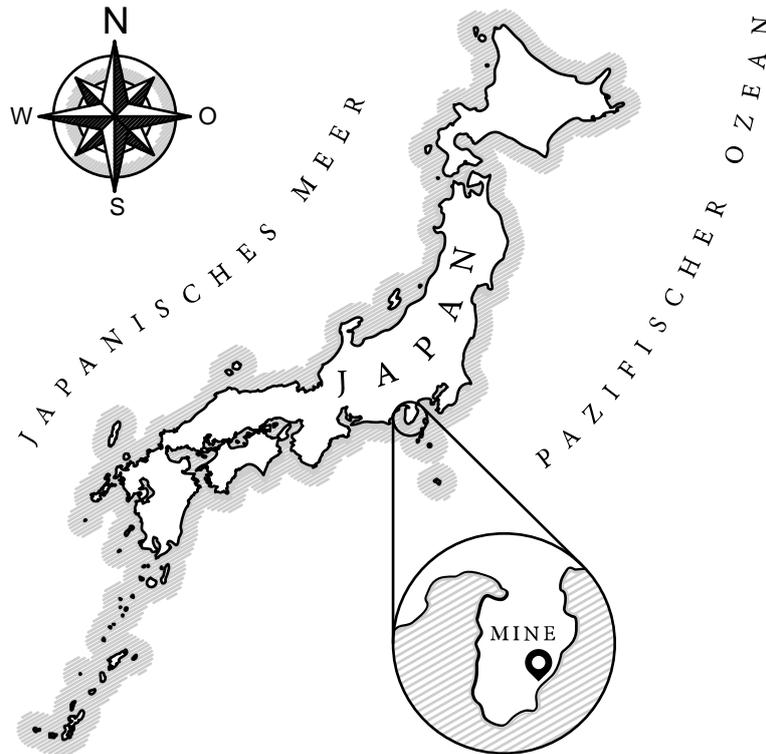


Abbildung 1: Isolationsort des extrem thermophilen Bakteriums *Thermus thermophilus*.

Es handelt sich um ein aerobes, stäbchenförmiges und Gram-negatives Bakterium, welches mit seiner optimalen Wachstumstemperatur von 70 °C den extrem Thermophilen angehört. Taxonomisch ist *T. thermophilus* dem Phylum *Deinococcus-Thermus* zuzuordnen⁴³. Die mit der hohen Wachstumstemperatur einhergehende, erhöhte thermische und chemische Proteinstabilität hat man sich zunutze gemacht, um mit *T. thermophilus* als Modellorganismus die Strukturbiologie voranzutreiben⁷. So war es z.B. dank *T. thermophilus* möglich, zum ersten Mal die Tertiärstruktur einer kompletten ribosomalen 30S Untereinheit und eines 70S Ribosoms kristallographisch zu bestimmen^{44,45}. Der Einzug in Labore rund um den Globus verdankt *T. thermophilus* unter anderem auch dem Shuttlevektor pMK18⁴⁶. Dieser von pUC18

Einleitung

abgeleitete Vektor ermöglicht die Verwendung von *E. coli* als Instrument der Herstellung diverser Plasmidkonstrukte und einer Gewinnung dieser in ausreichend großen Mengen. Dies erlaubt einen schnelleren und einfacheren Arbeitsablauf, da die Handhabung von *E. coli* im Gegensatz zu *T. thermophilus* in vielen Arbeitsgruppen Routine darstellt.

Industriell sind Thermophile und deren Enzyme von besonderem Interesse. Manche Reaktionen sind erst bei höheren Temperaturen realisierbar bzw. wirtschaftlich. Zudem erlaubt die Durchführung von biotechnologischen Prozessen mit Thermophilen einen (teilweisen) Verzicht auf eine Prozesskühlung, welche beim Einsatz von Mesophilen wegen der Aufheizung durch exotherme Stoffwechselreaktionen meist unabdinglich ist⁴⁷. Die hohe Kompetenz zur natürlichen Transformation und eine relativ hohe Wachstumsrate empfiehlt die Nutzung von *T. thermophilus* zur Entwicklung neuer Anwendungen^{7,48,49}. Einen großen Vorteil bietet *T. thermophilus* als Expressionswirt thermostabiler Proteine, denn die auf hohe Temperaturen ausgelegte Zellmaschinerie verspricht für bestimmte Zielgene aus Thermophilen eine höhere Erfolgchance auf eine funktionstüchtige rekombinante Expression⁵⁰. Dies steigert die Chancen einer erfolgreichen, funktionellen Analyse von Metagenombanken unter der Nutzung von *T. thermophilus* als alternatives Wirtsbakterium.

HENNE *et al.* gelang 2004 die Entschlüsselung des Genoms, welches durch ein Chromosom (TTC) von 1,89 MBp und einem Megaplasmid (pTT27) mit einer Länge von 233 kBp gebildet wird (Abbildung 2)⁵¹. Mit 69,4 % liegt der GC-Gehalt über dem für Bakterien üblichen Durchschnitt. Insgesamt liegen 2.227 Gene vor, wovon 2.126 für Proteine kodieren.

Einleitung

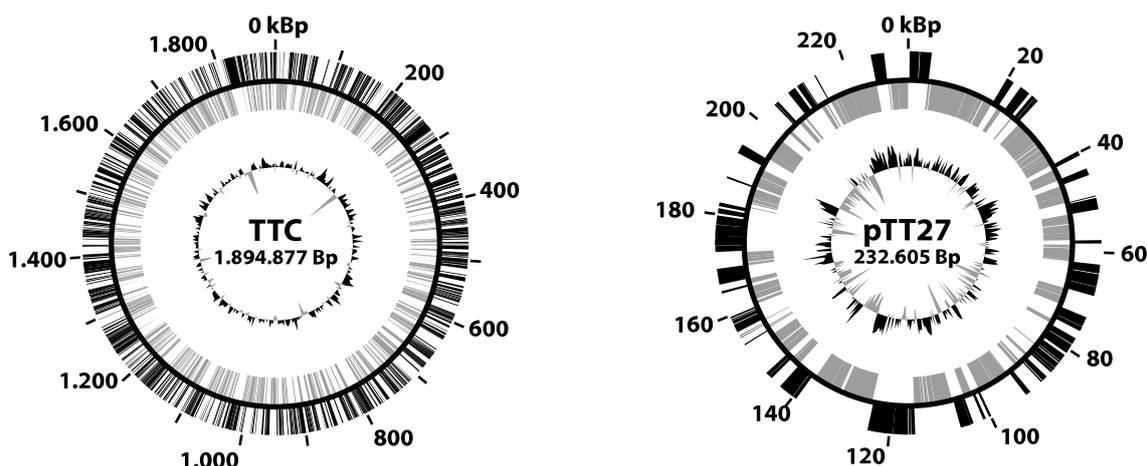


Abbildung 2: Karten des Chromosoms (TTC) und Megaplasמידs (pTT27) von *T. thermophilus* HB27. Die kodierenden Sequenzen liegen im äußeren Bereich und sind je nach Strangorientierung in Schwarz oder Grau gehalten. Die inneren Kreise geben den GC-Gehalt wieder (kleinere Werte liegen weiter innen).

Die Zellhüllstruktur von *T. thermophilus* besteht aus insgesamt vier Schichten⁵². Hierzu zählen eine Cytoplasmamembran, eine dünne Peptidoglycanschicht, eine Zwischenschicht bestehend aus Sekundärzellwandpolymer (*engl.*: SCWP –secondary cell wall polymer) und einem Komplex aus äußerer Membran mit aufgelagertem SlpA-Protein (S-layer protein)⁵³. Das Gen für dieses SlpA-Protein verfügt über einen starken Promotor (P_{slpA}), welcher ebenfalls für die rekombinante Expression von Proteinen – so auch in dieser Arbeit – in *T. thermophilus* eingesetzt wird^{7,54–56}.

1.3 Auf Massenspektrometrie basierende Proteomik

Die Möglichkeit, mit Massenspektrometrie ein hochkomplexes Gemisch unterschiedlichster chemischer Verbindungen in einem einzigen Experiment zu analysieren, macht diese Analysetechnik in vielen Wissenschaftszweigen so wertvoll. Entsprechend der Anforderungen an die Analyse von Protein- bzw. Peptidgemischen, wurde diese Technologie über die Zeit individuell weiterentwickelt.

Die methodischen Möglichkeiten sind inzwischen breit gefächert. Aus diesem Grund werden im Folgenden nur die für die durchgeführten Experimente dieser Arbeit relevanten Methoden näher ausgeführt.

Einleitung

Da bei Massenspektrometrie das Masse-zu-Ladungsverhältnis (m/z -Verhältnis) der Moleküle und nicht deren unmittelbare Masse bestimmt wird, ist vorgeschaltete Ionisierung essenziell. Mit der Entwicklung der *Elektronensprayionisation* (ESI) Ende der 1980er Jahre wurde die Ionisierung und der gleichzeitige Transfer in Lösung befindlicher Analyten in die Gasphase realisiert⁵⁷⁻⁶⁰. Der Analyt wird durch eine Kapillare geleitet, an welche eine Spannung von mehreren tausend Volt angelegt wird. Durch das hierbei entstehende elektrische Feld zwischen Kapillare und Gegenelektrode, werden positiv geladene Ionen Richtung Kathode gezogen. Infolge dessen bildet der Analyt einen Konus (TAYLOR-Konus), welcher sich so weit verjüngt, bis einzelne geladene Flüssigkeitströpfchen entstehen. Diese wiederum zerteilen sich in immer feinere Tröpfchen, bis am Ende einzelne, ein- und mehrfach geladene Ionen in Gasphase verbleiben. Nach dem Auftrennen der Ionen im Massenanalysator (diverse Methoden: s.u.), erfolgt die Erfassung dieser über einen Ionendetektor/-verstärker. Abhängig von der Bauart der Apparatur, wird zusätzlich ein Detektor benötigt (z.B. Sekundärelektronenvervielfacher oder FARADAY-Becher) oder ist baulich bedingt bereits im Massenanalysator enthalten (z.B. Orbitrap oder PAUL-Falle).

Am Ende einer jeden massenspektrometrischen Analyse steht der computergestützte Abgleich ermittelter Werte mit einer (großen) Datenbank^{61,62}. Hierfür ist es wichtig, so viele Informationen wie möglich über eine Probe zu erhalten. Im Falle der Proteomik wird daher die *Tandemmassenspektrometrie* (MS/MS) herangezogen⁶³. In einem ersten Schritt wird das m/z -Verhältnis eines intakten Peptids bestimmt (MS1), bevor es im Anschluss fragmentiert wird um aus den resultierenden Fragmenten die genaue Aminosäuresequenz in einer weiteren massenspektrometrischen Analyse (MS2) zu bestimmen. Da beide MS-Schritte nicht zwingend aneinandergelagert sind, können verschiedene Massenspektrometer für beide Messungen gekoppelt werden. QTOF (*engl.*: quadrupole-time-of-flight), QQQ (*engl.*: triple quadrupole) oder aber auch TOF-TOF sind häufig verwendete Kombinationen⁶⁴.

Im Rahmen dieser Arbeit wurde ein Hybrid-Quadrupol-Orbitrap Massenspektrometer verwendet. Es handelt sich um eine Kombination aus einem Massefilter (Quadrupol), einer HCD-Kammer (*engl.*: higher-energy collisional dissociation cell) zur Fraktionierung der

Einleitung

Molekülonen und einem Orbitrap-Massenspektrometer⁶⁵⁻⁶⁸. Letzteres dient der Aufzeichnung von MS²-Spektren. Die Ionen oszillieren hierbei spiralförmig um eine spindelförmige Elektrode und induzieren dadurch eine messbare Spannung an einer äußeren Elektrode. Mit Hilfe von Fourier-Transformation kann im Anschluss das m/z -Verhältnis der Ionen aus diesem Signal abgeleitet werden.

Durch eine der Ionenerzeugung (ESI) vorangestellte *Umkehrphasenchromatographie* (engl.: reversed phase liquid chromatography), lassen sich die Peptide im Vorfeld in einer weiteren Dimension auftrennen und die Gesamtleistung der massenspektrometrischen Analyse steigern. Der Analyt wird über eine unpolare stationäre Phase geleitet, während die mobile Phase polare Eigenschaften aufweist. Insgesamt spricht man auch von einem LC-MS/MS System.

1.4 Thermal proteome profiling (TPP)

Die Lebensfähigkeit von Organismen ist von der Temperatur ihrer Umgebung abhängig. Die enzymatischen Prozesse einer Zelle sind daher für einen bestimmten Temperaturbereich optimiert. Dies lässt sich im Falle von Bakterien anhand der Zellteilungsrate relativ einfach abbilden⁶⁹. Bedingt durch die *thermische Denaturierung* (im Folgenden als *Denaturierung* bezeichnet) von Proteinen tritt ab einer gewissen Temperatur oberhalb des Wachstumsoptimums der Zelltod ein. Bereits kleine thermische Änderungen können zur (partiellen) Entfaltung und somit zur Exposition hydrophober Aminosäurereste des Proteininneren führen. Als Folge davon aggregieren diese Proteine irreversibel, was somit zu deren Funktionsverlust führen.

Um die thermische Stabilität einzelner Proteine in gereinigter Form zu bestimmen, bieten sich diverse Methoden an, die Schmelztemperatur (T_m) zu ermitteln⁷⁰⁻⁷². Behandelt man Proben des untersuchten Proteins bei unterschiedlichen Temperaturen und zentrifugiert diese, so kann man über SDS-PAGE (SDS-Polyacrylamid-Gelelektrophorese) die verbleibende Menge nativen Proteins im Überstand bestimmen. Aus der Auswertung der Intensitäten der Proteinbanden lässt sich dann die Schmelztemperatur eines Proteins ableiten. In der Forschung

Einleitung

werden auch diverse spektroskopische Verfahren wie Fluoreszenzspektroskopie und Circular dichroismus (CD) für die Analyse von Proteinfaltung eingesetzt.

Strebt man jedoch an, die individuellen Schmelztemperaturen eines Gemischs von Proteinen bzw. eines ganzen Proteoms zu bestimmen, so gestaltet sich dies wesentlich komplexer. Da die Denaturierung eines jeden Proteins bei unterschiedlichen Temperaturen eintritt, stellt sich – betrachtet man das gesamte Proteom – die zellweite Denaturierung als kontinuierlicher Prozess dar. Üblicherweise verwendete Verfahren (s.o.) sind hierfür nicht hinreichend leistungsfähig.

Thermal proteome profiling (TPP) bietet hierbei einen kombinatorischen Ansatz aus Denaturierung bei unterschiedlichen Temperaturen und Proteomik unter Einsatz von Massenspektrometrie^{73,74}. Zellextrakt oder intakte Zellen werden aliquotiert und die einzelnen Proben für eine bestimmte Zeit bei unterschiedlichen Temperaturen inkubiert. Man geht hierbei i.d.R. von einem Temperaturbereich, welche von der optimalen Wachstumstemperatur bis 30 °C darüber reicht, aus. Daraus resultiert eine graduelle Aggregation aller Proteine, bis theoretisch keine nativen Proteine mehr in Lösung vorliegen. Mittels Ultrazentrifugation lassen sich die aggregierten Proteine pelletieren, wodurch die löslichen (nativen Proteine) im Überstand verbleiben.

Die unterschiedlichen Proben werden, um diese später nach der massenspektrometrischen Analyse unterscheiden zu können, mit *Tandem-Atommassenmarkierungen* (engl.: TMT – tandem mass tags) versehen⁷⁵⁻⁷⁸. Diese *Markierungen* (engl.: labels) bestehen im Wesentlichen aus (1) einer reaktiven Gruppe zur Kopplung des Labels an das Zielmolekül, (2) einem Massenreporterbereich, (3) einem Massennormalisierungsbereich und (4) einem Linker-Bereich welcher spaltbar ist. In dieser Arbeit wurde im Speziellen mit dem TMT10-plex System gearbeitet. Es handelt sich um 10 verschiedene Labels mit jeweils einer anderen Isotopenverteilung in der Reporterregion. Der Massennormalisierungsbereich sorgt wiederum dafür, dass alle Spezies isobar sind – sprich die gleiche Masse aufweisen. Während der Fraktionierung der Tandem-Massenspektrometrie, wird die Linker-Region gespalten und die Reporterionen unterschiedlicher Masse freigesetzt. Diese befinden sich im unteren m/z -

Einleitung

Bereich (Masse-zu-Ladungsverhältnis) im Spektrum und eignen sich zur indirekten, relativen Quantifizierung fraktionierter Peptide.

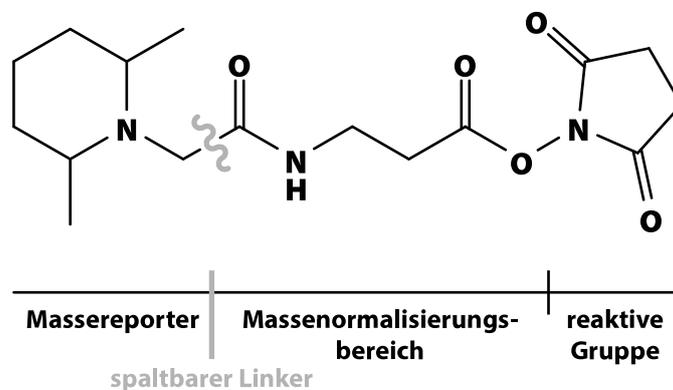


Abbildung 3: Grundlegender molekularer Aufbau eines TMT10plex-Labels. Die unterschiedlichen Isotopenspezies sind nicht dargestellt. Unterschiedliche Isotopenverteilungen zwischen Massereporter und Massennormalisierungsbereich resultieren in 10 unterschiedlichen Massereporterionen (126–131 m/z) nach der Kollisionsfraktionierung. In Anlehnung an die Bedienungsanleitung „Thermo Fisher Scientific TMT10plex™ Isobaric Label Reagent“.

Somit lassen sich in einem Schritt die Schmelztemperaturen von Proteinen eines komplexen Gemischs – wie es bei einer proteomweiten Analyse der Fall ist – bestimmen.

1.5 Cellular Thermal Shift Assay (CETSA)

Jüngst wurde die Methode des *cellular thermal shift assay* (CETSA) entwickelt⁷⁹. Hierbei können auf einer proteomweiten Ebene die Interaktion von Proteinen mit anderen Molekülen erfasst werden.

CETSA macht sich zu Nutzen, dass Proteine durch Wechselwirkungen mit anderen Molekülen (thermische) Stabilität gewinnen⁸⁰. Hierbei werden Effekte durch das Zusammenspiel wie etwa mit anderen Proteinen, niedermolekularen Verbindungen (z.B. Metabolite, Wirkstoffe, etc.) oder intrazellulären Polymeren (DNS, Kohlenhydrate, etc.) betrachtet. Man vergleicht ein Referenz-TPP-Experiment mit einem TPP-Experiment gewünschter Bedingungen (z.B. Zugabe von niedermolekularen Substanzen). Die Abweichungen der Schmelzpunkte von der Referenzprobe zeigen somit direkte bzw. indirekte Interaktionen mit diesen Proteinen. CETSA-Experimente könnten mit ganzen Zellen und mit Zellextrakten durchgeführt werden.

Einleitung

Aufwändiger ist die Verwendung ganzer Zellen, da die Zellaliquote nach der thermischen Behandlung z.T. einzeln aufgeschlossen werden müssen. Jedoch lassen sich in ganzen Zellen Interaktionen tiefergehend feststellen, da z.B. Enzyme betroffener Stoffwechselwege gemeinsam sichtbar beeinflusst werden. Die direkte Wechselwirkung mit einem Protein bedingt ebenfalls eine Änderung der Schmelztemperatur der Interaktionspartner. Im Falle von Zellextrakt sind meist nur die primären Auswirkungen sichtbar, da unter den *in vitro*-Bedingungen die Wechselwirkungen zwischen den Proteinen nicht unbedingt so bestehen wie *in vivo*.

In jüngster Zeit gewann diese Technologie besonders in der Wirkstoffentwicklung, der Untersuchung von Stoffwechselwegen und der Betrachtung von Proteinfaltungsdynamik verstärkt an Bedeutung^{74,81-85}.

1.6 Ziel der Arbeit

Eines der Hauptziele dieser Arbeit war die Entwicklung eines komplementationsbasierten Selektionssystems für den Einsatz von *T. thermophilus* HB27 als Wirtsbakterium für funktionelle Metagenomanalyse zur Gewinnung neuartiger, thermostabiler Glycosidhydrolasen (GH). Als Quelle dieser Enzyme sollten komplexe Metagenombanken aus heißen Habitaten oder Genbanken, welche aus Mutageneseexperimenten hervorgehen, dienen. Mithilfe eines solchen Selektionssystems sollten – im Vergleich zu Durchmusterungsverfahren – erheblich Zeit und Ressourcen eingespart werden können. Zugleich erfolgt durch dieses *in vivo*-System eine direkte Identifikation biologisch aktiver Enzymvarianten.

Für die Etablierung eines solchen Systems war zunächst die Entwicklung eines geeigneten Minimalmediums für *T. thermophilus* grundlegend. Außerdem wurde eine GH-Deletionsmutante, welche differenzierbares Wachstum in diesem Minimalmedium zeigt, benötigt. Der prinzipielle Beweis für die Funktionalität sollte durch die Komplementation des Mutantenphänotyps mittels heterologer Expression eines geeigneten GH-Gens, sowie eines Gens für eine thermostabilere Variante der GH, erbracht werden.

Einleitung

Für die künftige Nutzung von *T. thermophilus* als extrem thermophilen Genexpressionswirt bei dessen optimaler Wachstumstemperatur, aber ggf. auch bei noch höheren Temperaturen, ist es von Interesse zu untersuchen, wie sich das Proteom dieses Wirtsbakteriums bei hohen Temperaturen verhält. Im zweiten Teil dieser Arbeit stand die Untersuchung proteomweiter Thermostabilität von *T. thermophilus* und zusätzlich einer Auswahl von Organismen eines weit gefächerten Spektrums unterschiedlicher Wachstumstemperaturen im Fokus. Die deskriptive und analytische Exploration dieser sogenannten Meltsome sollte unter anderem die Frage beantworten, welche Charakteristika zu höherer Thermostabilität führen. Hierfür sollten Proteine im unteren und oberen Temperaturbereich anhand unterschiedlicher physikalisch-chemischer Eigenschaften einander gegenübergestellt werden. Ein Vergleich von intakten Zellen mit Zellextrakten sollte die Frage beantworten, welche Proteine unter *in vivo*-Bedingungen thermisch stabilisiert werden. Ebenso stand ein grundlegender Erkenntnisgewinn über Proteinstabilität mit der angewendeten und relativ jungen Methodik im Fokus. Die Korrelation von Aminosäurefrequenzen mit Proteinschmelzpunkten sollte die Frage klären, ob thermisch stabilere Proteine eine Über- oder Unterrepräsentation einer oder mehrerer Aminosäuren zeigen. Tiefe Einblicke in die evolutive Konservierung von Thermostabilität sollten durch die Untersuchung von Paralogen und Orthologen erlangt werden. Dies spielt auch im Kontext des ersten Teils der Arbeit eine wichtige Rolle, zur Klärung der Frage, in wie weit es möglich ist ein „beliebiges“ Enzym zu thermostabilisieren. Abschließend sollte anhand der erhobenen Meltsomedaten, in Kombination mit Zellinaktivierungsexperimenten und dynamischer Differenzfluorimetrie (*engl.*: DSF – differential scanning fluorimetry), nach der Ursache für die thermische Inaktivierung von Mikroorganismen gesucht werden.

2 MATERIAL & METHODEN

2.1 Chemikalien

Tabelle 3: Auflistung verwendeter Chemikalien.

Name	Hersteller
Agarose LE	Biozym Scientific GmbH (Hessisch Oldenburg, Deutschland)
Bleomycinsulfat	Tokyo Chemical Industry Co., Ltd. (Tokio, Japan)
Bradford Reagenz, 5 ×	SERVA Electrophoresis GmbH (Heidelberg, Deutschland)
CaCl ₂	J.T. Baker, Inc. (Phillipsburg, NJ, USA)
CaSO ₄	
CoCl ₂	Sigma-Aldrich Corp. (St. Louis, MO, USA)
CuSO ₄	Sigma-Aldrich Corp. (St. Louis, MO, USA)
D-Cellobiose	Sigma-Aldrich Corp. (St. Louis, MO, USA)
D-Maltose	Sigma-Aldrich Corp. (St. Louis, MO, USA)
DMF	AppliChem GmbH (Darmstadt, Deutschland)
DMSO	AppliChem GmbH (Darmstadt, Deutschland)
D-Saccharose	Merck KGaA (Darmstadt, Deutschland)
EDTA	AppliChem GmbH (Darmstadt, Deutschland)
Essigsäure	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
FeCl ₃	Merck KGaA (Darmstadt, Deutschland)
Fleischextrakt	VWR International (Radnor PA, USA)
Gold Agarose	Biozym Scientific GmbH (Hessisch Oldenburg, Deutschland)
H ₃ BO ₃	AppliChem GmbH (Darmstadt, Deutschland)
Hefeextrakt	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
KCl	J.T. Baker, Inc. (Phillipsburg, NJ, USA)
KH ₂ PO ₄	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
KNO ₃	J.T. Baker, Inc. (Phillipsburg, NJ, USA)

Material & Methoden

Name	Hersteller
L-Lysin	Sigma-Aldrich Corp. (St. Louis, MO, USA)
L-Methionin	SERVA Electrophoresis GmbH (Heidelberg, Deutschland)
L-Prolin	AppliChem GmbH (Darmstadt, Deutschland)
L-Serin	SERVA Electrophoresis GmbH (Heidelberg, Deutschland)
MgCl ₂	VWR International (Radnor PA, USA)
MgSO ₄	Sigma-Aldrich Corp. (St. Louis, MO, USA)
MnSO ₄	Merck KGaA (Darmstadt, Deutschland)
MOPS	AppliChem GmbH (Darmstadt, Deutschland)
Na ₂ HPO ₄	J.T. Baker, Inc. (Phillipsburg, NJ, USA)
Na ₂ MoO ₄	Merck KGaA (Darmstadt, Deutschland)
NaCl	AppliChem GmbH (Darmstadt, Deutschland)
NaH ₂ PO ₄	Merck KGaA (Darmstadt, Deutschland)
NaNO ₃	Merck KGaA (Darmstadt, Deutschland)
Nonidet P-40	Sigma-Aldrich Corp. (St. Louis, MO, USA)
NTA	Sigma-Aldrich Corp. (St. Louis, MO, USA)
Pepton	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
Pepton (aus Casein)	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
Pepton (aus Sojamehl)	VWR International (Radnor PA, USA)
PIPES	Sigma-Aldrich Corp. (St. Louis, MO, USA)
<i>p</i> -Nitrophenyl β-D-glucopyranoside	Carbosynth Ltd. (Compton, England)
Tris-HCl	AppliChem GmbH (Darmstadt, Deutschland)
Trypton	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
X-Gal	Carbosynth Ltd. (Compton, England)
ZnSO ₄	Sigma-Aldrich Corp. (St. Louis, MO, USA)

2.2 Nährmedien

Alle unterstehenden Medien wurden – sofern nicht abweichend angegeben – zur Sterilisation bei 121 °C für 15 min (Systec VX-150) bzw. 20 min (VARIOKLAV® 25T) autoklaviert.

Tabelle 4: Zusammensetzung verwendeter Nährmedien (Voll- und Minimalmedien).

Name	Bestandteile	Mengenangabe
LB-Medium (pH 7,2)	Trypton	10 g
	Hefeextrakt	5 g
	NaCl	5 g
	H ₂ O (VE)	ad 1.000 mL
Medium 1 (pH 7,0)	Trypton	5 g
	Fleischextrakt	3 g
	H ₂ O (VE)	ad 1.000 mL
Medium 220 (pH 7.3)	Pepton (aus Casein)	15 g
	Pepton (aus Sojamehl)	5 g
	NaCl	5 g
	H ₂ O (VE)	ad 1.000 mL
Medium 514 (pH 7,6)	Trypton	5,00 g
	Hefeextrakt	1,00 g
	Fe(III)-Citrat	0,10 g
	NaCl	19,45 g
	MgCl ₂ (Anhydrat)	5,90 g
	Na ₂ SO ₄	3,24 g
	CaCl ₂	1,80 g
	KCl	550,00 mg
	NaHCO ₃	160,00 mg
	KBr	80,00 mg
	SrCl ₂	34,00 mg
	H ₃ BO ₃	22,00 mg
	Na-silicate	4,00 mg
	NaF	2,40 mg
	(NH ₄)NO ₃	1,60 mg
	Na ₂ HPO ₄	8,00 mg
H ₂ O (VE)	ad 1.000,00 ml	

Material & Methoden

Name	Bestandteile	Mengenangabe
SOB (pH 7,0)	Trypton	20,0 g
	Hefeextrakt	5,0 g
	NaCl	0,5 g
	H ₂ O (Reinstwasser)	ad 1.000,0 mL
	Nach dem Autoklavieren wurden untenstehende Lösungen (sterilfiltriert) hinzugegeben	
	KCl	2,5 mM
	MgCl ₂	10,0 mM
	MgSO ₄	10,0 mM
TB-Medium (pH 7,5)	Trypton	8 g
	Hefeextrakt	4 g
	NaCl	3 g
	H ₂ O (Aqua Purania)	ad 1.000 mL
TMB-Medium (Thermus minimal broth) (pH 8,2) nach CASTENHOLZ ⁸⁶ mit Änderungen Siehe untenstehenden Kommentar zur Herstellung	Lösung A (Endkonzentrationsangabe)	
	NaNO ₃	689 mg
	Na ₂ HPO ₄ · 2H ₂ O	111 mg
	KNO ₃	103 mg
	MgSO ₄ · 7H ₂ O	100 mg
	NTA	100 mg
	CaSO ₄ · 2H ₂ O	60 mg
	NaCl	800 µg
	H ₃ BO ₃	500 µg
	ZnSO ₄ · 7H ₂ O	500 µg
	FeCl ₃ · 6H ₂ O	300 µg
	Lösung B (Endkonzentrationsangabe)	
	MnSO ₄ · H ₂ O	228 µg
	CoCl ₂ · 6H ₂ O	45 µg
CuSO ₄	25 µg	
Na ₂ MoO ₄ · 2H ₂ O	25 µg	
sterilfiltriert, nach dem Autoklavieren		
L-Serin	250 mg	
	H ₂ O (Aqua Purania)	ad 1.000 mL

Anmerkungen zur Herstellung von TMB-Medium: Lösung A wurde als konzentrierte 12,5 × Stammlösung hergestellt und autoklaviert. Diese Stammlösung wurde im Anschluss 1:10

Material & Methoden

verdünnt ($1,25 \times$ konzentriert; entspricht 80 % vom Endvolumen) und Lösung B ($10.000 \times$ Stammlösung) hinzugegeben. Hierbei musste ebenfalls der Konzentrationsfaktor von $1,25 \times$ berücksichtigt werden. Nach dem Autoklavieren wurde konzentrierte, sterilfiltrierte L-Serin-Lösung (100 g L^{-1}) hinzugefügt. Das fehlende Restvolumen wurde mit sterilem Aqua Purania bzw. konzentrierter Zucker-/Bakterienlösung auf das gewünschte Endvolumen aufgefüllt. Die durch den Autoklavierprozess ausfallende Salze lösen sich in $1 \times$ Medium erneut.

2.3 Puffer & Lösungen

Tabelle 5: Zusammensetzung verwendeter Puffer und Lösungen.

Name	Bestandteile	Mengenangabe
CTAB-Puffer	CTAB	2 %
	PVP	1 %
	NaCl	1,4 M
	Tris-HCl (pH 8,0)	100 mM
	EDTA (pH 8,0)	20 mM
HPO ₄ -Citrat-Puffer	Na ₂ HPO ₄	60 mM
	NaH ₂ PO ₄	40 mM
	Citrat	200 mM
Phosphatpuffer (PBS) (pH 7,4)	NaCl	137,0 mM
	KCl	2,7 mM
	Na ₂ HPO ₄	10,0 mM
	KH ₂ HPO ₄	1,8 mM
TBE-Puffer (pH 8,3)	Tris-HCl	45 mM
	Borsäure	45 mM
	EDTA	1 mM
TB-Puffer (pH 6,7)	PIPES	10 mM
	CaCl ₂	15 mM
	KCl	250 mM
	Einstellen des pH-Werts auf 6,7 (mit HCl bzw. KOH); dann Zugabe von MgCl ₂	
	MgCl ₂	55 mM
	Auffüllen auf gewünschtes Volumen (Reinstwasser) und Sterilfiltrieren. Lagerung bei 4 °C.	

Material & Methoden

Name	Bestandteile	Mengenangabe
TEA-Puffer (pH 8,0)	Tris-HCl	40 mM
	EDTA	1 mM
	Essigsäure	20 mM
	H ₂ O (Reinstwasser)	ad 1.000 mL
Z-Puffer (pH 6,5)	MOPS (pH 6,5)	0,1 mM
	NaCl	50,0 mM
	CaCl ₂	10,0 mM

2.4 Kits, kommerzielle Lösungen & Reagenzien

Tabelle 6: Auflistung verwendeter kommerzieller Kits.

Name	Hersteller
DNA Clean & Concentrator	Zymo Research Corp. (Irvine CA, USA)
Monarch [®] Plasmid Miniprep Kit	New England Biolabs, Inc. (Ipswich MA, USA)
NEBuilder HiFi DNA Assembly Master Mix	New England Biolabs, Inc. (Ipswich MA, USA)

Tabelle 7: Auflistung verwendeter kommerzieller Lösungen und Reagenzien.

Name	Hersteller
Bradford Reagent (5×)	SERVA Electrophoresis GmbH (Heidelberg, Deutschland)
GeneRuler™ 1 kb DNA Ladder PLUS, ready-to-use (0.1 µg/µL)	Thermo Fisher Scientific, Inc. (Waltham MA, USA)
GeneRuler™ 1 kb DNA Ladder, ready-to-use (0.1 µg/µL)	Thermo Fisher Scientific, Inc. (Waltham MA, USA)
Lambda PFG Ladder	New England Biolabs, Inc. (Ipswich MA, USA)

2.5 Technische Gerätschaften

Tabelle 8: Auflistung verwendeter technischer Gerätschaften.

Name	Hersteller
Alphaimager MINI	ProteinSimple (San Jose CA, USA)
Analysewaage Precisa® 125A SCS	Precisa Gravimetrics AG (Dietlikon, Schweiz)
Axio Scope.A1 (Mikroskop) mit AxioCam ICc 1 (Mikroskopkamera)	Carl Zeiss Microimaging GmbH (Göttingen, Deutschland)
Bandelin Sonorex RK 31 (Ultraschallbad)	BANDELIN electronic GmbH & Co. KG (Berlin, Deutschland)
Biometra TAdvance, Thermocycler	Analytic Jena AG (Jena, Deutschland)
Bosch GSP 30A21 (Gefrierschrank)	Robert Bosch GmbH (Stuttgart, Deutschland)
CFX96 Touch™ Real-Time PCR Detection System	Bio-Rad Laboratories, Inc. (Hercules CA, USA)
CHEF-DR® III Variable Angle System 962BR, Cooling Module 816BR und Variable Speed Pump 260BR (PFGE-System)	Bio-Rad Laboratories, Inc. (Hercules CA, USA)
Gene Pulser IITM Electoporation System	Bio-Rad Laboratories, Inc. (Hercules CA, USA)
Haake DC1 (Umwälzthermostat)	Thermo Haake GmbH (Karlsruhe, Deutschland)
Hamilton™ Microliter® #710, 100 µL (Mikroliterspritze)	Hamilton Bonaduz AG (Bonaduz, Schweiz)
HERAsafe HSP 18 (Sicherheitswerkbank)	Heraeus Holding GmbH (Hanau, Deutschland)
HTL Discovery Comfort (Pipette) (2 µL, 20 µL, 200 µL, 1 mL)	PZ HTL SA (Warschau, Polen)
Integra Voyager II Mehrkanalpipette (125 µL, 300 µL)	INTEGRA Biosciences AG (Zizers, Schweiz)
Klein-Zentrifuge Mikro 22R	Hettich AG (Bäch, Schweiz)
Messbecher, Polypropylen (2 L, 3 L, 5 L)	VITLAB GmbH (Grossostheim, Deutschland)
Messzylinder, SILBERBRAND, Borosilikatglas (50 mL, 100 mL, 250 mL, 500 mL, 1.000 mL)	BRAND GmbH & Co. KG (Wertheim, Deutschland)
Mittel-Zentrifuge Universal 320R	Hettich AG (Bäch, Schweiz)
Multipette® plus (Mehrfachdispenser)	Eppendorf AG (Hamburg, Deutschland)
MyCycler™ Thermal Cycler	Bio-Rad Laboratories, Inc. (Hercules CA, USA)

Material & Methoden

Name	Hersteller
NanoDrop™ ND-1000	NanoDrop Technologies LLC (Wilmington DE, USA)
New Brunswick™ Innova® 42/42R (Inkubationsschüttler)	New Brunswick Scientific Co., Inc. (Edison NJ, USA)
pH-Meter 902	Berrytec GmbH (Grünwald, Deutschland)
Präzisionswaage EW 600-2M	KERN & SOHN GmbH (Balingen-Frommern, Deutschland)
Reagenzflasche, Borosilikatglas (50 mL, 100 mL, 250 mL, 500 mL, 1.000 mL)	Schott AG (Mainz, Deutschland)
SLM Aminco FA-078 (French pressure cell press)	American Instruments Company Inc. (Silver Spring MD, USA)
Sorvall™ LYNX™ Superspeed-Zentrifuge mit Fiberlite™ F9-6×1000 LEX Festwinkelrotor / Fiberlite™ F14-6×250y LEX Festwinkelrotor / A27-8×50 Festwinkelrotor	Thermo Fisher Scientific, Inc. (Waltham MA, USA)
Sorvall™ RC6+ Zentrifuge mit Fiberlite™ F12S-6×500 LEX Festwinkelrotor	Thermo Fisher Scientific, Inc. (Waltham MA, USA)
Systec V-150, vertikaler Standautoklav 150 L	Systec GmbH (Linden, Deutschland)
Thermomixer comfort	Eppendorf AG (Hamburg, Deutschland)
TKA® MicroPure (Reinstwasseranlage)	TKA Wasseraufbereitungssysteme GmbH (Niederelbert, Deutschland)
Trockenschrank UM	Memmert GmbH & Co. KG (Schwabach, Deutschland)
Ultraschallprozessor UP200S	Hielscher Ultrasonics GmbH (Teltow, Deutschland)
Ultrospec 2100 <i>pro</i> (UV/VIS-Spektralphotomet)	GE Healthcare Bio Sciences AB (Uppsala, Schweden)
VARIOKLAV® 25 T, Dampfsterilisator, 25 L	H+P Labortechnik GmbH (Oberschleißheim, Deutschland)
VARIOKLAV® Dampftopf	H+P Labortechnik GmbH (Oberschleißheim, Deutschland)
VARIOMAG® Mobil 60 (Magnetrührer)	H+P Labortechnik GmbH (Oberschleißheim, Deutschland)
Vortex-Genie 2 (Reagenzglasschüttler)	Scientific Industries, Inc. (Bohemia NY, USA)
WTC Binder BD 240 (Brutschrank)	BINDER GmbH (Tuttlingen, Deutschland)
WTW Inolab pH 720 (pH-Meter)	Xylem Analytics Germany Sales GmbH & Co. KG (Weilheim, Deutschland)

2.6 Verbrauchsmaterialien

Tabelle 9: Auflistung verwendeter Verbrauchsmaterialien.

Name	Hersteller
Deckgläser, VWR Collection	VWR International (Radnor PA, USA)
Einmal-Impfösen, steril, VWR Collection	VWR International (Radnor PA, USA)
Elektroporationsküvetten, 2 mm Spalt (VWR Collection)	VWR International (Radnor PA, USA)
Eppendorf Combitips advanced® Multipipettenspitze (2,5 mL)	Eppendorf AG (Hamburg, Deutschland)
Glasperlen (M 2,85-3,45 mm)	Carl Roth GmbH + Co. KG (Karlsruhe, Deutschland)
Halb-Mikro-Küvette, Polystyrol (10×4×45 mm)	Sarstedt AG & Co (Nümbrecht, Deutschland)
HSW SOFT-JECT® Einmalspritze (steril) (10 mL, 20 mL, 50 mL)	Henke-Sass, Wolf GmbH (Tuttlingen, Deutschland)
Low-Profile PCR Tubes 8-tube strip, weiß	Bio-Rad Laboratories, Inc. (Hercules CA, USA)
Mikroröhre mit Verschluss, Polypropylen (2 mL)	Sarstedt AG & Co (Nümbrecht, Deutschland)
Mikrotestplatte 96 Well, Flachboden (steril)	Sarstedt AG & Co (Nümbrecht, Deutschland)
Minisart® NML Spritzenvorsatzfilter, 0.2 µm Netzmittel-freies Celluloseacetat, steril	Sartorius AG (Göttingen, Deutschland)
NATURE Star® BIO Zahnstocher, 8 cm	Franz Mensch GmbH (Buchloe, Deutschland)
Objektträger, VWR Collection	VWR International (Radnor PA, USA)
Optical Flat 8-Cap Strips (für 0,2 mL Reaktionsgefäße/-platten)	Bio-Rad Laboratories, Inc. (Hercules CA, USA)
PCR SoftTubes, DNS-, DNase-, RNasefrei (0.2 ml)	Biozym Scientific GmbH (Hessisch Oldenburg, Deutschland)
Petrischale, Polyesterol, 92 mm (steril mit Nocken)	Greiner Bio-One International GmbH (Kremsmünster, Österreich)
Polypropylen Röhrchen (15 mL, 50 mL)	Sarstedt AG & Co (Nümbrecht, Deutschland)
Reaktionsgefäß, Polypropylen, 1,5 mL	Zefa-Laborservice GmbH (Harthausen, Deutschland)
Reaktionsgefäß, Polypropylen, 2,0 mL	Greiner Bio-One International GmbH (Kremsmünster, Österreich)
TipOne® Pipettenspitzen (2 µL, 200 µL, 1.250 µL)	STARLAB International GmbH (Hamburg, Deutschland)

2.7 Computersoftware

Tabelle 10: Verwendete Computersoftware und deren Versionsnummer. Es handelt sich bei allen Programmen um Versionen für Microsoft® Windows 10.

Programm	Versionsnummer
Bio-Rad CFX Manager 3.1	3.1.1517.0823
Clone Manager Professional 9	9.51
GraphPad PRISM	7.0
Microsoft® Excel® 2016	16.0.9001.2080
R	3.5.1
R Studio	1.1.463
Zusätzliche R Studio Bibliotheken:	
Biostrings ⁸⁷	2.48.0
broom	0.5.1
dplyr ⁸⁸	0.8.0.1
ggplot2	3.1.0
ggridges	0.5.1
gtools	3.8.1
Peptides ⁸⁹	2.4
purrr	0.2.5
readr	1.3.1
readxl	1.1.0
stringr	1.4.0
tibble	2.0.1
tidyr	0.8.2
topGO ⁹⁰	2.32.0

2.8 Mikroorganismen

Tabelle 11: Auflistung verwendeter Mikroorganismen.

Stamm	Genotyp und Anmerkungen	Referenz
<i>Bacillus subtilis</i> 168	freundlicherweise zur Verfügung gestellt durch die Arbeitsgruppe Dr. Armin Ehrenreich (hausintern)	DSMZ 402
<i>Escherichia coli</i> EPI300-T1R	F ⁻ mcrA Δ(<i>mrr-hsdRMS-mcrBC</i>) (StrR) φ80 <i>dlacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1 endA1 araD139 (ara, leu)</i> 7697 <i>galU galK</i> λ ⁻ <i>rpsL nupG trfA tonA dhfr</i>	Epicentre Technologies Corp. (Madison, WI, USA)
<i>Escherichia coli</i> K12 MG1655	F ⁻ , λ ⁻ , <i>rph-1</i> . <i>Fnr</i> ⁺	DSMZ 18039
<i>Escherichia coli</i> XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI</i> ^q ZΔ <i>M15</i> Tn10 (Tet ^r)]	Stratagene Corp. (La Jolla, CA, USA)
<i>Escherichia coli</i> XL10-Gold	Tet ^r Δ(<i>mcrA</i>)183 Δ(<i>mcrCB-hsdSMR-mrr</i>)173 <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte</i> [F' <i>proAB lacI</i> ^q ZΔ <i>M15</i> Tn10 (Tet ^r) Amy Cam ^r]	Agilent Technologies, Inc. (Santa Clara, CA, USA)
<i>Geobacillus stearothermophilus</i> NCA 26	Typstamm	DSMZ 22
<i>Oleispira antarctica</i> RB-8	Typstamm	DSMZ 14852
<i>Picrophilus torridus</i> KAW 2/3	Typstamm	DSMZ 9790
<i>Thermus thermophilus</i> HB27	Typstamm (Laborstamm Göttingen)	DSMZ 7039
<i>Thermus thermophilus</i> Δ42	ΔTTP0042 (hausintern)	Li ⁹¹
<i>Thermus thermophilus</i> ΔGH	ΔTTP0042 ΔTTP0220-TTP0222 Blm ^R	Im Rahmen dieser Arbeit

2.8.1 Kultivierung von Mikroorganismen

Für Festmedien wurden den entsprechenden Flüssigmedien vor dem Autoklavieren 1,5 % Agar zugegeben. Im Falle von TB-Medium wurde die zugesetzte Menge auf 2 % erhöht, um dem Austrocknen des Festmediums bei höheren Temperaturen entgegenzuwirken.

Wurden die Organismen in Flüssigmedium kultiviert, so wurden Erlenmeyerkolben mit hinreichendem Volumen verwendet. Verschlossen mit einem Baumwollstopfen erfolgte die Kultivierung bei 180 rpm und konstant bleibender Temperatur in einem Schüttelinkubator

Material & Methoden

(New Brunswick™ Innova® 42). Bei kleineren Volumina (5 mL) wurden Reagenzgläser mit Aluminiumkappen verwendet.

Aufgrund der höheren Temperaturen bei der Kultivierung von *T. thermophilus*, wurden 15 bzw. 50 mL Polypropylen Röhrchen mit Schraubverschluss verwendet, um mögliche Evaporation des Mediums zu verhindern. Hierbei wurden die Röhrchen etwa zu einem Drittel gefüllt, so dass den Zellen ausreichend Sauerstoff zur Verfügung stand.

Untenstehend sind die Kultivierungsbedingungen der verwendeten Mikroorganismen aufgeführt. Die Zusammensetzung der Medien kann Tabelle 4 entnommen werden.

Tabelle 12: Verwendeten Medien in Bezug auf die zu kultivierenden Mikroorganismen.

Name	Medium	Kultivierungstemperatur
<i>O. antarctica</i>	Medium 514 mit 40 g L ⁻¹ Pyruvat	15 °C
<i>B. subtilis</i>	Medium 1	30 °C
<i>E. coli</i>	LB	37 °C
<i>G. stearothermophilus</i>	Medium 220	55 °C
<i>P. torridus</i>	Modifiziertes BROCK-Medium (2.8.3)	60 °C
<i>T. thermophilus</i>	TB (Vollmedium) oder TMB (Minimalmedium)	70 °C 60 °C (sofern nicht anders angegeben)

Material & Methoden

2.8.2 Antibiotika

Untenstehender Tabelle können die Konzentrationen aller verwendeten Antibiotika entnommen werden.

Tabelle 13: Verwendete Antibiotika und deren eingesetzte Konzentration.

Name	Abkürzung	Konzentration	Lösungsmittel
Ampicillin	amp	100 $\mu\text{g mL}^{-1}$	H ₂ O
Bleomycin (A2)	blm	12,5 $\mu\text{g mL}^{-1}$	H ₂ O
Kanamycin	kan	60 $\mu\text{g mL}^{-1}$	H ₂ O
Tetracyclin	tet	12,5 $\mu\text{g mL}^{-1}$	EtOH
Chloramphenicol	cam	12,5 $\mu\text{g mL}^{-1}$	EtOH

Im Falle von Festmedien wurde das entsprechende Antibiotikum dem noch flüssigen, handwarmem Medium vor dem Gießen von Agarplatten zugegeben.

2.8.3 Kultivierung von *Picrophilus torridus*

Die Kultivierung von *P. torridus* erfolgte in modifiziertem BROCK-Medium⁹² bei einem pH-Wert von 0,7.

Tabelle 14: Einzelbestandteile des modifizierten BROCK-Mediums.

Lösung	Bestandteil	Mengenangabe
A (100 ×) sterilfiltriert	(NH ₄) ₂ SO ₄	130 g
	MgSO ₄ · 7H ₂ O	25 g
	FeCl ₃ · 6H ₂ O	2 g
	H ₂ O (VE)	ad 1.000 mL
B (200 ×) autoklaviert	KH ₂ PO ₄	56 g
	MnCl ₂	360 mg
	ZnSO ₄	44 mg
	CuCl ₂	10 mg
	Na ₂ MoO ₄	6 mg
	VO ₂ SO ₄	6 mg
	CoSO ₄	2 mg
	Na ₂ B ₄ O ₇	900 µg
	H ₂ SO ₄	50 mL
	H ₂ O (VE)	ad 1.000 mL
C (1000 ×) autoklaviert	CaCl ₂ · 2H ₂ O	14 g
	H ₂ O (VE)	ad 200 mL
D (200 ×) sterilfiltriert	Hefeextrakt	10 g
	H ₂ O (VE)	ad 100 mL
E (1 ×) autoklaviert	H ₂ SO ₄	11 mL
	H ₂ O (VE)	968 mL

Für einen Liter BROCK-Medium werden Lösung E folgende Mengen der Restlösungen zugegeben: Lsg. A 10 mL, Lsg. B 5 mL, Lsg. C 1 mL und Lsg. D 5 mL. Der pH-Wert ist hierbei durch das Mischungsverhältnis bereits auf 0,7 eingestellt.

2.9 Bestimmung der Zelldichte

Die Bestimmung der optischen Dichte einer Bakterienkultur stellt eine schnelle Methode zur Quantifizierung der (indirekten) Zellzahl dar. Somit kann z.B. Aussagen über die Wachstumsphase getroffen werden oder Proben anhand dieser normalisiert werden. In eine Küvette wurden 1 mL Bakterienkultur gegeben und die Messung erfolgte im Spektralphotometer bei einer Wellenlänge von 600 nm. Bei einer optischen Dichte über 1,0 wurden geeignete Verdünnungen für die Messung angelegt.

2.10 Plasmide

Tabelle 15: Auflistung verwendeter Plasmide.

Name	Anmerkungen	Referenz
pET21a(+)- <i>cglT</i>	pET21a(+) mit der Glycosidhydrolase (GH1) CglT aus <i>Thermoanaerobacter brockii</i> ⁹³⁻⁹⁵	Lehrstuhlintern (CLAUDIA HELD)
pET21a(+)- <i>cglT</i> KIR	pET21a(+) mit thermostabiler Mutante der Glycosidhydrolase CglT (s.o.). Zu CglT KIR: siehe angemeldetes Patent WO2017063787A1 ⁹⁶	Lehrstuhlintern (BJÖRN ANDREESSEN), angemeldetes Patent: WO2017063787A1 ⁹⁶
pMK18	Ein von pUC18 abgeleiteter Shuttle-Vektor zwischen <i>E. coli</i> und <i>T. thermophilus</i> . Km ^R -Kassette für Selektion von Transformanten auf Kanamycinresistenz	DE GRADO <i>et al.</i> ⁴⁶
pMK18- <i>bgl</i>	Shuttle-Vektor pMK18 mit der aus <i>T. th.</i> HB27 stammenden β -Glucosidase	Arbeitsgruppenintern (HAIJUAN LI)
pMK18- <i>cglT</i>	Shuttle-Vektor pMK18 mit der aus <i>Thermoanaerobacter brockii</i> stammenden Glycosidhydrolase (GH1)	Im Rahmen dieser Arbeit
pMK18- <i>cglT</i> KIR	Shuttle-Vektor pMK18 mit der aus <i>T. brockii</i> stammenden, thermostabilisierten GH1 (triple-Mutante, s.o.)	Im Rahmen dieser Arbeit

2.11 Isolierung genetischen Materials aus Mikroorganismen

2.11.1 Genomische DNS

Die Isolierung genomischer DNS erfolgte mit dem MasterPure™ DNA Purification Kit (Epicentre Technologies Corp.) nach Herstellerangaben. Abweichend hiervon wurde CTAB-Puffer (Kapitel 2.3) an Stelle des „MPC Protein Precipitation Reagent“ verwendet.

2.11.2 Plasmide

Die Kultivierung plasmidtragender Stämme erfolgte i.d.R. über Nacht in Vollmedium (5 mL) unter Selektionsdruck (z.B. Antibiotika). Die Plasmide wurden mit Hilfe des Monarch® Plasmid Miniprep Kits (New England Biolabs, Inc.) gewonnen. Die Anwendung erfolgte nach Herstellerangaben.

2.12 Polymerase-Kettenreaktion

Für die Vervielfältigung genetischen Materials *in vitro* wurde die *Polymerase-Kettenreaktion* (PCR) verwendet. Grundlegend handelt es sich um drei zentrale Schritte, welche für eine gewisse Anzahl zyklisch wiederholt werden: (1) Die Denaturierung aller DNS-Stränge, (2) die Hybridisierung von Matrizen-DNS mit einem Oligonukleotidpaar (*Forward Primer* und *Reverse Primer*) und (3) die DNS-Strangelongation durch eine DNS-Polymerase. Alle Reaktionen wurden mit Hilfe des MyCycler™ Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules CA, USA) durchgeführt. Die Zusammensetzung eines standardgemäß verwendeten Reaktionsansatzes kann der unten aufgeführten Tabelle 16 entnommen werden. Die verwendeten Oligonukleotide wurden von der Eurofins Genomics GmbH (Ebersberg, Deutschland) bezogen. Eine Auflistung aller verwendeter Oligonukleotide befindet sich im Anhang in Tabelle 31. Bei den hinzugegebenen Nukleotiden handelte es sich um eine Mischung aus dATP, dGTP, dCTP und dTTP zu je 10 mM.

Material & Methoden

Tabelle 16: Einzelbestandteile eines Q5 PCR-Ansatzes (25 µL Gesamtvolumen).

Bestandteil	Mengenangabe
Reaktionspuffer (Q5, 5×)	5,00 µL
dNTPs (10 mM)	0,50 µL
<i>Forward-Primer</i> (10 µM)	1,25 µL
<i>Reverse-Primer</i> (10 µM)	1,25 µL
Matrizen-DNS	1,00 µL*
DNS-Polymerase (Q5 <i>High-Fidelity</i>)	0,25 µL
<i>High GC Enhancer</i> (Q5, 5×, optional)	5,00 µL
Reinstwasser	ad 25,00 µL

* Genomische DNS wurde vor der Verwendung 1:100 und Plasmid-DNS 1:1.000 verdünnt.

Als Negativkontrolle wurde jedes Mal ein Ansatz ohne Matrize, dafür mit gleichem Volumen Reinstwasser, mitgeführt.

Wie bereits weiter oben erwähnt, handelt es sich bei einer PCR im Wesentlichen um drei sich wiederholende Schritte. Die detaillierte Abfolge kann Tabelle 17 entnommen werden.

Tabelle 17: Bedingungen einer standardmäßigen Q5 PCR.

Schritt	Temperatur	Zeit
Initiale Denaturierung	98 °C	30 s
30 Zyklen		
Denaturierung	98 °C	10 s
Primerhybridisierung	60 °C*	20 s
Elongation	72 °C	variabel**
Finale Elongation	72 °C	120 s
Warten	4 °C	∞

* Standardgemäß wurde die Primerhybridisierung bei 60 °C durchgeführt. Sofern unspezifische Produkte auftraten oder eine stark abweichende Hybridisierungstemperatur durch Clone Manager 9 empfohlen wurde, erfolge eine entsprechende Temperaturanpassung. Mit Ausnahmen wurde eine Gradienten-PCR (Temperaturgradient von 60–72 °C mit 8 PCR-Ansätzen zu je 10 µL) zur Bestimmung der optimalen Hybridisierungstemperatur durchgeführt.

** Die Elongationszeit wurde anhand der Länge des Zielamplifikats gewählt. Diese wurde anhand der Amplifikationsgeschwindigkeit der Q5 Polymerase festgesetzt (2 kBp min⁻¹).

Nach jeder PCR (sofern nicht rein analytisch) folgte abschließend ein Reinigungsschritt (DNA Clean & Concentrator Kit).

2.13 Reinigung von Plasmiden und PCR-Produkten

Meist liegt DNS durch vorhergehende Prozesse wie etwa präparativer Plasmidverdau oder PCR in speziellen Puffern oder mit anderen chemischen Verbindungen in Lösung vor. Diese Bestandteile sind oftmals mit nachfolgenden Prozessen nicht kompatibel, wodurch die DNS gereinigt werden muss. Hierbei wurde das DNA Clean & Concentrator Kit (Zymo Research Corp.) nach Herstellerangaben verwendet.

2.14 DNS-Quantifizierung

Zur schnellen Konzentrationsbestimmung von Nucleinsäuren wurde ein Mikrovolumen-Spektralphotometer (NanoDrop™ ND-1000) verwendet. Die Messung zur Konzentrationsbestimmung erfolgte bei 260 nm. Um die Reinheit der Probe zu bestimmen, wurde zudem bei Wellenlängen von 230 nm und 280 nm gemessen und das Verhältnis zu 260 nm bestimmt. Mit dem Verhältnis 260 nm:280 nm wird bestimmt, wie groß der Anteil von Proteinen ist und wird mit einem Wert von ~1,8 als rein betrachtet. Ob Verunreinigungen – wie etwa durch organische Substanzen – vorhanden sind, zeigt das Verhältnis 260 nm:230 nm. Hierbei sollte der Wert zwischen 1,8 und 2,2 liegen.

Alternativ zur spektroskopischen Methode kann die DNS-Konzentration über eine Agarose-Gelelektrophorese (Kapitel 2.17) bestimmt werden. Der eingesetzte Längenstandard wird mehrfach in unterschiedlichen Mengen aufgetragen und dient nach dem Färben zur Abschätzung der Nucleotidkonzentration anhand der Bandenintensität.

2.15 Restriktionsverdau

Der Restriktionsverdau wurde nach den Herstellerangaben der jeweiligen Restriktionsenzyme durchgeführt. Die Inkubation analytischen Verdau betrug ≥ 15 min, wohingegen präparative Verdau für 3–16 h bis zur Hitzeinaktivierung inkubiert wurden. Sollte mit zwei Restriktionsenzymen verdaut werden, und es war entsprechende Pufferkompatibilität gegeben, so wurde ein Doppelverdau durchgeführt. Wiesen die Puffer hingegen keine Kompatibilität auf, so wurde sequenziell mit einem dazwischenliegenden Reinigungsschritt (Kapitel 2.12) verdaut.

2.16 DNS-Fragmentassemblierung nach GIBSON

GIBSON beschrieb mit seiner Methode der DNS-Fragmentassemblierung eine effizientere Alternative zur herkömmlichen Ligation zuvor verdauter DNS-Moleküle⁹⁷. In nur einem einzigen Schritt können *in vitro* nahtlos eine Vielzahl von DNS-Fragmenten in geordneter Reihenfolge zusammengefügt werden.

Die Fragmente wurden mittels Q5-PCR (Kapitel 2.11), unter Verwendung spezieller Primer, amplifiziert: Der *Forward*-Primer trägt an seinem 5'-Ende zusätzlich 20 Bp, welche dem 3'-Ende der im Zielkonstrukt vorausgehenden Sequenz entsprechen. Respektiv wird das 5'-Ende des *Reverse*-Primers um 20 Bp des Anfangs der nachfolgenden Sequenz erweitert. Nach der Amplifikation weisen somit alle Fragmente 20 Bp spezifische Überlappung zu angrenzenden Abschnitten auf. Die Assemblierungsreaktion erfolgt im Anschluss durch einen Enzymmix bestehend aus einer 5'-Exonuklease, einer DNS-Polymerase und DNS-Ligase. Durch die 5'-Exonukleaseaktivität werden die kompatiblen Abschnitte zwischen den Fragmenten freigelegt, welche somit in der Lage sind, zu hybridisieren. Lücken am Fragmentübergang werden durch die vorhandene Polymerase aufgefüllt und DNS-Sequenzen schlussendlich durch die Ligase verbunden.

Die Assemblierung erfolgte mit dem NEBuilder[®] HiFi DNA Assembly-Enzymmix (New England Biolabs, Inc.) nach Herstellerangaben.

2.17 Agarose-Gelelektrophorese

Um Nukleinsäuren (z.B. Plasmide nach einem Restriktionsverdau oder PCR-Amplifikate) hinsichtlich ihrer Länge zu analysieren, wurde die Agarose-Gelelektrophorese verwendet. Es wurde ein Agarosegel von 0,8 % sowie gekühlter TAE-Puffer (Tabelle 3) verwendet. Die Fragmente (Länge bis 20 kBp) wurden mit Ladepuffer (6× DNA Loading Dye, Thermo Fisher Scientific, Inc.) gemischt und bei 180 V über einen Zeitraum von etwa 20 min aufgetrennt. Im Fall genomischer DNS wurde die Spannung auf 100 V reduziert und die Trennzeit auf 1 h verlängert. Als Größenstandards dienten GeneRuler™ 1 kb bzw. GeneRuler™ 1 kb Plus (Thermo Fisher Scientific, Inc.).

Material & Methoden

Nach elektrophoretischer Trennung der Fragmente wurde das Agarosegel für 10 min mit einer Ethidiumbromidlösung (1 mg L^{-1}) gefärbt, danach kurz in Wasser (VE) entfärbt und im Anschluss unter UV-Licht betrachtet bzw. fotografisch festgehalten (Alphaimager MINI).

2.18 DNS-Sequenzierung

Die Sequenzierung von DNS-Abschnitten erfolgte durch die GATC Biotech AG (LightRun).

2.19 Transformation von Genmaterial

2.19.1 Chemische Transformation von *E. coli*

Basierend auf INOUE *et al.*⁹⁸ (mit einigen Modifikationen) erfolgte die Herstellung chemisch kompetenter *E. coli* XL-1 Blue Zellen.

Ausgehend von einer Übernachtskultur auf Festmedium (LB-Medium mit Tetracyclin), wurden etwa 10–12 große Kolonien verwendet, um 250 mL SOB-Medium (Kapitel 2.2) in einem 1 L Erlenmeyerkolben anzupflanzen. Die Inkubation erfolgte bei $19 \text{ }^{\circ}\text{C}$ und 150 rpm, bis die Flüssigkultur eine optische Dichte ($\text{OD}_{600\text{nm}}$) von etwa 0,5 betrug (24–36 h). Nach einer Inkubation des Kulturkolbens für 10 min auf Eis, wurden die Zellen bei $870 \times g$ und $4 \text{ }^{\circ}\text{C}$ für 10 min zentrifugiert. Der Überstand wurde verworfen, die Zellen in 80 mL eiskaltem TB-Puffer (Kapitel 2.3) resuspendiert und erneut für 10 min auf Eis inkubiert. Nach erneuter Zentrifugation ($3.400 \times g$, $4 \text{ }^{\circ}\text{C}$, 10 min), wurde erneut der Überstand verworfen und die Zellen in 20 mL TB-Puffer mit 1,4 mL DMSO resuspendiert. Die Zellsuspension wurde im Anschluss zu je $100 \text{ }\mu\text{L}$ aliquotiert (1,5 mL Reaktionsgefäße) und mittels flüssigen Stickstoffs schockgefroren. Die Lagerung bis zur Verwendung der Zellen erfolgte bei $-80 \text{ }^{\circ}\text{C}$.

Für die Verwendung der chemisch kompetenten Zellen, wurden diese auf Eis aufgetaut. Im Anschluss wurde das zu transformierende Genmaterial hinzugegeben und für eine halbe Stunde auf Eis inkubiert. Es folgte eine Hitzeschockbehandlung der Zellen in einem Wasserbad bei $42 \text{ }^{\circ}\text{C}$ für 1 min und eine Inkubation auf Eis für 10 min. Anschließend wurden die Zellen mit 1 mL LB-Flüssigmedium gemischt und für eine Stunde bei $37 \text{ }^{\circ}\text{C}$ und 250 rpm inkubiert. Abschließend erfolgte das Ausplattieren der Zellen auf Agarplatten mit Selektionsmedium

Material & Methoden

(i.d.R. LB-Medium mit entsprechendem Antibiotikum) und der Inkubation über Nacht bei 37 °C.

2.19.2 Transformation von *E. coli* mittels Elektroporation

Von einer 1 mL *E. coli* XL-1 Blue Übernachtskultur in LB-Medium (in einem 15 mL Polypropylenröhrchen; Inkubation bei 37 °C und 180 rpm) wurden 30 µL Bakteriensuspension in 1,4 mL LB-Medium überführt (2 mL Reaktionsgefäß) und für 2–3 h bei 37 °C und 100 rpm inkubiert. Die Zellesuspension wurde danach bei $21.000 \times g$ und 4 °C für 5 min zentrifugiert. Der Überstand wurde verworfen und die Zellen in eiskaltem Reinstwasser resuspendiert. Dieser Waschschrift wurde ein weiteres Mal wiederholt. Der Überstand wurde bis auf etwa 30 µL entfernt, die Zellen im Restvolumen resuspendiert und das zu transformierende Genmaterial hinzugegeben (bei Bedarf wurde dieses zuvor dialysiert). Nachdem alles in eine vorgekühlte Elektroporationsküvette (2 mm) überführt wurde, erfolgte die Elektroporation bei 2.500 V, 200 Ω und 25 µF. Um eine hohe Transformationseffizienz zu gewährleisten, wurden während der Elektroporation Zeitkonstanten von ≤ 5 ms angestrebt. Die Zellen wurden in der Elektroporationsküvette mit 1 mL LB-Medium gemischt und anschließend in einem 2 mL Reaktionsgefäß für 1 h bei 37 °C und 180 rpm inkubiert. Abschließend wurden die Zellen auf Agarplatten mit Selektionsmedium (i.d.R. LB-Medium mit entsprechendem Antibiotikum) ausplattiert und über Nacht bei 37 °C inkubiert.

2.19.3 Transformation von *T. thermophilus*

Ausgehend von einer *T. thermophilus* HB27 Kultur auf TB-Festmedium, wurden 5 mL TB-Flüssigmedium (in einem 15 mL Polypropylenröhrchen) mit einer einzelnen Bakterienkolonie inokuliert und bei 70 °C und 180 rpm über Nacht inkubiert. Von dieser Vorkultur wurden 500 µL mit weiteren 500 µL TB-Medium in einem 15 mL Polypropylenröhrchen gemischt und für 1 h bei 70 °C und 180 rpm inkubiert. Nach Zugabe des zu transformierenden Genmaterials (≥ 1 µg DNS), erfolgte eine zweite Inkubation bei 70 °C und 180 rpm für 2 h. Die Zellsuspension wurde im Anschluss auf Agarplatten mit Selektionsmedium (i.d.R. TB-Medium

Material & Methoden

mit entsprechendem Antibiotikum) ausplattiert. Die darauffolgende Inkubation erfolgte bei 70 °C – im Falle von Kanamycin oder Bleomycin als Selektionsmarker bei 60 °C.

2.20 Kolonie-Cracking

Dieses modifizierte Verfahren nach BARNES⁹⁹ dient der schnellen Analyse von *E. coli*-Transformanten, um deren Plasmide hinsichtlich ihrer Größe zu charakterisieren. Auf diese Weise lässt sich zeitsparend klären, ob der vorliegende Vektor einen eingefügten DNS-Abschnitt trägt. Diese Methode lässt sich nur bei Vektoren hoher Kopiezahl anwenden.

Nach einer erfolgten Klonierung wurde von einer Agarplatte der Großteil einer einzelnen Kolonie abgetragen und in 50 µL EDTA (10 mM, pH 8,0) resuspendiert. Nachdem 50 µL frisch angesetzter Aufschlusspuffer (Tabelle 18) hinzugegeben wurde, erfolgte eine Inkubation für 5 min bei 70 °C.

Tabelle 18: Zusammensetzung des Aufschlusspuffers für Kolonie-Cracking. Der Puffer wurde immer frisch angesetzt.

Bestandteil	Mengenangabe
NaOH (2 M)	100,0 µL
SDS (10 %)	50,0 µL
Saccharose	0,2 g
Reinstwasser	ad 1,0 mL

Nachdem die Probe auf Raumtemperatur abgekühlt war, wurden 1,5 µL KCl (4 M) zugegeben und 5 min auf Eis inkubiert. Diese wurde darauffolgend für 5 min bei 21.000 × g und 4 °C zentrifugiert und vom Überstand 20 µL zur elektrophoretischen Trennung auf ein Agarosegel gegeben (Kapitel 2.17).

Als Referenz diente das ursprüngliche Plasmid (ohne zusätzlichen DNS-Abschnitt) in superspiralierter Form.

2.21 Herstellung eines GH-defizienten *T. thermophilus* Stammes

Die Deletion verbleibender Glycosidhydrolasen erfolgte durch deren Austausch mittels homologer Rekombination mit einem Resistenzmarker. Bei den deletierten Genen handelt es sich um alle im Genom (auf Megaplasmid pTT27) verbleibenden Glycosidhydrolasen mit bekannter Aktivität gegenüber β -glycosidischen Bindungen (TTP0220, TTP0221, TTP0222). Zur Selektion auf rekombinante Klone wurde auf eine Bleomycinkassette zurückgegriffen, welcher ein *slp*-Promotor (P_{slp}) vorausgeht (*blm*).

Ausgehend von genomischer DNS (*T. thermophilus* HB27 wt) als Matrize, wurden die beiden Elemente „TTP0220 upstream“ (1.525 Bp, Primerpaar P0220-f.g, P0220-r.g.3) und „TTP0222 downstream“ (1.434 Bp, Primerpaar P0222-f.g, P0222-r.g.3) Mittels Q5-PCR amplifiziert. Zur Amplifikation des Resistenzmarkers diente pUC18_ΔparAc::bleo (laborintern: HAIJUAN LI) als Matrize (649 Bp, Primerpaar *slp*.f.g, *bleo*.r.g). Für alle drei PCRs wurde eine Primerhybridisierungstemperatur von 60 °C und eine Elongationszeit von 45 s gewählt.

Material & Methoden

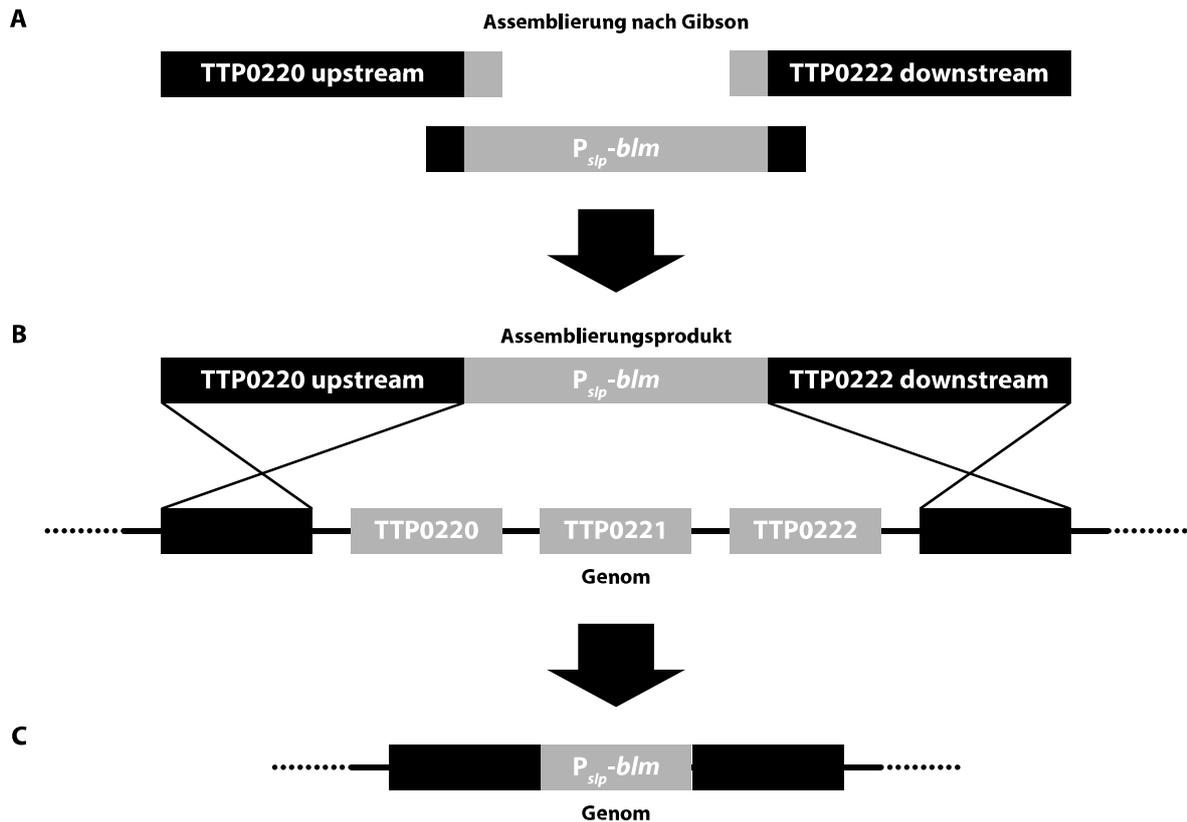


Abbildung 4: Schema zur Herstellung eines GH-defizienten *T. thermophilus*-Stammes. **A:** Assemblierung nach Gibson zweier der zu deletierenden Sequenz flankierende Bereiche (TTP0220 upstream und TTP0222 downstream) mit *slp*-Promotor und Bleomycin(A2)resistenzkassette (*blm*) als Selektionsmarker. **B:** Transformation von *T. th.* $\Delta 42$ mit dem linearen Assemblierungskonstrukt und darauffolgende homologe Rekombination mit dem genomischen Megaplasmid pTT27. **C:** Finale Darstellung des Stammes ΔGH . (Die relative Größe der Genabschnitte ist nicht proportional dargestellt.)

Die drei Amplifikate wurden mittels Gibson-Assemblierung verbunden (Kapitel 2.16) und im Anschluss in kompetente *T. thermophilus* HB28 $\Delta 42$ Zellen⁹¹ transformiert (Kapitel 2.19.3). Durch die homologen Bereiche findet in machen Transformanten eine homologe Rekombination statt, wodurch die zu deletierenden GHs (TTP220-TTP222) durch die Bleomycinresistenzkassette ausgetauscht wurden. Die Selektion dieser Transformanten erfolgte auf TB-Medium mit $12,5 \mu\text{g mL}^{-1}$ Bleomycinsulfat. Eine Bestätigung des Genotyps wurde von Sevgi Tansu Bağdatlı mittels Q5-PCR und Agarosegel-Elektrophorese bestätigt (Primerpaare 0042-1F/0042-1R, 220-wt-1/220-wt-2 und 220-del-1/220-del-2)¹⁰⁰.

2.22 Herstellung rekombinanter Stämme mit neu konstruierten Plasmiden

Aus einer 5 mL Übernachtskultur *E. coli* XL10 pMK18 wurde pMK18 isoliert (Kapitel 2.11.2). Die Linearisierung des Vektors erfolgte über sequenziellen Restriktionsverdau (Kapitel 2.14) mit HindIII und XbaI. Es wurde nach jedem Restriktionsverdau eine Reinigung des Restriktionsprodukts durchgeführt (Kapitel 2.12).

Für die Amplifikation der zu insertierenden Genabschnitte, wurden als Matrize die Plasmide pET21a(+)-CglT bzw. pET21a(+)-CglT KIR, sowie das Primerpaar cglT.f.g.V2/cglT.r.g (Anhang: Tabelle 31), verwendet. Die restlichen Parameter für die Q5-PCR-Reaktion können Kapitel 2.11 entnommen werden.

Der verdaute Vektor und die amplifizierte DNA wurden via Gibson-Assembly assembliert (Kapitel 2.16) und anschließend wurden mit den zirkulären DNA-Konstrukten chemisch kompetente *Escherichia coli* XL-1 Blue-Zellen transformiert (Kapitel 2.19.1). Nachdem die Zellen auf Vollmedium mit Kanamycin ausplattiert und für eine Nacht inkubiert wurden, erfolgte die Überprüfung einiger Kolonien auf korrekte Konstrukte via Kolonie-Cracking (Kapitel 2.20). Ein korrekter Klon wurde gewählt, eine 5 mL Übernachtskultur angelegt und die Plasmide extrahiert (Kapitel 2.11.2). Abschließend erfolgte die Transformation des Bakterienstamms *T. thermophilus* Δ GH (Kapitel 2.19.3) mit diesen Plasmiden.

2.23 GH-Aktivitätsbestimmung mit pNP-Substraten

Durch den Umsatz *p*-Nitrophenol-gekoppelter Substrate (*p*NP-Substrate), lässt sich die Aktivität eines Enzyms durch fotometrische Messung quantitativ bestimmen. Durch die Spaltung des Substrats wird *p*-Nitrophenol freigesetzt, dessen Absorption bei 405 nm gemessen werden kann. Im Rahmen dieser Arbeit – für die Aktivitätsbestimmung von β -Glucosidasen – wird auf *p*-Nitrophenyl- β -D-Glucopyranosid (*p*NPG) (Abbildung 5) zurückgegriffen.

Material & Methoden

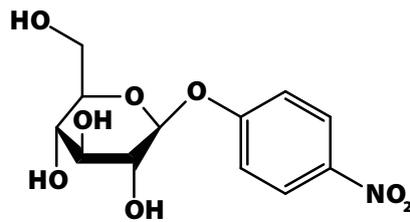


Abbildung 5: *p*-Nitrophenyl- β -D-Glucopyranosid.

Für einen Reaktionsansatz von 70 μ L wurden gereinigtes Enzym bzw. Zellextrakt mit 20 μ L Substratlösung (0,1 m *p*NPG in DMF : HPO₄-Citrat-Puffer, 1 : 14) gemischt und mit Z-Puffer auf ein Gesamtvolumen von 70 μ L aufgefüllt. Die Inkubation erfolgte in PCR-Reaktionsgefäßen für 30 min bei unterschiedlichen Temperaturen in einem Thermocycler (Biometra TAdvance, Analytic Jena AG). Im Anschluss wurden die Proben auf Eis gestellt und die Reaktion durch Zufügen von 150 μ L 1 M NaCO₃ gestoppt. Für die fotometrische Quantifizierung wurden 100 μ L dieses Ansatzes in einer Mikrotiterplatte überführt und die Absorption bei einer Wellenlänge von 405 nm bestimmt.

$$\text{spezifische Aktivität} = \frac{A_{405 \text{ test}} - A_{405 \text{ negativ}}}{t \times \text{Enzymkonzentration}} \quad (1)$$

2.24 Herstellung von Proteinextrakten für TPP Experimente

B. subtilis, *E. coli*, *G. stearothermophilus* und *T. thermophilus* wurden in entsprechenden Vollmedien, bei optimaler Wachstumstemperatur (Kapitel 2.2 und 2.8.1) und 180 rpm über Nacht kultiviert. Hierfür wurden jeweils 250 mL Flüssignährmedium in 1 L Erlenmeyerkolben mit einer Einzelkolonie (von Festmedium) beimpft. Für *P. torridus* erfolgte die Kultivierung – ausgehend von einer Flüssigvorkultur – in 100 mL Vollmedium (Kapitel 2.8.3) in einem 300 mL Erlenmeyerkolben für zwei Tage bei 60 °C und 50 rpm. Im Falle von *O. antarctica* wurden die Bakterien in 1 L Vollmedium (Kapitel 2.2) in einem 5 L Erlenmeyerkolben mit einer Einzelkolonie angeimpft und für 5 Tage bei 15 °C und 120 rpm kultiviert. Die Zellkulturen wurden für die Ernte in 500 mL Zentrifugenbecher (bei *P. torridus* 50 mL

Material & Methoden

Polypropylenröhrchen) überführt und für 20 min bei $3.100 \times g$ und $4 \text{ }^\circ\text{C}$ zentrifugiert. Der Überstand wurde verworfen und die Zellen in eiskaltem PBS resuspendiert. Für *P. torridus* wurde – um ein Lysieren der Zellen zu vermeiden – an Stelle von PBS, eisgekühlte Natriumacetatlösung (50 mM, pH 4,6) verwendet. Dieser Waschprozess wurde insgesamt zweimal durchgeführt. Final wurden Zellen in 50 mL gekühltem PBS (bzw. in Natriumacetatlösung) resuspendiert und auf sechs 15 mL Polypropylenröhrchen aufgeteilt. Nach einem letzten Zentrifugationsschritt bei $3.100 \times g$ für 20 min bei $4 \text{ }^\circ\text{C}$, wurde der Überstand verworfen und die Zellen zur Lagerung bei $-20 \text{ }^\circ\text{C}$ bis zur weiteren Verwendung verwahrt.

Für den Zellaufschluss wurden die Zellen in 3 mL eiskaltem PBS mit 0,5 % Nonidet P-40 resuspendiert und für 15–20 min bei einer Frequenz von 0,5 s und einer Amplitude von 30 % sonifiziert (Ultraschallprozessor UP200S). Zusätzlich wurden die *B. subtilis*-Proben dreimal mit einer French Press (SLM Aminco FA-078) bei etwa 170 MPa behandelt.

Die Zellextrakte wurden im Anschluss für 30 min bei $21.000 \times g$ ($4 \text{ }^\circ\text{C}$) zentrifugiert und der Überstand an Dr. Anna Jarzab (Lehrstuhl für Proteomik und Bioanalytik, Technische Universität München) zur weiteren Probenaufbereitung im Rahmen von thermal proteome profiling (TPP) und massenspektrometrischer Analyse übergeben. Eine Proteinkonzentrationsbestimmung des Überstands nach BRADFORD (Kapitel 2.25) versicherte, dass stets mindestens $2.000 \mu\text{g mL}^{-1}$ Protein vorlagen.

2.25 Proteinkonzentrationsbestimmung nach Bradford

Zur schnellen Bestimmung der Proteinkonzentration wurde die Analyse nach BRADFORD¹⁰¹ in unterschiedlichen Maßstäben verwendet. Es wurden 5× Bradfordreagenz (SERVA Electrophoresis GmbH) 1:4 mit Reinstwasser verdünnt und 990 μL dieses Reagenz mit 10 μL der zu analysierenden Probe gemischt. Nach einer Inkubation für 10 min bei Raumtemperatur im Dunklen, erfolgte die photometrische Absorptionsmessung bei 595 nm (Ultrospec 2100 pro). Als Standard diente Rinderserumalbumin (BSA) mit Konzentrationen von 0, 25, 125, 250,

Material & Methoden

500, 750, 1.000, 1.500 und 2.000 $\mu\text{g mL}^{-1}$ (BSA aus Micro BCA Protein Assay Kit, Thermo Fisher Scientific, Inc.).

Sollte die Proteinkonzentration nur abgeschätzt werden, wurde – an Stelle unmittelbar einen Proteinstandard mitzuführen – ein zuvor gemessener Proteinstandard zur Berechnung herangezogen.

Zur genaueren Bestimmung der Proteinkonzentration wurden Ansätze in Mikrotiterplatten durchgeführt. Es wurden 5 μL der Probe/des Proteinstandards vorgelegt und 250 μL Bradfordreagenz (1 \times) mit Hilfe einer Mehrfachdispenserpipette zugegeben. Nach 10 min Inkubation bei Raumtemperatur im Dunklen, erfolgte die Messung bei 595 nm in einem Multiplattenleser (FLUOstar Omega). Proben und Standard wurden hierbei immer als Triplikaten geführt.

Die Berechnung der Konzentration anhand des BSA-Standards erfolgte mittels nichtlinearer Regression einer quadratischen Funktion (Restriktion: $y = 0$) in GraphPad PRISM.

2.26 Dynamische Differenzfluorimetrie

Die Messung thermische Proteinverschiebung mittels dynamischer Differenzfluorimetrie (*engl.*: DSF – differential scanning fluorimetry) erfolgte nach NIESEN *et al.*⁷⁰ mit Modifikationen.

Eine *E. coli* und *T. thermophilus* Übernachtskultur in Vollmedium-Medium (5 mL; Inkubation bei 37 °C bzw. 70 °C und 180 rpm) wurde in ein 15 mL Polypropylenröhrchen überführt und für 20 min bei 3.100 \times g und 4 °C zentrifugiert (Medien: Tabelle 4 und Tabelle 12). Der Überstand wurde verworfen und die Zellen in vorgekühltem PBS resuspendiert. Dieser Waschschritt wurde ein weiteres Mal wiederholt. Die Zellen wurden ein weiteres Mal in gekühltem PBS resuspendiert und auf zwei Reaktionsgefäße aufgeteilt:

- (1) Die Zellen wurden über einen Zeitraum von 15 min mittels Ultraschall (Ultraschallprozessor UP200S) aufgeschlossen (Frequenz 0,5, Amplitude 30 %). Der Proteinrohextrakt wurde bei 21.000 \times g und 4 °C für 15 min zentrifugiert und der

Material & Methoden

Überstand in ein neues Reaktionsgefäß überführt. Die Proteinkonzentration wurde mit Hilfe des Bradford-Tests (Kapitel 2.25) bestimmt (1 mL Küvetten) und im Folgenden mit vorgekühltem PBS auf 2 mg mL^{-1} verdünnt.

- (2) Es wurde die optische Dichte der Zellsuspension bei 600 nm bestimmt und diese im Anschluss durch Verdünnen mit vorgekühltem PBS auf einen Wert von 2,0 eingestellt.

In einem PCR-Reagenzgefäß (8er Streifen, weiß) wurden $10 \mu\text{L}$ Proteinextrakt bzw. Zellsuspension mit $2 \mu\text{L}$ SYPRO™ Orange (50× Stammlösung in DMSO) gemischt und mit PBS auf ein Gesamtvolumen von $20 \mu\text{L}$ aufgefüllt. Die Reaktionsgefäße wurden mit einem Deckelstreifen (RT-PCR geeignet) verschlossen und bis zur Messung auf Eis gelagert.

Die Proben wurden während der Messung (CFX96 Real-Time System) die ersten 3 min konstant auf einer Temperatur von $30,0 \text{ }^\circ\text{C}$ gehalten und im Anschluss isotherm über einen Gradienten von $30,0\text{--}48,7 \text{ }^\circ\text{C}$ und $51,3\text{--}70,0 \text{ }^\circ\text{C}$ mit je 8 Proben für 10 min inkubiert. Die Fluoreszenzmessung erfolgte auf Kanal 2 (HEX; Anregung bei $525 \text{ nm} \pm 10 \text{ nm}$, Emissionsmessung bei $570 \text{ nm} \pm 10 \text{ nm}$).

2.27 Thermische Inaktivierung von Bakterien anhand der Überlebensrate

Die Zellen einer Übernachtskultur *E. coli* bzw. *T. thermophilus* (5 mL ; Inkubation bei $37 \text{ }^\circ\text{C}$ bzw. $70 \text{ }^\circ\text{C}$ und 180 rpm) wurden zweimal in eiskaltem PBS gewaschen und auf einen OD_{600} -Wert von 0,5 eingestellt (Medien: Tabelle 4 und Tabelle 12). Es folgte die isotherme Inkubation mehrerer Aliquote zu je $100 \mu\text{L}$ bei unterschiedlichen Temperaturen (*E. coli*: $37\text{--}67 \text{ }^\circ\text{C}$, *T. thermophilus*: $70\text{--}99 \text{ }^\circ\text{C}$) für 10 min in einem Thermocycler (Biometra TAdvance, Analytic Jena AG). Die Zellen wurden im Anschluss sofort auf Eis gestellt und eine Verdünnungsreihe (10^{-1} bis 10^{-8}) einer jeden Temperaturstufe angelegt. Von jeder Verdünnungsstufe wurden $5 \mu\text{L}$ punktförmig in Triplikaten auf Agarplatten mit Vollmedium ausgebracht und die Kulturen über Nacht inkubiert. Als Kontrolle dienten Probe thermisch unbehandelter Zellen. Am Folgetag wurden die Kolonien ausgezählt um die KBE mL^{-1} zu berechnen.

2.28 Logistische Regression

Bei Modellen der logistischen Regression handelt es sich um ein vielseitiges und robustes Klassifikationsverfahren. Es kommt meist dann zum Einsatz, wenn zwischen Zuständen unterschieden werden soll. Beispielhaft im Falle TPP, ob ein Protein, abhängig von der Temperatur, nativ oder aggregiert vorliegt. Die Anwendung logistischer Regression findet heutzutage in vielen verschiedenen Bereichen wie etwa der Finanzwelt, im Marketing oder in der Medizin Anwendung. Entsprechend werden die Funktionen für logistische Regression entsprechend der Anwendung (leicht) angepasst.

Formel 1 beschreibt hierbei die in dieser Arbeit verwendete Art logistischer Regression im Falle von DSF (Kapitel 2.26) und thermischem Abtöten von Mikroorganismen (Kapitel 2.27), wobei für x die Temperatur in °C eingesetzt wird.

$$f(x) = \frac{O - U}{1 + e^{(S \times (\log x - \log E))}} + U \quad (2)$$

Beschreibung der Variablen

O y-Wert des oberen Plateaus

U y-Wert des unteren Plateaus

S bestimmt die Steigung am steilsten Punkt der Kurve

E x-Wert am steilsten Punkt der Kurve, auch Schwellenwert genannt

Eine exemplarische Darstellung einer solchen Kurve befindet sich im Anhang (Abbildung 29).

3 ERGEBNISSE

3.1 Komplementierung eines GH-defizienten *Thermus thermophilus*-Stammes

Teile der nachfolgend beschriebenen Ergebnisse wurden auf der Tagung „*Microbial Interactions 2018*“ in Rom, Italien (19. –20. Juli 2018), präsentiert.

3.1.1 Aminosäuresupplementierung von basalem Minimalmedium

Das System zur komplementationsbasierten Selektion in *T. thermophilus* setzt ein Minimalmedium voraus, welches Wachstum in Anwesenheit von geeigneten Zuckersubstraten als Kohlenstoff- und Energiequelle gewährleistet. Als erster Schritt galt es, ein solches Minimalmedium (TMB – *Thermus minimal broth*, Tabelle 4) für diesen Zweck zu entwickeln.

Inspiziert durch eine abgewandelte Form des SH-Mediums nach KOYAMA⁴⁹, wurden verschiedene Aminosäuren basalem TMB in unterschiedlichen Kombinationen zugesetzt (Abbildung 6).

Ergebnisse

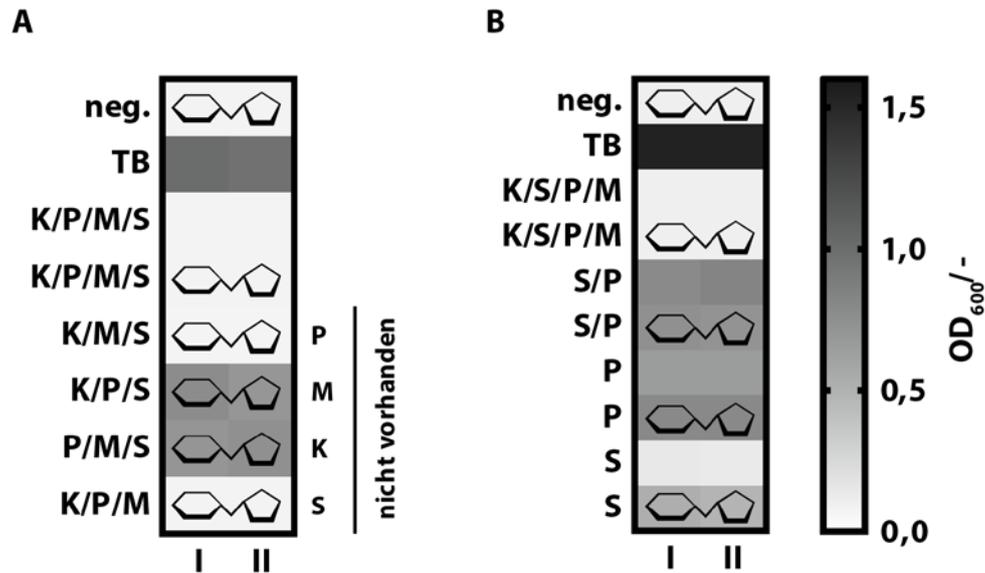


Abbildung 6: Wachstum von *T. thermophilus* in basalem TMB mit unterschiedlicher Aminosäuresupplementierung. **A:** Supplementierung mit L-Lysin (K), L-Prolin (P), L-Methionin (M) und L-Serin (S) in unterschiedlichen Kombinationen (1,0 g L⁻¹). Die Kultivierung erfolgte über Nacht. Rechts stehen die Aminosäuren, welche in der Kombination jeweils weggelassen wurden. **B:** Einzel- und Kombinationsansätze mit L-Serin und L-Prolin. Die Kultivierung erfolgte für 20 h. Die Negativkontrolle (neg.) bestand in beiden Fällen aus TMB mit 0,5 % Saccharose, die Positivkontrolle aus Vollmedium (TB). Ansätze mit 0,5 % Saccharose sind durch stilisierte Disaccharide markiert. Die Messungen erfolgten in Duplikaten (römische Ziffern). Die Legende ist für beide Graphen gültig.

Wurden L-Lysin, L-Prolin, L-Methionin und L-Serin in gleichen Mengen (1 g L⁻¹) basalem TMB beigegeben, so war mit und ohne Zugabe von 0,5 % Saccharose kein Wachstum zu beobachten (Abbildung 6A und Abbildung 6B). Im Falle einer Kombination aller Aminosäurezusätze, jedoch ohne L-Methionin bzw. L-Lysin, konnte dagegen Wachstum beobachtet werden (Abbildung 6A). Da in Abwesenheit von L-Prolin bzw. L-Serin kein Wachstum zu beobachten war und diese Aminosäuren daher für das Wachstum bedeutsam schienen, wurden gezielt Kombinationen und Einzelansätze untersucht (Abbildung 6B). Unspezifisches Wachstum (unabhängig ob mit oder ohne Saccharose) war bei Ansätzen, welche L-Prolin enthielten, zu beobachten. Die Ansätze rein mit L-Serin zeigten mit Saccharose Wachstum und ohne Zuckerzusatz keines.

Ergebnisse

3.1.2 Wachstum von *T. thermophilus* und GH-defizienten Stämmen in TMB

Nachdem für *T. thermophilus* ein entsprechendes Minimalmedium (TMB) entwickelt wurde (Kapitel 3.1.1), wurde damit das Wachstum unterschiedlicher *T. thermophilus*-Stämme untersucht. Damit sollte festgestellt werden, ob und in welchem Ausmaß die Komplementation von Δ GH (GH-defizienter *T. thermophilus*-Stamm mit unterschiedlichen GHs) Einfluss auf die Wachstumskinetik hat.

Es wurden der Wildtyp, Δ GH, Δ GH pMK18 (Leervektor), Δ GH pMK18-*bgl* (Bgl aus *T. th.*), Δ GH pMK18-*cglT* (GH1 aus *T. Brockii*) und Δ GH pMK18-*cglT KIR* (eine thermisch stabilere Variante von CglT) in TMB kultiviert⁹³. Als Kohlenstoffquelle diente 0,5 % D-Cellobiose; die Positivkontrolle wurde mit 0,5 % D-Maltose durchgeführt (im Folgenden als Cellobiose und Maltose bezeichnet). Zur Aufrechterhaltung des Selektionsdrucks wurden allen Kulturen mit plasmidhaltigen Stämmen Kanamycin zugesetzt.

Ergebnisse

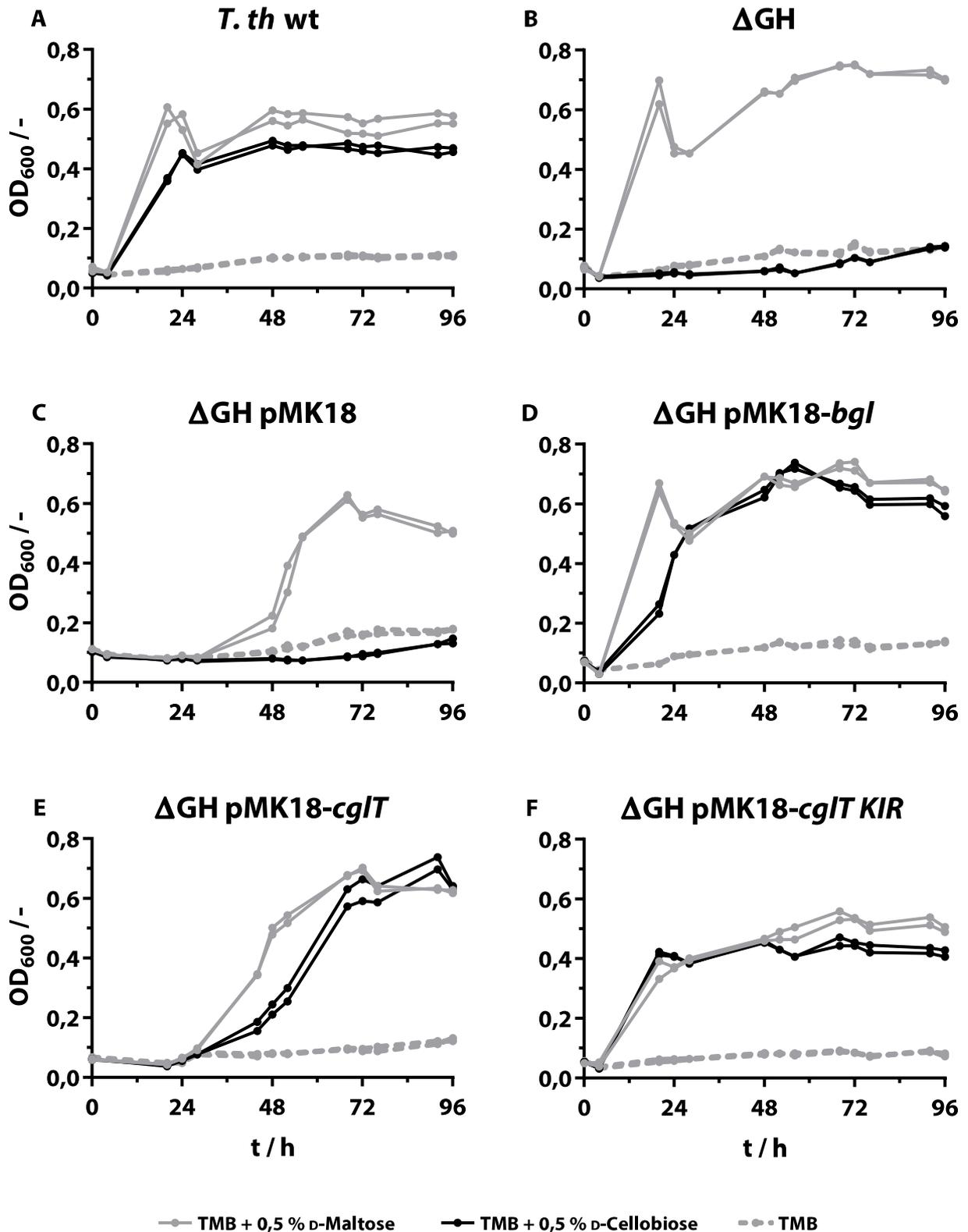


Abbildung 7: Wachstum unterschiedlicher *T. thermophilus*-Stämme in TMB bei 60 °C mit D-Cellobiose (Schwarz). Als Positivkontrolle diente D-Maltose (Grau), in der Negativkontrolle waren keine zusätzlichen Disaccharide enthalten (Grau, gestrichelt). Alle Stämme mit Plasmiden wurden mit Kanamycin kultiviert. Die Kultivierung erfolgte in Duplikaten.

Ergebnisse

Der Wildtyp zeigte für Cellobiose und Maltose bereits nach 24 h in einem Bereich von etwa 0,4 bis 0,6 seine maximale optische Dichte (OD_{600}). Im Falle der GH-defizienten Mutante (ΔGH) blieb für Cellobiose das Wachstum aus; die Probe mit Maltose zeigt einen vergleichbaren Verlauf wie der Wildtyp. Die Variante ΔGH pMK18 (Leervektor) wuchs in Anwesenheit von Cellobiose nicht an, die Positivkontrolle zeigte nach 48 h beginnendes Wachstum und erreichte maximale Kulturdichte nach etwa 56 h.

Der mit Bgl komplementierte Stamm (ΔGH pMK18-*bgl*) zeigte vergleichbares Verhalten wie der Wildtyp. Abweichend zeigte sich, dass mit Cellobiose das Wachstum etwas verzögert begann und ebenfalls später das Maximum erreicht wurde. Die mit CglT komplementierte Variante (ΔGH pMK18-*cglT*) zeigte für Cellobiose wie auch für die Positivkontrolle einen späteren Beginn es Wachstums. Hierbei ähnelte die Positivkontrolle (Stamm ΔGH pMK18-*cglT* auf Maltose) dem Wachstum von ΔGH pMK18 auf Maltose. Mit Cellobiose wurde das Maximum der optischen Dichte etwa 24 h später erreicht. Mit der thermostabilisierten Mutante CglT KIR war die Kultur mit Cellobiose von der Positivkontrolle im Verlauf kaum zu unterscheiden. Bereits nach 24 h war die maximale optische Dichte nahezu erreicht. Für diesen Ansatz fielen die maximalen OD-Werte im Vergleich etwas niedriger aus. Zu erwähnen ist in diesem Fall, dass es sich beim verwendeten Minimalmedium um eine andere Charge handelte, was sich auf das Wachstumsergebnis auswirken kann.

3.1.3 Aktivitäten der Enzyme Bgl, CglT und CglT KIR

Mit Hilfe von *pNP*-Glycosiden (hier *pNPG*) als chromogenen Substraten kann die Aktivität von Glycosidhydrolasen kolorimetrisch bestimmt werden. Die Inkubation mehrerer Reaktionsansätze über einen Temperaturgradienten ermöglicht die Erfassung der Abhängigkeit der Enzymaktivität von der Temperatur. Eine Bestimmung der temperaturabhängigen Aktivität von CglT, CglT KIR und Bgl soll hierbei den möglichen Zusammenhang zwischen Enzymaktivität/-stabilität und den zugehörigen Wachstumskinetiken entsprechender Bakterienstämme beleuchten. Die detaillierte

Ergebnisse

Durchführung sowie weitere Hintergrundinformationen zu *pNP*-Assays können Kapitel 2.23 entnommen werden.

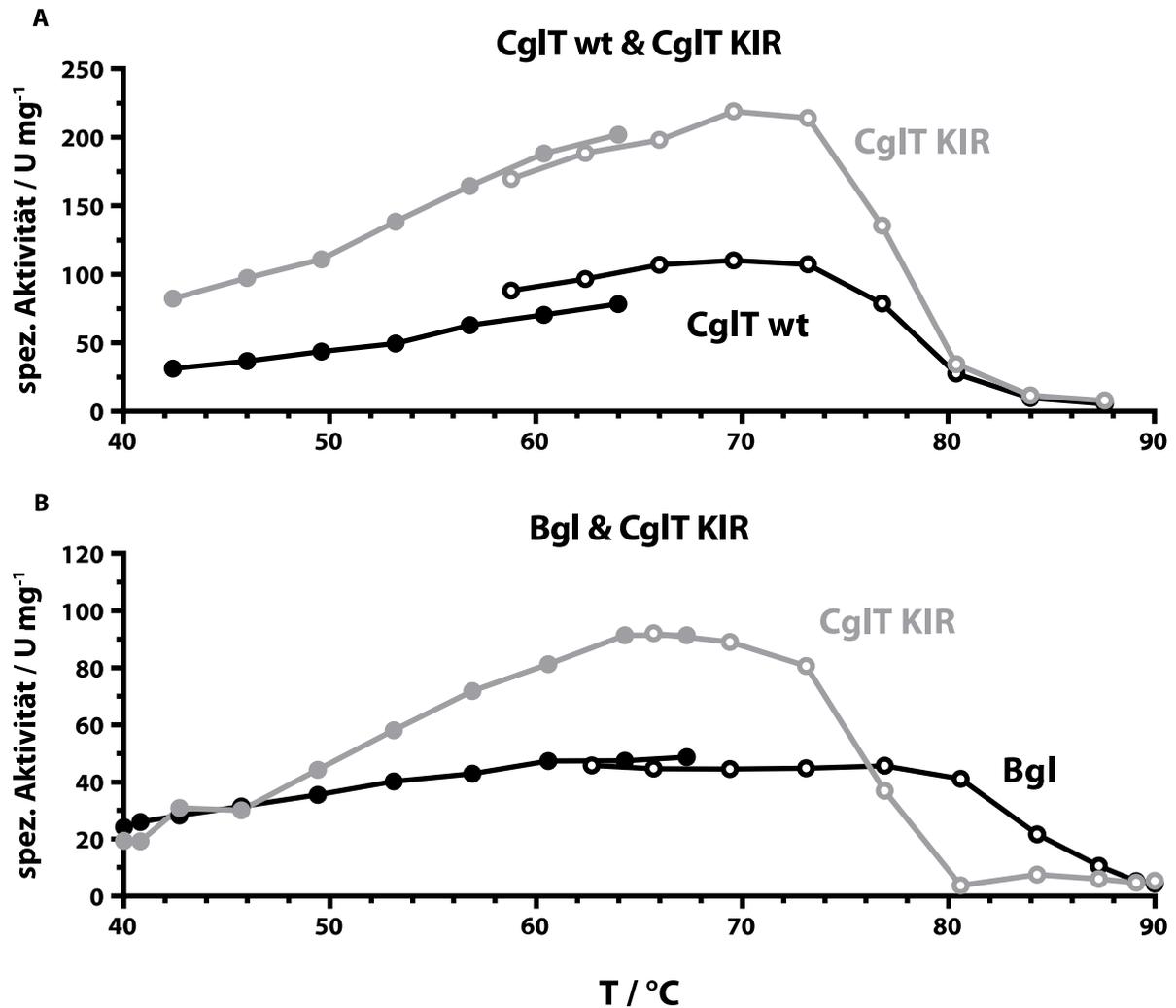


Abbildung 8: Temperaturabhängige Aktivität von Bgl, CglT und CglT KIR anhand eines *pNP*-Assays mit *pNPG*. **A:** Vergleich der Aktivität des CglT Wildtyps mit der Dreifachmutante CglT KIR. Die Messdaten wurden durch die Arbeitsgruppe Schwarz/Zverlov aufgenommen und freundlicherweise zur Verfügung gestellt. **B:** Unabhängige, eigenständig durchgeführte Messung von Bgl (freundlicherweise durch Yajing Liu zur Verfügung gestellt) und CglT KIR (ging aus gleicher Präparation wie A hervor). Alle Reaktionen erfolgten bei pH 6,5 in HPO₄-Citrat-Puffer (Zusammensetzung: Tabelle 5) über einen Zeitraum von 30 min. Die Messung erfolgte im Duplikat; dargestellt sind Durchschnittswerte.

Für CglT und CglT KIR (GH1 aus *T. brockii* und eine Dreifachmutante dieser) zeigen sich ähnliche Aktivitätsprofile mit einem Aktivitätsmaximum bei etwa 70 °C (Abbildung 8A). Beide Aktivitäten steigen mit zunehmender Temperatur annähernd linear bis zum Maximum an um dann bei etwa 80 °C wieder den Großteil der Aktivität zu verlieren. CglT KIR zeigt eine doppelt

Ergebnisse

so hohe, maximale spezifische Aktivität wie der Wildtyp (219 U mg^{-1} im Vergleich zu 110 U mg^{-1}).

Der Wiederholungsversuch (Abbildung 8B) zeigt für CglT KIR ein ähnliches Profil, jedoch mit niedrigerem Temperaturoptimum (etwa $65 \text{ }^\circ\text{C}$) und niedrigerer maximaler spezifischer Aktivität (92 U min^{-1}). Bgl zeigt ein deutlich flacheres Aktivitätsprofil. Von etwa $40\text{--}60 \text{ }^\circ\text{C}$ steigt die spezifische Aktivität linear an und geht dann in ein Plateau bis rund $80 \text{ }^\circ\text{C}$ über. Bis zum Minimum von $90 \text{ }^\circ\text{C}$ nimmt die Aktivität daraufhin linear ab. Das Aktivitätsmaximum für das Plateau beträgt im Durchschnitt 46 U mg^{-1} .

3.2 Proteomweite Untersuchung der Proteinstabilität bei *T. thermophilus* und anderen Mikroorganismen

Teile der nachfolgenden Ergebnisse wurden zur Veröffentlichung eingereicht: JARZAB, A. *et al.*, Meltome Atlas – thermal proteome stability across the tree of life, *Nature Methods* (2019)

Für die Untersuchung von Meltomen (Gesamtheit aller Schmelztemperaturen eines Organismus), wurden insgesamt sechs Mikroorganismen betrachtet (geordnet nach steigender, optimaler Wachstumstemperatur): *Oleispira antarctica*, *Bacillus subtilis*, *Escherichia coli*, *Geobacillus stearothermophilus*, *Picrophilus torridus* und *Thermus thermophilus*. Mit Ausnahme von *P. torridus* (Archaeon) handelt es sich bei allen Organismen um Bakterien. Die optimalen Wachstumstemperaturen und gewählten Anzuchtbedingungen können Kapitel 2.8.1 bzw. Tabelle 12 entnommen werden.

Alle Mikroorganismen wurden in der späten exponentiellen Phase geerntet nach mehrfachem Waschen in eiskaltem Phosphatpuffer (bei *P. torridus* Natriumacetatlösung) mit einem Ultraschallprozessor aufgeschlossen. Da es sich bei *B. subtilis* um ein Gram-positives Bakterium handelt, war vor dem Aufschluss mit Ultraschall eine dreimalige Behandlung mit der French Press notwendig. Obwohl *G. stearothermophilus* ebenfalls Gram-positiv ist, war hier dieser zusätzliche Schritt nicht notwendig. Bei den Proteinextrakten wurde darauf geachtet, dass eine Mindestproteinmenge von $2.000 \text{ } \mu\text{g mL}^{-1}$ erreicht wurde. Eine detaillierte Vorgehensweise zur Präparation von Proteinextrakten kann Kapitel 2.24 entnommen werden.

Ergebnisse

Die hergestellten Proteinextrakte wurden von Dr. Anna Jarzab (Lehrstuhl für Proteomik und Bioanalytik, Technische Universität München) für eine TPP-Analyse aufbereitet (thermische Behandlung, Ultrazentrifugation, TMT-Markierung, MS-Analyse und Bestimmung der Schmelztemperaturen). Im Rahmen dieser Arbeit erfolgte die eigenständige Analyse der auf diese Weise gewonnenen Proteom-/Meltom-Daten.

3.2.1 Übersicht aller gemessenen Proteome

Abbildung 9 zeigt die Abdeckung der durch massenspektrometrische Methoden erhaltenen Proteomdaten. Hierfür wurde der Anteil aller gemessenen Proteine am Gesamtproteom (im Genom kodiert) sowie der Proteine, für welche tatsächlich eine Schmelztemperatur berechnet werden konnte, bestimmt (Tabelle 19).

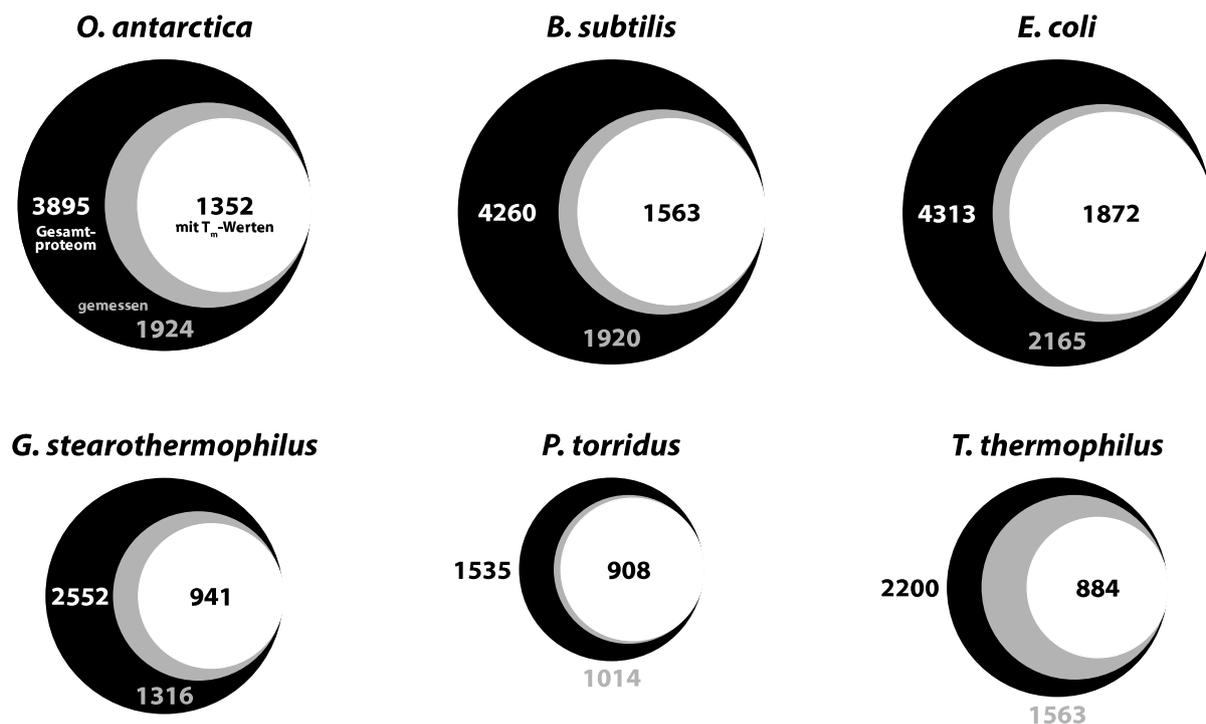


Abbildung 9: Venn-Diagramm aller massenspektrometrisch analysierten Proteome. Für die jeweiligen Organismen ist die Proteomgröße in Dunkelgrau (Zahl links), die Anzahl detektierter Proteine in Hellgrau (Zahl unten) und die Anzahl der Proteine, für welche ein T_m-Wert bestimmt werden konnte, in Weiß dargestellt (Zahl rechts). Angaben zur Proteomgröße wurden der UniProt-Datenbank entnommen³⁶.

Ergebnisse

Die psychrophilen und mesophilen Bakterien (*O. antarctica*, *B. subtilis* und *E. coli*) haben die größten Gesamtproteome (Abbildung 9). Von den 3.895 bis 4.313 Proteinen, wurde jeweils rund die Hälfte detektiert; für 70–86 % konnte ein T_m -Wert bestimmt werden (Tabelle 19). Die Thermophilen (*G. stearothermophilus*, *P. torridus* und *T. thermophilus*) hingegen haben kleinere Proteome als die eben beschriebenen (1.535–2.552 Proteine). Zwar wurden insgesamt weniger Proteine im Gesamten detektiert (884–941 Proteine), jedoch ist Ihr Anteil am Gesamtproteom mit 52–71 % höher. Für *T. thermophilus* liegt der Anteil, für welchen Schmelzkurven bestimmt werden konnten, im Vergleich niedriger (57 %) und bei *P. torridus* vergleichsweise hoch (90 %). *G. stearothermophilus* weist mit 72 % Abdeckung einen mit den Organismen im unteren Temperaturspektrum vergleichbaren Wert auf.

Detaillierte Anteilswerte können Tabelle 19 entnommen werden.

Tabelle 19: Unterschiedliche prozentuale Proteomabdeckungen aller Organismen. Abkürzungen: *detekt. Proteine*: alle detektierten Proteine (mit und ohne T_m -Werte); *Schmelzk.*: Proteine, für welche eine Schmelzkurve aufgenommen und somit ein T_m -Wert bestimmt werden konnte.

	detekt. Proteine in Gesamtproteom	mit Schmelzk. in Gesamtproteom	mit Schmelzk. in detekt. Proteinen
<i>O. antarctica</i>	49 %	35 %	70 %
<i>B. subtilis</i>	45 %	37 %	81 %
<i>E. coli</i>	50 %	43 %	86 %
<i>G. stearothermophilus</i>	52 %	37 %	72 %
<i>P. torridus</i>	66 %	59 %	90 %
<i>T. thermophilus</i>	71 %	40 %	57 %

Ergebnisse

3.2.2 Verteilung der proteomweiten Thermostabilität

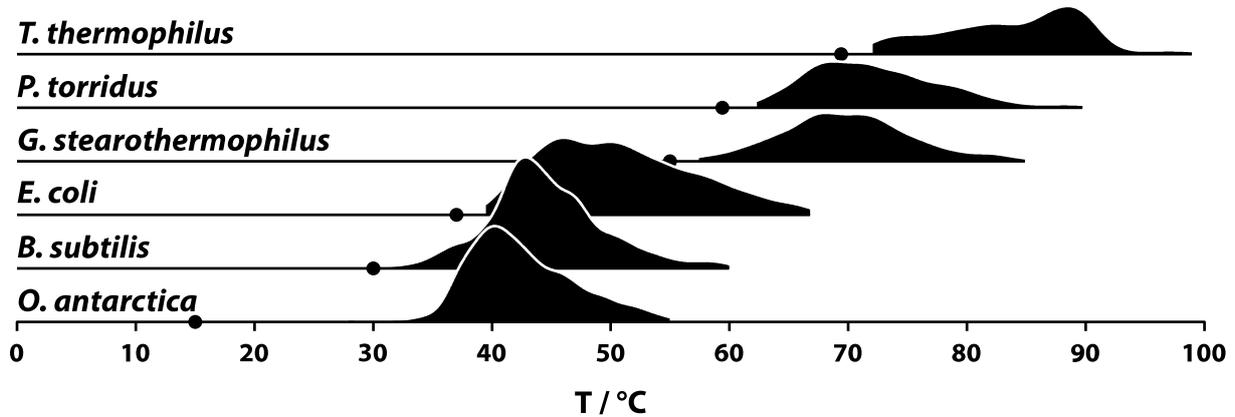


Abbildung 10: Verteilung der Proteinaggregationstemperaturen (T_m) nach Organismus. Die Wachstumstemperaturen der Organismen sind als schwarze Punkte kenntlich gemacht. In Anlehnung an JARZAB *et al.*¹⁰²

Für alle Mikroorganismen beginnt die Verteilung der T_m -Werte etwa 2°C oberhalb der Wachstumstemperatur¹⁰². Die Ausnahme bildet *O. antarctica* mit einem Beginn der Verteilung mehr als 12°C über dem T_{opt} des Wachstums.

Überwiegend gestaltete sich die Werteverteilung symmetrisch mit unterschiedlich ausgeprägter Verschiebung hin zu niedrigeren Temperaturen. Von diesem Verhalten weicht als einziges *T. thermophilus* ab: Die Verteilung ist zu höheren Temperaturen hin verschoben und zeigt sich im unteren Bereich annähernd gleichförmig mit leichtem Anstieg. Einige wenige Proteine liegen im obersten Temperaturbereich.

Ergebnisse

3.2.3 Proteomweite Verteilung von pI-Werten

Der pI-Wert eines Proteins beschreibt den pH-Wert, bei welchem dessen Nettoladung neutral und die Löslichkeit somit am geringsten ist. Da die direkte Bestimmung eines jeden pI-Werts innerhalb eines Proteoms einen zu großen Aufwand darstellt, kann dieser anhand der Primärsequenz nach BJELQVIST berechnet werden¹⁰³. Hierdurch soll in Erfahrung gebracht werden, ob sich die Organismen in der Verteilung der isoelektrischen Punkte ihrer Proteine unterscheiden und ob ein Zusammenhang mit der Temperatur vorliegt.

Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „*entry identifier*“ den gemessenen Proteinen zugeordnet^{36,88}. Die Berechnung des isoelektrischen Punkts eines jeden Proteins erfolgte anhand dessen Sequenz mit der Funktion *pI* (R-Bibliothek: *Peptides*)⁸⁹. Der Parameter *pKscale* wurde hierbei auf „Bjellqvist“ gesetzt.

Ergebnisse

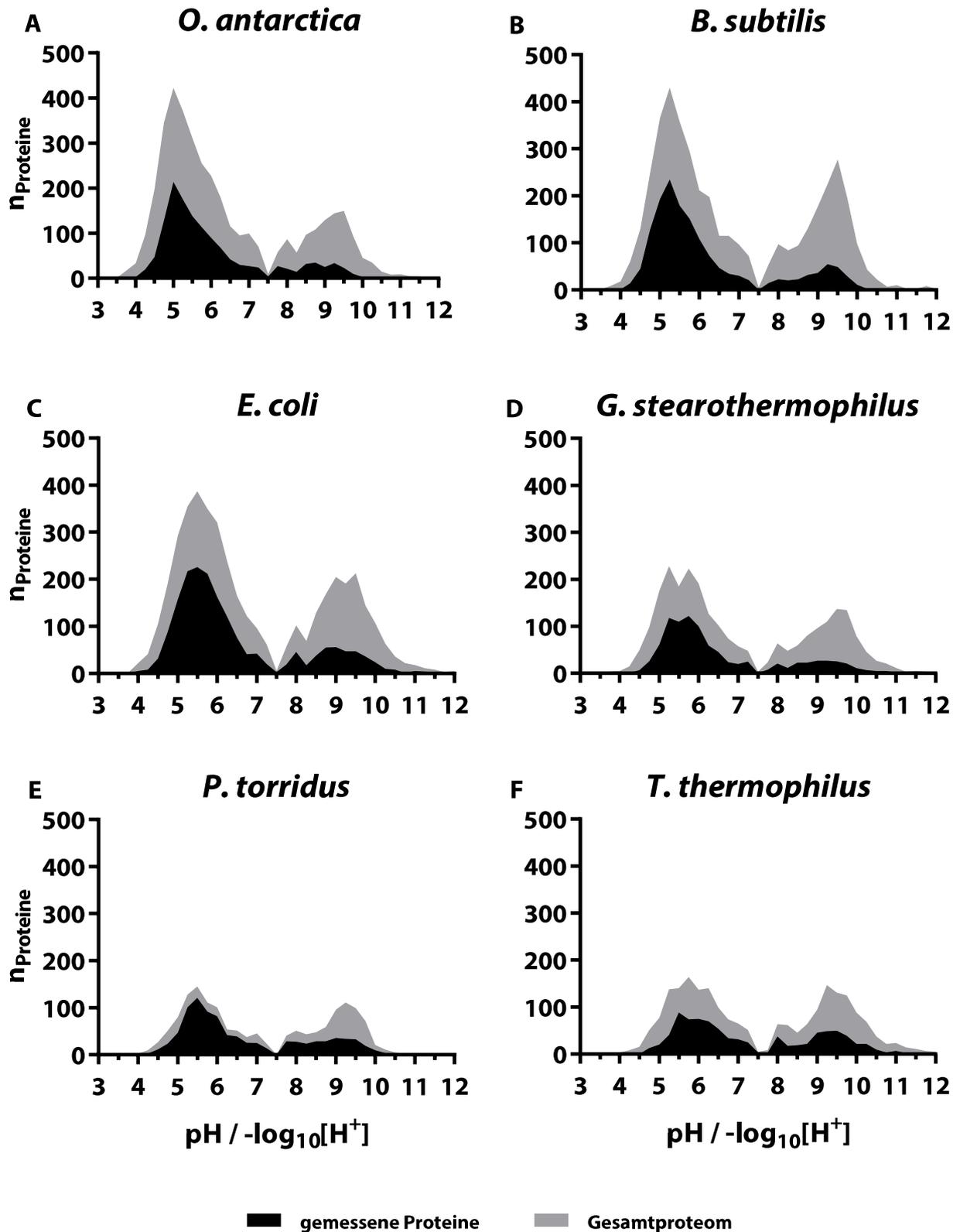


Abbildung 11: Verteilung der pI-Werte für Proteine mit Schmelzkurve und des Gesamtproteoms. Die Berechnung der isoelektrischen Punkte erfolgte anhand der Proteinsequenz nach BJELLQVIST¹⁰³. Bei der Berechnung entsprach die Bin-Weite einem pH-Bereich von 0,25.

Ergebnisse

Bei allen Organismen wurde eine bimodale Verteilung (Abbildung 11) für alle im Genom kodierten Proteine (Grau) und für diejenigen, für welche ein Schmelzpunkt bestimmt werden konnte (Schwarz), beobachtet. Bei einem pH-Wert von 7,5 sind keine Proteine mit entsprechendem pI-Wert vorhanden. Über alle Organismen hinweg zeigte sich, dass mehr Proteine einen pI-Wert unter- als oberhalb dieses charakteristischen Werts von 7,5 haben. Hinzu kommt außerdem, dass in diesem unteren Bereich der relative Anteil aller kodierten Proteine größer war als bei einem pI-Wert $> 7,5$.

3.2.4 Vergleich thermostabiler und instabiler Proteine

Die Klassifikation in thermisch stabile bzw. labile Proteine erfolgte in dieser Arbeit anhand der oberen bzw. unteren 10 % der gemessenen Proteine, entsprechend ihrer Rangfolge nach T_m . Diese beiden Gruppen können im Anschluss gegenübergestellt werden, um mit statistischen Mitteln Unterschiede diverser Charakteristika (z.B. Proteinelänge, Hydrophobizität, usw.) festzustellen. Hierfür wurde ein WILCOXON-MANN-WHITNEY-Test mit einem Konfidenzintervall von 95 % gewählt. Hierfür wurde die R-eigene Funktion *wilcox.test* mit folgenden Parametern verwendet: `alt = 'two.sided'`, `paired = F`, `exact = F`, `correct = T`, `conf.int = T`, `conf.level = 0.95`.

Proteinelänge

Es wurde die Länge eines jeden Proteins anhand seiner Sequenz berechnet, um den Einfluss dieser auf den T_m -Wert durch einen statistischen Vergleich thermisch labiler und stabiler Proteine zu untersuchen.

Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „*entry identifier*“ den gemessenen Proteinen zugeordnet^{36,88}. Die Berechnung der Proteinelänge eines jeden Proteins erfolgte anhand dessen Sequenz mit der R-eigenen Funktion *nchar* – einer Funktion zur Bestimmung der Länge von String-Variablen. Die Unterschiede zwischen thermostabilen und -labilen Proteinen (obere bzw. untere 10 %) wurden mit Hilfe eines WILCOXON-MANN-WHITNEY-Tests bestimmt (Details zur Funktion und gewählten Parametern: siehe Seite 57).

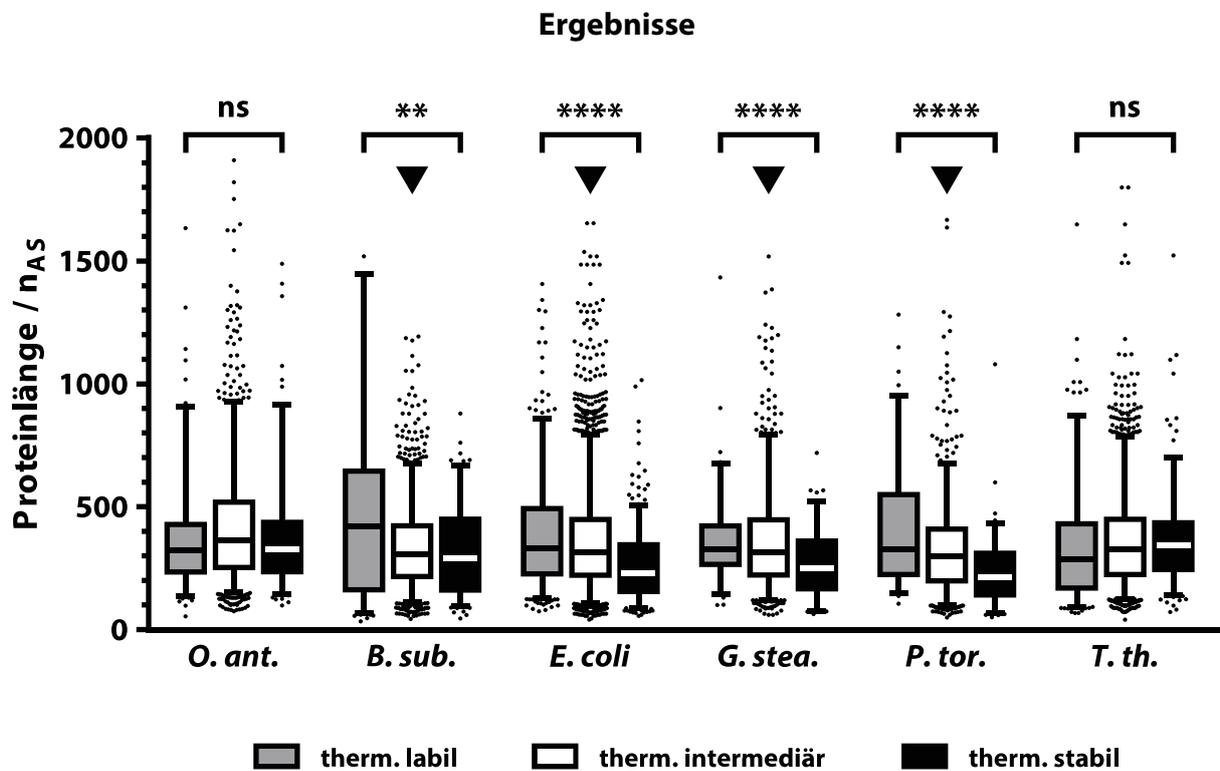


Abbildung 12: Proteinlänge in AS gruppiert nach thermischer Proteinstabilität. Gezeigt ist das 5–95 Perzentil. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$. Die Pfeilrichtung unterhalb der Signifikanzgrade gibt an, ob der Median der thermisch stabilen Fraktion über (▲) oder unter (▼) dem der labilen Fraktion liegt. Die y-Achse wurde beschnitten, um eine detailliertere Darstellung zu ermöglichen.

Es stellte sich heraus, dass die Proteine der thermisch stabilen Fraktion, im Vergleich zu thermisch labilen Proteinen, signifikant kürzer sind (Abbildung 12). Ausnahmen bildeten *O. antarctica* und *T. thermophilus*, bei denen hierbei kein signifikanter Unterschied auftrat.

Tabelle 20: Vergleich thermostabiler und -labiler Proteine anhand ihrer Länge. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$. Länge stabiler Proteine: Liegt der Median der thermisch stabilen Fraktion über (▲) oder unter (▼) dem der labilen Fraktion.

Organismus	p-Wert	Signifikanz	Länge stabiler Proteine
<i>O. antarctica</i>	8,67E-01	ns	▼
<i>B. subtilis</i>	3,53E-03	**	▼
<i>E. coli</i>	2,49E-13	****	▼
<i>G. stearothermophilus</i>	2,28E-05	****	▼
<i>P. torridus</i>	4,15E-08	****	▼
<i>T. thermophilus</i>	1,69E-01	ns	▲

Ergebnisse

Einen alternativen Ansatz bietet die Betrachtung der Korrelation nach PEARSON von T_m -Werten und Proteinlänge. Hierbei wird der lineare Zusammenhang zweier Parameter widergegeben. Der ermittelte Korrelationskoeffizient kann Werte in einem Bereich von +1 bis -1 annehmen. Bei diesen Extremwerten spricht man von vollständiger positiver bzw. negativer Korrelation. Werte nahe 0 lassen auf eine nur geringe bzw. keine Abhängigkeit schließen. Die Berechnung von Korrelationswerten erfolgte mit der R-eigenen Funktion *cor.test*. Hierbei wurde der Parameter *method* auf „pearson“ gesetzt.

Tabelle 21: PEARSON-Korrelationskoeffizienten von T_m und Proteinlänge. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Organismus	PEARSON-Korr.koeff.	p-Wert	Signifikanz
<i>O. antarctica</i>	-0,089	1,05E-03	**
<i>B. subtilis</i>	-0,144	1,13E-08	****
<i>E. coli</i>	-0,243	4,95E-26	****
<i>G. stearothermophilus</i>	-0,172	1,20E-07	****
<i>P. torridus</i>	-0,229	2,71E-12	****
<i>T. thermophilus</i>	-0,024	4,68E-01	ns

Fast alle Organismen zeigten eine hochsignifikante, negative Korrelation zwischen T_m -Werten und der Proteinlänge (*O. antarctica* signifikant, wenn nicht gar so stark ausgeprägt wie der Rest). Die einzige Ausnahme bildet *T. thermophilus*, welcher einen nicht signifikanten Korrelationskoeffizienten nahe 0 aufwies.

Hydrophobizität

Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „*entry identifier*“ den gemessenen Proteinen zugeordnet^{36,88}. Die Berechnung der Hydrophobizität eines jeden Proteins erfolgte anhand dessen Sequenz mit der Funktion *hydrophobicity* (R-Bibliothek: *Peptides*)⁸⁹. Der Parameter *scale* wurde hierbei auf „KyteDoolittle“ gesetzt. Die Unterschiede zwischen thermostabilen und -labilen Proteinen (obere bzw. untere 10 %) wurden mit Hilfe eines WILCOXON-MANN-WHITNEY-Tests bestimmt (Details zur Funktion und gewählten Parametern: siehe Seite 57).

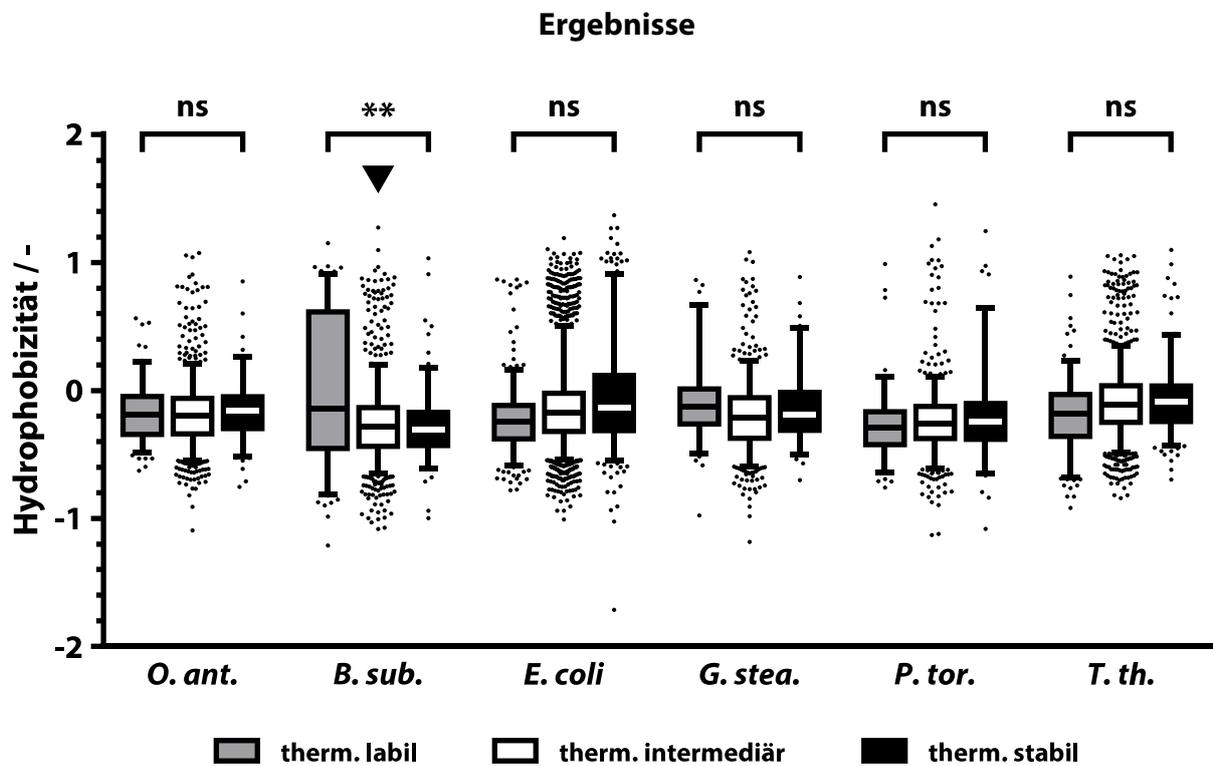


Abbildung 13: Proteinhydrophobizität gruppiert nach thermischer Proteinestabilität. Die Berechnung der Hydrophobizität erfolgte anhand der Proteinsequenz nach KYTE & DOOLITTLE¹⁰⁴. Gezeigt ist das 5–95 Perzentil. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$. Die Pfeilrichtung unterhalb der Signifikanzgrade gibt an, ob der Median der thermisch stabilen Fraktion über (\blacktriangle) oder unter (\blacktriangledown) dem der labilen Fraktion liegt.

Der Vergleich thermisch labiler und stabiler Proteine unter dem Gesichtspunkt der Hydrophobizität zeigte (mit Ausnahme von *B. subtilis*) keinen signifikanten Unterschied der Verteilungen. Die Mittelwerte aller Organismen und thermischen Fraktionen lagen im leicht hydrophilen Bereich und zeigten eine größere Streuung hin zu hydrophoberen Werten. Die bereits angesprochene Ausnahme bildet *B. subtilis*. Hier zeigte die labile Fraktion eine tendenziell verstärkt hydrophobe Verteilung auf, wodurch sich diese signifikant von den thermisch stabilen Proteinen abgrenzt.

Ergebnisse

Tabelle 22: Vergleich thermostabiler und -labiler Proteine anhand ihres Hydrophobizität. Die Berechnung der Hydrophobizität erfolgte anhand der Proteinsequenz nach KYTE & DOOLITTLE¹⁰⁴. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant); $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$. Hydroph. stabiler Proteine: Liegt der Median der thermisch stabilen Fraktion über (▲) oder unter (▼) dem der labilen Fraktion.

Organismus	p-Wert	Signifikanz	Hydroph. stabiler Proteine
<i>O. antarctica</i>	3,53E-01	ns	▲
<i>B. subtilis</i>	8,51E-04	***	▼
<i>E. coli</i>	6,97E-01	ns	▼
<i>G. stearothermophilus</i>	2,39E-01	ns	▼
<i>P. torridus</i>	8,79E-02	ns	▲
<i>T. thermophilus</i>	7,33E-01	ns	▲

Neben dem Vergleich thermisch labiler und stabiler Proteine, kann – wie schon bei der Proteinlänge zuvor – die Korrelation nach PEARSON zwischen der Hydrophobizität und T_m -Werten betrachtet werden (Tabelle 23). Die Berechnung von Korrelationswerten erfolgte mit der R-eigenen Funktion *cor.test*. Hierbei wurde der Parameter *method* auf „pearson“ gesetzt.

Tabelle 23: PEARSON-Korrelationskoeffizienten von T_m und Hydrophobizität. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant); $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Organismus	PEARSON-Korr.koeff.	p-Wert	Signifikanz
<i>O. antarctica</i>	0.059	2.91E-02	*
<i>B. subtilis</i>	-0.188	7.48E-14	****
<i>E. coli</i>	-0.031	1.85E-01	ns
<i>G. stearothermophilus</i>	-0.062	5.83E-02	ns
<i>P. torridus</i>	0.020	5.52E-01	ns
<i>T. thermophilus</i>	0.013	7.00E-01	ns

Überwiegend waren nur gering bzw. nicht signifikante Korrelationen zu beobachten, welche zudem nahe 0 lagen. Wie schon im vorherigen Vergleich der unterschiedlichen thermischen Fraktionen bildet *B. subtilis* eine Ausnahme und zeigte eine signifikante, leicht negative Korrelation zwischen Hydrophobizität und Schmelztemperatur.

Ergebnisse

Aminosäurefrequenz

Unterschiede in der Aminosäurezusammensetzung thermostabiler und -labiler Proteine wurden anhand des Vergleichs des relativen Anteils einer jeden Aminosäure in der Sequenz bestimmt. Anhand des p-Werts lässt sich somit ermitteln, ob eine signifikante Anreicherung einer Aminosäurespezies im Falle höherer bzw. niedrigerer Aggregationstemperaturen vorliegt (Abbildung 14). Die Analyse lehnt sich hierbei an die Arbeit von LEUENBERGER *et al.* an¹⁰⁵.

Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „*entry identifier*“ den gemessenen Proteinen zugeordnet^{36,88}. Die Berechnung der Aminosäurefrequenz eines jeden Proteins erfolgte anhand dessen Sequenzen mit der Funktion *letterFrequency* (R-Bibliothek: *Biostrings*)⁸⁷. Der Parameter *letters* wurde hierbei auf „AA_STANDARD“ gesetzt – eine Liste aus der Bibliothek *Biostrings*, welche alle proteinogenen Aminosäuren abbildet. Relative Frequenzen wurden durch die Division der zuvor erhaltenen Werte mit der Länge des entsprechenden Proteins (R-eigene Funktion *nchar*) erhalten. Die Unterschiede zwischen thermostabilen und -labilen Proteinen (obere bzw. untere 10 %) wurden mit Hilfe eines WILCOXON-MANN-WHITNEY-Tests bestimmt (Details zur Funktion und gewählten Parametern: siehe Seite 57).

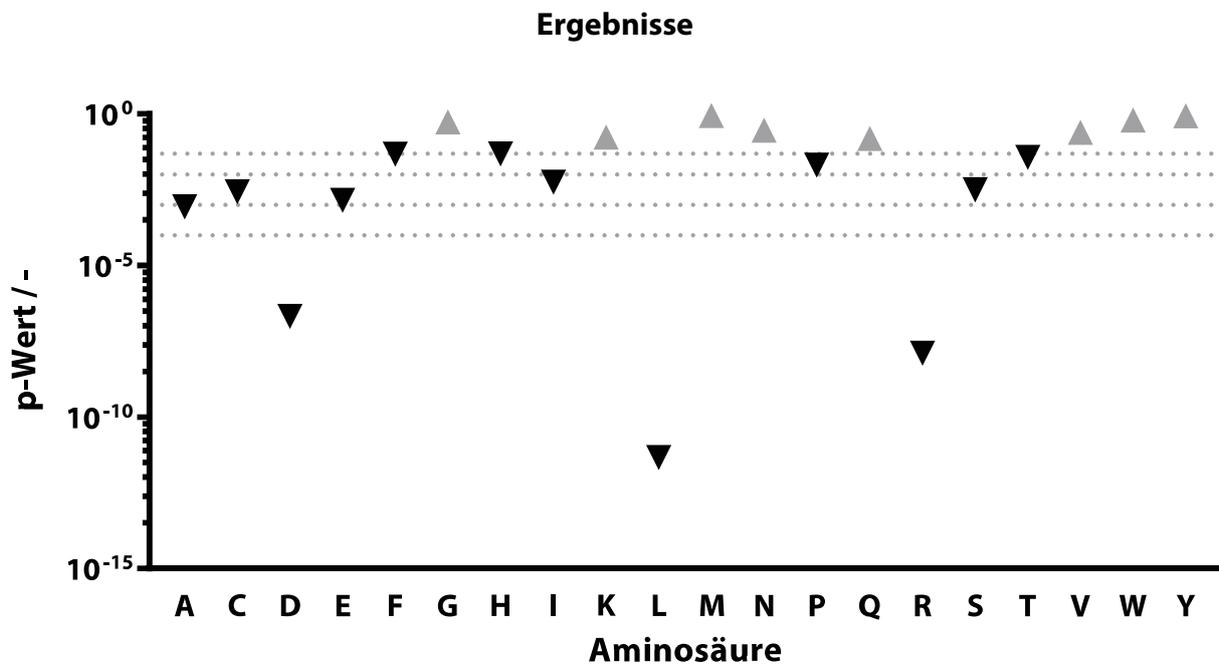


Abbildung 14: Aminosäurefrequenzvergleich thermostabiler und -labiler Proteine. Signifikanzgrade (p-Werte von 0,05; 0,01; 0,001; 0,0001) sind durch gepunktete Linien kenntlich gemacht. Signifikante Werte ($p \leq 0.05$) sind in Schwarz, nicht signifikante Werte in Grau gehalten. Die Pfeilrichtung gibt an, ob der Median der thermisch stabilen Fraktion über (▲) oder unter (▼) dem der labilen Fraktion liegt.

Im Vergleich der Aminosäurefrequenzen aller sechs Organismen zeigt sich hochsignifikant (p-Wert < 0,0001) weniger Aspartat (D), Leucin (L) und Arginin (R) in der Fraktion stabiler Proteine (Abbildung 14). Alle restlichen, signifikanten Unterschiede thermisch labiler und stabiler Proteine ($0,05 \leq p\text{-Wert} > 0,0001$) weisen auf eine Frequenzabnahme von insgesamt 9 Aminosäuren (A, C, E, F, H, I, P, S, T) in stabileren Proteinen hin. Alle verbleibenden 8 Aminosäuren (G, K, M, N, Q, V, W, Y) sind nicht signifikant häufiger in stabilen als in labilen Proteinen vorzufinden.

Detaillierte Werte pro Aminosäure, aufgeschlüsselt nach Organismus, können Tabelle 38 im Anhang entnommen werden.

3.2.5 Analyse von Paralogen und Orthologen

Paraloge

Paraloge Gene haben ihren Ursprung in einer Genverdopplung innerhalb eines Genoms im Lauf der Evolution¹⁰⁶. Auch wenn sich die Kopie, da diese nun (zunächst) keinem Selektionsdruck mehr unterworfen ist, unabhängig vom originalen Gen entwickeln kann, weisen Paraloge meist eine hohe Sequenzübereinstimmung auf. Somit ähneln sich paraloge Proteine oftmals strukturell sowie funktionell und finden sich häufig in gleichen Stoffwechselwegen oder Proteinkomplexen wieder.

Wie weit die Thermostabilität von Proteinen evolutiv konserviert ist, sollte mit dem Vergleich von Paralogen überprüft werden. Anhand ihrer eggNOG-Klassifikation¹⁰⁷ nach LUCA (*engl.*: last universal common ancestor), wurden entsprechende Proteine einander zugeordnet, die Differenz ihrer T_m -Werte (ΔT_m) gebildet und diese nach ihrer berechneten Sequenzidentität (nach NEEDLEMAN & WUNSCH¹⁰⁸) gruppiert.

Für die Analyse wurden die eggNOG-Klassifikatoren mit UniProt „entry identifier“ vom Server des EMBL bezogen (http://eggnogdb.embl.de/download/eggnog_4.5/id_mappings/uniprot/uniprot-15-May-2015.tsv.gz). Die Zuordnung dieser Klassifikatoren zu den gemessenen Proteinen erfolgte in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „entry identifier“^{36,88}. Paraloge wurden durch das Abgleichen eines jeden Organismus mit sich selbst über *left_join* anhand der eggNOG-Klassifikatoren gefunden. Die hierbei entstehenden Duplikate wurden im Anschluss herausgefiltert.

Da für *O. antarctica* und *G. stearotermophilus* noch keine eggNOG-Klassifikationen vorhanden waren, wurden diese mit Hilfe des eggNOG-Mappers zugeordnet¹⁰⁹. Als „Mapping Methode“ wurde DIAMOND gewählt und der Taxonomiebereich für *O. antarctica* auf γ -Proteobacteria sowie für *G. stearotermophilus* auf Bacilli eingeschränkt.

Ergebnisse

Bei der Erstellung randomisierter Paare, wurde die Paarung von Proteinen mit sich selbst ausgeschlossen. Die Unterschiede zwischen gepaarten und randomisierten Paralogen wurden mit Hilfe eines WILCOXON-MANN-WHITNEY-Tests bestimmt (Details zur Funktion und gewählten Parametern: siehe Seite 57).

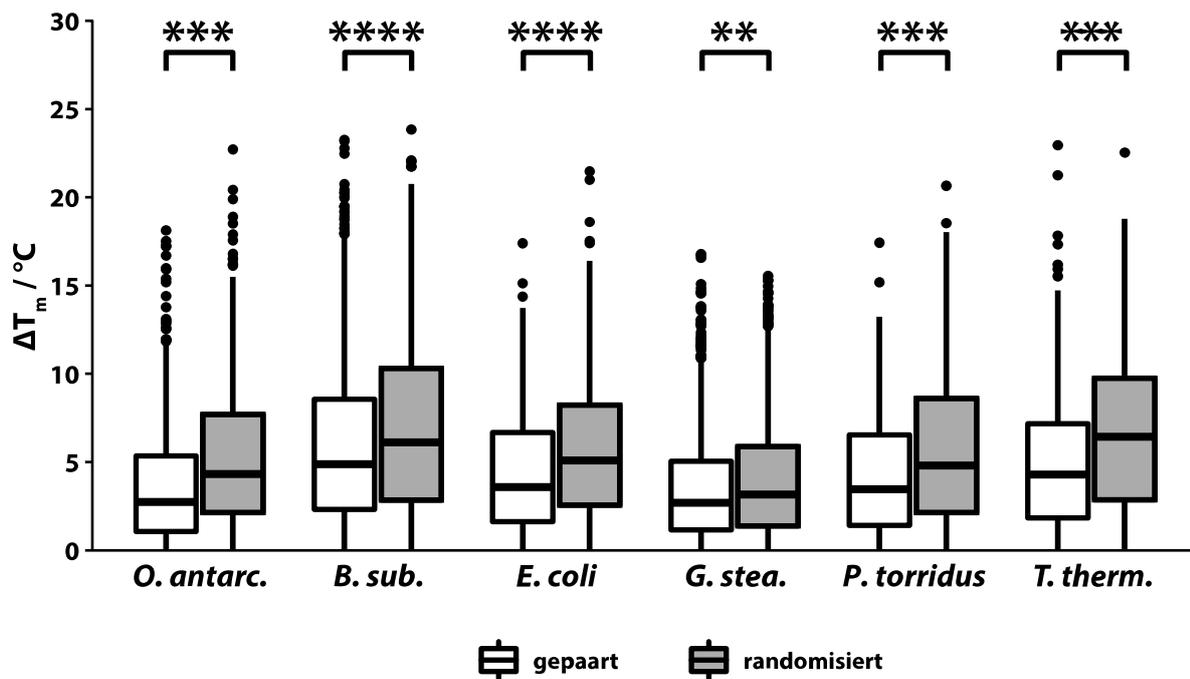


Abbildung 15: T_m -Differenzwerte von gepaarten und randomisierten Paralogen nach Organismus. Darstellung nach TUKEY. Die Signifikanzgrade wurden anhand eines einseitigen WILCOXON-Tests berechnet und sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Gepaarte Paraloge zeigen einen T_m -Werteunterschied mit einem Durchschnitt von etwa 3,5–6 °C (Abbildung 15). *O. antarctica* befindet sich hierbei am unteren und *E. coli* am oberen Ende. Bei allen Organismen liegen die kleinsten Differenzwerte von Paralogen nahe 0 °C und reichen bis etwa 17–23 °C. Randomisiert man die Zuordnung zwischen Paralogen, so zeigt ein Vergleich beider Verteilungen anhand eines WILCOXON-Tests in allen Fällen einen signifikant höheren ΔT_m -Wert (Abbildung 15 und Tabelle 25). Der Umfang des T_m -Wertebereichs bleibt trotz Randomisierung annähernd gleich.

Ergebnisse

Tabelle 24: Signifikanzwerte eines einseitigen WILCOXON-Tests von gepaarten und randomisierten Paralogen. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Organismus	p-Wert	Signifikanz
<i>O. antarctica</i>	4,10E-04	***
<i>B. subtilis</i>	2,21E-15	****
<i>E. coli</i>	1,16E-05	****
<i>G. stearothermophilus</i>	2,87E-03	**
<i>P. torridus</i>	1,40E-04	***
<i>T. thermophilus</i>	1,51E-04	***

Ein möglicher Zusammenhang zwischen der Vererbbarkeit von Thermostabilität und Primärsequenz kann über die Sequenzidentität von Paralogen erbracht werden. Der Vergleich von Paralogen und willkürlich zugeordneten Paaren, unter diesem Aspekt, soll Hinweis geben, wie stark sich die zuvor beobachteten Effekte vom Zufall abheben.

Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „*entry identifier*“ den paralogen Paaren zugeordnet^{36,88}. Die Sequenzidentität wurde über die Funktion *pairwiseAlignment* (R-Bibliothek: *Biostrings*) berechnet⁸⁷.

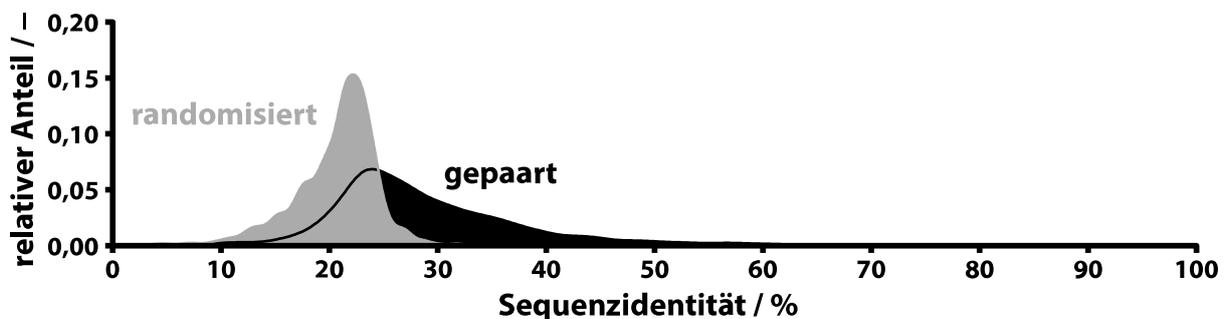


Abbildung 16: Gepaarte und randomisierte Paraloge nach Sequenzidentität (alle Mikroorganismen).

Die Verteilung der Sequenzidentitäten aller Organismen zeigen eine rechtsschiefe Verteilung mit einem Mittelwert von 29,2 % (Abbildung 16). Wenn auch sehr niederfrequent, reicht der obere Wertebereich bis 99,7 % Sequenzidentität (bei *E. coli*). Werden die ursprünglichen Paraloge zufällig zugeordnet, so sinkt der Durchschnitt auf 20,7 %. Ebenso gibt sich die Verteilung schmaler und (leicht linksschief) symmetrisch. Der Signifikanztest nach WILCOXON

Ergebnisse

weist mit einem p-Wert von $< 2,2 \times 10^{-16}$ einen eindeutigen Unterschied auf. Details zum verwendeten WILCOXON-MANN-WHITNEY-Tests und zu gewählten Parametern: siehe Seite 57.

Orthologe

Bei orthologen Genen handelt es sich um Homologe, welche auf ein einzelnes Gen im letzten gemeinsamen Vorfahren zweier zu vergleichenden Organismen zurückgehen¹¹⁰. Sie bilden das Gegenstück zu paralogenen Genen, welche im selben Organismus vorliegen und auf eine Genduplikation zurückzuführen sind. Aus diesem Grund sind Orthologe in der Regel älter als Paraloge.

Für die Analyse wurden die eggNOG-Klassifikatoren mit UniProt „entry identifier“ vom Server des EMBL bezogen (http://eggnogdb.embl.de/download/eggnog_4.5/id_mappings/uniprot/uniprot-15-May-2015.tsv.gz). Die Zuordnung dieser Klassifikatoren zu den gemessenen Proteinen erfolgte in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) anhand der UniProt „entry identifier“^{36,88}. Orthologe wurden durch das Abgleichen aller Organismen untereinander mit *left_join* anhand der eggNOG-Klassifikatoren gefunden.

Die Unterschiede zwischen gepaarten und randomisierten Orthologen wurden mit Hilfe eines WILCOXON-MANN-WHITNEY-Tests bestimmt (Details zur Funktion und gewählten Parametern: siehe Seite 57).

Ergebnisse

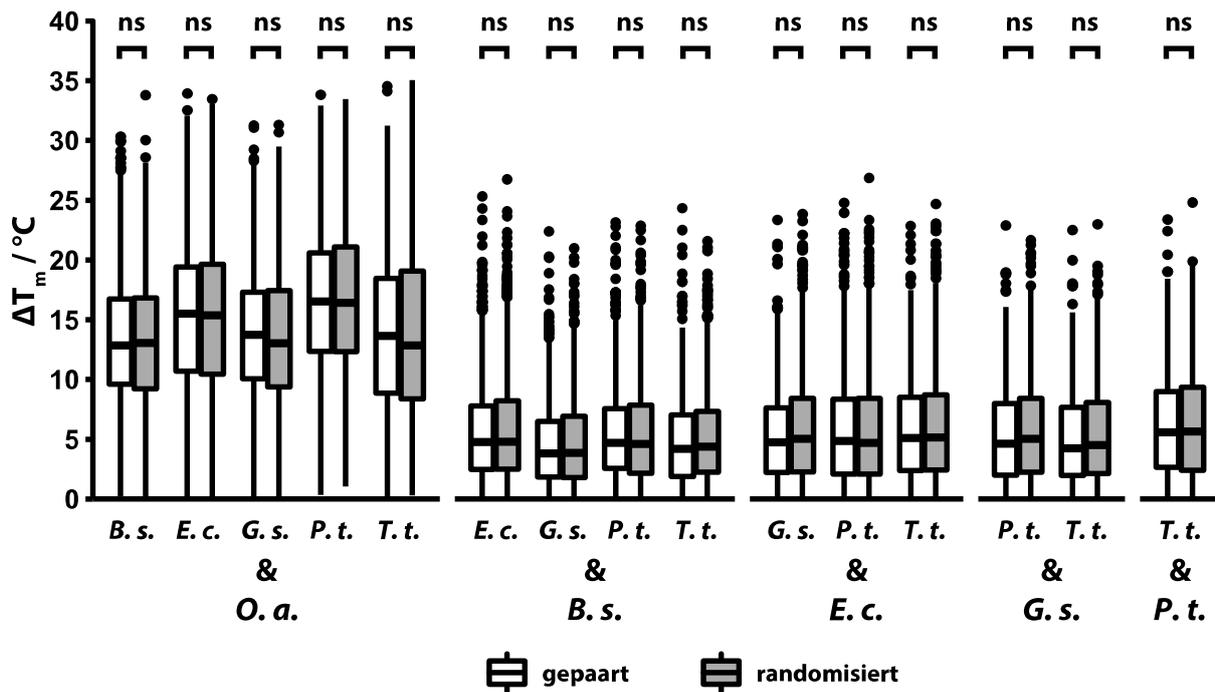


Abbildung 17: T_m -Differenzwerte von gepaarten und randomisierten Orthologen zwischen allen Organismen. Die Namen der Mikroorganismen wurden wie folgt abgekürzt: *O. a.*: *O. antarctica*, *B. s.*: *B. subtilis*, *E. c.*: *E. coli*, *G. s.*: *G. stearothermophilus*, *P. t.*: *P. torridus*, *T. t.*: *T. thermophilus*. Darstellung nach TUKEY. Die Signifikanzgrade wurden anhand eines einseitigen WILCOXON-Tests berechnet und sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Die Verteilungen der Schmelzpunktdifferenzen von Orthologen zeigten in den Mittelwerten zwischen den unterschiedlichen Verteilungen keine größeren Unterschiede auf und lagen um einen Bereich von 4–5 °C (Abbildung 17). Die Orthologenpaare mit *O. antarctica* waren im Vergleich zu allen anderen Orthologenpaarungen zu größeren Differenzen hin verschoben. Hier schwankten zudem die Mittelwerte zwischen den Paaren mit den unterschiedlichen Organismen. Der Verteilungsvergleich zwischen Orthologen und randomisierten Paaren zeigte in keinem der Fälle einen signifikanten Unterschied.

Wie schon zuvor bei Paralogen (Abbildung 16) können Orthologe ebenfalls anhand ihrer Sequenzidentität analysiert werden. Für die Analyse wurden die Primärsequenzen der UniProt-Datenbank entnommen und in R über die Funktion `left_join` (R-Bibliothek: `dplyr`) anhand der UniProt „*entry identifier*“ den orthologen Paaren zugeordnet^{36,88}. Die Sequenzidentität wurde über die Funktion `pairwiseAlignment` (R-Bibliothek: `Bistrings`) berechnet⁸⁷.

Ergebnisse

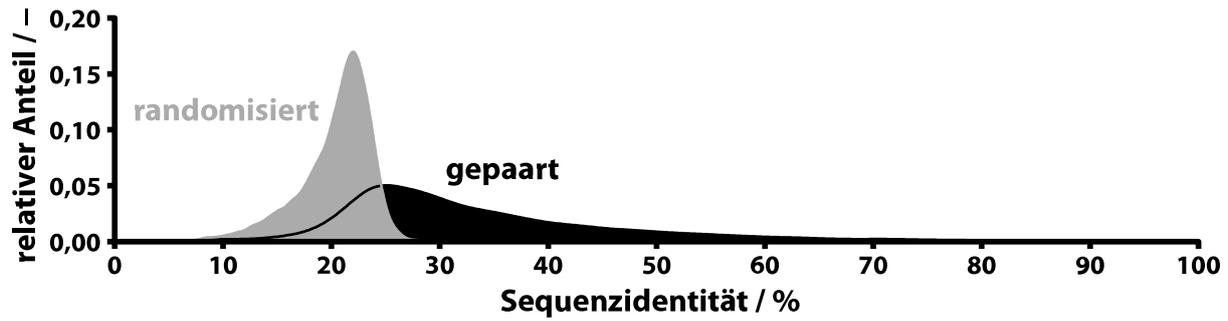


Abbildung 18: Gepaarte und randomisierte Orthologe nach Sequenzidentität (alle Mikroorganismen).

Die Verteilungen von Orthologen und randomisierten Paaren ähnelten den der Paralogen stark. Es zeigte sich eine rechtsschiefe Verteilung mit einem (vergleichbaren) Mittelwert von 33,6 %. Die maximale Übereinstimmung von 98,3 % wird zwischen *B. subtilis* und *G. stearothermophilus* erreicht. Die willkürliche Zuordnung von Paaren ergab eine schmale, symmetrische, leicht linksschiefe Verteilung. Der Mittelwert von 20,4 % ist mit dem der randomisierten Paralogen beinahe identisch (20,3 %). Der Signifikanztest nach WILCOXON weist mit einem p-Wert von $< 2,2 \times 10^{-16}$ einen eindeutigen Unterschied auf. Details zum verwendeten WILCOXON-MANN-WHITNEY-Tests und zu gewählten Parametern: siehe Seite 57.

Ergebnisse

Tabelle 25: Signifikanzwerte eines einseitigen WILCOXON-Tests von gepaarten und randomisierten Orthologen. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Organismus 1	Organismus 2	p-Wert	Signifikanz
<i>O. antarctica</i>	<i>B. subtilis</i>	9,80E-01	ns
	<i>E. coli</i>	9,13E-01	ns
	<i>G. stearothermophilus</i>	5,79E-01	ns
	<i>P. torridus</i>	9,57E-01	ns
	<i>T. thermophilus</i>	5,41E-01	ns
<i>B. subtilis</i>	<i>E. coli</i>	6,04E-01	ns
	<i>G. stearothermophilus</i>	3,90E-01	ns
	<i>P. torridus</i>	6,30E-01	ns
	<i>T. thermophilus</i>	6,67E-02	ns
<i>E. coli</i>	<i>G. stearothermophilus</i>	6,18E-02	ns
	<i>P. torridus</i>	9,56E-01	ns
	<i>T. thermophilus</i>	7,03E-01	ns
<i>G. stearothermophilus</i>	<i>P. torridus</i>	1,88E-01	ns
	<i>T. thermophilus</i>	1,06E-01	ns
<i>P. torridus</i>	<i>T. thermophilus</i>	9,12E-01	ns

3.2.6 Gegenüberstellung von intakten Zellen und Zellextrakten

Der Vorteil von TPP gegenüber anderen Methoden, wie etwa LiP-MS¹⁰⁵, ist die Möglichkeit, Schmelzkurven für Proteine in intakten Zellen zu bestimmen. Zusätzlich zu den Zellextrakten wurden für *E. coli* und *T. thermophilus* auch intakte Zellen gemessen.

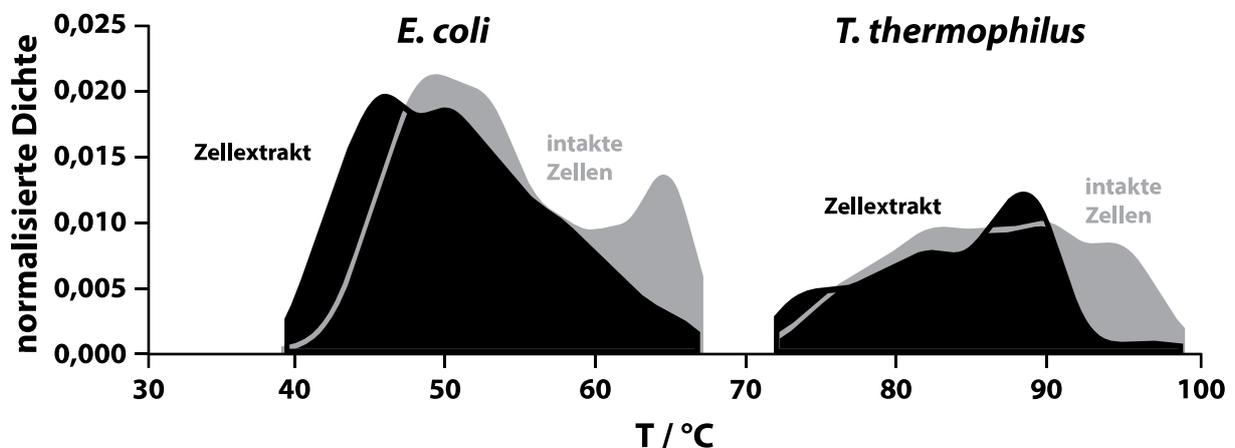


Abbildung 19: Verteilung der T_m -Werte von *E. coli* und *T. thermophilus* für Zellextrakte und intakte Zellen. Die Verteilung wurde anhand der jeweiligen Proteinanzahl normalisiert.

Ergebnisse

Für *E. coli* und *T. thermophilus* zeigte sich eine unterschiedliche Verteilung der proteomweiten T_m -Werte im Vergleich von Zellextrakt mit Werten aus intakten Zellen (Abbildung 19). Während der *E. coli* Zellextrakt überwiegend eine rechtsschiefe Verteilung ergab, zeigte die Verteilung intakter Zellen einen eindeutig bimodalen Charakter mit einer relativ großen zweiten Population im oberen Temperaturbereich. Das *T. thermophilus* Proteom verhielt sich – abgesehen von den naturgemäß großen Unterschieden in den Temperaturbereichen der T_m -Werte – im Vergleich zwischen Zellextrakt und intakten Zellen anders als das von *E. coli*. Während der Zellextrakt eine linksschiefe Verteilung mit einem starken Abfall im oberen Temperaturbereich ergab, erwiesen sich die T_m -Werte der intakten Zellen überwiegend gleichmäßig und nur sehr leicht linksschief verteilt.

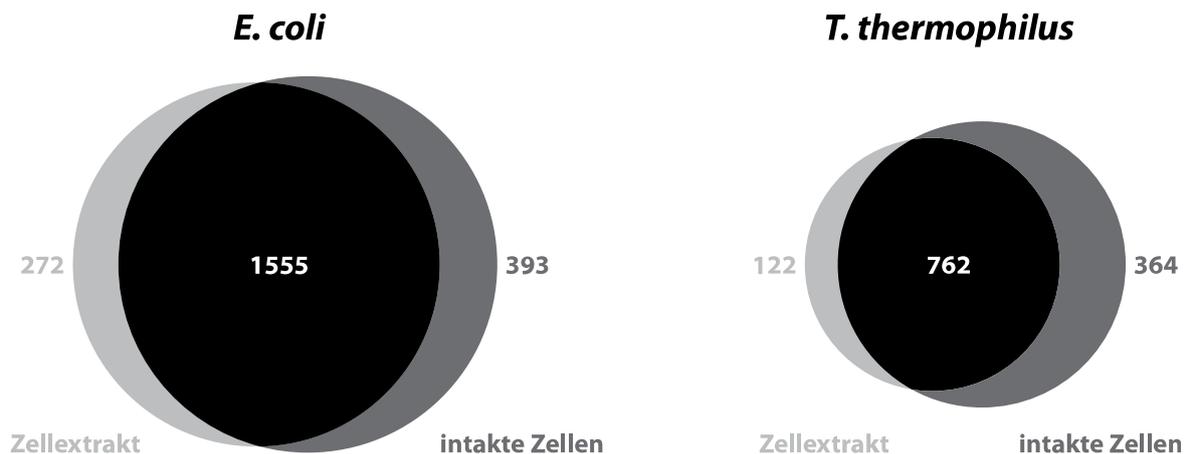


Abbildung 20: Venn-Diagramm für Proteine aus Zellextrakten und intakten Zellen für *E. coli* und *T. thermophilus*. Die Anzahlen entsprechen den Proteinen, für welche ein T_m -Wert bestimmt werden konnte.

Zusammengenommen umfassten die gemessenen Proteome von Zellextrakt und intakten Zellen von *E. coli* 2.220 unterschiedliche Proteine (Abbildung 20). Der Anteil gemeinsam vorkommender Proteine lag mit 1.555 Proteinen bei 70,0 %. Im Falle von *T. thermophilus* wurden insgesamt 1.248 individuelle Proteine gemessen, wovon 61,1 % von Zellextrakt und intakten Zellen geteilt wurden.

Durch eine Auftragung der Schmelztemperaturen eines jeden Proteins aus Zellextrakt und intakten Zellen sollte ein Überblick über den Zusammenhang bzw. die Unterschiede erhalten

Ergebnisse

werden. Zusätzlich wurden die Proteine anhand ihrer zellulären Lokalisation (Cytoplasma, Cytoplasmamembran, Periplasma und äußere Membran) anhand der Annotationen der PSORTdb¹¹¹⁻¹¹³ weiter unterteilt.

Für die Analyse wurden die subzellulären Lokalisationsdaten für *E. coli* und *T. thermophilus* (Chromosom und Megaplasmid separat) von PSORTdb (<http://db.psort.org>) bezogen. Die Zuordnung dieser Lokalisationsinformationen zu den Proteinen erfolgte über einen Umweg, da diese nicht über UniProt „entry identifier“ verfügen. Über den UniProt „Retrieve/ID mapping“-Dienst wurde eine Liste erstellt, welche UniProt „entry identifier“ die entsprechende RefSeq-ID (*E. coli*) bzw. NCBI GI (*T. th.*) zuordnet. Mit dieser Liste als Bindeglied erfolgte mit der Funktion *left_join* (R-Bibliothek: *dplyr*) die zweistufige Zuordnung der PSDORTdb Daten zu den Proteinen.

Ergebnisse

E. coli

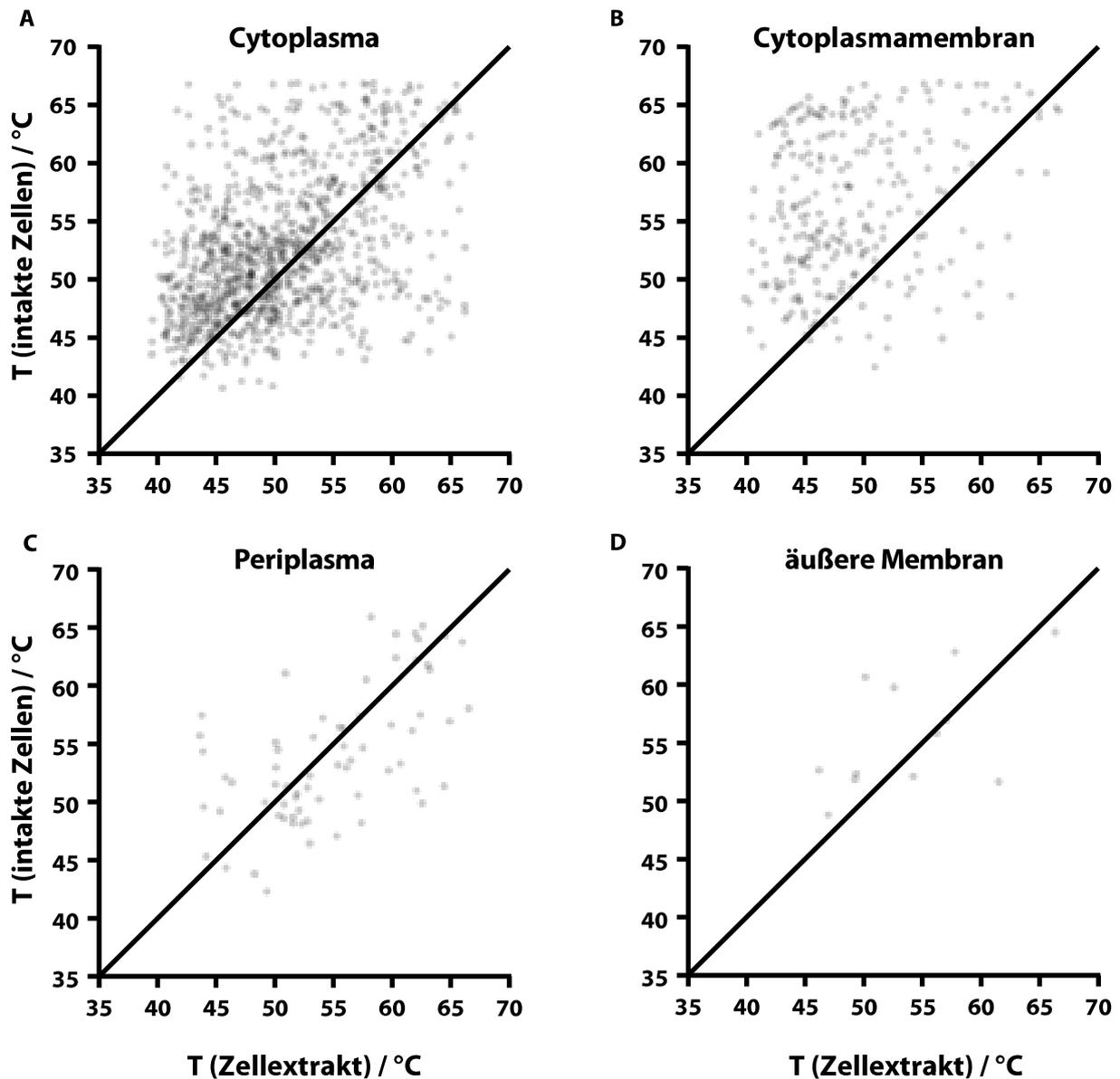


Abbildung 21: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in *E. coli* nach Lokalisation. Die schwarze Linie ist eine Ursprungsgerade mit einer Steigung von 1.

Die cytoplasmatischen Proteine, welche den Großteil des Proteoms definieren, zeigten eine weitgehend symmetrische Verteilung entlang und anscheinend leichte Häufung oberhalb der Winkelhalbierenden (Abbildung 21A), ihre T_m -Werte lagen vermehrt in einem Bereich von 40–55 °C. Proteine der Cytoplasmamembran (Abbildung 21B) zeigten im Vergleich höhere Schmelztemperaturen für intakte Zellen (Großteil der Werte oberhalb der Winkelhalbierenden). Für im Periplasma oder in der äußeren Membran (Abbildung 21C und

Ergebnisse

Abbildung 21D) befindliche Proteine, ergibt sich wieder ein symmetrisch verteiltes Bild – wenn auch mit deutlich reduzierter Proteinanzahl.

T. thermophilus

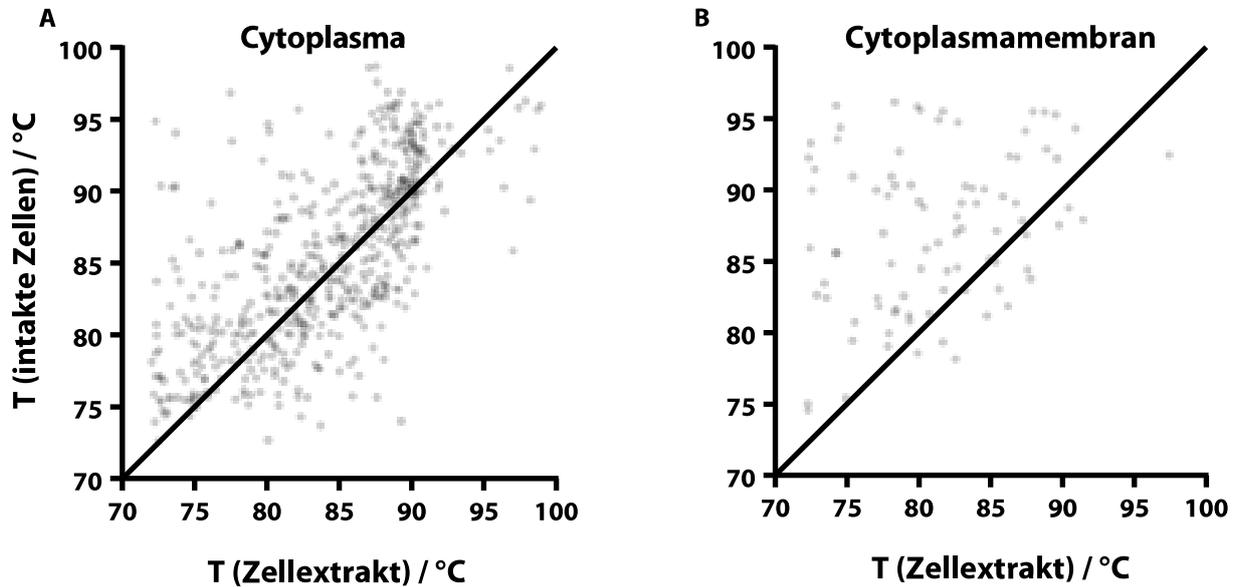


Abbildung 22: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in *T. thermophilus* nach Lokalisation. Die Graphen für Proteine des Periplasmas und der äußeren Membran sind im Anhang vorzufinden (Abbildung 33). Die schwarze Linie ist eine Ursprungsgerade mit einer Steigung von 1.

Ähnlich wie bei *E. coli* lagen auch bei *T. thermophilus* die meisten cytoplasmatischen Proteine, die den Großteil des Proteoms stellen, symmetrisch entlang der Winkelhalbierenden verteilt (Abbildung 22A). Hier fiel eine auffällige Anhäufung einiger Proteine, die in intakten Zellen stabiler als im Zellextrakt waren, entlang 90 °C der Zellextraktprobe auf, was für deren Stabilisierung in intakten Zellen spricht. Wie schon bei *E. coli* war auch bei *T. thermophilus* bei Proteinen der Cytoplasmamembran eine thermische Stabilisierung in intakten Zellen zu sehen (Abbildung 22B). Für Proteine des Periplasmas und der äußeren Membran wurde jeweils nur ein Protein identifiziert. Graphen hierzu finden sich im Anhang (Abbildung 33).

Für die oben gezeigten Graphen (Abbildung 21 und Abbildung 22) ergaben sich die unten stehenden PEARSON-Korrelationskoeffizienten (Tabelle 26). Da bei *T. thermophilus* jeweils nur ein im Periplasma und in der äußeren Membran befindliches Protein identifiziert wurde, konnte für diese keine Korrelationskoeffizienten berechnet werden. Die Berechnung von

Ergebnisse

Korrelationswerten erfolgte mit der R-eigenen Funktion *cor.test*. Hierbei wurde der Parameter *method* auf „pearson“ gesetzt.

Tabelle 26: PEARSON-Korrelationskoeffizienten von T_m -Werten aus Zellextrakt und intakten Zellen nach Lokalisation. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

Organismus	Lokalisation	PEARSON- Korr.koeff.	p-Wert	Sig.
<i>E. coli</i>	Cytoplasma	0,431	6,83E-47	****
	Cytoplasmamembran	0,257	3,70E-05	****
	Periplasma	0,587	1,83E-07	****
	Äußere Membran	0,557	5,99E-02	ns
<i>T. thermophilus</i>	Cytoplasma	0,674	1,92E-71	****
	Cytoplasmamembran	0,271	1,05E-02	*

Die Korrelationskoeffizienten (alle signifikant außer Äußere Membran bei *E. coli*) zeigen für beide Organismen eine höhere Korrelation bei cytoplasmatischen Proteinen als bei jenen, welche in der Cytoplasmamembran lokalisiert sind. Die *E. coli* angehörigen periplasmatischen und in der äußeren Membran vorhandenen Proteine zeigen ebenfalls eine hohe Korrelation; letztere jedoch statistisch nicht signifikant.

3.2.7 *T. thermophilus*: Vergleich von Chromosom und Megaplasmid

Das Genom von *T. thermophilus* besteht, wie bereits in Kapitel 1.2.6 beschrieben, aus einem Chromosom (TTC, 1,89 MBp) und einem Megaplasmid (pTT27, 0,23 MBp). Die Proteine können daher nicht nur anhand ihrer zellulären, sondern auch der genomischen Lokalisation der für sie kodierenden Gene unterschieden werden.

Das Megaplasmid umfasst mit 202 Genen 9,5 % des Gesamtgenoms. Von den massenspektroskopisch im Zellextrakt und in den intakten Zellen detektierten Proteinen waren 9,8 bzw. 9,2 % Megaplasmid-kodiert.

Ergebnisse

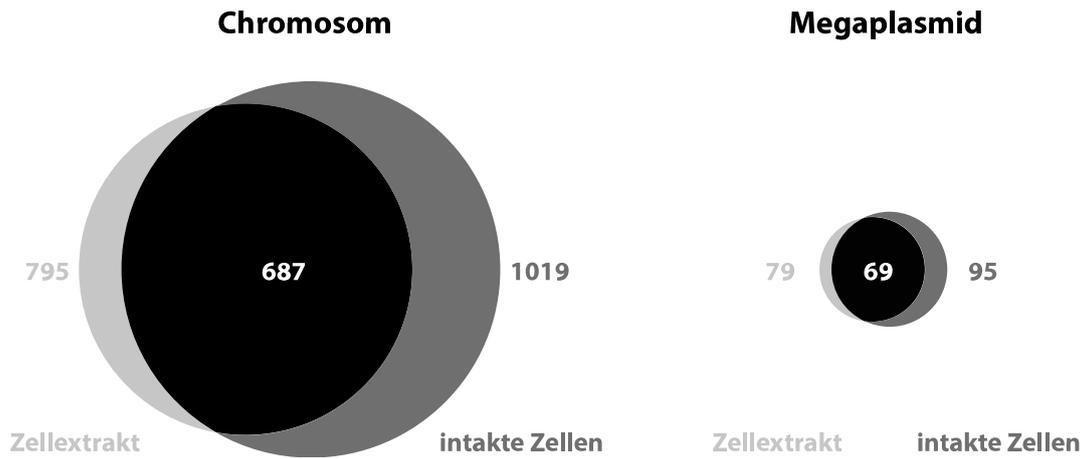


Abbildung 23: Venn-Diagramm: Proteine aus *T. th.* nach genomischer Lokalisation (Zellextrakt und intakte Zellen). Die Anzahlen entsprechen den Proteinen, für welche ein T_m -Wert bestimmt werden konnte.

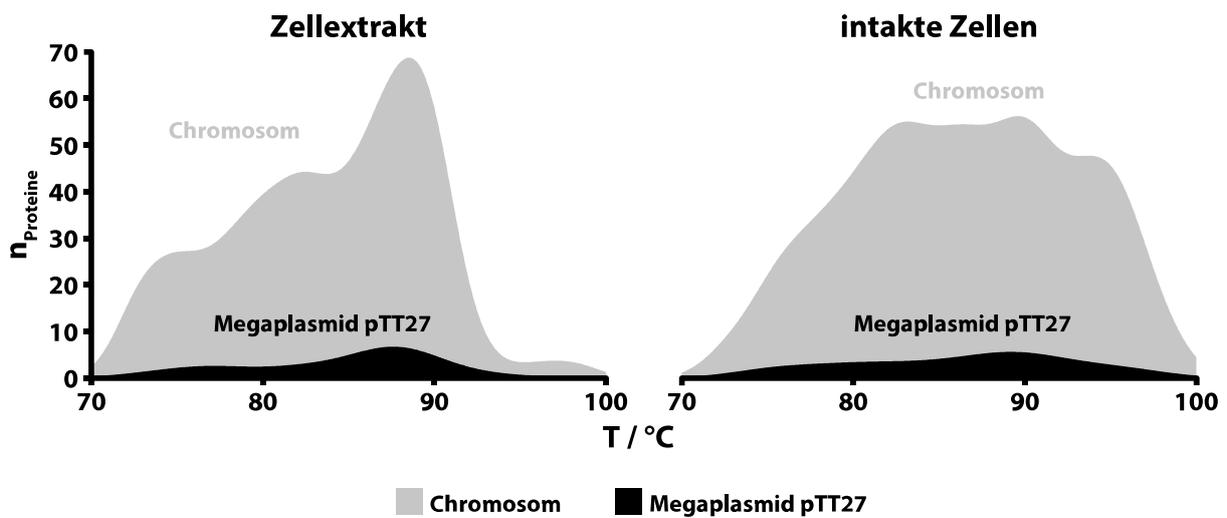


Abbildung 24: T_m -Werteverteilung für *T. thermophilus* in Zellextrakt und intakter Zellen. Die Verteilungen sind nach genomischer Lokalisation der hierbei kodierenden Gene aufgeteilt und geschichtet dargestellt.

Ergebnisse

In beiden Fällen zeigte sich für die Verteilung der vom Megaplasmid kodierten Proteine ein ähnliches Bild (Abbildung 24). Der häufigste Wert lag leicht unter 90 °C (bei den intakten Zellen etwas höher). Die Hauptvariabilität war bei den chromosomal kodierten Proteinen festzustellen. Die charakteristische Form der jeweiligen T_m -Verteilung wird maßgeblich durch diese geprägt. Sie sind diejenigen, die die meisten Unterschiede zwischen Zellextrakt und intakten Zellen aufweisen. Eine detailliertere Darstellung findet sich untenstehend (Abbildung 25).

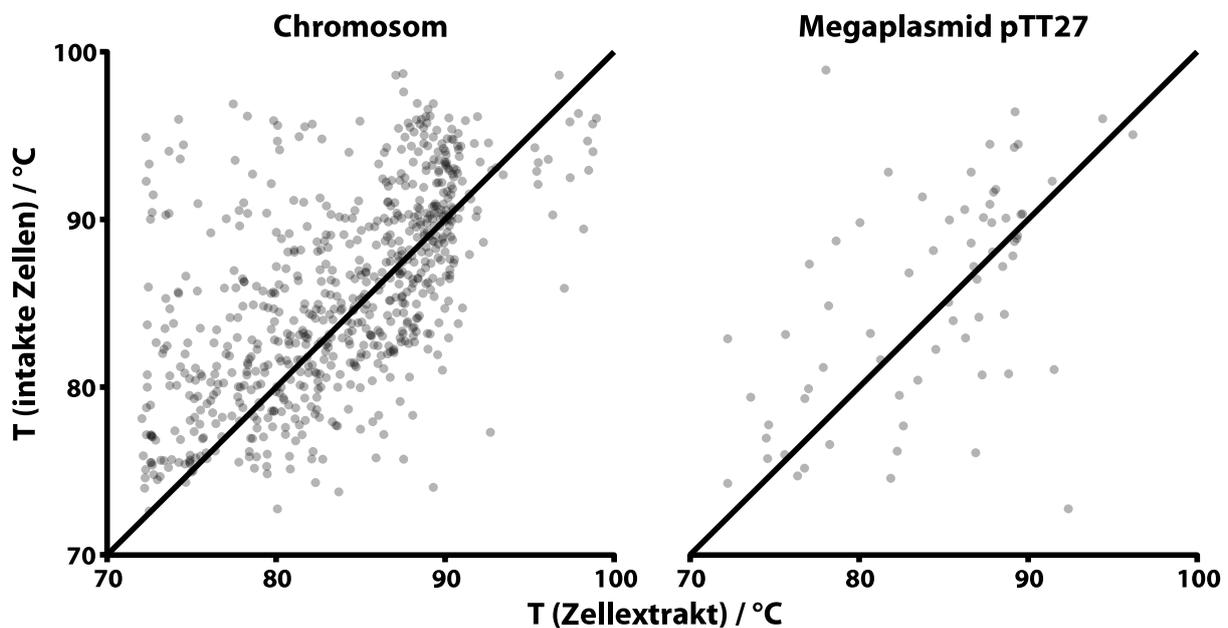


Abbildung 25: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in *T. thermophilus*. Die T_m -Werte sind nach genomischer Lokalisation der hierbei kodierenden Gene aufgeteilt. Die schwarze, diagonale Linie ist eine Ursprungsgerade mit einer Steigung von 1.

3.2.8 Thermische Inaktivierung von *E. coli* und *T. thermophilus*

Jeder Organismus besitzt eine charakteristische Höchsttemperatur, bei welcher noch Wachstum möglich ist. Diese wurde für *E. coli* und *T. thermophilus* bereits bestimmt^{114,115}. Diese maximale Wachstumstemperatur deckt sich nicht notwendigerweise mit einer (höheren) Temperatur, der ein Organismus für einen gewissen Zeitraum unbeschadet ausgesetzt werden kann^{114,116}.

Um *E. coli* und *T. thermophilus* in Bezug auf ihre Überlebensfähigkeit bei Temperaturen über dem jeweiligen Wachstumsoptimum näher zu charakterisieren und die so erhaltenen Daten

Ergebnisse

mit den T_m -Werten der korrespondierenden Proteome in Bezug setzen zu können, wurden Bakterienkulturen mit *E. coli* und *T. thermophilus* unterschiedlichen Temperaturen (Bereich: optimale Wachstumstemperatur bis etwa 30 °C darüber) für 10 min ausgesetzt und der Anteil überlebender Zellen bestimmt (Kapitel 2.27). Die gewählte Inkubationszeit entspricht der Dauer der Probeninkubation zur thermischen Denaturierung bei der Aufbereitung der Proben für TPP-Experimente.

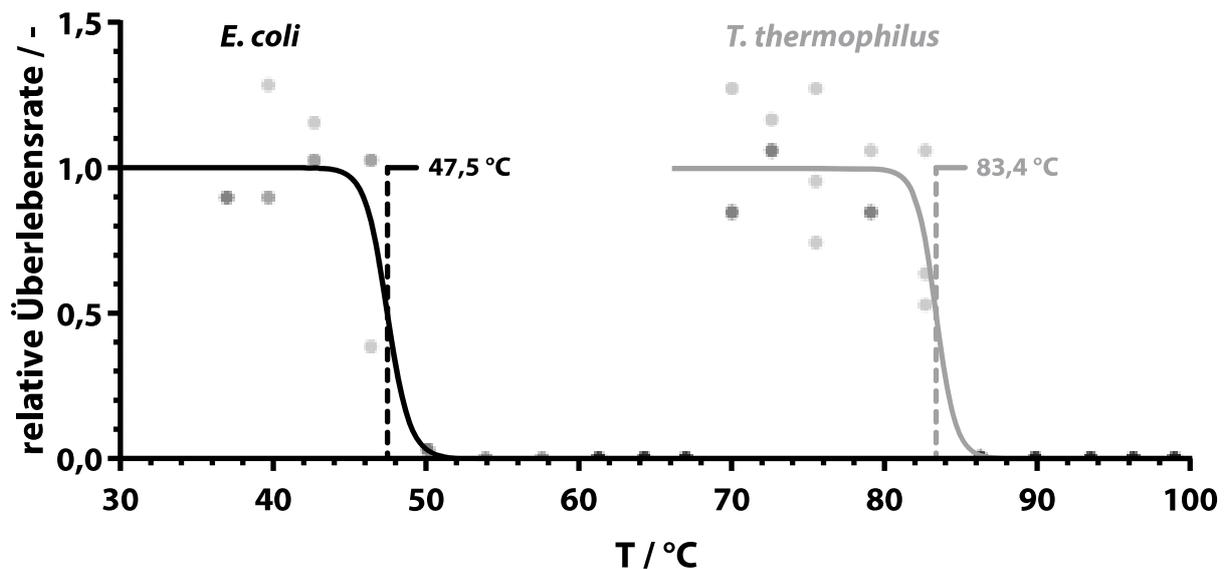


Abbildung 26: Relative Überlebensrate von *E. coli* und *T. thermophilus* in Abhängigkeit von der Temperatur. Links ist *E. coli* in Schwarz, rechts *T. thermophilus* in Grau dargestellt. Eine Funktion logistischer Regression ist jeweils überlagert. Die vertikalen Linien markieren die Temperatur, bei welcher 50 % aller Zellen inaktiviert wurden. Die Messung erfolgte jeweils in Triplikaten.

Beide Organismen zeigten einen schmalen Übergang von vitalen hin zu nicht mehr lebenden Zellen und ein vergleichbar ausgeprägtes Profil. Bei *E. coli* waren bereits bei 47,5 °C (nach 10 min) die Hälfte der Zellen nicht mehr lebensfähig (Abbildung 26). Dieser Wert deckt sich fast genau mit der maximalen Wachstumstemperatur von 47,4 °C, die in der Literatur beschrieben wurde¹¹⁴. Im Fall von *T. thermophilus* waren 50 % aller Bakterien bei 83,4 °C abgestorben. Dieser Wert liegt leicht unter dem gemessenen Wert der maximalen Wachstumstemperatur von 84 °C^{115,117}.

Ergebnisse

Um einen Temperaturbereich festzulegen, in welchem die Zellen stark in ihrer Überlebensfähigkeit beeinträchtigt sind, wurde das Intervall der 90–10 %igen Überlebensrate berechnet.

Tabelle 27: Berechnete Temperaturen für Überlebensraten von 10, 50 und 90 % von *E. coli* und *T. thermophilus* auf Grundlage der in Abbildung 26 dargestellten Experimente.

Überlebensrate	90 %	50 %	10 %
<i>E. coli</i>	45,9 °C	47,5 °C	49,1 °C
<i>T. thermophilus</i>	82,0 °C	83,4 °C	84,8 °C

Dieser Temperaturbereich wurde nun herangezogen, um im Folgenden nach Proteinen zu suchen, welche einen entsprechenden T_m -Wert aufweisen. Die auf diese Weise gefilterten Proteine lassen sich über eine Anreicherung von *Gene Ontology*-Begriffen (GO) näher kategorisieren. Hiermit sollte überprüft werden, ob eine gewisse Gruppe von Proteinen speziell in diesem Temperaturbereich denaturiert.

Für die Analyse wurde die R-Bibliothek *topGO* (ein Teil des Bioinformatik-Softwarepakets *Bioconductor* – <https://bioconductor.org>) verwendet⁹⁰. Die GO-Begriffanreicherung erfolgte nach Angaben des Referenzhandbuchs (PDF kann unter folgender URL abgerufen werden: <https://www.bioconductor.org/packages/release/bioc/html/topGO.html>). Die GO-Begriffe wurden von UniProt bezogen und die Zuordnung dieser zu den gemessenen Proteinen, erfolgte in R über die Funktion *left_join* (R-Bibliothek: *dplyr*) mit den UniProt „*entry identifier*“^{36,88}. Die schlussendliche Anreicherung mit *topGO* erfolgte anhand der Anzahl von GO-Begriffen, mit einem Proteinuniversum aus allen gemessenen Proteinen, einer *nodeSize* von „10“ und *ontology* von „MF“ (*molecular function*). Der Test wurde mit folgenden Parametern ausgeführt: *algorithm* = "classic", *statistic* = "fisher".

Ergebnisse

Tabelle 28: GO-Begriffanreicherung des thermischen Inaktivierungsbereichs von *E. coli*. Gezeigt sind die 5 Begriffe molekularer Funktion mit größter Signifikanz. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

E. coli

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0003676	nucleic acid binding	353	113	63,11	8,20E-14	****
GO:0003723	RNA binding	131	47	23,42	2,00E-07	****
GO:0003677	DNA binding	217	66	38,79	7,60E-07	****
GO:0005198	structural molecule activity	46	22	8,22	2,10E-06	****
GO:0003735	structural constituent of ribosome	40	20	7,15	2,60E-06	****

Tabelle 29: GO-Begriffanreicherung des thermischen Inaktivierungsbereichs von *T. thermophilus*. Gezeigt sind die 5 Begriffe molekularer Funktion mit größter Signifikanz. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

T. thermophilus

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0003924	GTPase activity	11	5	1,69	0,017	*
GO:0001883	purine nucleoside binding	19	7	2,92	0,017	*
GO:0005525	GTP binding	19	7	2,92	0,017	*
GO:0019001	guanyl nucleotide binding	19	7	2,92	0,017	*
GO:0032550	purine ribonucleoside binding	19	7	2,92	0,017	*

3.2.9 DSF mit intakten Zellen und Zellextrakten

Im Rahmen dieser Arbeit wurden Messungen von thermisch behandelten Bakterien und Zellextrakten mittels dynamischer Differenzfluorimetrie (*engl.*: DSF – differential scanning fluorimetry) durchgeführt. Dabei wird nicht die Überlebensrate durch thermische Inaktivierung von Zellen erfasst (Kapitel 3.2.8), sondern die Exposition hydrophober Strukturen in Abhängigkeit von der Temperatur. Dadurch kann die teilweise Entfaltung von Proteinen oder auch die Desintegration zellulärer Strukturen quantifiziert werden. Hier wurde die ansteigende Fluoreszenzintensität von SYPRO Orange bei der Interaktion mit hydrophoben Kontaktflächen für intakte Zellen und Zellextrakte von *E. coli* und *T. thermophilus* gemessen.

Ergebnisse

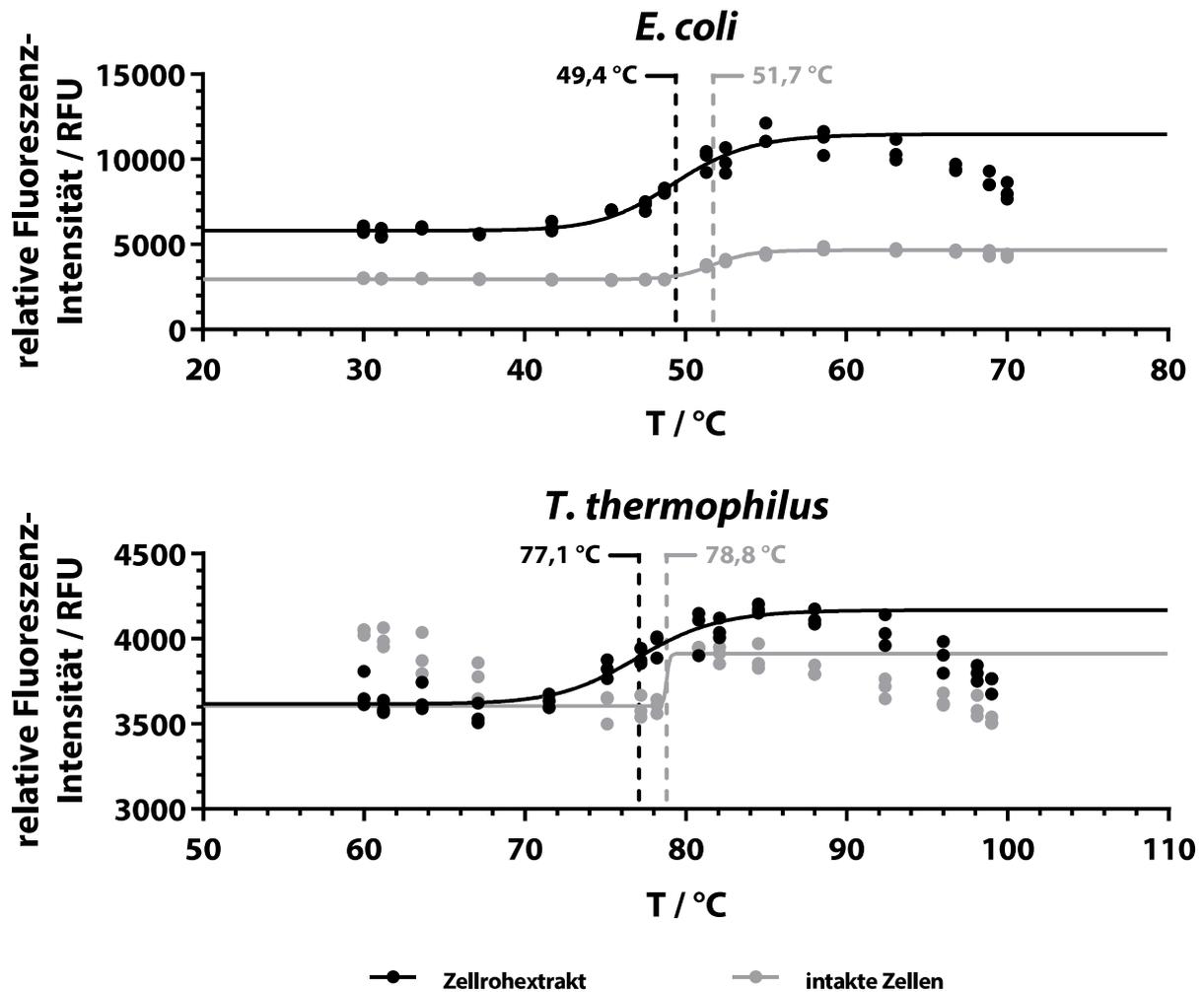


Abbildung 27: Isotherme, dynamische Differenzfluorimetrie von *E. coli* und *T. thermophilus*. Gezeigt ist die SYPRO Orange-Fluoreszenzintensität von Zellextrakt (schwarz) und intakten Zellen (grau) in Abhängigkeit von der Temperatur nach 10 min isothermer Inkubation. Eine Funktion logistischer Regression ist jeweils überlagert. Die vertikalen Linien markieren die Temperatur genau zwischen Minimal- und Maximalfluoreszenz. Die Messungen erfolgten jeweils in Triplikaten.

Alle Proben ergaben ein typisches Fluoreszenzprofil für dynamische Differenzfluorimetrie: Zunächst bleiben die Fluoreszenzwerte konstant um dann bei höheren Temperaturen – wie für DSF charakteristisch – sigmoidal anzusteigen. Erhöht man weiter die Temperatur, so nimmt die Fluoreszenz wieder ab. Letzteres lässt sich auf ein Verdrängen des Fluoreszenzfarbstoffs zurückführen, da denaturierte Proteine und andere molekulare Strukturen beginnen überwiegend untereinander zu wechselwirken.

Die Messungen mit den intakten *T. thermophilus*-Zellen wiesen eine Besonderheit auf: Hier nahmen die Fluoreszenzwerte im unteren Temperaturbereich (60–75 °C) zunächst ab, um darauffolgend wieder sigmoidal anzusteigen.

Ergebnisse

Vergleicht man die Messreihen der beiden Bakterien, so erkennt man, dass im Falle von *E. coli* jeweils ein höherer Fluoreszenzwert und ein höherer Dynamikumfang als bei *T. thermophilus* erreicht wurde. Auch zwischen Zellextrakt und intakten Zellen waren Unterschiede festzustellen. Bei *E. coli* wies der Zellextrakt eine etwa doppelt so hohe Fluoreszenzintensität im Vergleich zu Lebendzellen auf. Bei *T. thermophilus* lagen beide Proben dagegen im gleichen Intensitätsbereich.

Die Temperatur der 50 %igen Zunahme des Fluoreszenzsignals lag bei *E. coli* über, bei *T. thermophilus* hingegen unter der Temperatur der jeweiligen 50 %igen Zellinaktivierung (Tabelle 27 und Tabelle 30).

Tabelle 30: Temperaturen bei gegebener, normalisierter DSF-Fluoreszenzrate für *E. coli* und *T. thermophilus*.

normalisierte Fluoreszenzrate	Zellextrakt/ intakte Zellen	10 %	50 %	90 %
<i>E. coli</i>	Zellextrakt	44,5	49,4	54,8
	intakte Zellen	49,1	51,7	54,4
<i>T. thermophilus</i>	Zellextrakt	71,7	77,1	82,9
	intakte Zellen	78,6	78,8	79,0

4 DISKUSSION

4.1 Komplementierung eines GH-defizienten *Thermus thermophilus*-Stammes

Teile der nachfolgenden Diskussion basieren auf Ergebnissen, welche im Rahmen der Tagung „*Microbial Interactions 2018*“ in Rom, Italien (19.–20. Juli 2018) präsentiert wurden.

Die Intention dieses Projekts war die Entwicklung einer *T. thermophilus*-Deletionsmutante mit Wachstumsdefekt bzgl. der Verwertung von Cellobiose als Kohlenstoff- und Energiequelle, die als Wirt dienen kann, um damit rekombinante Genbibliotheken mittels Komplementierung des Wachstums in einem geeigneten Minimalmedium (TMB) auf der Suche nach Genen für neue β -Glucosidasen zu durchforsten. Gene für neue β -Glucosidasen können hierbei potenziell aus Metagenombanken, Mutagenesebanken o.ä. hervorgehen. Da es sich bei *T. thermophilus* um ein extrem thermophiles Wirtsbakterium handelt, kann ein für die komplementationsbasierte Selektion geeigneter Wirtsstamm sowohl zur Identifizierung neuartiger, thermostabiler Enzyme als auch zur Thermostabilisierung bereits charakterisierter Enzyme dienen. Denkbar ist zudem die Nutzung eines solchen Wirts zur Änderung oder Erweiterung des Substratspektrums einer bekannten GH, indem man eine Mutantenbibliothek eines Gens, dessen kodiertes Enzym Wachstum des *T. thermophilus* Wirtsstamms auf einem Substrat (z.B. ein β -Glucosid) erlaubt, herstellt und dann nach Wachstum auf einem anderen Substrat (z.B. ein β -Galactosid) selektiert. Sofern im Zuge der Selektion Wachstum stattgefunden hat, können die kultivierten Zellen aufgeschlossen und die Plasmide, Fosmide o.a. isoliert und sequenziert werden.

Ein alternativer Ansatz ist das Durchmustern (*engl.*: screening) von Genbanken unter Betrachtung einer jeden Genvariante. Dieser Ansatz ermöglicht zwar die detaillierte Betrachtung aller enthaltenen Gene, ist jedoch zeit- und kostenintensiv. Aufgrund der selektiven Kultivierungskomponente kann mit diesem System eine Vielzahl unterschiedlicher Enzymvarianten in nur einem Schritt analysiert werden. Nur, wenn der Phänotyp

Diskussion

wiederhergestellt wird, ist Wachstum möglich; Gene, welche nicht von Interesse sind, fallen bei dem selektiven Kultivierungsschritt heraus, weil die entsprechenden rekombinanten Klone nicht wachsen können. Weiterhin ist hervorzuheben, dass es sich bei komplementationsbasierter Selektion um eine *in vivo* Methodik handelt. Chaperone und ein breites Spektrum evtl. benötigte Kofaktoren, die im Wirtsbakterium vorhanden sind, erhöhen die Chance, eine funktionelle Variante positiv zu testen, die bei einem *in vitro* Enzymtest mit gleichem Enzym eventuell nicht erkannt worden wäre¹¹⁸. Allerdings hat auf der anderen Seite das *in vitro* Screening mit Enzymassays den Vorteil, dass relativ problemlos ein anderes Substrat herangezogen werden kann, während die selektive Komplementation abhängig ist von dem ihr zugrundeliegenden Stoffwechselweg.

4.1.1 Das Minimalmedium (TMB)

Die selektive Anreicherung funktioniert nur dann zuverlässig, wenn ausschließlich unter Anwesenheit von entsprechendem Substrat Wachstum stattfindet. Hierfür wurde ein neues Minimalmedium entwickelt und erprobt. Diverse Minimalmedien für *Thermus* sind bereits seit längerem bekannt, waren in diesem Fall jedoch nicht unmittelbar zielführend^{86,119-121}. Das Wachstum des Wildtyps war nicht oder nicht selektiv (nur in Anwesenheit von Cellobiose) zu beobachten. Einer Vermutung nach könnten die heute im Handel angebotenen Chemikalien, welche zur Herstellung der Medien benötigt werden, eine für unsere Zwecke zu hohe Reinheit besitzen. Etwaige Verunreinigungen, welche vor Jahrzehnten noch im höheren Grade vorlagen, könnten zur Katalyse unterschiedlichster zellulären Reaktion beigetragen haben. Selbst die Supplementierung mit unterschiedlichen Mengen Vollmedium (TB-Medium) führte lediglich von keinem Wachstum zu undifferenziertem Wachstum in An- und Abwesenheit einer Kohlenstoffquelle (Daten nicht dargestellt).

Reproduzierbar differenziertes Wachstum in Flüssigmedium, welches Abhängigkeit von der Zugabe eines verwertbaren Disaccharids zeigte, wurde mit einem auf Medium nach CASTENHOLZ⁸⁶ basierenden Minimalmedium unter Zugabe von L-Serin (im Folgenden wird auf die Angabe zur Chiralität verzichtet) erreicht. Das hierbei entwickelte Medium erhielt die

Diskussion

Bezeichnung TMB (*Thermus minimal broth*). Inspiration für die Serin-Zugabe war ein abgewandeltes Medium für *Thermus* nach KOYAMA, welchem nach dem Autoklavieren diverse Aminosäuren zugesetzt werden⁴⁹. Im Vergleich der Supplementation mit verschiedenen Aminosäuren (Lysin, Prolin, Methionin und Serin) zeigte sich, dass alleinig Serin das Wachstum in Minimalmedium begünstigte (Abbildung 6). In Abwesenheit von Prolin und Serin war kein Wachstum möglich. Außerdem scheint es, als würden Lysin und Methionin in Kombination, zumindest bei den hier eingesetzten Konzentrationen, das Wachstum von *T. thermophilus* unterbinden. Die gezielte, vereinzelt Erprobung der Zugabe von Prolin und Serin zeigte, dass nur im Falle von Serin differenzierbares Wachstum, abhängig vom Zusatz von Saccharose, zu beobachten war. In nachfolgenden Tests erzielte eine Konzentration von 250 mg L⁻¹ Serin bei unterschiedlichen Kohlenstoffquellen (D-Glucose, D-Mannose, D-Saccharose, D-Cellobiose und D-Maltose) die durchschnittlich höchste Zelldichte (Daten nicht dargestellt). Auf Basis der annotierten Genomsequenz von *T. thermophilus* kann abgeleitet werden, dass Stoffwechselwege sowohl für die Biosynthese als auch für den Abbau von Serin vorhanden sind⁵¹. Demnach sollte es sich bei Serin um keine für *T. thermophilus* essenzielle Aminosäure handeln. Eine mögliche Erklärung zur Abhängigkeit des Wachstums von Serin stellt eine Mutation im Serin-Biosyntheseweg dar, welche somit mit einer Serin-Auxothrophie einhergeht. Alternativ kann das beobachtete Verhalten dadurch erklärt werden, dass Serin aus 3-Phosphoglycerat hervorgeht – einem Stoffwechselprodukt des EMBDEN-MEYERHOF-PARNAS-Wegs¹²². In Minimalmedium (mit D-Cellobiose als Kohlenstoffquelle) steht dieser zentrale Stoffwechselweg im Zentrum des Abbaus von D-Glucose und ist damit wesentlich für die Energiegewinnung. Die Biosynthese von Serin könnte somit zu großen Mengen 3-Phosphoglycerat aus der Primärenergiegewinnung abziehen, wodurch das Wachstum zum Erliegen kommt. Dieser Effekt wird zusätzlich durch die Biosynthese von Glycin und (zum Teil) Cystein verstärkt, da diese aus Serin hervorgehen. Der Zusatz von Serin im Minimalmedium scheint diesen Stoffwechselengpässen entgegenzuwirken.

In Kombination mit Agar (KOBE I), wurde kein differenzierbares Wachstum in Abhängigkeit von der Zugabe eines verwertbaren Disaccharids (wie etwa D-Cellobiose) erreicht. Selbst ohne

Diskussion

zusätzliche Kohlenstoffquelle war Wachstum zu beobachten (Daten nicht dargestellt). Das Wachstum geht dabei vermutlich auf niedermolekulare Zucker zurück, welche in diesem weniger reinen Agar vorhanden sind. Nach einem Wechsel hin zu bakteriologischem Agar (höherer Reinheitsgrad) bzw. Gelrite™ (Polysaccharid: Tetrasaccharidabfolge aus Rhamnose-, Glucuronsäure- und zwei Glucose-Einheiten) um das Medium zu verfestigen, blieb das Wachstum wieder aus. Die Supplementierung mit TB-Medium bzw. einer Ammoniumquelle (NH_4NO_3) in unterschiedlichen Konzentrationen bewirkten nicht den gewünschten Effekt. Aus diesem Grund erfolgten alle Folgeexperimente in Flüssigmedium.

4.1.2 Herstellung einer für die komplementationsbasierte Selektion geeigneten GH-Deletionsmutante

Für die komplementationsbasierte Selektion stellt das oben beschriebene Minimalmedium nur eine der benötigten Grundlagen dar. Zusätzlich wird ein entsprechender GH-Deletions-Wirtsstamm benötigt, der durch die Einführung rekombinanter Zielgene komplementierbar ist. Der Stoffwechsel von *T. thermophilus* ist überwiegend auf die Verwertung von Proteinen bzw. Peptiden ausgelegt, wodurch die Unterbrechung der Verwertung von Oligosacchariden mit nur wenigen Deletionen im Genom umgesetzt werden kann⁵¹. Dies bedeutet andererseits jedoch auch, dass die Entwicklung von *T. thermophilus* Wirtsstämmen für komplementationsbasierte Selektion nach Genen für Oligosaccharid-spaltende Enzyme ohne großen Aufwand nur für ein eingeschränktes Spektrum von Oligosaccharid-Substraten möglich ist.

Ausgehend von der Deletionsmutante *T. thermophilus* $\Delta 42$ ($\Delta\text{TTP0042}$: Deletion eines Gens für β -Glucosidase (GH1) auf dem Megaplasmid; im Folgenden als $\Delta 42$ bezeichnet), welche aus einer früheren Arbeit hervorging, wurde als Ausgangspunkt gewählt⁹¹. Obwohl es sich bei TTP0042 um das einzige im Genom bekannte β -Glucosidasegen handelte, war im Minimalmedium mit D-Cellobiose (im Folgenden ohne Angabe zur Chiralität) Wachstum zu beobachten. Somit mussten eine oder mehrere Cellobiose-hydrolysierende Enzyme vorliegen, die eine (möglicherweise nur geringe) Aktivität für dieses Substrat aufweisen, deren Gene aber nicht als β -Glucosidasegene annotiert sind. Auf dem Megaplasmid wurden drei weitere GH-

Diskussion

Gene gefunden, welche günstigerweise gruppiert vorliegen. Es handelt sich hierbei um TTP0220, TTP0221 und TTP0222, deren Genprodukte für Enzyme der GH-Familien GH2, GH31 bzw. GH2 kodieren (GH2: β -Galactosidase, GH31: α -Glucosidase). Die Arbeit von FUJITA *et al.* zeigte, dass eine Deletion aller vier Gene möglich ist und dadurch die β -Galactosidaseaktivität vollständig unterbunden werden kann¹²³. In der vorliegenden Arbeit wurden ebenfalls alle genannten GH-Gene deletiert, indem in Stamm Δ 42 eine Deletion der drei anderen Gene erfolgte (Abbildung 4 und Abbildung 28)¹²³.

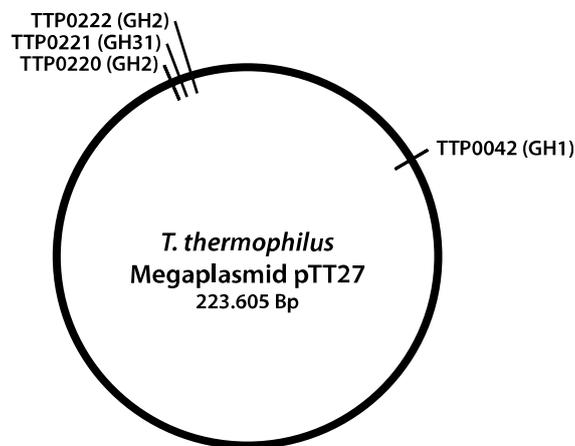


Abbildung 28: *T. thermophilus* Megaplasmid pTT27. Die Positionen der zusätzlich zu TPP0042 in dieser Arbeit deletierten GH-Gene sind gekennzeichnet.

Die Deletion erfolgte durch den Austausch aller drei Gene in einem Schritt durch ein Antibiotikumresistenzgen in Stamm Δ 42. Dazu wurde durch GIBSON-Assemblierung ein lineares DNA-Fragment aus einer Bleomycinresistenzkassette (A2) mit rund 1.500 Bp der flankierenden Genombereiche zu TPP0220 und TPP0222 vorbereitet und durch natürliche Transformation und homologer Rekombination anstelle der GH-Gene in das Megaplasmid von Stamm Δ 42 integriert. Durch die eingeführte Antibiotikumresistenzkassette konnte direkt auf entsprechende bleomycinresistente Genaustauschmutanten selektiert werden. Es wurde ein Bleomycinresistenz- und nicht der alternativ verfügbare Kanamycinresistenzmarker verwendet, da die Selektion mit Kanamycin den Klonierungen mit dem Shuttlevektor pMK18 vorbehalten bleiben musste.

Diskussion

Nach der Bestätigung des Genotyps mittels PCR wurde das Wachstum dieser vierfachen Deletionsmutante (im weiteren als Δ GH bezeichnet) in Minimalmedium untersucht¹⁰⁰. Wider Erwarten konnte im Falle von Δ GH in Minimalmedium (TMB) mit Cellobiose nach einigen Tagen Wachstum festgestellt werden. Die Überprüfung dieser angewachsenen Kulturen durch PCR ergab den Genotyp des Wildtyps¹⁰⁰. Die Erklärung dieses Phänomens liegt vermutlich in der Polyploidie (Vorliegen multipler Kopien des Genoms in den Zellen) von *T. thermophilus* begründet. Prinzipiell sind alle Mikroorganismen zu einem gewissen Grad polyploid und weisen besonders bei raschem Wachstum eine höhere Kopiezahl des Genoms auf^{124–126}. Bei *T. thermophilus* wurde eine Chromosomenkopiezahl von etwa 4-5 Kopien pro Zelle beschrieben¹²⁵. Im Zuge der Deletion durch Transformation werden sehr wahrscheinlich nicht in allen Genomkopien gleichzeitig alle parallel vorhandenen Ziel-Genloci durch homologe Rekombination ausgetauscht. Für die Erzielung der Bleomycinresistenz, auf welche selektiert wurde, genügt möglicherweise schon eine geringe Kopiezahl des Resistenzmarkers im Genom. Der PCR-Nachweis für den Genaustausch wird andererseits auch dann bereits positiv ausfallen, wenn nur eine Kopie des Resistenzmarkers im Genom vorhanden ist. Offenbar waren in den oben beschriebenen Stämmen, die zunächst isoliert wurden, zusätzlich zum erwünschten Genaustausch durch den Bleomycinresistenzmarker auch noch Wildtyp-Allele ohne Genaustausch vorhanden. Bei der Kultivierung in TMB mit Cellobiose wurde dann ein Selektionsdruck in Richtung des Wildtyps ausgeübt, der dann nach einigen Tagen zu Wachstum führte. Im Folgenden war auf Agarplatten mit erhöhter X-Gal-Konzentration tatsächlich bei diesen Stämmen eine grün-bläuliche Färbung der Kolonien zu beobachten.

Für die zuletzt deletierten Gene (TTP0220–TTP0222) stellt zwar die Bleomycinresistenz eine Möglichkeit dar, das Wiederauftreten der Wildtypgene zu unterdrücken, doch für TTP0042 gestaltete sich die markerlose Deletion nicht so einfach. Zusätzlich zur Selektion auf Bleomycin wurde ebenfalls mit hohen X-Gal Konzentrationen gegenselektiert. Frühere Arbeiten in unserer Gruppe zeigten, dass X-Gal nicht nur zur klassischen Verwendung der indirekten Selektion β -Galactosidase-aktiver Klone anhand einer Farbnachweisreaktion verwendet werden kann. Eingesetzt in hohen Konzentrationen ($250\text{--}500\ \mu\text{g L}^{-1}$), wirkt das

Diskussion

Spaltungsprodukt (durch β -Galactosidaseaktivität) toxisch auf Zellen^{127,128}. Das Wachstum wird konzentrationsabhängig eingeschränkt oder komplett unterbunden. Somit kann neben der Selektion auf Bleomycinresistenz durch X-Gal gegen das Vorliegen des Wildtypallels für TPP0042 (GH1; das davon kodierte Enzym besitzt auch β -Galactosidase-Aktivität) und gleichzeitig gegen den Wildtypzustand am Locus TPP0220–TPP0222 selektiert werden.

Durch wiederholtes Ausstreichen einer Einzelkolonie auf Agarplatten (Vollmedium mit X-Gal: 250 $\mu\text{g mL}^{-1}$ und Bleomycin: 12,5 $\mu\text{g mL}^{-1}$) konnte letztlich ausreichend Selektionsdruck erzeugt werden, um Klone mit dem gewünschten (homozygoten) Genotyp an den mutierten GH-Genloci zu erhalten. Während die ersten Ausstriche eine Inkubation von über 2 Tagen benötigten und die Kolonien eine grün-bläuliche Färbung aufwiesen, wuchsen nach einigen Wiederholungsausstrichen die Kulturen vollständig über Nacht sowie farblos heran. Das wiederholte Ausstreichen bewirkte somit ein schrittweises Verdünnen des Wildtyp-Genotyps unter Selektionsdruck, wodurch die toxische Wirkung des hochdosierten X-Gals ausblieb. Die Wachstumsgeschwindigkeit und Koloniefärbung waren vom Wildtyp auf einfachem Vollmedium nicht zu unterscheiden. Der so erhaltene Stamm konnte anschließend für die Analyse des Wachstumsverhaltens in TMB und für die Komplementation des Cellobiose-negativen Phänotyps durch Einführung eines geeigneten GH-Gens zur Validierung des Gesamtsystems eingesetzt werden.

4.1.3 Validierung des Multideletionsstamms ΔGH für die komplementationsbasierte Selektion

Bei der in Kapitel 4.1.2 beschriebene Herstellung der vierfach-Deletionsmutante ΔGH , konnte letztlich ausreichend Selektionsdruck erzeugt werden, um einen Bakterienstamm zu erhalten, welche wie erhofft in Minimalmedium mit Cellobiose selbst nach vier Tagen kein Wachstum zeigte (Abbildung 7B). Dies bestätigt die Vermutung, dass bei der ΔGH -Stammherstellung ursprünglich (ohne wiederholtes Kultivieren mit X-Gal und Bleomycin) zunächst eine heterozygote Mutante entstand, bei der während einer Inkubation auf Medium mit Cellobiose sich der Genotyp des Wildtyps wieder durchsetzen konnte (Kapitel 4.1.2). Die in der vorliegenden Arbeit aufgetretene Problematik bei der Generierung eines homozygoten

Diskussion

Genotyps des Δ GH-Stamms verdeutlicht die Wichtigkeit der genauen Überprüfung des Genotyps bei der Einfügung von Genommodifikationen in einem polyploiden Bakterium wie *T. thermophilus*.

Für die Komplementation des zuvor erhaltenen homozygoten Deletionsstamms, durch Einführung und Expression eines geeigneten GH-Gens und so den grundsätzlichen Funktionsnachweis des Systems zu erbringen, wurde das Gen für die β -Glucosidase CglT (GH1) aus *Thermoanaerobacter brockii* ausgewählt⁹³⁻⁹⁵. Die Temperatur für maximale Umsatzrate liegt für CglT bei 75 °C (pH 5,5) und dessen Länge von 449 Aminosäuren sind ähnlich den Eigenschaften des *T. thermophilus* eigenen Enzyms Bgl (431 AS). CglT sollte sich daher eignen, um nach rekombinanter Expression den Phänotyp des Stamms Δ GH zu komplementieren. An unserem Lehrstuhl wurde CglT durch den Austausch von drei Aminosäuren weiter thermostabilisiert sowie dessen Aktivität erhöht; das so erhaltene Enzym wird im Folgenden CglT KIR genannt⁹⁶. Durch einen Vergleich der Expression der Gene von CglT Wildtyp und der Mutante CglT KIR sollte es des Weiteren möglich sein, zu untersuchen, ob bei der Komplementation die unterschiedliche Thermostabilität sowie Aktivität dieser beiden Enzyme auch *in vivo* phänotypisch diskriminierbar sind.

Nachdem die Gene für beide Enzymvarianten in den Shuttlevektor pMK18 kloniert und der *T. thermophilus*-Stamm Δ GH mit diesen Plasmiden transformiert worden war, erfolgte die Kultivierung in TMB mit Cellobiose.

Wie zu erwarten, war ohne Zusatz von Disacchariden kein Wachstum zu beobachten, wohingegen alle Positivkontrollen mit 0,5 % Maltose (Maltose führte im Vergleich zu anderen Kohlenstoffquellen in Vorexperimenten zur höchste Zelldichte) Wachstum aufwiesen (Abbildung 7). Der Wildtyp wuchs in Anwesenheit von Cellobiose, wohingegen bei Δ GH – wie gewünscht – kein Anstieg der optischen Dichte (bei 600 nm) zu beobachten war. Generell lag die finale optische Dichte aller Kulturen, welche Wachstum aufwiesen, nach vier Tagen zwischen 0,4 und 0,8.

Diskussion

Abweichend von den restlichen Proben war beim Δ GH Deletionsstamm mit pMK18 ohne Insert ein um etwa 24 h verzögertes Wachstum mit Maltose zu beobachten (Abbildung 7C). Vermutlich begründet sich dies durch einen unterschiedlichen physiologischen Zustand der Vorkulturen der untersuchten Stämme. Sollte sich die Vorkultur von Δ GH pMK18 länger als die anderen Bakterienstämme in der stationären Phase befunden haben, so wäre dies eine Erklärung für die verlängerte Anlaufphase (lag-Phase) bei der Kultivierung in Minimalmedium.

Das Einbringen des Gens für Bgl auf pMK18-*bgl* in den vierfach-Deletionsstamm führte zur Komplementation des Cellobiose-negativen Phänotyps und raschem Anwachsen von Δ GH pMK18-*bgl* in cellobiosehaltigem TMB-Medium (Abbildung 7D). Im Falle von Δ GH pMK18-*cglT* (Wildtyp) war ebenfalls Komplementation zu beobachten (Abbildung 7E), allerdings erreichten die Zellen etwa 48 h später ihre maximale optische Dichte. Die lässt sich – wie schon weiter oben bei Δ GH pMK18 beschrieben – durch einen unterschiedlichen physiologischen Zustand der Vorkultur erklären. Der Stamm Δ GH pMK18-*cglT* KIR zeigte eine vergleichbar schnelle Wachstumskinetik wie Δ GH pMK18-*bgl* (Abbildung 7F). Zwar war die maximale Zelldichte nicht so hoch wie in den meisten anderen Proben (es handelte sich um eine andere TMB-Charge), zeigte jedoch vergleichend mit der Positivkontrolle Maltose ein etwas schnelleres Anwachsen als der Bgl-bildende Stamm.

Zusätzlich zu den Wachstumskinetiken wurden *pNP*-Assays mit *pNPG* zur Erfassung der temperaturabhängigen Aktivität von Bgl, CglT und CglT KIR in einem Temperaturbereich von 40-90 °C durchgeführt (Abbildung 8). Anhand der Aktivitäten sollte bestimmt werden, ob ein Zusammenhang zwischen Wachstumskinetik und spezifischen Enzymaktivität besteht. Sowohl Bgl und CglT zeigten hierbei nur rund die Hälfte der Aktivität von CglT KIR bei 60 °C. In den Wachstumskinetiken sieht man sowohl bei den Stämmen mit Bgl und CglT ein leicht verzögertes Anwachsen mit Cellobiose im Vergleich zu den Positivkontrollen (Abbildung 8D und Abbildung 8E). Dieser zeitliche Versatz ist beim Stamm mit CglT KIR nicht vorhanden (Abbildung 8F). Ob tatsächlich dieser Unterschied in der Enzymaktivität von Bgl und CglT zu

Diskussion

CglT KIR zu einer leicht verlängerten lag-Phase führt, ist zum aktuellen Zeitpunkt nicht eindeutig zu beantworten und benötigt weitere Untersuchung.

Nachdem mit Hilfe von CglT und CglT KIR der Grundsatzbeweis zur funktionierenden, komplementationsbasierten Selektion erbracht wurde, ist für die Zukunft noch der Nachweis der Anwendung des Systems bei der Durchmusterung von Metagenom- und/oder Mutantenbibliotheken zu erbringen.

4.1.4 Über die Anwendung komplementationsbasierter Selektion

Eine der möglichen Anwendungen komplementationsbasierter Selektion ist die selektive Anreicherung neuartiger GHs aus Metagenombibliotheken. Übliche Durchmusterungsverfahren erfordern die Analyse von mehreren Tausend bis mehreren Hunderttausend Klonen^{129,130}. Mithilfe von komplementationsbasierter Selektion ist theoretisch das Auffinden eines Gens für ein neuartiges Enzym in nur einem Schritt und zudem in einem Zeitraum von nur einigen Tagen möglich. Da durch Ausnutzung des Wachstumsphänotyps des Wirts – im vorliegenden Fall der Wachstumsdefekt einer β -Glycosidasemutante auf der C-Quelle Cellobiose – nur solche rekombinanten Klone auf Cellobiose-Minimalmedium wachsen können, die ein β -Glucosidasegen tragen und funktionell exprimieren. Es handelt sich also um ein positives Selektionsverfahren.

Die Anwendung dieses System reicht womöglich über die Analyse von Metagenombanken aus Umweltproben hinaus. Denkbar ist der Einsatz für die Thermostabilisierung von Glycosidhydrolasen oder für die Änderung bzw. Erweiterung der Substratspezifität dieser. Unterschiedliche Mutageneseverfahren wie z.B. fehlererzeugender PCR (*engl.*: Error-Prone PCR) oder *DNA-shuffling* könnten verwendet werden, um eine Mutationsbibliothek eines ausgewählten Enzym-kodierenden Gens zu generieren^{131,132}. Durch die hohe Temperatur während der Kultivierung des Wirtsbakteriums *T. thermophilus* kann auf thermostabilere Varianten eines mesophilen oder moderat thermophilen Enzyms selektiert werden. Auch die Suche nach mutierten Enzymen mit veränderter Substratspezifität ist so denkbar, sofern ein entsprechendes Substrat eingesetzt wird.

Diskussion

Für das Auffinden von Aminosäuresequenzpositionen, welche ggf. erfolgsversprechend mutagenisiert werden können, können multiple Sequenz-Alignierungen (*engl.*: multiple sequence alignment) durchgeführt werden^{133,134}. Ein Nachteil dieser Methode ist jedoch, dass *a priori* keine Kenntnis über die Relation der favorisierten Aminosäureaustausche zueinander besteht. Manche Mutationen können zwar im Einzelnen vorteilhaft sein, reduzieren jedoch evtl. in Kombination den gewünschten Effekt. Mit einer ungerichteten und genweiten Mutagenese kann diesem Problem entgegengewirkt werden, da bei einer hochdiversen Mutantenbibliothek ein wesentlich größerer Sequenzraum abgedeckt werden kann. Unvorteilhafte Sequenzkombinationen gehen automatisch im Zuge der Zellkultivierung verloren. Andererseits hat die ungerichtete Mutagenese den Nachteil, dass ggf. viel mehr Mutanten in die Selektion einbezogen werden müssen, um den gewünschten Phänotyp zu erhalten. Entsprechend hoch muss dann auch die Transformationseffizienz bei der Einführung mutierter Genvarianten in den Wirtsorganismus sein.

4.2 Untersuchung der Proteinstabilität unterschiedlicher Mikroorganismen

Teile der nachfolgenden Diskussion basieren auf Ergebnissen einer zur Veröffentlichung eingereichten Publikation: JARZAB, A. *et al.*, Meltome Atlas – thermal proteome stability across the tree of life, *Nature Methods* (2019)

Betrachtet man die Anzahl detektierter Proteine, für welche Schmelzkurven bestimmt werden konnten, so liegt die Proteomabdeckung in einem Bereich von etwa 35–40 % (Abbildung 9 und Tabelle 19). Die Ausnahme bildet *P. torridus*, bei dem die Proteomabdeckung knapp 60 % erreichte. Vermutlich begründet sich dies durch das verhältnismäßig kleine Proteom dieses Archaeums, da für eine autonom lebende Prokaryontenzelle eine gewisse Grundausstattung an Proteinen exprimiert werden muss (ähnliche Anzahl wie bei *G. stearothermophilus* und *T. thermophilus*), was wiederum bei einer kleineren Genomgröße zu einer größeren Abdeckung führt.

Diskussion

Bezüglich aller detektierten Proteine (mit und ohne Schmelzkurven), weicht *T. thermophilus* neben *P. torridus* (letzterer aus den eben genannten Gründen) von den restlichen Organismen ab. Während generell etwa 45–50 % des Gesamtproteoms detektiert werden, wurde bei *T. thermophilus* mit 70 % ein deutlich größerer Umfang erreicht. Dieses Phänomen scheint entweder biologisch oder methodisch begründet zu sein. Eine Analyse der detektierten Proteine ohne T_m -Werte, anhand einer Anreicherung von Gene Ontology-Begriffen (*engl.*: Gene Ontology Term Enrichment), ergab keine signifikanten Treffer.

Die T_m -Werte der untersuchten Organismen waren überwiegend symmetrisch verteilt (Abbildung 10)¹⁰². Lediglich *T. thermophilus* wich von diesem Verhalten ab und zeigte eine stark linksschiefe Verteilung und einige wenige Proteine in den oberen 5 °C des Temperaturbereichs. Diese charakteristische Verteilung scheint durch einige cytoplasmatische Proteine bedingt, welche in Zellextrakt destabilisiert werden (Abbildung 22 und weitere Ausführung in Kapitel 4.2.6). Bei intakten Zellen von *T. thermophilus* zeigte sich eine überwiegend symmetrische Verteilung (Abbildung 19). Auffällig ist bei der T_m -Verteilung zudem die große Differenz bei *O. antarctica* zwischen der optimalen Wachstumstemperatur und dem Schmelzpunkt der thermolabilsten Proteine. Während üblicherweise rund 2 °C Unterschied beobachtet wurden, waren es hier 12 °C. Es wäre anzunehmen gewesen, dass die durch niedrigere Temperaturen erhöhte Rigidität der Proteine – zur Kompensation – eine grundsätzlich höhere Flexibilität voraussetzt. Hier ist jedoch das Gegenteil der Fall. Dies mag sich möglicherweise mit Konservierung von Thermostabilität in den Sequenzen und/oder durch einen evolutiven Vorteil, (kurzzeitig) höhere Umgebungstemperaturen überstehen zu können, begründen.

4.2.1 pI-Werte

Betrachtet man die bimodale Verteilung der theoretischen pI-Werte (anhand der Proteinsequenz nach BJELLQVIST¹⁰³), so lässt sich wie erwartet erkennen, dass kein Organismus ein Protein mit einem pI von 7,5 besitzt (Abbildung 11). Dieser Wert entspricht dem intrazellulären pH; Proteine mit entsprechendem isoelektrischen Punkt hätten sonst *in vivo*

Diskussion

ihre geringste Löslichkeit¹³⁵. Wie bereits in Kapitel 1.2.5 erwähnt, liegt der intrazelluläre pH-Wert von *P. torridus* ungewöhnlicher Weise bei 4,6³⁹. Man erkennt in der pI-Werte-Verteilung, dass auch bei diesem Organismus fast keine Proteine im Bereich des cytoplasmatischen pH-Werts ihren pI haben. Interessant ist jedoch, dass man – wie bei allen anderen Organismen – ebenfalls keine Proteine bei einem pI bei pH 7,5 beobachtet. Es scheint sich hierbei um ein evolutiv konserviertes Merkmal zu handeln.

Größere Abweichungen im Verhältnis von detektierten Proteinen zum Gesamtprotein gab es zwischen den Organismen nicht, weder für die Population unter- noch für die oberhalb eines pH-Werts von 7,5. Generell wurde jedoch ein größerer Abdeckungsgrad für die Proteine mit pI im Bereich unter pH 7,5 erreicht. Bei *P. torridus* werden hierbei die meisten Proteine (relative Anteil) gemessen. Dies war allerdings abzusehen, da bei diesem Organismus auch bereits eine hohe Gesamtproteomabdeckung zu sehen war (Abbildung 9).

4.2.2 Proteinlänge

Eine Möglichkeit thermostabile mit thermolabilen Proteinen innerhalb eines Proteoms zu vergleichen, besteht darin, die unteren und oberen 10 % gegenüberzustellen¹⁰⁵. Dies hat den Vorteil, dass statistische Analysen wie der WILCOXON-MANN-WHITNEY-Test ohne Probleme durchzuführen sind, da diese unter Umständen (wie der zuvor genannte Test) gleichgroße Populationen voraussetzen. Im Folgenden werden die Proteine der unteren 10 % im T_m -Spektrum als thermolabil, die oberen 10 % als thermostabil bezeichnet.

Bei den meisten Organismen wurde ermittelt, dass thermostabile Proteine signifikant kürzer sind als thermolabile (Abbildung 12). Dies ist ein schon in der Vergangenheit beobachteter Effekt^{74,105}. Je länger ein Protein ist, desto höher ist die Wahrscheinlichkeit, dass eine thermisch bedingte Teilentfaltung des Proteins zu weiterer bis hin zu irreversibler Entfaltung führt. dieser Entfalteten Subdomänen sind in der Lage Wechselwirkung mit anderen molekularen Strukturen einzugehen und fördern dadurch die weitere Denaturierung.

Ausnahmen beim Vergleich dieser beiden Gruppen bilden *O. antarctica* und *T. thermophilus*. Beide Organismen zeigten keinen signifikanten Längenunterschied für den oberen und

Diskussion

unteren Temperaturbereich. Im Falle von *T. thermophilus* sind die thermostabilen Proteine sogar im Durchschnitt länger als die Thermolabilen; wenn auch nicht statistisch signifikant.

Relativ auffällig ist zudem der größere Längenumfang der thermolabilen Proteine bei *B. subtilis*. Betrachtet man die generelle Längenverteilung aller Organismen (Anhang: Abbildung 30), so sieht man keine größeren Unterschiede zwischen den Proteinlängenverteilungen aller Organismen. Dementsprechend handelt es sich um eine Umverteilung längerer Proteine von thermisch intermediären Proteinen hin zu niedrigeren T_m -Werten; ein Trend, welcher uns später erneut begegnen wird.

Alternativ kann ein Zusammenhang zwischen der Schmelztemperatur und Länge eines Proteins anhand der PEARSON-Korrelation beschrieben werden (Tabelle 21). Anders als beim alleinigen Vergleich von thermolabilen und thermostabilen Proteinen werden dabei auch thermisch intermediäre Proteine mit einbezogen. Die berechneten PEARSON-Korrelationskoeffizienten bestätigen den zuvor beobachteten, negativen Zusammenhang zwischen T_m -Wert und Proteinlänge. Nun, nach Einbezug aller Datenpunkte, zeigt sich eine signifikante Abnahme der Proteinstabilität mit zunehmender Länge auch für *O. antarctica*. Für *T. thermophilus* ergab sich dagegen, wie bereits zuvor, kein signifikanter Unterschied. Möglicherweise ist hier die erhöhte Thermostabilität längerer Proteine Chaperonen zu verdanken oder dies ist in ihrer Aminosäuresequenz begründet^{136,137}.

4.2.3 Hydrophobizität

Vergleicht man, wie zuvor, thermolabile und thermostabile Proteine, jedoch unter dem Gesichtspunkt von Hydrophobizität, so lässt sich kein signifikanter Unterschied feststellen (Abbildung 13). Die Ausnahme bildet *B. subtilis*: Hier sieht man, wie schon bei der Proteinlänge, dass die thermolabile Fraktion einen größeren Wertebereich aufweist.

Betrachtet man die generelle Hydrophobizitätsverteilung der Proteine (Anhang: Abbildung 31), so unterscheiden sich die aus den Proteomen ermittelten Kurven zwischen den Organismen nicht sichtlich. Auffällig ist die Bimodalität aller Verteilungen. Aufgrund der abweichenden Verteilung von *B. subtilis* wurde die zweite, kleinere Population hydrophoberer

Diskussion

Proteine mit einer Hydrophobizität $\geq 0,5$ gesondert betrachtet. Eine Anreicherung zugehöriger *Gene Ontology*-Begriffen ergab für alle Organismen signifikante Treffer für membranassoziierte Proteine (Anhang: Tabelle 32 bis Tabelle 37; technische Details zur Analyse finden sich in Kapitel 3.2.8). Dies ist nicht überraschend, da Proteine, welche membranverankert sind oder mit der Membran interagieren, hierfür hydrophober sein müssen. Die Ausnahme bildet *G. stearothermophilus* ohne signifikante Treffer. Dies ist wohl auf die geringe *Gene Ontology*-Annotationsqualität zurückzuführen, da diese automatisch anhand der Proteinsequenz erstellt wurden (eggNOG-Mapper¹⁰⁹).

Betrachtet man die Verteilung der unterschiedlichen Temperaturfraktionen, getrennt nach Hydrophobizität (\geq bzw. $< 0,5$), so sieht man, dass *B. subtilis* von den restlichen Verteilungen der anderen Organismen abweicht. Während die restlichen Organismen die gewählten Einteilung nach untere/obere 10 % und 80 % für durchschnittlich thermisch stabile Proteine größtenteils widerspiegeln, sieht man, dass in der hydrophoben Fraktion von *B. subtilis* eine Umverteilung vom mittleren in den unteren Schmelztemperaturbereich vorliegt (Anhang: Abbildung 32). Der Anteil thermolabiler Proteine liegt für hydrophobe Proteine in einem Bereich von etwa 10–20 %, wohingegen der Anteil von *B. subtilis* bei über 50 % liegt. Dementsprechend scheinen membranassoziierte Proteine in *B. subtilis* weniger stabil zu sein als bei den restlichen Organismen. Eine mögliche Erklärung könnte der Aufbau der Zellwand sein. Da es sich bei *B. subtilis* um ein Gram-positives Bakterium handelt, fehlt diesem die äußere Membran. MATEUS *et al.* zeigten, dass die Thermostabilität vom Inneren der Zelle nach außen hin zunimmt¹³⁸. Die Abwesenheit der äußeren Membran könnte daher das Verhältnis zwischen den unterschiedlich thermischen Fraktionen nach ihrer Lokalisation hin verschieben. Dieses Verhalten zeichnet sich ebenfalls leicht beim ebenso Gram-positiven *G. stearothermophilus* ab. Auch hier zeigen sich etwas mehr membranständige Proteine in der thermolabilen Fraktion (Anhang: Abbildung 32).

Diskussion

4.2.4 Aminosäurefrequenz

Beim Vergleich thermolabiler mit thermostabilen Proteinen zeigten sich Aspartat, Leucin und Arginin hochsignifikant angereichert im Bereich thermolabiler Proteine (Abbildung 14). Hinzu kommen weitere – wenn auch nicht ganz so signifikante – Aminosäuren, welche ebenfalls für thermisch labile Proteine angereichert scheinen.

LEUENBERGER *et al.* kamen bei vergleichenden Analysen mehrerer Proteome unterschiedlicher Organismen (welche *E. coli* und *T. thermophilus* beinhalten) zur Erkenntnis, dass Aspartat in weniger stabilen und Lysin vermehrt in thermostabilen Proteinen vorkommt¹⁰⁵. Betrachtet man die Frequenzwerte der einzelnen Aminosäuren, aufgeschlüsselt nach Organismus (Anhang: Tabelle 38), so sieht man eine starke Variabilität bereits unter den Organismen. Durch diese stark vereinfachte Betrachtung ist es daher fraglich, ob es überhaupt möglich ist, innerhalb eines Proteoms anhand der Aminosäurefrequenz verlässlich eine Aussage bzgl. der Thermostabilität zu treffen. Analysen in der Vergangenheit zeigten, dass anhand der Aminosäuresequenz (jedoch unter Einbezug vieler anderer physikalisch-chemischer Parameter) zuverlässig unterschieden werden konnte, ob es sich um ein meso- oder thermostabiles Protein handelt¹³⁹. Jedoch waren Versuche, eine allgemeingültige Aussage bzgl. genauer Aminosäuren zu treffen, nicht immer übereinstimmend^{140–142}.

Denkbar ist auch, dass die generelle Aminosäurezusammensetzung nicht nur mit der Thermostabilität, sondern ebenfalls mit anderen Faktoren, wie beispielhaft Umweltbedingungen oder dem Organismus eigenen Stoffwechselwegen, in Zusammenhang stehen. Die Position einer Aminosäure im Protein, wie etwa in unterschiedlichen Sekundärstrukturelementen oder auch deren Lage an der Oberfläche oder im Inneren des Proteins, erhöhen zudem die Komplexität, so dass von der Möglichkeit einer allgemeingültigen Aussage nicht auszugehen ist.

Diskussion

4.2.5 Betrachtung von Paralogen und Orthologen

Paraloge sind Gene bzw. Proteine, welche aus einer Genduplikation in der Vergangenheit hervorgingen. Im Gegensatz zum Original unterliegt die Kopie keinem evolutionären Druck mehr und kann sich unabhängig weiterentwickeln. Sie kommen i.d.R. im gleichen Organismus vor und übernehmen oftmals unterschiedliche Funktionen.

Orthologe sind hingegen Gene bzw. Proteine aus unterschiedlichen Organismen, welche auf einen gemeinsamen Vorfahren zurückgehen. Da Orthologe per Definition im Zuge von Speziation entstehen, sind diese i.d.R. älter als Paraloge. Aus diesem Grund haben Paraloge meist eine höhere Sequenzübereinstimmung, da ihre Entstehung zeitlich weniger weit zurückliegt.

Eine vergleichende Betrachtung dieser beiden Gruppen, hinsichtlich ihrer Thermostabilität, sollte Aufschluss über die Vererbbarkeit dieser Eigenschaft geben.

Der signifikante Unterschied der ΔT_m -Werte zwischen Paralogen und randomisierten Paaren anhand des WILCOXON-Tests zeigt, dass ihr Zusammenhang nicht willkürlich ist (Abbildung 15 und Tabelle 25). Ebenfalls zeigt der Vergleich der Sequenzidentitätsverteilungen einen signifikanten Unterschied zwischen gepaarten und randomisierten Paralogen; letztere mit geringerer Sequenzübereinstimmung (Abbildung 18).

Für die gesonderte Betrachtung aller jüngst entstandenen Paraloge, kann die Sequenzidentitätsverteilung von randomisierten Paaren herangezogen werden. Eine gute Abschätzung, wann sich die Paraloge vom Zufall abheben, stellt die obere Grenze des 95 % Konfidenzintervalls (entspricht dem 97,5 %-Perzentil) aller Randomisierten dar und liegt bei 27,0 % Sequenzidentität. Paraloge oberhalb dieses Grenzwertes zeigten eine stark signifikante, negative Korrelation zwischen ΔT_m -Werten und ihrer Sequenzidentität (PEARSON-Korrelationskoeffizient: $-0,104$; p-Wert: $6,99 \times 10^{-5}$). Für die verbleibenden, randomisierten Paare war keine signifikante Korrelation zu verzeichnen (p-Wert: $> 0,05$).

Es zeigt sich somit, dass tatsächlich mit steigender Sequenzidentität sich die Schmelztemperaturen von Proteinen einander annähern, was wiederum für eine

Diskussion

Konservierung von Thermostabilität in den Sequenzen spricht. Zu dieser Schlussfolgerung kamen ebenfalls LEUENBERGER *et al.* mit LiP-MS Experimenten (u.a. mit *E. coli* und *T. thermophilus*)¹⁰⁵. In ihren Analysen zeigten sie im Vergleich von Fraktionen niedriger mit hoher Sequenzidentität (> 60 % im Vergleich zu > 90 % Sequenzidentität) signifikant ähnlichere Schmelzwerte mit zunehmender Sequenzübereinstimmung.

Wie oben methodisch beschrieben, können Orthologe auf die gleiche Weise analysiert werden. Ein Blick auf die Verteilung der ΔT_m -Werte zeigte für Orthologe aus *O. antarctica* generell höhere Werte. Dies begründet sich auf die Normalisierung der Schmelztemperaturen anhand der Wachstumstemperatur für die Berechnung von Differenzen. Während für die meisten Organismen der Schmelzbereich aller Proteine relativ nah an der optimalen Wachstumsbereich beginnt, liegt der von *O. antarctica* deutlich höher (Abbildung 10).

Betrachtet man die Verteilungen kleinster ΔT_m -Werte, so zeigten *B. subtilis* und *G. stearothermophilus* von allen Paarungen im Durchschnitt die ähnlichsten Schmelztemperaturen. Dies ist nicht verwunderlich, da beide der Familie der *Bacillaceae* angehören.

Im Gegensatz zu Paralogen, besteht kein signifikanter Unterschied im Schmelztemperaturunterschied, wenn man Orthologe mit Zufallspaaren vergleicht. Diese Beobachtung machten ebenfalls LEUENBERGER *et al.*¹⁰⁵.

Anhand der Verteilung willkürlich zugeordneter Paare kann, wie schon bei Paralogen, ein Grenzwert errechnet werden, um die Orthologen von Zufallsereignissen abzugrenzen. Wie bereits angesprochen, weichen die Orthologe aus *O. antarctica* von den Restlichen in ihren ΔT_m -Werten ab. Lässt man diese außen vor, um einen Grenzwert zu berechnen, so liegt dieser bei 25,1 % (entspricht dem 97,5 %-Perzentil). Für tatsächliche Orthologe besteht eine Korrelation nach PEARSON von $-0,063$ (p-Wert: $3,85 \times 10^{-7}$). Für randomisierte Paare ist diese mit $-0,143$ zwar stärker, wenn auch mit nur schwacher Signifikanz (p-Wert: $3,50 \times 10^{-2}$).

Zusammenfassend lässt sich sagen, dass besonders bei den Paralogen eine nachweisliche Korrelation zwischen Sequenzübereinstimmung und konvergierenden Schmelzwerten

Diskussion

vorhanden ist. Berücksichtigt man, dass der Austausch bereits einer oder weniger Aminosäuren zu starker Thermo(de)stabilisierung führen kann¹⁴³⁻¹⁴⁶, ist es beachtlich, dass dieser Effekt in diesem Umfang anhand eines globalen Sequenzabgleiches hier nachgewiesen werden konnte.

Die im Falle der Orthologe beobachtete, sehr schwache, Korrelation lässt sich vermutlich mit der zeitlich weiter zurückliegenden Entstehung erklären. Dies spricht dafür, dass sich einige wenige Schlüsselpositionen in der Primärsequenz maßgeblich auf die Thermostabilität auswirken.

Die hier untersuchten Organismen umspannen in ihrer Wachstumstemperatur einen insgesamt sehr großen Bereich. Hierbei zeigen ihre Orthologe große Unterschiede in der Schmelztemperatur, während die Sequenzübereinstimmung stark der von Paralogen ähnelt. Da sich die Sequenzen global für Para- und Orthologe überwiegend gleichen, muss der Unterschied, welcher zu Thermostabilisierung führt, im Detail und somit in der Identität einzelner Aminosäuren liegen. Die bereits mehrfach angesprochene, zeitliche Komponente könnte dafür sorgen, dass die wesentlich jüngeren Paraloge an genau diesen „Schlüsselpositionen“ verstärkt übereinstimmen, was mit einem globalen Sequenzabgleich nur gering sichtbar wäre. Zudem stellt, da Paraloge einem einzigen Organismus entstammen, die konstante Umgebungstemperatur geschichtlich gesehen über einen langen Zeitraum einen evolutiven Druck dar, genau diese ausschlaggebenden Aminosäuren auf entsprechenden Schlüsselpositionen zu erhalten.

Die große Spannweite der Wachstumsbereiche wurde erreicht, obwohl sich die Proteine global nicht merklich in ihrer Sequenzidentität von Paralogen abheben (Vergleich von Abbildung 16 mit Abbildung 18). Eine Anpassung an die unterschiedlichen, weit auseinanderliegenden Temperaturbereiche, scheint durch die Änderung einiger weniger Aminosäuren im Laufe der Evolution erfolgt zu sein.

Es lässt sich somit aus dieser Untersuchung ableiten, dass Thermostabilität evolutiv konserviert ist und diese Eigenschaft von einigen ausschlaggebenden Aminosäurepositionen bestimmt

Diskussion

wird. Diese Erkenntnis ist z.B. für die Mutagenese zur Thermostabilisierung von Enzymen von Bedeutung. Es kann generell davon ausgegangen werden, dass ein großes Spektrum von Proteinen – weitestgehend unabhängig von seiner Primärsequenz – thermostabilisiert werden kann.

4.2.6 Gegenüberstellung von intakten Zellen und Zellextrakten

Vergleichende TPP Experimente – kurz CETSA (*engl.*: cellular thermal shift assay) – finden aktuell immer mehr Anwendung in der Wirkstoffentwicklung^{79,81}. Durch die Interaktion eines Proteins mit anderen molekularen Strukturen ändert sich dessen Schmelztemperatur nachweislich. Vergleicht man behandelte und unbehandelte Zellen, so lässt eine Verschiebung der Schmelztemperatur auf eine Interaktion mit entsprechender Substanz schließen. Eine direkte Interaktion wird überwiegend bei der Verwendung von Zellextrakt beobachtet, wohingegen zusätzlich Auswirkungen entlang von Stoffwechselwegen in intakten Zellen betrachtet werden können⁷⁴. Da im Zuge eines Zellaufschlusses die Proteine weiter von ihren Wechselwirkungspartnern in Lösung entfernt sind, entfallen weitestgehend stabilisierende Effekte. Die beobachteten Unterschiede zwischen Zellextrakt und intakten Zellen kann, da Proteine oft als funktionelle Multiproteinkomplexe vorliegen, unter Umständen recht groß sein¹⁴⁷.

Diese Verschiebungen der Schmelztemperatur sind somit ebenfalls zwischen Zellextrakt und intakten Zellen vorhanden, da *in vivo* zwischen den Proteinen und anderen intrazellulären Strukturen Wechselwirkungen stattfinden. Zur Klärung der Frage, welche Protein(-gruppen) intrazellulär am stärksten wechselwirken, wurden für einen Vergleich ebenfalls intakte *E. coli*- und *T. thermophilus*-Zellen massenspektrometrisch nach TPP analysiert.

Anhand der T_m -Werte-Verteilung (Abbildung 19) sieht man *in vivo*, im Falle beider Organismen, eine Verschiebung hin zu höheren Schmelztemperaturen. Untersuchungen in der Vergangenheit haben gezeigt, dass z.B. Proteine, welche an der Atmungskette beteiligt sind, eine durchschnittlich höhere Schmelztemperatur besitzen¹⁰². Da diese Proteine meist

Diskussion

membranassoziiert sind, liegt es nahe, dass Unterschiede zwischen Zellextrakt und intakten Zellen primär durch ihre intrazelluläre Lokalisation bestimmt werden.

Da sich ein Großteil der gemessenen Proteome unter beiden Bedingungen überlappen (Abbildung 20), ist eine aussagekräftige Gegenüberstellung möglich. Die Gegenüberstellung von T_m -Werten aus Zellextrakt und intakten Zellen zeigt, wenn man die Proteine nach ihrer Lokalisation anhand der Datenbank PSORTdb¹¹¹⁻¹¹³ getrennt betrachtet, eine größere Abweichung für membranassoziierte Proteine (Abbildung 21 und Abbildung 22). Da bei *T. thermophilus* jeweils nur eine Beobachtung für Proteine des Periplasmas und der äußeren Membran vorhanden waren, wurden diese außen vor gelassen.

Während für lösliche Proteine (Cyto- und Periplasma), sowohl für *E. coli* als auch *T. thermophilus*, weitgehend ein linearer Zusammenhang ersichtlich ist, weichen die Membranproteine (Cytoplasmamembran und äußeren Membran) von diesem Verhalten ab und demonstrieren eine höhere Stabilität unter *in vivo* Bedingungen.

Diese Beobachtungen werden ebenfalls durch die Quantifizierung anhand von Korrelationsanalysen bestätigt (Tabelle 26). Für alle löslichen Proteine ergeben sich höhere PEARSON-Korrelationskoeffizienten als bei Membranständigen. Für die äußere Membran von *E. coli* liegt zwar eine vergleichsweise hohe Korrelation vor, doch ist diese nicht signifikant.

Es zeigt sich somit, dass membranassoziierte Proteine am stärksten in ihrer Schmelztemperatur durch das Auflösen zellulärer Strukturen beeinflusst werden. Dies ist nicht verwunderlich, da die Membranen – somit auch Membranproteine – im Zuge des Zellaufschlusses, durch Ultraschall bzw. French Press, stark in Mitleidenschaft gezogen werden.

Im Gegensatz zu *E. coli* umfasst das Genom von *T. thermophilus* ein Chromosom (TTC) mit 1,89 MBp (1.924 Proteine) und ein zusätzlich ein Megaplasmid (pTT27) mit 233 kBp (202 Proteine). Dies ermöglicht, neben einer Betrachtung nach intrazellulärer Lokalisation, eine getrennte Unterscheidung nach genomischer Lokalisation.

Das Chromosom spiegelt das zuvor beobachtete Verhalten wider: Ein Großteil der Proteine zeigt sich in intakten Zellen stabiler als in Zellextrakt (Abbildung 25). Abweichend hiervon

Diskussion

beobachtet man für die Proteine des Megaplasמידs in beiden Proben eine vergleichbare Verteilung der T_m -Werte (Abbildung 24 und Abbildung 25).

Insgesamt hat das Megaplasמיד einen Anteil am Gesamtproteom von 9,5 %. Für beide Proben lag dieser Wert im gleichen Bereich (Zellextrakt: 9,2 %, intakte Zellen: 9,8 %); ebenso liegt die Schnittmenge beider mit 10,0 % ebenfalls in diesem Bereich (Abbildung 23). Demnach ist Seitens der Proteinzahl keine Verzerrung zu beobachten. Eine weitere Auflösung nach der intrazellulären Lokalisation, zeigt keinen merklichen Unterschied zum Chromosom auf. Prozentual sind gleichviele Proteine und Cytoplasma und Cytoplasmamembran vorhanden. Vermutlich ergibt sich das erhaltene Bild aus der geringen Grundgesamtheit der Megaplasמידproteine. Wie bereits zuvor zu sehen war, sind nicht zwangsläufig alle Proteine der Cytoplasmamembran in intakten Zellen stabilisiert (Abbildung 22). Anhand vorhandener Stoffwechselwege sind keine offensichtlichen Unterschiede zwischen Chromosom und Megaplasמיד zu beobachten. Es ist davon auszugehen, dass die beobachtete Verteilung der Megaplasמידproteine auf ihre geringe Anzahl zurückzuführen ist.

4.2.7 Thermische Zellinaktivierungskinetiken und DSF

Proteomweite Schmelzpunkte sind nicht nur für Anwendungen wie z.B. CETSA, sondern ebenfalls für die Klärung biologischer Zusammenhänge von Interesse. Die thermisch induzierte, zelluläre Inaktivierung scheint in direktem Zusammenhang mit Proteindenaturierung zu stehen^{148,149}. Aus diesem Grund wurden Inaktivierungsprofile für *E. coli* und *T. thermophilus* angefertigt. Zellen aus Vollmedium wurden für 10 min bei unterschiedlichen Temperaturen inkubiert und die Überlebensrate anhand der Keimzahl auf Festmedium bestimmt (Kapitel 3.2.8). Komplementierend sollte dynamische Differenzfluorimetrie (*engl.*: DSF – differential scanning fluorimetry) indirekt Auskunft über die Exposition hydrophober Strukturen entlang eines Temperaturgradienten für Zellextrakt und intakte Zellen geben (Kapitel 3.2.9).

Aus den Inaktivierungskurven geht hervor, dass die Temperatur 50 %er Zellinaktivierung nahe der durch die Literatur beschriebene, maximalen Wachstumstemperatur liegt. Für *E. coli* liegt

Diskussion

dieser Wert bei 47,5 °C (maximale Wachstumstemperatur: 47,4 °C) und bei *T. thermophilus* bei 83,4 °C (maximale Wachstumstemperatur: 84 °C). Zur Klärung der Frage, welche Proteine in diesem Bereich denaturieren und welche (essenzielle) Proteingruppe für die thermische Inaktivierung Verantwortung tragen, wurde eine Anreicherung von Gene Ontology-Begriffen (GO-Begriffe) mit topGO durchgeführt⁹⁰. Als Temperaturbereich wurde ein Intervall der 90% bis 10%iger Überlebensrate entsprechend herangezogen.

Sowohl für *E. coli* als auch für *T. thermophilus* zeigten sich Proteine signifikant erhöht, welche an Nukleotidbindung beteiligt sind. Zusätzlich zeigen sich bei *E. coli* strukturgebende Elemente von Ribosomen in diesem Temperaturbereich in Mitleidenschaft gezogen. Die geringere Signifikanz bei *T. thermophilus* lässt sich zum einen auf das kleinere Proteom sowie auf die im Vergleich zu *E. coli* weniger umfangreiche GO-Annotierung zurückführen. Dennoch sieht man sehr gut, dass Parallelen zwischen beiden Organismen bestehen.

Einer anderen Theorie zufolge tritt die thermische Inaktivierung bzw. der Zelltod durch die Beschädigung von Membranen ein. Nebst dem Verlust mit Membranen in Verbindung stehender Prozesse, tritt Cytoplasma aus, was in größerem Umfang zur Zellinaktivierung führt.

DSF erlaubt, die indirekte, quantitative Messung hydrophober Wechselwirkungen. Als Fluorophor dient SYPRO Orange, welches durch die Bindung hydrophober Strukturen bei einer Wellenlänge von 525 nm angeregt und bei 575 nm gemessen werden kann. Für intakte Zellen ist somit primär die Beschädigung von Membranen sowie für Zellextrakt überwiegend die Denaturierung von Proteinen nachweisbar. Ein Vergleich von thermischer Zellinaktivierung und DSF kann somit zeigen, ob die Membranen und ein Großteil des Proteoms vor, während oder nach der Inaktivierung in Mitleidenschaft gezogen werden. Dies wird ebenfalls durch andere Studien mit *E. coli* gestützt, welche erhöhte Membranpermeabilität erst bei Temperaturen weit oberhalb der thermischen Zellinaktivierung nachweisen konnten^{150,151}.

Im Falle von *E. coli* liegt der Temperaturwert 50%iger Fluoreszenz, sowohl für Zellextrakt als auch für intakte Zellen, oberhalb der Temperatur 50%iger Zellinaktivierung (Abbildung 26

Diskussion

und Abbildung 27). In der Tat ist zu diesem Zeitpunkt sogar keine (intakte Zellen) oder vernachlässigbar geringe (Zellextrakt) Fluoreszenz nachweisbar. Dies spricht gegen die Theorie, dass die thermische Beschädigung von Membranen für die zelluläre Inaktivierung ursächlich ist. Ebenfalls präsentiert sich der Großteil der Proteine des Zellextrakts hierbei im nativen Zustand. Es kann somit geschlossen werden, dass die beobachtete Inaktivierung nicht aus der Denaturierung der Proteinhauptmasse hervorgeht.

T. thermophilus zeigt hingegen einen früheren Übergang anhand der Fluoreszenzwerte im Vergleich zu den Inaktivierungsversuchen. Zudem beobachtet man für intakte Zellen ein für niedrige Temperaturen relativ hohe Fluoreszenzwerte, welche mit dem Anstieg der Temperatur abnehmen, um dann wieder das für DSF charakteristische Profil anzunehmen.

Für das atypische Verhalten der intakten Zellen könnten die S-Layer Proteine, welche sich auf der Zelloberfläche befinden, verantwortlich sein^{149,152}. Die Interaktionen des Fluoreszenzfarbstoffes mit diesen Oberflächenstrukturen könnten ursächlich für dieses Artefakt sein. Ebenso müsste laut DSF der Hauptanteil aller Proteine am Temperaturpunkt der zellulären Inaktivierung denaturiert sein. Vergleichen lässt sich diese Beobachtung mit der Proteingehamtsmenge in Abhängigkeit von der Temperatur aus TPP-Experimenten. Eine gute Näherung für diese Gesamtsmenge stellt die Intensität der Reporterionen aus der MS2-Messung der niedrigsten Temperatur dar. Diese Reporterionen spiegeln die Proteinmenge in der thermisch unbehandelten Probe wider (die unterste Temperaturbehandlung entspricht der Wachstumstemperatur des Organismus). Hier zeigt sich, dass bei 50 % Zellinaktivierung 63 % aller Proteine immer noch nativ vorliegen (bei *E. coli* sind es sogar 83 %). Aus diesem Grund scheinen die DSF-Daten für *T. thermophilus* nur bedingt vertrauenswürdig. Bei *E. coli* bestätigen sich das Bild durch diese Herangehensweisen.

Da weder eine Beschädigung der Membranen beobachtet wird und ein Großteil der Proteine noch nicht denaturiert ist, liegt es nahe, dass der Funktionsverlust eines essenziellen Proteins bzw. einer essenziellen Proteingruppe für die zelluläre Inaktivierung als Ursache infrage kommt. Die Anreicherung von GO-Begriffen zeigt, dass bei *E. coli* signifikant mehr DNS- und RNS-bindende Proteine sowie strukturelle Ribosomenbestandteile im Bereich der thermischen

Diskussion

Inaktivierung beobachtet werden (Tabelle 28). Für *T. thermophilus* erzielt die gleiche Analyse überwiegend Proteine, welche in der Bindung von Nukleotiden involviert sind (Tabelle 29). Es zeigt sich somit anhand der GO-Begriffe eine weitgehende Übereinstimmung zwischen *E. coli* und *T. thermophilus*. Auswirkungen durch thermisch bedingte, strukturelle Änderungen der DNS wurden ausgeschlossen, da diese erst oberhalb der zellulären Inaktivierungstemperatur denaturiert¹⁵³.

Zusammenfassend lässt sich sagen, dass die thermische Zelinaktivierung weder auf eine übermäßige Schädigung der äußeren Membran noch auf die Denaturierung des Proteomgroßteils zurückführbar ist. Vielmehr ist der spezifische Funktionsverlust DNS- und RNS-bindender Proteine, sowie die Beeinträchtigung von Ribosomen, ausschlaggebend. Hiervon abgeleitet scheinen Transkription und Translation (bzw. Teilprozesse) *die* essenziellen Prozesse zu sein, welche zu einer zellweiten thermischen Inaktivierung führen.

5 AUSBLICK

Nachdem das System zur komplementationsbasierten Selektion anhand des Wildtyps und einer thermostabileren Variante einer GH1 seine Funktionsweise unter Beweis gestellt hat, sollte auf die Skalierbarkeit des Systems hingearbeitet werden. Da bisher nicht geklärt ist, ob eine vorteilhafte Variante eines Enzyms ebenfalls in einer komplexen Genbank angereichert werden kann, muss mit einer solchen Modellgenbank dieser Nachweis noch erbracht werden. Hierfür könnte einer Genbank großen Umfangs, welche wohl wissend zu keiner Komplementation führen würde, die zuvor verwendete GH1 aus *T. Brockii* zugegeben werden, um erneut Wachstumsexperimente durchzuführen. Ergänzend hierzu könnten Mutageneseexperimente mit fehlererzeugender PCR oder etwa *DNA-shuffling* genutzt werden, um bekannte Enzyme zu thermostabilisieren. Denkbar wäre auch die zusätzlich Spezifitätsänderung durch die Wahl eines anderen Substrats während der Selektion.

Im Falle der Meltomanalysen sollten weitere Organismen untersucht werden. Spannend könnte sich die Betrachtung weiterer Bacillen gestalten. Hier finden sich einige sequenziell stark übereinstimmende Arten, welche z.T. stark abweichende Wachstumstemperaturen aufweisen. Eine Betrachtung dieser könnte unser Verständnis um die Zusammenhänge von Thermostabilität und Proteinsequenz erweitern.

Ein vorstellbares Projekt für die Zukunft könnte die Analyse heterologer Expression mithilfe von TPP sein. Genbanken von Organismen wie *B. subtilis* oder *T. thermophilus* in *E. coli* könnten beleuchten, welche Proteine überhaupt exprimiert und in wie weit diese von der fremden Umgebung in ihrer Stabilität beeinflusst werden. Diese Erkenntnisse könnten in Forschung und Industrie neue Impulse für die Optimierung heterologer Expression geben. Aus diesem Grund wären Genbanken einiger Organismen in *E. coli* sowie Meltomanalysen dieser Organismen in intaktem Zustand Gegenstand eines solchen möglichen Projekts.

LITERATUR- VERZEICHNIS

1. Nguyen, S. T. C., Freund, H. L., Kasanjian, J. & Berlemont, R. Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. *Appl. Microbiol. Biotechnol.* **102**, 1629–1637 (2018).
2. Henrissat, B. *et al.* Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci.* **92**, 7090–7094 (1995).
3. Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316 (1991).
4. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, 490–495 (2014).
5. Grogan, D. W. Extreme Thermophiles. in *eLS* 1–12 (John Wiley & Sons, Ltd, 2013). doi:10.1002/9780470015902.a0000392.pub2.
6. Takai, K. *et al.* Cell proliferation at 122 degrees C and isotopically heavy CH₄ production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10949–54 (2008).
7. Cava, F., Hidalgo, A. & Berenguer, J. *Thermus thermophilus* as biological model. *Extremophiles* **13**, 213–31 (2009).
8. Brock, T. D. & Freeze, H. *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *J. Bacteriol.* **98**, 289–97 (1969).
9. Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335–50 (1987).
10. Yakimov, M. M. *et al.* *Oleispira antarctica* gen. nov., sp. nov., a novel hydrocarbonoclastic marine bacterium isolated from Antarctic coastal sea water. *Int. J. Syst. Evol. Microbiol.* **53**, 779–85 (2003).
11. Kube, M. *et al.* Genome sequence and functional genomic analysis of the oil-degrading bacterium *Oleispira antarctica*. *Nat. Commun.* **4**, 2156 (2013).
12. Guibert, L. M. *et al.* Alkane biodegradation genes from chronically polluted subantarctic coastal sediments and their shifts in response to oil exposure. *Microb. Ecol.* **64**, 605–16 (2012).
13. Gentile, G. *et al.* Biodegradation potentiality of psychrophilic bacterial strain *Oleispira antarctica* RB-8(T). *Mar. Pollut. Bull.* **105**, 125–30 (2016).
14. Hemphill, H. E. & Whiteley, H. R. Bacteriophages of *Bacillus subtilis*. *Bacteriol. Rev.* **39**, 257–315 (1975).
15. Nakamura, L. K., Roberts, M. S. & Cohan, F. M. Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. *Int. J. Syst. Bacteriol.* **49**, 1211–1215 (1999).
16. Earl, A. M., Losick, R. & Kolter, R. Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol.* **16**, 269–75 (2008).

Literaturverzeichnis

17. Kunst, F. *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–56 (1997).
18. Bückner, H. & Horneck, G. The biological effectiveness of HZE-particles of cosmic radiation studied in the Apollo 16 and 17 Biostack experiments. *Acta Astronaut.* **2**, 247–264 (1975).
19. Kubo, Y. *et al.* Phylogenetic analysis of *Bacillus subtilis* strains applicable to natto (fermented soybean) production. *Appl. Environ. Microbiol.* **77**, 6463–6469 (2011).
20. Blount, Z. D. The unexhausted potential of *E. coli*. *Elife* **4**, 1–12 (2015).
21. Nirenberg, M. *et al.* RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proc. Natl. Acad. Sci. U. S. A.* **53**, 1161–8 (1965).
22. Crick, F. H., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* **192**, 1227–32 (1961).
23. Meselson, M. & Stahl, F. W. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **44**, 671–82 (1958).
24. Hayashi, K. *et al.* Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol. Syst. Biol.* **2**, 1453–62 (2006).
25. Riley, M. *et al.* *Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Res.* **34**, 1–9 (2006).
26. Angelov, A., Fu, O., Valerius, O., Braus, G. H. & Liebl, W. Properties of the recombinant glucose/galactose dehydrogenase from the extreme thermoacidophile, *Picrophilus torridus*. **272**, 1054–1062 (2005).
27. Angelov, A. *et al.* A metagenome-derived thermostable β -glucanase with an unusual module architecture which defines the new glycoside hydrolase family GH148. *Sci. Rep.* **7**, 17306 (2017).
28. Donk, P. J. A Highly Resistant Thermophilic Organism. *J. Bacteriol.* **5**, 373–4 (1920).
29. Downey, R. J., Kiskiss, D. F. & Nuner, J. H. Influence of oxygen on development of nitrate respiration in *Bacillus stearothermophilus*. *J. Bacteriol.* **98**, 1056–62 (1969).
30. Nazina, T. N. *et al.* Taxonomic study of aerobic thermophilic bacilli: Descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzonensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus th.* *Int. J. Syst. Evol. Microbiol.* **51**, 433–446 (2001).
31. Watanabe, S., Kirikae, T. & Miyoshi-Akiyama, T. Complete Genome Sequence of *Streptococcus dysgalactiae* subsp. *equisimilis* 167 Carrying Lancefield Group C Antigen and Comparative Genomics of *S. dysgalactiae* subsp. *equisimilis* Strains. *Genome Biol. Evol.* **5**, 1644–1651 (2013).
32. Bhalla, A., Kainth, A. S. & Sani, R. K. Draft Genome Sequence of Lignocellulose-Degrading Thermophilic Bacterium *Geobacillus* sp. Strain WSUCF1. *Genome Announc.* **1**, 2013 (2013).
33. Feng, L. *et al.* Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 5602–7 (2007).
34. Rozali, S. N. M., Milani, E. A., Deed, R. C. & Silva, F. V. M. Bacteria, mould and yeast spore inactivation studies by scanning electron microscope observations. *Int. J. Food Microbiol.* **263**, 17–25 (2017).
35. Benson, D. A. *et al.* GenBank. *Nucleic Acids Res.* **41**, D36–D42 (2012).
36. Bateman, A. *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
37. Fütterer, O. *et al.* Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9091–6 (2004).

Literaturverzeichnis

38. Schleper, C., Pühler, G., Kühlmorgen, B. & Zillig, W. Life at extremely low pH. *Nature* **375**, 741–2 (1995).
39. Van de Vossenberg, J. L. C. M., Driessen, A. J. M., Zillig, W. & Konings, W. N. Bioenergetics and cytoplasmic membrane stability of the extremely acidophilic, thermophilic archaeon *Picrophilus oshimae*. *Extremophiles* **2**, 67–74 (1998).
40. Thürmer, A. *et al.* Proteomic analysis of the extremely thermoacidophilic archaeon *Picrophilus torridus* at pH and temperature values close to its growth limit. *Proteomics* **11**, 4559–68 (2011).
41. Oshima, T. & Imahori, K. Isolation of an extreme thermophile and thermostability of its transfer ribonucleic acid and ribosomes. *J. Gen. Appl. Microbiol.* **17**, 513–517 (1971).
42. Oshima, T. & Imahori, K. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. *Int. J. Syst. Bacteriol.* **24**, 102–112 (1974).
43. Omelchenko, M. V. *et al.* Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans*: divergent routes of adaptation to thermophily and radiation resistance. *BMC Evol. Biol.* **5**, 57 (2005).
44. Yusupov, M. M. *et al.* Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–96 (2001).
45. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–39 (2000).
46. de Grado, M., Castán, P. & Berenguer, J. A high-transformation-efficiency cloning vector for *Thermus thermophilus*. *Plasmid* **42**, 241–5 (1999).
47. Matsushita, K. *et al.* Genomic analyses of thermotolerant microorganisms used for high-temperature fermentations. *Biosci. Biotechnol. Biochem.* **8451**, 1–14 (2016).
48. Friedrich, A., Prust, C., Hartsch, T., Henne, A. & Averhoff, B. Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. *Appl. Environ. Microbiol.* **68**, 745–55 (2002).
49. Koyama, Y., Hoshino, T., Tomizuka, N. & Furukawa, K. Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other *Thermus* spp. *J. Bacteriol.* **166**, 338–40 (1986).
50. Angelov, A., Mientus, M., Liebl, S. & Liebl, W. A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. *Syst. Appl. Microbiol.* **32**, 177–185 (2009).
51. Henne, A. *et al.* The genome sequence of the extreme thermophile *Thermus thermophilus*. *Nat. Biotechnol.* **22**, 547–553 (2004).
52. Castón, J. R., Berenguer, J., de Pedro, M. A. & Carrascosa, J. L. S-layer protein from *Thermus thermophilus* HB8 assembles into porin-like structures. *Mol. Microbiol.* **9**, 65–75 (1993).
53. Cava, F., Pedro, M. A. De, Schwarz, H., Henne, A. & Berenguer, J. Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. **52**, 677–690 (2004).
54. Wang, L., Hoffmann, J., Watzlawick, H. & Altenbuchner, J. Markerless Gene Deletion with Cytosine Deaminase in *Thermus thermophilus* Strain HB27. *Appl. Environ. Microbiol.* **82**, 1249–55 (2015).
55. Cava, F. *et al.* Control of the respiratory metabolism of *Thermus thermophilus* by the nitrate respiration conjugative element NCE. *Mol. Microbiol.* **64**, 630–646 (2007).
56. Fridjonsson, O., Watzlawick, H. & Mattes, R. Thermoadaptation of alpha-galactosidase AgaB1 in *Thermus thermophilus*. *J. Bacteriol.* **184**, 3385–91 (2002).
57. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71 (1989).
58. Iribarne, J. V. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287 (1976).

Literaturverzeichnis

59. Thomson, B. A. & Iribarne, J. V. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* **71**, 4451–4463 (1979).
60. Mack, L. L., Kralik, P., Rheude, A. & Dole, M. Molecular Beams of Macroions. II. *J. Chem. Phys.* **52**, 4977–4986 (1970).
61. Cox, J. *et al.* Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
62. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–72 (2008).
63. McLafferty, F. W. *Tandem Mass Spectrometry*. (Wiley, 1983).
64. Mallick, P. & Kuster, B. Proteomics: A pragmatic perspective. *Nat. Biotechnol.* **28**, 695–709 (2010).
65. Sleno, L. & Volmer, D. A. Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **39**, 1091–1112 (2004).
66. Makarov, A. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Anal. Chem.* **72**, 1156–1162 (2000).
67. Fohlman, J. & Roepstorff, P. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11**, 601 (1984).
68. Shao, C., Zhang, Y. & Sun, W. Statistical characterization of HCD fragmentation patterns of tryptic peptides on an LTQ Orbitrap Velos mass spectrometer. *J. Proteomics* **109**, 26–37 (2014).
69. Ratkowsky, D. A., Olley, J. & Ross, T. Unifying temperature effects on the growth rate of bacteria and the stability of globular proteins. *J. Theor. Biol.* **233**, 351–362 (2005).
70. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
71. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–90 (2006).
72. Johnson, C. M. Differential scanning calorimetry as a tool for protein folding and stability. *Arch. Biochem. Biophys.* **531**, 100–9 (2013).
73. Franken, H. *et al.* Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* **10**, 1567–93 (2015).
74. Savitski, M. M. *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).
75. Werner, T. *et al.* Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal. Chem.* **86**, 3594–601 (2014).
76. Thompson, A. *et al.* Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904 (2003).
77. McAlister, G. C. *et al.* Increasing the Multiplexing Capacity of TMTs Using Reporter Ion Isotopologues with Isobaric Masses. *Anal. Chem.* **84**, 7469–7478 (2012).
78. Werner, T. *et al.* High-Resolution Enabled TMT 8-plexing. *Anal. Chem.* **84**, 7188–7194 (2012).
79. Martinez Molina, D. *et al.* Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84–7 (2013).
80. Reinhard, F. B. M. *et al.* Thermal proteome profiling monitors ligand interactions with cellular membrane proteins. *Nat. Methods* **12**, 1129–1131 (2015).

Literaturverzeichnis

81. Jafari, R. *et al.* The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat. Protoc.* **9**, 2100–2122 (2014).
82. Martinez Molina, D. & Nordlund, P. The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target Engagement and Mechanistic Biomarker Studies. *Annu. Rev. Pharmacol. Toxicol.* **56**, 141–61 (2016).
83. Huber, K. V. M. *et al.* Proteome-wide drug and metabolite interaction mapping by thermal-stability profiling. *Nat. Methods* **12**, 1055–7 (2015).
84. Mateus, A., Määttä, T. A. & Savitski, M. M. Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome Sci.* **15**, 13 (2016).
85. Miettinen, T. P. & Björklund, M. NQO2 Is a Reactive Oxygen Species Generating Off-Target for Acetaminophen. *Mol. Pharm.* **11**, 4395–4404 (2014).
86. Castenholz, R. W. Aggregation in a Thermophilic Oscillatoria. *Nature* **215**, 1285–1286 (1967).
87. Pagès, H., Aboyou, P., Gentleman, R. & DebRoy, S. Biostrings: Efficient manipulation of biological strings. (2018).
88. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A Grammar of Data Manipulation. (2018).
89. Osorio, D., Rondon-Villarreal, P. & Torres, R. Peptides: Calculate Indices and Theoretical Physicochemical Properties of Protein Sequences. (2018).
90. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. (2016).
91. Haijuan, L. Chromosome and Megaplasmid partitioning in *Thermus thermophilus* HB27. (Lehrstuhl für Mikrobiologie, Technische Universität München, 2014).
92. Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* **84**, 54–68 (1972).
93. Breves, R. *et al.* Genes encoding two different beta-glucosidases of *Thermoanaerobacter brockii* are clustered in a common operon. *Appl. Environ. Microbiol.* **63**, 3902–10 (1997).
94. Zeikus, J. G., Hegge, P. W. & Anderson, M. A. *Thermoanaerobium brockii* gen. nov. and sp. nov., a new chemoorganotrophic, caldoactive, anaerobic bacterium. *Arch. Microbiol.* **122**, 41–48 (1979).
95. Lee, Y.-E., Jain, M. K., Lee, C. & Zeikus, J. G. Taxonomic Distinction of Saccharolytic Thermophilic Anaerobes: Description of *Thermoanaerobacterium xylanolyticum* gen. nov., sp. nov., and *Thermoanaerobacterium saccharolyticum* gen. nov., sp. nov.; Reclassification of *Thermoanaerobium brockii*, *Clostridium*. *Int. J. Syst. Bacteriol.* **43**, 41–51 (1993).
96. Zverlov, V. *et al.* Mutant beta-glucosidase variants with increased thermostability (Patent: WO2017063787A1, Lehrstuhl für Mikrobiologie, Technische Universität München). 51 (2017).
97. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–5 (2009).
98. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–8 (1990).
99. Barnes, W. M. Plasmid detection and sizing in single colony lysates. *Science* **195**, 393–4 (1977).
100. Bagdatli, S. T. Identification of novel genes for thermostable β -glucosidases using a β -glucosidase-deficient *Thermus thermophilus* strain. (Lehrstuhl für Mikrobiologie, Technische Universität München, 2017).
101. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–54 (1976).

Literaturverzeichnis

102. Jarzab, A. *et al.* Meltome Atlas – thermal proteome stability across the tree of life. *Nature Methods* (eingereichte Publikation, 2019).
103. Bjellqvist, B. *et al.* The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **14**, 1023–31 (1993).
104. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–32 (1982).
105. Leuenberger, P. *et al.* Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **355**, eaai7825 (2017).
106. Koonin, E. V. Orthologs, paralogs, and evolutionary genomics. *Annu. Rev. Genet.* **39**, 309–38 (2005).
107. Huerta-Cepas, J. *et al.* eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286–93 (2016).
108. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
109. Huerta-Cepas, J. *et al.* Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017).
110. Petersen, M. *et al.* Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics* **18**, 111 (2017).
111. Rey, S. *et al.* PSORTdb: a protein subcellular localization database for bacteria. *Nucleic Acids Res.* **33**, D164–8 (2005).
112. Yu, N. Y., Laird, M. R., Spencer, C. & Brinkman, F. S. L. PSORTdb--an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res.* **39**, D241–4 (2011).
113. Peabody, M. A., Laird, M. R., Vlasschaert, C., Lo, R. & Brinkman, F. S. L. PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* **44**, D663–D668 (2016).
114. Van Derlinden, E. & Van Impe, J. F. Modeling growth rates as a function of temperature: Model performance evaluation with focus on the suboptimal temperature range. *Int. J. Food Microbiol.* **158**, 73–78 (2012).
115. Sonnleitner, B., Cometta, S. & Fiechter, A. Microbiology and growth kinetics of *Thermus thermophilus*. *Eur. J. of Appl. Microbiol. Biotechnol.* **15**, 75–82 (1982).
116. Cebrián, G., Condón, S. & Mañas, P. Physiology of the Inactivation of Vegetative Bacteria by Thermal Treatments: Mode of Action, Influence of Environmental Factors and Inactivation Kinetics. *Foods (Basel, Switzerland)* **6**, 107 (2017).
117. Demirtas, M. U., Kolhatkar, A. & Kilbane, J. J. Effect of aeration and agitation on growth rate of *Thermus thermophilus* in batch mode. *J. Biosci. Bioeng.* **95**, 113–117 (2003).
118. Leis, B. *et al.* Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. *Front. Microbiol.* **6**, 275 (2015).
119. Degryse, E., Glansdorff, N. & Pierard, A. Comparative analysis of extreme thermophilic bacteria belonging to genus *Thermus*. *Arch. Microbiol.* **117**, 189–196 (1978).
120. Single, F. O. R. A. *et al.* Respiratory Proteins from the Extremely Thermophilic Aerobic. **853**, 153–185 (1984).

Literaturverzeichnis

121. Bjornsdottir, S. H. *et al.* *Thermus islandicus* sp. nov., a mixotrophic sulfur-oxidizing bacterium isolated from the Torfajokull geothermal area. *Int. J. Syst. Evol. Microbiol.* **59**, 2962–6 (2009).
122. Antranikian, G. *Angewandte Mikrobiologie*. (Springer-Verlag, 2006). doi:10.1007/3-540-29456-2.
123. Fujita, A., Sato, T., Koyama, Y. & Misumi, Y. A reporter gene system for the precise measurement of promoter activity in *Thermus thermophilus* HB27. *Extremophiles* **19**, 1193–201 (2015).
124. Ebihara, A., Manzoku, M., Fukui, K. & Shimada, A. Roles of Mn - catalase and a possible heme peroxidase homologue in protection from oxidative stress in *Thermus thermophilus*. *Extremophiles* **19**, 775–785 (2015).
125. Ohtani, N., Tomita, M. & Itaya, M. An Extreme Thermophile, *Thermus thermophilus*, Is a Polyploid Bacterium. **192**, 5499–5505 (2010).
126. Sun, L. *et al.* Effective polyploidy causes phenotypic delay and influences bacterial evolvability. *PLoS Biol.* **16**, e2004644 (2018).
127. Angelov, A., Li, H., Geissler, A., Leis, B. & Liebl, W. Toxicity of indoxyl derivative accumulation in bacteria and its use as a new counterselection principle. *Syst. Appl. Microbiol.* **36**, 585–592 (2013).
128. Xie, Z. *et al.* An external substrate-free blue/white screening system in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **101**, 3811–3820 (2017).
129. Duque, E. *et al.* Ruminal metagenomic libraries as a source of relevant hemicellulolytic enzymes for biofuel production. *Microb. Biotechnol.* **11**, 781–787 (2018).
130. Rashamuse, K. *et al.* Metagenomic mining of glycoside hydrolases from the hindgut bacterial symbionts of a termite (*Trinervitermes trinervoides*) and the characterization of a multimodular β -1,4-xylanase (GH11). *Biotechnol. Appl. Biochem.* **64**, 174–186 (2017).
131. Zaccolo, M., Williams, D. M., Brown, D. M. & Gherardi, E. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**, 589–603 (1996).
132. Stemmer, W. P. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10747–51 (1994).
133. Porebski, B. T. & Buckle, A. M. Consensus protein design. *Protein Eng. Des. Sel.* **29**, 245–251 (2016).
134. Lehmann, M. *et al.* The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* **15**, 403–11 (2002).
135. Arakawa, T. & Timasheff, S. N. Theory of protein solubility. in *Methods in Enzymology* vol. 114 49–77 (1985).
136. Klostermeier, D., Seidel, R. & Reinstein, J. Functional properties of the molecular chaperone DnaK from *Thermus thermophilus*. *J. Mol. Biol.* **279**, 841–853 (1998).
137. Groemping, Y. & Reinstein, J. 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 (2001).
138. Mateus, A. *et al.* Thermal proteome profiling in bacteria: probing protein state in vivo. *Mol. Syst. Biol.* **14**, e8242 (2018).
139. Ebrahimi, M., Lakizadeh, A., Agha-Golzadeh, P., Ebrahimie, E. & Ebrahimi, M. Prediction of thermostability from amino acid attributes by combination of clustering with attribute weighting: a new vista in engineering enzymes. *PLoS One* **6**, e23146 (2011).
140. Yokota, K., Satou, K. & Ohki, S. ya. Comparative analysis of protein thermostability: Differences in amino acid content and substitution at the surfaces and in the core regions of thermophilic and mesophilic proteins. *Sci. Technol. Adv. Mater.* **7**, 255–262 (2006).

Literaturverzeichnis

141. Farias, S. T. & Bonato, M. C. M. Preferred amino acids and thermostability. *Genet. Mol. Res.* **2**, 383–393 (2003).
142. Ponnuswamy, P. K., Muthusamy, R. & Manavalan, P. Amino acid composition and thermal stability of proteins. *Int. J. Biol. Macromol.* **4**, 186–190 (1982).
143. Pucci, F., Bourgeas, R. & Rooman, M. Predicting protein thermal stability changes upon point mutations using statistical potentials: Introducing HoTMuSiC. *Sci. Rep.* **6**, 23257 (2016).
144. Williams, J. C. *et al.* Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power. *Protein Eng.* **12**, 243–50 (1999).
145. Perez-Jimenez, R. *et al.* Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat. Publ. Gr.* **18**, 592–596 (2011).
146. Mansfeld, J. *et al.* Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J. Biol. Chem.* **272**, 11152–6 (1997).
147. B. Alberts. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **92**, 291–294 (1998).
148. Chang, R. L. *et al.* Structural systems biology evaluation of metabolic thermotolerance in Escherichia coli. *Science* **340**, 1220–3 (2013).
149. Farci, D. *et al.* BBA - Biomembranes On the S-layer of Thermus thermophilus and the assembling of its main protein SlpA. *BBA - Biomembr.* **1860**, 1554–1562 (2018).
150. Kramer, B. & Thielmann, J. Monitoring the live to dead transition of bacteria during thermal stress by a multi-method approach. *J. Microbiol. Methods* **123**, 24–30 (2016).
151. Marcén, M., Ruiz, V., Serrano, M. J., Condón, S. & Mañas, P. Oxidative stress in E. coli cells upon exposure to heat treatments. *Int. J. Food Microbiol.* **241**, 198–205 (2017).
152. Olabarría, G., Castón, J. R., Lasa, I., Casto, R. & Lasa, I. G. O. Horizontal transference of S-layer genes within Thermus thermophilus. *J. Bacteriol.* **177**, 5460–5466 (1995).
153. Lee, J. & Kaletunç, G. Evaluation of the heat inactivation of Escherichia coli and Lactobacillus plantarum by differential scanning calorimetry. *Appl. Environ. Microbiol.* **68**, 5379–86 (2002).

ANHANG

Tabelle 31: Sequenzen aller verwendeter Oligonukleotide.

Bezeichnung	Sequenz
cgl.t.f.g.V2	TAAAACGACGGCCAGTGCCAAAGAAGGAGATATACATATGGC
cgl.t.r.g	GCTCGGTACCCGGGGATCCTCTTTCCTTTTCGGGCTTTGTTA
P0220-f.g	TAAAACGACGGCCAGTGCCATCGGCTTCTACCTGGACCTC
P0220-r.g.3	GGTTTCTGTTATACTCCCGGGCCGCATCCTCACTCCTTGAC
slp.f.g	GTCAAGGAGTGAGGATGCGGCCCGGGAGTATAACAGAAACC
bleo.r.g	GAGAAGATGAGCCGGGCCAGCGGTATCCGTTATGCTAAGG
P0222-f.g	CCTTAGCATAACGGATAACCGCTGGCCCGGCTCATCTTCTC
P0222-r.g.2	GCTCGGTACCCGGGGATCCTCTCCCGGATGGCCAGGATC
0042-1F	CTTCGCCTCCATTCCCAAG
0042-1R	GAGGCCATCTCCGAGGGAAAG
220-wt-1	GGTACCTGGCCTCTACGAGAAC
220-wt-2	TCCGAGCCCACGAAGTAGAC
220-del-1	GGGACGACGTGACCCTGTTTCATC
220-del-2	TCAATGACCTCCGCCAGGAGGAC

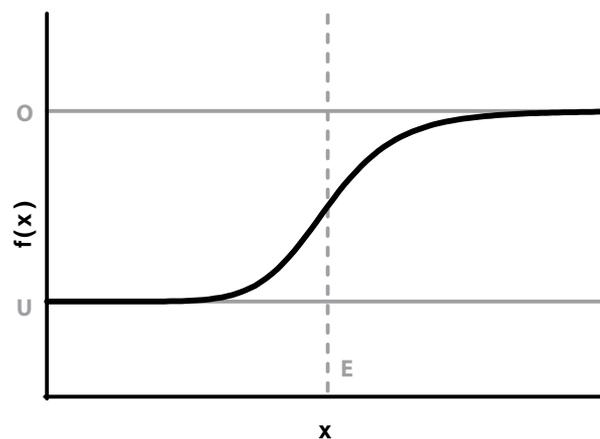


Abbildung 29: Beispielhafte Darstellung einer Kurve logistischer Regression. O – y-Wert des oberen Plateaus, U – y-Wert des unteren Plateaus, E – x-Wert an der steilsten Stelle der Kurve.

Anhang

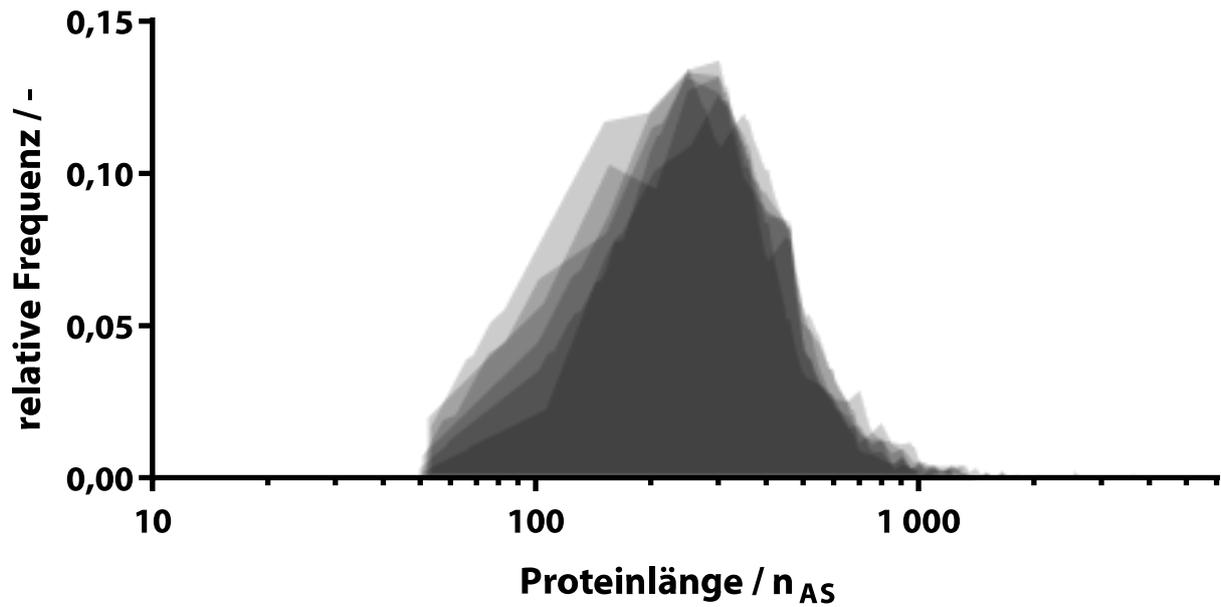


Abbildung 30: Überlagerte Darstellung der Proteinlängenverteilung aller Organismen.

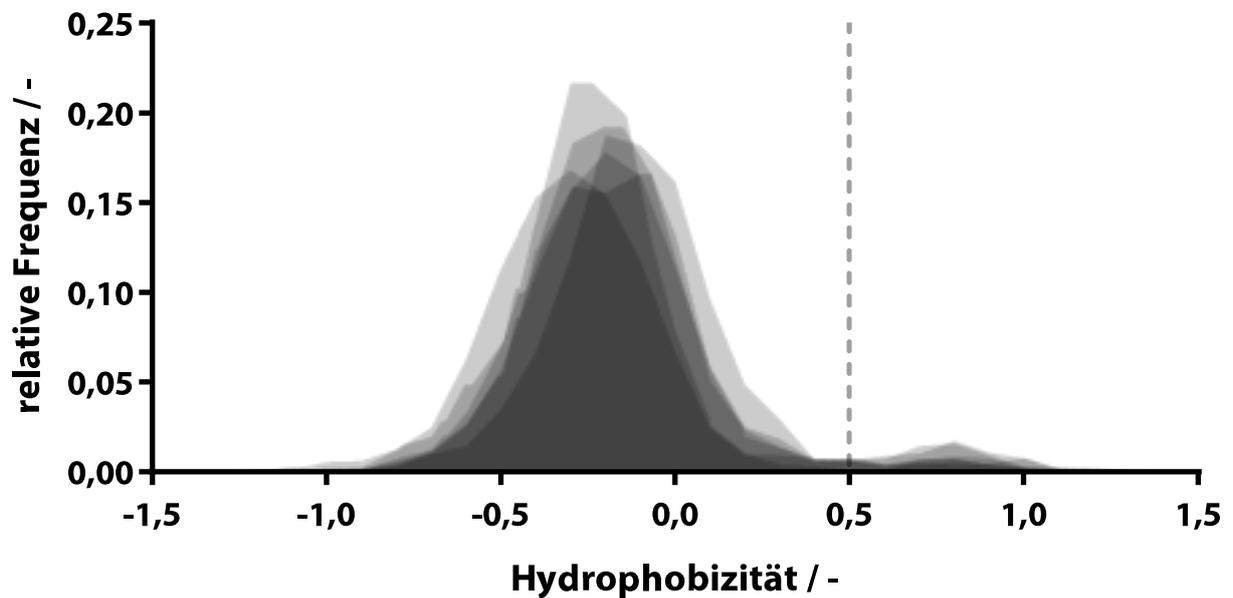


Abbildung 31: Überlagerte Darstellung der Hydrophobizitätsverteilung aller Organismen. Die Berechnung der Hydrophobizität erfolgte anhand der Proteinsequenz nach KYTE & DOOLITTLE¹⁰⁴. Die vertikale Linie markiert den gewählten Grenzwert um die Haupt- von der Nebenpopulation (Hydrophobizität > 0,5) zu trennen.

Anhang

Tabelle 32: GO-Begriffanreicherung von Proteinen aus *O. antarctica* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

O. antarctica

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0006810	transport	53	13	1,80	2,00E-10	****
GO:0051234	establishment of localization	56	13	1,90	4,40E-10	****
GO:0051179	localization	63	13	2,14	2,20E-09	****
GO:0006811	ion transport	21	8	0,71	5,30E-08	****
GO:0015672	monovalent inorganic cation transport	11	6	0,37	2,70E-07	****
GO:0055085	transmembrane transport	30	8	1,02	1,30E-06	****
GO:0006812	cation transport	15	6	0,51	2,70E-06	****
GO:0034220	ion transmembrane transport	15	5	0,51	6,30E-05	****
GO:0006091	generation of precursor metabolites and energy	26	6	0,88	9,90E-05	****
GO:0015980	energy derivation by oxidation of organic compounds	17	5	0,58	1,20E-04	***

Tabelle 33: GO-Begriffanreicherung von Proteinen aus *B. subtilis* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

B. subtilis

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0006810	transport	118	41	6,58	7,40E-29	****
GO:0051234	establishment of localization	118	41	6,58	7,40E-29	****
GO:0051179	localization	122	41	6,80	3,60E-28	****
GO:0071702	organic substance transport	75	29	4,18	1,70E-20	****
GO:0071705	nitrogen compound transport	49	22	2,73	6,50E-17	****
GO:0055085	transmembrane transport	30	17	1,67	3,00E-15	****
GO:0015711	organic anion transport	19	12	1,06	1,10E-11	****
GO:0015849	organic acid transport	19	12	1,06	1,10E-11	****
GO:0046942	carboxylic acid transport	19	12	1,06	1,10E-11	****
GO:0006811	ion transport	54	18	3,01	4,00E-11	****

Anhang

Tabelle 34: GO-Begriffanreicherung von Proteinen aus *E. coli* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

E. coli

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0006810	transport	206	67	12,07	<1,0E-30	****
GO:0051234	establishment of localization	211	67	12,36	<1,0E-30	****
GO:0051179	localization	239	68	14,00	<1,0E-30	****
GO:0071702	organic substance transport	147	49	8,61	1,60E-29	****
GO:0055085	transmembrane transport	102	41	5,98	8,10E-28	****
GO:0006811	ion transport	104	38	6,09	9,50E-24	****
GO:0015849	organic acid transport	36	21	2,11	3,60E-18	****
GO:0046942	carboxylic acid transport	36	21	2,11	3,60E-18	****
GO:0071705	nitrogen compound transport	80	29	4,69	1,00E-17	****
GO:0034220	ion transmembrane transport	53	24	3,11	2,50E-17	****

Tabelle 35: GO-Begriffanreicherung von Proteinen aus *G. stearothermophilus* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

G. stearothermophilus

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0009141	nucleoside triphosphate metabolic process	10	1	0,28	0,28	ns
GO:0009144	purine nucleoside triphosphate metabolic process	10	1	0,28	0,28	ns
GO:0009199	ribonucleoside triphosphate metabolic process	10	1	0,28	0,28	ns
GO:0009205	purine ribonucleoside triphosphate metabolic process	10	1	0,28	0,28	ns
GO:0006163	purine nucleotide metabolic process	11	1	0,31	0,31	ns
GO:0009150	purine ribonucleotide metabolic process	11	1	0,31	0,31	ns
GO:0009259	ribonucleotide metabolic process	11	1	0,31	0,31	ns
GO:0019693	ribose phosphate metabolic process	11	1	0,31	0,31	ns
GO:0072521	purine-containing compound metabolic process	11	1	0,31	0,31	ns
GO:0006753	nucleoside phosphate metabolic process	12	1	0,33	0,33	ns

Anhang

Tabelle 36: GO-Begriffanreicherung von Proteinen aus *P. torridus* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

P. torridus

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0055085	transmembrane transport	16	8	0,44	6,60E-10	****
GO:0006810	transport	27	9	0,75	6,60E-10	****
GO:0051179	localization	27	9	0,75	4,00E-03	****
GO:0051234	establishment of localization	27	9	0,75	4,00E-03	****
GO:0045333	cellular respiration	13	3	0,36	6,10E-03	**
GO:0046034	ATP metabolic process	13	3	0,36	7,40E-03	**
GO:0015980	energy derivation by oxidation of organic compounds	15	3	0,41	7,40E-03	**
GO:0009144	purine nucleoside triphosphate metabolic process	16	3	0,44	1,05E-02	**
GO:0009205	purine ribonucleoside triphosphate metabolic process	16	3	0,44	6,60E-10	**
GO:0009199	ribonucleoside triphosphate metabolic process	18	3	0,50	6,60E-10	*

Tabelle 37: GO-Begriffanreicherung von Proteinen aus *T. thermophilus* mit einer Hydrophobizität $\geq 0,5$. Gezeigt sind die 10 signifikantesten Begriffe. Abkürzungen: *an.*: annotierte Proteine, *sig.*: Signifikanzschwelle, *erw.*: Anzahl erwarteter Proteine, *Sig.*: Signifikanzgrad. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$.

T. thermophilus

GO-ID	GO-Begriff	an.	sig.	erw.	p-Wert	Sig.
GO:0055085	transmembrane transport	12	6	0,33	1,00E-07	****
GO:0045333	cellular respiration	13	5	0,36	7,80E-06	****
GO:0015980	energy derivation by oxidation of organic compounds	14	5	0,39	1,20E-05	****
GO:0006810	transport	25	6	0,69	1,60E-05	****
GO:0051179	localization	25	6	0,69	1,60E-05	****
GO:0051234	establishment of localization	25	6	0,69	1,60E-05	****
GO:0006091	generation of precursor metabolites and energy	17	5	0,47	3,50E-05	****
GO:0055114	oxidation-reduction process	22	5	0,61	1,40E-04	***
GO:0046034	ATP metabolic process	10	3	0,28	1,77E-03	**
GO:0009144	purine nucleoside triphosphate metabolic process	11	3	0,30	2,40E-03	**

Anhang

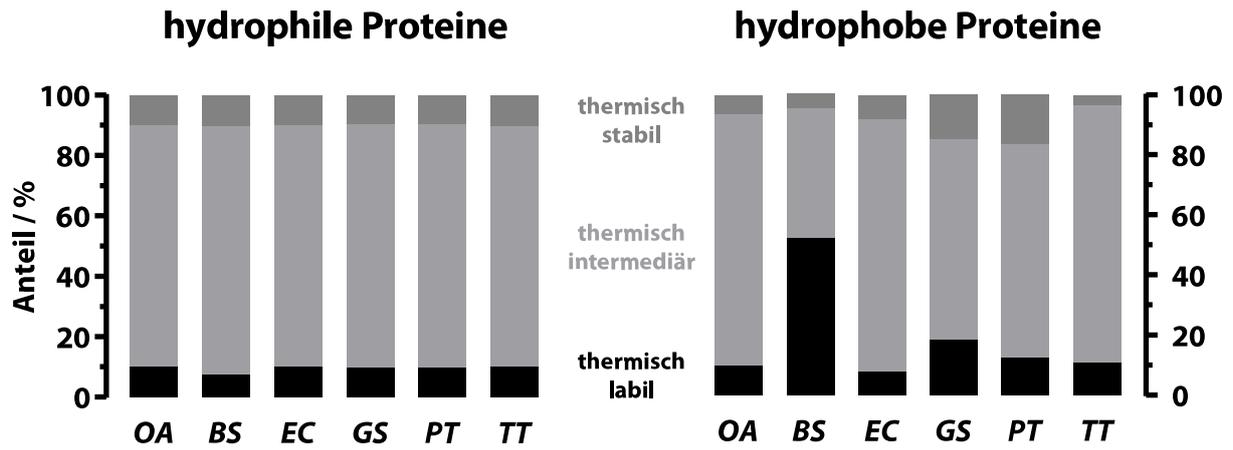


Abbildung 32: Anteile hydrophober und hydrophiler Proteine nach Thermostabilitätsbereich. Die Trennung in hydrophob und hydrophil erfolgte anhand eines Grenzwertes von 0,5.

Error! Use the Home tab to apply Überschrift 1 to the text that you want to appear here.

Tabelle 38: Aminosäurefrequenzunterschiede im Vergleich thermolabiler und thermostabiler Proteine. Signifikanzgrade sind folgendermaßen definiert: ns (nicht signifikant): $p > 0,05$; *: $p \leq 0,05$; **: $p \leq 0,01$; ***: $p \leq 0,001$; ****: $p \leq 0,0001$. Die Pfeilrichtung recht der Signifikanzgrade gibt an, ob der Median der thermisch stabilen Fraktion über (▲) oder unter (▼) dem der labilen Fraktion liegt. Schwarze Pfeile symbolisieren statistisch signifikante, Weiße statistisch nicht signifikante Werte.

	<i>O. antarctica</i>			<i>B. subtilis</i>			<i>E. coli</i>			<i>G. stearotherophilus</i>			<i>P. torridus</i>			<i>T. thermophilus</i>		
AS	p-Wert	Sig.		p-Wert	Sig.		p-Wert	Sig.		p-Wert	Sig.		p-Wert	Sig.		p-Wert	Sig.	
A	2.04E-04	***	▼	3.39E-02	*	▼	7.85E-03	**	▼	8.41E-02	ns	△	8.00E-01	ns	△	4.47E-02	*	▼
C	7.89E-01	ns	△	1.93E-06	****	▼	5.60E-01	ns	△	4.02E-02	*	▼	5.48E-02	ns	△	2.56E-01	ns	△
D	4.64E-01	ns	△	3.84E-09	****	▼	9.15E-03	**	▼	1.33E-02	*	▼	8.29E-01	ns	△	3.54E-01	ns	△
E	7.17E-01	ns	△	4.12E-06	****	▼	1.19E-01	ns	△	7.42E-01	ns	△	4.94E-01	ns	△	7.63E-02	ns	△
F	3.54E-03	**	▼	4.97E-03	**	▼	3.44E-01	ns	△	9.85E-01	ns	△	6.71E-01	ns	△	9.58E-03	**	▼
G	6.16E-02	ns	△	3.29E-01	ns	△	1.85E-01	ns	△	5.99E-01	ns	△	1.23E-01	ns	△	8.35E-01	ns	△
H	2.55E-01	ns	△	1.01E-06	****	▼	2.17E-02	*	▼	7.26E-01	ns	△	2.09E-03	**	▼	2.84E-01	ns	△
I	7.87E-01	ns	△	2.81E-04	***	▼	1.78E-01	ns	△	9.70E-01	ns	△	3.77E-01	ns	△	2.08E-03	**	▼
K	4.53E-01	ns	△	1.47E-02	*	▼	4.85E-02	*	▼	5.47E-01	ns	△	1.25E-02	*	▼	1.07E-01	ns	△
L	4.10E-03	**	▼	5.43E-09	****	▼	6.04E-04	***	▼	1.83E-06	****	▼	5.38E-01	ns	△	2.74E-01	ns	△
M	1.96E-01	ns	△	3.90E-01	ns	△	7.22E-01	ns	△	6.02E-02	ns	△	1.24E-03	**	▼	9.10E-01	ns	△
N	6.15E-01	ns	△	4.28E-03	**	▼	5.97E-01	ns	△	1.20E-01	ns	△	6.42E-02	ns	△	5.51E-01	ns	△
P	2.22E-01	ns	△	1.60E-02	*	▼	4.27E-01	ns	△	7.41E-01	ns	△	6.25E-01	ns	△	4.19E-02	*	▼
Q	1.29E-02	*	▼	6.21E-01	ns	△	2.14E-01	ns	△	6.00E-02	ns	△	4.78E-01	ns	△	9.52E-02	ns	△
R	2.53E-01	ns	△	4.66E-01	ns	△	5.38E-12	****	▼	3.87E-01	ns	△	5.12E-03	**	▼	2.96E-03	**	▼
S	3.28E-02	*	▼	7.72E-03	**	▼	2.44E-02	*	▼	1.40E-03	**	▼	4.79E-02	*	▼	3.25E-01	ns	△
T	1.24E-02	*	▼	5.34E-01	ns	△	3.32E-01	ns	△	4.54E-01	ns	△	3.26E-01	ns	△	5.58E-01	ns	△
V	6.84E-02	ns	△	2.05E-01	ns	△	2.37E-01	ns	△	2.75E-01	ns	△	2.99E-01	ns	△	1.65E-01	ns	△
W	3.92E-01	ns	△	5.38E-01	ns	△	2.25E-01	ns	△	7.25E-01	ns	△	1.75E-01	ns	△	4.51E-03	**	▼
Y	5.55E-02	ns	△	1.41E-02	*	▼	5.83E-01	ns	△	6.72E-01	ns	△	1.62E-01	ns	△	2.24E-02	*	▼

Error! Use the Home tab to apply Überschrift 1 to the text that you want to appear here.

T. thermophilus

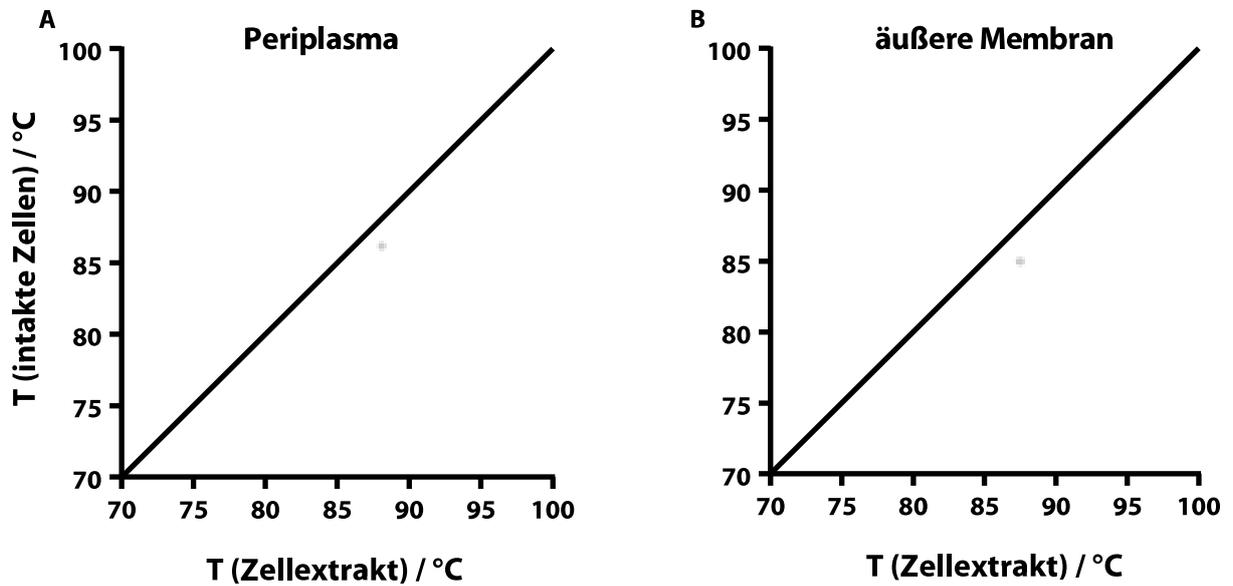


Abbildung 33: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in *T. thermophilus* nach Lokalisation. Die Graphen für Proteine des Cytoplasmas und der Periplasmamembran sind unter 3.2.6 zu finden (Abbildung 22). Die schwarze Linie ist eine Ursprungsgerade mit einer Steigung von 1.

ABBILDUNGS- VERZEICHNIS

Abbildung 1: Isolationsort des extrem thermophilen Bakteriums <i>Thermus thermophilus</i>	6
Abbildung 2: Karten des Chromosoms (TTC) und Megaplasmid (pTT27) von <i>T. thermophilus</i> HB27	8
Abbildung 3: Grundlegender molekularer Aufbau eines TMT10plex-Labels	12
Abbildung 4: Schema zur Herstellung eines GH-defizienten <i>T. thermophilus</i> -Stammes	38
Abbildung 5: <i>p</i> -Nitrophenyl- β -D-Glucopyranosid.....	40
Abbildung 6: Wachstum von <i>T. thermophilus</i> in basalem TMB mit unterschiedlicher Aminosäuresupplementierung	46
Abbildung 7: Wachstum unterschiedlicher <i>T. thermophilus</i> -Stämme in TMB bei 60 °C mit D-Cellobiose	48
Abbildung 8: Temperaturabhängige Aktivität von Bgl, CglT und CglT KIR anhand eines <i>p</i> NP-Assays mit <i>p</i> NPG	50
Abbildung 9: Venn-Diagramm aller massenspektrometrisch analysierten Proteome	52
Abbildung 10: Verteilung der Proteinaggregationstemperaturen (T_m) nach Organismus	54
Abbildung 11: Verteilung der <i>pI</i> -Werte für Proteine mit Schmelzkurve und des Gesamtproteoms	56
Abbildung 12: Proteinlänge in AS gruppiert nach thermischer Proteinstabilität.....	58
Abbildung 13: Proteinhydrophobizität gruppiert nach thermischer Proteinstabilität	60
Abbildung 14: Aminosäurefrequenzvergleich thermostabiler und -labiler Proteine	63
Abbildung 15: T_m -Differenzwerte von gepaarten und randomisierten Paralogen nach Organismus	65
Abbildung 16: Gepaarte und randomisierte Paraloge nach Sequenzidentität (alle Mikroorganismen)	66
Abbildung 17: T_m -Differenzwerte von gepaarten und randomisierten Orthologen zwischen allen Organismen.....	68
Abbildung 18: Gepaarte und randomisierte Orthologe nach Sequenzidentität (alle Mikroorganismen).....	69
Abbildung 19: Verteilung der T_m -Werte von <i>E. coli</i> und <i>T. thermophilus</i> für Zellextrakte und intakte Zellen	70
Abbildung 20: Venn-Diagramm für Proteine aus Zellextrakten und intakten Zellen für <i>E. coli</i> und <i>T. thermophilus</i>	71
Abbildung 21: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in <i>E. coli</i> nach Lokalisation.....	73
Abbildung 22: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in <i>T. thermophilus</i> nach Lokalisation.....	74
Abbildung 23: Venn-Diagramm: Proteine aus <i>T. th.</i> nach genomischer Lokalisation (Zellextrakt und intakte Zellen).....	76
Abbildung 24: T_m -Werte Verteilung für <i>T. thermophilus</i> in Zellextrakt und intakter Zellen	76
Abbildung 25: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in <i>T. thermophilus</i>	77
Abbildung 26: Relative Überlebensrate von <i>E. coli</i> und <i>T. thermophilus</i> in Abhängigkeit von der Temperatur	78
Abbildung 27: Isotherme, dynamische Differenzfluorimetrie von <i>E. coli</i> und <i>T. thermophilus</i>	81
Abbildung 28: <i>T. thermophilus</i> Megaplasmid pTT27.....	87
Abbildung 29: Beispielhafte Darstellung einer Kurve logistischer Regression.....	XX
Abbildung 30: Überlagerte Darstellung der Proteinlängenverteilung aller Organismen	XXI

Abbildungsverzeichnis

Abbildung 31: Überlagerte Darstellung der Hydrophobizitätsverteilung aller Organismen	XXI
Abbildung 32: Anteile hydrophober und hydrophiler Proteine nach Thermostabilitätsbereich	XXV
Abbildung 33: T_m -Werte von Proteinen aus Zellextrakten und intakten Zellen in <i>T. thermophilus</i> nach Lokalisation....	XXVII

TABELLEN- VERZEICHNIS

Tabelle 1: Abkürzungen und deren ausgeschriebene Form	V
Tabelle 2: Summenformeln chemischer Verbindungen und deren ausgeschriebene Form	VI
Tabelle 3: Auflistung verwendeter Chemikalien	15
Tabelle 4: Zusammensetzung verwendeter Nährmedien	17
Tabelle 5: Zusammensetzung verwendeter Puffer und Lösungen	19
Tabelle 6: Auflistung verwendeter kommerzieller Kits.....	20
Tabelle 7: Auflistung verwendeter kommerzieller Lösungen und Reagenzien	20
Tabelle 8: Auflistung verwendeter technischer Gerätschaften.....	21
Tabelle 9: Auflistung verwendeter Verbrauchsmaterialien.....	23
Tabelle 10: Verwendete Computersoftware und deren Versionsnummer.....	24
Tabelle 11: Auflistung verwendeter Mikroorganismen	25
Tabelle 12: Verwendeten Medien in Bezug auf die zu kultivierenden Mikroorganismen	26
Tabelle 13: Verwendete Antibiotika und deren eingesetzte Konzentration.....	27
Tabelle 14: Einzelbestandteile des modifizierten BROCK-Mediums.....	28
Tabelle 15: Auflistung verwendeter Plasmide	29
Tabelle 16: Einzelbestandteile eines Q5 PCR-Ansatzes.....	31
Tabelle 17: Bedingungen einer standardmäßigen Q5 PCR.....	31
Tabelle 18: Zusammensetzung des Aufschlusspuffers für Kolonie- <i>Cracking</i>	36
Tabelle 19: Unterschiedliche prozentuale Proteomabdeckungen aller Organismen	53
Tabelle 20: Vergleich thermostabiler und -labiler Proteine anhand ihrer Länge	58
Tabelle 21: PEARSON-Korrelationskoeffizienten von T_m und Proteinlänge	59
Tabelle 22: Vergleich thermostabiler und -labiler Proteine anhand ihres Hydrophobizität	61
Tabelle 23: PEARSON-Korrelationskoeffizienten von T_m und Hydrophobizität	61
Tabelle 24: Signifikanzwerte eines einseitigen WILCOXON-Tests von gepaarten und randomisierten Paralogen.....	66
Tabelle 25: Signifikanzwerte eines einseitigen WILCOXON-Tests von gepaarten und randomisierten Orthologen.....	70
Tabelle 26: PEARSON-Korrelationskoeffizienten von T_m -Werten aus Zellextrakt und intakten Zellen nach Lokalisation.....	75
Tabelle 27: Berechnete Temperaturen für Überlebensraten von 10, 50 und 90 % von <i>E. coli</i> und <i>T. thermophilus</i> auf Grundlage der in Abbildung 26 dargestellten Experimente	79
Tabelle 28: GO-Begriffanreicherung des thermischen Inaktivierungsbereichs von <i>E. coli</i>	80
Tabelle 29: GO-Begriffanreicherung des thermischen Inaktivierungsbereichs von <i>T. thermophilus</i>	80

Tabellenverzeichnis

Tabelle 30: Temperaturen bei gegebener, normalisierter DSF-Fluoreszenzrate für <i>E. coli</i> und <i>T. thermophilus</i>	82
Tabelle 31: Sequenzen aller verwendeter Oligonukleotide	XX
Tabelle 32: GO-Begriffanreicherung von Proteinen aus <i>O. antarctica</i> mit einer Hydrophobizität $\geq 0,5$	XXII
Tabelle 33: GO-Begriffanreicherung von Proteinen aus <i>B. subtilis</i> mit einer Hydrophobizität $\geq 0,5$	XXII
Tabelle 34: GO-Begriffanreicherung von Proteinen aus <i>E. coli</i> mit einer Hydrophobizität $\geq 0,5$	XXIII
Tabelle 35: GO-Begriffanreicherung von Proteinen aus <i>G. stearothermophilus</i> mit einer Hydrophobizität $\geq 0,5$	XXIII
Tabelle 36: GO-Begriffanreicherung von Proteinen aus <i>P. torridus</i> mit einer Hydrophobizität $\geq 0,5$	XXIV
Tabelle 37: GO-Begriffanreicherung von Proteinen aus <i>T. thermophilus</i> mit einer Hydrophobizität $\geq 0,5$	XXIV
Tabelle 38: Aminosäurefrequenzunterschiede im Vergleich thermolabiler und thermostabiler Proteine	XXVI

EIGENSTÄNDIGKEITS- ERKLÄRUNG

Ich erkläre hiermit, dass ich meine Dissertation mit dem Titel „Entwicklung eines Systems zur komplementationsbasierten Selektion thermostabiler Glycosidhydrolasen in *Thermus thermophilus*“ (engl.: „Development of a system for complementation-based selection of thermostable glycoside hydrolases in *Thermus thermophilus*“) selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt habe. Alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, wurden durch Angaben von Quellen als Entlehnung kenntlich gemacht.

Matthias Mörch, M.Sc. (TUM)

Freising, den 01. Juli 2019

DANKSAGUNG

Ich danke meinem Doktorvater Prof. Dr. Wolfgang Liebl, dass ich meine Dissertation unter seiner Leitung anfertigen und an einem sehr spannenden Forschungsthema teilhaben durfte.

Dem Zweitprüfer Prof. Dr. Rudi Vogel sowie der Prüfungsvorsitzenden Prof. Dr. Aphrodite Kapurniotu möchte ich herzlich für ihre Bereitschaft an der Teilnahme zur Doktorprüfung danken.

Besonderer Dank gilt Dr. Angel Angelov als Labormanager und direkter fachlicher Betreuer. Er wusste immer neue Impulse zu geben und regte stets an, außerhalb üblicher Strukturen zu denken. Ohne Ihn hätte ich es sicherlich auch nie geschafft, eine 7+ zu klettern.

Ich danke Dr. Vladimir Zverlov, Dr. Benedikt Leis, Dr. Claudia Held und Dr. Björn Andreeßen der Arbeitsgruppe Schwarz/Zverlov für fachlichen Input sowie für die Bereitstellung diverser Enzyme und Plasmidkonstrukte.

Ich möchte Dr. Trang Pham-Weiß für die kompetente und reibungslose Projektübergabe danken. Zudem hatte sie immer ein offenes Ohr für Probleme wissenschaftlicher oder privater Natur und stand mir mit guten Ratschlägen zur Seite.

Für die großartige Atmosphäre im Labor 002 danke ich meinen Mitstreitern Antoni Lichev und Dr. Maximilian Surger.

Jahrelang saßen Maria Übelacker und ich täglich nebeneinander im Labor. Ohne sie hätte ich die üblichen Durststrecken einer Doktorarbeit nicht durchgestanden. Durch ihre positive Art und ihr soziales Wesen war sie mehr als eine Kollegin und wurde über diese Zeit zu einer guten Freundin. Dafür danke ich ihr von Herzen!

Ich danke Kornelia Garus; der guten Seele unseres Lehrstuhls. Sie hat mich nicht nur immer tatkräftig bei der Arbeit unterstützt, sondern war auch stets bereit mit mir anregende Gespräche zum Thema Kochen sowie Backen (und selbstverständlich auch Essen) zu führen.

Dank geht zudem an die von mir betreuten Master- und Bachelorstudenten Tobias Hecht, Alberto Salas Ruíz, Sevgi Tansu Bağdatlı und Katharina Eyring. Diese vier haben immer wieder frischen Wind in unser Labor gebracht.

Ich bedanke mich ganz herzlich bei meinen Freunden Isabella Carolina Kreszentia Merlin und Florian Hans Joseph Heinritzi für den Beistand in den letzten Stunden der Entstehung dieser Arbeit.

Abschließend bedanke ich mich herzlich bei meinen Eltern und meinem Bruder für die uneingeschränkte, liebevolle und vielseitige Unterstützung während meines Studiums und meiner Zeit als Doktorand. Auch wenn ich oft an mir gezweifelt habe, haben sie stets an mich geglaubt.