



Fakultät Wissenschaftszentrum Weihenstephan  
für Ernährung, Landnutzung und Umwelt  
Lehrstuhl für Mikrobiologie

# Charakterisierung thermostabiler Enzyme aus funktionell durchforsteten metagenomischen Genbanken

Vu Thuy Trang Pham

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.

Vorsitzender:

Prof. Dr. Erwin Grill

Prüfer der Dissertation

1. Prof. Dr. Wolfgang Liebl

2. Prof. Dr. Rudi F. Vogel

Die Dissertation wurde am 18.12.2017 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 22.02.2018 angenommen.



*Ich widme diese Doktorarbeit meiner  
Phamily Vu*

## I. Zusammenfassung

In der vorliegenden Arbeit wurden neuartige Enzyme untersucht, deren Gene in metagenomischen Genbanken aufgefunden werden konnten. Das Gen mit der Bezeichnung *engU* (ursprünglich *bga48*) auf dem Fosmid FosK48H3 stammt von Umweltproben am Avachinsky Vulkan (Kamtschatka) und war im Rahmen früherer Arbeiten identifiziert worden. Die Funktionsanalyse des rekombinant hergestellten Genprodukts weist auf eine Carboxymethylcellulase (CMCase) hin, während die sequenzbasierten Analysen EngU als eine  $\beta$ -Galaktosidase kennzeichnen. Der modulare Aufbau von EngU ist außergewöhnlich. Aminosäuresequenzvergleiche ließen eine geringfügige Ähnlichkeit zur Glykosidhydrolase Familie 42 (GH42) zu, allerdings war in EngU das katalytische Modul durch eine Kohlenhydrat-Bindedomäne (CBM4\_9) unterbrochen. Schematisch kann die Modulstruktur von EngU wie folgt beschrieben werden: N-GH42a-CBM-GH42b-CTERM, wobei CTERM für einen 240 Aminosäuren großen Abschnitt unbekannter Funktion am C-Terminus steht. Die beiden Teile des GH42 Moduls, GH42a und GH42b, besitzen jeweils einen katalytischen Rest, nämlich die Positionen E239 (putative katalytische Säure/Base) und E581 (putatives katalytisches Nukleophil). Die Funktionen der einzelnen Enzymmodule konnten durch simultane Expression verschiedener Kombinationen von Teilabschnitten der *engU* Gensequenz im pETDuet-Vektor nicht eindeutig zugewiesen werden. Jedoch können sich in einer Zelle die separat aber gleichzeitig synthetisierten Polypeptidketten GH42a-CBM und GH42b-CTERM *co-trans* komplementieren und bildeten zusammen ein aktives Enzym. Neben dem aus früheren Arbeiten bekannten Substratspektrum ( $\beta$ -Glukan, Lichenan und CMC) wurden Aktivitäten mit Glukomannan und Xyloglukan festgestellt. EngU besitzt eine *endo*-spaltende Aktivität gegenüber gemischt  $\beta$ -1,3-/ $\beta$ -1,4-verknüpften Glukanketten und weiterhin eine vermutlich *exo*-spaltende Aktivität an  $\beta$ -1,3-glykosidischen Bindungen, wenn diese nahe an reduzierenden Enden oder an Seitenketten liegen, welche von  $\beta$ -1,4-Bindungen umgeben sind. EngU stellt den Prototyp einer neuen Glykosidhydrolase-Enzymfamilie dar, die die Bezeichnung GH148 erhielt.

Zwei weitere für thermostabile Enzyme kodierende Fosmide wurden im Rahmen der vergleichenden funktionellen Durchmusterungen von Genbanken hyperthermophiler Archaeen (HA) und thermophiler Bakterien (TB) in zwei verschiedenen Wirtsorganismen, *Escherichia coli* und *Thermus thermophilus*, identifiziert und untersucht. HA-CMC-1 kodiert für eine bei 70 °C aktive Cellulase, deren Aktivität nur im Wirt *E. coli* beobachtet wurde. Im Fall von Fosmid TB-EST-7 wurde die kodierte lipolytische Aktivität bei 60 °C in *T. thermophilus* gemessen, jedoch nur sehr schwach in dem mesophilen Bakterium *E. coli*. Zur Durchleuchtung dieser Aktivitätsunterschiede in den beiden Expressionswirten wurden zunächst Wachstumsexperimente und Real-Time PCR-Analysen mit den Fosmiden HA-CMC-

1 und TB-EST-7 durchgeführt. Gemessen in der frühen stationären Wachstumsphase der jeweiligen Wirtsbakterien ergab die Auswertung der Real-Time PCR Läufe ein allgemein höheres Transkriptionsniveau in *E. coli* als in *T. thermophilus*. Die beiden Kandidaten-ORFs der Fosmide HA-CMC-1 (HA23) und TB-EST-7 (TB32) werden demnach in den *E. coli* EPI300 Zellen häufiger transkribiert. Trotz des höheren Transkriptionslevel in *E. coli* wird die Esteraseaktivität verstärkt in *T. thermophilus* gemessen, was auf eine effiziente Proteinbiosynthese im thermophilen Wirt hinweist.

Die auf dem Fosmid HA-CMC-1 kodierte thermostabile Cellulase Cel12E war bereits im Vorfeld dieser Arbeit charakterisiert worden. Ausgehend von dem 534 Aminosäuren großen Cel12E wurde eine verkürzte, 272 Aminosäuren lange Version (GH12E) hergestellt. Diese Enzymvariante trägt nur mehr die beiden katalytischen GH12 Module. Das aufgereinigte Protein GH12E wurde an unsere Kooperationspartner an der Swedish University of Agricultural Science für eine Proteinkristallisation versendet. Die ebenfalls thermostabile, auf dem Fosmidinsert TB-EST-7 kodierte Esterase wurde erfolgreich rekombinant hergestellt und deren Aktivität mit pharmazeutisch relevanten Substraten getestet. Das Protein Est7 zeigte mit Ibuprofen eine geringe spezifische Aktivität von  $3,33 \times 10^{-2}$  mU/mg bei 60 °C und 0,54 mU/mg bei 80 °C Inkubationstemperatur.

In Rahmen dieser Dissertation wurden eigene Metagenom-Genbanken hergestellt und die Durchmusterungs-Methoden weiter entwickelt. Aus „Mexico“- und „Russia“-Umweltproben wurden zwei Metagenom-Genbibliotheken mit insgesamt 3322 bzw. 6645 Fosmiden erstellt und deren Diversität mittels 16S rDNA-Amplikonsequenzierung analysiert. Das komparative Durchmusterungsprotokoll für die beiden Expressionswirte wurde mithilfe eines halb-automatischen Pipettierroboters (VIAFLO 96) optimiert und so eine 300%ige Durchsatzsteigerung erreicht. Des Weiteren konnte die Effizienz der Übertragung der Fosmidbibliotheken in *T. thermophilus* mittels der Verwendung von Vakuum Manifold und spezieller Lysat-Bindeplatten signifikant von 29 % auf 88 % verbessert werden. Zudem wurde ein Protokoll für die Durchmusterung von gemischten Fosmidklonen (Pool-Screening) entwickelt, um eine hohe Erfolgsrate bei der Durchforstung der Klonbibliotheken zu erzielen. Es wurde ein neuer Vektor pRetro-B, der für den Transfer einer pCC1FOS-Metagenom-Genbank in *Bacillus subtilis* benötigt wird, hergestellt. Schließlich wurden, um einen BAC-Vektor für die Einführung großer Metagenominsets in *T. thermophilus* zu erhalten, pRetro-BAC1 und -BAC2 konstruiert, welche unterschiedliche Segmente des *T. thermophilus* Megaplasmids pTT27 enthalten. Die Vektoren sollten als extrachromosomale DNA-Elemente in der Zelle vorliegen, integrierten jedoch ins *Thermus* Genom.

## II. Summary

In this work new enzymes encoded by genes identified in metagenomic gene libraries were characterized. The gene *engU* (originally named *bga48*) was obtained from the fosmid FosK48H3, which was isolated from an Avachinsky crater (Kamtschatka) environmental sample. Functional characterization of the recombinantly produced EngU enzyme revealed carboxymethylcellulase (CMCase) activity, while sequence-based analysis predicted its specificity as a  $\beta$ -galactosidase. The module composition of EngU is unusual and indicates a catalytic module distantly related to glycoside hydrolase family 42 (GH42), which is split in half by a carbohydrate binding module (CBM4\_9). Schematically, the module structure of EngU can be written as follows: N-GH42a-CBM-GH42b-CTERM, whereby CTERM stands for a 240 amino acid residue C-terminal region with unknown function. The two parts of the GH42 module, GH42a and GH42b, each carry one catalytic residue: E239 (putative acid/base catalyst) and E581 (putative nucleophile catalyst), respectively. Differently combined sections of *engU* were co-expressed in pETDuet vectors, but the functions of the enzyme modules could not be completely elucidated. However, separate but simultaneous expression of the polypeptide chains GH42a\_CBM and GH42b\_CTERM could co-trans complement each other and together yielded an active enzyme. The previously known substrate spectrum of EngU, which includes  $\beta$ -glucan, lichenan, and carboxymethylcellulose, could be extended with additional substrates, glucomannan and xyloglucan. EngU displays *endo*-activity towards  $\beta$ -1,3-1,4-mixed-linkage glucan chains and in addition revealed an *exo*-activity for  $\beta$ -1,3-glycosidic bonds, if they are close to a reducing-end sugar or at side chains with neighboring  $\beta$ -1,4-bonds. EngU represents the prototype of a new glycoside hydrolase enzyme family, which is named GH148.

Two further thermostable proteins were identified using comparative functional screenings of “hyperthermophilic archaea” (HA) and “thermophile bacterium” (TB) metagenomic libraries in the two different hosts *Escherichia coli* and *Thermus thermophilus*. HA-CMC-1 encoded a cellulase active at 70 °C. This activity was observed only in the host *E. coli*. The lipolytic activity of TB-EST-7 was detectable in *T. thermophilus* at 60 °C but negligible in the mesophilic bacterium *E. coli*. To investigate the differences in expression in these two expression hosts, growth and quantitative PCR (qPCR) analysis were carried out. The evaluation of the qPCR experiments revealed a higher transcriptional level in *E. coli* than in *T. thermophilus*. Both candidate ORFs HA23 and TB32 of the respective fosmids were highly expressed in EPI300 cells. Despite the transcriptional level the esterase activity appears to be better in *T. thermophilus*, which could indicate a better protein biosynthesis in the thermophilic host.

The thermostable cellulase Cel12E from HA-CMC-1 has been characterized in previous work. Starting from the 534 amino acid Cel12E, a truncated 272-residue version (GH12E) was

designed, which contains only the two catalytic GH12 domains. The purified GH12E protein has been sent for crystallization analysis to our cooperation partners at the Swedish University of Agricultural Science. Furthermore, the esterase encoded on fosmid TB-EST-7 was expressed and screened towards pharmaceutically relevant substrates. The Est7 protein was expressed in *E. coli* and revealed activity towards Ibuprofen with  $3.33 \times 10^{-2}$  mU/mg at 60 °C and 0.54 mU/mg at an incubation temperature of 80 °C.

This thesis includes the establishment of metagenomic gene libraries and the development of improved screening methods. From the environmental samples „Mexico“ and „Russia“ two metagenomic gene libraries with 3322 and 6645 single fosmid clones, respectively, were generated and their diversity was analyzed via 16S rDNA sequencing. The comparative screening protocol for the two host bacteria was optimized with the aid of a semi-automatic pipetting robot (VIAFLO 96), which enhanced the screening throughput up to 300 %. Furthermore, the success of transformation in *T. thermophilus* was improved significantly from 29 % to 88 % by using special lysate plates with the associated vacuum manifold. Moreover, an effective pool screening protocol for pools of mixed fosmid clones was established, which enhanced the screening success. A new shuttle vector pRetro-B was constructed to simplify the transfer of a pCC1FOS-based metagenome gene library to *Bacillus subtilis*. Finally, to allow for the transfer of larger fosmid inserts into *T. thermophilus*, the new vectors pRetro-BAC1 und BAC2 were constructed, which hold different parts of the *T. thermophilus* pTT27 megaplasmid and should be established as external DNA elements in the cell. However, the vectors seem to integrate into the *T. thermophilus* genome.

# Inhaltsverzeichnis

|  |           |
|--|-----------|
| <b>I. ZUSAMMENFASSUNG.....</b>   | <b>4</b>  |
| <b>II. SUMMARY.....</b>  | <b>6</b>  |
| <b>1 EINLEITUNG.....</b>   | <b>19</b> |
| 1.1 „METAGENOMIK“ - DIE SUCHE NACH DER STECKNADEL IM METAGENOMISCHEN HEUHAUFEN.....                    | 19        |
| 1.2 UNTERSCHIEDLICHE EXPRESSIONSWIRTE FÜR DIE SUCHE NACH ENZYMEN AUS EXTREMOPHILEN.....                | 21        |
| 1.3 METAGENOMIK KOMBINIERT MIT DEN ENTWICKLUNGEN DES NEXT GENERATION SEQUENCINGS (NGS) .....           | 24        |
| 1.4 DER NUTZEN VON THERMOSTABILEN ENZYMEN FÜR BIOTECHNOLOGISCHE PROZESSE...                            | 24        |
| 1.5 KLASSIFIZIERUNG DER GLYKOSIDHYDROLASEN .....   | 28        |
| <b>2 ZIEL DER ARBEIT:.....</b>   | <b>31</b> |
| 2.1 EXPRESSION UND CHARAKTERISIERUNG DER THERMOSTABILEN B-GLUKANASE ENGU...                            | 31        |
| 2.2 EXPRESSION UND TRANSKRIPTIONELLE ANALYSE ZWEIER AKTIVER FOSMIDE IN ZWEI VERSCHIEDENEN WIRTEN ..... | 31        |
| 2.3 OPTIMIERUNG DER FUNKTIONELLEN SCREENINGS IN <i>E. COLI</i> UND <i>T. THERMOPHILUS</i> .....        | 32        |
| <b>3 MATERIAL UND METHODEN .....</b>   | <b>33</b> |
| 3.1 VERWENDETE STÄMME UND WACHSTUMSBEDINGUNGEN .....   | 33        |
| 3.1.1 Stämme und Plasmide .....  | 33        |
| 3.1.2 Nährmedien, Zusätze und verwendete Kits.....   | 36        |
| 3.1.3 Kultivierung von Bakterien.....  | 39        |
| 3.1.4 Verwendete Puffermischungen .....  | 39        |
| 3.1.5 Bestimmung der Zelldichte.....   | 40        |
| 3.1.6 Zellernte und Gewinnung von Zellextrakten .....  | 40        |
| 3.1.7 Herstellung von Glycerolstocks .....   | 41        |
| 3.2 MOLEKULARGENETISCHE METHODEN .....   | 41        |
| 3.2.1 Nukleinsäuren Analytik.....  | 41        |



|          |  |           |
|----------|--|-----------|
| 3.2.2    | Molekulare Arbeiten und Modifikationen der DNA .....   | 48        |
| 3.2.3    | Transformation von DNA.....  | 51        |
| 3.2.4    | Konstruktion und funktionelle Analysen von Metagenom-Genbanken .....   | 53        |
| 3.2.5    | Durchmusterungsmethoden .....  | 57        |
| 3.3      | PROTEINBIOCHEMISCHE ANALYSEN .....   | 60        |
| 3.3.1    | Quantifizierung der Proteinkonzentration .....   | 60        |
| 3.3.2    | Aufkonzentrierung und Dialyse von Proteinen.....   | 61        |
| 3.3.3    | Formulierung von Proteinlösungen.....  | 61        |
| 3.3.4    | Heterologische Proteinexpression und Zellaufschluss mit Ultraschall und mechanisch (Bead-beater).....          | 61        |
| 3.3.5    | Proteinaufreinigung .....  | 62        |
| 3.4      | BIOCHEMISCHE CHARAKTERISIERUNG.....  | 67        |
| 3.4.1    | $\beta$ -Glukosidase- und $\beta$ -Galaktosidase Test.....   | 67        |
| 3.4.2    | Bestimmung reduzierter Zuckerenden mittels DNSA-Test.....  | 67        |
| 3.4.3    | para-Nitrophenol (pNP) Test .....  | 69        |
| 3.4.4    | Lipase Assay.....  | 71        |
| 3.4.5    | Qualitative Analyse der Hydrolyseprodukte .....  | 71        |
| 3.4.6    | Analyse mit der HPAE-PAD .....   | 72        |
| 3.4.7    | Zymogramm zur Detektion des enzymatischen Abbaus von Polysacchariden .   | 73        |
| 3.4.8    | Bindungsverhalten von Proteinen mit unlöslichen Polysacchariden .....  | 74        |
| 3.5      | DATENDOKUMENTATION UND -VERARBEITUNG .....   | 74        |
| <b>4</b> | <b>ERGEBNISSE</b> .....  | <b>75</b> |
| 4.1      | GLYKOSIDHYDROLASE ENGU .....   | 75        |
| 4.1.1    | Isolierung und Charakterisierung des thermostabilen EngU Proteins .....  | 75        |
| 4.1.2    | Mutagenese der katalytischen Reste (Nukleophil- und Säure/Base-Rest).....                                      | 78        |
| 4.1.3    | pETDuet-Expressionskonstrukte und trans Komplementation .....  | 80        |
| 4.1.4    | Substratspektrum und Substratspezifität .....  | 83        |
| 4.2      | EXPRESSIONSANALYSE METAGENOMISCHER GENE IN DEN WIRTSORGANISMEN <i>E. COLI</i> UND <i>T. THERMOPHILUS</i> ..... | 88        |

|          |   |            |
|----------|---|------------|
| 4.2.1    | Allgemeine Informationen zu den Fosmiden HA-CMC-1 (CelE12) und TB-EST-7 (Est7) .....  | 88         |
| 4.2.2    | Wachstumskurven der Expressionsstämme.....  | 91         |
| 4.2.3    | Vorbereitungen für die Real-time PCR Läufe .....  | 92         |
| 4.2.4    | HA-CMC-1 und TB-EST-7 qPCR Ergebnisse.....  | 94         |
| 4.3      | BIOCHEMISCHE UNTERSUCHUNGEN DER KODIERTEN ENZYMEN AUF DEN FOSMIDEN PCT3FK-HA-CMC-1 UND PCT3FK-TB-EST-7.....   | 101        |
| 4.3.1    | Endoglukanase Cel12E (Fosmid HA-CMC-1).....   | 101        |
| 4.3.2    | Weitere Untersuchungen von auf TB-EST-7 kodierten Enzymen .....   | 106        |
| 4.3.3    | Rekombinante Herstellung von Est7 und Aktivitätsmessung mit Lipase-Substraten .....   | 106        |
| 4.4      | DURCHMUSTERUNG METAGENOMISCHER GENBIBLIOTHEK.....   | 111        |
| 4.4.1    | Aufbau von pCC1FOS-basierten Metagenom-Genbibliotheken aus „Russia“- und „Mexico“- Umweltproben.....  | 111        |
| 4.4.2    | Übertragung verschiedener Metagenom-Genbibliotheken von <i>E. coli</i> zu <i>T. thermophilus</i> .....  | 114        |
| 4.4.3    | Optimierung des Durchmusterungsprozesses und Steigerung der Transformationseffizienz beim komparativen Screening in <i>E. coli</i> und <i>T. thermophilus</i> ..... | 117        |
| 4.4.1    | In vivo Modifizierungs-Methoden, um Fosmid-Genbibliotheken in andere Expressionswirte ( <i>B. subtilis</i> und <i>M. luteus</i> ) übertragen zu können.....         | 124        |
| 4.4.2    | Bacterial artificial chromosome (BAC) in <i>T. thermophilus</i> .....   | 126        |
| <b>5</b> | <b>DISKUSSION</b> .....   | <b>130</b> |
| 5.1      | EIGENSCHAFTEN DER NEUARTIGEN GLYKOSIDHYDROLASE EngU.....  | 130        |
| 5.1.1    | Sequenz- und Modulvergleiche von EngU.....  | 130        |
| 5.1.2    | Optimierung der EngU-Proteinbiosynthese .....   | 134        |
| 5.1.3    | Putative katalytische Reste des EngU Proteins.....  | 136        |
| 5.1.4    | Substratspektrum und -spezifität von EngU.....  | 137        |
| 5.1.5    | Biotechnologisches Potential von EngU .....   | 139        |
| 5.1.6    | EngU als Prototyp für die neue GH-Familie (GH148) .....   | 139        |

|           |  |            |
|-----------|--|------------|
| 5.2       | EXPRESSIONSANALYSE METAGENOMISCHER GENE IN DEN WIRTSORGANISMEN <i>E. COLI</i> UND <i>T. THERMOPHILUS</i> ..... | 140        |
| 5.2.1     | Vergleich und Interpretation der relativen qPCR Ergebnisse in <i>E. coli</i> ( $\Delta$ Ct)...                 | 143        |
| 5.2.2     | Vergleich und Interpretation der relativen qPCR Ergebnisse in <i>T. thermophilus</i> ( $\Delta$ Ct) .....      | 144        |
| 5.2.3     | Interpretation der $\Delta\Delta$ Ct Ergebnisse zwischen <i>E. coli</i> und <i>T. thermophilus</i> .....       | 145        |
| 5.3       | BIOCHEMISCHE UNTERSUCHUNGEN ZU AUF DEN FOSMIDEN PCT3FK-HA-CMC-1 UND PCT3FK-TB-EST-7 KODIERTEN ENZYMEN .....    | 146        |
| 5.3.1     | Gerichtete Mutagenese des nukleophilen katalytischen Rests von GH12E, um eine Glykosynthase zu erhalten .....  | 146        |
| 5.3.2     | Produktion der thermostabilen Esterase Est7 und ihre Substratspezifität.....                                   | 148        |
| 5.4       | DURCHMUSTERUNG METAGENOMISCHER GENBANKEN.....  | 150        |
| 5.4.1     | Metagenom Isolierungsstrategien .....  | 153        |
| 5.4.2     | Interpretation der 16s rDNA Analyse der „Russia“- und „Mexico“-Umweltproben und der Isolate.....               | 153        |
| 5.4.3     | Funktionelle Screeningmethoden .....   | 156        |
| 5.4.4     | Durchmusterungseffizienz.....  | 158        |
| 5.4.5     | Entwicklung und Evaluierung neuer Expressionswirte.....  | 160        |
| 5.4.6     | BAC-basierte Metagenom-Genbanken.....  | 162        |
| <b>6</b>  | <b>AUSBLICK</b> .....  | <b>164</b> |
| <b>7</b>  | <b>ANHANG</b> .....  | <b>165</b> |
| <b>8</b>  | <b>LITERATURVERZEICHNIS</b> .....  | <b>178</b> |
| <b>9</b>  | <b>PUBLIKATIONEN UND KONFERENZEN</b> .....   | <b>201</b> |
| <b>10</b> | <b>EIGENSTÄNDIGKEITSERKLÄRUNG</b> .....  | <b>202</b> |
| <b>11</b> | <b>DANKSAGUNG</b> .....  | <b>203</b> |

## Abbildungsverzeichnis

|  |     |
|--|-----|
| ABBILDUNG 1: TRANSFER EINER PCC1FOS-METAGENOMBIBLIOTHEK IN VERSCHIEDENE<br>EXPRESSIONSWIRTE.....                       | 23  |
| ABBILDUNG 2: INVERTIERENDER UND RETENDIERENDER MECHANISMUS BEI<br>GLYKOSIDHYDROLASEN .....                             | 29  |
| ABBILDUNG 3: GIBSON ASSEMBLY SCHEMA .....  | 50  |
| ABBILDUNG 4: ENGU ALIGNMENT GEGEN DIE SEEDS DER GLYKOSIDHYDROLASEN-FAMILIEN .....                                      | 77  |
| ABBILDUNG 5: ÜBERBLICK ÜBER ENGU UND DIE PETDUET-EXPRESSIONSKONSTRUKTE .....   | 79  |
| ABBILDUNG 6: SDS-PAGE DER HITZEBEHANDELTEN ENGU-EXPRESSIONSPROTEINE.....   | 81  |
| ABBILDUNG 7: SPALTUNGSMECHANISMUS VON ENGU MIT MIXED-LINKAGE GLUKOTETRAOSE. ....                                       | 83  |
| ABBILDUNG 8: DÜNNSCHICHTCHROMATOGRAPHIE DER ABBAUPRODUKTE DER GLUKOTETRAOSEN<br>.....                                  | 84  |
| ABBILDUNG 9: HPAEC-PAD CHROMATOGRAPHIE DER ABBAUPRODUKTE DER GLUKOTETRAOSEN<br>.....                                   | 84  |
| ABBILDUNG 10: DNSA UND PNP-ASSAY MIT HA-CMC-1 UND TB-EST-7 IN <i>E. COLI</i> UND<br><i>T. THERMOPHILUS</i> .....       | 90  |
| ABBILDUNG 11: WACHSTUMSKURVE MIT DEN FOSMIDEN.....   | 91  |
| ABBILDUNG 12: VERIFIZIERUNG DER PRIMERPAARE HA23 UND TB32 .....  | 92  |
| ABBILDUNG 13: RNA ISOLIERUNG .....   | 93  |
| ABBILDUNG 14: GRAPHIK MIT DEN VERGLEICHENDEN TRANSKRIPTIONSERGEBNISSEN DER 36<br>ORFS DES FOSMIDES HA-CMC-1 .....      | 96  |
| ABBILDUNG 15: GRAPHIK MIT DEN VERGLEICHENDEN TRANSKRIPTIONSERGEBNISSEN DER 16<br>ORFS AUS FOSMID TB-EST- .....         | 97  |
| ABBILDUNG 16: GRAPHISCHE DARSTELLUNG DES CEL12E PROTEINS SOWIE DIE VERKÜRZTE<br>VERSION GH12.....                      | 102 |
| ABBILDUNG 17: SDS-PAGE GELBILDER DES VERKÜRZTEN CEL12E PROTEINS GH12E UND DIE<br>MUTANTEN E171A, E171S UND E171G ..... | 103 |
| ABBILDUNG 18: GLYKOSYNTHESE-MECHANISMUS MIT EINEM MIT FLUOR SUBSTITUIERTEN<br>SUBSTRAT. ....                           | 105 |

|   |     |
|---|-----|
| ABBILDUNG 19: ÜBERSICHT ÜBER DIE EXPRESSION VON EST7 AUS DEN DREI PLASMIDEN<br>PET21A_EST7 .....  | 108 |
| ABBILDUNG 20: LIPASESUBSTRATE VOM KOOPERATIONSPARTNER PROF. DR. J. PIETRUSZKA .   | 110 |
| ABBILDUNG 21: GELELEKTROPHORESE DER UMWELTPROBEN „MEXICO“ UND „RUSSIA“ .....  | 112 |
| ABBILDUNG 22: SCHEMA ZUR IN VIVO FOSMID-MODIFIZIERUNG. ....   | 116 |
| ABBILDUNG 23: GELBILD VON DEN VIER VERSCHIEDENEN EXTRAKTIONSPROTOKOLLEN .....   | 118 |
| ABBILDUNG 24: DURCHFÜHRUNG DES POOL-SCREENINGS .....  | 120 |
| ABBILDUNG 25: SUBSTRATPLATTEN LB-XYLAN ODER LB-CMC MIT DEM B8 KLON UND DEM POOL<br>M12-22 .....   | 121 |
| ABBILDUNG 26: AUF TRIBUTYRIN AKTIVER KLON (L01) AUS DER HA- UND TB-FOSMIDBANK.....  | 123 |
| ABBILDUNG 27: VEKTORKARTE VON PRETRO-B UND PRETRO-M .....   | 125 |
| ABBILDUNG 28: VEKTORKARTEN VON PRETRO-BAC1 UND PRETRO-BAC2 .....  | 127 |
| ABBILDUNG 29: SOUTHERN BLOT DER PRETROBAC1 UND BAC2 KONSTRUKT. ....   | 128 |
| ABBILDUNG 30: SOUTHERN BLOT MIT GDNA AUS ZELLEN, DIE BEI STEIGENDER KANAMYCIN<br>KONZENTRATION KULTIVIERT WURDEN.....                         | 129 |
| ABBILDUNG 31: ECKHARDTS GEL AND PFGE-GEL ZUR VISUALISIERUNG DER PRETRO-BAC1 UND<br>PRETRO-BAC2 VEKTOREN IN <i>T. THERMOPHILUS</i> HB27WT..... | 177 |

## Tabellenverzeichnis

|  |    |
|--|----|
| TABELLE 1: WEITERE BEISPIELE VON INDUSTRIE-RELEVANTEN THERMOSTABILEN ENZYMEN .....               | 25 |
| TABELLE 2: THERMOSTABILE METAGENOMISCHE ENZYME AUS FUNKTIONELLEN SCREENINGS ..                   | 27 |
| TABELLE 3: VERWENDETE STÄMME .....   | 33 |
| TABELLE 4: VERWENDETE PLASMIDE .....   | 34 |
| TABELLE 5: ZUSAMMENSETZUNG LB- UND TB-MEDIUM .....   | 37 |
| TABELLE 6: SH MINIMALMEDIUM .....  | 37 |
| TABELLE 7: HERSTELLUNG VON YENB, SOB UND SOC MEDIUM .....  | 38 |
| TABELLE 8: VERWENDETE ANTIBIOTIKA UND MEDIENZUSÄTZE .....  | 38 |
| TABELLE 9: ZUSAMMENSETZUNG DER VERWENDETEN PUFFER.....   | 39 |
| TABELLE 10: PUFFERZUSAMMENSETZUNG FÜR DIE ISOPROPANOLFÄLLUNG .....                               | 42 |
| TABELLE 11: PUFFER UND PARAMETER FÜR DIE PULSE FIELD GELELEKTROPHORESE .....                     | 43 |
| TABELLE 12: qPCR PROBENANSATZ.....   | 47 |
| TABELLE 13: qPCR PROGRAMM FÜR CDNA, PLASMID DNA UND GENOMISCHE DNA .....                         | 47 |
| TABELLE 14: PCR-ANSATZ FÜR Q5 HIGH-FIDELITY POLYMERASE.....                                      | 48 |
| TABELLE 15: PCR LAUFPROGRAMM .....   | 49 |
| TABELLE 16: ANSATZ FÜR EINEN RESTRIKTIONSVERDAU.....   | 49 |
| TABELLE 17: ÜBERSICHT DER UNTERSUCHTEN METAGENOM-GENBANKEN.....                                  | 53 |
| TABELLE 18: ZUSAMMENSETZUNG DER SUBSTRAT-AGARPLATTEN .....                                       | 57 |
| TABELLE 19: PARAMETER UND SÄULEN FÜR DIE FPLC-LÄUFE .....  | 63 |
| TABELLE 20: ZUSAMMENSETZUNG FÜR 4 SDS-GELE.....  | 65 |
| TABELLE 21: ANSÄTZE FÜR DEN DNSA-ASSAY .....   | 68 |
| TABELLE 22: LISTE DER PNP-SUBSTRATE .....  | 70 |
| TABELLE 23: ANSÄTZE FÜR DEN PNP-ASSAY .....  | 70 |
| TABELLE 24: LAUF- UND FÄRBE MITTELZUSAMMENSETZUNG FÜR DIE<br>DÜNN SCHICHTCHROMATOGRAPHIE.....    | 72 |
| TABELLE 25: GRADIENTENLAUF FÜR DIE HPAE-PAD .....  | 73 |
| TABELLE 26: CELLULASE-AKTIVITÄTEN MUTIERTER ENGU VARIANTEN IM VERGLEICH ZUM ENGU<br>WILDTYP..... | 78 |

|   |     |
|---|-----|
| TABELLE 27: ÜBERBLICK ÜBER DIE GETESTETEN SUBSTRATE .....   | 85  |
| TABELLE 28: ÜBERBLICK ÜBER DIE FOSMIDE HA-CMC-1 UND TB-EST-7 .....  | 89  |
| TABELLE 29: ÜBERSICHT ZU DEN ANALYSIERTEN ORFS DER HA-CMC-1 UND TB-EST-7 FOSMIDE<br>MIT DEN DAZUGEHÖRIGEN CT-WERTEN. ....   | 98  |
| TABELLE 30: REINIGUNG VON GH12E, EINER VERKÜRZTEN VERSION VON CEL12 .....   | 102 |
| TABELLE 31: ZUSAMMENFASSUNG DER ÜBER 16SRNA IDENTIFIZIERTEN PHYLA AUS DEN<br>UMWELTPROBEN „MEXICO“ UND „RUSSIA“ .....   | 113 |
| TABELLE 32: AKTIVITÄT ANALYSE DER ANREICHERUNGSKULTUREN AUS DEN UMWELTPROBEN<br>„MEXICO“ UND RUSSIA“ .....  | 114 |
| TABELLE 33: ZUSAMMENFASSUNG DER VIER PROTOKOLLVARIATIONEN ZUR FOSMIDAUFREINIGUNG<br>IM 96-WELL FORMAT UND IHRE TRANSFORMATIONSEFFIZIENZ IN <i>T. THERMOPHILUS</i> ..... | 118 |
| TABELLE 34: ZUSAMMENFASSUNG DER FOSMID-BANKEN UND DEN DURCHFORSTETEN FOSMIDE IN<br><i>E. COLI</i> UND <i>T. THERMOPHILUS</i> .....                                      | 122 |
| TABELLE 35: ÜBERSICHT DER KOMPONENTEN DER IN VIVO MODIFIZIERUNG DES LEERVEKTORS<br>PCC1FOS. ....  | 125 |
| TABELLE 36: ÜBERBLICK ÜBER DIE BESTEN TREFFER DER BLAST ANALYSE ZUR SUCHE NACH<br>PROTEINSEQUENZEN MIT ÄHNLICHKEITEN ZU DER GESAMTEN ENGÜ LÄNGE. ....                   | 130 |
| TABELLE 37: ÜBERBLICK ÜBER DIE ORF GENE UND IHRE $\Delta$ CT-WERTE IN JEWELIGEN<br>EXPRESSIONSWIRT .....  | 141 |
| TABELLE 38: PRIMERLISTE DER ENGÜ EXPRESSIONSKONSTRUKTE. ....  | 165 |
| TABELLE 39: PRIMERLISTE DER TRANSKRIPTOMANALYSE DES PCT3FK-HA-CMC-1 FOSMIDS .   | 169 |
| TABELLE 40: PRIMERLISTE DER TRANSKRIPTOMANALYSE DES PCT3FK-TB-EST-7 FOSMIDS ..  | 173 |
| TABELLE 41: GEGEBÜBERSTELLUNG VON CEL12E UND GH12E.....   | 176 |

## Abkürzungs- und Einheitenverzeichnis

|                          |  |
|--------------------------|--|
| ::                       | Insertion in der Gensequenz  |
| °C                       | Grad Celsius   |
| x g                      | Vielfaches der Erdbeschleunigung   |
| µ-                       | Mikro- ( $1 \times 10^{-6}$ )  |
| A                        | Ampere   |
| AA                       | Aminosäure/n (engl.: <i>amino acid/s</i> , AS)   |
| AIS                      | Autoinduction Solution (Autoinduktionslösung)  |
| Amp                      | Ampicillin   |
| APS                      | Ammoniumpersulfat  |
| BAC                      | <i>Bacterial Artificial Chromosome</i>   |
| BCI-                     | 5-Bromo-4-Chloro-3-Indoxyl-  |
| BH                       | BL03_pCT3FK-HA-CMC-1   |
| Bidest. H <sub>2</sub> O | Deionisiertes Wasser   |
| bp                       | Basenpaar  |
| BP                       | BL03_pCT3FK  |
| BT                       | BL03_pCT3FK-TB-EST-7   |
| c-                       | Centi- ( $1 \times 10^{-2}$ )  |
| C-terminal               | Carboxyterminal  |
| ca.                      | circa  |
| Car                      | Carotinoid-Cluster, Teil des Carotinoid-Biosynthese ( <i>crt</i> ) Clusters in <i>T. thermophilus</i> (kodierende <i>car</i> -Gene sind TT_P0055 bis TT_P0058) |
| CBM                      | Kohlenhydrat-Bindemodul  |
| CE                       | <i>crude extract</i> (deutsch: Rohextrakt)   |
| Cm / Chl                 | Chloramphenicol  |
| CMC                      | Carboxymethyl-Cellulose  |
| CV                       | <i>column volume</i> (deutsch: Säulenvolumen)  |
| d.h.                     | das heißt  |
| Da                       | Dalton   |
| ddH <sub>2</sub> O       | doppelt demineralisiertes, deionisiertes Wasser  |
| DMSO                     | Dimethylsulfoxid   |
| DNA                      | <i>Desoxyribonucleic acid</i> (deutsch: Desoxyribonukleinsäure)  |
| DNSA                     | Dinitrosalicyl acid (deutsch: Dinitrosalicylsäure)   |
| dNTP                     | Desoxyribonukleotidtriphosphat   |



|               |  |
|---------------|--|
| EC            | Nummer der Enzym-Kommission  |
| EDTA          | Ethylendiamintetraacetat   |
| EH            | EPI300_pCT3FK-HA-CMC-1   |
| EP            | EPI300_pCT3FK  |
| ET            | EPI300_pCT3FK-TB-EST-7   |
| <i>et al.</i> | und weitere ( <i>et alter</i> )  |
| FPLC          | <i>Fast Protein Liquid Chromatography</i>  |
| g             | Gramm  |
| gDNA          | genomische DNA   |
| GC            | Prozentualer Gehalt an Guanin und Cytosin in DNA: $(G+C)/(A+T+G+C) \times 100$                                       |
| GO            | <i>Gene ontology</i> , Vereinheitlichung der Annotationen von Genen und Genprodukten, Ashburner <i>et al.</i> (2000) |
| h             | Stunde   |
| His-tag       | Hexahistidin(6xHis)-tag  |
| HT-SN         | <i>heat-treated supernatant</i> (deutsch: Hitze-behandelter Überstand)   |
| IMAC          | Immobilisierte Metallionen-Affinitäts-Chromatographie  |
| IPTG          | Isopropyl- $\beta$ -D-thiogalactopyranosid   |
| k-            | Kilo- ( $1 \times 10^3$ )  |
| <i>kat</i>    | Kanamycin-Resistenzkassette  |
| kb            | Kilobasenpaare   |
| KBE           | Kolonie bildende Einheit   |
| L             | Liter  |
| LB            | <i>lysogeny broth</i> (auch als Luria-Bertani Medium bekannt)  |
| m-            | Milli- ( $1 \times 10^{-3}$ )  |
| M-            | Mega- ( $1 \times 10^6$ )  |
| m             | Meter  |
| M             | Molar ( $\text{mol} \times \text{L}^{-1}$ )  |
| MCS           | Multiple Klonierungsstelle ( <i>multiple cloning site</i> )  |
| MES           | 2-(N-Morpholino)ethansulfonsäure   |
| MQ            | Milli-Q-Wasser   |
| min           | Minute(n)  |
| n-            | Nano- ( $1 \times 10^{-9}$ )   |
| N-terminal    | aminoterminal  |
| NCBI          | <i>National Center for Biotechnology Information</i>   |
| nt            | Nukleotide   |
| o.g.          | oben genannt   |

|                  |  |
|------------------|--|
| OD <sub>nm</sub> | Optische Dichte bei einer bestimmten Wellenlänge (nm)              |
| ORF              | Offener Leserahmen ( <i>open reading frame</i> )                   |
| P <sub>X</sub>   | Promotor des Gens X  |
| PAGE             | Polyacrylamid Gelelektrophorese                                    |
| PCR              | Polymerase-Kettenreaktion ( <i>polymerase chain reaction</i> )     |
| pH               | negativer dekadischer Logarithmus der Protonenkonzentration        |
| pNP-<br>R        | <i>para</i> -Nitrophenyl-<br>Resistenz                             |
| RBS              | <i>Ribosom Binding Site</i> (deutsch: Ribosomale Bindestelle)      |
| RNA              | <i>Ribonucleic acid</i> (deutsch: Ribonukleinsäure)                |
| rpm              | Rounds per minutes   |
| RT               | Raumtemperatur   |
| s                | Sekunde  |
| SDS              | Natriumlaurylsulfat  |
| SN               | <i>Supernatant</i> (deutsch: Überstand)                            |
| SOB              | Super Optimal Broth  |
| SOC              | Super Optimal Broth mit Glukose                                    |
| SOE-PCR          | <i>Splicing by overlap extension</i> PCR (Heckman and Pease, 2007) |
| sp.              | Spezies  |
| TB               | <i>Thermus broth</i>   |
| TEMED            | Tetramethylethylendiamin   |
| Tris             | Tris-(hydroxymethyl)-aminomethan                                   |
| U                | Einheit der Enzymaktivität ( <i>Unit</i> )                         |
| u.a.             | unter anderem  |
| ÜN               | über Nacht   |
| Upm              | Umdrehungen pro Minute   |
| UV               | Ultraviolett   |
| V                | Volt   |
| VE               | Vollentsalztes Wasser  |
| v/v              | Volumen pro Volumen  |
| w/v              | Gewicht pro Volumen  |

# 1 Einleitung

## 1.1 „Metagenomik“ - Die Suche nach der Stecknadel im metagenomischen Heuhaufen

Mikroorganismen leben seit 3,8 Milliarden Jahren auf unserer Erde und haben sich aufgrund ihrer schnellen Evolution zu einer großen Zahl (ca. 4 bis  $6 \times 10^{30}$  Zellen) und Artenvielfalt diversifiziert (Simon and Daniel, 2009a; Whitman et al., 1998). Nicht nur die Diversität der Arten, sondern auch die extremen Habitats, die sie besiedeln können, sind beeindruckend. Dank ihrer schnellen Anpassungsfähigkeit überstehen sie widrige Umwelteinflüsse wie extreme Temperaturen, pH-Werte, Druck, Strahlung, hohe Salzkonzentrationen, Metallkontaminationen und Nährstoffmangel (Amann et al., 1995; Simon and Daniel, 2009a; Whitman et al., 1998). Seit Beginn der Kultivierung von Mikroorganismen ist es bis zur heutigen Zeit gelungen nur einen winzigen Anteil dieser Vielfalt im Labor zu züchten (Amann et al., 1995; Torsvik et al., 1990). Dieses Problem kann umgangen werden, indem man die Genome direkt aus Umweltproben isoliert und mit dem heterogenen Erbgut weiter arbeitet (Moran et al., 1993). Zunächst muss die DNA aus den Proben isoliert, in Vektoren kloniert und in geeignete Expressionswirte übertragen werden. Allein diese drei Schritte zur Vorbereitung einer metagenomischen Genbank sind schon mit Problemen bzw. Limitierungen behaftet (Ferrer et al., 2005). Trotz der Suche nach der sprichwörtlichen Stecknadel im metagenomischen Heuhaufen („Finding the Needles in the Metagenome Haystack“, Kowalchuk et al., 2007) ist es möglich auf diese Art Gene für neue Biokatalysatoren zu finden.

Das Durchforsten (engl.: *Screening*) der Genbibliotheken wird in zwei Kategorien eingeteilt: die sequenzbasierte und funktionsbasierte Methode. Für die sequenzabhängige Variante wird die metagenomische DNA zunächst sequenziert und die Funktion auf Grundlage von Ähnlichkeitsvergleichen mit bereits bekannten Genen, Proteindomänen oder Motiven vorhergesagt (Kim et al., 2007). Der Vorteil liegt darin, dass mit den neuen Sequenzierungsmethoden große Genbanken schnell und einfach durchforstet werden können. Der große Nachteil hingegen ist die Tatsache, dass viele neuartige Enzyme unentdeckt bleiben, da die Sequenzen nur mit bereits charakterisierten/annotierten Proteinen in den Datenbanken abgeglichen werden können (Ferrer et al., 2012). Das funktionsabhängige Verfahren setzt eine erfolgreiche heterologe Expression der metagenomischen Gene im Wirt voraus. Drei mögliche funktionsbasierte Methoden sind bekannt (Ekkers et al. 2012): (i) phänotypische Detektion: Enzymaktivität wird durch Indikatorreagenzien nachgewiesen (Beloqui et al., 2010; Gloux et al., 2011; Liaw, 2010). (ii) modulierte Detektion: Verwendung von Reportersystemen oder heterologe Komplementation (Leis, 2015a; Simon and Daniel,

2009a) (iii) Substrat- oder Produkt-induziertes Genexpressionsscreening (Uchiyama and Miyazaki, 2010; Williamson et al., 2005).

Neben den unterschiedlichen Screeningmethoden gibt es verschiedene Genbanken, die sich in der Größe des metagenomischen Inserts unterscheiden lassen. Die kleinsten Inserts tragen Standard-Plasmidvektoren mit einer maximalen Kapazität von etwa 15 kb, gefolgt von Cosmiden und Fosmiden mit etwa 30 - 42 kb (Lam et al., 2014; Leis et al., 2015a). Einen großen Sprung machen hier die „Bacterial Artificial Chromosomes (BACs)“ und „Yeast Artificial Chromosomes (YACs)“, die 350 kb bzw. 3000 kb der metagenomischen DNA aufnehmen können (Kakirde et al., 2010, 2012). Je größer die Genbank ist, desto weniger Klone müssen getestet werden, um die gesamte Umweltprobe abzudecken und desto größer wird die Wahrscheinlichkeit, ganze Gencluster und Stoffwechselwege zu finden (Daniel, 2005). Ein wesentlicher Punkt bei dem funktionellen Screening ist die erfolgreiche Expression der Gene. Dafür müssen die Vektoren in geeigneten Expressionsstämmen kultiviert werden, d.h. die Proteinbiosynthese-Mechanismen im Stamm sollten mit den zu exprimierenden metagenomischen Genen zusammenpassen. Beispielsweise muss der Wirt in der Lage sein, die heterologen Transkriptionssignale des Fremdgens zu nutzen (Warren et al., 2008), eine ausreichende mRNA-Stabilität zu gewährleisten (Kudla et al., 2009), für die nachfolgende Translation entsprechende Ribosom-Bindestellen (RBS), die für die Ribosomenbindung und Translationsinitiation notwendig sind, zu erkennen und für die korrekte Proteinfaltung entsprechende Chaperone bereitzustellen (Culligan et al., 2014; Hidalgo et al., 2004; Leis, 2015a; Mujacic et al., 2003; Zerbs et al., 2014).

## 1.2 Unterschiedliche Expressionswirte für die Suche nach Enzymen aus Extremophilen

In den meisten funktionsbasierten Metagenom-Studien wurde *E. coli* als heterologer Expressionswirt verwendet (Ekkers et al., 2012; Simon and Daniel, 2011; Taupp et al., 2011). Der Gram-negative, stäbchenförmige und fakultativ anaerobe Modellorganismus ist gut untersucht und aufgrund der für ihn etablierten Methoden und für ihn vorhandenen Mutantenstämme und Plasmide einfach im Labor zu handhaben (Helena and Hård, 2002; Sastry et al., 2004; Terpe, 2003; Zerbs et al., 2014). Dennoch ist *E. coli* nicht unbedingt die beste Wahl, wenn es um die Untersuchung von extremophilen Metagenomproben geht. Bei einer vergleichenden Durchmusterung einer mit genomischer DNA aus *Spirochaeta thermophila* hergestellten Fosmid-Genbibliothek in *E. coli* und *T. thermophilus* wurde ein Klon mit thermostabiler Xylanaseaktivität in *T. thermophilus*, jedoch nicht in *E. coli*, gefunden (Angelov et al., 2009). Dies zeigt, dass ein großer Bedarf an neuen und vielseitigen Expressionswirten besteht, um das ganze Potential der Metagenom-Genbibliotheken auszuschöpfen.

Der Organismus *T. thermophilus* wurde aus heißen Quellen isoliert und als ein Gram-negatives, nicht Sporenbildendes und nicht motiles Bakterium beschrieben (Oshima and Imahori, 1974), welches sich als neuer Expressionswirt sehr gut eignet (Liebl et al., 2014b). Phylogenetisch ist *Thermus* mit der Gattung *Deinococcus* eng verwandt. Sein Wachstumsoptimum liegt in einem Temperaturbereich zwischen 62 - 75 °C und die Generationszeit beträgt nur 40 Minuten. Die Genome der Stämme HB8 und HB27 wurden von Masui (2005) und Henne (2004) sequenziert (Genbank: AP008226.1 und AE017221.1). Das HB27-Chromosom ist 1,89 Mbp groß und die Zellen besitzen ein Megaplasmid (pTT27) mit 232 kbp, zudem ist *Thermus* polyploid (Ohtani et al., 2010). Der GC-Gehalt beträgt 69,4 mol % und etwa 95 % des Genoms kodieren für ca. 2200 Proteine (Liebl, 2004). *Thermus* ist natürlich kompetent und seine Transformationshäufigkeit mit chromosomaler DNA liegt bei  $10^{-1}$  (Hidaka et al., 1994; Koyama et al., 1986; Schwarzenlander and Averhoff, 2006). Weiterhin gibt es Plasmide und genetische Werkzeuge, wie Selektionsmarker und steuerbare Promotoren für *Thermus spp.* (Koyama 1990, Koyama 1990). Zum Beispiel kann der modifizierte HB27::nar Stamm mit dem Nar-Operon zur induzierbaren Proteinexpression in *Thermus* eingesetzt werden (Moreno et al., 2005; Ramírez et al., 2000; Zafra et al., 2002). Das Shuttlevektor-System ermöglicht die Übertragung einer auf pCC1FOS-basierenden Metagenom-Genbank von *E. coli* auf *T. thermophilus* (Angelov et al., 2009). Weiterhin sind von den thermostabilen Proteinen dieses Bakteriums schon viele Proteinkristallstrukturen ermittelt worden (Bae et al., 2015; Chang et al., 2013; Miyazaki et al., 2015). Aufgrund seiner extrem thermophilen

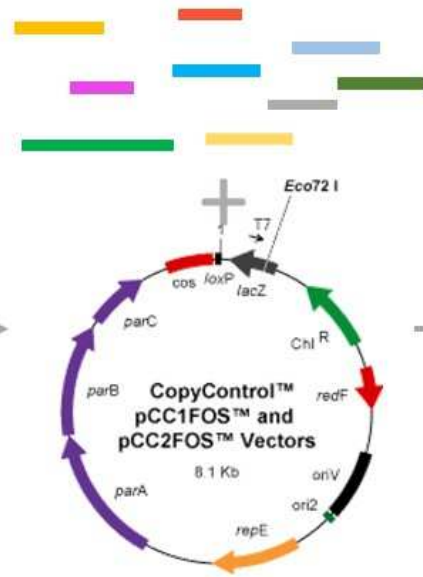
Eigenschaften ist es naheliegend, dass *T. thermophilus* ein vielversprechender Expressionswirt für Gene aus Höchsttemperatur-Umweltproben sein könnte.

Weitere ebenfalls natürlich kompetente Bakterien sind *Micrococcus luteus* und *Bacillus subtilis*. Erstgenannter ist ein Gram-positives, Kokken-bildendes und strikt aerobes Bakterium. Sein Genom besitzt einen hohen GC-Gehalt (73 mol%) und als Selektionsmarker können Erythromycin, Kanamycin und Hygromycin verwendet werden (Angelov et al., 2015; Liebl et al., 2002). Im Labor ist das Arbeiten mit *Micrococcus* unkompliziert und im Großmaßstab möglich. *B. subtilis* ist ein Gram-positives Bakterium mit niedrigem GC-Gehalt (43,52 mol%) und außerdem ein Modellorganismus für den Zelldifferenzierungsvorgang der Sporulation. Alle drei Mikroorganismen besitzen unterschiedliche Eigenschaften, die für die funktionelle Durchforstung metagenomischer Genbanken von Vorteil sein können. Nun sollten Mobilisierungsverfahren entwickelt werden, wie bereits bei *T. thermophilus* angewandt, die eine Metagenom-Genbibliothek für mehrere Expressionswirte zugänglich machen. In Kombination mit der Entwicklung und Anwendung neuer oder modifizierter Substrate können derartige alternative Wirtsorganismen dafür eingesetzt werden, um neue attraktive Biokatalysatoren für die weiße und rote Biotechnologie bereitzustellen.

## Umweltproben

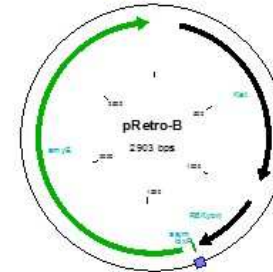
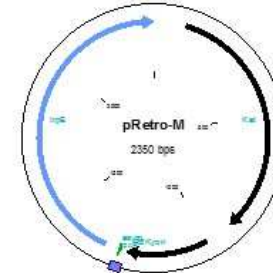
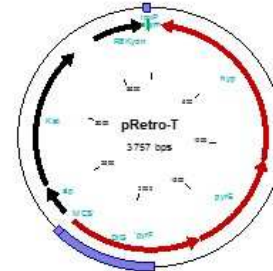


## Metagenombank



in *E. coli* als primärer Wirt

## In vivo Modifikation



## Expressionswirte

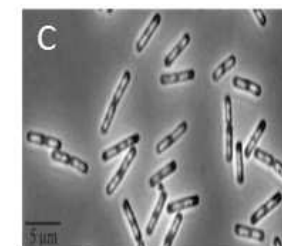
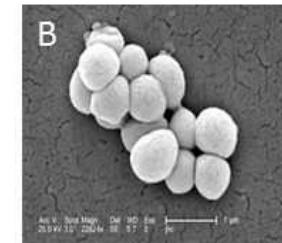
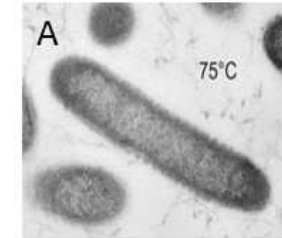


Abbildung 1: Transfer einer pCC1FOS-Metagenombibliothek in verschiedene Expressionswirte. *T. thermophilus* (A), *M. luteus* (B) und *B. subtilis* (C). Aus den Umweltproben wird die metagenomische DNA isoliert, die auf eine Größe von ca. 40 kb verkleinert wird. Die Fragmente werden in den CopyControl™ pCC1FOS Vektor reinkloniert, um eine Fosmid-Genbank zu erstellen. Die Genbanken können durch die verschiedenen wirtsspezifischen pRetro-Vektoren und mithilfe der Cre-Rekombinase aus dem pJM2545 Plasmid modifiziert werden. Nach der in vivo Modifizierung besitzen die Fosmide neben dem metagenomischen Insert, einem Selektionsmarker auch eine homologe Sequenz (rot = pyr Gen von *T. thermophilus*; blau = trpE Gen von *M. luteus*, grün = amyE Gen von *B. subtilis*), die es ihnen ermöglicht, in die Genome der einzelnen Wirte zu integrieren.

### 1.3 Metagenomik kombiniert mit den Entwicklungen des Next Generation Sequencings (NGS)

Die ständig weiter entwickelten Sequenzieretechniken ermöglichen und erleichtern die Arbeiten in der Metagenomforschung. Mit der Analyse der 16S rRNA kann man die verwandtschaftliche Beziehung der Organismen kultivierungsunabhängig erfassen und die ökologischen Funktionen innerhalb einer mikrobiellen Lebensgemeinschaft genauer betrachten (Handelsman et al., 1998; Lane et al., 1985; Healy, 1995; Schmidt et al., 1991). Durch Schrotschuss-Sequenzierung, bei der die DNA vervielfältigt und in zufällige 300 bis 1000 bp lange Fragmente zerlegt, sequenziert und anschließend bioinformatisch wieder zusammengesetzt wird, können umfangreiche Sequenzdatensätze metagenomischer DNA erzeugt werden (Handelsman, 2004; Venter, 2004). Die dabei manchmal eingesetzte Multidisplacement Amplifikation (engl.: *multiple displacement amplification*, MDA) läuft ähnlich wie eine zelluläre DNA-Replikation ab und benötigt randomisierte, am 3' Ende Phosphorothioat-verlinkte Oligos und die  $\phi$ 29 Polymerase, die eine Strangverdrängung (engl.: *strand displacement*) initiieren kann. Die Polymerase erkennt einen Primer und vervollständigt den Strang, bis sie auf den downstream Primer trifft, und bildet so eine Vergabelung, so dass die Einzelstrang-DNA freigelegt wird und für die nächste Amplifikation bereit steht. Dadurch entstehen viele Vergabelungen und es können große DNA-Fragmente von 12 bis 100 kb generiert werden (Dean et al., 2001; Hosono et al., 2003). Diese Technik kommt ebenfalls bei der Sequenzierung der Genome von einzelnen Zellen zum Einsatz (Lasken, 2009). Mit der Kombination aus den Daten der „shotgun“ und „single cell“ –Sequenzierung können die identifizierten Gene aus einer Metagenom-Genbank einer Spezies zugeordnet werden (Eloe et al., 2011). Dies führt zu neuen Erkenntnissen über die Synergien der durchmischten Mikroorganismen-Population, ihre extremen Lebensräume und die dazu benötigten Enzyme.

### 1.4 Der Nutzen von thermostabilen Enzymen für biotechnologische Prozesse

Viele verschiedene Industriezweige haben großes Interesse an thermostabilen Enzymen, die biotechnologische Prozesse selektieren, beschleunigen oder gar erst ermöglichen. Durch die erhöhten Temperaturen sind auch weniger Kontaminationen mit anderen mesophilen Mikroorganismen möglich und man erreicht eine bessere Löslichkeit von Substraten bzw. Produkten sowie die Erhöhung der Diffusionsraten und die Verringerungen der Viskosität (Sternler and Liebl, 2001). Darüber hinaus sind hitzeresistente Biokatalysatoren robuster und über mehrere Zyklen wieder einsetzbar, was zu einer Zeit- und Kostenersparnis der Produktionsprozesse führt (Littlechild, 2015).



Tabelle 1: Weitere Beispiele von Industrie-relevanten thermostabilen Enzymen

| Enzym     | Organismus                            | Opt. Temp. [°C] | Referenz                       |
|-----------|---------------------------------------|-----------------|--------------------------------|
| α-Amylase | <i>Bacillus cereus</i>                | 65              | (Annamalai et al., 2011)       |
|           | <i>Bacillus subtilis</i>              | 55              | (Das et al., 2004)             |
|           | <i>Pyrococcus furiosus</i>            | 100             | (Laderman et al., 1993)        |
|           | <i>Paecilomyces variotii</i>          | 60              | (Michelin et al., 2010)        |
|           | <i>Rhodothermus marinus</i>           | 85              | (Gomes et al., 2003)           |
| Cellulase | <i>Anaerocellu thermophilum</i>       | 85-90           | (Zverlov et al., 1998a)        |
|           | <i>Bacillus subtilis</i>              | 65-70           | (Mawadza et al., 2000)         |
|           | <i>Pyrococcus furiosus</i>            | 102-105         | (Kengen et al., 1993)          |
|           | <i>Bacillus vallismortis</i>          | 65              | (Gaur and Tiwari, 2015)        |
|           | <i>Thermobifida halotolerans</i>      | 55              | (Yin et al., 2015)             |
| Esterase  | <i>Bacillus acidocaldarius</i>        | 70              | (Manco et al., 1998)           |
|           | <i>Streptomyces lividans</i>          | 50              | (Wang et al., 2015)            |
| Lipase    | <i>Bacillus sp.</i>                   | 70              | (Shariff et al., 2011)         |
|           | <i>Geobacillus thermodentrificans</i> | 65              | (Balan et al., 2012)           |
|           | <i>Pseudomonas sp.</i>                | 65              | (Yang et al., 2015)            |
| Proteasen | <i>Bacillus subtilis</i>              | 60              | (Adinarayana et al., 2003)     |
|           | <i>Bacillus stearothermophilus</i>    | 70-90           | (Sookkheo et al., 2000)        |
|           | <i>Geobacillus toebii</i>             | 90              | (Thebti et al., 2016)          |
|           | <i>Pyrococcus sp.</i>                 | 110             | (Morikawa et al., 1994)        |
| Xylanase  | <i>Bacillus stearothermophilus</i>    | 60              | (Nanmori et al., 1990)         |
|           | <i>Pyrococcus furiosus</i>            | 100             | (Bauer et al., 1999)           |
|           | <i>Thermatoga maritima</i><br>MSB8    | 92              | (Winterhalter and Liebl, 1995) |
|           | <i>Thermotoga neapolitana</i>         | 102             | (Zverlov et al., 1996)         |

Die Kraftstoffindustrie bewegt sich immer weiter von den Erdölressourcen weg und hin zu den erneuerbaren Energieträgern. In der zweiten Generation der Biokraftstoffe werden nicht mehr Zuckerrohr oder Mais als Ausgangsmaterialien verwendet, sondern cellulose- oder lignocellulosehaltige Rohstoffe. Cellulose ist der Hauptbestandteil der pflanzlichen Zellwände und folglich der weltweit häufigste nachwachsende Rohstoff. Das unverzweigte Polymer besteht aus linearen Ketten von ca. 15.000 D-Glukoseeinheiten, die durch  $\beta$ -1,4-glykosidische Verbindungen verknüpft sind. Die Glukoseringe in der Glukankette sind um  $180^\circ$  verdreht und ermöglichen dadurch Wasserstoffbrückenbindungen zwischen der OH-Gruppe des C3 und dem Ringsauerstoff des folgenden Restes. Stapel aus Lagen von  $\beta$ -Glukanketten werden wiederum über Van-der-Waals-Wechselwirkungen und intermolekulare Wasserstoffbrücken zusammengehalten. Diese komplexe Matrix gibt der Cellulose ihre starke Zugfestigkeit und hohe Belastbarkeit, erschwert jedoch den Abbau durch die drei abbauenden Enzym-Komponenten: (i) Endoglukanasen (1,4(1,3;1,4)- $\beta$ -D-Glukan-4-Glukanhydrolase, EC 3.2.1.4) schneiden zufällig innerhalb der amorphen Bereiche langer Celluloseketten (ii) Exoglukanasen (1,4- $\beta$ -D-Glukan-Cellobiohydrolase, EC 3.2.1.91) spalten Cellobiose oder Glukose ab (iii)  $\beta$ -Glukosidasen (EC 3.2.1.21) hydrolysieren die gelösten Cellodextrine und Cellobiose in Glukose (Li et al., 2011; Parry et al., 2001). Es ist wichtig, neue thermostabile Cellulasen zu finden (Hong et al., 2003; Leis et al., 2015b; Ruttersmith and Daniel, 1991; Tuohy et al., 2002), die die biotechnologischen Prozesskosten senken können.

Eine weitere Gruppe biotechnologisch bedeutsamer Enzyme sind lipolytische Enzyme. Diese gehören zu den alpha/beta Hydrolasen und werden in zwei Gruppen, die Lipasen (EC 3.1.1.3) und die Carboxylesterasen (EC 3.1.1.1), eingeteilt. Der Unterschied liegt in ihrer Substratpräferenz: Esterasen spalten bevorzugt lösliche, kurzkettige Substrate mit weniger als 12 Kohlenstoffatomen, während Lipasen die langkettigen ( $\geq 12$  Kohlenstoffatome) Fettsäureester mit geringer Löslichkeit spalten. Beide hydrolysieren in Anwesenheit von Wasser Esterbindungen. Ihr Aktivitätszentrum ist charakterisiert durch eine katalytische Triade (Serin, Aspartat/Glutamat und Histidin) mit dem Konsensusmotiv Gly-X-Ser-X-Gly. Zunächst bindet das katalytische Serin an das Carbonylkohlenstoffatom der Lipidesterbindung und bildet dadurch ein tetraedrisches Intermediat, welches durch die katalytischen Reste Histidin und Aspartat/Glutamat stabilisiert wird. Die alkoholische Komponente der Esterbindung wird abgespalten und es bildet sich eine neue Esterbindung aus der Säurekomponenten mit dem katalytischen Serin. Daraufhin hydrolysiert ein Wassermolekül das kovalente Intermediat, wobei der Acylrest freigesetzt wird. Bei einer Lipase-katalysierten Veresterungsreaktion wird das Wasser bei der Reaktion durch einen Alkohol oder einen Ester ersetzt (Jaeger et al., 1999). Es konnten bereits viele hitzestabile Esterasen/Lipasen aus diversen Metagenom-Genbibliotheken isoliert werden wie z.B. aus heißen Quellen (Rhee et al., 2005; Tirawongsaroj

et al., 2008), Kompost (Kang et al., 2011), Erd- und Meeresproben (Hu et al., 2010; Ko et al., 2012).

Man kann neue thermostabile Enzyme über Protein-Engineering herstellen, indem man *bereits* bekannte und sequenzierte Proteinsequenzen modifiziert oder man durchforstet Metagenom-Genbanken nach unbekanntem hitzestabilen Enzymen. Die Cellulase und Xylanase aus der Tabelle 2 sind Beispiele für thermostabile Proteine, die durch funktionelles Screening am Lehrstuhl für Mikrobiologie an der Technischen Universität München entdeckt wurden.

*Tabelle 2: Thermostabile metagenomische Enzyme aus funktionellen Screenings*

| <b>Enzym</b>      | <b>Umweltprobe</b>    | <b>Temperaturen [°C]</b> | <b>Referenz</b>         |
|-------------------|-----------------------|--------------------------|-------------------------|
| $\alpha$ -Amylase | Boden und Wasser      | 95                       | Richardson et al., 2002 |
| Cellulase         | Tiefseespalten        | 92                       | Leis et al., 2015       |
| Esterase          | Schlammreiches Wasser | 95                       | Rhee et al., 2005       |
| Lipase            | Boden                 | 60                       | Faoro et al., 2012      |
| Xylanase          | Vulkanmaterial        | 95                       | Mientus et al., 2013    |

## 1.5 Klassifizierung der Glykosidhydrolasen

Die Familiengruppen der Glykosidhydrolasen wurden aufgrund ihrer Aminosäure-Sequenzähnlichkeit eingeordnet (Henrissat et al., 1991). Da es einen direkten Zusammenhang zwischen Sequenz- und Faltungsähnlichkeiten gibt, spiegeln die strukturellen Eigenschaften dieser Enzyme besser eine Ähnlichkeitsgruppe wider, als z.B. nur die Substratspezifität. Die drei-dimensionalen Strukturen der Glykosidhydrolasen und ihre damit verbundenen katalytischen Zentren sind stark konserviert und können in drei topologische Formen eingeteilt werden (Henrissat et al., 1995). Die Struktur der „Tasche oder Krater“ eignet sich sehr gut für die Erkennung von Sacchariden mit nicht reduzierenden Enden, jedoch weniger für den Abbau von langfaserigen Substraten wie Cellulose, die kaum freie Seitenketten besitzen. Anders verhält es sich bei der Struktur der „Spalte oder Furche“, die eine zufällige Bindung an Zuckereinheiten in polymeren Substraten ermöglichen können. Häufig wird diese Form bei *endo*-aktiven Enzymen wie  $\beta$ -1,3-1,4-Glukanasen,  $\beta$ -1,4-Glukanasen, Lysozym, Endocellulase, Chitinase, Xylanase und  $\alpha$ -Amylasen gefunden. Eine weitere topologische Form, die sich aus der vorherigen entwickelt hat, ist der „Tunnel“ (Rouvinen et al., 1967). Durch die Öffnung können Polysaccharide eingefädelt, im katalytischen Zentrum gespalten und die Produkte wieder frei gesetzt werden, während die Kette weiterhin fest gebunden ist. Das wiederum ermöglicht eine kontinuierliche Reaktionskette und ist ein Schlüsselfaktor für einen effizienten enzymatischen Abbau von unlöslicher kristalliner Cellulose (Davies und Henrissat, 1995).

Die Beschaffenheit des katalytischen Zentrums beeinflusst die Art und Weise des Spaltungsmechanismus. Enzyme mit einem *endo*-Mechanismus besitzen ihr Aktivitätszentrum in einem offenen Spalt für die zufällige Bindung und Degradation der Polymerkette (Sulzenbacher et al., 1996). Die *exo*-Enzyme, die sehr spezifisch an den Polysaccharidkettenenden spalten, tragen ihr katalytisches Zentrum innerhalb der Taschenform. Dabei bestimmt die Tiefe und Ausbuchtung der Tasche, wo an dem Substrat gebunden wird und wie lang die Abspaltungsprodukte sind (Davies and Wilkson, 1997).

Die stereochemische Anordnung der Produkte ist vom Mechanismus der Hydrolase abhängig, d.h. ob die Konfiguration des anomeren Kohlenstoffatoms nach der Umsetzung umgekehrt wird (Inversion) oder unverändert bleibt (Retention) (siehe Abbildung 2). Bei invertierenden Glykosidhydrolasen gibt es nur eine Reaktion, bei der die katalytische Säure ein Proton an das zum anomeren Kohlenstoff benachbarte Sauerstoffatom abgibt, während die katalytische Base einem Wassermolekül im katalytischen Zentrum ein Proton entzieht. Dadurch wird die Nukleophilie des Wassers erhöht und der Angriff auf das anomere Zentrum vereinfacht. Bei dem nach Dan Koshland benannten „Koshland“ retaining Mechanismus liegt dagegen eine

zweistufige Reaktion vor, wobei der katalytische Säure-/Baserest zuerst als eine Säure und dann als Base fungiert.

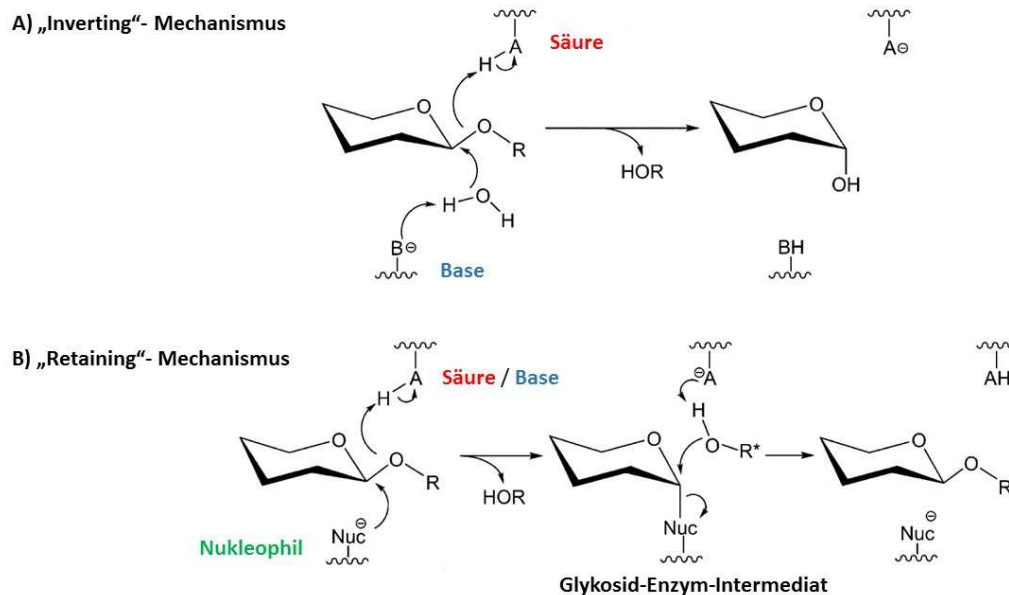


Abbildung 2: Invertierender und retendierender Mechanismus bei Glykosidhydrolasen (nach Vuong et al., 2010, modifiziert).

Die „retaining“-Reaktion besteht aus zwei aufeinanderfolgenden nukleophilen Substitutionen am anomeren Kohlenstoffatom. Bei dem ersten Schritt übergibt der katalytische Säurerest ein Proton an das freie Elektronenpaar des Sauerstoffatoms der glykosidischen Bindung. Die zunehmende Polarisierung der C-O-Bindung hat eine erhöhte Elektrophilie am anomeren Kohlenstoff zur Folge. Danach erfolgt ein nukleophiler Angriff am anomeren Kohlenstoffatom durch den nukleophilen Rest ( $B^-$ ). Die Orientierung der glykosidischen Bindung ( $\alpha$  oder  $\beta$ ) wird zu einer entgegengesetzt orientierten Glykosyl-Esterbindung umgewandelt. Dadurch entsteht das sogenannte „Glykosyl-Enzym-Intermediat“. Im zweiten Schritt der Reaktion wird durch den negativ geladenen Säure/Base ( $A^-$ ) Rest einem Wassermolekül ein Proton entzogen. Das Wasser wird zum Hydroxid-Ion, das stark nukleophil ist und einen nukleophilen Angriff am anomeren Kohlenstoff vollzieht. Die Glykosyl-Esterbindung zwischen dem Intermediat wird aufgebrochen und durch eine neue Hydroxylgruppe gleicher Orientierung relativ zum Ausgangssubstrat substituiert (Thoma and Koshland, 1960).

Die Glykosidhydrolasen werden bis heute in 145 verschiedene Familien unterteilt, gemäß aktuellem Stand in der CAZy-Datenbank (für *Carbohydrate-Active enZymes*; [www.cazy.org](http://www.cazy.org), Stand 11.2017). Dank der schnellen Entwicklung und der Zusammenarbeit der verschiedenen wissenschaftlichen Bereiche wie der Sequenzierung, funktionellen Metagenomik und Proteinkristallographie, können heutzutage viele neue Daten generiert und miteinander verknüpft werden, sodass das Potenzial für die Entdeckung neuer effizienterer Biokatalysatoren immer besser ausgeschöpft werden kann.

## 2 Ziel der Arbeit:

Die Hauptziele dieser Arbeit waren neben der Charakterisierung einer neuartigen hitzestabilen Endoglukanase, die bei der Durchforstung einer Metagenom-Genbibliothek identifiziert worden war, die Untersuchung der Expression metagenomischer Gene in zwei heterologen Expressionswirten, sowie die Erweiterung der Möglichkeiten zur Nutzung alternativer Wirtsorganismen zur funktionellen Durchforstung von Metagenom-Genbanken.

### 2.1 Expression und Charakterisierung der thermostabilen *endo*- $\beta$ -Glukanase EngU

In einer vorangegangenen Dissertation aus der Arbeitsgruppe von Prof. Liebl (Wittenberg, 2009) wurde die thermostabile  $\beta$ -Glukanase EngU (frühere Bezeichnung: Bgl48) in der metagenomischen Genbank FosK48H3 (Avachinsky Krater, Kamtschatka) entdeckt und biochemisch charakterisiert. Die Arbeiten sollten mit Hilfe neuer Sequenzdaten fortgeführt und ein Domänen-Modell von EngU erstellt werden. Die Funktionen der Domänen sollten durch verschieden konzipierte pETDuet-Expressionsplasmide experimentell untersucht und die katalytischen Reste über gezielte Punktmutationen identifiziert werden. Darüber hinaus sollte über verschiedene analytische Methoden (Dünnschichtchromatographie und HPAEC-PAD) das Substratspektrum, die Substratspezifität und der Spaltungsmechanismus genauer charakterisiert werden.

### 2.2 Expression und transkriptionelle Analyse zweier aktiver Fosmide in zwei verschiedenen Wirten

Bei der Durchmusterung von Metagenom-Genbibliotheken in *E. coli* und *T. thermophilus* werden manche metagenomischen Gene nicht gleichermaßen in beiden Wirten exprimiert. Dieses Phänomen sollte im Rahmen der vorliegenden Arbeit exemplarisch an ausgewählten Beispielen untersucht werden. Die beiden Fosmide HA-CMC-1 und TB-EST-7 wurden in EPI300 (*E. coli*) und BL03 (*T. thermophilus*) transformiert und exprimiert. Die Cellulase-Aktivität von HA-CMC-1 kann in *E. coli* nachgewiesen werden, in *T. thermophilus* jedoch nicht. Entgegengesetzt verhält es sich mit der Esterase-Aktivität von TB-EST-7, die nur in *T. thermophilus* nachgewiesen werden kann. Für die Aufklärung dieses Phänotyps sollten Transkriptionsanalysen zu den beiden Fosmiden in den zwei verschiedenen Expressionswirten durchgeführt werden. Dabei sollten nicht nur die Kandidatengene ORF HA23 und TB32 untersucht, sondern auch die anderen Gene auf dem Metagenominsert miteinander verglichen werden.

### 2.3 Optimierung der funktionellen Screenings in *E. coli* und *T. thermophilus*

Neben der *in vivo* Mobilisierung von pCC1FOS ist es möglich über den Shuttlevektor pCT3FK metagenomische Genbibliotheken von *E. coli* in *T. thermophilus* zu übertragen. Die Transformationseffizienz der Methode ist jedoch nicht ausreichend und die Handhabung noch zu umständlich, um diese in einem Hochdurchsatz-Screening anzuwenden. Daher sollte in dieser Arbeit die Modifizierung, die Übertragung und das Screening der Fosmid-Genbanken optimiert und auf andere Expressionswirte (*M. luteus* und *B. subtilis*) erweitert werden. Dazu wurden verschiedene Geräte wie z. B. der halbautomatische Pipettierroboter ViaFlow96 oder der Vakuum Manifold mit den dazugehörigen Präparationsplatten getestet. Zudem sollte eine sogenannte „Pool-Modifizierung“ helfen, eine große Anzahl von Klonen zu transformieren und zu selektieren.



### 3 Material und Methoden

#### 3.1 Verwendete Stämme und Wachstumsbedingungen

##### 3.1.1 Stämme und Plasmide

Die in dieser Arbeit verwendeten bakteriellen Stämme sind in Tabelle 3 zusammengefasst.

Tabelle 3: Verwendete Stämme

| Stamm                              | Genotyp  | Referenz                           |
|------------------------------------|--|------------------------------------|
| <i>E. coli</i> :                   |  |                                    |
| XL1-Blue                           | <i>recA-</i> , <i>thi</i> , <i>hsdR1</i> , <i>supE44</i> , <i>relA1</i> , <i>lacF'</i> , <i>proAB</i> , <i>lacIq</i> , <i>lacZΔM15</i> , <i>Tn10</i> [Tet] | Stratagene (La Jolla, USA)         |
| DH10B                              | <i>F-</i> <i>endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-</i>                                | Invitrogen (Carlsbad, USA)         |
| JM109                              | <i>endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB+ Δ(lac-proAB) e14- [F' traD36 proAB+ lacIq lacZΔM15] hsdR17(rK-mK+)</i>                                      | Promega (Fitchburg, USA)           |
| EPI300-T1R                         | <i>F-</i> <i>mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara,leu)7697 galU galK λ- rpsL (StrR) nupG trfA tonA</i>                  | Epicentre (Madison, USA)           |
| BL21 (DE3)                         | <i>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHI ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5</i>   | New England Biolabs (Ipswich, USA) |
| <i>E. coli</i> NEB 5-α             | <i>fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>   | New England Biolabs (Ipswich, USA) |
| <i>E. coli</i> Top 10              | <i>F-</i> <i>mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-</i>                          | Invitrogen                         |
| <i>E. coli</i> Rosetta 2           | <i>F-</i> <i>ompT hsdSB(rB- mB-) gal dcm pRARE2 (CamR)</i>   | Novagen                            |
| <i>E. coli</i> Artic Express Cells | <i>B F-</i> <i>ompT hsdS (rB- mB-) dcm+ Tetr gal endA Hte [cpn10 cpn60 Gentr] [argU proL Strr]</i>   | Agilent                            |

| Stamm                          | Genotyp   | Referenz                          |
|--------------------------------|---|-----------------------------------|
| <u><i>T. thermophilus</i>:</u> |   |                                   |
| HB27                           | Wildtyp Laborstamm Göttingen, DSMZ 7039                                 | Oshima and Imahori 1974;<br>DSMZ  |
| $\Delta$ eng                   | $\Delta$ TT_P0042   | Ohta <i>et al.</i> 2006;<br>H. Li |
| HB27::nar                      | nar-Operon  | Ramírez-Arcos <i>et al.</i> 1998  |
| BL03                           | $\Delta$ TT_C42; $\Delta$ TT_C904; $\Delta$ TT_C340-1; $\Delta$ TTC1787 | B. Leis <i>et al.</i> 2015        |

Die Liste der verwendeten Plasmide ist in Tabelle 4 abgebildet.

Tabelle 4: Verwendete Plasmide

| Bezeichnung | Beschreibung  | Referenz                           |
|-------------|---|------------------------------------|
| pCC1FOS     | 8,1 kb großes Fosmid für die Herstellung einer CopyControl Fosmid-Bibliothek für die Klonierung von 40 kb großen genomischen DNA-Fragmenten<br>Nur <i>E. coli</i> kompatibel<br>CmR; single (F-Plasmid) oder high ( <i>oriV</i> , induzierbar) copy ori   | Epicentre                          |
| pCT3FK      | 11,6 kb großes <i>E. coli</i> / <i>T. thermophilus</i> Shuttle fosmid, basierend auf dem pCC1FOS Fosmid; CmR; KmR; besitzt flankierende Regionen für die Integration in <i>pyrE</i> Locus in <i>T. thermophilus</i>                                       | Angelov <i>et al.</i> (2009)       |
| pRetro-T2   | 3,7 kb großes Plasmid mit flankierenden Regionen zum <i>pyr</i> -Locus in <i>T. thermophilus</i> . Dient als Vektor für die chromosomale Integration des Metagenoms in <i>T. thermophilus</i> ; basierend auf pScpyr_T2; KmR                              | A. Angelov                         |
| pJM2545     | Cre-Recombinase Expressionsplasmid für <i>in vivo</i> Modifizierung und Cre-abhängige Manipulationen; lacZ-Cre auf pTSK29, ts pSC101 low copy origin (30 °C), induzierbar mit 0,5 mM IPTG; exprimiert nur in <i>E. coli</i> Stamm DH10B; 6,8 kb groß, KmR | Margin-Lachmann <i>et al.</i> 2003 |

| Bezeichnung        | Beschreibung  | Referenz                        |
|--------------------|---|---------------------------------|
| pMK18              | 5 kb großes Plasmid, das von pUC abstammt; besitzt Sequenzen von <i>T. thermophilus</i> , wodurch es als Shuttle Vektor zwischen <i>E. coli</i> und <i>T. thermophilus</i> fungiert; KmR Kasette für thermostabile Kanamycin-Selektion unter bifunktionaler Promotorkontrolle | De Grado <i>et al.</i> 1999     |
| pMKE2              | Expressionsvektor für <i>Thermus</i> -Stamm mit <i>nar</i> -Operon, <i>oriE</i> , <i>repA</i> , <i>kat</i> (Kan <sup>R</sup> ), P <sub>nar</sub>  | Moreno <i>et al.</i> (2005)     |
| pUC18              | Klonierungsvektor, rep(pMB1), bla(AmpR), lacPOZ', MCS   | Fermentas (Vilnius, Litauen)    |
| pJET1.2/blunt      | 3,0 kb großes Plasmid mit pMB1 <i>ori</i> (rep), ApR; <i>eco47IR</i> ;  | Thermo Fischer Scientific (USA) |
| pET21a-HACMC1      | Expressionsvektor zur biochemischen Charakterisierung der archaeellen metagenomischen Endoglukanase Cel12E  | B. Leis (2015)                  |
| pET21a-spHACMC1    | Expressionsvektor zur biochemischen Charakterisierung der archaeellen metagenomischen Endoglukanase Cel12E mit nativem Signalpeptid   | B. Leis (2015)                  |
| pET21a(+)          | Amp100-Resistenz; Genexpression über lacZ induzierbar   | Merck-Millipore (USA)           |
| pETDuet1           | Amp100-Resistenz; Genexpression über lacZ induzierbar; zwei MCS ermöglichen eine Coexpression von zwei Genen  | Merck-Millipore (USA)           |
| pET21a_engUfull    | pET21a(+) mit eingefügter EngU-Sequenz  | Silja Brady (2007)              |
| pET21a_gh42a       | pET21a(+) mit eingefügter gh42a-Sequenz von EngU  | Diese Arbeit                    |
| pET21a_[gh42a_cbm] | pET21a(+) mit eingefügter [gh42a_cbm]-Sequenz von EngU  | Diese Arbeit                    |
| pET21a_cterm1      | pET21a(+) mit eingefügter cterm1-Sequenz von EngU   | Diese Arbeit                    |
| pETDuet_engUfull   | pETDuet mit eingefügter EngU-Sequenz in MCS1  | Diese Arbeit                    |
| pETDuet_cterm1     | pETDuet1 mit cterm1 in MCS2 durch Gibson Assembly   | Diese Arbeit                    |

| Bezeichnung                | Beschreibung   | Referenz     |
|----------------------------|--|--------------|
| pETDuet_cterm2             | pETDuet1 mit cterm2 in MCS2 durch Gibson Assembly  | Diese Arbeit |
| pETDuet_[gh42a_cbm]        | pETDuet1 mit [gh42a_cbm] in MCS1 durch Ligation  | Diese Arbeit |
| pETDuet_[gh42a_cbm]_cterm1 | pETDuet1 mit [gh42a_cbm] in MCS1 und cterm in MCS1 durch Ligation  | Diese Arbeit |
| pETDuet_gh42a              | pETDuet1 mit gh42a in MCS1 durch Ligation  | Diese Arbeit |
| pETDuet_gh42a_cterm1       | pETDuet1 mit gh42a in MCS1 und cterm1 in MCS2 durch Ligation   | Diese Arbeit |
| pETDuet_gh42ab             | pETDuet1 mit gh42ab in MCS1 durch Gibson Assembly  | Diese Arbeit |
| pETDuet_gh42ab_cterm2      | pETDuet1 mit gh42ab in MCS1 durch Gibson Assembly und cterm2 in MCS2 durch Ligation  | Diese Arbeit |
| pRetro-B                   | pRetro mit loxP, <i>ori</i> , R6K $\gamma$ ori, Kan-Kassette (aus pWLTK6) und 1980 bp <i>amyE</i> (aus <i>B. subtilis</i> Stamm 168) | Diese Arbeit |
| pRetro-M                   | pRetro mit loxP, <i>ori</i> , R6K $\gamma$ ori, Kan-Kassette (aus pWLTK6) und 1520 bp <i>trpE</i> (aus <i>M. luteus</i> ATCC 27141)  | Diese Arbeit |
| pRetro-BAC1                | pRetro mit R6K $\gamma$ ori, Kan-Kassette und 2062 bp aus pTT27 ( <i>repA</i> , <i>ori</i> )   | Diese Arbeit |
| pRetro-BAC2                | pRetro mit R6K $\gamma$ ori, Kan-Kassette und 4074 bp aus pTT27 ( <i>repA</i> , <i>ori</i> und <i>par locus</i> )                    | Diese Arbeit |

### 3.1.2 Nährmedien, Zusätze und verwendete Kits

Die verwendeten Medien wurden nach folgenden Angaben (Tabelle 5, Tabelle 6, Tabelle 7) hergestellt und bei 121 °C für 20 Minuten sterilisiert. Für Festmedien wurde zu der Lösung 1,5 % (w/v) Agar-Agar zugegeben. Weitere Zusätze wie Antibiotika, Zucker, Substrate, Mineralien oder Vitamine wurden zunächst steril filtriert und zu dem auf 60 °C abgekühlten autoklavierten Medium gegeben. Unter der Sterilbank wurden 25 mL des flüssigen Festmediums in eine Petrischale (Greiner Bio-One International AG, Kremsmünster, Österreich) gegossen und nach dem Aushärten weiter verwendet.

Tabelle 5: Zusammensetzung LB- und TB-Medium

|             | LB-Medium ( <i>E. coli</i> ) | TB-Medium ( <i>T. thermophilus</i> ) |
|-------------|------------------------------|--------------------------------------|
| Trypton     | 10 g/L                       | 8 g/L                                |
| Hefeextrakt | 5 g/L                        | 4 g/L                                |
| NaCl        | 5 g/L                        | 3 g/L                                |
| Wasser      | VE-Wasser                    | Aqua pure                            |
| pH          | keine                        | 7,5                                  |
|             | (Luria <i>et al.</i> , 1960) | (De Grado <i>et al.</i> , 1999)      |

Tabelle 6: SH Minimalmedium

| SH Minimalmedium   |           |   |          |
|--|-----------|---|----------|
| Tributyrin   | 1 % (v/v) | Casamino acids  | 5 g/L    |
| K <sub>2</sub> HPO <sub>4</sub>                                | 0,75 g/L  | 100 µg/mL Biotin                                      | 0,1 mg/L |
| KH <sub>2</sub> PO <sub>4</sub>                                | 0,25 g/L  | 1 mg/mL Thiamin                                       | 1 mg/L   |
| NaCl   | 2 g/L     | Agar-Agar   | 15 g/L   |
| (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>                | 2,5 g/L   |   |          |
| pH-Einstellung auf 7,2 und autoklavieren bei 121 °C für 10 min |           |   |          |
| Zugabe steriler Spurenelementlösungen (Endkonzentrationen)     |           |   |          |
| CaCl <sub>2</sub>  | 170 µM    | MnCl <sub>2</sub> x 4 H <sub>2</sub> O                | 2,5 µM   |
| CoCl <sub>2</sub> x 6 H <sub>2</sub> O                         | 3,4 µM    | Na <sub>2</sub> MoO <sub>4</sub> x 2 H <sub>2</sub> O | 5 µM     |
| CuSO <sub>4</sub> x 5 H <sub>2</sub> O                         | 0,1 µM    | NiCl <sub>2</sub> x 6 H <sub>2</sub> O                | 0,1 µM   |
| FeSO <sub>4</sub> x 7 H <sub>2</sub> O                         | 21,6 µM   | VO <sub>2</sub> SO <sub>4</sub> x 3 H <sub>2</sub> O  | 0,5 µM   |
| MgCl <sub>2</sub>  | 615 µM    | ZnSO <sub>4</sub> x 7 H <sub>2</sub> O                | 0,2 µM   |

Tabelle 7: Herstellung von YENB, SOB und SOC Medium

| Medium                                | Zusatz                               |         |
|---------------------------------------|--------------------------------------|---------|
| YENB                                  | Hefextrakt                           | 7,5 g/L |
|                                       | Bacto Nutrient Broth                 | 8,0 g/L |
| SOB<br>(Hanahan <i>et al.</i> , 1983) | Hefeextrakt                          | 0,5 %   |
|                                       | Trypton                              | 2 %     |
|                                       | NaCl                                 | 10 mM   |
|                                       | KCl                                  | 2,5 mM  |
|                                       | autoklavieren und dann:              |         |
| SOC                                   | MgCl <sub>2</sub>                    | 10 mM   |
|                                       | MgSO <sub>4</sub>                    | 10 mM   |
|                                       | Siehe SOB:                           |         |
|                                       | Zugabe von sterilfiltrierter Glukose | 20 mM   |

Die Chemikalien und Verbrauchsmaterialien stammen von den Firmen Applichem, Sigma, Merck und Fluka, wenn nicht anders angegeben. Die Antibiotikastammlösungen und andere Medienzusätze wurden entsprechend Tabelle 8 unten gelöst und steril filtriert. Die Antibiotika wurden bei -20 °C gelagert und erst kurz vor dem Gebrauch zu den Medien gegeben.

Tabelle 8: Verwendete Antibiotika und Medienzusätze

| Zusatz                             | Abkürzung | Konzentration | Lösungsmittel |
|------------------------------------|-----------|---------------|---------------|
| Ampicillin                         | Amp       | 100 µg/mL     | Wasser        |
| Chloramphenicol                    | Chl       | 12,5 µg/mL    | Ethanol       |
| Kanamycin                          | Kan       | 20 µg/mL      | Wasser        |
| Glukose                            | Glc       | 1 % (w/v)     | Wasser        |
| Isopropyl-β-D-thiogalaktopyranosid | ITPG      | 0,1 – 1 mM    | Wasser        |
| Fosmid Autoinduction Solution      | AIS       | 1 % (v/v)     | Wasser        |

### 3.1.3 Kultivierung von Bakterien

Die Anzucht von *E. coli* erfolgte in LB- oder SOB/SOC-Medium für 24 h bei 37 ° C, außer es besaß ein *in vivo* modifiziertes Fosmid, dann wurde bei 42 °C inkubiert. *T. thermophilus* benötigt zum optimalen Wachstum TB-Medium und wurde bei Anwesenheit von Kanamycin bei 60 °C und bei Abwesenheit bei 70 °C inkubiert. Die Anzucht erfolgte entweder auf Festmedien in Brutschränken oder in Flüssigkulturen, die für eine ausreichende Sauerstoffzufuhr bei 180 rpm geschüttelt werden mussten. Hierbei wurden ab einem Kulturvolumen von 30 mL Erlenmeyer-Kolben mit Schikane verwendet und für kleinere Kulturvolumina Reagenzgläser beziehungsweise 15 mL Falcon-Tubes eingesetzt.

### 3.1.4 Verwendete Puffermischungen

In der Tabelle 9 sind alle Puffer, die in dieser Arbeit verwendet wurden, aufgelistet.

Tabelle 9: Zusammensetzung der verwendeten Puffer

| Puffer                 | Zusammensetzung und Herstellung |                                  |
|------------------------|---------------------------------|----------------------------------|
| McIlvaine pH 6         | 0,1 M                           | Citrat                           |
|                        | 0,2 M                           | Na <sub>2</sub> HPO <sub>4</sub> |
| PBS 50 mM pH 6         | 0,6 %                           | NaH <sub>2</sub> PO <sub>4</sub> |
|                        | 0,16 %                          | Na <sub>2</sub> HPO <sub>4</sub> |
| MES-Puffer pH 5,5      | 50 mM                           | 2-(N-Morpholino)ethansulfonsäure |
| Sammelgelpuffer pH 6,8 | 61 g/L                          | Tris                             |
|                        | 4 g/L                           | SDS                              |
| Trenngelpuffer pH 8,8  | 182 g/L                         | Tris                             |
|                        | 4 g/L                           | SDS                              |
| TB-Puffer pH 6,7       | 10 mM                           | PIPES                            |
|                        | 15 mM                           | CaCl <sub>2</sub>                |
|                        | 250 mM                          | KCl                              |
|                        | 50 mM                           | MnCl <sub>2</sub>                |

|                           | pH Wert einstellen und sterilfiltrieren |  |
|---------------------------|---|--|
| 4x SDS Loading Dye        | 200 mM                                  | Tris-HCl                                       |
|                           | 4 %                                     | B-Mercaptoethanol                              |
|                           | 8 %                                     | SDS  |
|                           | 0,4 %                                   | Bromphenolblau                                 |
|                           | 40 %                                    | Glycin   |
| 10x SDS Laufpuffer        | 5 mM                                    | Tris   |
|                           | 192 mM                                  | Glycin   |
|                           | 0,1 %                                   | SDS  |
| 50 mM Sorensen Puffer pH8 | 50 mM                                   | Na <sub>2</sub> H <sub>2</sub> PO <sub>4</sub> |
|                           | 0,1 %                                   | Gum Arabis                                     |
|                           | 5 mM                                    | Natriumdeoxycholat                             |

### 3.1.5 Bestimmung der Zelldichte

Um den Wachstumsverlauf einer Bakterienkultur zu überprüfen, wurde die optische Dichte von 1 mL der Suspension im Spektralphotometer gemessen. Die Wellenlänge wurde bei einer *E. coli* Kultur auf 600 nm und bei *T. thermophilus* auf 550 nm eingestellt. Das entsprechende Medium diente als Referenz und wurde ebenfalls für die Verdünnung eingesetzt, wenn sich der OD-Wert über 1,5 befand.

### 3.1.6 Zellernte und Gewinnung von Zellextrakten

Für eine Vorkultur mit *E. coli* wurde eine Einzelkolonie von der Platte in 5 mL oder 30 mL Selektionsmedium angeimpft und ÜN bei 37 °C und 180 rpm geschüttelt. Mit dieser Vorkultur wurde die Expressionskultur mit einer OD<sub>600</sub> von 0,1 gestartet, bei 37 °C geschüttelt bis zu einer OD<sub>600</sub> von ca. 0,7 inkubiert. Für eine Expression ÜN wird IPTG mit einer Endkonzentration von 0,1 mM induziert während für eine 4 h Expression die Endkonzentration 1 mM benötigt wird. Die Kultivierung wurde jeweils unter Schütteln entweder für 4 h bei 37 °C oder ÜN bei 37 °C oder 20 °C durchgeführt. Nach der Expression wurde die Kultur bei 1000 rpm und 4 °C für 10 Minuten abzentrifugiert. Der klare Überstand wird abgenommen und verworfen, die Pellets können für den Zellaufschluss verwendet oder bei -20 °C oder über



längere Zeit bei -80 °C gelagert werden. Bei einer Kultur mit *T. thermophilus* wurden dieselben Arbeitsschritte eingeschaltet, jedoch wurden die Temperaturen angepasst. Bei einer Kultur ohne Antibiotikum wird bei 70 °C und mit Kanamycin 20 bei 60 °C inkubiert.

### 3.1.7 Herstellung von Glycerolstocks

Die Stammhaltung von *E. coli* und *Thermus* Kulturen erfolgte über Glycerinstocks. Die über Nacht gewachsenen Bakteriensuspensionen wurden mit 25 % Glycerin (v/v) versetzt und in Reaktionsgefäßen (1,5 mL) mit Schraubdeckel aliquotiert, in flüssigem Stickstoff Schock gefroren und bei -80 °C gelagert. Die Positionen der Stocks sind in Quartz (Trang Pham Box 1 - 4) hinterlegt.

## 3.2 Molekulargenetische Methoden

Alle Materialien wie Eppendorf Tubes, Pipettenspitzen, Gefäße etc., wurden vor dem Gebrauch entweder autoklaviert, mit 70%igen Ethanol (w/v) abgewaschen oder mit einem Bunsenbrenner abgeflammt.

### 3.2.1 Nukleinsäuren Analytik

#### 3.2.1.1 Sequenzieren

Die Sequenzierung von DNA-Bereichen wurden mithilfe von EZ-Seq durch die Firma Macrogen durchgeführt. Dafür wurden jeweils 5 µL der zu sequenzierenden DNA (100 ng/µL) und 5 µL des Sequenzierprimers (5 ng/µL) in ein Eppendorf-Reaktionsgefäß gegeben und zur Sequenzierung an die Firma Macrogen geschickt.

#### 3.2.1.2 Isolation und Aufreinigung von (meta)genomischen Nukleinsäuren und Plasmiden (DNA)

Zur Isolierung von Plasmiden für die Klonierung, Transformation oder Lagerung wurde das EasyPrep® Pro Plasmid Miniprep von Biozym gemäß Herstellerangaben benutzt. Die Elution erfolgte mit MilliQ Wasser statt Elutionspuffer.

Für die Isolation von Plasmiden für Analysen, die keine hohe Reinheit erforderten, wurde die Methode der Isopropanol Fällung („dirty prep“) genutzt. Hierzu wurden 4 mL einer ÜN-Kultur bei 13.000 rpm für 1 Minute abzentrifugiert und der Überstand verworfen. Das Pellet wurde in

250 µL Puffer 1 (P1) resuspendiert und anschließend 250 µL Puffer 2 (P2) zugegeben. Das Gefäß wurde 10-mal invertiert und die Suspension maximal 5 Minuten bei RT inkubiert. Es wurden 350 µL Puffer 3 (P3) hinzugegeben, mehrmals invertiert, 5 Minuten auf Eis inkubiert und dann bei 13.000 rpm und 4 °C für 15 Minuten zentrifugiert. Der klare Überstand wurde in ein neues Reaktionsgefäß überführt, mit dem 0,7-fachen Volumen Isopropanol durch Invertieren gemischt und 5 Minuten bei RT inkubiert. Die Suspension wurde bei 13.000 rpm und 4 °C bis zu 1 h abzentrifugiert, der Überstand abgezogen, das Pellet mit 1 mL 70 %igem Ethanol (EtOH) gewaschen und bei 13.000 rpm und 4 °C für 2 Minuten abzentrifugiert. Der Überstand wurde vorsichtig abgezogen und das Pellet bei geöffnetem Gefäßdeckel und RT getrocknet. Zuletzt wurde die DNA in 35 - 50 µL MilliQ-Wasser resuspendiert.

*Tabelle 10: Pufferzusammensetzung für die Isopropanolfällung*

| <b>Puffer 1 (Resuspension)</b> | <b>Puffer 2 (Lyse)</b> | <b>Puffer 3 (Neutralisation)</b> |
|--------------------------------|------------------------|----------------------------------|
| 10 mM EDTA                     | 0,2 M NaOH             | 3,2 M Kaliumacetat               |
| 50 mM Tris-HCl                 | 1 % SDS                |                                  |
| auf 100 mL mit MQ              | auf 100 mL mit MQ      | auf 100 mL mit MQ                |
| pH 8 einstellen                |                        |                                  |
| autoklavieren                  |                        |                                  |
| 100 µL RNase A (10 mg/mL)      |                        |                                  |
| bei 4 °C lagern                |                        |                                  |

### 3.2.1.3 Gelelektrophorese

Zur Bestimmung der Größe und Konzentration von DNA-Fragmenten wurde eine Gelelektrophorese durchgeführt. Die hierzu benötigte Agarose (DNA: 0,8 % und RNA: 3 %) wurde mit 1x TAE-Puffer in der Mikrowelle aufgeköcht und konnte nach dem Auflösen bei 60 °C zur weiteren Verwendung gelagert werden. Die flüssige Agarose wurde in eine dafür vorgesehene Plastikwanne mit Gelkamm gegossen. Nach dem Aushärten des Gels wurde es mit 4 °C gekühltem 1x TAE-Puffer bedeckt und mit den Proben, die mit 1x Loading Dye gemischt wurden, beladen. Zur Identifizierung der Banden wurde der 1 kB Marker der Firma Fermentas verwendet. Die Gelelektrophorese erfolgte bei 180 V für 22 Minuten. Zum Anfärben der Banden wurde das Gel für mindestens 10 Minuten in ein Ethidiumbromidbad gegeben und anschließend unter UV-Licht betrachtet und mit dem Alphamager (Protein simple) dokumentiert.

### 3.2.1.4 Pulse field Gelelektrophorese (PFGE)

Gegenüber der konventionellen Agarosegelelektrophorese können mittels Pulsed Field-Gelelektrophorese (PFGE) Fragmente von 10 Kilobasen bis 1 Megabasen aufgetrennt werden. Dabei werden komplette Bakterienzellen in Gelblöcke eingebettet und in diesen mit einem häufig schneidenden Restriktionsenzym verdaut. Nach dieser Präparation werden die Gelblöcke in die Gelslots eingebracht und es folgt die Elektrophorese. Der Strom bei der PFGE läuft nicht nur in eine Richtung, sondern das elektrische Feld ändert nach einer kurzen Zeit (Pulszeit) immer wieder seine Richtung in einem Winkel von maximal 90°. Durch den ständigen Richtungswechsel des elektrischen Feldes sind die DNA-Moleküle gezwungen eine längere Strecke durchs das Gel zu laufen. Dabei wandern die kleineren DNA-Moleküle schneller durch das Gel als die größeren, wodurch eine Auftrennung solch großer Fragmente möglich wird. Die Auftrennung der DNA wird durch viele Faktoren beeinflusst wie z.B. der eingeschlossene Winkel, die Feldstärke, die Gelkonzentration und die Temperatur. Die Puffer und die Parameter für die PFGE sind in der Tabelle 11 aufgelistet. Das Gel wurde nach 24 Stunden für eine halbe Stunde in einer 1%igen Ethidiumbromidlösung gefärbt und anschließend dokumentiert.

Tabelle 11: Puffer und Parameter für die Pulse Field Gelelektrophorese

| <b>Puffer</b> | <b>Zusammensetzung</b>   |
|---------------|--|
| TNE Puffer    | 10 mM Tris<br>200 mM NaCl<br>100 mM EDTA (pH 7,2)  |
| TE Puffer     | 10 mM Tris<br>1 mM EDTA  |
| EC Lysepuffer | 6 mM Tris-HCl (pH7,2)<br>200 mM NaCl (w/v)<br>100 mM EDTA<br>0,5 Brij-58 (polyoxyethylen-20-cetylester, Sigma)<br>0,2 % Deoxycholat<br>0,5 % N-Laurylsarcosin<br>1 mg/mL Lysozym<br>20 µg/mL RNase |
| NDSK Puffer   | 0,5 M EDTA   |

|                      |   |
|----------------------|---|
|                      | 1 % (w/v) N-Laurylsarcosine   |
|                      | 1 mg/mL Proteinase K (kurz vor Gebrauch)  |
| 0,1 M PMSF Lösung    | in Isopropanol gelöst. Bei -20 °C lagern (Vorsicht: Phenylmethansulfonylfluoride ist giftig! Schutzkleidung anziehen) |
| Restriktionslösung   | Restriktionsenzym: 4 µL<br>10-facher Puffer: 20 µL<br>0,02 % Triton X-100: 4 µL<br>H <sub>2</sub> O: 172 µL           |
| 10-facher TBE Puffer | 108 g Tris<br>55 g Borsäure<br>40 mL von 0,5 M EDTA   |

| Parameter  | Einstellungen |
|------------|---------------|
| Spannung   | 1,5 v/cm      |
| Zeit       | 24 Stunden    |
| Temperatur | 14 °C         |
| Startzeit  | 1 s           |
| Endzeit    | 25 s          |

### 3.2.1.5 DNA Gel Purifikation

Um mit einzelnen auserwählten DNA-Banden weiter zu arbeiten, müssen diese aus einem Agarosegel aufgereinigt werden. Hierzu wurde zunächst eine Gelelektrophorese durchgeführt. Da Mutationen durch die Ethidiumbromid-Färbung oder UV-Strahlung vermieden werden sollten, wurden die aufzureinigenden Proben in eine verbreiterte Kammer in der Mitte und je 2 µL rechts und links davon aufgetragen. Vor dem Färben wurde der Teil mit den verbreiterten Kammern vorsichtig mit einem Skalpell ausgeschnitten und nur die 2 µL-Probenbanden inklusive Marker ins Ethidiumbromidbad gegeben. Unter UV-Licht wurden anschließend die Banden, die die gewünschte Fragmentgröße enthielten, zur Markierung eingeschnitten. Nach Abschalten des UV-Lichts wurde das komplette Gel wieder zusammengefügt und die gewünschten Fragmente anhand der Markierungen ausgeschnitten. Die Aufreinigung erfolgte durch das AccuPrep® DNA Gel Purification Kit von Biozym anhand des beigelieferten Protokolls, anstelle des Elutionspuffer wurde MilliQ-Wasser verwendet.

#### *3.2.1.6 DNA Quantifizierung*

Um die Konzentration und die Reinheit der isolierten DNA zu überprüfen wurde 1 µL der Probe am NanoDrop ND-1000 Spectrophotometer (NanoDrop Products, Wilmington, USA) gemessen. Das DNA/Proteinverhältnis (OD 260/280) sollte zwischen 1,8 und 2,0 liegen. Werte, die nicht in dem Bereich liegen, könnten auf eine Verunreinigung hinweisen. Die Konzentration von einzelnen linearen DNA-Fragmenten kann auch über die Gelelektrophorese bestimmt werden, indem man unterschiedliche Mengen des Markers aufträgt und die Intensitäten der Banden miteinander vergleicht.

#### *3.2.1.7 Arbeiten mit mRNA*

Die Quantifizierung der messenger RNA (mRNA) sagt aus, wie oft ein Gen in der Zelle umgeschrieben wird. Die mRNA ist einzelsträngig und daher instabiler als die doppelsträngige DNA. Daher sollte man alle Tubes und Pipettenspitzen zweimal autoklavieren und beim Arbeiten mit RNA immer Handschuhe tragen, um einen Verdau durch RNase zu vermeiden.

#### *3.2.1.8 Isolation von RNA*

Für die Isolierung der Gesamt-RNA wurde das MasterPURE® von Epicentre verwendet und nach Herstellerangaben gearbeitet. Dem Protokoll wurde ein weiterer Schritt mit der Roche DNaseI von der Firma Roche hinzugefügt, um die DNA-Kontamination zu verringern.

#### *3.2.1.9 Umschreibung der mRNA in cDNA*

Die Reverse Transkriptase (RT) schreibt die einsträngige RNA in eine doppelsträngige cDNA um. Dazu erstellen die Enzyme zuerst einen RNA-DNA-Hybridstrang und bauen dann später den RNA-Anteil ab. Zum Schluss folgt die Vervollständigung des einzel- zum doppelsträngigen DNA-Strang. Für diese Reaktion wurde das iScript™ Select cDNA Synthesis Kit von Biozym verwendet und nach folgendem Protokoll durchgeführt:

#### *3.2.1.10 Quantifizierung der RNA (qPCR)*

Nachdem die mRNA zur cDNA umgeschrieben wurde, kann mithilfe der quantitativen Real-time PCR die Transkriptmenge eines Gens durch das Fluoreszenzsignal in Echtzeit bestimmt werden. Dazu wurden die spezifischen Primer, die in der Mitte der mRNA liegen, erstellt und die Reagenzien aus dem Sso Advanced™ Universal SYBR® Green Supermix Kit von Biozym verwendet. Der Farbstoff SYBR Green bindet sequenzunabhängig an doppelsträngiger DNA, wobei die Stärke des Signals proportional mit dem Gehalt des Amplifikats zusammenhängt. Der Zyklus, bei dem das Signal zum ersten Mal über der Hintergrundfluoreszenz liegt und

somit den Schwellenwert schneidet wird als Threshold-Zyklus (engl.: *threshold cycle*; Ct) bezeichnet. Der Ct-Wert ist von der Ausgangsmenge der cDNA abhängig d.h. je mehr cDNA am Anfang der Amplifikation in der Probe vorhanden ist, desto kleiner ist der Ct-Wert. Die qPCR-Ansätze und auch die Programmierung des Laufes können aus Tabelle 12 und Tabelle 13 abgelesen werden.

Tabelle 12: qPCR Probenansatz

| Komponenten                       | Volumen / Rkt. | Endkonzentration               |
|-----------------------------------|----------------|--------------------------------|
| Sso Advanced™ SYBR Green supermix | 10 µL          | 1x                             |
| Forward Primer                    | variable       | 500 µL                         |
| Reverse Primer                    | variable       | 500 µL                         |
| RNase/DNase-freies Wasser         | variable       | -                              |
| DNA Template                      | variable       | cDNA: 1 µg<br>gDNA: 50 ng–5 pg |
| Gesamtvolumen                     | 20 µL          | -                              |

Tabelle 13: qPCR Programm für cDNA, Plasmid DNA und genomische DNA

| cDNA und Plasmid DNA  |                                 |                   |              |
|-----------------------|---------------------------------|-------------------|--------------|
| Schritt               | Temperatur                      | Zeit              | Zyklusanzahl |
| Enzymaktivierung      | 95 °C                           | 30 s              | 1            |
| Denaturierung         | 95 °C                           | 5 s               | 35 bis 40    |
| Annealing / Extension | opt. Temp.                      | 10 – 30 s         |              |
| Schmelzkurve          | 65–95 °C<br>(in 5 °C Schritten) | 2 – 5 s / Schritt | 1            |
| Genomische DNA        |                                 |                   |              |
| Schritt               | Temperatur                      | Zeit              | Zyklusanzahl |
| Enzymaktivierung      | 98 °C                           | 2 min             | 1            |
| Denaturierung         | 98 °C                           | 5 s               | 35 bis 40    |
| Annealing / Extension | opt. Temp.                      | 10–30 s           |              |
| Schmelzkurve          | 65–95 °C<br>(in 5 °C Schritten) | 4–5 s / Schritt   | 1            |

## 3.2.2 Molekulare Arbeiten und Modifikationen der DNA

### 3.2.2.1 Polymerase-Kettenreaktion

Die Polymerase-Kettenreaktion (engl.: *Polymerase-Chain-Reaction*, PCR) wird für die *in vitro* Amplifizierung von Genen verwendet. Die Reaktion besteht aus der Denaturierung, dem Annealing und der Elongation. Diese drei Phasen werden zyklisch wiederholt (Loops), wodurch das DNA-Fragment spezifisch vervielfältigt wird. Die PCR-Reaktionen wurden im Thermocycler von BioRad durchgeführt. Die Q5 High-Fidelity DNA-Polymerase von New England BioLabs und das dazu gehörige Protokoll des Herstellers wurden hierzu verwendet. Die Reaktionen wurden in 25 µL Ansätzen durchgeführt und als Negativkontrolle diente je ein Ansatz mit MilliQ Wasser anstelle der Template-DNA.

Der PCR-Ansatz und das dazugehörige Programm kann aus der unteren Tabelle abgelesen werden (Tabelle 14, Tabelle 15). Die Hybridisierungstemperatur X richtete sich nach dem CG-Gehalt des Primer-Paares und die Elongationszeit Y nach der Länge des Amplifikats. Die PCR-Produkte wurden mit dem AccuPrep® PCR Purification Kit der Firma Bioneer anhand des beigelieferten Protokolls durchgeführt.

Tabelle 14: PCR-Ansatz für Q5 High-Fidelity Polymerase

| Komponente               | Eingesetztes Volumen |
|--------------------------|----------------------|
| Enhancer                 | 5 µL                 |
| Puffer                   | 5 µL                 |
| DNTPs                    | 0,5 µL               |
| Forward Primer (10 pmol) | 1,25 µL              |
| Reverse Primer (10 pmol) | 1,25 µL              |
| Q5-Polymerase            | 0,25 µL              |
| H <sub>2</sub> O         | 10,75 µL             |



Tabelle 15: PCR Laufprogramm

| Schritt                | Temperatur     | Zeit  |      |
|------------------------|----------------|-------|------|
| Initiale Denaturierung | 98 °C          | 30 s  |      |
| 30x                    | Denaturierung  | 98 °C | 10 s |
|                        | Hybridisierung | X °C  | 20 s |
|                        | Elongation     | 72 °C | X s  |
| Finale Elongation      | 72 °C          | 120 s |      |
| Kühlen                 | 4 °C           | ∞ s   |      |

### 3.2.2.2 Restriktionsverdau

Bei einem Verdau werden die Palindrome der doppelsträngigen DNA, die sog. Restriktionsschnittstellen, von Endonukleasen spezifisch erkannt und geschnitten. Dabei können zwei verschiedene Enden (*sticky* und *blunt ends*) entstehen, die von der Lage der gespaltenen Phosphodiesterbindungen abhängen. Es wurden für das jeweilige Enzym oder auch Kombinationen von Enzymen die vom Hersteller (Fermentas) empfohlenen Puffer verwendet. Die Inkubationszeit betrug 1 - 3 Stunden bei 37 °C. Der Ansatz wurde für mindestens 60 Minuten bei 37 °C inkubiert. Bei höherer DNA Menge wurden die Volumina der anderen Komponenten entsprechend erhöht.

Tabelle 16: Ansatz für einen Restriktionsverdau

| Komponente        | Eingesetzte Menge |
|-------------------|-------------------|
| DNA               | 2-3 µL            |
| Restriktionsenzym | 1 µL              |
| Puffer            | 1 µL              |
| H <sub>2</sub> O  | X µL              |
| Gesamtvolumen     | 10 µL             |

### 3.2.2.3 Dephosphorylierung und Blunting

Nach einem Verdau bleiben an den 5' Enden der DNA Phosphatreste zurück, die für eine Ligation benötigt werden. Die alkalische Phosphatase von Fermentas wurde verwendet, die die Dephosphorylierung (des Vektors) katalysiert und das Religieren von linearisierten

Plasmiden verhindert. Der Prozess dauerte 10 Minuten bei 37 °C und wurde bei 75 °C (5 Minuten) gestoppt. Zum Glätten (engl.: *blunt*) der Überhänger (engl.: *sticky ends*) wurde das Quick Blunt Kit von Epicentre verwendet.

### 3.2.2.4 Ligationen

Bei der Ligation wird ein geöffneter, dephosphorylierter Vektor mit einem DNA-Fragment durch ihre kompatiblen Enden zusammengefügt. Mithilfe der Ligase wird das 3' Hydroxy-Ende mit dem 5' Phosphat-Ende des Fragments durch die Ausbildung einer Phosphodiesterbindung verbunden. Das Verhältnis zwischen den beiden eingesetzten Komponenten kann aus folgender Formel entnommen werden:

$$m(\text{Insert}) = \frac{m(\text{Vektor}) \times \text{Größe}(\text{Insert})}{\text{Größe}(\text{Vektor})} \times \text{Verhältnisfaktor}$$

### 3.2.2.5 Gibson Assembly

Diese Methode bietet die Möglichkeit, bis zu fünf verschiedene DNA-Fragmente simultan zu verbinden, ohne die Notwendigkeit von Restriktionsendonukleasen. Für Gibson-Assembly ist es zunächst notwendig, durch entsprechenden Primer-Design und PCR die zu verknüpfenden DNA-Fragmente so zu erstellen, so dass

die zu verbindenden Fragmente eine Sequenzüberlappung von etwa 20 bp aufweisen. Während der Reaktion verdaut eine Exonuklease dann die 5'-Enden der Fragmente, dass die freiliegenden 3'-Enden überlappen und assembliert werden können (Abbildung 3). Für die Assemblierungsreaktion wurden zunächst 10 µL eines Mixes aus den zu assemblierenden Fragmenten hergestellt, indem die Fragmente jeweils im molaren Verhältnis von 1:1 vorlagen. Daraufhin wurden 10 µL des Gibson-Assembly® Mastermix zugegeben und bei 50 °C für 120 Minuten im Thermocycler inkubiert. Anschließend konnte ein Teil des Reaktionsansatzes in kompetente Zellen eines Klonierungsstammes transformiert werden.

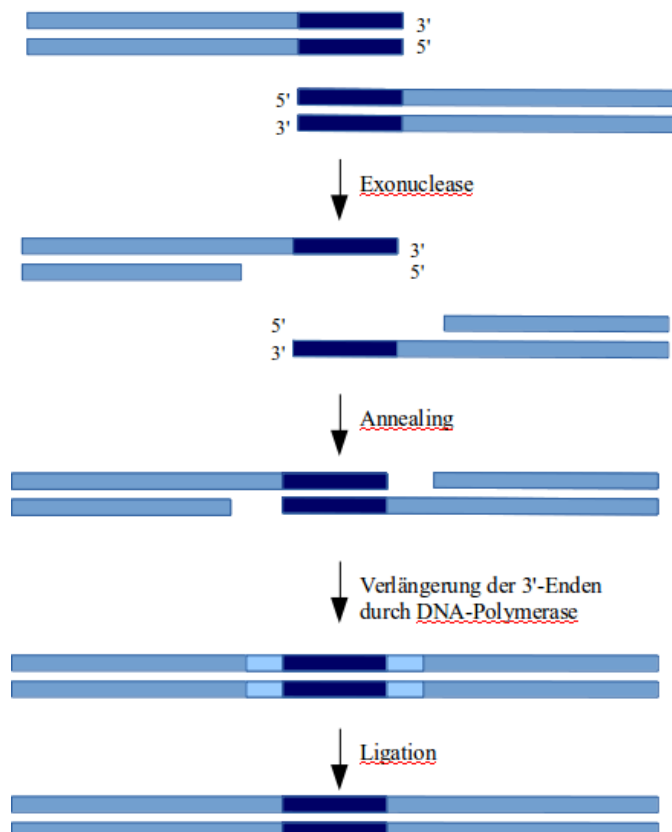


Abbildung 3: Gibson Assembly Schema

### 3.2.2.6 *Site-directed mutagenesis*

Das ChangeIT™ Multiple Mutation Site Directed Mutagenesis Kit von Affymetrix bietet die Möglichkeit, zielgerichtete Punktmutationen mittels PCR in Plasmide einzufügen. Dazu wird eine Vektor-PCR mit dem mutagenisierten Primer (z.B. Aminosäurenaustausch) und mit dem komplementären Amp Primer aus dem Kit (forward oder reverse) durchgeführt. Die Software Clone Manager 9.0 wurde zum Erstellen der Primer verwendet. Die mutagenisierten Basenpaare sollten sich in der Mitte der 20 - 25 Basenpaare befinden und das 5' Ende sollte phosphoryliert sein, damit sich das Plasmid komplett schließen kann. Die Ansätze und das PCR Programm wurden nach Herstellerangaben durchgeführt. Die 5 - 10 µL der PCR wurden in kompetente *E. coli* Zellen transformiert und auf Ampicillin 100 Platten ausplattiert. 5 - 10 Einzelklone wurden ausgewählt, die Plasmide isoliert und durch Sequenzierung verifiziert.

## 3.2.3 Transformation von DNA

### 3.2.3.1 *Herstellung von Ultra-kompetenten E. coli Zellen*

Zur Herstellung ultra-kompetenter *E. coli*-Zellen wurden 250 mL SOB-Medium mit 1 mL einer Übernachtskultur von *E. coli* angeimpft und bei 4 °C zu einem OD<sub>600</sub>-Wert von 0,6 inkubiert. Danach wurde die Zellsuspension in gekühlte 50 ml Falcons überführt und 10 Minuten auf Eis gekühlt. Bei 5000 rpm, 4 °C wurde 15 Minuten lang abzentrifugiert und das Zellpellet vorsichtig durch Schwenken in insgesamt 50 mL Transformationspuffer resuspendiert. Dieser Schritt wurde zweimal wiederholt und die Suspension zwischendurch immer wieder für 10 Minuten auf Eis gekühlt. Zum Schluss wurden die Zellen in 17 ml Transformationspuffer und 1,19 mL DMSO (in zwei Portionen mit einer 5-minütigen Pause dazwischen dazugeben) aufgenommen und aliquotiert. Nach dem Schockfrieren im flüssigen Stickstoff wurden die Zellen bei -80 °C aufbewahrt.

### 3.2.3.2 *Chemische Transformation von E. coli*

Ein Aliquot kompetenter Zellen (50 - 100 µL) wurde auf Eis langsam aufgetaut, 1 - 3 µL des Plasmids dazu pipettiert, gemischt und die Zellen 30 Minuten auf Eis inkubiert. Es folgte der Hitzeschock für 20 - 40 Sekunden (je nach Herstellerangaben) im 42 °C-warmen Wasserbad auf den gleich eine zweiminütige Inkubation auf Eis folgte. Durch die Hitze öffnen sich die Poren und die Zellen können die DNA leichter aufnehmen, die anschließende Kälte stabilisiert die Zellmembran und schließt die Poren wieder. Nun wurden 900 µL LB- oder SOC-Medium zu den transformierten Zellen hinzugegeben und zur Expression der Antibiotikaresistenz 1 h bei 37 °C inkubiert. Unterschiedliche Mengen der Bakteriensuspension wurden dann steril auf Agarplatten mit den entsprechenden Antibiotika ausplattiert und ÜN 37 °C bebrütet. Für die

Transformation von Plasmiden in kommerziellen *NEB-5 $\alpha$* -Zellen wurde nach dem Transformations-Protokoll der Firma New England Biolabs verwendet.

### 3.2.3.3 Elektroporation von *E. coli*

Für die Herstellung elektrokompetenter *E. coli*-Stämme wurde 1 L YENB-Medium mit 5 mL ÜN-Kultur inokuliert und für 3 - 6 h bei 37 °C geschüttelt. Wenn die Zellen eine optische Dichte von 0,5 - 0,9 bei einer Wellenlänge von 600 nm erreicht hatten, wurde die Kultur zunächst für 10 Minuten auf Eis gelegt und dann bei 10.000 rpm bei 4 °C abzentrifugiert. Das Pellet wurde zweimal in 200 mL eiskaltem Wasser resuspendiert und gewaschen. Danach wurden die Zellen in 50 mL eiskalten 10 % (w/v) Glycerol resuspendiert und erneut abzentrifugiert. Zum Schluss wurde mit 2 mL 10 % Glycerol resuspendiert, in 50  $\mu$ L aliquotiert und in flüssigem Stickstoff weggefroren. Die elektrokompetenten Zellen können für mehrere Monate bei -80 °C gelagert werden und bei Bedarf auf Eis aufgetaut werden. Zu den Zellen wurde 2 - 3  $\mu$ L der zu transformierenden DNA hinzugefügt. Dabei ist es sehr wichtig, dass die DNA-Lösung vorher dialysiert wurde (1 Stunde auf Dialysefilterpapier von der Firma Millipore) und salzfrei ist, damit die Elektroporation funktioniert.

Die Elektroporation wurde in einer Elektroporationsküvette mit einem Elektrodenabstand von 1 mm (PeqLab Biotechnologie GmbH, Erlangen) mit einem Bio-Rad Gene Pulser II™ Electoporation System durchgeführt. Die Küvetten wurden gekühlt, bevor die aufgetauten Zellen mit der DNA hinein pipettiert wurden. Die verwendeten Elektroporations-Parameter waren 2,5 kVmm, 1,25  $\mu$ Farad und 200 Ohm mit Zeitkonstanten um 5 ms. Nach erfolgreichem Elektroschock wurde 1 mL warmes SOC-Medium zu den Zellen gegeben und die Suspension eine Stunde schüttelnd bei 37 °C inkubiert, ehe der Transformationsansatz auf Selektionsmedien ausplattiert wurde.

### 3.2.3.4 Natürliche Transformation von *T. thermophilus*

Die Transformation von Zellen des natürlich kompetenten Stammes *T. thermophilus* HB27 mit Fremd-DNA wurde nach dem Protokoll von Angelov *et. al.*, 2009 durchgeführt. Es wurde eine 5 mL Vorkultur in TB-Medium mit einer Einzelkolonie des Stammes angeimpft und bei 70 °C und 180 rpm über Nacht inkubiert (15 mL Falcon). Am nächsten Tag wurde mit 500  $\mu$ L der ÜN-Kultur, 5 mL frisches TB-Medium angeimpft und wieder bei 70 °C und 180 rpm, bis zu einer OD<sub>550</sub> von ca. 0,7, inkubiert. Anschließend wurde diese Kultur 1:1 mit frischem TB-Medium (Gesamtvolumen von 1 mL) in ein neues Reaktionsgefäß überführt und für 1 Stunde weiter inkubiert. Anschließend wurde mindestens 1  $\mu$ g DNA (maximal 35  $\mu$ L Gesamtvolumen DNA in Wasser) zugegeben und die Inkubation für 2 Stunden fortgesetzt. Zum Schluss wurden

die Verdünnungsstufen  $10^{-5}$ ,  $10^{-6}$  und  $10^{-7}$  erstellt und je 100  $\mu\text{L}$  auf purem TB-Agar zur Gesamtzellzahlbestimmung aufgetragen. Von der unverdünnten Kultur wurden auch 100  $\mu\text{L}$  auf je eine TB-Agarplatte mit Kanamycin 20 (TB-Kan20-Agar) ausplattiert. Man kann im Anschluss die Transformationseffizienz berechnen, indem man die Transformanden auf der Kanamycinplatte durch die Menge [ $\mu\text{g}$ ] an eingesetzter DNA teilt. Alle Platten werden zwei Tage bei 60 °C inkubiert.

### 3.2.4 Konstruktion und funktionelle Analysen von Metagenom-Genbanken

#### 3.2.4.1 Verwendete Genbanken und Umweltproben

Für die funktionelle Metagenomanalyse standen verschiedene Genbanken in *E. coli* von Kooperationspartnern und aus eigener Herstellung zur Verfügung (Tabelle 17). Die Umweltproben stammen hauptsächlich aus extrem thermophilen Regionen und bieten dadurch eine große genetische Variabilität an hitzestabilen Enzymen. Das Metagenom der meist nicht kultivierbaren Bakterien kann direkt isoliert und in Fosmide verpackt werden. Das hat den Vorteil, dass man die Proben vorher nicht aufkonzentrieren muss und die genetische Diversität erhalten bleibt.

Tabelle 17: Übersicht der untersuchten Metagenom-Genbanken

| Probenbezeichnung                          | Größe der metagen. Genbank   | Herkunft   |
|--|--|--|
| Mesophile Bakterien (MB)                   | ca. 100.000 gepoolte Klone in pCC1FOS  |  |
| Thermophile Bakterien (TB)                 | ca. 100.000 gepoolte Klone in pCC1FOS  | P. Golyshin<br>(Bangor University, UK)                     |
| Hyperthermophile Archaeen (HA)             | ca. 12.000 gepoolte Klone in pCC1FOS   |  |
| Umweltproben von heißen Quellen der Azoren | Wasser, Sedimente und Biofilm Probenmaterial, Azoren (Portugal), teils in pCC1FOS und teils pCT3FK   | L. Conçaves, H. Santos<br>(Universität Lissabon, Portugal) |
| „Mexico“                                   | Sediment aus heißer Schwefelquelle; 50 °C; pH 6,7; ca. 3330 Klone in pCC1FOS                         | Ag Angelov   |
| „Russia“                                   | Biologisches Material aus Biofilter einer Trocknungsanlage 60 °C; pH 7,73; ca. 6650 Klone in pCC1FOS | Ag Angelov   |

#### 3.2.4.2 Erstellung von Fosmidbanken aus (meta)genomischer DNA

Die Klonierung großer (meta)genomischer DNA-Fragmente in Fosmide bietet die Möglichkeit, größere Sequenzabschnitte (etwa 40 kb) pro Klon funktionell zu untersuchen und eine höhere Wahrscheinlichkeit, zusammenhängende Gene zu identifizieren. Voraussetzung für die Klonierung dieser sogenannten „*large-insert*“ Genbanken ist, entsprechend lange und intakte DNA-Fragmente aus Mikroorganismen und Umweltproben zu erhalten. Für die direkte Isolierung der metagenomischen DNA aus der Umweltprobe wurde das “PowerSoil® DNA Isolation Kit“ der Firma Mo Bio Laboratories, Inc. verwendet. Es wurde nach Herstellerangaben gearbeitet. Die metagenomische DNA wurde in ungefähr 40 kb große Fragmente geschert und in pCC1FOS ligiert. Die Übertragung der Fosmide erfolgte durch Transduktion von *E. coli* EPI300 mit MaxPlax™ Lambda Bakteriophagen. Nach diesen Prozeduren wurden die transduzierten Fosmidklone auf LB-Chloramphenicol (12,5 µg/mL) ausplattiert und mindestens 10 Einzelkolonien mittels analytischem Restriktionsverdau mit EcoRI und BamHI auf ihre Diversität hin analysiert. Nach erfolgreicher Überprüfung wurden je nach Genbankgröße mehrere hundert bis tausend gewachsene Kolonien in 96-Well Mikrotiterplatten übertragen. Jede Kolonie entspricht damit einer spezifischen Position in der Mikrotiterplatte. Die mit 180 µL LB-Medium und Chloramphenicol (12,5 µg/mL) über Nacht bei 37 °C inkubierten Mikrotiterplatten wurden schließlich zur Langzeitlagerung mit 20 µL DMSO versetzt, mit einer sterilen Kunststoffolie (Tape pads 50, Qiagen, Hilden) verschlossen und bei –80 °C weggefroren. Neben den pCC1FOS Fosmidbanken, gab es auch pCT3FK Genbanken. Dieser Vektor wurde von Angel *et. al.* 2009 hergestellt und enthält neben den genetischen Elementen des pCC1FOS Vektors zusätzliche Selektions- und Mobilisierungselemente ( $P_{slp}$  Promotor und thermostabile Kanamycin-Adenyltransferase des *E. coli*-*Thermus Shuttle*-Plasmids pMK18 und chromosomale Sequenzelemente des *pyr*-Locus in *Thermus*). Durch die neuen Elemente war es möglich, Genbanken direkt in *T. thermophilus* zu transformieren und zu verifizieren.

#### 3.2.4.3 Subklonierung von Fosmid-DNA

Einzelne Fosmide wurden für die Sequenzierung in *shotgun*-Genbanken subkloniert. Dabei wurde die Fosmid-DNA durch wenige Ultraschall-Pulse fragmentiert und die 1,5 – 3 kb großen DNA-Stücke nach gelelektrophoretischer Trennung aufgereinigt. Die DNA-Enden wurden mit dem End-It™ DNA End-Repair Kit (Epicentre) behandelt. Das Anfügen von 3'-dA-Überhängen (*A-tailing*) wurde mit der DreamTaq DNA Polymerase (Fermentas) gemäß Herstellerangaben ohne Primer und lediglich unter Zugabe von 0,2 mM dATP bei 72 °C für 30 Minuten durchgeführt. Die Klonierung der Insert-DNA in den pCR2.1-XL-TOPO Vektor (Invitrogen) wurde laut Herstellerangaben durchgeführt.

#### 3.2.4.4 Fosmid-DNA-Präparation im 96-Well Format

Für die Fosmidisolierung wurden 2 Deep-well Blöcke mit 1,2 mL zweifachen LB-Medium, Chloramphenicol- und AIS-Zusatz (Autoinduction Solution) mit 20 µl Zellsuspension angeimpft und die Zellen für 48 h bei 37 °C und 350 rpm angezogen. Für die Fosmidpräparation aus *in vivo* modifizierten, gesicherten *E. coli* EPI300-Zellen wurde genau wie oben beschrieben vorgegangen nur, dass dem Medium zusätzlich Kanamycin 60 hinzugefügt und bei 42 °C inkubiert wurde. Die Zellen wurden bei 4000 rpm und 4 °C zehnminütig abzentrifugiert, der Überstand verworfen und in 200 µl Puffer P1 (Resuspensionspuffer) resuspendiert. Durch Zugabe von 200 µl Puffer P2 (Lysepuffer) erfolgte nun die Lyse der Zellen, welche durch Zugabe von Puffer P3 (Neutralisierungspuffer) gestoppt wurde (Puffer: Tabelle 10). Nachdem gut gemischt wurde, folgte ein 30-minütiger Zentrifugationsschritt bei 4000 rpm und 4 °C, wobei Zellbestandteile pelletiert wurden, während die DNA im Überstand verblieb. Der Überstand wurde nun in einen neuen Deep-well Block überführt, in welchem 650 µl Isopropanol vorgelegt waren, und gut vermischt. Erneut wurde für eine Stunde bei 4000 rpm und 4 °C zentrifugiert, der Überstand verworfen und die pelletierte DNA durch Zugabe von 200 µl 70 %-igem Ethanol gewaschen. Anschließend wurde das Pellet bei Raumtemperatur getrocknet und in 35 µl MilliQ-Wasser gelöst.

#### 3.2.4.5 *In vivo* Mobilisierung im 96-Well Format

Um ein und dieselbe pCC1FOS-Metagenom-Genbank vergleichend in *E. coli* und in *T. thermophilus* untersuchen zu können, muss diese modifiziert und für *T. thermophilus* kompatibel gemacht werden. Hierfür wird die Genbank über die Plasmide pRetro-T2 und pJM2545 in *E. coli* genetisch so verändert, dass ein Shuttle-Vektor geschaffen wird, der durch homologe Rekombination die in ihm enthaltene metagenomische DNA in das *Thermus*-Chromosom integrieren kann.

Von der in *E. coli* basierenden pCC1FOS-Genbank im DMSO-Stock wurden 10 - 20 µl der Zellsuspension durch einen Stempel in 200 µl LB-Medium und Chloramphenicol neu angeimpft und bei 37 °C über Nacht inkubiert. Am nächsten Tag wurde ein Deep-well Block mit 1,2 mL SOB-Medium und Chloramphenicolzusatz mit 20 µl der Übernachtskultur angeimpft, dieser zuerst für 24 Stunden bei Raumtemperatur und anschließend 24 Stunden bei 20 °C unter ständigem Schütteln bei 350 rpm inkubiert bis die OD<sub>600</sub> im Bereich von 2,5 - 3,5 lag. Nach zehnminütiger Inkubation des Deep-well Blocks auf Eis, wurde für 10 Minuten bei 4000 rpm und 4 °C zentrifugiert, der Überstand verworfen und 350 µl TB-Puffer hinzu pipettiert. Anschließend wurde vorgetextet und weitere 10 Minuten auf Eis inkubiert. Es folgte

ein weiterer zehnmütiger Zentrifugationsschritt bei 4000 rpm und 4 °C, bevor erneut der Überstand verworfen und vorgetextet wurde. Die Zellen waren nun bereit für die Transformation und es konnten 80 - 100 µl TB-Puffer mit vorher beigemischter Plasmid-DNA (100 ng/Vektor/Well, Verhältnis pRetro-T2/pJM2545: 1:1) hinzugegeben werden.

Nach der Zugabe von pRetro-T2 und pJM2545 wurde 30 Minuten auf Eis inkubiert und dann ein einmütiger Hitzeschock im Wasserbad bei 48 °C durchgeführt. Anschließend folgte erneut ein dreiminütiger Inkubationsschritt auf Eis. Nun wurde zu den Zellen 1 mL SOB-Medium mit 1,0 mM IPTG zugegeben und eine Stunde bei 37 °C und 350 rpm geschüttelt. Durch den Zusatz von IPTG wird ermöglicht, dass die Expression der Cre-Rekombinase auf dem Plasmid pJM2545 induziert wird, welche wiederum die Rekombination von pCC1FOS und pRetro-T2 zu pCC1FOS-retro über die *loxP*-Seiten katalysiert. Abschließend wurden davon 100 µl in einen neuen Deep-well Block mit 1,1 µl 2x LB-Medium, Chloramphenicol-, Kanamycin und AIS-Zusatz überführt und für 48 h bei 42 °C unter ständigem Schütteln bei 350 rpm inkubiert. Aufgrund seines temperatursensitiven *oris* (bis 30 °C) kann pJM2545 bei 42 °C nicht repliziert werden und geht verloren. Somit ist eine Rückreaktion von pCC1FOS-retro zu pCC1FOS und pRetro-T2 weitgehend ausgeschlossen.

#### 3.2.4.6 Transformation von Zellen des Expressionswirts *T. thermophilus* im 96-Well Format mit Fosmidbanken

Die Transformation in *T. thermophilus* im 96-Well Format erfolgte nach der Methode von Berenguer. Hierfür wurden zunächst mit einer Einzelkolonie von *T. thermophilus* BL03 30 ml TB-Flüssigmedium angeimpft und über Nacht bei 70 °C inkubiert. Mit 15 mL der Übernachtskultur wurden erneut 150 mL frisches TB-Medium angeimpft und bei 70 °C und unter ständigem Schütteln bei 180 rpm solange inkubiert, bis eine OD<sub>550</sub> von 0,7 erreicht wurde. Anschließend wurden 150 mL frisches TB-Medium hinzugegeben und eine weitere Stunde bei 70 °C und 180 rpm inkubiert. Daraufhin wurden 1,2 mL/Well dieser *Thermus*-Suspension auf, die vorher ebenfalls im 96-Well Format extrahierte DNA, gegeben und für weitere zwei Stunden bei 70 °C und 350 rpm inkubiert. In einem letzten Schritt wurden davon 100 µl/Well entnommen und in einen neuen Deep-well Block mit 1,2 mL/Well vorgelegtem TB-Kanamycin-Medium überführt und 48 h bei 60 °C inkubiert.



### 3.2.5 Durchmusterungsmethoden

#### 3.2.5.1 Platten-Screenings

Das Hauptaugenmerk dieser Arbeit lag auf der Suche nach neuen, biotechnologisch anwendbaren Enzymen. Dazu wurde ein Platten-Screening auf den Substraten Carboxymethylcellulose (CMC), fettarme Milch, Xylan, Stärke, Triolein und Tributyrin durchgeführt.

#### 3.2.5.2 Vorbereitung der Substratplatten

Für die Festmedien wurden die spezifischen Medien mit 15 g/l Agar und Antibiotikazusatz erstellt. Festmedien mit den Substraten CMC, Xylan, Stärke, Triolein und Tributyrin konnten auf herkömmliche Weise bei 121 °C für 20 Minuten autoklaviert werden. Das Medium für fettarme Milch wurde dagegen nur für 10 Minuten bei 80 °C autoklaviert, da ansonsten die im Milchpulver enthaltenen Proteine denaturieren würden. Separat davon wurden 7,5 g Agar in 250 mL bidestilliertes Wasser gelöst und für 20 Minuten bei 121 °C autoklaviert. Anschließend konnten Medium und Agar zusammengeführt werden. Bevor Triolein- und Tributyrin-Substratplatten gegossen werden konnten, musste das autoklavierte Medium emulgiert werden, da sich die darin enthaltenen Fette zu Micellen zusammenlagern würden. Zusätzlich wurden dem Trioleinmedium vor dem Gießen der Platten 0,001 % Rhodamin B zugesetzt, welches sich an Triolein anlagert. Ein positiver Klon wäre in der Lage, Triolein und Rhodamin B zu spalten, welches dann durch Bestrahlung mit UV-Licht bei 365 nm nachgewiesen werden könnte.

Tabelle 18: Zusammensetzung der Substrat-Agarplatten

| Substrat                   | Zusammensetzung   | Anmerkung   |
|----------------------------|---|---|
| Xylan                      | LB- oder TB-Medium<br>1 g/L Xylan<br>15 g/L Agar<br>Auf 1 L mit bidest. H <sub>2</sub> O              | pH 7 (bei LB)<br>pH 7,5 (bei TB)<br>autoklavieren   |
| Fettarme Milch (Skim Milk) | 20 g/L Skim milk<br>5 g/L Hefe-Extrakt<br>0,8 g/L NaCitrat<br>Auf 500 mL mit bidest. H <sub>2</sub> O | pH= 7<br>Für 1 – 2 h rühren, 10 min bei 80 °C erhitzen<br>2-fachen 7,5 g Agar in 500 mL bidest. H <sub>2</sub> O autoklavieren und vor dem Platten gießen mit Substrat vermischen |

| <b>Substrat</b>         | <b>Zusammensetzung</b>   | <b>Anmerkung</b>   |
|-------------------------|--|--|
| Stärken                 | 10 g/L Pepton<br>5 g/LK <sub>2</sub> HPO <sub>4</sub><br>2,5 g/L Hefe-Extrakt<br>3 g/L lösl. Stärke<br>15 g/L Agar<br>Auf 1 L mit bidest. H <sub>2</sub> O | pH 7,2 und autoklavieren   |
| Carboxymethyl-cellulose | LB- oder TB-Agar<br>1 g/L CMC<br>15 g Agar<br>Auf 1 L mit bidest. H <sub>2</sub> O   | pH 7 (bei LB)<br>pH 7,5 (bei TB)<br>autoklavieren  |
| Triolein                | LB- und TB-Medium<br>10 g/L Triolein<br>15 g/L Agar<br>Auf 1L mit bidest. H <sub>2</sub> O   | pH 7 (bei LB)<br>pH 7,5 (bei TB)<br>autoklavieren<br>vor dem Gießen emulgieren und<br>0,001 % Rhodamin B zugeben |
| Tributyryn              | LB- und TB-Medium<br>10 g/L Tributyrin<br>15 g/L Agar<br>Auf 1L mit bidest. H <sub>2</sub> O   | pH 7 (bei LB)<br>pH 7,5 (bei TB)<br>autoklavieren<br>vor dem Gießen emulgieren                                   |

Die Fosmidklone in *E. coli* EPI300 beziehungsweise in *T. thermophilus* wurden mithilfe eines 46 Metallstempels auf die verschiedenen Substratplatten übertragen und für *E. coli*-Zellen bei 37 °C und 60 °C jeweils 24 Stunden beziehungsweise für *T. thermophilus*-Zellen für 48 Stunden bei 60 °C inkubiert.

### 3.2.5.3 CMCasen und Xylanasen

Um Klone zu finden die in der Lage sind, die Celluloseverbindungen CMC und Xylan abzubauen, wurden die Substratplatten zunächst durch Abschaben von den darauf befindlichen Zellen befreit und die Platten für 20 Minuten mit 1 %-iger Kongorotlösung eingefärbt. Anschließend wurden die Platten durch 1 M NaCl-Lösung über Nacht entfärbt. Besaß ein Klon eine CMCase- beziehungsweise Xylanase-Aktivität hatte er das jeweilige Substrat lokal abgebaut, sodass an dieser Stelle ein heller Halo auftrat.

#### 3.2.5.4 *Esterasen, Lipasen und Proteasen*

Der Nachweis von Esterase-, Lipase und Protease-Aktivität erfolgte durch trübe, lichtundurchlässige Agarplatten. Klone, die in der Lage waren, die Substrate fettarme Milch, Triolein und Tributyrin abzubauen, konnten durch einen Halo um die Bakterienkultur identifiziert werden. Bei Aktivität auf dem Substrat Triolein wurde zusätzlich Rhodamin B freigesetzt, welches durch seine Fluoreszenz bei 365 nm nachweisbar war.

#### 3.2.5.5 *Amylasen*

Um Amylase-Aktivität nachzuweisen, wurden die Stärkeplatten zunächst durch Abschaben mit 70%igem Ethanol von den darauf befindlichen Zellen befreit, mit Wasser neutralisiert und anschließend mit Lugolscher Lösung etwa 20 Minuten lang inkubiert. Das in der Lugolschen Lösung befindliche Jod lagert sich hierbei in die Stärke-Moleküle ein, so dass sie dunkelviolett angefärbt werden. Bei Amylaseaktivität wird die Stärke allerdings abgebaut, wodurch Jod nicht interkalieren kann und weiße Halos auf den Platten ersichtlich werden.

#### 3.2.5.6 *Vakuumgerät-Manifold (PALL)*

Zur Optimierung der Fosmidisolation im 96 Deep-well Format wurde die Multi-Well Plate Vacuum Manifold von Pall angewendet. Die Apparatur wird in Kombination mit 96 Deep-well Platten, die verschiedene Bindemembranen besitzen, benutzt. Zum Beispiel wird nach dem Resuspendieren und Lysieren der Zellen die Suspension auf die „Lysate Clarification“ Platte pipettiert. Mithilfe des angeschlossenen Vakuums, wird der klare Überstand in die neue Deep-well Platte gezogen und die ausgefallenden Proteine und Zellrückstände in der Membran zurückgehalten. Mit dem gleichen Prinzip funktioniert auch die „DNA-binding“ Platte, wobei hier die aufzureinigende DNA an der Silica-Membran bindet und mit Wasser wieder eluiert wird.

#### 3.2.5.7 *ViaFlo 96 (INTEGRA)*

Zur Effizienzsteigerung wurde der Pipettierroboter Viaflo 96 von der Firma INTEGRA für die Arbeiten im 96 Deep-well Blöcken ausgetestet. Das Gerät kann einerseits manuell gesteuert werden oder Pipettierabfolgen automatisch durchführen; dadurch sollte der Durchsatz der Fosmidpräparationen und vor allem die Reinheit der Fosmide gesteigert werden. Darüber hinaus können durch das präzise und standardisierte Pipettieren Kontaminationen und Pipettierfehler vermieden werden.

### 3.2.5.8 Durchmusterung im Pool

Die vergleichende Durchmusterung von Fosmid-Bibliotheken in *E. coli* und *T. thermophilus* ist mit dem Problem behaftet, dass zwar das Screening der Fosmid-Bibliothek in *E. coli* meist keine Probleme darstellt und gut durchführbar ist, die anschließende Transformation in *T. thermophilus* besonders mit vorhergehender *in vivo* Modifikation jedoch keine gute Effizienz aufweist und damit immer weniger Klone in *T. thermophilus* durchmustert werden als es in *E. coli* der Fall ist. Zusätzlich stellt es einen enormen Zeit- und Kostenfaktor dar. Diesbezüglich könnte das Pool-Screening eine enorme Erleichterung darstellen.

## 3.3 Proteinbiochemische Analysen

### 3.3.1 Quantifizierung der Proteinkonzentration

Durch den Bradford Assay kann mithilfe eines Photometers die Proteinkonzentration im Bereich mg/mL bestimmt werden. Der Coomassie-Brillant-Blau Farbstoff bildet Komplexe mit den unpolaren und kationischen Seitenketten der Proteine und verschiebt dadurch das Absorptionsmaximum von 479 nm auf 595 nm. Die Farbreaktion vom rötlichen zum bläulichen hängt von der Menge der Proteine ab, deshalb ist für die Konzentrationsbestimmung eine Kalibrierung mit einem Protein notwendig. In dieser Arbeit wurde eine Kalibrierungskurve mit dem Standardprotein Bovines Serumalbumin (BSA) durchgeführt. Zur Bestimmung von Proteinkonzentrationen wurde ein Bradford-Reagenz der Firma Serva verwendet. Dazu wurden 10 - 100  $\mu$ L der Probe in eine 1 mL-Küvette gegeben und auf 1 mL mit Bradford-Reagenz aufgefüllt. Nach dem Mischen wurde die Probe für 5 Minuten lichtgeschützt inkubiert und anschließend die OD<sub>595</sub> gemessen. Als Referenz diente 1 mL pures Bradford-Reagenz. Anhand einer zuvor mit BSA erstellten Eichgerade konnte dann die Konzentration der in der Küvette befindlichen Proteine berechnet werden. Die Eichgerade wurde mit BSA-Konzentrationen von 0 - 150  $\mu$ g/Küvette erstellt. Dafür wurde BSA mit einer Konzentration von 10  $\mu$ g/ $\mu$ L verdünnt und in die Küvetten pipettiert, so dass die gewünschte Menge an BSA vorlag. Anschließend wurde mit Bradford-Reagenz auf 1 mL aufgefüllt und nach sorgfältigem Durchmischen und Inkubation bei Raumtemperatur für 5 Minuten die OD<sub>595</sub> gemessen. Die OD<sub>595</sub> wurde dann gegen die BSA-Menge aufgetragen und anhand des linearen Bereichs eine Formel geschrieben, mit der später die Proteinmenge in den Proben berechnet werden konnte. Um die Proteinkonzentration in  $\mu$ g/ $\mu$ L zu erhalten wurde dann durch das eingesetzte Probenvolumen geteilt:

Proteinkonzentration (in µg/Küvette) = OD<sub>595</sub> 164

$$\text{Proteinkonzentration } [\mu\text{g}/\mu\text{L}] = \frac{\text{Proteinkonzentration } [\mu\text{g}/\text{Küvette}]}{\text{Eingesetztes Probenvolumen } [\mu\text{L}/\text{Küvette}]}$$

### 3.3.2 Aufkonzentrierung und Dialyse von Proteinen

Um die Proteinkonzentration zu erhöhen wurde mithilfe von Ultrafiltrationssäulen (Vivaspin 15, MCO 10 kDa, Sartorius Stedim Biotech GmbH, Göttingen) auf ein kleines Volumen eingeeengt. Viele Salzkomponenten können die Aufreinigungen über die Affinitätssäulen stören und sollten vor dem Beladen durch eine PD-10 Säule (Sephadex™ G-25M, GE Healthcare) entfernt werden.

### 3.3.3 Formulierung von Proteinlösungen

Für eine längere Haltbarkeit und bessere Lagerung der aufgereinigten Proteine wurde 20%iges Glycerol hinzugegeben. Bei einer Lagerung bei 4 °C konnte die Proteinlösung für mehrere Wochen verwendet werden, bei -20 °C wurden die aktiven Proteine längerfristig konserviert.

### 3.3.4 Heterologische Proteinexpression und Zellaufschluss mit Ultraschall und mechanisch (Bead-beater)

Zur Gewinnung von Proteinen wurden zunächst 30 mL LB-Medium mit entsprechendem Antibiotikum in 100 mL-Schikanekolben mit Einzelkolonien oder mit einer Übernacht-Flüssigkultur der klonierten Rosetta2-Zellen angeimpft. Beim Animpfen der 1 - 3 Liter großen Expressionskultur wurde zunächst die OD<sub>600</sub> der Vorkultur gemessen und so angeimpft, dass eine Start-OD<sub>600</sub> von 0,1 - 0,2 erreicht wurde. Anschließend wurde bei 37 °C und 180 rpm so lange inkubiert, bis die OD<sub>600</sub> der Zellsuspension bei etwa 0,7 lag. Die Induktion erfolgte durch die Zugabe von 0,1 mM IPTG, woraufhin über Nacht bei 37 °C und 180 rpm weiter inkubiert wurde. Zur Proteingewinnung wird ein Pellet geernteter Zellen in 2 mL 1 M Tris-HCl-Puffer (pH 8) resuspendiert und die Zellen über Sonifizierung aufgeschlossen (Einstellungen: 30-35 %, 5 - 15 Minuten, Intervall 0,5). Eine weitere Möglichkeit der Zellaufschlüsselung ist mithilfe des „Bead-beaters“-Gerät und den dazugehörigen FastPROTEIN™ BLUE KIT Reaktionsgefäßen von der Firma MP Biomedicals. Die Suspension auch Rohextrakt genannt (engl.: *crude extract*, CE), wird bei 13.000 rpm und 4 °C für 5 Minuten zentrifugiert und der

klare Überstand wird in ein neues Reaktionsgefäß überführt (Überstand; engl.: *supernatant*, SN). Wenn mit einem hitzestabilen Protein weiter gearbeitet werden soll, kann ein zusätzlicher Hitzedenaturierungsschritt bei 80 - 85 °C für 10 Minuten eingefügt werden. Danach wird die etwas weißliche ausgeflockte Suspension wieder bei 13000 rpm zentrifugiert und das Pellet verworfen, der mit Hitze behandelte Überstand (engl.: *heat-treated supernatant*, HT-SN) wird weiter verwendet.

### 3.3.5 Proteinaufreinigung

#### 3.3.5.1 Hitzefällung von Zellextrakten

In dieser Arbeit wurde hauptsächlich mit thermostabilen Proteinen gearbeitet, die bei 80 - 95 °C ihre richtig gefaltete Tertiärstruktur und Aktivität nicht verlieren. Diese Eigenschaft wurde genutzt, um das Zielprotein von den restlichen Proteinen im Rohextrakt zu entfernen. Dazu wurden die durch Ultraschall aufgeschlossene Zellsuspension bei 13.000 rpm und 4 °C zentrifugiert. Der Überstand wird dann für 10 Minuten bei 80 °C inkubiert und wieder unter den gleichen Bedingungen zentrifugiert. Der hitzebehandelte Überstand kann nun weiter über Affinitätsäulen aufgereinigt werden oder für Enzymtests eingesetzt werden.

#### 3.3.5.2 Aufreinigung mittels Fast Protein Liquid Chromatographie (FPLC)

Das Zielprotein wurde mithilfe der ÄKTA™explorer System (Amersham Pharmacia Biotech, Uppsala, Sweden) mit P-920 Pumpe, UPC-900 Monitor, INV-907 Verteiler und den Fraktionskollektor Frac-900 aufgereinigt. Über die UNICORN Software konnte die Äkta bedient werden und die Laufbedingungen wurden mit dem WIZARD Programm eingestellt. Störende Partikel wurden durch einen Filter (Minissart 0,2 µm Porengröße, Statorius-Stedim, Göttingen) von Proben, Lösungen oder Puffern entfernt. Alle Anschlüsse wurden durch ständigen Flüssigkeitskontakt auch „drop-to-drop“ genannt, miteinander verbunden, um das Eindringen von Luftblasen zu verhindern. Zwischen dem Umpuffern von Ethanol und dem Laufpuffern wurde stets mit gefiltertem MilliQ-Wasser gespült. Die Säule sollte nach Benutzung zum Regenerieren mit 0,5 mM NaOH und zur Lagerung mit 20 %igem Ethanol (v/v) gespült werden. Die verwendeten Säulen und ihre Parameter sind in Tabelle 19 aufgeführt.

Tabelle 19: Parameter und Säulen für die FPLC-Läufe

|                  | <b>Kationen-<br/>austauscher</b> | <b>Anionen-<br/>austauscher</b> | <b>Gelfiltration</b>                   |
|------------------|----------------------------------|---------------------------------|--|
| Säulentyp        | Source 15S                       | Source30Q                       | HiLoad XK 16/60<br>Superdex200         |
| Puffer A         | 50 mM<br>Phosphatpuffer pH 7     | 50 mM<br>Tris-HCl pH 8          | 50 mM<br>Tris-HCl pH 8,<br>150 mM NaCl |
| Puffer B         | Puffer A + 1 M NaCl              | Puffer A + 1 M NaCl             | -                                      |
| Gradient         | 20 CV                            | 10 CV                           | Isokratisch; 1,2 CV<br>Fraktionierung  |
| Flussrate        | 2 mL/min                         | 2 mL/min                        | 1 mL/min                               |
| Fraktionsvolumen | 2 mL                             | 2 mL                            | 2 mL                                   |

### 3.3.5.3 Aufreinigung mittels Size Exclusion (Größenausschluss)

Nachdem die Proteinlösung aufgrund ihrer Affinität zum Anionen- bzw. Kationenaustauscher aufgereinigt wurde, wird die Reinheit des Zielproteins durch den Größenausschluss über die Size Exclusion Säule erhöht. Dafür wurde eine Hi Load 16/60 Superdex 200 Säule (GE Healthcare) mit dem Puffer 1 (50 mM Tris-HCl; 150 mM NaCl) verwendet. Die Flussrate der Säule wurde auf konstante 1 mL/min eingestellt und zunächst ein Säulenvolumen (ca. 120 mL) mit filtriertem MilliQ-Wasser und dann mit Puffer 1 equilibriert. Über einen 1 mL Loop wird die Probe auf die Säule aufgetragen und die Eluate in 2 mL Fraktionen vom Fraktionskollektor gesammelt. Die Proteine mit einem größeren Molekulargewicht werden zuerst eluiert, da sie leichter durch das Säulenmaterial fließen können, während die kleineren Proteine in die Poren des Säulenmaterials eindringen und einen längeren Weg über die Säule durchlaufen müssen.

### 3.3.5.4 Aufreinigung mittels immobilisierter Metallionen-Affinitäts-Chromatographie

Es wurden Konstrukte mit dem pET21a Vektor erstellt, der ein rekombinantes Protein mit einem HIS-tag exprimieren kann. Die Aufreinigung solcher Proteine erfolgte über eine Metallionen-Affinitätschromatographie (IMAC). Dazu wurde das Protino® Ni-IDA Protein Aufreinigungs-kit (Macherey-Nagel) verwendet und nach Herstellerangaben gearbeitet. Hierzu wurde zunächst ein Pellet der geernteten Zellen in 1 mL 1x LEW-Puffer resuspendiert. Zum Zellaufschluss wurde solange sonifiziert (Einstellungen: 30 - 35 %, 5 - 15 min, Intervall 0,5, für 5 - 15 min) bis die Trübung sichtbar nachließ. Anschließend wurde eine Probe des so

gewonnen "crude extracts" für spätere Analysen abgenommen und aufbewahrt. Der Rest wurde bei 10.000 x g und 4 °C für 30 Minuten zentrifugiert und der Überstand in ein neues Reaktionsgefäß überführt. Die Säule wurde mit 4 mL 1x LEW-Puffer equilibriert. Dann wurde das klare Lysat aufgetragen und der Durchlauf aufgefangen. Es wurde zwei Mal mit 4 mL 1x LEW-Puffer gewaschen und die Fraktionen getrennt aufgefangen. Zuletzt wurde drei Mal mit 3 mL 1x Elutionspuffer das gebundene Protein eluiert und die drei Fraktionen wiederum getrennt gesammelt.

#### *3.3.5.5 Polyacrylamid-Gelelektrophorese*

Für die Auftrennung von Proteinen anhand ihres Molekulargewichts wurden Tris-Glycerin-SDS-Polyacrylamid Gele gegossen. Die Gele liefen durch ein Elektrophoresegerät von Biorad und wurden anschließend mit Coomassie Blue gefärbt und mit einer Lösung aus 30 % Methanol und 10%iger Essigsäure wieder entfärbt (jeweils 10 Minuten bei 60 °C). Als Marker wurde der PageRuler Prestained Protein Ladder von Fermentas verwendet. Man unterscheidet zwischen einem denaturierenden Polyacrylamidgel und einem nativen Gel, bei dem die Proteine sowohl strukturell als auch funktionell intakt bleiben.

Für die Auftrennung von Proteinen nach ihrer molekularen Masse wurden SDS-Gele nach unten stehendem Pipettierschema hergestellt. Die Gele wurden mithilfe von Glaskammern der Firma Biorad so gegossen, dass das Sammelgel ca. 3 cm lang war. Die erstarrten Gele wurden in die Mini-PROTEAN® Tetra System Gelapparaturen eingespannt. Die Proteinproben wurden mit 4x SDS-Ladepuffer gemischt und bei 96 °C für 5 Minuten denaturiert. Nach kurzem Abkühlen wurden sie dann in die Taschen des SDS-Gels pipettiert. Als Größenstandard wurde PageRuler Plus™ Prestained Protein Ladder der Firma Thermo Scientific verwendet. Die Gelelektrophorese wurde bei 15 mA gestartet, bis die Banden am unteren Ende des Sammelgels eine gerade Linie bildeten. Anschließend wurde auf 30 mA erhöht, bis die unterste Bande des Markers den unteren Gelrand erreicht hatte. Das Gel wurde dann vorsichtig aus der Kammer entfernt und für 20 Minuten mit einer Coomassie-Färbelösung (25 % Isopropanol, 10 % Essigsäure, 0,5 % Coomassie Blue) angefärbt und anschließend mit 10%iger Essigsäure entfärbt, bis deutliche Banden sichtbar waren. In der unteren Tabelle 20 sind die Ansätze für jeweils 5 Gele aufgelistet.



Tabelle 20: Zusammensetzung für 4 SDS-Gele

| <b>Trenngel (10 %)</b> | <b>Volumen</b> |
|------------------------|----------------|
| Trenngelpuffer         | 5 mL           |
| H <sub>2</sub> O       | 10 mL          |
| (Bis-) Acrylamid       | 5 mL           |
| APS                    | 10 µL          |
| TEMED                  | 30 µL          |

| <b>Sammelgel (4 %)</b> | <b>Volumen</b> |
|------------------------|----------------|
| Sammelgelpuffer        | 2,4 µL         |
| H <sub>2</sub> O       | 6,6 µL         |
| (Bis-)Acrylamid        | 1 mL           |
| APS                    | 100 µL         |
| TEMED                  | 60 µL          |

#### 3.3.5.6 Silver Staining

Die Silberfärbung ist ein Verfahren, um Proteine, Kohlenhydrate und Nukleinsäuren in einem Polyacrylamidgel sichtbar zu machen. Mit dieser Methode wird eine viel höhere Sensitivität (Nachweisgrenze: 0,1 ng - 1 ng/ Bande) im Vergleich zu Coomassie-gefärbten Gelen oder anderen Proteinfärbungen erreicht. Die Silberionen in der Silbernitratlösung lagern sich an die negativ geladenen Seitenketten der Proteine an und werden durch die anschließende Zugabe von Formaldehyd zu elementarem Silber reduziert. Dadurch werden die Proteine im Gel schwarz gefärbt. In dieser Arbeit wurden die Lösungen und auch das Protokoll des PageSilver™ Silver Staining Kits von Fermentas verwendet.

#### 3.3.5.7 Southern Blot

Die Southern-Blot Methode dient dazu gezielte DNA-Fragmente zu identifizieren und zu quantifizieren. Die gesamte genomische DNA wurde mit einem geeigneten Restriktionsenzym, bei dem die zu identifizierende Bande nicht zerstört wurde und eine definierte Länge besitzt, über Nacht verdaut und für 2 - 3 Stunden (max. 80 V) auf einem 0,8%igen Agarosegel aufgetrennt. Auf die Vakuum-Blot Apparatur wurde nun die Nylonmembran (SensiBlot™ Plus Nylon Membrane, Fermentas) aufgelegt und die restliche Fläche mit einer Kunststoffolie versiegelt, die eine Aussparung für die Membran und das Agarosegel besitzt. Nachdem das

Gel über der Membran platziert wurde, wurden die Gelkammern mit 2%igem Agarosegel verschlossen und das Vakuum (5 Hg) eingeschaltet. Um eine hohe Transfereffizienz zu erreichen, erfolgte zunächst eine Depurinierung der DNA für 15 Minuten mit einer 250 mM HCl Lösung, dabei schlug die bläuliche Farbe des Loading Dyes ins Gelbliche um. Im Anschluss wurde zunächst zweimal für 1 Minute mit MilliQ-Wasser gespült und mit der Denaturierung (20 min in Lösung 2: 1,5 M NaCl; 0,5 M NaOH) fortgesetzt. Nach der Neutralisierung (30 min in Lösung 3: 1,5 M NaCl, 0,5 M Tris, pH 7,5) erfolgte der Transfer für 2 Stunden mit 20x SSC Puffer (0,3 M Na<sub>3</sub>Citrat, 3 M NaCl, pH 7). Der nächste wichtige Schritt war das „Crosslinking“, bei dem mithilfe einer einmütigen UV-Strahlung die DNA kovalent an die Nylonmembran verknüpft wurde. Die Membran wurde anschließend mit der Hybridisierungslösung (1,5 mL 20x SSC; 0,5 mL 50x Denhardt's Solution; 0,25 mL 10 % SDS; 2,5 mL 100 % Formamide (v/v); 0,25 mL MQ) für 1 Stunde bei 42 °C vorhybridisiert. Danach wurde die markierte Sonde (Biotin DecaLabel Labeling Kit, Fermentas) in Hybridisierungslösung gegeben, für 5 Minuten bei 100 °C aufgekocht und über Nacht mit der Membran bei 42 °C hybridisiert. Am nächsten Tag wurde die Membran zweimal für 10 Minuten mit Waschlösung 1 (2x SSC, 0,1 % SDS) bei RT und dann zweimal für 20 Minuten bei 65 °C in Waschlösung 2 (0,1x SSC, 0,1 % SDS) gewaschen. Die Detektion der hybridisierten DNA-Fragmente erfolgte genau nach den Angaben des Biotin Chromogenic Detection Kits. Die Membran wurde getrocknet und zwischen Filterpapieren zur Fotodokumentation aufbewahrt.

## 3.4 Biochemische Charakterisierung

### 3.4.1 $\beta$ -Glukosidase- und $\beta$ -Galaktosidase Test

1,2 mL einer  $\ddot{U}$ N-Kultur wurde pelletiert und in 1,2 mL 0,1 M Phosphatpuffer (60,1 mM  $\text{Na}_2\text{HPO}_4$ , 45,8 mM  $\text{NaH}_2\text{PO}_4$ ) resuspendiert. 200  $\mu\text{L}$  wurden davon abgenommen, um die optische Dichte bei  $\text{OD}_{600}$  zu bestimmen und der Rest der Suspension wurde mit 1 % (v/v) Toluol gemischt. Die Chemikalie permeabilisiert die Zellwand und setzt die Proteine frei. Für die Bestimmung von  $\beta$ -Glukosidase wurde *Para*-Nitrophenyl- $\beta$ -D-Glukopyranosid (pNP-Glu) verwendet und *para*-Nitrophenyl- $\beta$ -D-Galaktopyranosid (pNP-Gal) für die  $\beta$ -Galaktosidase Messung. Drei Ansätze mit jeweils 500  $\mu\text{L}$  der permeabilisierten Zellen, 500  $\mu\text{L}$  Phosphatpuffer und 50  $\mu\text{L}$  einer 50 mg/mL Substratlösung wurden angefertigt und bei 80 °C inkubiert, bis sich die Lösung gelblich (Freisetzung des *para*-Nitrophenols) färbte. Danach wurde die Reaktion mit 200  $\mu\text{L}$   $\text{Na}_2\text{CO}_3$  abgestoppt und für 5 Minuten bei 21.000 x g abzentrifugiert. Der Überstand wurde bei einer Wellenlänge von 420 nm und zur Absorptionskorrektur durch Zellbestandteile bei 550 nm im Photometer gemessen. Mit folgender Formel werden die Miller Units für die Enzymaktivität bestimmt, wobei die Zeit (t) in Minuten angegeben wird und das Küvettenvolumen (V) immer 1 mL entsprach.

$$\text{Eng-Aktivität (Miller Units)} = \frac{1000 \times [\text{OD}_{420\text{nm}} - (1,75 \times \text{OD}_{550\text{nm}})]}{(t \times V \times \text{OD}_{600\text{nm}})}$$

### 3.4.2 Bestimmung reduzierter Zuckerenden mittels DNSA-Test

Der DNSA-Assay ist eine Methode, um die Konzentration an freien Zuckerenden in einer Lösung zu bestimmen. Er kann somit für die Bestimmung von Enzymaktivitäten eingesetzt werden, welche Polysaccharide spalten und somit freie Zuckerreste freisetzen. In dieser Arbeit wurde 1%iges Substrat mit Mcllvaine- oder MES Puffer mit pH 6 verwendet.

Das Experiment wurde folgendermaßen angesetzt (Tabelle 21).

Tabelle 21: Ansätze für den DNSA-Assay

| Lösungen         | Probe  | Referent | Negativprobe |
|------------------|--------|----------|--------------|
| Puffer           | 100 µL | 100 µL   | 100 µL       |
| Zellextrakt      | 20 µL  | -        | 20 µL        |
| Substrat         | 250 µL | 250 µL   | -            |
| H <sub>2</sub> O | 130 µL | 150 µL   | 380 µL       |

Zunächst wurden alle Komponenten nach obigem Schema bis auf das Substrat gemischt und bei 60 °C bzw. 80 °C für 10 Minuten vorgewärmt. Anschließend wurde das Substrat zugegeben und für 10 - 65 Minuten bei gleicher Temperatur inkubiert. Die Reaktion wurde durch Zugabe von 750 µL DNSA-Reagenz abgestoppt und für weitere 5 Minuten bei 95 °C aufgeköcht. Um eventuelle Zellrückstände zu entfernen, wurde dann nach zweiminütiger Inkubation auf Eis für 2 Minuten bei 4 °C und 13000 rpm zentrifugiert. 1 mL des Überstands wurde dann in Küvetten überführt und die OD<sub>575</sub> gemessen. Als Referenz dienten die Proben ohne Proteinzugabe. Anhand einer mit Maltose vorher erstellten Eichgerade konnten dann die in der Küvette befindlichen freien reduzierten Zuckerenden berechnet werden und so schließlich auch die Enzymaktivität.

Die Eichkurve wurde mit Maltose-Konzentrationen von 0 - 1,2 µmol/Küvette erstellt. Dafür wurde eine Verdünnungsreihe von Maltose hergestellt und jeweils so viel der jeweiligen Verdünnung in die Küvette gegeben, dass die gewünschte Maltose-Menge vorlag. Anschließend wurde auf ein Volumen von 1 mL mit DNSA-Reagenz aufgefüllt und wie oben beschrieben für 5 Minuten bei 95 °C aufgeköcht, abgekühlt und anschließend die OD<sub>575</sub> gemessen. Die OD<sub>575</sub> wurde dann gegen die Maltose-Menge aufgetragen und anhand des linearen Bereichs eine Formel erstellt, mit der die Menge an reduzierten Zuckerenden in den Küvetten der Proben berechnet werden konnte:

Reduzierten Zuckerenden (in  $\mu\text{mol/Küvette}$ ) =  $\text{OD}_{575}$  0,449

Um daraus die Aktivität und die spezifische Aktivität berechnen zu können, müssen die Inkubationszeit sowie die Proteinmenge beachtet werden:

$$\text{Proteinmenge [mg/Küvette]} = \frac{\text{Proteinkonzentration } [\mu\text{g}/\mu\text{L}]}{1000} \times 20 \mu\text{L}$$

$$\text{Spezifische Aktivität [U]} = \frac{\text{Reduzierende Zuckerenden } [\mu\text{mol/Küvette}]}{\text{Inkubationszeit [min] / Proteinmenge } [\mu\text{g}]}$$

### 3.4.3 *para*-Nitrophenol (*p*NP) Test

Die Esterase- / Lipase-Aktivität diverser Proteine wurde mit zahlreichen Substraten, die mit *para*-Nitrophenyl (*p*NP) verknüpft sind, getestet. Bei einer Enzymaktivität wird das Chromogen vom Substrat gespalten und die Lösung färbt sich gelb. Die Menge der freigesetzten *p*NP-Moleküle kann im Spektralphotometer bei der Wellenlänge von 400 nm gemessen werden. Für das Experiment wurde entweder 50 mM PBS pH 6 oder Tris-HCl Puffer pH 7,5 verwendet. Das Substrat wurde in unterschiedlichen Lösungsmitteln zu einer Konzentration von 0,1 M aufgelöst. Eine Auflistung der getesteten Substrate und die Probenvorbereitung können in der folgenden Tabelle 22 und Tabelle 23 abgelesen werden.

Tabelle 22: Liste der pNP-Substrate

| Substratname                       | Abkürzung           | Herkunft            |
|------------------------------------|---------------------|---------------------|
| <i>para</i> -Nitrophenol-propionat | pNP-C <sub>3</sub>  | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-butyrat   | pNP-C <sub>4</sub>  | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-valerat   | pNP-C <sub>5</sub>  | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-caproat   | pNP-C <sub>6</sub>  | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-caprylat  | pNP-C <sub>8</sub>  | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-capat     | pNP-C <sub>10</sub> | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-laurat    | pNP-C <sub>12</sub> | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-myristat  | pNP-C <sub>14</sub> | Sigma-Aldrich, USA  |
| <i>para</i> -Nitrophenol-palmitat  | pNP-C <sub>16</sub> | Sigma-Aldrich, USA  |
| 2-Methyldecansäure-pNP-Ester       | -                   | B. Lauinger, Jülich |
| Indancarbonsäure-pNP-Ester         | -                   | B. Lauinger, Jülich |
| Ibuprofen-pNP-Ester                | -                   | B. Lauinger, Jülich |
| Naproxen-pNP-Ester                 | -                   | B. Lauinger, Jülich |
| Mulzer-pNP-Ester                   | -                   | B. Lauinger, Jülich |

Tabelle 23: Ansätze für den pNP-Assay

| Lösungen    | Probe  | Referenz | Negativprobe |
|-------------|--------|----------|--------------|
| Puffer      | 930 µL | 950 µL   | 980 µL       |
| Zellextrakt | 20 µL  | -        | 20 µL        |
| Substrat    | 50 µL  | 50 µL    | -            |

Die Komponenten wurden in Eppendorf-Reaktionsgefäße pipettiert und anschließend bei 60 °C bzw. 80 °C so lange inkubiert, bis eine Gelbfärbung sichtbar war. Die Reaktion wurde mit 200 µL 1 M Na<sub>2</sub>CO<sub>3</sub> abgestoppt und für 5 Minuten auf Eis gestellt. Die Inkubationszeit wurde notiert und am Ende in die Formel mit einberechnet. Um Zellrückstände zu entfernen wurde dann für 2 Minuten bei 4 °C und 13000 rpm zentrifugiert und 1 mL des Überstands wurde anschließend in Kunststoffküvetten bei OD<sub>400nm</sub> gemessen. Als Referenz dienten die Proben ohne Proteinsubstitutionen. Anhand des Extinktionskoeffizienten von *para*-Nitrophenol

(Extinktionskoeffizient bei 400 nm:  $\epsilon = 18300$ ) konnte dann die Konzentration an enzymatisch umgesetztem Substrat berechnet werden und somit auch die Enzymaktivität.

$$\text{pNP [in } \mu\text{mol/Küvette]} = \text{OD}_{400} (\epsilon^{-1} \cdot 10^6) = \text{OD}_{400} \cdot 54,654 \mu\text{mol}$$

$$\text{Aktivität [U]} = \frac{\text{pNP } [\mu\text{mol/Küvette}]}{\text{Inkubationszeit [min]}}$$

$$\text{Spezifische Aktivität [U/mg]} = \frac{\text{Aktivität U}}{\text{Proteinmenge } [\mu\text{g}]}$$

#### 3.4.4 Lipase Assay

Für eine 10 mM Substrat-Stammlösung wurden 2 - 3 mg der jeweiligen Substrate (siehe Tabelle 22) in ein 2 mL Reaktionsgefäß abgewogen und mit DMSO gelöst. Die Methylumbelliferon-Ester Lösung wurde lediglich 1:10 verdünnt, um eine Konzentration von 1 mM zu erhalten. Für die Pufferherstellung werden zunächst 50 mM Dinatriumdihydrogenphosphat Lösung angefertigt und mit Kaliumhydrogenphosphat auf einen pH von 8 eingestellt. Danach werden zu 50 mL jener Lösung 104 mg Natriumdeoxycholat (5 mM) und 50 mg Gummiarabikum (0,1 %) hinzugefügt und vermischt, um 50 mM Sørensen Puffer zu erhalten. Für die Enzymreaktion wurde eine Mischung aus 100  $\mu\text{L}$  Sørensen Puffer (50 mM), 15  $\mu\text{L}$  DMSO und 5  $\mu\text{L}$  der Substrat-Stammlösung pro Reaktion erstellt und dann zum Schluss 30  $\mu\text{L}$  der Enzymlösung dazu gegeben. Die Reaktionslösung wurde bei 40, 60 und 80 °C inkubiert und auf Eis wieder gestoppt. Die Enzymaktivität wurde durch den gelblichen Farbumschlag nachgewiesen und protokolliert.

#### 3.4.5 Qualitative Analyse der Hydrolyseprodukte durch Dünnschichtchromatographie

Die Umsetzung der Poly- und Oligosaccharide durch das EngU Enzym wurde anhand ihrer Hydrolyseprodukte auf der Dünnschichtchromatographie (engl.: *thin layer chromatography*, TLC) untersucht. Die Moleküle wandern aufgrund ihrer Löslichkeit mit unterschiedlicher Geschwindigkeit in der stationären (hydrophilen) und der mobilen (hydrophoben) Phase und werden somit getrennt. Zuerst wurden Reaktionen mit unterschiedlichen Substraten

durchgeführt. Ein Reaktionsansatz enthielt 2 µg aufgereinigtes Enzym, 0,1 % Oligosaccharid (aus einer 1%igen Lösung) bzw. 0,5 % Polysaccharid und wurde auf 600 µL mit MES-Puffer pH 6 aufgefüllt. Es wurde für 4 h bzw. über Nacht bei der für das Enzym optimalen Temperatur inkubiert, wobei ggf. zur Untersuchung der Kinetik zu verschiedenen Zeitpunkten Proben genommen wurden. Die Negativkontrollen enthielten jeweils statt des Enzyms bzw. des Substrats zusätzlichen Puffer. Die Reaktion des hitzestabilen Enzyms EngU konnte bei Raumtemperatur gestoppt und später auf eine mit Kieselgel beschichtete Aluminiumplatte (Typ Siliagel 60 F254, Merck) aufgetragen werden. Mit einem Bleistift wurde eine Linie mit 2 cm Entfernung zum Rand der Platte gezogen. Auf dieser Linie wurden dann 1,5 - 6 µL der Proben pipettiert, wobei stets 1 cm Abstand zu den Rändern und 0,75 cm Abstand zwischen den einzelnen Proben gelassen wurde. Wenn die Probenvolumina mehr als 1,5 µL betragen, musste die Platte zwischen den Auftragsschritten getrocknet werden. Die Chromatographie-Platte wurde dann in eine mit 100 mL Laufmittel befüllte Laufkammer mit Deckel gestellt. Sobald die mobile Phase bis auf 2 cm unter den Plattenrand gewandert war, wurde diese für 10 - 60 Minuten getrocknet und im Anschluss daran durch gleichmäßiges Besprühen mit frisch hergestellter Färbelösung gefärbt. Nach 5 Minuten Trocknungszeit im Abzug wurde die Platte für 15 Minuten bei 120 °C inkubiert. Die Mischungen des Laufmittels sowie der Färbelösung sind in Tabelle 24 aufgelistet.

*Tabelle 24: Lauf- und Färbemittelzusammensetzung für die Dünnschichtchromatographie*

| <b>Laufmittel für DC</b>       | <b>Färbemittel für DC</b> |
|--------------------------------|---------------------------|
| Acetonitril:Wasser (80:20)     | 1 mL Anilin               |
| oder                           | 1 g Diphenylalanin        |
| Butanol:Ethanol:Wasser (5:5:4) | 100 mL Aceton             |
|                                | 85 % Orthophosphorsäure   |

### 3.4.6 Analyse mit der HPAE-PAD (High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection)

Neben der Dünnschichtchromatographie wurden die Hydrolyseprodukte der Oligosaccharide über die HPAE-PAD ermittelt. Getestet wurden Cellotetraose und die beiden Glukotetraosen B und C von Megazyme. 40 mg/L von dem jeweiligen Substrat und 40 mg/L EngU Protein wurden mit 10-fachen MES Puffer pH 5.5 vermischt und bei 90 °C 12 h inkubiert. Bei den Negativkontrollen wurde das jeweilige Substrat weggelassen. Die einzelnen Komponenten des Markers wurden zusammen pipettiert und ergaben eine Mischung aus 10 mg/L Glukose,



Cellulose, Cellotriose und Cellotetrose (C1 - 4) in MES Puffer mit dem pH-Wert von 5,5. 25 µL der Probe wurde zur Analyse in eine Dionex ICS-3000 gespritzt, die auf 30 °C und auf eine Flussgeschwindigkeit von 1 mL/min eingestellt war. Das Programm für die Auftrennung der Abbauprodukte und dem Standard sind aus Tabelle 25 abzulesen.

Tabelle 25: Gradientenlauf für die HPAE-PAD

| <b>Zeit</b> | <b>NaOH</b> | <b>NaAcetat</b> | <b>Zweck</b>                    |
|-------------|-------------|-----------------|---------------------------------|
| 0 - 45 min  | 100 mM      | 0 → 100 mM      | Elution der Oligos              |
| 45 – 47 min | 100 mM      | 0 → 500 mM      |                                 |
| 47 – 49 min | 100 mM      | 500 mM          | Standardreinigung nach dem Lauf |
| 49 – 50 min | 100 mM      | 500 → 0 mM      |                                 |
| 50 – 60 min | 100 mM      | 0 mM            | Reequilibrieren der Säule       |

### 3.4.7 Zymogramm zur Detektion des enzymatischen Abbaus von Polysacchariden

Die Aktivitäts-PAGE dient dazu, hydrolytische Proteine in einem SDS-Gel nachzuweisen. Dazu wurden die löslichen Substrate wie z.B. Lichenan (5 % Endkonzentration) und β-Glukan (0,15 % Endkonzentration) in ein 10%iges SDS-Trenngel gegossen. Zwei Gele wurden identisch beladen und zusammen in einer Gelelektrophoresekammer angeschlossen und bei konstanten 30 mA laufen gelassen. Eins der Gele wird mit Coomassie-Blue gefärbt, um die Proteinbanden zu identifizieren und das andere Gel dient der Aktivitätsbestimmung. Am Ende der Elektrophorese wird die Aktivitäts-PAGE mit Wasser gespült, dann 5 Minuten mit McIlvaine-Puffer und 25 % (v/v) Isopropanol und anschließend 30 Minuten in McIlvaine Puffer renaturiert. Die PAGE wurde über Nacht bei 80 °C in 0,1%igem Kongorot inkubiert und danach für 20 Minuten in 1 M Natriumchloridlösung entfärbt. Das Gel sollte rot gefärbt sein, während die hydrolytisch aktive Proteinbande als heller Hof sichtbar sein sollte.

### 3.4.8 Bindungsverhalten von Proteinen mit unlöslichen Polysacchariden

Kohlenhydrat-prozessierende Enzyme besitzen meistens eine Bindestelle für die Substrate, so genannte Kohlenhydrat-Bindestellen (engl. *carbohydrate binding site*, CBM). Um die Bindungsfähigkeit verschiedener Enzyme nachzuweisen wurde eine 30 mL große Expressionskultur angesetzt. Die Proteinkonzentration des Rohextraktes wurde bestimmt und 300 µg mit 5%igem (w/v) unlöslichen Avicel in 50 mM MES Puffer pH 5 und 200 mM NaCl gemischt. Das gesamte Volumen von 200 µL wurde über Nacht bei 6 °C unter ständigem Schütteln inkubiert. Als Negativkontrolle diente eine Probe ohne Zugabe des Substrates. Nach dem Abzentrifugieren bei 4 °C, 21.000x g für 30 Minuten konnten die ungebundenen Substrate von dem Protein im Pellet getrennt werden. Das Proteinpellet wurde 4 - 5 Male mit dem Bindepuffer gewaschen und wieder in 30 µL Wasser resuspendiert. Die Proben wurden mit 4-fachen Ladebuffer versetzt und elektrophoretisch aufgetrennt. Die Proteinbanden, die Kohlenhydrate gebunden haben und somit eine Bindestelle für solche besitzen sind schwerer als die Proteine, die ohne Substrat inkubiert wurden. Dieser Unterschied wird durch das SDS-Gel sichtbar gemacht.

### 3.5 Datendokumentation und -verarbeitung

Alle für diese Arbeit benötigten Primer wurden mit CloneManager9 hergestellt. Mit diesem Programm ließen sich auch die gewünschten Mutationen für die changeIT-Mutagenese einfügen und die Konstrukte *in silico* erstellen. Zur Identifizierung potentieller katalytischer Reste des EngU-Proteins wurden multiple Sequenz-Alignments gegen Homologie-Modelle aus Pfam (<http://pfam.xfam.org/>) durchgeführt. Dazu wurden die Modelle der Pfam-Familien PF02449 (GH42), PF00150 (GH5) und PF02018 (CBM4) mit ClustalW alignt, stark konservierte Glutamat- und Aspartatreste ausgewählt und mutagenisiert.

## 4 Ergebnisse

### 4.1 Glykosidhydrolase EngU

#### 4.1.1 Isolierung und Charakterisierung des thermostabilen EngU Proteins

Im Rahmen einer Doktorarbeit an der Universität Göttingen (Schuldes, 2008) wurde eine Metagenom-Genbibliothek FosK aus Umweltproben des Avachinsky Kraters in Kamtschatka (Sibirien) hergestellt, die im Rahmen der Doktorarbeit von Dr. Brady (geb. Wittenberg) und nachfolgenden Arbeiten nach Klonen mit Cellulase- und Hemicellulaseaktivität funktionell untersucht wurde (Mientus et al., 2013; Wittenberg, 2009). Der Klon FosK48H3 wies eine Aktivität bei 85 - 90 °C mit den Substraten Carboxymethyl-Cellulose (CMC) und 4-Methylumbelliferyl- $\beta$ -D-cellobiosid (4-MUC) auf. Die Sequenzanalyse führte zur Identifizierung des ORF *bga48*, der für das Protein Bga48 mit einer 38%igen Ähnlichkeit zu einer Glykosidhydrolase aus *Thermotoga petrophila* RKU-1 (Accession no. YP\_001244493) kodiert. Bga48 hat eine vorhergesagte Molekularmasse von 102,7 kDa und ist aus mehreren Modulen aufgebaut (Abbildung 4): Am N-Terminus konnte eine marginale Sequenzähnlichkeit zu einer Glykosidhydrolase der Familie 5 aufgefunden werden (83. - 168. Aminosäure der Gesamtproteinsequenz mit einem e-Wert von  $9.9 \cdot 10^{-7}$ ). Daneben wurden zwei durch etwa 200 Aminosäuren voneinander getrennte Bereiche mit schwacher Ähnlichkeit zu GH42-Familie entdeckt. Der erste Bereich davon GH42a von Aminosäure 150 bis 241 und der zweite Bereich GH42b von der 476. bis zu 616. Aminosäure sind durch eine Kohlenhydrat-Bindedomäne (CBM) voneinander getrennt. In vorangegangenen Studien konnte keine  $\beta$ -Glukosidase Aktivität (GH42) auf Laktose oder 5-Brom-4-Chlor-3-Indoxyl- $\beta$ -D-Galaktopyranosid (X-Gal) gemessen werden. Die ersten 22 Aminosäuren bilden das Signalpeptid (SP). Die Expression ohne das Signalpeptid ermöglichte eine größere Ausbeute von löslichem Protein. Das verkürzte Protein ohne Signalpeptid hat eine Molekularmasse von 100,6 kDa. Die höchste Aktivität wurde mit  $\beta$ -Glukan gemessen (132,9 U/mg, als 100 % relative Aktivität definiert). Mit Lichenan war die relative Aktivität 80 %, während mit den modifizierten Cellulosesubstraten wie CMC und HEC nur 10 % Aktivität gemessen wurde. Bei 90 °C und pH 6 werden die optimalen Bedingungen für das Enzym geschaffen und sogar bis zu einer Temperatur von 105 °C wurde noch signifikante Aktivität gemessen. Mittels Dünnschichtchromatographie wurde der Abbau von  $\beta$ -Glukan verfolgt. Bereits nach 30 Minuten hatten die vorwiegenden Zwischenprodukte einen Polymerisationsgrad von 4 bis 7 und am Ende der Inkubation blieben nur Cellotetraosen. Diese Ergebnisse ließen vermuten, dass Bga48 eine *endo*-Aktivität gegenüber  $\beta$ -1,3-1,4-verknüpften Glukanketten besitzt (Wittenberg, 2009). Aufgrund dieser Ergebnisse wurde das Bga48 Protein in dieser Arbeit auf den Namen EngU umbenannt.

Abbildung 4 zeigt den Sequenzvergleich zwischen EngU und den Pfam-Seeds von GH42 (PF02449), GH5 (PF00150) und dem CB Modul 4\_9 (PF02018). Der GH5-ähnliche Bereich zeigt mit einem E-Wert von  $9,9 \cdot 10^{-7}$  eine Abdeckung des GH5-Seeds im Bereich von der 12. bis zur 82. Aminosäure. In diesem N-terminalen Bereich überlappt der GH5-ähnliche Bereich mit 18 Aminosäuren das GH42a Segment, das sich über den Bereich von der 150. bis zur 241. Aminosäure in EngU erstreckt, während GH42b von Aminosäure 473 bis 585 reicht. Die beiden GH42a und GH42b decken im GH42 Pfam-Seed die Bereiche von der 49–172 bzw. 196–299 mit E-Werten von  $4,1 \cdot 10^{-10}$  bzw.  $0,0041$  ab. Zwischen den Bereichen GH42a und GH42b liegt von der 293. bis zur 430. Aminosäure in EngU das Cellulosebindemodul, welche Ähnlichkeit mit dem gesamten 131 Aminosäuren langen Pfam-Seed von CBM4\_9 aufweist (E-Wert von  $1,1 \cdot 10^{-22}$ ). Der C-Terminus, der sich von Position 665 - 905 der Proteinsequenz ausdehnt, wies trotz Einbeziehung aller verfügbaren Datenbanken keine Ähnlichkeit zu bisher annotierten Enzymen auf.

Nach den aktuellen Sequenzanalysen ist die Ähnlichkeit von EngU zur GH5 Familie zu vernachlässigen und zur GH42 Familie sehr schwach. Der mosaik-ähnliche Aufbau des Enzyms mit den zwei getrennten katalytischen Resten zeigt einen komplett neuen Proteinaufbau (Prof. Bernard Henrissat und Prof. Wolfgang Liebl, persönliche Mitteilung).

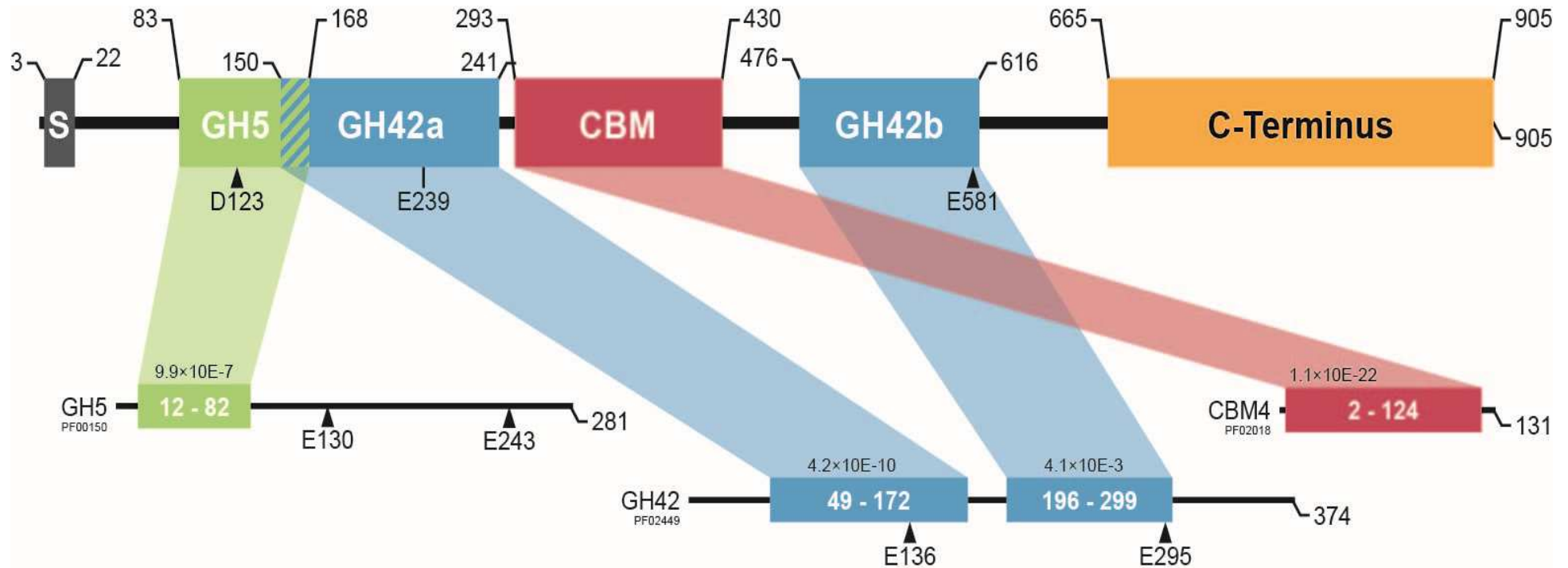


Abbildung 4: EngU Alignment gegen die Seeds der Glykosidhydrolasen-Familien GH5, GH42 (sehr schwache Sequenzähnlichkeit) und der Cellulose-Bindemodul-Familie CBM4 (sehr hohe Sequenzähnlichkeit). Der vorhergesagte Modulaufbau von EngU ist mit den Domänen Lokalisationen (erste Aminosäure  $\uparrow$  letzte Aminosäure) aufgezeichnet und mit den Alignments der jeweiligen Seeds verbunden. Die katalytischen Reste sind mit Dreiecken (▲) gekennzeichnet und die hmmscan E-Werte zu den Domänen sind über den übereinstimmenden Bereichen der Seeds vermerkt. S, Signalpeptid.

#### 4.1.2 Mutagenese der katalytischen Reste (Nukleophil- und Säure/Base-Rest)

Die Mutagenese der katalytischen Reste von EngU führt zum Funktionsverlust des Proteins und damit zur Identifizierung derjenigen Aminosäurereste, welche die Aktivität des Proteins vermitteln. Mit Hilfe von BLAST konnten Proteinsequenzen mit Sequenzähnlichkeiten zu EngU aufgefunden werden. Von den 25 Treffern mit den höchsten Sequenzidentitäten zu EngU wurde ein Alignment erstellt, um hoch konservierte Sequenzbereiche zu identifizieren. Darauf basierend wurden Glutamat- und Aspartatreste des vermuteten katalytischen Zentrums ausgewählt, um sie durch ein katalytisch inaktives Alanin zu ersetzen. Dazu wurden folgende Positionen mithilfe von ortsgerichteter Mutagenese verändert und die Aktivität der erhaltenen mutierten Enzyme via DNSA-Test nachgewiesen und mit dem Wildtyp verglichen (Tabelle 26). Die Proteinvarianten E239A und E581A verloren durch die Punktmutation ihre Aktivität. E289 ist wahrscheinlich der katalytische Säure/Base-Rest, während E581 der nukleophile Rest ist. Die restlichen mutierten Enzymvarianten zeigten entweder einen leichten Anstieg oder eine Reduktion der enzymatischen Aktivität auf (siehe Tabelle 26).

*Tabelle 26: Cellulase-Aktivitäten mutierter EngU Varianten im Vergleich zum EngU Wildtyp (EngUwt). DNSA Assay mit CMC (Endkonzentration 0,2 % w/v), bei 90 °C für 10 Minuten*

| <b>EngU-Varianten</b> | <b>Cellulase-Aktivität</b> |
|-----------------------|----------------------------|
| wt                    | 100,00 %                   |
| D123A                 | 2,97 %                     |
| D151A                 | 104,01 %                   |
| E183A                 | 124,02 %                   |
| E235A                 | 27,84 %                    |
| E239A                 | 0,00 %                     |
| E581A                 | 0,00 %                     |
| E593A                 | 51,35 %                    |
| D605A                 | 32,43 %                    |

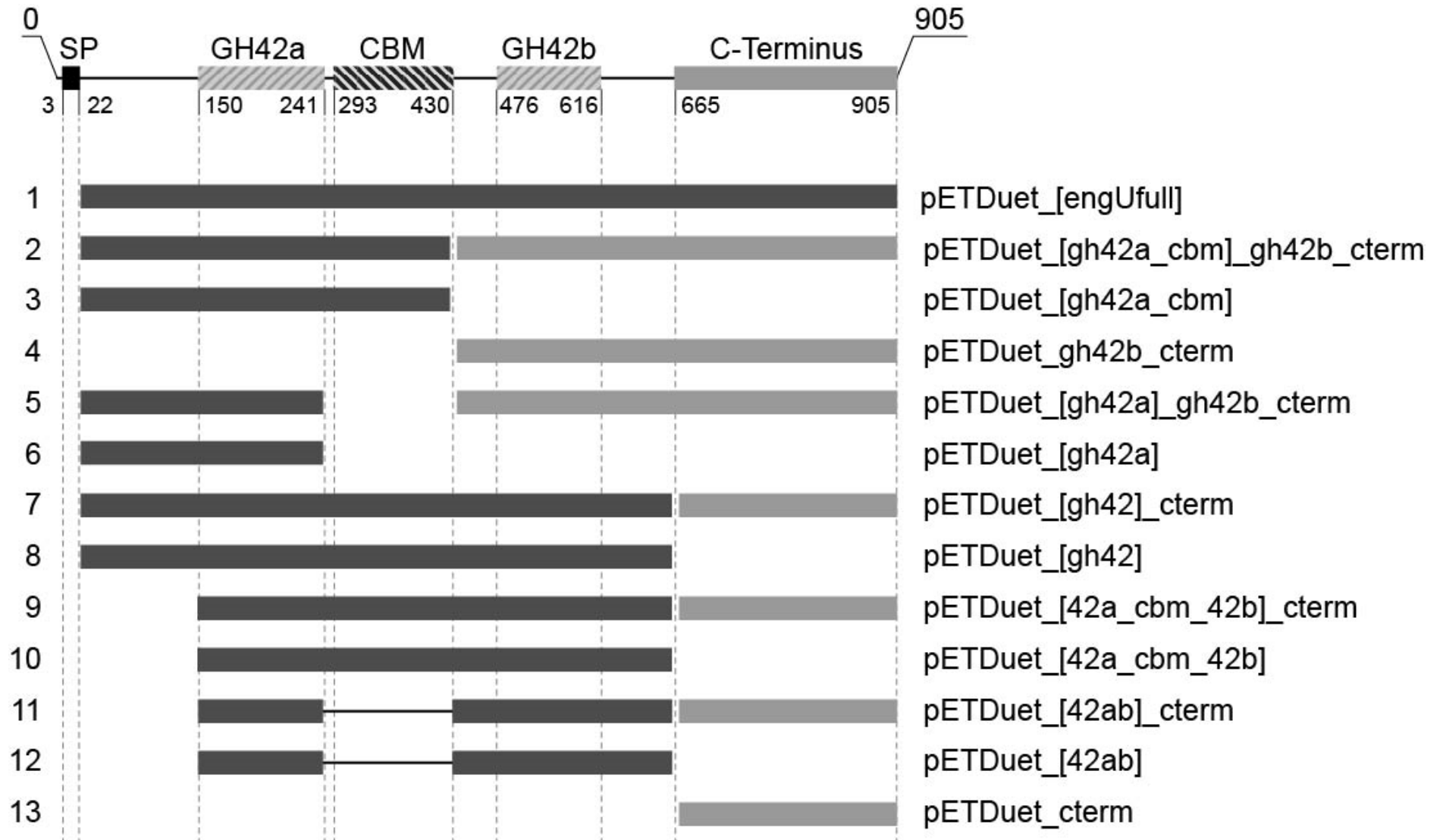


Abbildung 5: Überblick über EngU und die pETDuet-Expressionskonstrukte. SP, Signalpeptid; GH42a und b, zu GH42 ähnliche Segmente; CBM, Kohlenhydrat-Bindemodul; C-Term, C-Terminus. Die dunkelgrauen Balken repräsentieren die kodierenden Bereiche, die in die erste Multiple Cloning Site (MCS) des Vektors pETDuets inseriert wurden, während die mit hellgrauen Balken repräsentierenden Bereiche in der zweiten MCS liegen. Rechts stehen die dazugehörigen Namen, die eckige Klammer beschreibt das Fragment in der MCS1.

### 4.1.3 pETDuet-Expressionskonstrukte und *trans* Komplementation

Die verschiedenen verkürzten und kombinierten Abschnitte der kodierenden Bereiche von *engU* wurden mit entsprechenden Primern (siehe Primerliste im Anhang: Tabelle 38) amplifiziert und in den pETDuet Vektor inseriert. Um die Funktionen der einzelnen Domänen zu erforschen, wurden über klassische Klonierungsmethoden oder über Gibson Assembly Expressionsplasmide hergestellt, die in Abbildung 5 grafisch dargestellt sind. Sie werden immer nach den Genabschnitten, aus denen sie gebaut sind, benannt. Dabei bedeuten die eckigen Klammern, dass dieses Fragment in die erste MCS der pETDuet-Vektor kloniert wurde. pETDuet\_[engUfull] trägt die ganze Gensequenz ohne das Signalpeptid (23. - 905. Aminosäure) in der ersten Klonierungsstelle des Vektors. In Plasmid pETDuet\_[gh42a\_cbm]\_gh42b\_cterm wurde die EngU-kodierende Sequenz nach der 430. Aminosäure in zwei Fragmente aufgespalten und jeweils getrennt in die beiden Klonierungsstellen eingefügt. Durch diese Separation werden im *E. coli* Wirt zwei einzelne Polypeptidketten ausgebildet, die ggf. *in vivo* miteinander assoziieren können. Die Aktivitätsassays bei 60 und 80 °C auf CMC zeigten geringere Aktivitäten (41,02 % bzw. 82,5 %) im Vergleich zum EngUfull-Konstrukt (100 %). Zudem wurden die beiden Teilfragmente auch einzeln im pETDuet-Vektor exprimiert (siehe pETDuet\_[gh42a\_cbm] und pETDuet\_gh42b\_cterm, Abbildung 5, Nr. 3 und 4). Die klonierten Genabschnitte konnten erfolgreich exprimiert und die gebildeten Enzymvarianten gereinigt werden (Abbildung 6), zeigten in den Enzymassays jedoch keine Aktivität. Wie sich das Protein in Abwesenheit des CBM-Moduls verhält, wurde mit den beiden Plasmidkonstrukten pETDuet\_[gh42a]\_gh42b\_cterm und pETDuet\_[gh42a] (Abb. 5, Nr. 5 und 6) überprüft. Die Plasmide pETDuet\_[gh42]\_cterm und pETDuet\_[gh42] (Abb. 5; Nr. 7 und 8) dienten der genaueren Bestimmung der Rolle des C-Terminus. pETDuet\_[42a\_cbm\_42b]\_cterm und die Variante ohne den C-Terminus (Abb. 5; Nr. 9 und 10) wurden konstruiert, um die Funktionalität der zwei Abschnitte mit GH42-Ähnlichkeit Domänen und der Bindedomäne zu testen. Der in den Expressionsplasmiden Nr. 10 und 11 inserierte DNA-Bereich mit der Bezeichnung „gh42ab“ wurde via Gibson Assembly erstellt und fusioniert die beiden zu GH42 ähnlichen Domänen a und b zu einem zusammenhängenden kodierenden Bereich (pETDuet\_[gh42ab] und pETDuet\_[gh42ab]\_cterm). Dadurch wurde untersucht, ob die beiden Teile ursprünglich zusammenhängendes katalytisches Modul verkörperten und lediglich durch das CBM getrennt wurden.



Viele dieser Expressionskonstrukte führten nicht zur Bildung von löslichen Proteinen, so dass es nicht immer möglich war, den einzelnen Bereichen Funktionalitäten zuzuordnen. Eine Proteinprobe aus dem Expressionsstamm pETDuet\_[gh42a\_cbm]\_gh42b\_cterm zeigte nach Auftrennung durch SDS-PAGE zwei Banden der separat exprimierten Fragmente bei 46,9 kDa und 54,5 kDa, die den kodierenden Abschnitten gh42a\_cbm und gh42b\_cterm entsprachen (Abbildung 6; Probe 8). Enzymaktivitätsassays zeigten, dass diese zwei Peptidketten offenbar bei Genexpression in derselben Zelle *in vivo* zu einem funktionalen Enzym assemblierten, jedoch war die Aktivität auf 80,85 % der Ausgangsaktivität des Wildtyp-Proteins (100 %) reduziert. Eine sehr geringe, jedoch nachweisbare Aktivität hatte der Expressionsstamm mit dem Plasmid pETDuet\_[gh42]\_cterm, die 127 Aminosäuren kürzere Variante pETDuet\_[42a\_cbm\_42b]\_cterm (Abbildung 5; Nr. 7 und 9) jedoch nicht.

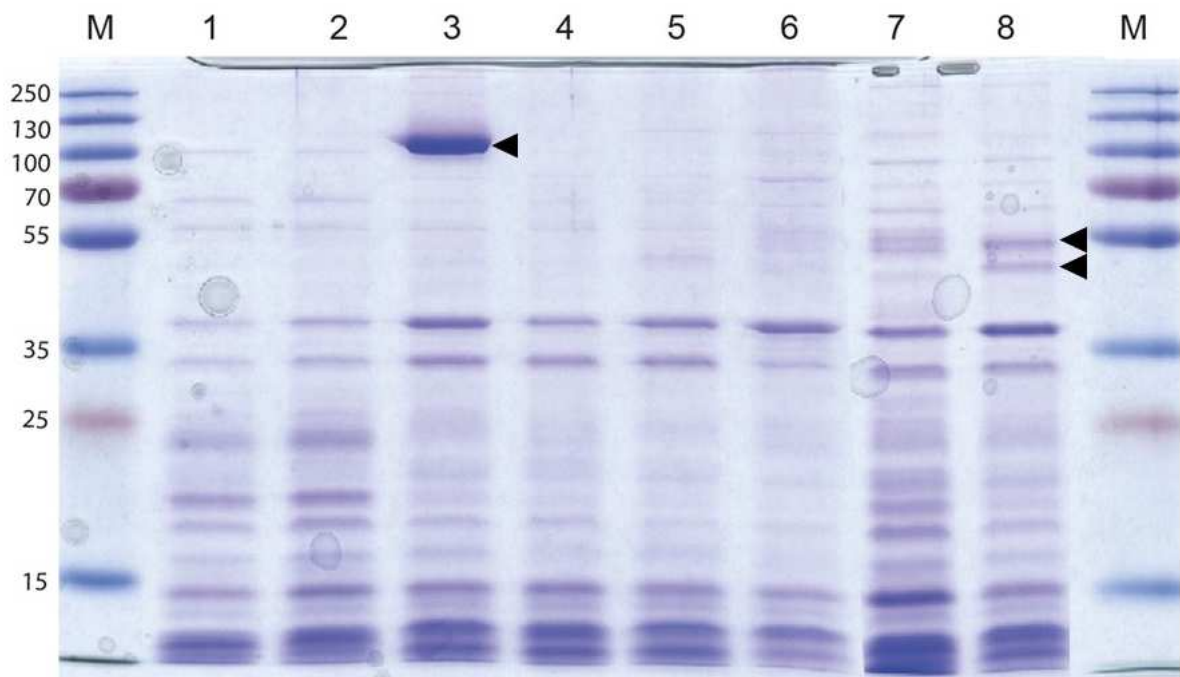


Abbildung 6: SDS-PAGE der hitzebehandelten EngU-Expressionsproteine. 5  $\mu$ L des Protein Markers (Fermentas) und 10  $\mu$ g der jeweiligen Proben. 1) pET21a 2) pETDuet 3) pET21a\_engUfull 4) pETDuet\_[gh42a] 5) pETDuet\_[gh42a\_cbm] 6) pETDuet\_ gh42b\_ cterm 7) pETDuet\_[gh42a]\_gh42b\_cterm 8) pETDuet\_[gh42a\_cbm]\_gh42b\_cterm. Nur EngUfull und pETDuet\_[gh42a\_cbm]\_gh42b\_cterm zeigen exprimierte Proteinbanden (Eng4full: 100 kDa; gh42a\_cbm: 46,9 kDa; gh42b\_cterm: 54,5 kDa mit Dreiecken gekennzeichnet).

Es wurde versucht, mit der Kombination der verschiedenen Fragmente nach der Expression in verschiedenen Stämmen die katalytische Fähigkeit des Proteins *in vitro* zu reaktivieren. Dazu wurden die Polypeptide GH42a\_cbm und GH42b\_cterm (Plasmid Nr. 3 und 4; Abbildung 5) einzeln in Rosetta 2 exprimiert. Dann wurde jeweils der Überstand mit den Proteinbestandteilen gewonnen und ein Teil des Überstandes bei 85 °C für 10 Minuten temperiert. Als Kontrollen dienten Überstände von Stämmen mit dem Leervektor pET21a und mit pET21a\_engUfull. Der hitzebehandelte Teil des Überstandes und der unbehandelte Überstand wurden in verschiedenen Kombinationen auf Aktivität hin getestet (Messung reduzierender Zucker auf 2 % CMC als Substrat). Bei keiner dieser Versuche zur *in vitro* Komplementation getrennt-synthetisierter Enzymteile war eine Cellulaseaktivität messbar.

#### 4.1.4 Substratspektrum und Substratspezifität

Neben den bereits in Vorarbeiten bestimmten Polysacchariden wie z.B.  $\beta$ -Glukan, Lichenan und CMC (Wittenberg, 2009) wurden weitere Hemicellulosesubstrate für EngU bestimmt (Tabelle 27). Dabei wurde auf Xyloglukan (7,38 %) und Glukomannan (27,83 %) eine geringe Aktivität im Vergleich zu  $\beta$ -Glukan (100 %) nachgewiesen.

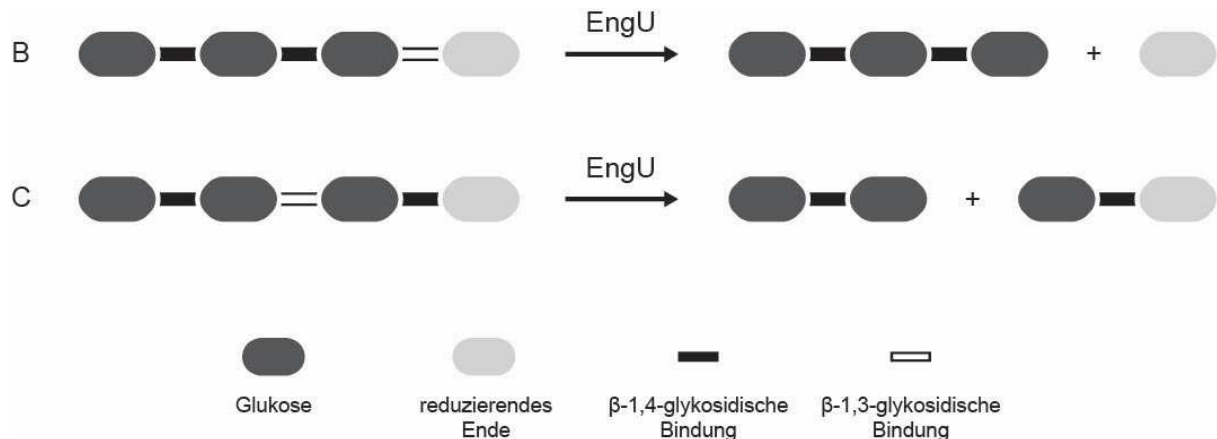


Abbildung 7: Spaltungsmechanismus von EngU mit mixed-linkage Glukotetraosen. Glukotetraose B besteht aus vier Glukosemolekülen. Die ersten drei Einheiten sind über zwei  $\beta$ -1,4-glykosidische Bindungen (schwarze Linie) verbunden und die letzte Einheit mit dem reduzierten Ende ist über eine  $\beta$ -1,3-glykosidische Bindung (weiße Linie) verknüpft. Glukotetraose C hat ebenfalls vier Glukosemonomere mit einem reduzierenden Ende. Hierbei sind zwei Cellobiose-Einheiten (je zwei  $\beta$ -1,4-verknüpfte Glukosereste) über eine  $\beta$ -1,3-glykosidische Bindung verbunden.

Der Spaltmechanismus von EngU konnte mithilfe der unterschiedlich verknüpften Glukotetraosen B und C (Megazym, USA) untersucht werden. Glukotetraose B Variante besteht aus vier Glukoseresten. Die ersten drei sind über zwei  $\beta$ -1,4-glykosidische Bindungen verknüpft und die vierte Glukose, am reduzierenden Ende, ist über eine  $\beta$ -1,3-glykosidische Bindung verknüpft. Glukotetraose C unterscheidet sich darin, dass sich die  $\beta$ -1,3-glykosidische Bindung in der Mitte der Glukotetraose befindet und somit zwei Cellobiosen-Disaccharide (mit  $\beta$ -1,4-glykosidische Bindungen)  $\beta$ -1,3-glykosidisch verbindet (Abbildung 7). An den Spaltprodukten dieser beiden Substrate kann man ablesen, wie spezifisch EngU die Oligosaccharide spaltet. Nach einer Inkubation von 4 Stunden wurden mit Glukotetraose B hauptsächlich Cellotriose und Glukose gefunden, während mit Glukotetraose C mittels Dünnschichtchromatographie nur Cellobiose nachgewiesen werden konnte (Abbildung 8). Die Zuckeranalyse mittels HPAEC-PAD zeigte das gleiche Muster der Abbauprodukte (Abbildung 9). Die erhaltenden Resultate weisen darauf hin, dass EngU mehr als  $\beta$ -Glukanase oder Lichenase zu bezeichnen ist.

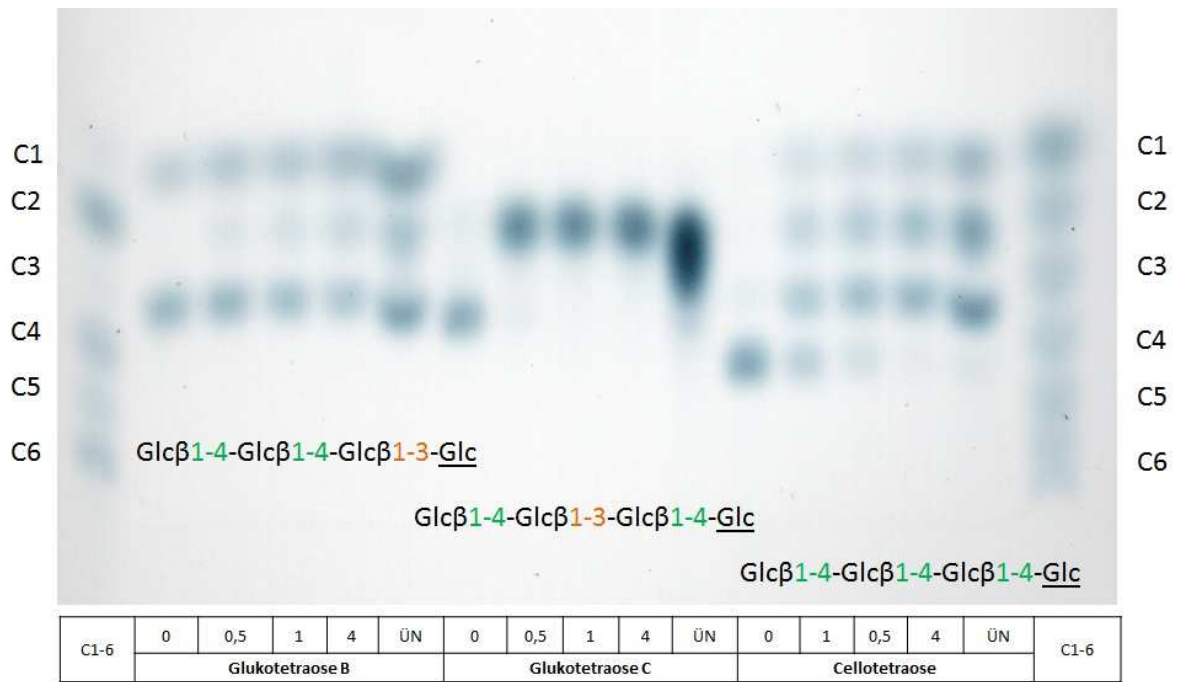


Abbildung 8: Dünnschichtchromatographie der Abbauprodukte der Glukotetraosen und der Kontrolle nach einer Inkubation mit EngU. Es wurden zwei verschiedene Cellooligosaccharid-Marker (C1-C6) aufgetragen.

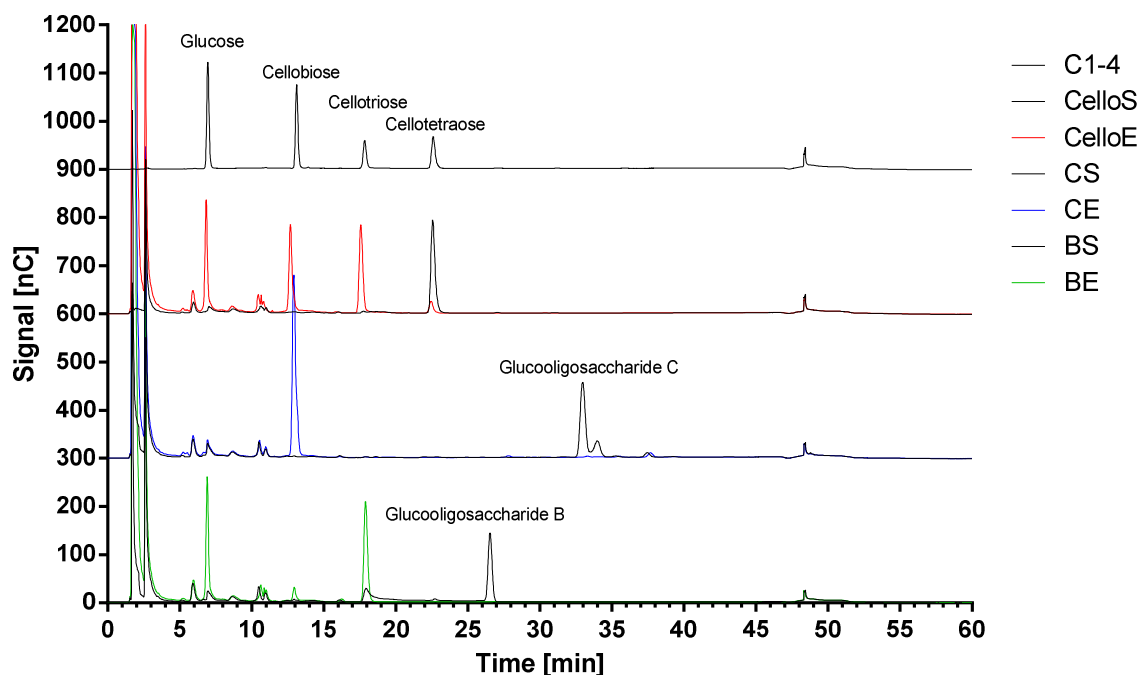


Abbildung 9: HPAEC-PAD Chromatographie der Abbauprodukte der Glukotetraosen und der Kontrolle nach einer Inkubation mit EngU (30 min, 1 h, 4 h und ÜN). Die Hauptprodukte der Spaltung Glukotetraose B waren die Cellotriose und Glukose. Glukotetraose C wurde bereits nach 30 Minuten komplett in Cellobiose ( $\beta$ -1,4-glykosidische Verbindungen) umgewandelt. Aus Cellotetraose entstanden nach Inkubation mit EngU Cellotriose, Cellobiose und Glukose. „S“ bedeutet die Kontrollprobe nur mit dem Substrat und „E“ ist mit Zugabe von EngU.

Tabelle 27: Überblick über die getesteten Substrate (0,2 % w/v Endkonzentration, Aktivitätstest über 10 Minuten bei 90 °C und Messung mittels Bildung reduzierender Zucker, DNSA)

| <b>Substrate</b>                 | <b>Einheiten</b>  | <b>Glykosidische Verbindungen</b> | <b>Rel. Aktivität (%)</b> |
|----------------------------------|---|-----------------------------------|---------------------------|
| β-Glukan (Gerste)                | β-D-Glukopyranose   | β-1,3; β-1,4 linear               | 100,00                    |
| Lichenin ( <i>C. islandica</i> ) | β-D-Glukopyranose   | β-1,3; β-1,4 linear               | 93,60                     |
| Glukomannan (Konjak)             | Mannose:Glukose (60:40)   | β-1,4 linear                      | 27,83                     |
| Carboxymethylcellulose           | β-D-Glukopyranose mit Carboxymethylgruppen                                      | β-1,4 linear                      | 23,77                     |
| Hydroxyethylcellulose            | β-D-Glukopyranose mit Hydroxethylgruppen  | β-1,4 linear                      | 1,93                      |
| Xyloglukan                       | β-D-Glukopyranose und α-D-Xylopyranose Seitenketten                             | β-1,4 linear                      | 0,25                      |
| Pachyman ( <i>P. cocos</i> )     | β-D-Glukopyranose   | β-1,3; β-1,6                      | 0                         |
| Zyosan ( <i>C. cerevisae</i> )   | β-D-Glukopyranose   | β-1,3 linear, Proteinkomplex      | 0                         |
| Curdlan ( <i>A. faecalis</i> )   | β-D-Glukopyranose   | β-1,3 linear                      | 0                         |
| Auxoferm ( <i>S. cerevisae</i> ) | β-D-Glukopyranose   | β-1,3; β-1,6                      | 0                         |
| Laminarin ( <i>L. digitata</i> ) | β-D-Glukopyranose   | β-1,3; 1,6 linear                 | 0                         |
| Mikrocrystalline Cellulose       | β-D-Glukopyranose   | β-1,4 kristalline Struktur        | 0                         |
| Arabinoxylan                     | Arabinose: 36 %, Xylose: 51 %, Glukose: 6.5 %, Mannose: 4.4 %, Galaktose: 1.6 % | β-1,4; β-1,3 verzweigt            | 0                         |
| Dextran                          | α-D-Glukopyranose   | α-1,6; α-1,4; α-1,3;              | 0                         |

|                                    |  |   |   |
|------------------------------------|--|---|---|
| Galaktan                           | Galaktose: 87 %, Arabinose: 5 %, Rhamnose: 1 %, Xylose: 1 %, Galakturonsäure: 5 % andere Zucker      | $\beta$ -1,4; $\beta$ -1,3 verzweigt                  | 0 |
| Galaktomannan                      | $\beta$ -D-Mannopyranose & $\beta$ -D-Galaktopyranose Seitenketten (62:38)                           | $\beta$ -D-1,4 & side chain $\alpha$ -D-1,6 verzweigt | 0 |
| Inulin (Dahlie)                    | Fructose & Glukose   | $\beta$ -2,1 linear                                   | 0 |
| Laminarin                          | $\beta$ -D-Glukopyranose   | $\beta$ -D-1,3 mit $\beta$ -D-1,6 verzweigt           | 0 |
| Levan ( <i>Erwinia herbicola</i> ) | $\alpha$ -D-Fruktopyranose   | $\alpha$ -D-2,6 linear                                | 0 |
| Mannan                             | Mannose: 98 %, Galaktose: ca. 1 %  | $\beta$ -D-1,4 linear                                 | 0 |
| Mannan (Steinnuss)                 | Mannose: 99 %, Arabinose, Xylose: wenig  | $\beta$ -D-1,4 linear                                 | 0 |
| „Pektik Galaktan“ (Lupine)         | Galaktose: 74 %, Arabinose: 17 %, Rhamnose: 3 %, Xylose: 1 %, Galakturonsäure: 5 %, Glukose (Spuren) | $\beta$ -D-1,4 linear                                 | 0 |
| „Pektik Galaktan“ (Kartoffel)      | Galaktose: 82 %, Arabinose: 6 %, Rhamnose: 3 %, Galakturonsäure: 9 %                                 | $\beta$ -D-1,4 linear                                 | 0 |
| Polygalakturonsäure                | $\alpha$ -D- Galakturonsäure   | $\alpha$ -D-1,4 linear                                | 0 |
| Polyoses (Hemicellulose)           | Xylose, Arabinose, Glukose, Mannose, Galaktose   |   | 0 |
| Pullulan                           | Maltotriose (3 x $\alpha$ -D-Glukopyranose)  | $\alpha$ -D-1,4/ Maltotriose: $\alpha$ -D-1,6 linear  | 0 |
| Rhamnogalacturonan I               | $\alpha$ -D-Galakturonsäure & $\alpha$ -L-Rhamnopyranose   | $\alpha$ -D-1,6/ Rha: $\alpha$ -L-1,2 verzweigt       | 0 |
| Sinistrin                          | $\beta$ -D-Fruktopyranose  | $\beta$ -D-1,2 & 1,6 verzweigt                        | 0 |
| Xylan (Birke)                      | $\beta$ -D-Xylopyranose  | $\beta$ -D-1,4 verzweigt                              | 0 |

|   |   |                          |   |
|---|---|--------------------------|---|
| Xylan (Hafer) [Arabino-4-O-methylglukuronoxylan]  | $\beta$ -D-Xylopyranose                               | $\beta$ -D-1,4 verzweigt | 0 |
| Xylan (Lärche) [Arabino-4-O-methylGlukuronoxylan] | $\beta$ -D-Xylopyranose                               | $\beta$ -D-1,4 verzweigt | 0 |
| Laktose   | Galactose, terminal Glukose                           | $\beta$ -1,4             | 0 |
| X-Gal   | 5-Brom-4-Chlor-3-Indoxyl- $\beta$ -D-Galaktopyranosid |                          | 0 |

---

## 4.2 Expressionsanalyse metagenomischer Gene in den Wirtsorganismen *E. coli* und *T. thermophilus*

### 4.2.1 Allgemeine Informationen zu den Fosmiden HA-CMC-1 (CelE12) und TB-EST-7 (Est7)

Identische Geninformation kann in verschiedenen Wirten auf unterschiedliche Weise umgesetzt werden. Diese Tatsache spielt beim funktionellen Screening metagenomischer Genbanken und bei der heterologen Genexpression eine große Rolle und sollte an einem Fallbeispiel mit zwei Fosmiden in den zwei Wirten *E. coli* und *T. thermophilus* genauer durchleuchtet werden. Das Fosmid HA-CMC-1 wurde aus einer Genbank aus metagenomischer DNA von hyperthermophilen Archaeen durch CMC-Platten-Assays entdeckt (Leis, 2015a; Leis et al., 2015b). Auf dem 38.175 bp langen Fosmidinsert befinden sich insgesamt 45 Gene. ORF 23 kodiert für eine Glykosidhydrolase der Familie 12 (GHF12, Pfam Familie PF01670). Die bei 60 - 70 °C aktive Cellulase mit dem Namen Cel12E besitzt eine Masse von 62,3 kDa und einen IEP bei 4,26. Die Aminosäuresequenz zeigt eine 45%ige Ähnlichkeit mit einer *endo*-1,4- $\beta$ -Glukanase b aus *Thermococcus sp.* AM4 (YP\_002581913.2). Fosmid TB-EST-7 hingegen wurde aus einer metagenomischen Genbank thermophiler Bakterien isoliert, dessen Insert metagenomischer DNA umfasst 37 Gene (29.693 bp). Der Kandidaten-ORF 32 kodiert für eine Tributyrase, die bei 60 - 80 °C aktiv ist. Das Enzym hat mit 314 Aminosäuren eine Molekularmasse von 34,92 kDa und einen IEP bei pH 7,86. Der Kandidaten-ORF 32 kodiert für eine Tributyrase, die bei 60 - 80 °C aktiv ist. Die Suche nach ähnlichen Proteinen mittels BLAST-Analysen ergeben eine Patatin-ähnliche Phospholipase aus *Rhodothermus marinus* (WP\_012844179; 90 % Aminosäuresequenz- Ähnlichkeit). Daten zu den beiden Fosmiden sind in Tabelle 28 zusammengefasst.



Tabelle 28: Überblick über die Fosmide HA-CMC-1 und TB-EST-7

|   | HA-CMC-1   | TB-EST-7   |
|---|--|--|
| Metagenombank   | HA<br>(Hyperthermophile Archaeen)                            | TB<br>(Thermophile Bakterien)  |
| Insertgröße   | 38.175 bp  | 29.693 bp  |
| Annotierte Gene   | 45   | 27   |
| ORF-Nr. der durch funktionelles Screening identifizierten Aktivität | 23   | 32c  |
| Ähnlichkeit   | Endo-1,4- $\beta$ -Glukanase aus <i>Thermococcus sp.</i> AM4 | Patatin-ähnliche Phospholipase Familie Protein aus <i>Rhodothermus marinus</i> |
| Aktivität   | Cellulase bei 60 - 70 °C                                     | Tributyrase bei 60 °C  |
| MW  | 62,3 kDa   | 34,92 kDa  |
| IEP   | 4,26   | 7,86   |

Beide Fosmide (pCT3FK\_HA-CMC-1 und pCT3FK\_TB-EST-7; im Folgenden wird „pCT3FK“ in der Bezeichnung der Fosmide weggelassen) wurden sowohl in *E. coli* (EPI300) als auch in *T. thermophilus* (BL03) Zellen eingeführt. Als Kontrollen dienten EPI300 und BL03 Zellen mit dem Leervektor pCT3FK. Die Cellulaseaktivität wurde mithilfe des DNSA-Assays mit dem Substrat CMC und die Tributyraseaktivität mit pNP-Substraten mit unterschiedlichen Kohlenstoffkettenlängen untersucht. Die von Fosmid HA-CMC-1 vermittelte cellulolytische Aktivität konnte nur in *E. coli*, jedoch nicht in *T. thermophilus* nachgewiesen werden. Genau umgekehrt verhielt es sich mit der Tributyraseaktivität von TB-EST-7, die in *T. thermophilus* stärker ausgeprägt war. In *E. coli* liegt das Fosmid als extrachromosomales Plasmid vor, während es in *T. thermophilus* im Chromosom integriert vorliegt. Die maximale mRNA-Produktion findet bei den beiden Organismen möglicherweise während unterschiedlichen Wachstumsphasen statt. Als Ausgangsbasis wurde eine optische Dichte OD<sub>600</sub> von 1

festgelegt. Zu diesem Erntezeitpunkt konnten keine Aktivitäten bei *T. thermophilus* gemessen werden, da sich die Proteine zunächst in den Zellen akkumulieren müssen, bis die Nachweisgrenze überschritten wurde. HA-CMC-1 bewirkte in *E. coli* EPI300 eine deutliche Cellulaseaktivität (orange), während TB-EST-7 in *T. thermophilus* BL03 lediglich eine etwas höhere Tributyraseaktivität als der Kontrollstamm BL03 mit Leervektor in BL03 aufwies (grün) (Abbildung 10). Hier ist anzumerken, dass *T. thermophilus* signifikante Esterase-Hintergrundaktivität aufweist.

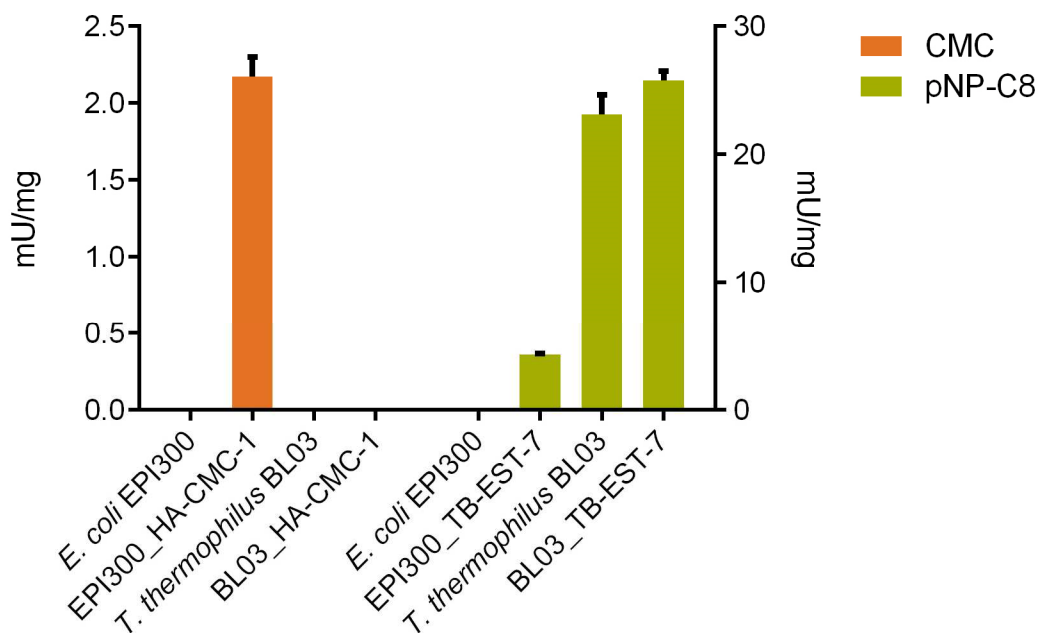


Abbildung 10: DNSA und pNP-Assay mit HA-CMC-1 und TB-EST-7 in *E. coli* und *T. thermophilus*. DNSA Assay mit CMC (Endkonzentration 0,2 % w/v) und 50mM Phosphatpuffer pH7,5, bei 70 °C für 10 Minuten. Die Reaktion wurde mit 750 µL DNSA und 5 min bei 90 °C abgestoppt. 1 mL des abzentrifugierten Überstands wurden bei 575 nm gemessen und mit der Kontrollprobe mit Wasser verglichen. pNP: Inkubation des Proteins, mit pNP-Substrat in 50 mM Phosphatbuffer pH 7,5 bei 60 °C bis Geldfärbung ersichtlich war. Abgestoppt mit 200µL Natriumcarbonat (1 M) für 5 Minuten. 1 mL des abzentrifugierten Überstands wurde bei 400 nm gemessen und mit der Kontrollprobe mit Wasser verglichen.

## 4.2.2 Wachstumskurven der Expressionsstämme

Um die Variablen innerhalb der Stämme für die beabsichtigte Transkriptomanalyse so gering wie möglich zu halten, wurde von allen EPI300 und BL03 Zellen jeweils mit dem Leervektor pCT3FK, pCT3FK\_HA-CMC-1 und pCT3FK\_TB-EST-7 eine Wachstumskurve erstellt, um die jeweilige späte log-Phase zu bestimmen. Während dieser Phase sollte die Transkriptionsmaschinerie die höchste Aktivität aufweisen und der maximale mRNA-Spiegel erreicht werden. Das Wachstum aller Proben wurden bei einer  $OD_{600}$  0,1 gestartet und die *E. coli* EPI300-Stämme bei 37 °C und die *T. thermophilus* BL03 Stämme bei 60 °C schüttelnd inkubiert. Der Übergang der beiden Phasen wurde bei den EPI300 Zellen schon nach ca. 3,5 – 4 Stunden Inkubationszeit erreicht. Die zu dem Erntezeitpunkt gemessenen  $OD_{600}$  Werte liegen bei 2,75 für EPI300\_pCT3FK, 2,68 für EPI300 mit dem HA-CMC-1 Fosmid und für TB-EST-7 lag der Wert bei 3,45. Der *T. thermophilus*-Stamm BL03 benötigte zum Erreichen der späten log-Phase ungefähr 16 - 17 Stunden. Für den Leervektor tragenden Stamm wurde zum Erntezeitpunkt der Wert  $OD_{600}$  2,34, mit HA-CMC-1 der Wert 1,90 und mit TB-EST-7 der Wert 2,10 ermittelt.

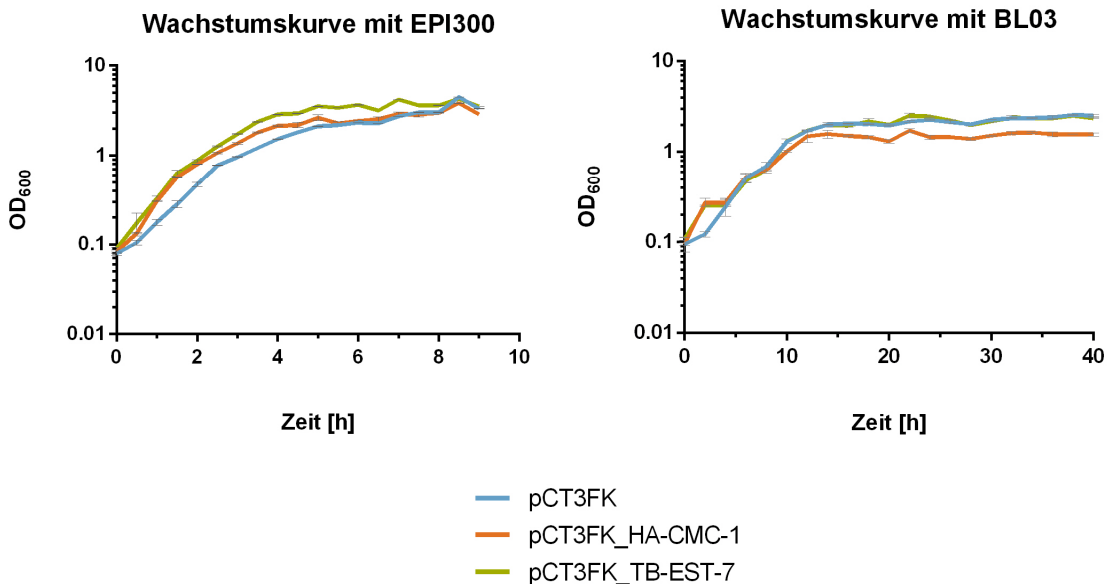


Abbildung 11: Wachstumskurve mit den Fosmiden pCT3FK (Leervektor), pCT3FK\_HA-CMC-1 und pCT3FK\_TB-EST-7 in *E. coli* EPI300 (37 °C) und *T. thermophilus* BL03 Zellen (60 °C). Die späte log bzw. frühe stationäre Phase wird bei den EPI300 Zellen schon nach ca. 3,5 - 4 Stunden Inkubationszeit erreicht. Der *Thermus*-Stamm BL03 benötigt für die richtige Phase ungefähr 16 - 17 Stunden.

#### 4.2.3 Vorbereitungen für die Real-time PCR Läufe

Für die Fosmid-spezifische Transkriptomanalyse wurden für alle 45 ORFs auf HA-CMC-1 und 27 ORFs auf TB-EST-7 Primerpaare, die ein zwischen 100–150 bp großes Amplifikat produzieren, erstellt (Tabelle 39; Tabelle 40). Die Durchführung dieser Versuchsreihe wird im Kapitel „Material und Methoden“ unter dem Punkt 2.2.1.7 beschrieben. Durch PCR-Amplifikation erzeugten 37 von 45 Primerpaaren auf HA-CMC-1 und 17 von 27 bei TB-EST-7 Produkte. Die ORFs, bei denen mit den jeweiligen Primerpaaren kein PCR-Produkt gebildet wurde, wurden für die weiteren Experimente nicht weiter berücksichtigt. Die Spezifität der Primer für die Kandidatengene HA23 und TB32 wurden mit PCR in EPI300 und BL03 Zellen mit und ohne Vektoren getestet. Das Gelbild zeigt die spezifischen Banden mit den richtigen Fosmiden und keine Amplifikate bei den Negativkontrollen (Abbildung 12).

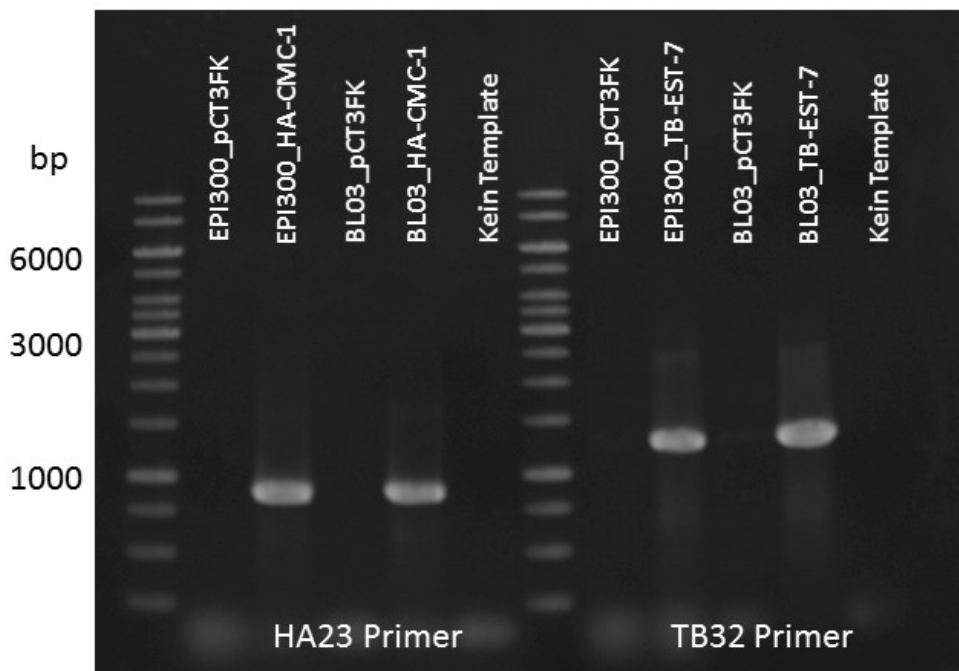
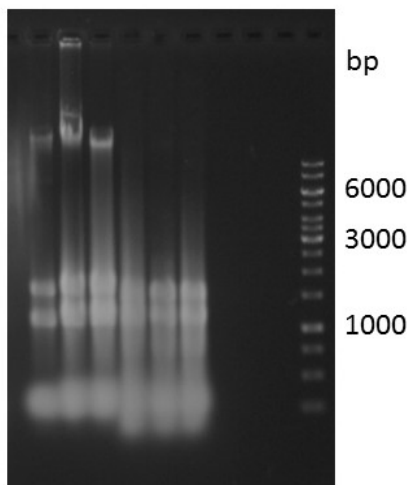


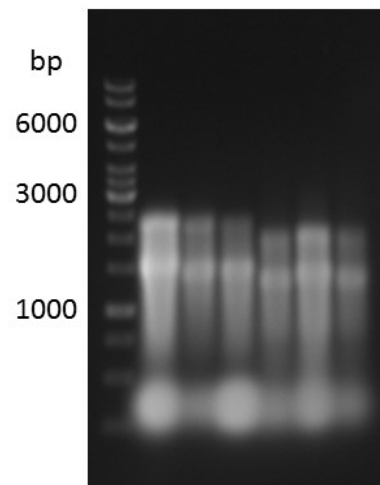
Abbildung 12: Verifizierung der Primerpaare HA23 und TB32 via PCR und Gelanalyse. Die Primerpaare (HA23F/HA23R und TB32F/TB32R) binden spezifisch an den ORFs der Fosmide (HA23 und TB32) in den jeweiligen Wirtsorganismen (*E. coli* und *T. thermophilus*).

### RNA Isolation (MasterPure Kit)



EPI300\_pCT3FK  
EPI300\_HA-CMC-1  
EPI300\_TB-EST-7  
BL03\_pCT3FK  
BL03\_HA-CMC-1  
BL03\_TB-EST-7

### Zusätzliche DNaseI Behandlung (Roche)



EPI300\_pCT3FK  
EPI300\_HA-CMC-1  
EPI300\_TB-EST-7  
BL03\_pCT3FK  
BL03\_HA-CMC-1  
BL03\_TB-EST-7

Abbildung 13: RNA Isolierung mit einer DNaseI Behandlung von Epicentre (links) und mit einer zusätzlichen einstündigen Inkubation mit DNaseI von Roche. RNA-Isolationen des Leervektor pCT3FK und den Fosmiden pCT3FK-HA-CMC-1 und pCT3FK-TB-EST-7 aus *E. coli* (EPI300) und *T. thermophilus* (BL03) Zellen.

Die RNA-Aufreinigung aus eingefrorenen Zellpellets wurde mit den Reagenzien aus dem MasterPure™ RNA Purification Kit von Epicentre durchgeführt. Darin ist eine DNase vorhanden, die mit insgesamt 5 Units pro Probe für 2 Stunden bei 37 °C inkubiert wird. Auf den Gelbildern (siehe Abbildung 13, links) sieht man über 10 kb große Banden, die für eine DNA-Kontamination sprechen. Die RNA-Proben wurden einem weiteren DNase Verdau (Firma Roche; 6 Units/Probe) für 1 Stunde bei 37 °C mit anschließender Inaktivierung (bei 90 °C für 5 min) unterzogen, sodass keine hochmolekularen Banden mehr detektierbar waren (Abbildung 13, rechtes Bild). Ein weiterer Hinweis auf die Reinheit der Proben kann durch einen Testlauf mit oder ohne die Zugabe von Reverser Transkriptase (RT+ und RT-) aufgedeckt werden. Dabei weist eine hohe Differenz zwischen den Ct-Werten von RT+ und RT- Proben darauf hin, dass wenig DNA-Kontamination nach der RNA-Isolierung vorhanden ist und dass das Umschreiben in copyDNA (cDNA) erfolgreich verlief. Bei den Proben aus HA-CMC-1 Fosmid enthaltenden Zellen war ein Unterschied von  $6,635 \pm 0,89$  und  $9,79 \pm 0,55$  in *E. coli* bzw. in *T. thermophilus* zu messen, während bei den Proben aus TB-EST-7 Zellen eine Differenz von 15,71 und 9,12 (keine Mehrfachbestimmung vorhanden) verzeichnet wurde.

#### 4.2.4 HA-CMC-1 und TB-EST-7 qPCR Ergebnisse

Nach den Vorbereitungen und Testläufen wurden quantitative PCR (qPCR) Läufe mit den 37 HA-CMC-1-spezifischen Primerpaaren und den 17 TB-EST-7-spezifischen Primerpaaren durchgeführt, um die Transkriptmengen der einzelnen ORFs in *E. coli* und *T. thermophilus* vergleichen zu können. Für die relative Quantifizierung wird die Transkriptmenge im Vergleich zur jeweiligen Referenzmenge als Änderung des Vielfachen (engl.: *fold change*) ausgedrückt. In dieser Arbeit wurde *E. coli* EPI300 als Referenz verwendet (E) und das Vielfache der Transkripte in *T. thermophilus* BL03 (B) ermittelt. Als Referenz-ORFs wurden die jeweiligen Kandidatengene der Fosmide ausgewählt, welche für die beobachtete Aktivität (CMCase bzw. Tributerase) der jeweiligen Klone verantwortlich waren. Die Transkriptmengen für weitere Gensequenzen auf dem Fosmid HA-CMC-1 (H) wurden mit Referenz-ORF Nummer 23 verglichen, während die Gene auf TB-EST-7 (T) dem Referenz-ORF TB<sub>ORF32</sub> gegenüber gestellt wurden. Für die Berechnung der Unterschiedswerte (sogenannte *fold change*) wurde folgende Formel angewendet:

$$\text{Fold change} = 2^{(-\Delta\Delta Ct)}$$

$$\Delta\Delta Ct = (BH_{ORFX} - BH_{ORF23}) - (EH_{ORFX} - EH_{ORF23})$$

$$\Delta\Delta Ct = (BT_{ORFX} - BT_{ORF32}) - (ET_{ORFX} - ET_{ORF32})$$

Das  $\Delta\Delta Ct$  setzt sich aus den einzelnen  $\Delta Ct$  Differenzen zwischen dem Referenz- und dem Kandidatengen zusammen und sagt aus, wie sich ein bestimmter ORFX zu dem Referenz-ORF innerhalb einer Zelle verhält. Die Auswertungen der Ergebnisse für die Probe aus Zellen mit HA-CMC-1 sind in Abbildung 14 dargestellt und die Resultate für die Proben aus Zellen mit dem TB-EST-7 Fosmid in Abbildung 15 zusammengefasst. In EPI300 Zellen wurden die ORFs 29, 31 und 38 bis zu 4,06 - 5 mal weniger transkribiert als ORF HA23. HA14 besaß ungefähr die Hälfte der HA23 Transkriptmenge (0,42-fach), während alle anderen Insertgene mehr transkribiert wurden bis zu einem Höchstwert von 6,29 mal mehr bei dem Gen HA23. In BL03 wurde nur ORF 31 schwächer als HA23 transkribiert, ansonsten waren alle Transkripte häufiger in der Zelle vorzufinden. Im Vergleich der Transkriptmengen beider Wirte wurden die meisten Gene des HA-CMC-1 Fosmids in *E. coli* Zellen stärker transkribiert. Die mRNA des hypothetische Proteins HA13 (2,86-fach) und der Antiporter Untereinheit HA44 (1,13-fach) waren in *T. thermophilus* häufiger zu messen. Im Vergleich zu dem Fosmid TB-EST-7 war die Verteilung der Transkripte heterogener. Hier wurden in *E. coli* sechs Gene (TB13, TB 14, TB 16, TB21, TB27, TB33) schwächer transkribiert als das Gen für die Phoslipase (TB32). Bei

*T. thermophilus* wurde nur eine Cysteine Synthetase (TB27) mit 6,72-facher Menge weniger detektiert. Die relative Transkriptanalyse für ORF TB28 aus Fosmid TB-EST-7 ergab einen 1,48-fach geringeren  $\Delta\Delta\text{Ct}$ -Wert in *T. thermophilus*. Im Allgemeinen waren in *E. coli* mehr Transkripte zu verzeichnen als in *T. thermophilus*.

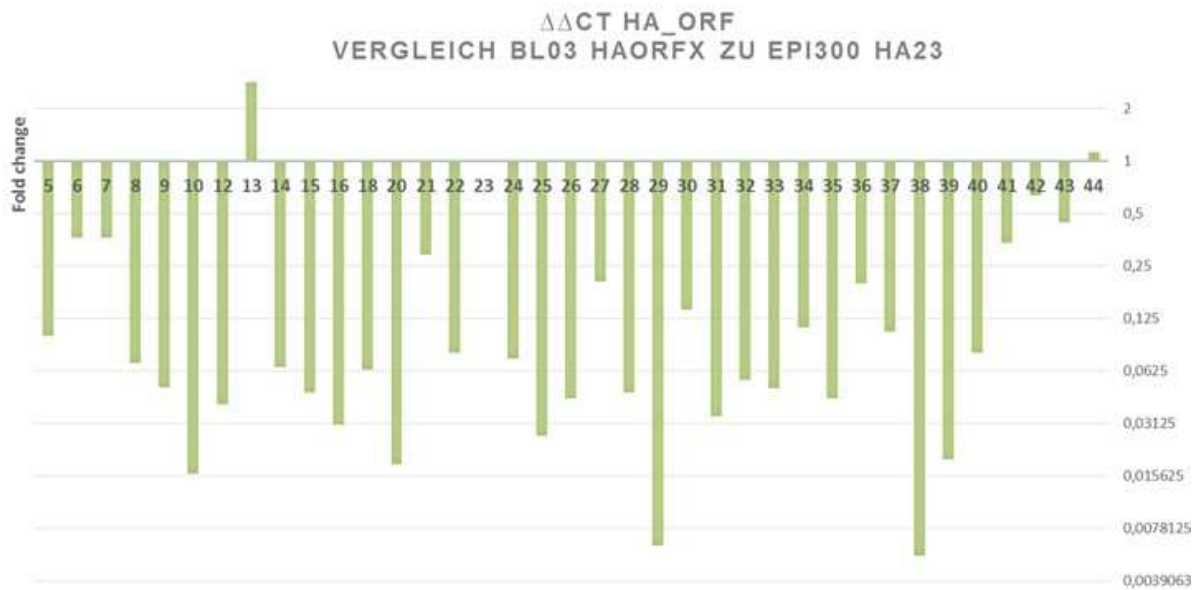
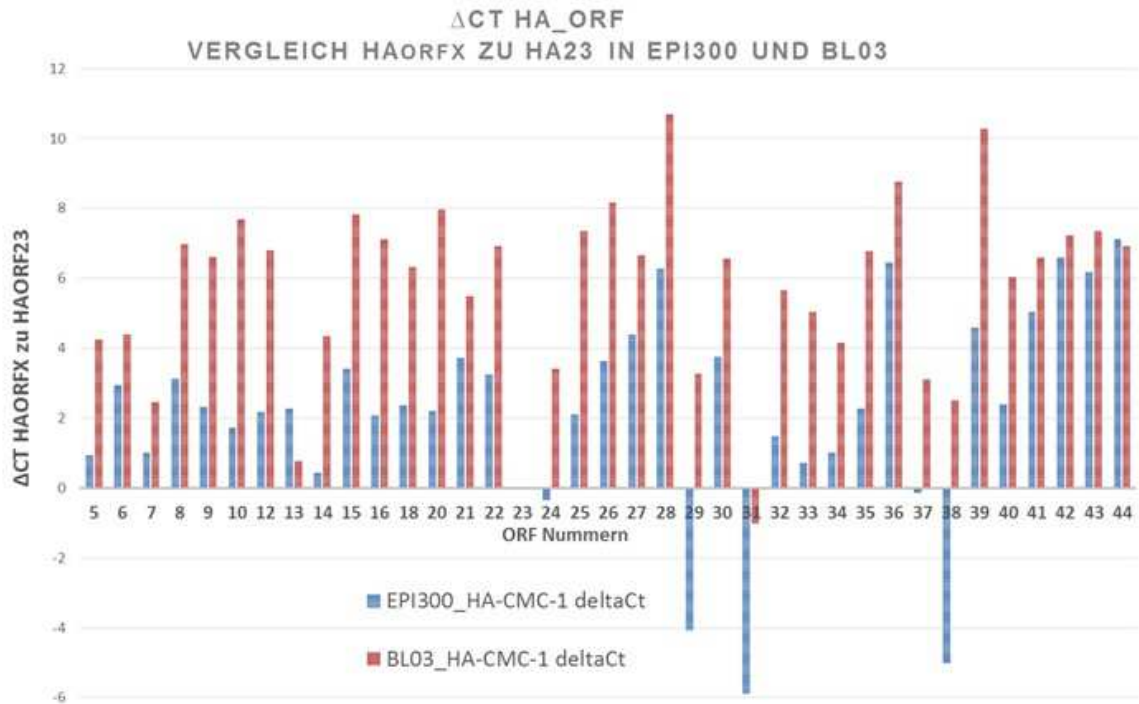


Abbildung 14: Im oberen Diagramm  $\Delta Ct$  HA\_ORF wird die Transkription von 36 ORFs des Fosmides HA-CMC-1 mit dem Endoglucanas-ORF HA23 relativ verglichen ( $\Delta Ct = HAORFX - HAORF23$ ). Die Auswertung für EPI300 (blau) und BL03 (rot) wurden in einen Graphen zusammengefasst. Das untere Balkendiagramm zeigt die vielfache Menge der Thermus-Transkripte im Vergleich zu E. coli EPI300, welches auch als  $\Delta\Delta Ct$  bezeichnet wird und folgendermaßen berechnet wird:  $Fold\ change = 2^{\exp(BHORFX - BHORF23) - (EHORFX - EHORF23)}$ .



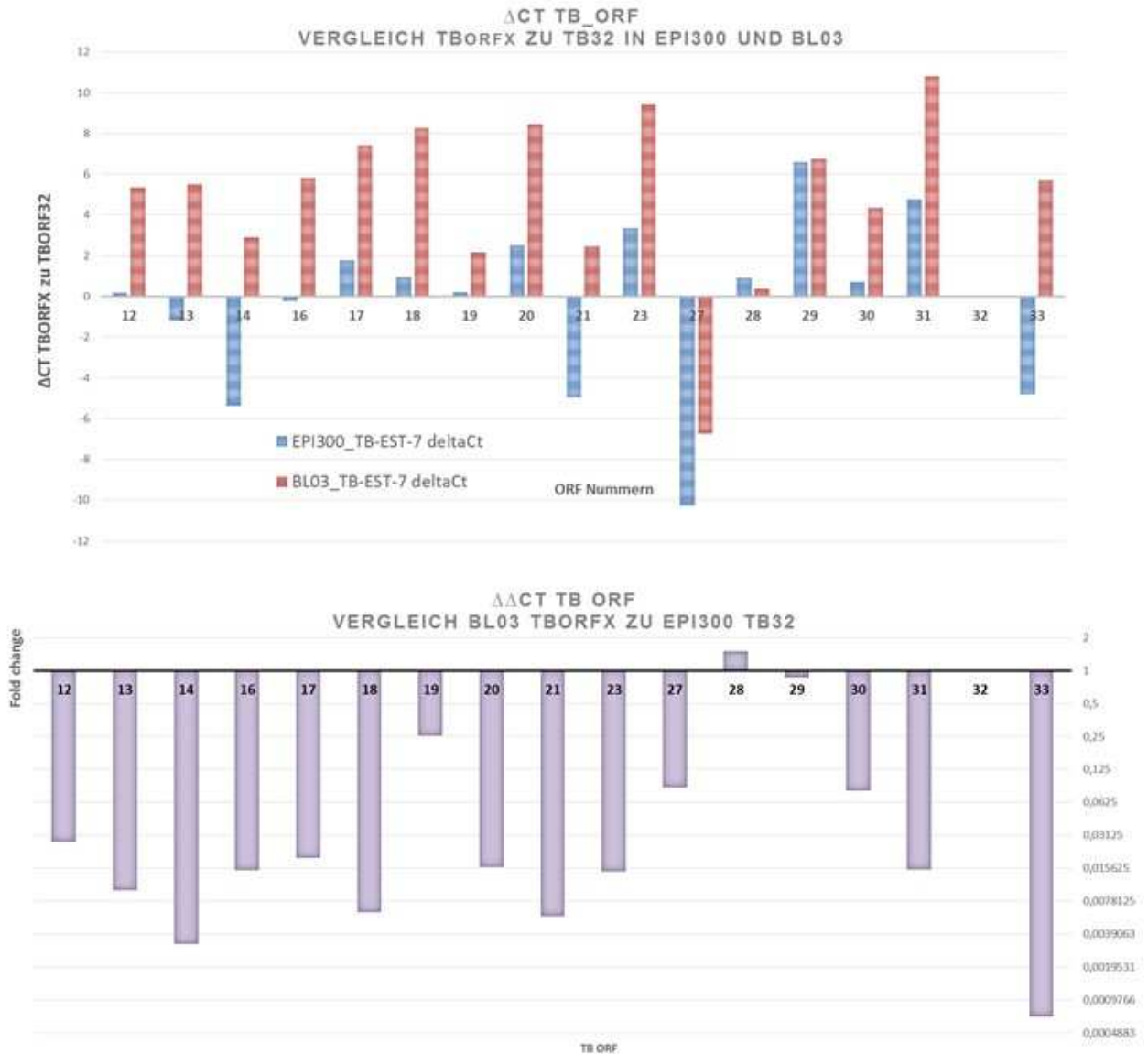


Abbildung 15: In der oberen Graphik werden 16 ORFs aus Fosmid TB-EST-7 mit dem Esterase-ORF TB32 relativ verglichen ( $\Delta Ct = TBORFX - TBORF32$ ). Die Auswertung für EPI300 (blau) und BL03 (rot) wurden in einen Graphen zusammengefasst. Das untere Diagramm zeigt die vielfache Menge der *Thermus*-Transkripte im Vergleich zu EPI300 an, welches auch als  $\Delta\Delta Ct$  bezeichnet wird und folgendermaßen berechnet wird:  $Fold\ change = 2^{\exp(BTORFX - BTORF32) - (ETORFX - ETORF32)}$ .

Tabelle 29: Übersicht zu den analysierten ORFs der HA-CMC-1 und TB-EST-7 Fosmide mit den dazugehörigen Ct-Werten. HA oder H für HA-CMC-1; TB oder T für TB-EST-7; E für *E. coli* und T für *T. thermophilus*.

| ORF Nummer | Protein   | Ct-Wert (EH) | Ct-Wert (BH) | ORF Nummer | Protein  | Ct-Wert (ET) | Ct-Wert (BT) |
|------------|---|--------------|--------------|------------|--|--------------|--------------|
| HA 05      | putative Aminotransferase                           | 28,38        | 26,51        | TB 12      | Methylcrotonyl-CoA Carboxylase Carboxyl-Transferase Untereinheit | 27,75        | 27,36        |
| HA 06      | Initiation Faktor 2 Untereinheit                    | 26,38        | 26,38        | TB 13      | putatives CsbD-ähnliches Protein                                 | 29,07        | 27,21        |
| HA 07      | rRNA Biogenese Protein Nop5/Nop65                   | 28,30        | 28,30        | TB 14      | hypothetisches Protein   | 33,29        | 29,81        |
| HA 08      | Fibrillarin   | 26,19        | 23,79        | TB 16      | Cystathionine gamma-Lyase  | 28,12        | 26,87        |
| HA 09      | Hydroxymethylglutaryl-CoA Synthase                  | 27,00        | 24,14        | TB 17      | Uroporphyrinogen III Lyase                                       | 26,12        | 25,26        |
| HA 10      | Nicht spezifisches Lipid-Transfer Protein           | 27,59        | 23,08        | TB 18      | Siroheme Synthese/Precorrin-2 Oxidase                            | 26,97        | 24,44        |
| HA 12      | hypothetisches Protein                              | 27,13        | 23,95        | TB 19      | Thiosulfate Sulfurtransferase, Rhodanase                         | 27,69        | 30,52        |
| HA 13      | hypothetisches. Protein                             | 27,03        | 30,00        | TB 20      | hypothetisches Protein   | 25,39        | 24,24        |
| HA 14      | hypothetisches Protein                              | 28,89        | 26,42        | TB 21      | 2,3,4,5-Tetrahydropyridine-2,6-dicarboxylate N-Acetyltransferase | 32,88        | 30,25        |
| HA 15      | hypothetisches Protein                              | 25,90        | 22,94        | TB 23      | UDP-Glukose Dehydrogenase  | 24,56        | 23,27        |
| HA 16      | Acetyltransferase verwandtes Protein (GNAT Familie) | 27,23        | 23,66        | TB 27      | Cysteine Synthase  | 38,17        | 39,43        |
| HA 18      | putativer HTH-type Transkriptionsregulator MJ0586   | 26,95        | 24,43        | TB 28      | Mov34/MPN/PAD-1 Familie  | 26,99        | 32,35        |

| ORF Nummer   | Protein  | Ct-Wert (EH)        | Ct-Wert (BH)        | ORF Nummer   | Protein  | Ct-Wert (ET)        | Ct-Wert (BT)        |
|--------------|--|---------------------|---------------------|--------------|--|---------------------|---------------------|
| HA 20        | hypothetisches protein                                       | 27,12               | 22,79               | TB 29        | ThiamineS Protein                                | 21,32               | 25,94               |
| HA 21        | TATA Box bindendes Protein                                   | 25,59               | 25,26               | TB 30        | Sulfur carrier Protein<br>Adenyltransferase ThiF | 27,18               | 28,34               |
| HA 22        | Histone Deacetylase Domäne                                   | 26,05               | 23,85               | TB 31        | putative Thiol-disulfide<br>Oxidoreduktase DCC   | 23,14               | 21,91               |
| <b>HA 23</b> | <b><u>putative Endoglucanase,<br/>GHF12</u></b>              | <b><u>29,31</u></b> | <b><u>30,76</u></b> | <b>TB 32</b> | <b><u>putative Patatin</u></b>                   | <b><u>27,91</u></b> | <b><u>32,71</u></b> |
| HA 24        | Na+/H+ Untereinheit  | 29,65               | 27,34               | TB 33        | Alpha-Glukosidase                                | 32,71               | 27,01               |
| HA 25        | Indolepyruvat Oxidoreduktase<br>Untereinheit lorB            | 27,21               | 23,42               |              |  |                     |                     |
| HA 26        | Indolepyruvat Oxidoreduktase<br>Untereinheit lorB            | 25,68               | 22,61               |              |  |                     |                     |
| HA 27        | Acetyl-CoA Synthetase (ADP-<br>bildend) alpha und beta Kette | 24,93               | 24,09               |              |  |                     |                     |
| HA 28        | Metall-abhängige Hydrolase<br>(beta-Lactamase)               | 23,03               | 20,06               |              |  |                     |                     |
| HA 29        | hypothetisches Protein                                       | 33,37               | 27,50               |              |  |                     |                     |
| HA 30        | putative C <sub>2</sub> H <sub>2</sub> -Typ Zinkfinger       | 25,57               | 24,19               |              |  |                     |                     |
| HA 31        | hypothetisches Protein                                       | 35,19               | 31,78               |              |  |                     |                     |
| HA 32        | hypothetisches Protein                                       | 27,83               | 25,12               |              |  |                     |                     |
| HA 33        | hypothetisches Protein                                       | 28,60               | 25,73               |              |  |                     |                     |

| ORF Nummer | Protein   | Ct-Wert (EH) | Ct-Wert (BH) | ORF Nummer | Protein | Ct-Wert (ET) | Ct-Wert (BT) |
|------------|---|--------------|--------------|------------|---------|--------------|--------------|
| HA 34      | Spermidine Synthetase                                       | 28,31        | 26,60        |            |         |              |              |
| HA 35      | Pyruvoyl-abhängige Arginine Decarboxylase                   | 27,05        | 23,98        |            |         |              |              |
| HA 36      | Natrium/Hydrogen Austausch                                  | 22,88        | 22,00        |            |         |              |              |
| HA 37      | Regulator des Aminosäuren Metabolismus, mit ACT domain      | 29,46        | 27,66        |            |         |              |              |
| HA 38      | Na (+) H(+) Untereinheit D                                  | 34,32        | 28,25        |            |         |              |              |
| HA 39      | Multieinheit Natrium/Hydrogen Antiporter, MnhC Untereinheit | 24,72        | 20,49        |            |         |              |              |
| HA 40      | putative Na (+) H(+) Untereinheit B                         | 26,92        | 24,72        |            |         |              |              |
| HA 41      | Multieinheit Natrium/Hydrogen Antiporter                    | 24,27        | 24,17        |            |         |              |              |
| HA 42      | Na (+) H(+) Untereinheit G                                  | 22,71        | 23,52        |            |         |              |              |
| HA 43      | putative Natrium/Protonen Antiporter Untereinheit MrpF      | 23,13        | 23,42        |            |         |              |              |
| HA 44      | putative Antiporter Untereinheit                            | 22,20        | 23,83        |            |         |              |              |

## 4.3 Biochemische Untersuchungen der kodierten Enzymen auf den Fosmiden pCT3FK-HA-CMC-1 und pCT3FK-TB-EST-7

### 4.3.1 Endoglukanase Cel12E (Fosmid HA-CMC-1)

#### 4.3.1.1 Überproduktion und Reinigung einer verkürzten Variante von Cel12E

In der hyperthermophilen Archaeen Bank wurde über funktionelle Durchforstung auf CMC-Platten ein bei 60 - 70 °C aktiver Klon gefunden. Sequenzanalysen mit dem Fosmid HA-CMC-1 (38.175 bp) ergaben, dass der ORF 23 (1701 bp) für eine Glykosidhydrolase der Familie 12 kodiert (Accession nr.: YP\_002581913.2). Die neue thermostabile Endo-1,4-Glukanase wurde in früheren Arbeiten der Gruppe bereits beschrieben und als Cel12E bezeichnet (Leis et al., 2015b). Das Protein hat eine Molekularmasse von 62,3 kDa (543 AS) und einen berechneten isoelektrischen Punkt bei pH 4,26. Das Gen *cel12E* wurde ohne den für das 24 Aminosäuren lange Signalpeptid in pET21a kloniert. Cel12E besitzt neben der katalytischen GH12 Domäne noch zwei CBM Module der Familie 2. Der nukleophile katalytische Rest ist E171 und der Säure/Base Rest ist E266. Nachdem die biochemischen Eigenschaften bereits experimentell untersucht worden sind, sollte die Tertiärstruktur des Proteins bestimmt werden.

#### 4.3.1.2 Verkürzte Version von HA-CMC (GH12E)

Um bessere Voraussetzungen für die Kristallisierung von Cel12E zu erreichen, wurden die beiden CBMs entfernt und nur das katalytische Modul exprimiert. Die verkürzte Version wurde als GH12E bezeichnet; sie umfasst den Bereich von der ersten Aminosäure bis zu Position 272 von Cel12E. Dazu wurde der Vektor pET21a mit NdeI und NotI geschnitten und über eine Gelextraktion aufgereinigt (5368 bp). Anschließend wurde mit dem Plasmid pCT3FK\_HA-CMC-1 als Template-DNA eine PCR mit den Primern HA-CMC-F und GH12E01 R durchgeführt. Das Amplifikat wurde mit dem geschnittenen pET21a Vektor mithilfe von Gibson Assembly Mix assembliert und das resultierende Plasmid pET21a\_GH12E01 genannt (6188 bp). Das verkürzte Protein hat eine berechnete Molekularmasse von 30,6 kDa und einen leicht veränderten kalkulatorischen pI von 4,33. Bei 80 °C lag die Enzymaktivität bei 7,2 U/mg mit 0,2%igem (w/v Endkonzentration) CMC als Substrat. (Vergleichsdaten von Cel12E und GH12E; Tabelle 41). Das Protein wurde in einer zwei Liter großen Kultur produziert, das Zellpellet in 50 mM Tris-HCl Puffer (pH 8) suspendiert, eine Protease-Inhibitor Tablette zugesetzt und die Zellen mit Ultraschall aufgebrochen. Nach einer 10-minütigen Hitzebehandlung bei 80 °C wurde denaturiertes Protein durch Zentrifugation abgetrennt und das Zielprotein GH12E zuerst über Chromatographie mit Anionenaustauscher Source30Q und anschließend durch Größenausschlusschromatographie Superdex 200 GF 16/60 Säule aufgereinigt.

Tabelle 30: Reinigung von GH12E, einer verkürzten Version von Cel12

| Fraktion                       | Volumen [mL] | Protein [mg] | Aufreinigungsfaktor | Total units X 10 <sup>3</sup> [U] | Spez. Aktivität [U/mg] |
|--------------------------------|--------------|--------------|---------------------|-----------------------------------|------------------------|
| Überstand                      | 17,5         | 191,9        | 1                   | 91,70                             | 477,87                 |
| hitzebehandelter Überstand     | 12,5         | 68,75        | 1,79                | 58,95                             | 857,43                 |
| Anionenaustauscher (Source30Q) | 1,2          | 8,8          | 1,33                | 5,60                              | 636,89                 |
| Gelfiltraton (Superdex 200)    | 2            | 7,8          | 2,56                | 9,55                              | 1224,56                |

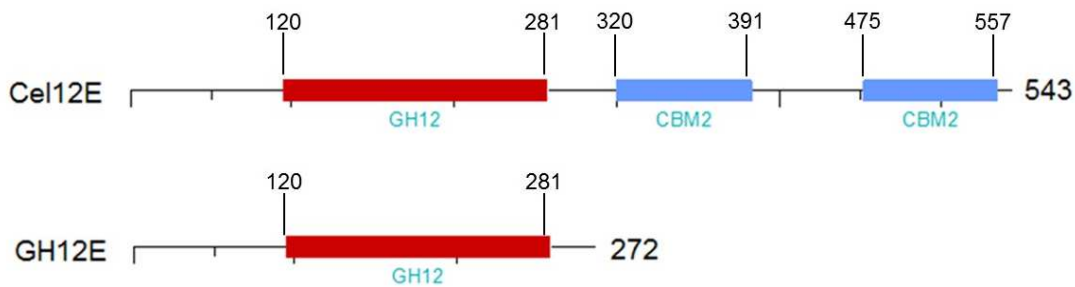


Abbildung 16: Graphische Darstellung des Cel12E Proteins (543 Aminosäuren) mit dem katalytischen GH12 Modul (roter Balken) und zwei Substratbindemodulen (CBM, blaue Balken), sowie die verkürzte Version GH12E mit einer Gesamtlänge von 272 Aminosäuren ohne die CBM-Module.

Die visuelle Kontrolle des Erfolgs der Produktion und Reinigung von GH12E erfolgte über SDS-PAGE Gele (Abbildung 17; GH12E). Für die katalytische Domäne GH12E wurde neben der erwarteten 30 kDa Proteinbande für GH12E auf dem SDS-PAGE Gel eine weitere Bande bei ca. 40 kDa sichtbar. Bei der zusätzlichen Bande handelt es sich vermutlich um unvollständig denaturiertes Enzym. Die Bande lässt sich nicht durch eine weitere Aufreinigung über Gelausschlusschromatographie entfernen.

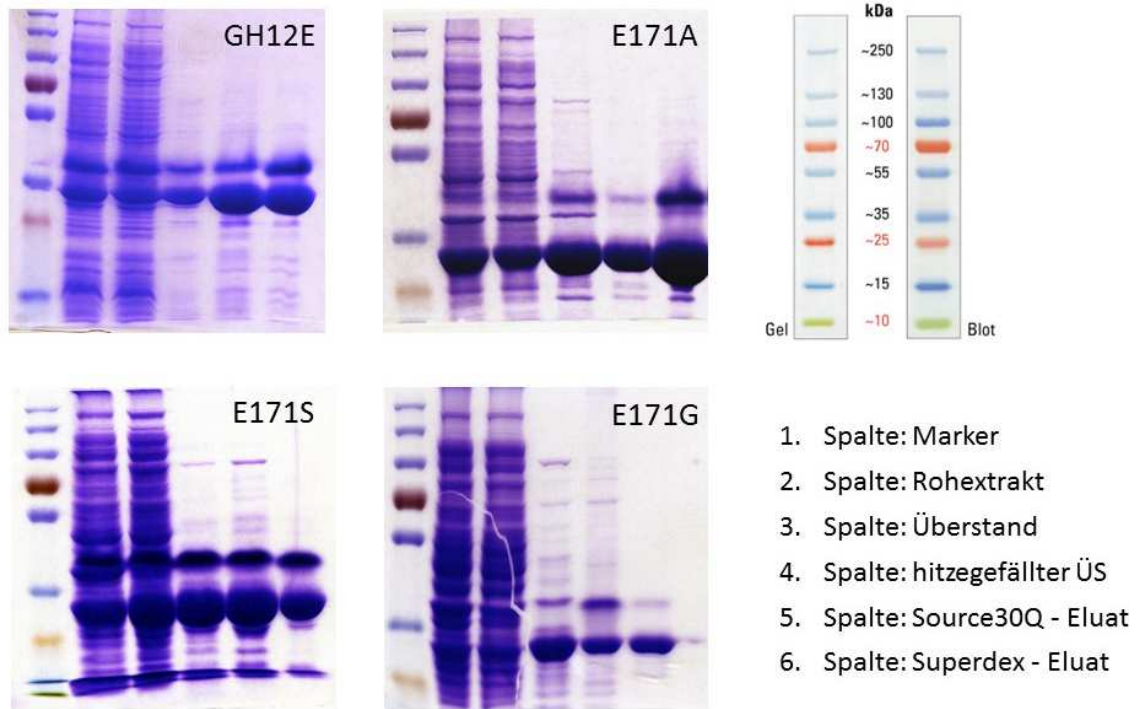


Abbildung 17: SDS-PAGE Gelbilder des verkürzten Cel12E Proteins GH12E und die Mutanten E171A, E171S und E171G. Als Standard wurde der Prestained Plus Marker von Fermentas verwendet, dessen Bandengrößen rechts oben abgelesen werden können. In der ersten Spalte der Gelbilder befindet sich der Marker (5  $\mu$ L), danach folgen die verschiedenen Fraktionen aus der Proteinaufreinigung: Rohextrakt, Überstand, hitzegefällter Überstand, das Eluat nach dem Anionenaustauscher Source30Q und das Eluat aus der Size Exclusion Säule Superdex 200 GF.

#### 4.3.1.3 Gezielte Mutagenese des nukleophilen katalytischen Rests von GH12E

GH12E nutzt wie auch andere retendierende („retaining“) Glykosidhydrolasen einen Mechanismus, der aus zwei aufeinanderfolgenden nukleophilen Substitutionen am anomeren Kohlenstoff besteht. Beim ersten Schritt entsteht aus der Spaltung der glykosidischen Bindung des Substrats und der Verbindung des nukleophilen Rests mit dem Glykon-Anteil des Substrats das sogenannte Glykosyl-Enzym-Intermediat. Bei der Deglykosylierung wird durch ein Hydroxid-Ion die Bindung wieder aufgebrochen und damit das zweite Hydrolyseprodukt vom Enzym wieder freigesetzt.

Glykosynthasen haben die entgegengesetzte Funktion; sie können artifizielle Kohlenhydrate mit bioorthogonalen funktionellen Gruppen auf komplexe Biomoleküle transferieren. Dadurch können neue Biomoleküle gebildet und getestet werden, die neue biologische Funktionen eröffnen. In der Regel werden dafür Glykosidhydrolasen modifiziert, indem der katalytische nukleophile Rest (meist Glutamat oder Aspartat) von „retaining“-Glykosidhydrolasen zu einer kleineren Aminosäure wie z.B. Glycin, Alanin oder Serin ausgetauscht wird, wodurch eine kleine Vertiefung im katalytischen Zentrum entsteht. Durch diese Modifikation können sich aktivierte Substrate, wie am anomeren Kohlenstoffatom fluorierte Monomere, in diese Vertiefung einfügen und das Glykosyl-Enzym-Intermediat nachahmen. Wenn nun die Hydroxylgruppe eines zu verknüpfenden Monomers in der Nähe des anomeren Kohlenstoffs positioniert ist, kann diese durch den katalytischen Säure/Base-Rest deprotoniert werden, wodurch ihre Nukleophilie erhöht wird. Der nachfolgende nukleophile Angriff auf das anomere Kohlenstoffatom ( $S_N2$ ) hat eine Invertierung zur Folge. Die Reaktion wird begünstigt, da Fluorid eine gute Abgangsgruppe ist. Die Konservierung des Säure/Base Restes ist dabei essentiell, um neue glykosidische Bindungen zu knüpfen.



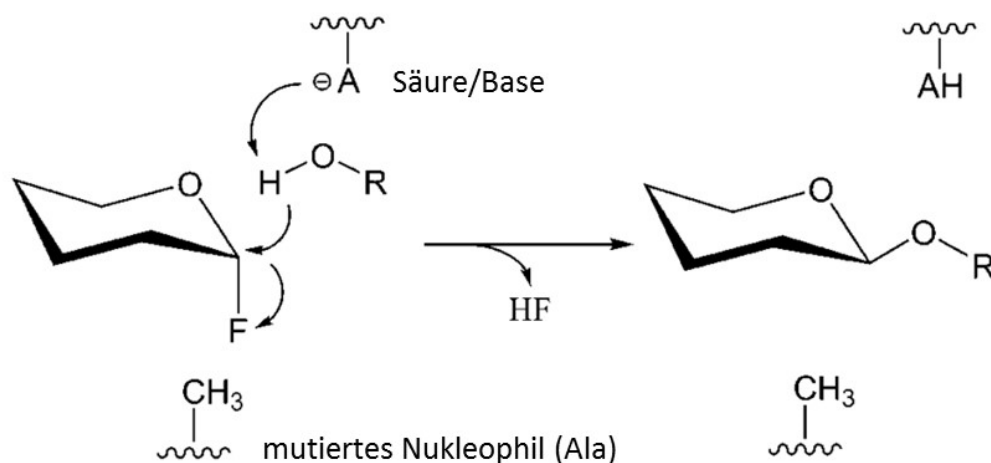


Abbildung 18: Glykosynthese-Mechanismus mit einem mit Fluor substituierten Substrat (nach Vuong et al., 2010, modifiziert).

Für die Umwandlung von GH12E in eine Glykosynthase wurde der nukleophile katalytische Rest E171 des GH12E Proteins durch drei verschiedene Aminosäuren ausgetauscht. Dazu wurde in einer Reaktion das Template pET21a\_GH12E01 mit dem Amp REV Primer aus dem ChangeIT™ Kit (Affymetrix) und den drei phosphorylierten Mutagenese-Primern E171A F, E171S F, E171G F gemischt und das Mutageneseprotokoll nach Angaben des Herstellers durchgeführt. Danach wurden einige Klone ausgewählt, deren Plasmide aufgereinigt und mittels Sequenzierung verifiziert. Danach wurden die Proteine nacheinander durch Expression der mutierten Gene produziert und wie GH12E über die Säulen Source30Q und 200GF aufgereinigt. Da die mutierten Enzyme keine Cellulaseaktivität mehr besaßen, wurden für die Proteinaufreinigung die gleichen Laufprogramme und -bedingungen wie bei GH12E verwendet. Die einzelnen Fraktionen wurden auf SDS-PAGE Gelbildern dokumentiert (Abbildung 17). Abschließend wurden mit den reinen Proteinen unter Einsatz von 0,2 % CMC als Substrat DNSA-Assays durchgeführt, um die Inaktivierung der Enzymaktivität zu verifizieren. Die Enzymmutanten E171S und E171G waren komplett inert, während die Alanin-Mutante (E171A) eine geringe Hintergrundaktivität von  $0,084 \pm 0,045$  U/mg aufwies, im Vergleich zum Wildtyp mit  $8,74 \pm 0,27$  U/mg. Die aufgereinigten Proteine wurden mit 20%igen Glycerol versetzt an die Swedish University of Agricultural Science (Abteilung der Chemie und Biotechnologie; Jerry Stahlberg) für die weiteren Schritte der Kristallisation verschickt.

#### 4.3.2 Weitere Untersuchungen von auf TB-EST-7 kodierten Enzymen

Zwei weitere interessante Gene, die neben *est7* (siehe unten) auf dem Fosmid TB-EST-7 lagen, waren die ORFs für UDP-Glukose Dehydrogenase (TB23) und die  $\alpha$ -Glukosidase (TB33). Die UDP-Glukose-6-Dehydrogenase (UDPGDH) wandelt UDP-Glukose zu UDP-Glukuronat um, während NAD<sup>+</sup> zu NADH reduziert wird. Für den Assay wurde das Substrat UDP-Glukose und NAD mit dem Rohextrakt (bzw. hitzebehandelten Überstand) des pCT3FK-TB-EST-7 Fosmids aus *E. coli* und *T. thermophilus* getestet. *E. coli* besitzt bereits eine hohe Hintergrundaktivität (EPI300\_pCT3FK), so dass keine verstärkte Reaktion in den Zellen mit dem Fosmid nachgewiesen werden konnte. In BL03 Stamm konnte mit dem UDP-Glukose-Dehydrogenase Assay weder bei der Negativkontrolle noch bei dem Fosmid-tragenden Stamm eine Veränderung im NADH Spiegel gemessen werden. Für den Nachweis einer  $\alpha$ -Glukosidase Aktivität wurde in einem Assaypuffer Rohextrakt mit *para*-Nitrophenol- $\alpha$ -Glukopyranose gemischt und der Abbau des Substrats anhand der Gelbfärbung im Photometer bei OD<sub>400</sub> gemessen. In diesem Fall war jedoch im Kontrollstamm *T. thermophilus* BL03\_pCT3FK bereits eine hohe Hintergrundaktivität (266,72 - 457,55 Miller Units bei verschiedenen Temperaturen zwischen 70 - 100 °C) messbar, welche nicht durch den Besitz des Fosmides pCT3FK-TB-EST-7 erhöht wurde. Währenddessen wurde weder mit dem rekombinaten Fosmid noch mit der pCT3FK-Kontrolle Aktivitäten in den *E. coli* Stämmen nachgewiesen (Daten nicht gezeigt).

#### 4.3.3 Rekombinante Herstellung von Est7 und Aktivitätsmessung mit Lipase-Substraten

Aus der Genbank der thermophilen Bakterien (TB) mit 100.000 Einzelklonen (P. Golyshin, Bangor University, England) wurde der lipolytisch aktive Klon TB-EST-7 entdeckt und im Zuge dieser Arbeit näher charakterisiert. Das Zielgen *est7* (ORF Nummer 32 der metagenomischen DNA-Insertion von TB-EST-7) umfasst 954 Basen und kodiert für ein 317 Aminosäuren langes Protein (MW 34,93 kDa). Die katalytischen Aminosäuren von Est7 sind Serin 29 und Asparaginsäure 161 und das in lipolytischen Enzymen konservierte Pentapeptidmuster (G-X-S-X-G) ist vorhanden (Leis, 2015a). Verschiedene Vektoren wurden mit der Gensequenz *est7* aus TB-EST-7 hergestellt und zur Genexpression in die kompatiblen Wirte eingeführt. Das Expressionsplasmid pET21a\_Est7 besteht aus dem Rückgrat pET21a (5345 bp) und dem Insert *est7* (974 bp) mit einer für einen N-terminalen His-Tag (6x His) kodierenden Sequenz und wurde im *E. coli* Stamm BL21 als Expressionswirt eingebracht. Nach der Proteinaufarbeitung wurde durch SDS-PAGE Analyse klar (Abbildung 19, A), dass viel Protein im Zellpellet (P) enthalten war und nach der Aufreinigung über Metall-

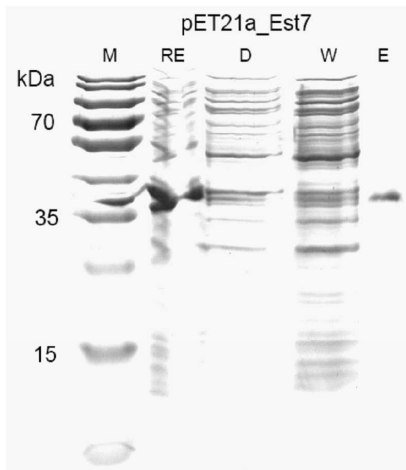
106

Affinitätschromatographie sich zwar sehr wenig aber sehr reines Protein mit einer Molekularmasse von 36 kDa im Eluat (E) befand. Bei den pNP-Assays gab es große Schwankungen zwischen den mit Substraten unterschiedlich langer Kohlenstoffkettenlängen ermittelten Aktivitäten. Dies lag teilweise an der schnellen Verfärbung der Assayansätze nach dem Hinzufügen des Proteins, vor allem bei den kurz- bis mittelkettigen Substraten von C4-C12. Dies wird durch die großen Standardabweichungen deutlich. Dennoch zeigten die Aktivitätsmessungen, dass Est7 die kurzkettigen Esterase-Substrate bevorzugte (Abbildung 17, A).

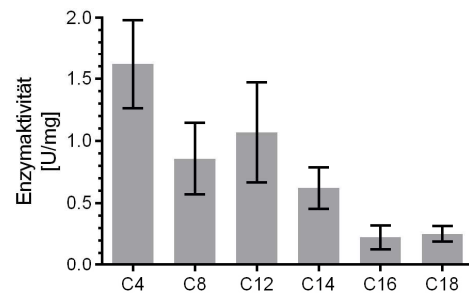
Als weitere Möglichkeit der Produktion der Esterase Est7 wurde die Expression in *T. thermophilus* erprobt, um zu untersuchen, ob die Proteinbildung und -faltung in dem thermophilen Stamm begünstigt wird. Als Vektor wurde pMKE2 (8237 bp) verwendet und das gleiche Insert wie zuvor bei der Expression in *E. coli* kloniert. pMKE2 wurde von Moreno *et al.* (2000) speziell als Expressionsvektor für den Stamm *T. thermophilus* HB27::nar entwickelt. Das Plasmid trägt eine Kanamycin-Kassette für die Selektion. Das Protein wird erst produziert, wenn Nitrat und eine sauerstoffarme Umgebung vorhanden sind (Kolbenstopfen und kein Schütteln der Erlenmeyergefäße). Die Proteinmuster nach SDS-PAGE (Abbildung 19, B) zeigten keine Überproduktion von Est7. Dies weist darauf hin, dass die Induktion in *T. thermophilus* erfolglos war. Auch die Aktivitäten waren an der Nachweisgrenze von 0,03 - 0,09 U/mg und unterstreichen die marginale Produktion von Est7 in Stamm HB27::nar.

Expressionswirt *E. coli* BL21

**A**

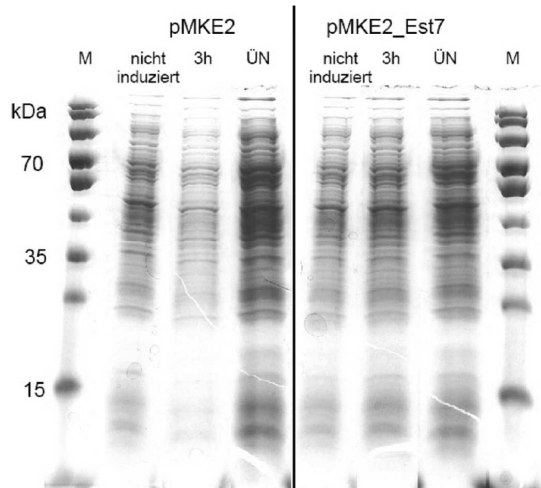


pET21a\_Est7 (BL21)

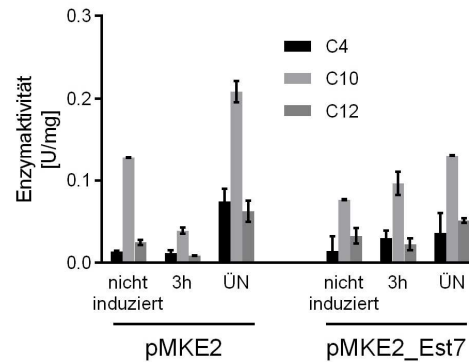


Expressionswirt *T. thermophilus* HB27::nar

**B**

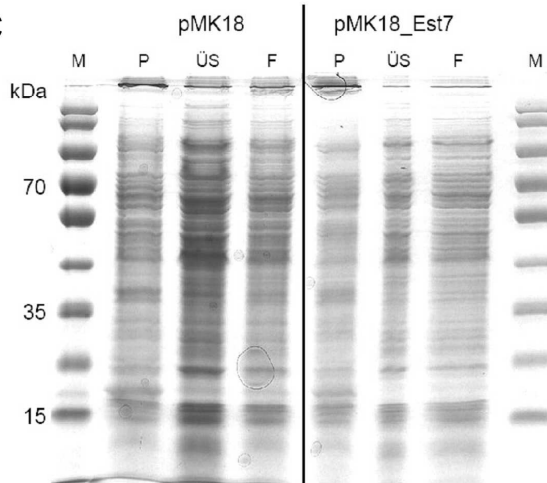


pMKE2\_Est7 (HB27::nar)



Expressionswirt *T. thermophilus* BL03

**C**



pMK18\_Est7 (BL03)

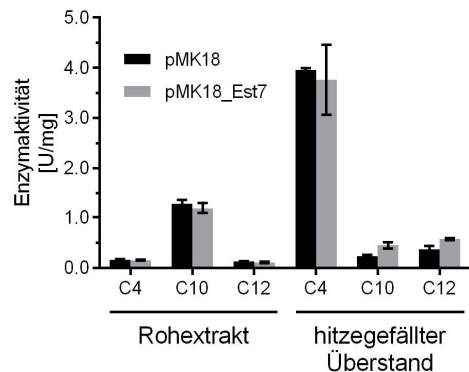
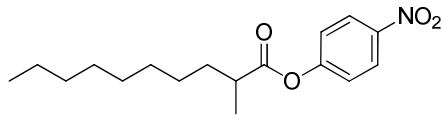


Abbildung 19: Übersicht über die Expression von Est7 aus den drei Plasmiden pET21a\_Est7 (ermöglicht induzierbare Expression in *E. coli* BL21), pMKE2\_Est7 (ermöglicht induzierbare Expression in *T. thermophilus* HB27::nar) und pMK18\_Est7 (Plasmid ohne Induktionsmöglichkeit in *T. thermophilus*) in verschiedenen Wirten mit den Ergebnissen aus den pNP-Aktivitätsassays. M, Marker; RE, Rohextrakt; P, Pellet; ÜS, Überstand; F, Flowthrough; ÜN, über Nacht, W, Wasch-Eluat; E, Eluat.

Schließlich wurde die Expression von *est7* in einem anderen *Thermus*-Vektor (pMK18: 4910 bp) in BL03 getestet. Stamm *T. thermophilus* BL03 mit pMK18\_Est7 wurde von Dr. Leis im Rahmen seiner Dissertation erstellt. Bei dem Wirtstamm handelt es sich um eine Vierfachmutante ( $\Delta$ TT\_P0042  $\Delta$ TT\_C1787  $\Delta$ TT\_C0340 - 0341  $\Delta$ TT\_C0904) mit einer verringerter Esteraseaktivität, dadurch eignet sich BL03 gut für Esterase-Screenings in *T. thermophilus* (Leis, 2015a). Zur Selektion wurde wieder Kanamycin als Marker verwendet, der gleichzeitig die konstitutive Expression steuern soll. Der Versuch, Est7 in *T. thermophilus* zu produzieren, war jedoch erfolglos. Auf den Gelbildern (Abbildung 19, C) war keine Proteinbande ersichtlich und Aktivitätsmessungen mit Rohextrakt und der aufgereinigten Fraktion lieferten keine eindeutigen Ergebnisse.

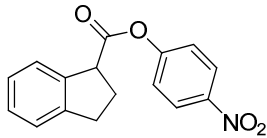
In Zusammenarbeit mit der Arbeitsgruppe von Prof. Dr. Pietruszka (Bioanorganische Chemie, Forschungszentrum Jülich) wurde die Fähigkeit von Est7 zur Spaltung spezieller, in Zusammenhang mit möglichen Anwendungen des Enzyms interessanten, Esterase-/Lipasesubstraten überprüft. In Abbildung 20 sind die getesteten Substrate die vom Kooperationspartner zur Verfügung gestellt wurden, aufgelistet und mit Strukturformeln dargestellt. Der Assay wurde wie bei Punkt 1.4.4 beschrieben durchgeführt, dabei wurde der Farbumschlag durch freigesetztes *para*-Nitrophenol gemessen. Als Protein wurde rekombinant hergestelltes Est7 aus *E. coli* BL21 mit pET21a-Est7 verwendet. Die Aktivitätsmessung mit diesen Lipasesubstraten ergab eine spezifische Aktivität von  $3,33 \times 10^{-2}$  und 0,536 mU/mg auf Ibuprofen-*p*NP-Ester bei 60 °C bzw. bei 80 °C.

In Abschnitt 1.3 wurden Beispiele von thermostabilen Enzymen genannt, welche durch unsere aktuell entwickelten Methoden bei der Durchforstung metagenomischer Genbibliotheken entdeckt wurden. Um das enorme Potenzial an neuen nützlichen Biokatalysatoren aus oft weitgehend unerforschten Habitaten unserer Umwelt noch weiter ausschöpfen zu können, müssen die Methoden für die Durchforstung der Metagenome solcher Umweltproben verbessert und die Auswahl an Expressionswirten erweitert werden.



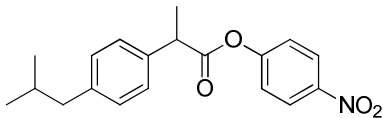
Molecular Weight: 307,38

2-Methyldecansäure-pNP-Ester



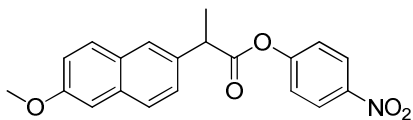
Molecular Weight: 283,28

Indancarbonsäure-pNP-Ester



Molecular Weight: 327,37

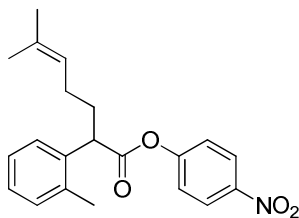
Ibuprofen-pNP-Ester



Molecular Weight: 351,35

"Naproxen-pNP-Ester"

der Handelsname Naproxen bezeichnet eigentlich nur das (S)-Enantiomer



Molecular Weight: 353,41

"Mulzer"-pNP-Ester  
Penifulvin A Vorstufe

Abbildung 20: Lipasesubstrate, die vom Kooperationspartner Prof. Dr. J. Pietruszka (Forschungszentrum Jülich) zur Verfügung gestellt wurden.

## 4.4 Durchmusterung metagenomischer Genbibliothek

In dieser Arbeit wurde mit verschiedenen metagenomischen Genbibliotheken gearbeitet, einige von ihnen wurden von Kooperationspartnern bereit gestellt, z. B. Genbibliotheken mit den Bezeichnungen „hyperthermophile Archaeen (HA)“, „mesophile Bakterien (MB)“ und „thermophile Bakterien (TB)“ stammten von Prof. Golyshin (Bangor University, UK). Andere Genbibliotheken wurden in unserer Arbeitsgruppe am Lehrstuhl für Mikrobiologie der Technischen Universität München erstellt. Die „Azoren“-Metagenom-Genbibliothek wurde beispielsweise von Herrn Dr. Leis aufgebaut und in dieser Arbeit wurde metagenomische DNA von Umweltproben aus Mexiko („Mexico“) und einer Tierfutterproduktionsanlage aus Russland („Russia“) für die Herstellung von pCC1FOS-basierten Metagenom-Genbanken verwendet.

### 4.4.1 Aufbau von pCC1FOS-basierten Metagenom-Genbibliotheken aus „Russia“- und „Mexico“- Umweltproben

Die Umweltprobe mit dem Namen „Russia“ wurde aus einer Trocknungsanlage für Tierfutterherstellung aus Stupino, einem Vorort von Moskau, entnommen (+54° 56'46,17'' geographischer Breite, +38° 2'56,51'' geographischer Länge). Es handelt sich hierbei um Material aus einem Luftreinigungssystem direkt vor dem Biofilter, der mit feuchter 50 - 60 °C heißer Luft durchströmt wurde. Der pH-Wert lag im neutralen Milieu bei 7,73. Die Umweltprobe aus Mexiko, die die gleichnamige Bezeichnung trägt, bestand aus schwefelhaltigen Sedimenten aus einer heißen Quelle in Chignahuapan, Puebla, Mexiko (19° 50' 15'' geographischer Breite und -97° 59'37'' geographischer Länge). Bei der Probenentnahme wurde eine Temperatur von 50 °C und der pH-Wert 6,7 protokolliert (Lachmann, 2013).

Metagenomische DNA wurde aus den Proben „Russia“ und „Mexico“ mittels des PowerSoil® DNA Isolation Kit (MoBio Laboratories, USA) isoliert. Nach gelelektrophoretischer Auftrennung (Abbildung 21) waren Banden mit hochmolekularer aufgereinigter DNA zu erkennen. Aus der Umweltprobe „Russia“ konnte deutlich mehr DNA gewonnen werden, mit einer Konzentration von 107,7 ng/µL als die „Mexico“-Probe mit 31,7 ng/µL (Gesamtmenge an DNA: 2,37 µg und 1,3 µg). Aufgrund der konzentrierteren DNA-Menge wurde bei der Herstellung der „Russia“-Fosmidbibliothek eine höhere Transduktionseffizienz von  $5,26 \times 10^4$  cfu/mg DNA und eine größere Metagenombank von 6645 Fosmiden erreicht (Material und Methoden, Abschnitt 1.2.3.2). Im Vergleich dazu beinhaltet die Bank „Mexico“ nur  $1,67 \times 10^4$  cfu/mg und 3322 Fosmide. Zur Validierung der erstellten Metagenom-Genbanken wurde über Restriktionsverdau sichergestellt, dass die Fosmid-Inserts über eine ausreichend hohe Diversität verfügen. Nach dem Verdau eines pCC1FOS-Fosmids mit EcoRI oder BamHI

müssen im Bandenmuster die für die jeweiligen Enzyme charakteristischen Vektor-Fragmente (2,3 kb für EcoRI; 9,3 kb für BamHI) wieder zu finden sein (Bild nicht gezeigt).

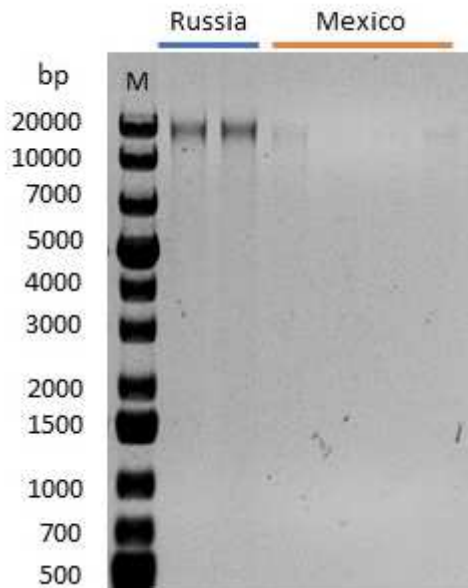


Abbildung 21: Gelelektrophorese der metagenomischen DNA aus den Umweltproben „Mexico“ und „Russia“

Um eine hohe Diversität der in den Fosmidklonen klonierten metagenomischen DNA zu erhalten, sollte bereits in den Umweltproben eine hohe Bakterienvielfalt gegeben sein. Bereits unter dem Mikroskop war eine größere morphologische Vielfalt der Bakterien in der „Russia“-Probe als in der „Mexico“-Probe zu erkennen. Für die Identifizierung der Bakterien in den Umweltproben wurde ein Bereich der 16S rRNA mittels einer PCR mit den universellen, bakteriellen 16S rDNA Primern „100K“ (5'-GGTTACCTTGTTACGTA CT-3') und „616 v\_alt“ (5'-AGAGTTT GAT(CT)(AC)TGGCTCAG-3') amplifiziert. Die PCR wurde zum einen direkt mit den Umweltproben durchgeführt und zum anderen mit den Isolaten, die auf Minimalmedien-Platten mit CMC oder Avicel als Kohlenstoffquelle (SH-CMC bzw. SH-Avicel) gewachsen waren. Somit ergaben sich insgesamt sechs Proben:

- Umweltprobe „Russia“
- Umweltprobe „Mexico“
- Isolat „Russia“ SH-CMC
- Isolat „Russia“ SH-Avicel
- Isolat „Mexico“ SH-CMC
- Isolat „Mexico“ SH-Avicel



Tabelle 31: Zusammenfassung der über 16sRNA identifizierten Phyla aus den Umweltproben „Mexico“ und „Russia“.

|                                       | Umweltprobe „Russia“ | Umweltprobe „Mexico“ |
|---------------------------------------|----------------------|----------------------|
| <i>Actinobacteria</i>                 | -                    | 1                    |
| <i>Bacteroidetes</i>                  | 3                    | 1                    |
| <i>Chloroflexi</i>                    | 2                    | -                    |
| <i>Gammaproteobacteria</i>            | 1                    | 1                    |
| <i>Firmicutes</i>                     | 7                    | 1                    |
| <i>Nitrospirae</i>                    | -                    | 2                    |
| <i>Spirochaetes</i>                   | 1                    | -                    |
| <i>Proteobacteria</i>                 | 7                    | 4                    |
| verschiedene Spezies aus<br>23 Proben | 21                   | 10                   |

Aus den Umweltproben wurden nur relativ wenige 16s rDNA Amplikons sequenziert, denn es sollte damit nur ein grober Einblick zur Vielfalt in den Proben genommen werden. Aus der Umweltprobe „Russia“ wurden aus 23 durchgeführten 16s rDNA-Sequenzreaktionen 21 unterschiedliche Bakterienspezies identifiziert, während die „Mexico“-Probe eine geringere Diversität von nur 10 verschiedenen Arten aus 23 Proben besaß (Tabelle 31). Die Identifikation der Isolate auf den CMC- und Avicel-Platten ergaben folgende Spezies mit jeweils 99 % maximaler Identität (Tabelle 32). Die Isolate wurden daraufhin auf Substratplatten mit Tributyrin, Triolein, Stärke, Milch, CMC, Xylan oder Avicel ausgestrichen und bei 50 °C (Mexico) und 60 °C (Russia) für 24 Stunden inkubiert. Wenn die Isolate auf den Platten gewachsen sind oder sich die Trübung beispielweise bei den Milchplatten geklärt hat wurde eine Aktivität auf das jeweilige Substrat vermutet.

Tabelle 32: Aktivität Analyse der Anreicherungskulturen aus den Umweltproben „Mexico“ und Russia“

| Isolat             | Spezies                                    | Aktivität                                     |
|--------------------|--|---|
| „Russia“ SH-CMC    | <i>Anoxybacillus flavithermus</i> DSM 2641 | Tributyryn, Stärke, Milch                     |
| „Russia“ SH-Avicel | <i>Bacillus smithii</i> NRS-173            | keine   |
| „Mexico“ SH-CMC    | <i>Brevibacillus borstelensis</i> DSM 6347 | Milch (sehr stark)                            |
| „Mexico“ SH-Avicel | <i>Bacillus sonorensis</i> NRRL B-23154    | Tributyryn, Stärke, Milch, CMC, Avicel, Xylan |

#### 4.4.2 Übertragung verschiedener Metagenom-Genbibliotheken von *E. coli* zu *T. thermophilus*

Das CopyControl Fosmid pCC1FOS ermöglicht die Integration von metagenomischer DNA bis zu einer Größe von etwa 40 kb. Bei dieser Insertgröße steigt die Wahrscheinlichkeit, dass eventuell ganze Gen-Cluster, die für komplexe Stoffwechselwege oder Operone kodieren, in einem Fosmid entdeckt werden können (Daniel et al., 2005). Das Plasmid pCC1FOS besitzt zwei Replikationsursprünge, zum einen den *ori* des F-Faktors (single-copy), der für die Insert-Stabilität und erfolgreiche Klonierung instabiler DNA-Sequenzen notwendig ist, und zum anderen einen high-copy *ori*, der nach Induktion mit Auto-Induktionslösung (AIS) eine Steigerung der Fosmid-Kopienzahl auf bis zu 50 Kopien pro Zelle erlaubt. Die hohe Kopienzahl der Fosmide steigert ggf. die Enzymaktivität durch klonierte Enzymgene und führt dadurch zu einer sensitiveren und damit effizienteren Nachweisbarkeit. Zudem besitzt pCC1FOS für eine positive Selektion von Fosmid-tragenden *E. coli* Zellen ein Chloramphenicol-Resistenzgen. Das pCC1FOS Fosmidssystem hat jedoch seine Limitierungen, da es sich auf den mesophilen Expressionswirt *E. coli* beschränkt. Dabei könnten in alternativen Wirtsbakterien wie z. B. *T. thermophilus* neue hitzestabile Enzyme entdeckt werden, die im klassischen Wirt *E. coli* nicht gebildet bzw. identifiziert würden. Für eine Evaluierung von *T. thermophilus* als neuen Expressionswirt für die funktionale Durchmusterung von Metagenom-Bibliotheken muss die Effizienz und Sensitivität der Durchmusterung derselben Genbibliotheken in *E. coli* und *T. thermophilus* von beiden Stämmen gegenüber gestellt werden. Um eine Mobilisierung einer pCC1FOS Metagenombank von *E. coli* zu *T. thermophilus* zu ermöglichen, bedarf es einer Modifikation des Fosmids.

Dies geschieht durch die in unserer Arbeitsgruppe entwickelte Methode der *in vivo* Modifikation rekombinanter Fosmide über das *Cre/loxP* System in *E. coli* (Abbildung 22). Zunächst werden dafür die beiden Plasmide pRetro-T2 und pJM2545 in eine Fosmid-tragende *E. coli* Zelle transformiert. Nach IPTG-Induktion wird das auf pJM2545 liegende Gen für *Cre*-Rekombinase exprimiert, welche pCC1FOS und pRetro-T2 an den jeweiligen *loxP*-Seiten schneidet und mittels Sequenz-spezifischer Rekombination eine Fusion beider Plasmide unter Bildung von pCC1FOS-retro (kurz für Retro Fitting) möglich macht. Die Rückreaktion wird durch den temperatursensitiven Replikationsursprung auf pJM2545 verhindert, wenn man die Transformanden bei 42 °C inkubiert. Bei dieser hohen Temperatur kann das Plasmid nicht mehr replizieren und infolgedessen wird keine weitere *Cre*-Rekombinase synthetisiert (Nagy et al., 2000). Über das zweite Plasmid pRetro-T2 werden ein zweiter Selektionsmarker (Kanamycin), der für die Selektion in *T. thermophilus* benötigt wird, sowie eine Integrationssequenz, welche für die Eingliederung der Metagenom-DNA in das *T. thermophilus* Genom erforderlich sind, in das Fusionsplasmid eingebaut (Angelov et al., 2009).

Eine weitere Möglichkeit für eine parallel durchgeführte Durchforstung in zwei verschiedenen Expressionsstämmen ist die Erstellung einer Metagenombank mit dem Rückgrat eines *E. coli*/*T. thermophilus* Shuttle Fosmids, die sich direkt (ohne Modifikation) in beide Stämme transformieren lässt. Dafür wurde von Angelov et al., (2009) das Fosmid pCT3FK konstruiert, welches für derartige komparative Metagenom-Studien eingesetzt werden kann.

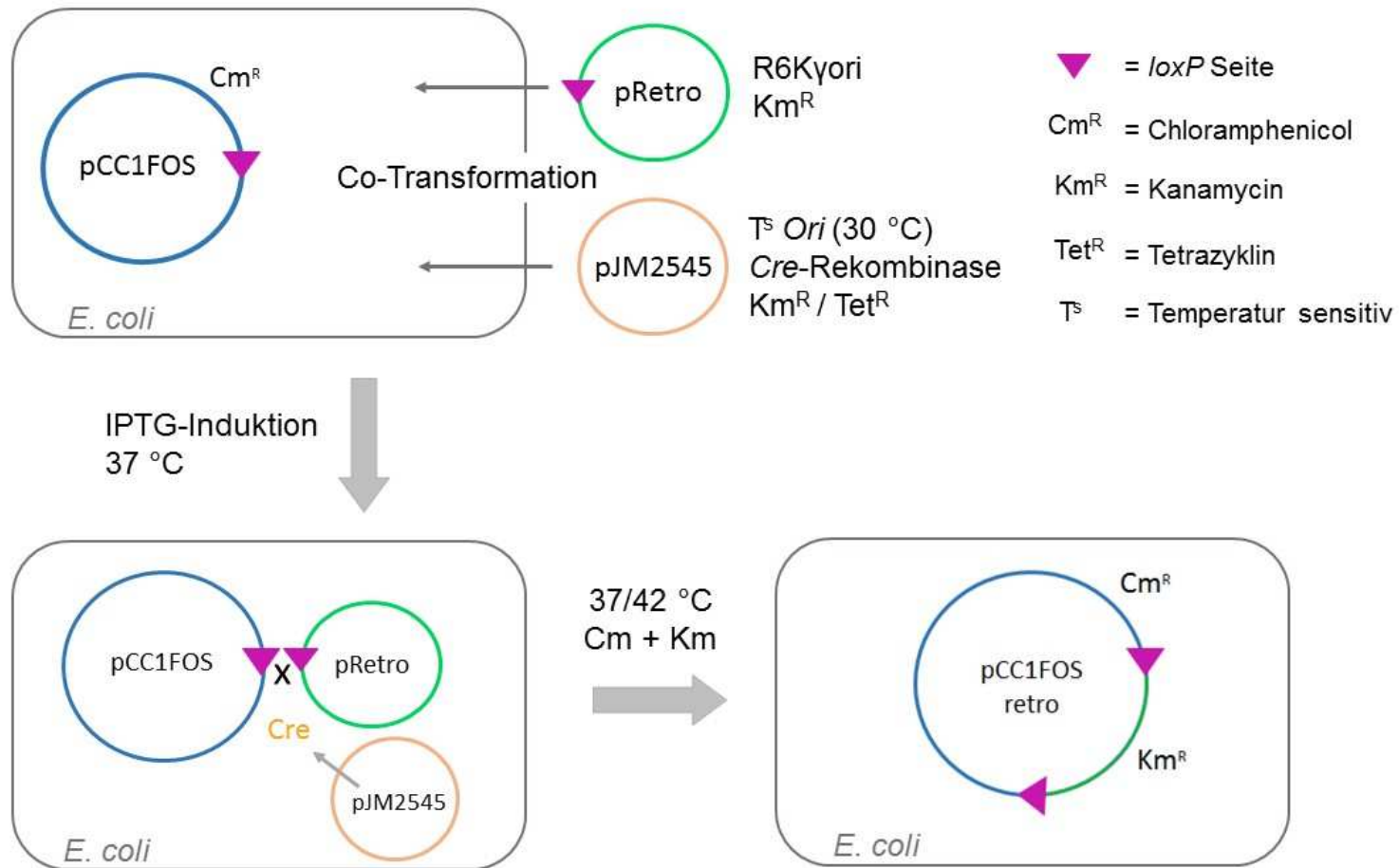


Abbildung 22: Schema zur *in vivo* Fosmid-Modifizierung in kompetente pCC1FOS tragende *E. coli* Zellen werden die Plasmide pRetro (loxP site, R6K ori und Km-Resistenz) und pJM2545 (30 °C sensitiver ori, Cre-Rekombinase, Kanamycin- und Tetrazyklin-Resistenz) co-transformiert. Nach einer IPTG Induktion und Inkubation bei 37 °C wird die Cre-Rekombinase exprimiert und verhilft den beiden Vektoren pCC1FOS und pRetro zur Rekombination zu pCC1FOS-retro. Um eine Rückreaktion zu verhindern werden die modifizierten Zellen bei 42 °C mit Chloramphenicol und Kanamycin kultiviert. Der R6K ori stellt sicher, dass die Vektor in *E. coli* Zellen, die das *pir* Genprodukt exprimieren bleibt und repliziert wird.

#### 4.4.3 Optimierung des Durchmusterungsprozesses und Steigerung der Transformationseffizienz beim komparativen Screening in *E. coli* und *T. thermophilus*

Für den Transformationserfolg in *T. thermophilus* sind Qualität und Quantität der Fosmid-DNA entscheidend. Die höchste Transformationseffizienz wird bei Zugabe von 1 µg Fosmid-DNA erreicht, die frei von Proteinen, RNA und anderen Zellbestandteilen ist. Zudem muss sich der Wirt in der exponentiellen Wachstumsphase befinden, um das Maximum an Nukleinsäuren aufnehmen zu können (Material und Methoden 1.2.2.4). Ein Hauptaugenmerk dieser Arbeit lag auf der Steigerung des Durchmusterungsdurchsatzes mithilfe zusätzlicher Geräte der Verbesserung der Sensibilität der Aktivitätsnachweise. Die metagenomischen Genbibliotheken „Russia“, „Mexico“, „Golyshin“ und die „Azoren (Wasser)“ wurden dabei in den beiden Wirtsorganismen *E. coli* und *T. thermophilus* funktionell durchforstet (Tabelle 34).

##### 4.4.3.1 Filterplatten und Vakuum Manifold

Die Fosmid-Isolation im 96 Deep-well Format kann auf zwei verschiedene Arten erfolgen. Zum einen durch eine alkalische Lyse mit anschließender Isopropanol-Fällung und zum anderen durch Verwendung spezieller Filter und Säulen, die DNA an ihre Silikamembran binden und von anderen Zellbestandteilen trennen. Für die Aufreinigungsschritte mit den Filterplatten wurde der Vakuum Manifold der Firma Pall verwendet (siehe Material und Methoden 2.2.4.2), ansonsten wurden die Überstände über Zentrifugation in Sorvall RC3B Plus (Kendro Laboratory Products) gewonnen. Es wurden vier verschiedene Protokolle mit Kombinationen durchgeführt, dabei dieselbe 96 Mikrotiter-Platte mit Fosmid-tragenden *E. coli* Zellen aufgereinigt und *Thermus*-Zellen mit der erhaltenden DNA transformiert. Die Protokollversion 1 wird im Material- und Methodenteil unter dem Punkt 3.4 „Fosmid-DNA-Präparation im 96-Well Format“ erläutert. Bei allen Experimenten erfolgte die Bestimmung der DNA-Menge und -Reinheit durch Restriktionsverdauansätze von jeweils fünf Fosmiden und anschließender analytischer Gelelektrophorese.

Tabelle 33: Zusammenfassung der vier Protokollvariationen zur Fosmidreinigung im 96-well Format und ihre Transformationseffizienz in *T. thermophilus*.

| Nr. | „Lysat“-Aufreinigungsschritt    | „DNA“-Aufreinigungsschritt | Transformation Effizienz |
|-----|---------------------------------|----------------------------|--------------------------|
| 1   | Zentrifugation                  | Isopropanol Fällung        | ~ 29 %                   |
| 2   | „Lysate Clarification“ - Platte | Isopropanol Fällung        | ~ 88 %                   |
| 3   | Zentrifugation                  | „DNA Binding“-Platte       | ~ 9~%                    |
| 4   | „Lysate Clarification“ - Platte | „DNA Binding“-Platte       | ~ 71 %                   |

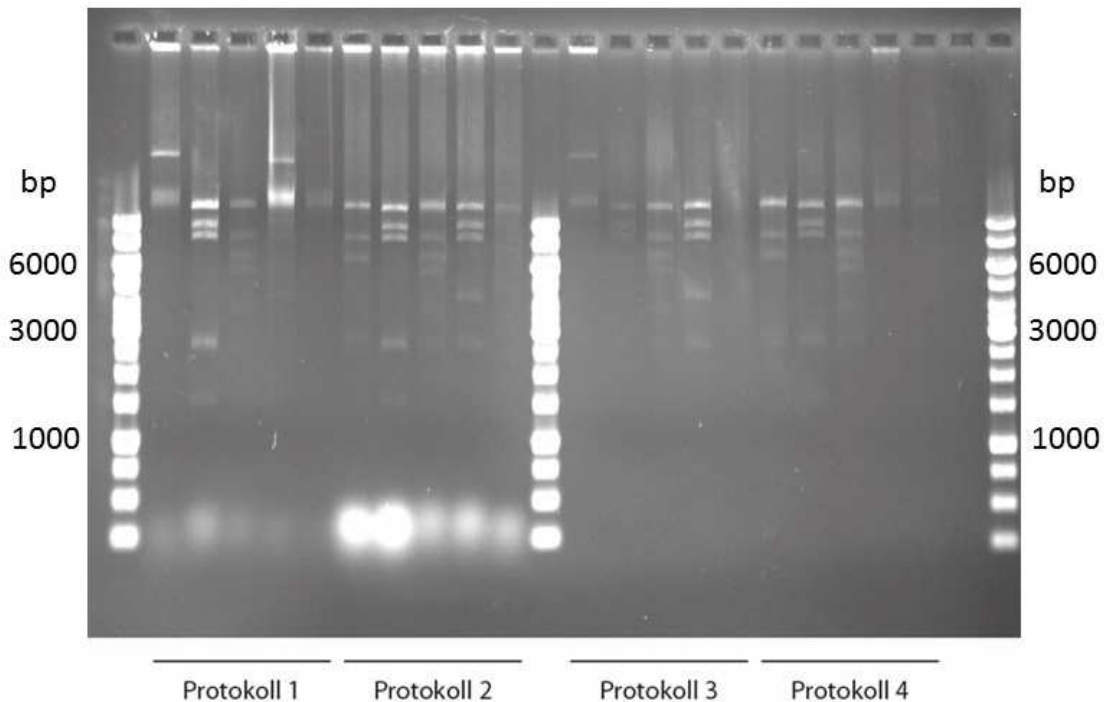


Abbildung 23: Gelbild von den vier verschiedenen Extraktionsprotokollen. 1) Fosmid-Isolation durch alkalische Lyse, 2) Verwendung einer „Lysate Clarification“-Platte, 3) Verwendung einer „DNA-binding“ Platte und 4) Verwendung von „Lysate Clarification“- und „DNA-binding“ Platte. Die Fosmide wurden mit *EcoRI* geschnitten.

Auf dem Gelbild sieht man größere DNA-Mengen bei Protokoll 1 und 2, aber auch signifikante Mengen an RNA, die bei 250 bp erkennbar ist (Abbildung 23). Bei den anderen beiden Protokollen waren die DNA-Konzentrationen geringfügig niedriger, jedoch enthielten diese Proben keine RNA-Kontamination. Für die Bestimmung der Transformationseffizienz im 96 Well-Format wurden *Thermus*-Kulturen in der exponentiellen Wachstumsphase zu der

aufgereinigten Fosmid-DNA hinzugegeben, zwei Stunden bei 70 °C inkubiert und anschließend für 48 Stunden bei 60 °C in einem TB-Kan20 Medium kultiviert. Tabelle 33 zeigt den Prozentsatz der angewachsenen Kulturen in der 96 Deep-well Platte. Aus diesen Ergebnissen lässt sich schließen, dass die Proben bei der Verwendung einer „DNA-Binding“-Platte bezüglich der RNA-Verunreinigungen besser aufgereinigt wurden. Die Ausbeute wurde durch diese Platte allerdings verringert. Die Transformationseffizienz als entscheidender Punkt war bei den Protokollen 2 und 4 mit Prozentsätzen von 88 bzw. 71 % am höchsten. In beiden Fällen wurde eine „Lysate Clarification“-Platte verwendet, um die DNA aus dem Zellysat zu gewinnen. Dieser Aufreinigungsschritt steigerte den Transformationserfolg enorm, während die RNA-Verunreinigung der DNA eher eine marginale Rolle spielte.

#### 4.4.3.2 VIAFLO 96

Die Aufreinigung von 192 Fosmiden konnte durch alkalische Lyse im 96-Well Format mithilfe einer 8-Kanalpipette an einem Tag durchgeführt werden (Material und Methoden Punkte 1.2.3.4 und 1.2.3.5). Eine enorme Arbeitserleichterung stellte bei der Extraktion die Verwendung der 96-Kanalpipette VIAFLO 96 der Firma INTEGRA Biosciences GmbH (Fernwald, Deutschland) dar. Damit konnte die Anzahl der DNA-Extraktionen auf 576 Präparationen pro Tag gesteigert werden. Die Reinheit der isolierten Fosmide wurde weder gesteigert noch gemindert, da die Bedingungen der kritischen Schritte wie die Lysataufreinigung und DNA-Fällung, durch den Pipettier-Roboter nicht verbessert wurden. Jedoch führt der Einsatz der VIAFLO 96 zu einer enormen Zeitersparnis, weil dadurch der Durchsatz pro Tag verdreifacht werden konnte.

#### 4.4.3.3 Validierung des Pool-Screening

Das funktions-basierte Screening von Fosmidklonen ist mit großem Zeit- und Arbeitsaufwand verbunden. Jeder einzelne Klon wird zunächst in *E. coli* untersucht und dann in *T. thermophilus* übertragen, worauf eine erneute Durchforstung auf Substratplatten erfolgte. Diese Schritte beim komparativen Screening können nicht umgangen oder vereinfacht werden, aber es besteht die Möglichkeit diese Prozeduren entweder einerseits an Einzelklonen in einem 96 Deep-well Blöcken durchzuführen oder andererseits 48 Klone in einem Pool zusammenzufassen, welcher dann wie eine Probe behandelt wird. Diese als „Pool-Screening“ bezeichnete Methode wurde für pCC1FOS-Genbibliotheken entwickelt und beinhaltet folgende Schritte: a) Erstellen des Fosmid-Pools b) *in vivo* Modifizierung im Pool-Format c) Fosmid-DNA-Extraktion d) Transformation in *T. thermophilus*.

Zunächst wurde eine pCC1FOS-Genbibliothek aus Kompost-Umweltproben (Ilmberger et al., 2012) verwendet, in der sich ein bereits validierter auf Xylan aktiver Klon befindetet. Der aktive

Klon befand sich auf der Position B8 in dem M12-22B-Pool, der insgesamt 48 Fosmidklone umfasst. Verschiedene Kultivierungsmethoden, wie die Zugabe von Autoinduction Solution zu unterschiedlichen Zeitpunkten, wurden untersucht, um ihren Einfluss auf die Diversität der Fosmide und die Effizienz der Modifizierung festzustellen (Abbildung 24).

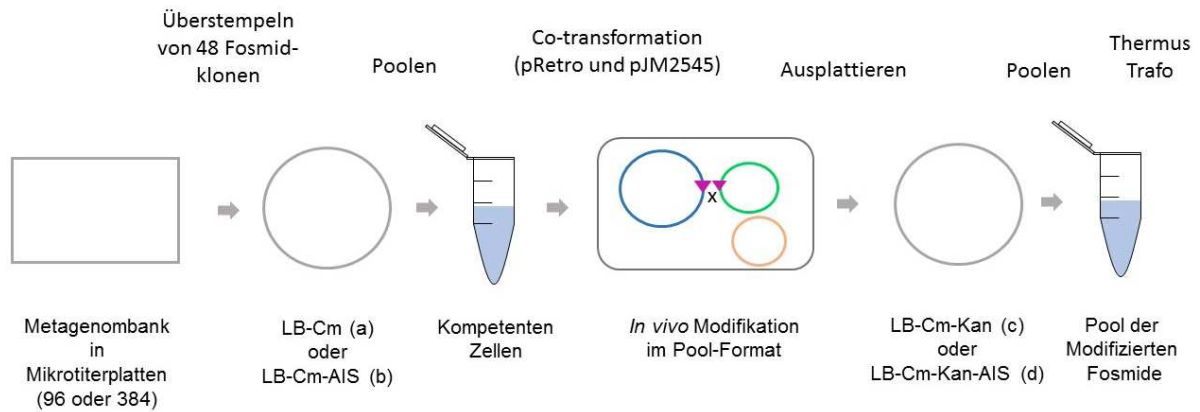


Abbildung 24: Durchführung des Pool-Screenings. Die Metagenombank im Mikrotiter-Format wurde mit einem Stempel mit 48 Pins entweder auf Platten nur mit Chloramphenicol (Cm) übertragen oder mit Cm und AIS (Autoinduction Solution). Die Zugabe von AIS erhöht die Kopienzahl des pCC1FOS Vektors in der Zelle. Nachdem die 48 Klone auf den Platten gewaschen sind, wurden sie in ein Reaktionsgefäß zusammengeführt (poolen) und chemisch kompetent gemacht. Daraufhin erfolgte eine in vivo Modifikation des gesamten Pools durch die beiden Plasmide pRetro und pJM2545. Die modifizierten Zellen wurden anschließend auf Platten mit beiden Antibiotika (Chloramphenicol und Kanamycin) und mit oder ohne AIS ausplattiert. Die Klone wurden wieder zusammengeführt, die Fosmide aufgereinigt und in *T. thermophilus* transformiert.

Für die Modifizierung des Pools wurden 5 µg der jeweiligen Plasmide pRetro-T2 und pJM2545 eingesetzt, während bei der Reaktion mit dem Einzelklon B8 nur ein Zehntel dieser Menge transformiert wurde. Bei anfänglicher Kultivierung mit AIS wuchsen mit einer Gesamtzahl von 848 deutlich mehr Klone auf der LB-Cm-Kan-AIS-Platte im Vergleich zu dem Fall, wo vorher keine AIS-Kultivierung stattfand (21 Klone). Zur Überprüfung der Diversität und der Modifizierungseffizienz wurden 5 - 6 Klone der jeweiligen Variationen mit einem EcoRI Restriktionsverdau der enthaltenden Plasmide kontrolliert. Eine spezifische Bande mit einer Größe von 2,3 kb, die vom Vektor herrührt, ist deutlich zu erkennen. Zwei von fünf Klonen ohne AIS und zwei aus sechs Klonen mit AIS waren identisch. Zur Überprüfung, ob die in vivo Modifizierung bzw. das Zusammenführen der Klone eine Auswirkung auf die Aktivität der Klone hat, wurden die Proben B8-M12-22 (mit AIS), Pool-M12-22B (ohne AIS) und Pool-M12-22B (mit AIS) im Anschluss an das Retrofitting auf Substratplatten mit Xylan und CMC ausplattiert.



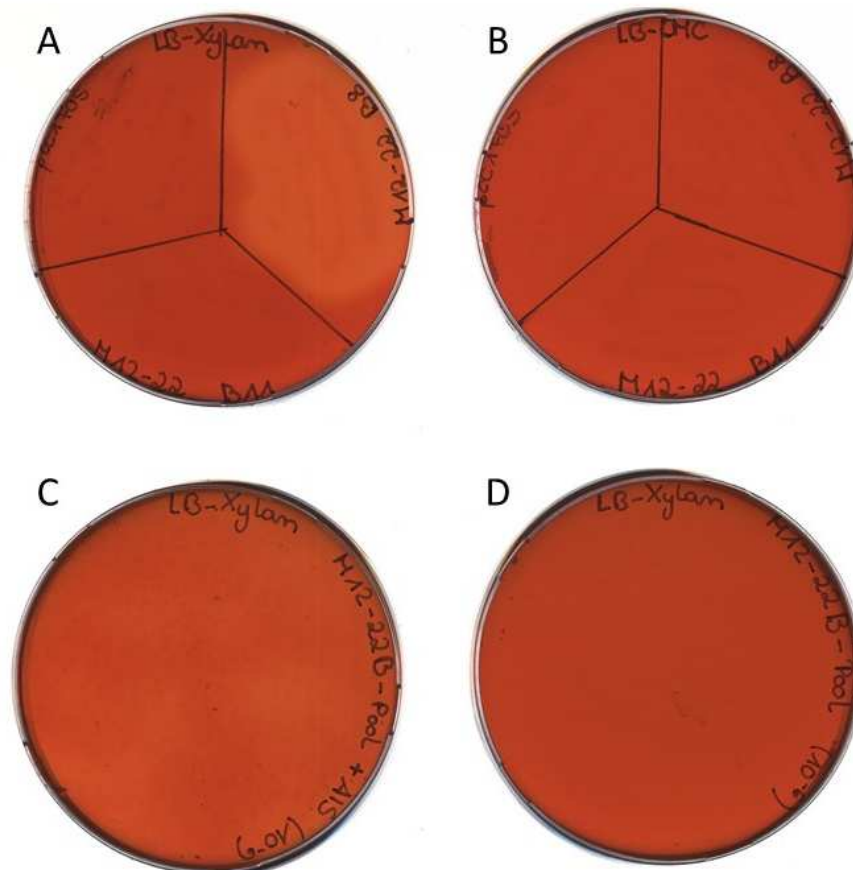


Abbildung 25: Substratplatten LB-Xylan (A) oder LB-CMC (B) mit dem B8 Klon (M12-22) und dem Pool M12-22 (48 Fosmide). Der Klon B8 weist eine Xylanaseaktivität durch einen hellen Hof nach der Kongorotfärbung auf. Der Pool, der zuvor mit AIS inkubiert wurde, zeigt helle Flecken auf der LB-Xylan AIS Platte (C). Die LB-Xylan Platte mit dem M12-22B Pool ohne AIS zeigt keine Aktivität (D).

Auf den Platten konnte man erkennen, dass die Xylanase B8 (aus dem Pool M12-22) nach der *in vivo* Fosmid-Modifizierung weiterhin aktiv war und diese somit keinen negativen Einfluss auf die Enzymaktivität hatte. Dies war an dem deutlichen Halo von B8 auf der LB-Xylan-Platte zu erkennen (Abbildung 25, A). Als Negativkontrolle diente eine LB-Platte mit CMC (B) auf dem kein Substratabbau nachgewiesen werden konnte. Darüber hinaus konnte gezeigt werden, dass auch innerhalb eines Pools die Aktivität weiterhin deutlich nachweisbar war und nicht im Gemisch aus 48 unterschiedlichen Klonen verloren ging (helle Bereiche auf der Platte, C). Bei der anfänglichen Kultivierung auf LB-Agarplatten musste AIS im Medium enthalten sein, da sonst der Abbau des Substrates zu schwach für einen Nachweis war. Abbildung 25 zeigt deutlich, dass beim Pool M12-22B, welcher nicht mit AIS kultiviert wurde, keine Aktivität durch Halos auszumachen ist (D). Die Transformationszahlen des Pools M12-22B (29 Klone) und die des Einzelklons B11-M12-22 (33 Klone), welche anfänglich mit AIS kultiviert wurde, sind nahezu gleich. In dem Pool ohne AIS konnten gar keine Transformanden nachgewiesen werden.

#### 4.4.3.4 Durchmusterung verschiedener Bibliotheken

In dieser Dissertation wurden für die Entwicklung und Optimierung der Durchforstungsmethoden wie z.B. Einsatz von Multipipetten (96 ViaFlo), Manifold Präparation mit Aufreinigungsplatten, Pool-Screening und komparatives Screening verschiedener Genbanken in *E. coli* und *T. thermophilus* untersucht. Aus Tabelle 34 kann abgelesen werden, wie viele Fosmide jene Bank insgesamt beinhaltet und welche Anzahl in *E. coli* bzw. in *T. thermophilus* nach dem Transfer durchforstet wurden.

Beim Überstempeln der Kulturen aus den Mikrotiterplatten oder Deep-well Blöcken auf die Agar-Platten wurde darauf geachtet werden, dass genug Zellmaterial übertragen wurde und dass die Positionen der einzelnen Wells wieder zurückverfolgt werden konnten. Die Substratplatten wurden bis zu 2/3 der Petrischale mit Agarmedium gefüllt, da sie für mehrere Tage bei hohen Temperaturen bis zu 60 °C inkubiert werden mussten und der Agar sonst ausgetrocknet wäre. Gegen die Verdunstung des Wassers half noch zusätzlich das Einwickeln in eine Kunststofftüte. Nach Identifizierung eines aktiven Klons auf einer Platte, wurde das Fosmid isoliert und wieder in eine kompetente EPI300 Zelle übertragen und mit dem Leervektor verglichen. Dies diente der Evaluierung des Fosmids und bestätigte, dass die dazu gewonnene Aktivität tatsächlich vom metagenomischen Insert stammte. Oftmals wurde der zunächst positive Phänotyp von Klonen nach der Retransformation nicht bestätigt. Das kann verschiedene Gründe haben: (i) mehrere Enzyme, die nicht ausschließlich auf dem Fosmid lokalisiert sind, interagieren zusammen (ii) Klone stellten keine Reinkulturen dar (iii) Verwechslung der Klone.

Tabelle 34: Zusammenfassung der Fosmid-Banken und den durchforsteten Fosmide in *E. coli* und *T. thermophilus*

| <b>Bibliothek</b>   | <b>Gesamtzahl</b> | <b>Durchforstet in<br/><i>E. coli</i></b> | <b>Durchforstet in<br/><i>T. thermophilus</i></b> |
|---|-------------------|---|---|
| Mexico  | 3322              | 480                                       | 274   |
| Russia  | 6645              | 1440                                      | 761   |
| Azoren (Wasser)   | 8160              | 2880                                      | 1590  |
| Hyperthermophile<br>Archaeaen und<br>Thermophile<br>Bakterien | 3840              | 768                                       | 60  |

Bei der Durchforstung der Genbibliothek aus der Azoren Wasserprobe im Wirt *E. coli* (Lt Platte 9/IV/F12) wurde ein Klon mit CMCase-Aktivität gefunden, die Aktivität war allerdings bei 37 °C höher als bei 60 °C. Dieser Klon mit mesophiler CMCase wurde nicht weiter untersucht, da hier das Hauptaugenmerk auf der Entdeckung von neuen thermostabilen Enzymen lag.

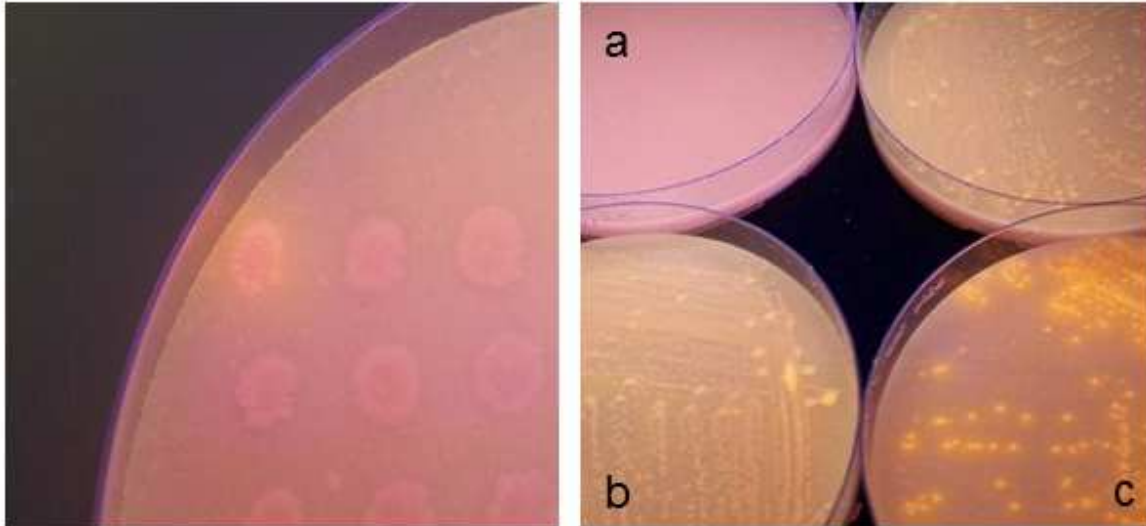


Abbildung 26: Auf Tributyrin aktiver Klon (L01) aus der HA- und TB-Fosmidbank (384 Mikrotiterplatten). (a) pCC1FOS Leervektor als Negativkontrolle (b) L01 bei 37 °C (c) L01 erst einen Tag bei 37 °C und dann einen weiteren Tag bei 60 °C.

Auch aus einer der 384 Mikrotiterplatte der HA- und TB-Metagenom-Genbibliothek wurde auf Position Lt/I/H1 ein auf Tributyrin aktiver Klon (L01) entdeckt (Abbildung 26). Nach der Retransformation behielt der Klon noch seine Enzymaktivität. Aus dem etwa 40 kb langen Metagenominsert wurde eine Shotgun-Genbibliothek erstellt, das Kandidatengen konnte jedoch nicht identifiziert werden.

Bei den Platten-Screenings wurden durch Halos um die Kolonien zwei Fosmidklone mit vermutlicher Avicel-Aktivität in *E. coli* entdeckt (A02 und A06). Daraufhin wurde ein DNSA-Assay mit den zwei Kandidaten und einer Kontrolle mit pCC1FOS mit 1 % Avicel und 1 % CMC durchgeführt. Die Negativkontrolle wies mit beiden Substraten eine höhere Aktivität auf als die beiden putativen Avicelasen. Es wurden zusätzlich noch fünf Klone in *T. thermophilus* gefunden, die auf Avicellaseaktivität hinwiesen (T\_rf08: E5, H1 und H2; T\_rf20: H11 und A7). Nach weiteren Untersuchungen wurden auch diese Klone vernachlässigt, da Avicelaktivitäten nicht eindeutig bestätigt werden konnte.

#### 4.4.1 *In vivo* Modifizierungs-Methoden, um Fosmid-Genbibliotheken in andere Expressionswirte (*B. subtilis* und *M. luteus*) übertragen zu können

In dieser Arbeit sollten die vorliegenden Metagenom-Genbanken in pCC1FOS Fosmiden aus *E. coli* in verschiedene Expressionswirte übertragen werden. Das Ziel dieses Teilprojekts war es bereits bestehende pCC1FOS Metagenom-Genbibliotheken in verschiedene Expressionswirte zu übertragen. Dies sollte mit der gleichen Methode wie bei Angelov *et al.* (2009) für *T. thermophilus* beschrieben, für andere ebenso natürlich kompetente Bakterien wie *B. subtilis* und *M. luteus* entwickelt werden.

Zur Erstellung des modifizierten pRetro-Plasmids für die beiden Stämme wurden zunächst die Kan-Kassette auf dem Plasmid pWLTK6 amplifiziert (911 bp; Primer KanF2/KanR3), während der Replikationsursprung R6K $\gamma$ ori und die darauffolgende 425 bp lange *loxP*-Sequenz von pRetro-T2 stammte (Primer: OriLoxF/ OriLoxR). Diese beiden Bestandteile sind gleich, jedoch wurden für die unterschiedlichen Expressionsstämme andere Sequenzen für homologe Rekombination ins Wirtschromosom eingefügt. Für pRetro-B1 (B für *Bacillus*) wurde ein 1980 bp langes Fragment des Amylasegens *amyE* aus dem Stamm 168 mit den Primern amyE F und amyE R mittels PCR vervielfältigt und mit dem Selektionsmarker *kan* und dem Ori-*loxP*-Amplifikat über Gibson Assembly zusammengefügt. Für das für *Micrococcus* spezifische Plasmid pRetro-M1 wurde analog aus dem *M. luteus* ATCC27141 *trpE* Gen ein PCR-Produkt gewonnen und mit *kan* und *ori-loxP* assembliert. Mithilfe der gewählten homologen Gensequenzen können sich die Metagenomsequenzen nach der *in vivo* Fosmidmodifikation und Transformation von *B. subtilis* bzw. in *M. luteus* ins Wirtsgenom integrieren und Klone mit chromosomaler Integration über Kanamycinresistenz selektiert werden. Bei *Bacillus* sollte bei einer Insertion der metagenomischen DNA die Amylase-Expression verhindert sein, was mit einem einfachen Test nachzuweisen ist. Dazu werden erhaltende *Bacillus* Klone und der Wildtyp zur Kontrolle auf einer LB-Platte mit Stärke kultiviert und dann mit Lugolscher Lösung bedeckt. Ein heller Hof um den Wildtyp zeigt den Abbau von Stärke im Medium durch die Amylaseaktivität, während sich um Kolonien von Klonen mit Insertion metagenomischer Fosmid-DNA ins *amyE* Gen kein Hof bilden kann. Bei *M. luteus*-Klonen äußert sich eine erfolgreiche Integration durch eine Tryptophanauxotrophie. Stämme mit chromosomal integrierter Metagenom-DNA sollten nur auf Minimalmedium-Platten mit Tryptophanzusatz wachsen können.

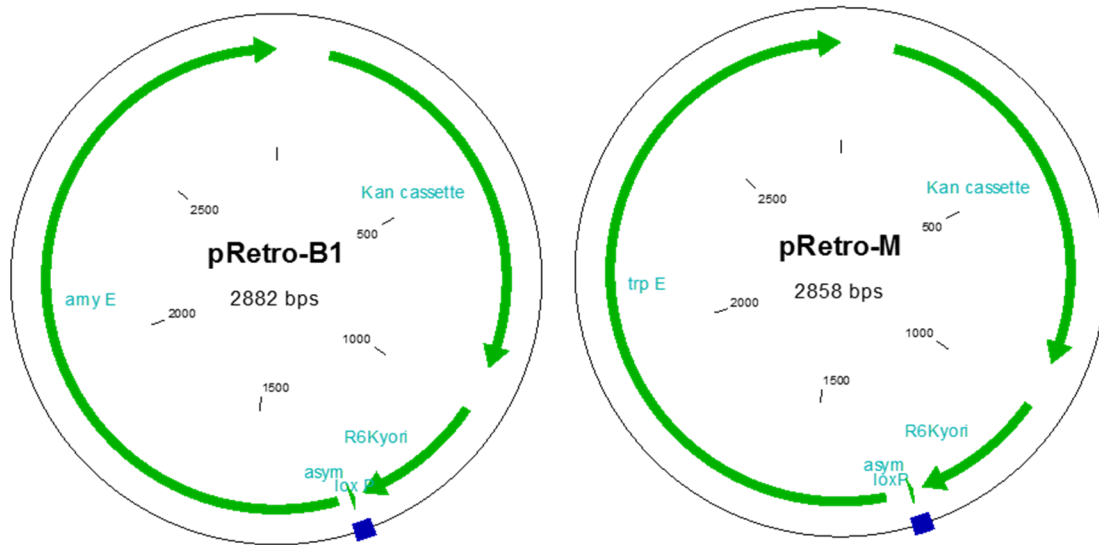


Abbildung 27: Vektorkarte von pRetro-B und pRetro-M. Die Kanamycinkassette, ori und der loxP-Seite ist bei beiden Vektoren identisch. pRetro-B trägt zusätzlich die Amylase Gensequenz amyE aus *B. subtilis* DB168 und pRetro-M das trpE Gen aus *M. luteus* ATCC 27141.

Das pRetro-M Plasmid konnte trotz Hertsellung verschiedener trpE-Amplifikate nicht generiert werden (trpE TP F/R: 1520 bp oder trpE F2/R2: 1013 bp). Nur die Konstruktionen von pRetro-B1 war erfolgreich. Dieser *Bacillus* spezifische Vektor und pJM2545 wurden zunächst in kompetente EPI300 Zellen mit dem Leervektor pCC1FOS transformiert und *in vivo* modifiziert. Zum Vergleich wurden noch der bereits evaluierte pRetro-T2 für *T. thermophilus* zusammen mit pCC1FOS im EPI300 transformiert und Negativkontrollen ohne die Zugabe von pJM2545. Es wurden jeweils die gleichen Mengen an DNA (1 µg) eingesetzt und später die Transformanden ausgezählt (Tabelle 35).

Tabelle 35: Übersicht der Komponenten der *in vivo* Modifizierung des Leervektors pCC1FOS und die Anzahl der daraus resultierten modifizierten EPI300 Kolonien (+ vorhanden; - nicht vorhanden).

| pCC1FOS | pJM2545 | pRetro-B1 | pRetro-T2 | <i>E. coli</i> EPI300 |
|---------|---------|-----------|-----------|-----------------------|
| +       | +       | +         | -         | ~ 40                  |
| +       | +       | -         | +         | ~ 700                 |
| +       | -       | +         | -         | 0                     |
| +       | -       | -         | +         | 0                     |

Bei dem zweiten Versuch der *in vivo* Modifizierung mit dem neuen pRetro-B1 Konstrukt wurde ein Metagenom-tragendes Fosmid ausgewählt: pCC1FOS-HA-CMC-1 und pCC1FOS-TB-EST-7. Die erhaltenen pRetro-B1 Plasmide wurden mit NcoI verdaut und die erhaltenen Fragmente gelelektrophoretisch aufgetrennt. Die Modifizierung mit dem pRetro-B Vektor war nicht erfolgreich, da sich das Bandenmuster vor und nach der Behandlung nicht unterschieden und die Bande von pRetro-B (2882 bp) in den retro gefitteten Proben nicht sichtbar ist (Daten nicht gezeigt).

#### 4.4.2 Bacterial artificial chromosome (BAC) in *T. thermophilus*

BAC ist die Abkürzung für *Bacterial Artificial Chromosome*. Bei BACs handelt es sich im Grunde um künstliche Chromosomen, die auf der Grundlage des low-copy F-Plasmids des Bakteriums *E. coli* entwickelt wurden. BACs dienen als Klonierungsvektoren und enthalten die Spaltstellen  $\lambda$ -*cosN*- und *P1-loxP*, einen Selektionsmarker sowie einen T7 oder Sp6 Promotor. BACs werden bei der Zellvermehrung stabil weitergegeben. Aufgrund ihrer großen Inserts von 150 bis 350 kb kann die Anzahl der zu durchforstenden Klone reduziert und auch große Gen-Cluster, die für ganze Stoffwechselwege kodieren, identifiziert werden. Zur stabilen Erhaltung als low-copy-number Replikons nach der Zellteilung dient ein *par*-System. Dieses System umfasst drei Genkomponenten: die ATPase (*parA*). Das DNA-bindende Protein ParB (*parB*) und eine Zentromer-ähnliche Bindestelle *parS*.

In dieser Arbeit wurde versucht BAC-Vektoren zu erstellen, die große Insertion metagenomischer DNA tragen können und entweder als zusätzliches externes Plasmid in *T. thermophilus* bestehen bleiben oder sich sogar in das Genom integrieren. Für die eigenständige Replikation des modifizierten BACs in *T. thermophilus* wurden die putativen Replikationselemente des Megaplasmids pTT27 (*repA*, *ori* und die *par*-Locus Gene *parA,B* und *parS*) in den Klonierungsvektor integriert.

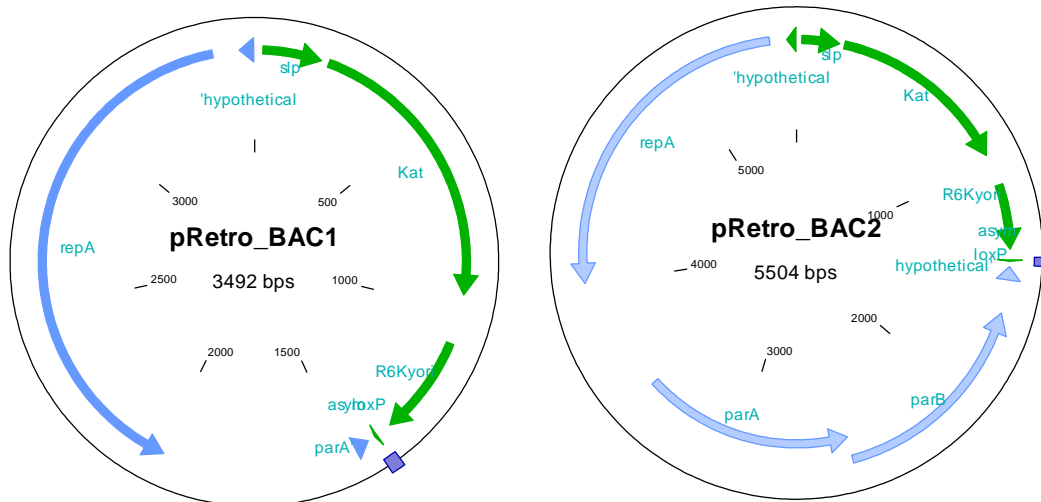


Abbildung 28: Vektorkarten von pRetro-BAC1 und pRetro-BAC2

Für die Konstruktion von pRetro-BAC1 wurde pRetro-T2 mit BamHI und SacI geschnitten und das 1434 bp lange Fragment aus dem Gel aufgereinigt. Dieses DNA-Stück enthält den Replikationsursprung R6Kyori und die Kanamycinresistenz Kasette. Eine weitere PCR wurde mit dem Primerpaar pRetro-BAC1 F/-BAC1 R und der genomischen DNA von *T. thermophilus* als Template durchgeführt. Dadurch wurde ein 2062 bp langes Amplifikat erhalten, welches das *repA* Gen und den putativen *ori* (hoher AT-Gehalt) des Megaplasmids umfasst. Für die zweite Variante pRetro-BAC2 wurde ein 4074 bp Fragment des Megaplasmids vervielfältigt, das neben dem *repA* Gen und dem *ori* noch zusätzlich den *par* Locus mit *parA*, *B* und *S* mit einschließt. Nach dem Gibson Assembly wurden kompetente EC100D (pir+) Zellen mit den Vektoren pRetro-BAC1 und pRetro- transformiert, die aus dem jeweiligen Megaplasmid-Fragmente BAC1 bzw. BAC2 und dem pRetro Rückgrat bestehen.

Nach der Validierung der Vektoren durch Sequenzanalysen (Sequenzprimer pRetro-B S1) wurden sie jeweils mit zwei verschiedenen Fosmiden Lt/2/cell A1 und Lt/2/cell P1 modifiziert. Der Kontrollverdau der Fosmide wies darauf hin, dass die *in vivo* Modifizierung erfolgreich war. Für die weiteren Experimente wurden nur die BAC-Vektoren mit integriertem Fosmid Lt/2/cell A1 weiter verwendet und als pRetro-rf-BAC1 und pRetro\_rf-BAC2 bezeichnet. Mit den vier DNA-Vektoren pRetro-BAC1, pRetro-BAC2, pRetro-rf-BAC1 und pRetro-rf-BAC2 (kurz: BAC1, BAC2, rf\_BAC1 und rfBAC2) wurde *T. thermophilus* Wildtyp HB27wt transformiert.

Für den Southern Blot wurde die Kanamycinkasette als Hybridisierungsprobe verwendet. Die Negativkontrolle mit dem Wildtypstamm HB27 (1) zeigt keine klare Bande, während die Stämme mit dem Vektor pRetro-BAC1 bzw. pRetro-BAC2 (2 und 3) oder mit den „retro gefitteten“ Fosmiden (4 und 5) ein deutliches Signal bei der genomischen DNA aufweisen.

Diese Ergebnisse deuten auf eine Integration der BAC-Vektoren ins *T. thermophilus* Genom hin. In der 6. Spalte kann man das linearisierte pRetro-BAC1 Plasmid sehen (Abbildung 29).

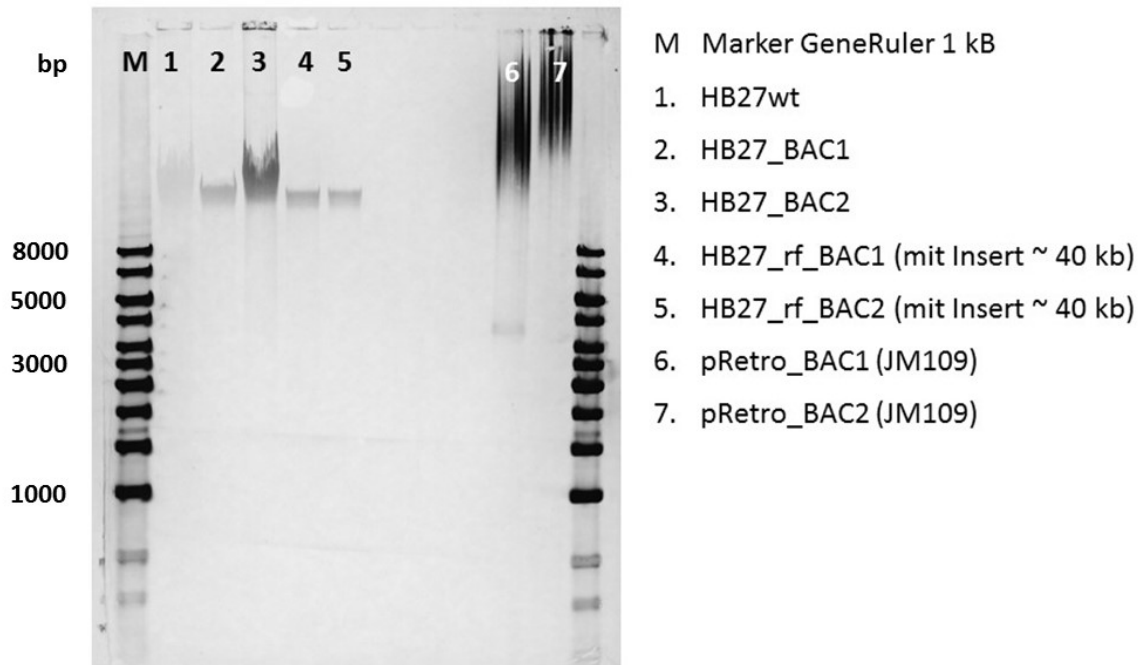


Abbildung 29: Southern Blot der pRetroBAC1 und BAC2 Konstrukt. Als Marker (M) dient der GeneRuler 1 kb von der Firma Fermentas.

Für den Nachweis der erfolgreichen Transformation und auch zur Lokalisationsbestimmung der pRetro-BAC Vektoren wurden zunächst Geldokumentationsmethoden angewendet, denn die Plasmide könnten sich extrachromosomal in den Zellen befinden oder über die homologen Sequenzen in das Megaplasmid pTT27 integriert worden sein. Sowohl bei der Eckhardts Gel- als auch bei der Pulsefield Gel-Methode werden ganze Bakterienzellen nach einer Zellysebehandlung auf ein Agarosegel aufgeladen, damit die gesamte Nukleinsäure durch das Gel aufgetrennt und analysiert wird. Jedoch wurden mit beiden Anwendungen keine Gelbilder mit deutlich sichtbaren Banden erzielt (siehe Anhang; Abbildung 31). Aus diesem Grund wurde eine weitere Visualisierungsmöglichkeit gesucht und ein Southern Blot mit der Kanamycin Kasette als Hybridisierungs-Probe durchgeführt. Die Sonde ist ein 823 bp Fragment, welches aus dem Plasmid pMOD3\_Kat (Doppelverdau mit EcoRI und BglII) gewonnen wurde. Die *Thermus*-Stämme mit den BAC Vektoren wurden mit unterschiedlichen Kanamycinkonzentrationen kultiviert (Kan 0, 50, 100, 200 µg/mL). Die aufsteigende Menge des Antibiotikums verstärkt dabei den Selektionsdruck auf Integration und kann die Kopienzahl des BAC-Vektors erhöhen.



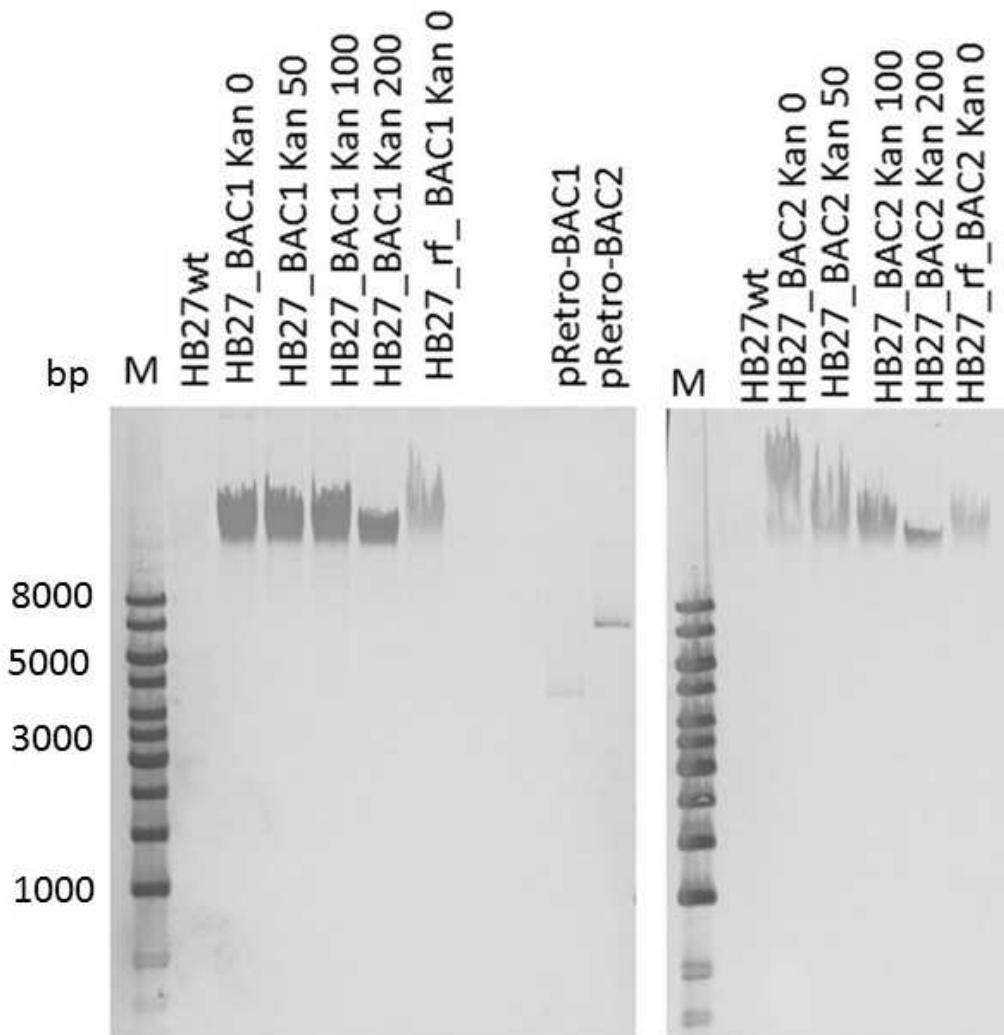


Abbildung 30: Southern Blot mit gDNA aus Zellen, die bei steigender Kanamycin Konzentration kultiviert wurden. Als Marker (M) wurde der GeneRuler 1 kb von der Firma Ferments verwendet. Auf dem linken Bild sind die *Thermus*-Stämme, die mit pRetro-BAC1 modifiziert wurden, das rechte Bild zeigt den pRetro-BAC2 Vektor. Die Stämme mit der Bezeichnung rf (retro-fitting) im Namen tragen ein mit pRetro-BAC1 bzw. BAC2 *in vivo* modifiziertes Fosmid. Der Wildtyp-Stamm HB27 wurde als Negativkontrolle verwendet.

Die Kultivierung von HB27\_BAC1 und HB27\_BAC2 war bei allen Kanamycinmengen erfolgreich, während die *Thermus*-Stämme mit dem retro-gefitteten BAC-Vektor nur bei Kan 0 und auf den Platten bei Kan 20 angewachsen sind. Das lässt vermuten, dass Kanamycinkassette in den Stämmen rf-BAC1 und rf-BAC2 nicht vervielfältigt wurde. Bei der Southern Blot Hybridisierung wurde die DNA des Wildtyps nicht angefärbt. Die Proben mit den BAC1 und BAC2 Vektoren zeigen diffuse Banden weit oberhalb der 10 kb Bande was ein Hinweis auf die Integration der Vektoren in das Megaplasmid darstellt (Abbildung 30).

## 5 Diskussion

### 5.1 Eigenschaften der neuartigen Glykosidhydrolase EngU

Die Untersuchung der Glykosidhydrolas EngU knüpfte an die Ergebnisse früherer Arbeiten (Wittenberg, 2009) an und die Enzymcharakterisierung wurde erweitert, sowie zwischenzeitlich neu verfügbare Sequenzdaten mit einbezogen.

#### 5.1.1 Sequenz- und Modulvergleiche von EngU

Frühere, zur Zeit der Isolierung und Sequenzierung des *engU*-Gens durchgeführte BLAST-Suchen mit der Sequenz von EngU gegen die verfügbaren, öffentlich zugänglichen Sequenzdatenbanken hatten nur schwache und nur partielle Sequenzähnlichkeit mit damals hinterlegten Sequenzen ergeben (Wittenberg, 2009). Daher bestand das Risiko das rekombinante Fosmid könne eventuell eine chimäre DNA-Insertion aus verschiedenen, eigentlich nicht zusammengehörigen, metagenomischen DNA-Fragmenten tragen. Eine neuerliche BLAST-Suche ergab jedoch später ca. 25 Sequenzen, die über die gesamte Proteinlänge Ähnlichkeiten zeigen und aus unterschiedlichen Organismen stammen (Tabelle 36). Dies ist ein Beweis dafür, dass die Gensequenz kein Artefakt ist und nicht durch zufällige Klonierungsereignisse entstanden ist.

Tabelle 36: Überblick über die besten Treffer der BLAST Analyse zur Suche nach Proteinsequenzen mit Ähnlichkeiten zu der gesamten EngU Länge.

| Name                                 | Beschreibung                         | Identität | Accession Nr.  |
|--------------------------------------|--------------------------------------|-----------|----------------|
| <i>Planctomycetes</i><br>Bakterium   | hypothetisches Protein               | 51%       | KPK65117.1     |
| <i>Bathyarchaeota</i><br>Archaeon    | Kohlenhydrat-bindende CenC<br>Domäne | 47%       | KYH37837.1     |
| <i>Latescibacteria</i><br>Bakterium  | hypothetisches Protein               | 48%       | KPJ58997.1     |
| <i>Armatimonadestes</i><br>Bakterium | Beta-Galaktosidase                   | 47%       | CUU00576.1     |
| <i>Chthoniobacter flavus</i>         | Kohlenhydrat-bindendes<br>Protein    | 46%       | WP_006980642.1 |
| <i>Ignisphaera aggregans</i>         | Kohlenhydrat-bindende CenC<br>Domäne | 45%       | ADM28658.1     |
| <i>Cellvibrio mixtus</i>             | hypothetisches Protein               | 41%       | WP_052417202.1 |
| <i>Opitutus terrae</i>               | Kohlenhydrat-bindende CenC<br>Domäne | 41%       | WP_012376627.1 |

|                                     |                                |     |                |
|-------------------------------------|--------------------------------|-----|----------------|
| <i>Paludibacterium yongneupense</i> | hypothetisches Protein         | 41% | WP_028534507.1 |
| <i>Dickeya sp.</i>                  | hypothetisches Protein         | 40% | WP_049853918.1 |
| <i>Teredinibacter turnerae</i>      | hypothetisches Protein         | 38% | WP_018275959.1 |
| <i>Opitutaceae</i> Bakterium        | Kohlenhydrat-bindendes Protein | 36% | WP_009508551.1 |
| <i>Diplosphaera colitermitum</i>    | hypothetisches Protein         | 36% | WP_043588348.1 |
| <i>Melioribacter roseus</i>         | Kohlenhydrat-bindendes Protein | 32% | WP_014855430.1 |

Die meisten Bakterienstämme in Tabelle 36 wurden aus mesophilen oder thermophilen Gebieten isoliert oder besitzen ihr Temperaturoptimum in diesem Bereich (Kadnikov et al., 2013; Kwon et al., 2008; Marrero et al., 2013). Die ideale Wachstumstemperatur von *Ignisphaera aggregans* liegt bei 92 - 95 °C, die von *Plantomyces bacterium* bei 35 - 55 °C und die von *Teredinibacter turnerae* liegt zwischen 30 und 35 °C (Distel et al., 2015; Fuerst et al., 1997; Niederberger et al., 2006). Auffällig ist, dass unter den annotierten Genen keine Lichenase- oder  $\beta$ -Glukanasegene sind, wohingegen EngU gemäß unserer funktionellen Assays diese Aktivitäten aufweist.

Viele Treffer kamen aus dem Phylum der *Verrucomicrobia*, welches aus fünf Untergruppen (V1 - V5) besteht und bspw. die Stämme *Chthoniobacter flavus* und *Opitutus terrae* mit einschließt (Chin and Janssen, 2002; Choo et al., 2007). *Chthoniobacter flavus* wird häufig in pflanzlicher Biomasse gefunden und besitzt daher die Fähigkeit viele Polysaccharide zu metabolisieren (Sangwan et al., 2004). Genauso wie *Ignisphaera aggregans*, der eine große Bandbreite verschiedener Substrate wie z.B. Stärke, Konjak Glukomannan, Pepton, Mannose, Laktose, Galaktose, Maltose, Glykogen und  $\beta$ -Cyclodextrin als Kohlenstoffquelle verwenden kann (Niederberger et al., 2006). Andere Mikroorganismen werden bereits in der Industrie genutzt wie *Cellvibrio mixtus*, dessen *endo*-1,4- $\beta$ -Xylanase im großen Maßstab hergestellt und vermarktet wird (Megazym, USA). Die Treffer der Homologieanalyse über die gesamte EngU-Proteinlänge beziehen sich auf Daten, die aus der single-cell Sequenzierung erhoben wurden. Zum jetzigen Kenntnisstand (Stand 11.2017) gibt es keine orthologen *engU* Gene, die aus kultivierbaren Mikroorganismen stammen.

Die PFAM Datenbank ist eine große Sammlung von Proteinfamilien, die anhand eines Multiplen Sequenzen Alignments oder eines Hidden Markov Modells dargestellt werden. Mithilfe jener Datenbankanalyse werden vorhandene Module bzw. Domänen in einer Proteinsequenz identifiziert, die wiederum einen Rückschluss auf die Funktion des Proteins

geben. Die Sequenzanalyse von EngU in PFAM ergaben Ähnlichkeit zu drei funktionellen Modulen (GH5, GH42 und CBM4) und die Alignments mit den jeweiligen Seeds zeigten einen für Glykosidhydrolasen sehr ungewöhnlichen modularen Aufbau.

Die GH5 Familie umfasst eine große Vielfalt an Enzymen: *endo*- $\beta$ -1,4-Glukanase/Cellulase (EC 3.2.1.4); *endo*- $\beta$ -1,4-Xylanase (EC 3.2.1.8);  $\beta$ -Glukosidase (EC 3.2.1.21), Lichenase (EC 3.2.1.73),  $\beta$ -Mannosidase (EC 3.2.1.25) und weitere Aktivitäten. Die Familie GH42 beinhaltet Enzyme mit  $\beta$ -Galaktosidase und  $\alpha$ -L-Arabinopyranosidase Aktivität, die zum Clan GH-A gehören und glykosidische Bindungen unter Beibehaltung der anomeren Konfiguration (Retention) hydrolysieren. Bei den Grundmerkmalen wie Mechanismus, Clan und Kristallstruktur besitzen beide Familien, GH5 und GH42, jedoch viele Gemeinsamkeiten. Ihre Sekundärstruktur wird als „TIM-barrel“ (TIM steht für Triosephosphat-Isomerase) bezeichnet und besteht aus acht parallelen  $\beta$ -Faltblattsträngen, die durch acht  $\alpha$ -Helices miteinander verbunden werden. Die Faltblattstränge bilden die Dauben einer Fassstruktur (engl.: *barrel*), die von den äußeren Helices festgehalten werden (Goldman et al., 2016; Nanda, 2016). Die Struktur-basierten Alignments der Sequenzen der zu der GH42 Familie ähnlichen Sequenzabschnitte von EngU (GH5-GH42a und GH42b) weisen darauf hin, dass diese Sequenzabschnitte zwei Hälften eines ( $\beta/\alpha$ )<sub>8</sub> TIM-Barrels darstellen. Zwischen den beiden Hälften und nach dem vierten  $\beta$ -Faltblattstrang ist ein Kohlenhydrat-Bindemodul (CBM) insertiert, welches im nächsten Absatz näher beschrieben wird. Eine „Sandwich“-ähnliche Struktur wurde auch bei einer GH42  $\beta$ -Galaktosidase aus *T. thermophilus* A4 entdeckt. In diesem Fall wurde die TIM-Barrel Struktur zwischen  $\beta$ -4 und  $\alpha$ -4 durch eine Subdomäne H unterbrochen (Hidaka et al., 2002a). In dieser Veröffentlichung wurden die zwei Barrel-Hälften *in vivo* co-exprimiert und *in vitro* zu einem funktionellen Enzym rekonstruiert (Höcker et al., 2001). Diese Strukturbeispiele, mit EngU eingeschlossen, bestärken die Theorie, dass die TIM Barrel Struktur aus einer Genduplikation und einer Genfusion aus Halb-Barrel Vorläufern entstanden sein könnte (Lang et al., 2000).

CBMs sind 30 - 200 Aminosäuren lange Module in Glykosidhydrolasen, die Oligo- oder Polysaccharide erkennen und binden können. Zudem können diese Module einer Disruption des Substrates entgegenwirken und diese für die Abbaureaktion stabilisieren (Shoseyov et al., 2006). Oft enthalten hydrolytisch aktive Proteine ein bis drei solcher Module, die über Verbindungssequenzen (engl.: *linker sequences*) verknüpft sind (Gilkes, et al. 1995). In Cellulosomen unterstützen CBMs die katalytischen Module (Ding et al., 2003; Zverlov et al., 1998b). Die Bindemodule wurden in vielen unterschiedlichen Enzymklassen und Wirten bereits entdeckt, wie z.B. Cellulasen und Xylanasen aus verschiedenen Mikroorganismen, Maltodextranase aus *Arthrobacter globiformis*, Arabinofuranosidasen aus *Aspergillus kawachii*,  $\beta$ -Glukosidasen aus *Phanerochaete chrysosporium*, Esterasen aus *Penicillium*

*funiculosum*, Endoglukanasen aus Pflanzen, Endomannanase aus *T. reesei*, oder Pektin Lyase aus *Pseudomonas cellulosa* (Charnock et al., 2001; Kroon et al., 2000; Lymar et al., 1995; Miyanaga et al., 2004; Shoseyov et al., 2006; Stålbrand et al., 1995; Yoshida et al., 2006). CBMs werden, ähnlich wie Glykosidhydrolasen, anhand ihrer Aminosäuresequenz, Tertiärstruktur und Substratspezifität in 83 Familien unterteilt (Stand 11.2017). Durch Sequenzanalysen wurde in EngU eine CBM der Familie 4\_9 mit einer 46 %igen Ähnlichkeit zu einem hypothetischen Protein aus *Planctomyces* Bakterium SM23\_32 identifiziert. Die Module dieser Klasse werden dem Typ B zugeordnet und binden hauptsächlich lösliche Glukanketten in ihrer Spalt-artigen Substratbindungstasche (Boraston et al., 2004). Die CBM4-Kristallstruktur aus Cellulase K von *Clostridium thermocellum* wurde untersucht und Alahuhta (2011) fand in einer Peptidschleife der Bindetasche einen in Clostridien hoch konservierten Tryptophanrest Trp138 (118 in CbhA CMB4). Im Ähnlichkeitsmuster („Seed“) der CBM4-Familie befindet sich ebenso ein konservierter Tryptophanrest an Position 17, der auch in EngU vorhanden ist (Pos. 310 in der EngU-Sequenz bzw. Pos. 20 der CBM4\_9 Moduls von EngU).

Bei diversen modularen aufgebauten Glykosidhydrolasen findet man flexible Linker zwischen manchen der Module, zum Beispiel die thermostabile *endo*-1,4 Glukanase (Cel7A) aus *Neosartorya fischeri* P1 enthält eine GH7 Domäne, die über eine 37 Aminosäuren lange Linkerregion (71 %iger Thr und Ser Gehalt) mit einem CBM der Klasse 1 verbunden ist (Liu et al., 2015). Ähnliche Verbindungsstrukturen findet man auch bei der *endo*-Glukanase in *Myceliophthora thermophila* und bei der GH61 Cellulase von *Sporotrichium thermophile* (Dimarogona et al., 2012; Karnaouri et al., 2014). Die Aminosäuresequenzen vor und nach der CBM Domäne von EngU weisen gemäß Analyse mit dem DLP-SVM Web Service (Ebina et al., 2009) keine signifikanten Linker Regionen auf. Dies könnte auf eine relativ starre Verbindung der Module von EngU hinweisen.

### 5.1.2 Optimierung der EngU-Proteinbiosynthese

Die Funktionalität der EngU Module konnte nicht vollständig identifiziert werden, da einige Konstrukte mit einzelnen Modulen nicht exprimiert wurden. Gründe dafür könnten der falsche Codon-Usage, Expression toxischer Produkte oder inkorrekte Faltung der Proteine sein. Für eine effizientere Auswertung dieser Versuche muss die Synthese von löslichen Proteinen erhöht und optimiert werden. Dafür wurden verschiedene *E. coli*-Stämme als Expressionszellen verwendet, die Plasmide für eine spezielle Codon Usage tragen. Rosetta sind Varianten von BL21 die sechs seltene tRNAs für die Codone AUA, AGG, AGA, CUA, GGA und CGG exprimieren können (bei Rosetta 2 zusätzlich CGG) (Tegel et al., 2010). Die Überexpressionsstämme BL21\_C41 und C43 sind speziell für die Produktion von toxischen Proteinen konzipiert. Die mutierten Stämme regulieren die T7 RNAP Aktivität herunter, steuern dadurch einer Anreicherung toxischer rekombinanter Proteine entgegen und verhindern somit den Zelltod. BL21\_C43 trägt mehr Mutationen als das Derivat C41 und ist noch toleranter gegenüber giftigen Polypeptiden. Von diesen beiden BL21 Stämmen gibt es noch eine weitere Variante mit dem zusätzlichen pLysS Plasmid mit einem Chloramphenicol-Resistenzgen. In diesen Zellen wird eine geringe Menge an T7-Lysozym gebildet, welches ein natürlicher Inhibitor der T7-RNA-Polymerase ist und wiederum entgegen der Akkumulation toxischer Proteine wirkt. Die Expression der pETDuet-Konstrukte hat in keinem der BL21-Derivate funktioniert. Auch die Anwendung von „Arctic Express“-Zellen, die eine Expression bei niedrigen Temperaturen von z.B. 10 °C aufgrund ihrer kälteadaptierten Chaperone (Cpn10, Cpn60) aus *Oleispira antarctica* ermöglichen (Belval et al., 2015), führte zu keiner gesteigerten Proteinproduktion.

Chaperone sind Proteine, die an der Faltung anderer Proteine beteiligt sind. Sie wirken der Aggregationsreaktion entgegen und beschleunigen die korrekte Faltung. Dabei können sie nicht nur neu synthetisierte Proteine, sondern auch Proteinaggregationen zu ihrer funktionellen Struktur verhelfen (Fink, 1999; Kellner et al., 2014). Viele Chaperone werden durch Stressfaktoren wie Temperatur (Hitzeschock- oder Kälteschock-Proteine), oxidativem Stress oder toxischen Substanzen aktiviert (Sekhar et al., 2012). Die Klassifizierung der Faltungshelfer erfolgt entweder nach ihrem Molekulargewicht, ihrer Sequenzhomologie oder Funktionalität. Demnach gibt es die Klasse der „Folding“-Chaperone, die durch ATP-angetriebene Konformationsänderung andere Proteine ent- und neufalten. Dazu gehören beispielsweise DnaK und GroEL. Die „Holding“-Chaperone (Hsp33, Hsp31 und IbpB) halten die zu modifizierenden Proteine auf ihrer Oberfläche fest und bereiten sie somit für andere Faltungshelfer vor. Die letzte Gruppe bilden Chaperone wie ClpB, die Proteinaggregate resolubilisieren und umstrukturieren können (Fink, 1999). Nach der Auswertung der SDS-Gele und Aktivitätstests wurde klar, dass der C-Terminus (665.-905. Aminosäure) von EngU

offenbar eine wichtige Rolle für die Bildung eines funktionalen, enzymatisch aktiven Proteins EngU spielt und eventuell eine Chaperon-ähnliche Funktion haben könnte.

Das gibt einen Hinweis darauf, dass die Genabschnitte gh42a\_cbm und gh42b\_cterm beide vorhanden sein müssen, um die Bildung eines funktionstüchtigen Enzyms zu garantieren. Die beiden Polypeptidketten GH42a-CBM und GH42b-CTERM komplementieren zu einem funktionellen Enzym, während durch Expression der einzelnen Genfragmente keine Aktivität entstand. Die Expression der Plasmidkonstrukte pETDuet\_[gh42a\_cbm]\_gh42b\_cterm und pETDuet\_[gh42]\_cterm führte zu löslichen und aktiven Enzymen (Abbildung 5; Nr. 2 und 7). Beide Enzyme wiesen eine reduzierte Aktivität von 82,5 % und 6,6 % im Vergleich zum EngU Protein auf. Die einzeln aber gleichzeitig synthetisierten Proteinketten konnten sich in der *E. coli* Zelle komplementieren und zu einem funktionsfähigen Enzym zusammenfallen. Ein bekanntes Beispiel für die Komplementation getrennt gebildeter Enzymteile ist  $\alpha$ -Komplementation bei der  $\beta$ -Galaktosidase (Ullmann et al., 1967). Dieser Mechanismus wird für die Blau/Weiß-Selektion in modifizierten *E. coli* Stämmen genutzt. Die Zellen tragen ein Plasmid, das einen kleinen Teil des Gens für die  $\beta$ -Galaktosidase (*lacZ*-Gen) mit einer der Multiple Cloning Site kodiert, der Rest des *lacZ* Gens liegt im Chromosom vor. Bei einer Insertion eines Transgens in die Multiple Cloning Site wird die Basensequenz des *lacZ*-Genfragments unterbrochen und das Enzym inaktiviert. Die Mutantenstämme können dann das in den Agarplatten enthaltene Substrat X-Gal (5-Brom-4-chlor-3-indoxyl- $\beta$ -D-Galaktopyranosid) nicht mehr zu dem blauen Farbstoff 5,5'-Dibromo-4,4'-dichlor-Indigo und Galaktose spalten. Mit Hilfe der unterschiedlichen Phänotypen können die Transgen-tragenden weißen Zellen von den blauen Zellen ohne die Insertion des Transgens schnell differenziert werden (Welply et al., 1981).

### 5.1.3 Putative katalytische Reste des EngU Proteins

Punktmutationen in der Aminosäuresequenz können gravierenden Einfluss auf die Eigenschaften des Proteins wie z.B. die Substratspezifität, das pH- oder das Temperaturoptimum nehmen (Jia et al., 2012; Matsuzawa et al., 2014; SD. Baird, M.A. Hefford, D. A. Johnson, 1990; Tishkov et al., 2013). Beispielsweise wurde die  $\beta$ -1,3 Glukanase Lam81A aus *Thermobifida fusca* auf ihre Substratspezifität und -bindung untersucht, indem 14 konservierte Aminosäuren mutiert wurden (McGrath et al., 2009). Mithilfe der Mutanten konnte die katalytische Base bestimmt werden, die bei der Substratorientierung eine Rolle spielt. Darüber hinaus führte die Substitution mit Alanin zu einer Verringerung der Hydrolyseaktivität, während durch einen Austausch mit Tyrosin die Spaltung von  $\beta$ -Glukan und CMC verstärkt wurde (McGrath et al., 2009). Nach der Mutation der beiden katalytischen Reste in der thermostabilen Exoinulase aus *Kluyveromyces cicerisporus* folgte der komplette Verlust der Aktivität (Ma et al., 2015).

Der GH42 *hmm* Alignment weist auf die zwei konservierte Glutamatreste E136 und E295 hin, welche die katalytischen Reste des Enzyms darstellen. Die entsprechenden Aminosäuren finden sich in der EngU Proteinsequenz an den Positionen E239 (INNEN) und E581 (VTEYN). Die beiden Positionen wurden jeweils mit der kleineren nicht polaren Aminosäure Alanin ausgewechselt, daraufhin verloren die mutierten Enzyme komplett ihre Aktivität. In Abbildung 4 ist gut zu erkennen, dass die katalytischen Reste im Alignment weit voneinander getrennt sind. Nach der Vorhersage der katalytischen Reste in A4- $\beta$ -Gal aus *T. thermophilus* (Hidaka et al., 2002b) wird vermutet, dass der nukleophile Rest E239 im ersten Teil der GH42 Moduls (GH42 a) liegt, während sich der Säure/Base Rest in dem zweiten Abschnitt des GH42 Seeds befindet (GH42 b). Dies ist wieder ein Hinweis für die postulierte Insertion des CB-Moduls in das katalytische Modul. Das Alignment der GH42  $\beta$ -Galaktosidasen von verschiedenen Mikroorganismen bildete eine Konsensussequenz für den Säure/Base (-GWH-SNE-) und für den nukleophilen (-LMES-) Rest, aber diese Aminosäureabfolgen sind nicht in dem EngU Modul zu finden (Di Lauro et al., 2008). Sakon *et al.* haben die Kristallstruktur der thermostabilen Endocellulase E1 von *Acidothermus cellulolyticus* untersucht (Sakon et al., 1996). Laut ihren Ergebnissen haben acht konservierte Reste, die in der Familie GH5 vorhanden sind, einen großen Einfluss auf die Cellulaseaktivität. In dieser Arbeit wurden auch weitere Reste in EngU mutiert, die entweder einen verstärkenden oder schwächenden Effekt auf die katalytische Reaktion ausübten: D123A (2,97 %), D151A (104,01 %), E183A (124,02 %), E235A (27,84 %), E593A (51,35 %) und E605 (32,43 %) im Vergleich zum Wildtyp (100%). Dennoch stimmt keines der mutierten Aminosäurenreste mit denen von Sakon *et al.* deklarierten essentiellen Resten in GH5 Familien überein. D123 reduzierte die Enzymaktivität am stärksten, was man durch die Nähe zu dem hoch konservierten Argininrest Arg62 des



GH5 Seeds erklären kann. Für eine weitere Analyse wird eine 3D Kristallstruktur benötigt, um die genauen Interaktionen der katalytischen Reste durchleuchten zu können.

#### 5.1.4 Substratspektrum und -spezifität von EngU

Bei dem Aktivitätsassay mit EngU wurden die höchsten Werte mit  $\beta$ -Glukan und Lichenan gemessen. Diese Polysaccharide bestehen beide aus 400 bis 1200 linear  $\beta$ -glykosidisch verknüpften Glukose-Einheiten. Ein bis drei  $\beta$ -1,4-verknüpfte Glukosemoleküle werden jeweils durch eine  $\beta$ -1,3-glykosidische Bindung an die nächste  $\beta$ -1,4-Glukosekette gebunden. Es gibt drei verschiedene Enzymgruppen, die für die  $\beta$ -Glukan Spaltung verantwortlich sind: (i) Cellulasen hydrolysieren die  $\beta$ -1,4-Bindungen (ii) Laminarinasen spalten  $\beta$ -1,3- und  $\beta$ -1,4-Verknüpfungen und (iii) Lichenasen können nur  $\beta$ -1,4-Verbindungen hydrolysieren, wenn diese auf  $\beta$ -1,3-Bindungen folgen. Neben den beiden Hauptsubstraten zeigt EngU auch eine geringe Aktivität auf Carboxymethylcellulose und Hydroxyethylcellulose, die nur aus  $\beta$ -1,4-glykosidisch verknüpften Glukosemolekülen bestehen. Nach einem vierstündigen Abbau des  $\beta$ -Glukans durch EngU unter den hier gewählten Bedingungen wurden hauptsächlich Cellotetraose und eine geringe Menge an Cellotriose nachgewiesen. Nach 24 Stunden war die Hydrolyse abgeschlossen und es wurden keine weiteren Spaltungsprodukte gebildet. Basierend auf diesen Ergebnissen wurde EngU als eine *endo*- $\beta$ -1,3- $\beta$ -1,4-Glukanase klassifiziert, da es in der Lage ist,  $\beta$ -1,4-verbundene Polysaccharide wie Carboxymethylcellulose,  $\beta$ -1,3-1,4-verknüpfte Polysaccharide wie Lichenan und  $\beta$ -Glukan zu hydrolysieren, jedoch keine rein  $\beta$ -1,3-verknüpften Polysaccharide wie Curdlan oder Polysaccharide mit  $\beta$ -1,3- und  $\beta$ -1,6-Bindungen wie Laminarin degradieren kann.

In dieser Arbeit wurden weitere Substrate unter Anwendung von DNSA-Assays und Dünnschichtchromatographie getestet (siehe Tabelle 27). Verglichen mit der Aktivität mit  $\beta$ -Glukan (100 %) wurden geringere Aktivitäten mit Glukomannan aus Konjak (27,83 %) und Xyloglukan (7,38 %) gemessen. Xyloglukan ist eine Hemicellulose und besitzt ein Rückgrat aus  $\beta$ -1,4-glykosidisch verknüpften Glukoseresten und Seitenketten aus  $\beta$ -1,6 verknüpften D-Xyloseresten. Zudem können die Reste mit D-Galaktose oder L-Fruktose verlängert sein. In höheren Pflanzen sind die Xyloglukane in der Primärzellwand eng mit Cellulosefibrillen assoziiert. Glukomannan ist unverzweigt und besteht aus Mannose und Glukose im Verhältnis 60 zu 40, die  $\beta$ -D-1,4-glykosidisch verknüpft sind.

Mithilfe der unterschiedlich verknüpften Glukotetraosen B und C wurde der Spaltmechanismus von EngU untersucht. Die Glukotetraose B wurde hauptsächlich in Trimere und Monomere gespalten, während aus Glukotetraose C nur Cellobiose entstand (Abbildung 8). Zur Vervollständigung dieses Experiment fehlt die Variante A, deren  $\beta$ -1,3-glykosidische Bindung

zwischen dem 1. und 2. Glukosemolekül lokalisiert ist. Dieses spezifische Substrat ist leider kommerziell nicht mehr verfügbar und stand daher für weitere Spaltungsanalysen nicht zur Verfügung. Wenn die Glukotetraose A nicht gespalten wird, würde es bedeuten, dass EngU die  $\beta$ -1,3-glykosidische Verbindung nur schneidet, wenn diese sich in der Nähe eines reduzierten Zuckerendes befindet. Eine andere Schlussfolgerung würde sich ergeben, wenn sich nur Disaccharide als Spaltprodukt ergäben. Das würde bedeuten, dass in diesem Fall nicht die  $\beta$ -1,3- sondern die mittlere  $\beta$ -1,4-Bindung der Tetrasaccharide gespalten wurde. Dieses Ergebnis könnte sich gut darstellen lassen, indem man die Lauflänge der Dünnschichtchromatographie verlängert. Dadurch würde das Disaccharid mit der  $\beta$ -1,3 Bindung vermutlich eine andere Laufweite aufweisen als das Disaccharid mit der  $\beta$ -1,4-glykosidischen Verbindung. Mit den neuen Erkenntnissen mit den Glukotetraose-Substraten B und C und dem sonstigen Substratspektrum (Hydrolyse von CMC, mixed-linkage  $\beta$ -Glukan und Lichenan) kann man über den Spaltungsmechanismus von EngU sagen, dass das Enzym *endo*- $\beta$ -1,3-spaltende Aktivität und schwache *endo*- $\beta$ -1,4-spaltende Aktivität besitzt, aber bei kurzkettigen Zuckermolekülen eine *exo*- $\beta$ -1,3-spaltende Aktivität aufweist, wenn auf die  $\beta$ -1,3-Bindung eine  $\beta$ -1,4 glykosidische Bindung folgt. Im Gegensatz zu typischen bisher beschriebenen Lichenasen hydrolysiert EngU vorwiegend die  $\beta$ -1,3-Bindungen und nicht die  $\beta$ -1,4-Bindungen von mixed-linkage  $\beta$ -1,3- $\beta$ -1,4-Glukan und Lichenan.

### 5.1.5 Biotechnologisches Potential von EngU

Aufgrund des hohen Temperaturoptimums bei 90 °C und seiner Substratspezifität ist die  $\beta$ -1,3-1,4-Glukanase EngU ein interessanter Biokatalysator für industrielle Prozesse. Bei der Bierherstellung ist Gerste- $\beta$ -Glukan ein großer Störfaktor. Durch die Viskosität der Maischelösung kann der Filterprozess verlangsamt werden und sich später im fertigen Bier absetzen (Walsh et al., 1994). Der Maischeprozess läuft bei hohen Temperaturen von bis zu 80 °C ab und bietet somit gute Bedingungen für die thermostabile  $\beta$ -1,3-1,4-Glukanase EngU. Weitere Optimierungsmöglichkeiten für die Wein- und Bierherstellung wurden in der Veröffentlichung von Van Rensburg et al. beschrieben (1997). Die *exo*- $\beta$ -1,3-Glukanase der Bierhefe *Saccharomyces cerevisiae* wurde durch die Überexpression der *endo*- $\beta$ -1,3-1,4-Glukanase von *Bacillus subtilis* und der *endo*- $\beta$ -1,4-Glukanase von *Butyrivibrio fibrisolvens* komplementiert.

### 5.1.6 EngU als Prototyp für die neue GH-Familie (GH148)

EngU ist eine Glukanase, die einen ungewöhnlichen modularen Aufbau besitzt und eine Substrataktivität, die in den bis dato zugänglichen Sequenzdaten nicht aufgezeichnet wurde. Auf der Sequenzhomologie und den katalytischen Resten basierend, wurde zunächst eine  $\beta$ -Glukosidase Aktivität vorhergesagt, die sich nach Experimenten als eine *endo*-aktive  $\beta$ -Glukanase erwiesen hat. Damit konnte demonstriert werden, dass die funktionelle Durchmusterung von Metagenom-Genbanken, im Gegensatz zu sequenzbasierten Methoden, zur Identifizierung vollkommen neuer Biokatalysatoren mit interessanten Eigenschaften führen kann. Bisher ist auch noch nicht bekannt, dass ein Kohlenhydrat-Bindungsmodul von zwei getrennten Aktivitäts-Teilmodulen mit ihren katalytischen Resten flankiert wird. Nach Rücksprache mit Prof. B. Henrissat (CAZy Datenbank) wurde postuliert, dass EngU den Prototyp einer neuen GH-Familie darstellt, die als GH148 bezeichnet wird (Angelov und Pham et al., 2017).

## 5.2 Expressionsanalyse metagenomischer Gene in den Wirtsorganismen *E. coli* und *T. thermophilus*

Die Fosmide HA-CMC-1 und TB-EST-7 wurden in Metgenom-Genbibliotheken durch funktionelle Platten-Screenings entdeckt und in dieser Arbeit in zwei verschiedenen Wirten gegenüber gestellt. Die spezifische CMCCase-Aktivität im hitze-behandelten Zellextrakt-Überstand von *E. coli* EPI300-HA-CMC-1 betrug 2,21 mU/mg, während in *T. thermophilus* BL03-HA-CMC-1 keine Aktivität festgestellt wurde (Abbildung 10). Im Falle des Esterase-kodierenden Fosmids wies umgekehrt der (TB-EST-7-tragende) *T. thermophilus* Stamm eine weit höhere spezifische *p*NP-Capylat (C8)-spaltende Aktivität auf als der entsprechende *E. coli* Stamm EPI300-TB-EST-7. Allerdings war ein Großteil der Aktivität in *T. thermophilus* auf Esterase-Hintergrundaktivität zurückzuführen. Für eine genauere Bestimmung der durch das rekombinante Fosmid vermittelten Aktivität wäre die Verwendung eines Wirtsstamms mit geringerer Hintergrundaktivität wünschenswert gewesen, ein solcher war jedoch nicht verfügbar.

Für die qPCR-Versuche wurden für alle annotierten Gene der Fosmide je ein Primerpaar bestellt und mithilfe einer PCR getestet. Bei HA-CMC-1 bildeten 27 von 45 Primerpaaren ein Produkt und bei TB-EST-7 17 aus insgesamt 33 Primerpaaren. Die ORFs für die ein hypothetisches Protein vorhergesagt wurde, wurden nicht weiter analysiert. Der Fokus lag auf den annotierten über- oder unterexprimierten Genen.

Zunächst werden die einzelnen ORFs von HA und TB relativ zu ihren Kandidatengenen im gleichen Wirt genauer betrachtet. Für die Cellulaseaktivität im HA-CMC-1 Fosmid ist der ORF23 (HA23 genannt) verantwortlich. TB32 ist das im TB-EST-7 Fosmid für die Esterase kodierende Kandidatengen. Die über Real-time PCR ermittelte relative Transkriptionsmenge wird als  $\Delta$ Ct-Wert bezeichnet und ergibt sich aus der Differenz der Werte des Referenz-ORFs (HA23 oder TB32) und des zu vergleichenden ORFx. Ein positiver Differenzwert weist daraufhin, dass der ORFx im Wirt öfter transkribiert wird als das Referenzgen. Ein negativer  $\Delta$ Ct-Wert entsteht, wenn der zu untersuchende ORFx einen höheren Ct-Wert besitzt als das Referenzgen und das bedeutet wiederum insgesamt weniger ORFx-Transkripte in der Zelle als der Referenz-ORF. Die Tabelle 37 zeigt eine Übersicht über die ORF Gene und ihre  $\Delta$ CT-Werte in den jeweiligen Expressionswirten. Bei diesem Versuch wurden Einzelbestimmungen der jeweiligen Gene durchgeführt, da die Ergebnisse als Anhaltspunkt zur Bewertung des Transkriptniveaus der Gene dienen sollte, die exakte Differenz der Transkriptmengen war nicht notwendig.

Tabelle 37: Überblick über ausgewählte ORF Gene und ihre  $\Delta Ct$ -Werte (Referenz-ORF - ORFX) in jeweiligen Expressionswirt (Einzelbestimmungen). *E. coli* mit dem Fosmid HA-CMC-1 (EH), BL03 bzw. *T. thermophilus* mit HA-CMC-1 (BH), *E. coli* mit TB-EST-7 (ET) und BL03 bzw. *T. thermophilus* mit TB-EST-7 (BT).

| ORF  | Bezeichnung   | <i>E. coli</i> | <i>T. thermophilus</i> |
|------|---|----------------|------------------------|
|      |   | EPI300         | BL03                   |
|      |   | $\Delta Ct$    | $\Delta Ct$            |
| HA24 | Na <sup>+</sup> /H <sup>+</sup> subunit               | -0,34          | 3,42                   |
| HA28 | Metal-dependent hydrolase ( $\beta$ -lactamase)       | 6,28           | 10,7                   |
| HA36 | Sodium/hydrogen exchanger                             | 6,43           | 8,76                   |
| HA38 | Na <sup>+</sup> /H <sup>+</sup> subunit D             | -5,01          | 10,7                   |
| HA39 | Multisubunit sodium/hydrogen antiporter, MnhC subunit | 4,59           | 10,27                  |
| HA42 | Na <sup>+</sup> /H <sup>+</sup> subunit G             | 5,79           | 7,24                   |
| HA43 | putative sodium/proton antiporter subunit MrpF        | 6,18           | 7,34                   |
| HA44 | putative antiporter subunit                           | 7,11           | 6,93                   |
| TB13 | putative CsbD-like protein                            | -1,16          | 5,5                    |
| TB16 | Cystathionine gamma-lyase                             | -0,21          | 5,84                   |
| TB17 | Uroporphyrinogen III lyase                            | 1,79           | 7,49                   |

|      |  | <i>E. coli</i> | <i>T. thermophilus</i> |
|------|--|----------------|------------------------|
|      |  | EPI300         | BL03                   |
| ORF  | Bezeichnung  | $\Delta$ Ct    | $\Delta$ Ct            |
| TB18 | Siroheme Synthese/Precorrin-2 oxidase                            | 0,94           | 8,27                   |
| TB21 | 2,3,4,5-Tetrahydropyridine-2,6-dicarboxylate N-Acetyltransferase | -4,97          | 2,46                   |
| TB23 | UDP-glucose dehydrogenase  | 3,35           | 9,44                   |
| TB27 | Cysteine synthase  | -10,26         | -6,72                  |
| TB29 | Thiamine S protein   | 6,59           | 6,77                   |
| TB31 | putative thiol-disulfide oxidoreductase DCC                      | 4,77           | 10,8                   |
| TB33 | $\alpha$ -Glucosidase  | -4,8           | 5,7                    |

### 5.2.1 Vergleich und Interpretation der relativen qPCR Ergebnisse in *E. coli* ( $\Delta$ Ct)

In EPI300 Zellen wurden im Vergleich zum Referenz-ORF23 die Gene HA28, 36, 42, 43 und 44 häufiger und die zwei Gene HA24 und 38 weniger transkribiert. Unter diesen Genen befinden sich insgesamt drei Na<sup>+</sup>/H<sup>+</sup> Antiporter. Die Untereinheit G (HA42) wird mit einem  $\Delta$ Ct-Wert von 5,79 gemessen, während die Untereinheit D (HA38) einen deutlichen negativen Wert von -5,81 aufweist. Die einzelnen Komponenten des Na<sup>+</sup>/H<sup>+</sup>-Komplexes werden anscheinend nicht äquimolar gebildet. Des Weiteren fallen die hohen Transkriptionsmengen der Na<sup>+</sup>/H<sup>+</sup>-Antiporter auf: HA36 (6,43), HA43 (6,18) und HA44 (7,11). Diese Enzyme tauschen schnell Protonen und Natrium-Ionen durch die Membranen aus und sorgen somit für die Homöostase (intrazellulärer pH, Zellvolumen und Ionenkonzentration) in der Zelle. Der Na<sup>+</sup>/H<sup>+</sup>-Antiporter NhaA aus *E. coli* ist in der Literatur detailliert beschrieben. Er wurde überexprimiert, aufgereinigt und in Proteoliposomen funktionell wieder hergestellt. Das Protein ist ein Dimer aus zwei Monomeren, die jeweils aus 12 Transmembran-Segmenten bestehen. In dieser Struktur wurde von Lee et al. eine Salzbrücke zwischen dem essentiellen Aspartat Asp163 und dem konservierten Lysin Lys300 beobachtet, die auch in dem homologen Austauschere NapA gefunden werden kann (Lee et al., 2014). Downstream des *nhaA* Gens befindet sich der positive Regulator *nhaR*, welcher durch die Na<sup>+</sup>-Ionen aktiviert wird. Aus DNA-Mobility Shift Assays wurde die Bindung des NhaR Regulators an der Promotorregion P1 (von -120 bis +14 des Initiationcodons) von *nhaA* deutlich gemacht. (Padan et al., 2001). Möglicherweise war die Natriumionen-Konzentration in den EPI300 ausreichend, um eine erhöhte Transkription der Gene für Na<sup>+</sup>/H<sup>+</sup>-Antiporterkomponenten auf dem HA-CMC-1 Fosmid im Wirt *E. coli* zu begünstigen. Aus der ORF Tabelle von HA-CMC-1 (Tabelle 29) ist der putative HTH-type Transkriptionsregulator MJ0586 aufgelistet (HA18). Durch eine Mutation des Regulators könnte man die Veränderung der Cellulaseaktivität untersuchen und Rückschlüsse ziehen, ob es sich um einen positiven bzw. negativen Regulator für das *endo*-Glukanasegen handelt. So wirkt der Helix-Turn-Helix Transkriptionsaktivator MalT, der den Maltose und Maltodextrine Metabolismus in *E. coli* reguliert. MalT ist ein allosterisches Protein, welches die Proteinexpression positiv auf Maltotriose und negative auf die Proteine, die in der Prozessierung von Maltose involviert sind (MalK, MalY und Aes), reguliert (Schlegel et al., 2006).

Auf dem TB-EST-7 Fosmid wurden die ORFs TB13, 16, 21, 27 und 33 mit einem  $\Delta$ Ct-Wertebereich von -0,21 bis -10,26 gemessen. Die  $\alpha$ -Glukosidase Aktivität im *E. coli*-Wirt wurde mit *para*-Nitrophenol- $\alpha$ -Glukopyranose als Substrat untersucht. Das Ergebnis war negativ und wird durch das niedrige Transkriptionsniveau ( $\Delta$ Ct = -4,8) bestätigt. Aus den qPCR Läufen ergibt sich ein  $\Delta$ Ct-Wert von 3,35. Das ThiaminS Protein und die Thiol-disulfid Oxidoreduktase DCC wurden mit den  $\Delta$ Ct-Werten 6,59 und 4,77 ausgewertet. Das zuletzt genannte Protein

gehört wahrscheinlich zu der DCC-Familie. Diese Familie besteht hauptsächlich aus uncharakterisierten und hypothetischen Proteinen, die das konservierte Motiv DxxCxxC tragen. Weitere Datenanalysen weisen darauf hin, dass die DCC Mitglieder eine weitere funktionelle Domäne brauchen, um ihre physiologische Funktion in ihrer Familie zu entfalten. Auf dem Fosmidinsert von HA-CMC-1 konnten keine zusätzlichen funktionellen Domänen vorhergesagt werden.

### 5.2.2 Vergleich und Interpretation der relativen qPCR Ergebnisse in *T. thermophilus* ( $\Delta Ct$ )

In dem Expressionswirt BL03 wurden alle Gene auf dem HA-CMC-1 Fosmid im Vergleich zum HA23 ORF häufiger transkribiert. Eine Ausnahme gibt es beim ORF HA31, welcher für ein hypothetisches Protein kodiert und somit nicht weiter analysiert wurde. Auffällig sind die ORFs HA28, 36 und 39 mit hohen  $\Delta Ct$ -Werten zwischen 8,76 bis 10,7. Diese ORFs sind annotiert als Gene für eine Metall-abhängige Hydrolase ( $\beta$ -Laktamase), eine  $Na^+/H^+$ -Austauscher- und für eine Antiporter-(MnhC) Untereinheit. Wie bereits oben bei *E. coli* erwähnt sind die Ionen/Protonen-Austauscher essentiell für die Ionen-Homeostase der Zelle. Das hohe Transkriptionslevel des putativen Antiporters (ORF HA39) in BL03 beweist, dass die Transkriptionsmaschinerie die Gensequenz erkennen konnte. Die 3D-Struktur der Metall-abhängigen  $\beta$ -Laktamase aus *T. thermophilus* HB8 wurde von Yamamura et al., 2009 analysiert. Auch in dieser Arbeit wurde eine Metall-abhängige  $\beta$ -Laktamase im thermophilen Expressionswirt gefunden und charakterisiert. Die verstärkte Transkription des Gens für eine putative Metall-abhängige Hydrolase/ $\beta$ -Laktamase (HA28) im thermophilen Expressionswirt könnte wieder eine Folge der optimierten Fosmid-DNA Sequenzerkennung in *T. thermophilus* sein.

Auch beim zweiten Fosmid TB-EST-7 bleiben die  $\Delta Ct$ -Werte positiv im Vergleich zur TB32-Referenz. Ein Sonderfall ist die Cysteine Synthetase (TB27) mit -6,72 und die HA31, die als ein hypothetisches Protein notiert wurde. Eine hohe Transkriptionmenge wird bei der Uroporphyrinogen III Lyase (TB17), Siroheme Synthese/Precorrin-2 Oxidase (TB18), UDP-Glukose Dehydrogenase (TB23) und der  $\alpha$ -Glukosidase (TB33) ermittelt. Die Aktivitäten der letzten beiden Proteine wurden in einem Assay, wie auch schon bei *E. coli* erwähnt, getestet. Dabei wurde eine hohe Hintergrundaktivität der UDP Dehydrogenase in *T. thermophilus*, aber keine Reaktion von der  $\alpha$ -Glukosidase gemessen. Vermutlich werden beide Gene im Wirt transkribiert, jedoch wird die  $\alpha$ -Glukosidase nicht korrekt zum funktionsfähigen Enzym gefaltet. Zhou et al. hat eine bis zu 90 °C hitzeresistente rekombinante  $\alpha$ -Glukosidase TtAG aus *T. thermophilus* (TC11) heterolog in *E. coli* exprimiert und aufgereinigt. Das Protein hat eine molekulare Masse von 184 kDa und besteht aus 59 kDa schweren Untereinheiten. Die



spezifische Aktivität auf *p*NP- $\alpha$ -D-Glucopyranoside (*p*NPG) beträgt 288,96 U/mg bei pH 5 und 90 °C. Die Substitution von Q10 mit Tyrosin führte zu einer zusätzlichen Wasserstoffbrücke und somit zu einer erhöhten Thermostabilität (Zhou et al., 2015). Das Interesse an thermostabilen  $\alpha$ -Glukosidasen mit einer Transglykosylierungsaktivität steigt in der Biotech-Industrie und daher wäre es interessant die  $\alpha$ -Glukosidase (TB33) in *T. thermophilus* unter anderen Bedingungen weiterhin zu untersuchen.

### 5.2.3 Interpretation der $\Delta\Delta$ Ct Ergebnisse zwischen *E. coli* und *T. thermophilus*

Nach der Auswertung der Transkriptionsverhältnisse in den einzelnen Expressionswirten, werden die Ergebnisse von *E. coli* und *T. thermophilus* miteinander verglichen.

In beiden Wirten werden die ORFs HA28, HA36 mehr als der Kandidaten-ORF HA23 transkribiert. Dahingegen wird TB27, eine Cystein Synthase, und HA31, ein hypothetisches Protein, in den Organismen weniger gut gebildet. Die Cystein Synthase gehört zu den Transferasen (EC 2.5.1.47) und katalysiert die Reaktion von O<sub>3</sub>-Acetyl-L-Serin und Schwefelwasserstoff zu L-Cystein und Acetat. Der Cystein Synthase Komplex (CSC) ist ein wichtiger Bestandteil des Schwefelmetabolismus in Bakterien. Dazu werden die zwei Enzyme Serine O-Acetyltransferase (SAT; EC 2.3.1.30) und O-Acetyl-L-Serine Sulfhydrylase (OASS, EC 2.5.1.47) benötigt, welche die letzten Schritte der Cystein Biosynthese katalysieren. Aus L-Serine und Acetyl-CoA bildet die Serine O-Acetyltransferase das Produkt O-Acetyl-Serine (OAS), einen Cystein Vorgänger mit zwei allosterischen Funktionen. CysB ist der transkriptionelle Aktivator des Cystein-Regulons. OAS verstärkt die CysB-Expression und balanciert gleichzeitig die freie Verfügbarkeit der SAT- und OASS-Enzyme, indem es selbst an dem CSC bindet. Die OASS scheint in dem Komplex inaktiv zu sein. Erst bei einer Dissoziation der Sulfhydrylase, werden der OAS-Verbrauch und die Cystein-Synthese stimuliert. Zum Abschluss dieses Regulationzyklus inhibiert das Endprodukt Cystein wiederum die Serine O-Acetyltransferase (Wang and Leyh, 2012). Aufgrund dieses engen Zusammenspiels werden alle CSC-Komponenten benötigt und ihre Transkription strengstens gesteuert. Neuste Erkenntnisse in der transkriptionellen Regulation in *E. coli* haben gezeigt dass die mRNA-Spiegel der SAT- und OASS-Transkripte unter den verschiedensten Wachstumsbedingungen immer sehr nah beieinander liegen (Zare et al., 2009). In dem Metagenom-Insert TB-EST-7 ist nicht das vollständige Cystein-Regulon kodiert, sondern mit ORF TB27 nur ein Bestandteil der Cystein Synthase. Die Abwesenheit der anderen Elemente könnte die Transkription des Gens unterbinden und das könnte wiederum die verringerte Transkriptionsmenge in beiden Wirten erklären.

Im Allgemeinen zeigen die  $\Delta\Delta\text{Ct}$ -Auswertungen, dass mehr Transkripte in *E. coli* im Vergleich zu *T. thermophilus* gebildet wurden, wobei der *E. coli*-Stamm mit dem jeweiligen Kandidatengen als Referenz herangezogen wurde. Die Ausnahmen bilden das hypothetische Protein (HA13), die putative Antiporter-Untereinheit (HA44) und die Mov34/MPN/PAD-1 Familie (TB28). Das bedeutet, dass mehr Transkripte von TB32, die für die Esteraseaktivität zuständig ist, in *E. coli* vorhanden waren. Dennoch gab es keinerlei Aktivität in dem mesophilen Stamm, sondern eher verstärkt in *T. thermophilus*. Das lässt vermuten, dass nicht die Menge der Transkripte, sondern die korrekte Faltung des Proteins ausschlaggebend für die Aktivität ist. Vermutlich konnten die verhältnismäßig wenigeren TB32-mRNAs in *T. thermophilus* besser translatiert und zu einem funktionellen Enzym gefaltet werden.

### 5.3 Biochemische Untersuchungen zu auf den Fosmiden pCT3FK-HA-CMC-1 und pCT3FK-TB-EST-7 kodierten Enzymen

#### 5.3.1 Gerichtete Mutagenese des nukleophilen katalytischen Rests von GH12E, um eine Glykosynthase zu erhalten

Um eine für die Strukturanalyse geeignete Form des katalytischen Moduls von Cel12 zu erhalten, wurde im Rahmen dieser Arbeit die Sequenz von Cel12 gekürzt, indem im Leserahmen die für und die beiden CBM2 Module kodierenden Bereiche entfernt wurden. Das verkleinerte Gen kodiert nur das katalytische GH12 Modul mit insgesamt 272 Aminosäuren. Das resultierende Enzym GH12E hat eine Molekularmasse von 30,6 kDa und war demnach nur etwa halb so groß wie das Ausgangsenzym Cel12 (62,3 kDa). SDS-PAGE-Analyse ergab das Vorliegen einer zusätzlichen Proteinbande bei 40 kDa (Abbildung 17). Das unbekannte Protein könnte die Kristallisierung stören, konnte aber durch Size Exclusion Chromatographie nicht entfernt werden. Aufgereinigte Proteinproben von Cel12 und GH12E wurden nach Schweden an unsere Kooperationspartner an der Universität (Agriculture Science) weitergeleitet, wo derzeit in der Forschungsgruppe von Dr. Jerry Stahlberg an der Kristallisierung der Proteine gearbeitet wird.

Zusätzlich zu dem verkürzten GH12E Enzym wurden drei Mutanten generiert, bei denen der nukleophile katalytische Glutamatrest E171 durch drei verschiedene Aminosäuren substituiert wurde. Die drei neuen Varianten E171A (Alanin), E171S (Serin) und E171G (Glycin) wurden mit dem gleichen Protokoll wie GH12E aufgereinigt und ihre Inaktivierung über einen Standard DNSA-Test überprüft.

Der Schlüssel für eine Glykosynthase-Reaktion ist die Mutation des nukleophilen Restes und die intakte Funktion des Säure/Base-Rests. Der Säure/Base-Katalysator tendiert weniger dazu

das Wassermolekül zu aktivieren, kann aber immer noch die Hydroxylgruppe eines Zuckerrests anregen. Diese Funktion ist wichtig, da ansonsten lediglich eine Hydrolyse des fluorierten Substrates erfolgen würde, nicht jedoch die glykosidische Verknüpfung mit einem weiteren Monomer. Bei einigen Glykosphase-Mutanten bleibt die Aktivität des Wildtyps erhalten, daher wurden auch drei unterschiedliche Mutationen an Position E171 von GH12E erprobt. Tatsächlich wurde bei der Alanin-Mutanten E171A noch eine geringe Hintergrundaktivität von 0,11 % im Vergleich zum Wildtyp nachgewiesen.

Es gibt viele Beispiele von Glykosphase-Mutanten wie z. B. die Endoglukanase I aus dem Pilz *Humicola insolens*, welche zu der GH Familie 7 gehört und Ähnlichkeiten zu EGI/Cel7B aus *Hypocrea jecorina* besitzt. Die Mutanten am Nukleophil des Enzyms von *H. insolens* sind effiziente Glykosphasen und es stellte sich heraus, dass die Serin-Substitution eine höhere Enzymaktivität hervor rief als der Austausch mit Alanin. Ob auch eine Glycin-Mutante ausprobiert wurde, ist nicht bekannt (Lantz et al., 2010; Xu et al., 2015b). Dr. Stahlbergs Gruppe arbeitete an der GH16 Laminarinase Lam16A aus dem Pilz *Phanerochaete chrysosporium*. Dieses Enzym kann jeweils  $\beta$ -1,3- und  $\beta$ -1,4-glykosidische Verbindungen in  $\beta$ -1,3 und gemischten  $\beta$ -1,3/1,4 Glukanen spalten. Bei Experimenten mit  $\alpha$ -Fluor Substraten zeigte sich dabei ebenfalls die Serin-Mutante im Vergleich zu der Glycin-Variante als effizienter (J. Stahlberg, persönliche Mitteilung).

Im Rahmen dieser Arbeit konnte nicht abschließend geklärt werden, welche der Mutanten eine höhere Stabilität oder Aktivität besitzt. Bisher liegen noch keine Ergebnisse von der Gruppe von Dr. Stahlberg bezüglich der Aktivität der GH12E-Mutanten mit Glycosylfluoride vor. Jedoch wird nach bisherigen anderweitigen Erfahrungen vermutet, dass die E172G-Mutante die geringste Aktivität aufweisen wird. Denn ein Merkmal des katalytischen Zentrums ist eine Methionin-Seitenkette in der Nähe des nukleophilen Rests. Wenn die ausgetauschte Aminosäure zu klein ist, könnte die Methionin-Seitenkette in die Vertiefung eintreten und eine unkontrollierte Enzymaktivität erzeugen. Das Risiko ist bei der kleinen Aminosäure Glycin am höchsten, andererseits könnte die größere Aminosäure Serin das  $\alpha$ -Fluor Atom sterisch hindern. Leider kann dies für GH12E erst nach Vorliegen entsprechender Messungen unserer Kooperationspartner weiter beurteilt werden.

### 5.3.2 Produktion der thermostabilen Esterase Est7 und ihre Substratspezifität

Die Expression des Est7-Leserahmens in *E. coli* mittels Expressionsvektor pET21a ergab die besten Ergebnisse für die Proteinausbeute. Im Rohextrakt wurde über SDS-PAGE eine starke Bande mit einer molekularen Masse von 36 kDa nachgewiesen und das rekombinante Protein konnte über Nickel-Affinitätschromatographie gereinigt werden. Mit dem pMKE2\_Est7 Vektor konnte im thermophilen Wirt *T. thermophilus* BL03 weder nach 3 Stunden noch nach 12 Stunden die Synthese der Esterase im SDS-Gel nachgewiesen werden, weiterhin wurden bei der Aktivitätsmessung geringere Aktivitäten als in der Kontrolle mit Leervektor gefunden. Dies weist darauf hin, dass die Induktion mit Kaliumnitrat unter Sauerstoffabschluss nicht erfolgreich war. Mangelnde Expression war außerdem bei pMK18\_Est7 in BL03 beobachtet worden. Zum Teil waren die Substrate bereits nach dem Wiederauflösen mit DMSO schon so stark gelblich verfärbt, dass der Messbereich des Photometers überschritten wurde.

Die Spaltungsspezifität von Est7 wurde zusätzlich noch mit synthetischen Lipasesubstraten untersucht. Diese speziellen Substanzen sind *p*NP-Derivate von Arzneimitteln. Weitere getestete Proteine waren die thermostabilen Esterasen EstA2 und EstB1, die mehrere *p*NP-Derivate hydrolysieren konnten (Leis et al., 2015a). Bei einer Inkubationstemperatur von 80 °C erreichte EstA2 eine spezifische Aktivität von 3,09 mU/mg, während bei Est7 nur 0,054 mU/mg gemessen werden konnte.

Nahezu 80 % der synthetischen chiralen Arzneistoffe werden als Racemate (S- und R-Enantiomere) im Handel angeboten. Das kann gravierende Folgen haben, da es pharmakodynamische und -kinetische Differenzen zwischen den beiden Formen gibt, die sich beispielsweise auf die unterschiedlichen Wechselwirkungen mit Proteinen, Ionenkanälen oder Rezeptoren zurückführen lassen. Das bekannteste Beispiel ist das Arzneimittel Thalidomid, auch bekannt unter dem Markennamen Contergan. Dem (R)-Enantiomer wird eine schlaffördernde, entzündungs- und tumorwachstumshemmende Wirkung zugesprochen. Bei schwangeren Frauen, die in den ersten drei Monaten das Medikament eingenommen haben, führte es zu vermehrten Fehlgeburten oder Fehlbildungen. Die Ursache ist das (S)-Enantiomer von Thalidomid, welches an das teratogene Protein Cereblon bindet und das Wachsen von Gliedmaßen in der embryonalen Entwicklung hemmt. Aus dem Grund ist die Enantiomerenanalytik für die Pharmaindustrie von großer Bedeutung. Dazu muss zunächst der Enantiomerenüberschuss (engl.: *enantiomer excess*, *ee*) berechnet werden, der sich aus folgender Formel ergibt (E1 = Enantiomer 1; E2 = Enantiomer 2):

$$ee [\%] = \frac{(E1 - E2)}{(E1 + E2)} \times 100$$

Bei einem Wert von 100 % ist die Substanz enantiomerrein, bei 0 % liegt es als Racemat (50:50) vor und bei beispielsweise 90 % liegt ein Gemisch aus 95:5 Anteilen der beiden Enantiomere vor. Eine Methode für die direkte Trennung ist eine HPLC mit einer chiralen stationären Phase (CSP's). Die CSP interagiert durch Wasserstoffbrücken,  $\pi$ - $\pi$ -Wechselwirkungen und Dipol-Dipol-Wechselwirkungen unterschiedlich stark mit dem jeweiligen Enantiomer, dadurch werden diese zu verschiedenen Zeitpunkten getrennt eluiert. Ibuprofen wird häufig zur Behandlung von Schmerzen und Entzündungen eingenommen. Das (R)-(-)-Enantiomer ist als Diastomer inaktiv, invertiert aber zum aktiven Eutomer Dexibuprofen ((S)-(+)-Enantiomer). So lagen auch das verwendete Ibuprofensubstrat und die anderen 4 Lipasesubstrate (2-Methyldecansäure, Indancarbonsäure, Naproxen und Mulzer) als Racemat vor und die reinen Enantiomere konnten nicht zur Verfügung gestellt werden. Daher konnte die Enantioselektivität und Stereoselektivität von Protein Est7 nicht geklärt werden.

## 5.4 Durchmusterung metagenomischer Genbanken

In dieser Doktorarbeit lag das Hauptaugenmerk auf thermostabilen Enzymen, welche nützliche Eigenschaften für biotechnologische Prozesse aufweisen. Die Durchführung biotechnologischer Industrieprozesse bei hohen Temperaturen führen zu weniger Kontaminationen mit mesophilen Mikroorganismen, zu einer besseren Bioverfügbarkeit und Löslichkeit der organischen Komponenten und zu einer höheren Reaktionsrate aufgrund der sinkenden Viskosität (Becker, 1997; Krahe et al., 1996; Kumar and Swati, 2001). Aus diesen Gründen ist die Entdeckung und Etablierung von neuen hitzeresistenten Proteinen ein interessantes und lukratives Forschungsgebiet. Im praktischen Teil der Dissertation wurden verschiedene Metagenom-Genbanken nach thermostabilen Amylasen, Cellulasen, Esterasen/Lipasen, Proteasen und Xylanasen durchforstet, die eine mögliche Anwendung in den unterschiedlichen Industriezweigen finden könnten. Nachfolgend werden ein paar Beispiele für Industrie-relevante Proteine aus Metagenom-Genbibliotheken in Verbindung mit den Proteinen, die im Rahmen dieser Doktorarbeit entdeckt wurden, genannt (Tabelle 1 und Tabelle 2).

In der Stärke-Industrie werden für den Stärkeabbau verschiedene *endo*- und *exo*-Enzyme benötigt. Die  $\alpha$ -Amylase ist beispielsweise ein *endo*-prozessierendes Enzym, spaltet an zufälligen Stellen im Polymer und bildet somit lineare oder verzweigte Oligosaccharide, die wiederum von *exo*-Enzymen am nicht-reduzierten Ende angegriffen werden und zu Oligo- oder Monosacchariden degradiert werden. Der Abbau beginnt mit der Gelatinierung der Stärke, welche die Auflösung, Verflüssigung und Verzuckerung des Substrates beinhaltet und durch das Erhitzen mit Wasser erreicht wird. Damit die Hydrolyse gleich nach der Gelatinierung anknüpfen kann und keine weitere Abkühlung erforderlich ist, werden thermostabile Amylasen oder Pullanasen verwendet (Ferrer et al., 2012). In dieser Arbeit wurden Metagenom-DNA tragende Organismen auf Stärkeplatten gestempelt, bebrütet und vor der Phänotyp-Analyse von der Platte mit Ethanol und Wasser abgewaschen. Dadruch soll der Stärkeabbau nach der Färbung mit Lugolscher Lösung besser visualisiert werden. Jedoch wurde in keiner der hier durchforsteten Metagenombanken ein Stärke-abbauender Klon entdeckt. Mit einer etwas abgewandelten Screeningmethode haben (Yun et al., 2005) eine Amylase AmyM aus einer Boden-Metagenombank entdeckt. Nach einer eintägigen Inkubation (LB-Agar mit Antibiotikum und 2%iger löslicher Stärke) wurden die Kolonien durch eine Topagarschicht mit D-Cycloserine (60  $\mu\text{g/mL}$ ) überdeckt, welches die Zellen permeabilisieren und somit die intrazellulären Enzyme freisetzen sollte. Die Platte wurde nochmal für 24 Stunden bei 37 °C inkubiert und danach mit einer Gram's Iodlösung (0,203 g I<sub>2</sub>, 5,2 g KI in 100 mL Wasser) eingefärbt und unter fluoreszierendem Licht betrachtet. Bei weiteren funktionellen Screenings

nach Amylasen könnte erprobt werden, ob dieses Verfahren empfindlicher ist als die übliche Methodik.

Die komplexe, schwer zersetzbare Cellulose taucht in vielen Herstellungsprozessen auf. Für ihren Abbau benötigt man Cellulasen, die sowohl in den groben Abbauschritten wie z.B. bei der Vorbehandlung von Futtermitteln oder bei der Müllverwertung, als auch bei den feinen Aufarbeitungsvorgängen der Textil- und Medikamentenherstellung Verwendungen finden (Ando et al. 2002; Nakamura et al., 2001; Xu et al., 2015a; Zhou et al. 2001). Die Produktion von Cellulasen ist teuer, deshalb stellt der Einsatz dieser Enzyme einen erheblichen Kostenfaktor in einer Produktionskette dar. Bei der Ethanolherstellung durch Biomasse beträgt die Bereitstellung des Biokatalysators bereits 40 % der Gesamtkosten und erhöht den Verkaufspreis des Endprodukts. Deshalb können Biokraftstoffe oft nicht mit den günstigeren fossilen Brennstoffen konkurrieren und verlieren dadurch an Attraktivität am Markt (Dwidar et al., 2012; Yeoman, 2015). Im Rahmen dieser Dissertation wurde in Kooperation mit der Auburn University in Alabama (USA) und der NTNU in Trondheim (Norwegen) mithilfe des Sequenz-basierten oder funktionellen Screenings von Metagenombanken nach neuen Enzymen gesucht. Der norwegische Kooperationspartner stellte dafür eine Metagenombank aus einem Oil Reservoir zur Verfügung. Die Proben stammten aus einer Tiefe von etwa 2,5 km unterhalb der norwegischen Seeplatte, wo Temperaturen von 85 °C und ein Druck von 250 bar herrschen. Die Metagenombank (pCC1FOS-basiert) dieser Proben sollten im Labor von Herrn Prof. Dr. Liles durchforstet werden. Mithilfe des funktionellen Cellulase-Screenings (Material und Methoden 3.2.5.3), konnte eine thermostabile Cellulase F1 detektiert werden. Das Protein besitzt zwei markante Cellulase Domänen, die auf eine Fusion von zwei archaellen Cellulasen schließen lässt. Solch eine Struktur ist bis jetzt noch nicht bekannt und könnte für die starke Aktivität und verbesserte Thermostabilität des entdeckten Enzyms verantwortlich sein. Diese Ergebnisse wurden kürzlich publiziert. (Lewin et al., 2017).

Eine Enzymklasse mit vielseitigen Biokonversionsmöglichkeiten (Acido-, Alkohol- und Aminolyse, Veresterungen) und einer oft hohen Substratspezifität (Stereospezifität, Regioselektivität) stellen die Esterasen und Lipasen dar (Arpigny and Jaeger, 1999; Jaeger and Reetz, 1998) Ihre Produkte finden Einsatz als Geschmacksaromen, Zusatzstoffe in Kosmetikartikeln oder als Biokraftstoffe. Darüber hinaus sind Esterasen/Lipasen nützlich, um Reststoffe oder Verunreinigungen in der Milch-, Medikamenten-, und Lederproduktion zu entfernen (Linko et al., 1994; Pandey et al., 1999; Traore, M.K. and Buschle-Diller, 2000). Im funktionellen Screening wurde aus einer 384 Mikrotiterplatte der HA- und TB-Metagenom-Genbibliothek bei 60 °C ein auf Tributyrin aktiver Klon entdeckt. Die Aktivität wurde nach der Retransformation bestätigt (Abbildung 26). Jedoch konnte das Lipasegen nach Erstellung der Shotgun-Genbibliothek nicht identifiziert werden. Mögliche Gründe hierfür könnte eine unvollständige Shotgun-Genbibliothek sein, die nicht das Tributyrin abdeckt und somit kein

aktives Protein erzeugen konnte. Des Weiteren könnte es durch die Erstellung der Shotgun-Bibliothek zu einer Verminderung der Lipaseaktivität kommen, die durch das herkömmliche funktionelle Screeningverfahren nicht entdeckt werden kann.

Proteasen werden in Exopeptidasen, die Aminosäuren am Ende von Peptiden abspalten, und Endopeptidasen, die Peptidbindungen innerhalb von Proteinen hydrolysieren, eingeteilt. Sie werden vielfältig, z.B. in der Pharma-, Lebensmittel-, Textil- und Lederindustrie eingesetzt. Im nächsten Kapitel wird ein in dieser Arbeit entdecktes Isolat „Mexico SH-Avicel“ beschrieben, welches starke Proteaseaktivität aufweist. Aus den Metagenombanken jedoch wurde weder in *E. coli* noch in *T. thermophilus* ein aktiver Klon mit Proteaseaktivität entdeckt.



#### 5.4.1 Metagenom Isolierungsstrategien

Die Metagenomproben, die für diese Arbeit relevant waren, stammen alle aus heißen Habitaten mit einer Temperatur von ca. 50 - 90 °C. Teilweise wurden die Metagenom-Genbibliotheken von Kooperationspartnern bereitgestellt, aber es wurden auch am Lehrstuhl für Mikrobiologie Umweltproben selbst aufgereinigt und deren DNA zu Metagenom-Genbibliotheken verarbeitet. Dabei ist von großer Bedeutung, dass die große Breite an Mikroorganismen in der Umweltprobe bestehen bleibt, damit die ursprüngliche mikrobielle Population in der extrahierten DNA repräsentiert ist. Zudem sollte das metagenomische Erbgut nicht zu stark fragmentiert werden. Die hoch molekulare DNA wird benötigt, um möglichst auch Aussagen über die Phylogenie der Organismen machen zu können und um zu vermeiden, dass chimäre Produkte entstehen können. Zuletzt sollte darauf geachtet werden, dass keine Kontaminationen in die Metagenom-DNA Probe eingebracht werden (Schmeisser et al., 2007). Auf dem Markt werden Kits für die unterschiedlichsten Umweltproben angeboten, dennoch entwickeln einige Arbeitsgruppen ihre eigenen Aufreinigungsprotokolle, um potentielle selektierende Reagenzien oder Arbeitsschritte zu vermeiden, die zu einer Verfälschung der Ergebnisse führen könnten (Frostegård et al., 1999; Miller et al., 1999). In dieser Arbeit wurde das PowerSoil® DNA Isolation Kit von der Firma Mo Bio benutzt, um aus den beiden Umweltproben „Russia“ und „Mexico“ DNA aufzureinigen. Auf dem Agarose-Gel zeigten beide isolierte Proben das Vorliegen von hochmolekularer DNA von ca. 20 kb (Abbildung 21). Die DNA musste nicht mehr zusätzlich geschert werden, sondern wurde direkt zur Herstellung von Metagenom-Genbibliotheken im Vektor pCC1FOS verwendet.

#### 5.4.2 Interpretation der 16s rDNA Analyse der „Russia“- und „Mexico“-Umweltproben und der Isolate

Für die 16S rDNA-Analyse der Bakterien in den Umweltproben und Anreicherungskulturen wurden 16 rDNA-Sequenzen über eine PCR mit universellen 16S rDNA Primern 100K und 616v\_alt amplifiziert. Aus jeder Probe wurden nur 23 Klone Analyse selektiert. Das ist eine zu geringe Anzahl, um die gesamte Diversität in den Proben gut abzudecken. Dennoch sollte dadurch eine grobe Abschätzung der Breite der bakteriellen Diversität in den Umweltproben erfolgen. In der „Russia“-Umweltprobe konnten von den 23 16S rDNA-Genen über BLAST-Analyse 21 Zuordnungen zu unterschiedlichen Bakterien gemacht werden. Das weist auf eine hohe Diversität der Metagenom-Genbank hin. Die interessanten thermophilen Bakterien darunter sind *Exilispa thermophila* (37 - 60 °C), *Caldalkalibacillus thermarum* (60 °C) und *Thermosediminibacter oceani* (52 - 76 °C), deren Genome für potentielle hitzestabile Proteine kodieren könnten (Imachi et al., 2008; Lee et al., 2005; Xue et al., 2006).

Aus der Sequenzanalyse aus der „Mexico“-Probe hingegen ergaben sich insgesamt nur 10 unterschiedliche Arten aus 23 Sequenzen, da die Stämme *Thiobacillus aquaesulis*, *Oxalicibacterium flavuum* und *Thiofaba tepidiphila* häufiger auftraten. Die drei zuletzt genannten Bakterien sind jedoch nicht thermophil, im Gegensatz zu *Thermodesulfovibrio islandicus* (45 - 70 °C), *Thermosulfovibrio hydrogeniphilus* (50 - 80 °C) und *Rhodothermus marinus* (54 - 77 °C) (Bjornsdottir et al., 2007; Jeanthon et al., 2002; Sonne-hansen and Westermann, 1999), die ebenfalls in der Mexico-Probe identifiziert wurden.

Die 16S rDNA Analyse der Anreicherungskulturen auf den Minimalmedien mit CMC oder Xylan mittels BLAST ergaben Treffer bei BLAST, die alle einen hohen Sequenzähnlichkeitswert (99 %) aufwiesen. Zusätzlich wurden diese auf Substratplatten (Tributyryn, Triolein, Stärke, Skim Milk, CMC, Aviel und Xylan) ausgestrichen und bei 50 - 60 °C inkubiert. Ein Wachstum auf den jeweiligen Platten kann auf Aktivität gegenüber den verwendeten Substraten zurückzuführen sein. Die Isolate wurden auf Substrataktivitäten getestet und weisen ein interessantes Spektrum auf (Tabelle 32). Nachfolgend werden die gefundenen Aktivitäten der Anreicherungen den literaturbekannten Eigenschaften der damit über 16S rRNA-Amplikonanalyse verknüpften Bakterienarten gegenübergestellt:

*Bacillus smithii* wurde aus Kompostproben isoliert und kann bei 55 - 65 °C aus Xylose und Glukose organische Säuren bilden. Der Stamm ist mit dem Plasmid pNW334 transformierbar und es wurde auch bereits eine Methode für eine Marker-freie Deletion entwickelt (Bosma et al., 2015a, 2015b). In dieser Arbeit zeigt das Isolat „Russia SH- Avicel“ auf keinerlei der getesteten Substraten Aktivitäten auf.

*Brevibacillus borstelensis* (Accession: AY764129) ist ein Gram-positives, thermophiles (50 °C Temperaturoptimum) und Sporen-bildendes Bakterium. Ihm wurde ein hoher Verwandtschaftsgrad mit *Bacillus centrosporus* und *Bacillus choshinensis* nachgewiesen. Besonders interessant ist der Stamm, da er verzweigte Hochdruckpolyethylene degradieren (Hadad et al., 2005) und industrierelevante Amylasen, Proteasen und Amidasen produzieren kann (Tripathy et al., 2016). Diese Aussage deckt sich auch zum Teil mit den hier generierten Ergebnissen. Auf den Agarplatten mit Skim Milk wurde bei 60 °C ein sehr starker Proteinabbau der Kultur „Mexico SH-CMC“ verzeichnet, da die durch das Milchpulver trübe Platte zu 85 % geklärt wurde.

Ein ebenso thermophiles Gram-positives Bakterium, das Sporen bilden kann, ist *Anoxybacillus flavithermus* (Berendsen et al., 2016). 2015 wurde eine Acetyl Xylan Esterase aus diesem Organismus isoliert, dessen Temperaturoptimum bei 50 °C und pH-Optimum bei 7,5 liegt (Eminoğlu et al., 2015). In dieser Arbeit wurde bei dem Isolat „Russia SH-Avicel“ eine schwache Tributyrin und eine Amylase- und Avicel Aktivität detektiert.

Der Gram-positive, aerobe Endosporen-Bildner *Bacillus sonorensis* wurde zum ersten Mal aus der Wüste in Sonora isoliert und ist nahe mit *Bacillus licheniformis* verwandt. Aus *B. sonorensis* werden häufig Lipasen gefunden, wie die hyperthermostabile alkaline Lipase, die bei 80 - 120 °C noch aktiv ist (Bhosale et al., 2016). Das Substratspektrum der in dieser Arbeit gefundenen Isolate „Mexico SH-Avicel“ zeigt jedoch ein viel breiteres Spektrum auf, denn alle Platten, außer die mit Triolein, konnten von dem Bakterium bewachsen und genutzt werden.

Aus dem letzten Abschnitt kann man herauslesen, wie vielversprechend die neu entdeckten Isolate für die Industrie werden können.

### 5.4.3 Funktionelle Screeningmethoden

Mithilfe von DNA-Sequenzen können Bakterienstämme identifiziert und der Verwandtschaftsgrad zueinander bestimmt werden. Durch Sequenzierungen können auch Gene für bestimmte Enzyme identifiziert werden, jedoch mit der Einschränkung, dass nur Proteine mit Ähnlichkeit zu bereits annotierten Sequenzen gefunden werden. Dagegen können durch ein funktionelles Durchforsten einer Metagenom-Genbank mit verschiedenen Substraten und unter verschiedenen Bedingungen auch wirklich neuartige industriellrelevante Enzyme entdeckt werden.

Die funktionsbasierte Durchmusterung kann in drei Kategorien eingeteilt werden (Simon and Daniel, 2009b). Die erste wird als phänotypische Insert-Detektion bezeichnet, die sich wiederum in die direkte und indirekte Detektion untergliedert. Bei der direkten Methode (1a) erkennt man durch die veränderte Morphologie, Pigmentierung oder Hofbildung des Klons, die direkt durch das heterolog gebildete Produkt des positiven Klons gebildet wird. Diese Methode ist allerdings nicht besonders sensitiv. Bei der indirekten Variante (1b) geht das gesuchte Genprodukt eine Reaktion mit einem Indikator, der sich beispielsweise im Medium befindet, ein. Die Indikatoren können Substrate sein, deren Abbau auf der Platte sichtbar gemacht wird (Craig et al., 2010). Die heterologen Proteine werden intrazellulär gebildet und sind nicht in der Lage durch die Zellwand des Wirtes zum Indikator zu gelangen. Daher muss eine Zellyse (1c) durchgeführt und darauf geachtet werden, dass das erwünschte Protein nicht denaturiert (Li et al., 2007). Die zweite Kategorie ist die Modulare Detektion (2), bei der beispielsweise Reportergene wie *lacZ* verwendet werden. Die metagenomischen Proteine reagieren mit dem Quorum Sensing (QS) Molekül, welches als Trigger für das *lacZ* Operon benötigt wird. Durch die Inhibition kann X-Gal nicht gespalten werden und die positiven Klone erscheinen weiß (Schipper et al., 2009). Wenn das Genprodukt dazu beiträgt, dass der Wirt unter selektiven Bedingungen wachsen kann, wird dies als eine heterologe Komplementation bezeichnet (Leis 2015). Die letzte der drei Kategorien des funktionellen Screenings ist die Substrat-induzierte Variante (engl: *substrate-induced gene expression*, SIGEX; 3). Hier wird die metagenomische DNA mit „operon-trap GFR“-Expressionsvektor kloniert. Bei einer erfolgreichen Expression wird GFP auch gebildet und die Zelle wird über FACS (engl.: *fluorescence-assisted cell sorting*) aussortiert (Uchiyama and Miyazaki, 2010).

In dieser Arbeit wurde hauptsächlich die erste Methode (1a) angewandt. Die Hofbildung war bei den Platten mit Magermilchpulver durch die Erhellung des bräunlichen Agar leicht zu erkennen. Die Xylan-, CMC-, Avicel- sowie Stärkeaktivität konnte man erst nach einer Färbung mit Kongorot oder Lugolscher Lösung feststellen. Insgesamt ist diese Vorgehensweise kostengünstig, da die verwendeten Substrate und ihre Färbemittel nicht teuer sind und man die Metagenom-DNA nicht in speziell modifizierte Expressionswirte übertragen muss. Die

Sensitivität kann man erhöhen, indem man, wie in Methode 1c beschrieben, die Zellen aufschließt ohne die intrazellulären Proteine zu denaturieren. Dies wäre eine Optimierungsmöglichkeit für die weiteren funktionellen Screenings der Metagenombanken.

Der Transfer von theoretischen Ansätzen über zu den praktischen Arbeiten ist oftmals eine Herausforderung im Labor, denn die Umsetzung der zum Teil manuellen, zum Teil automatisierten Schritte des funktionellen Screenings ist arbeits- und zeitaufwendig. Daher ist es sehr wichtig die Arbeitsschritte zu optimieren, um die Effizienz der Suche nach neuen Genprodukten zu steigern.

#### 5.4.4 Durchmusterungseffizienz

Die Verwendung von Filterplatten sollte das vergleichende Screening in *E. coli* und *T. thermophilus* verbessern und eine Zeitersparnis bringen. Dazu wurden vier verschiedene Fosmid-DNA Aufreinigungsprotokolle im 96-Well Mikrotiterplatten-Format durchgeführt und die Transformationseffizienz gemessen. Die niedrigste Transformationseffizienz von 9 % zeigt dabei Experiment III (Ergebnisse 4.4.3.1; Abbildung 23), in dem ein DNA-Bindefilter eingesetzt wurde. Die zweitniedrigste Effizienz von unter 3 % wird bei der herkömmlichen Fosmid-Extraktion durch alkalische Lyse erreicht. Die beiden Methoden haben als gemeinsamen Schritt die Lysat-Aufreinigung über Zentrifugation. Hierbei bleiben störende Faktoren, wie zum Beispiel RNA, Proteine und andere Zellbestandteile in der DNA-Lösung enthalten und verschlechtern den Transformationserfolg in *T. thermophilus*. Die erfolgversprechendste Extraktionsmethode stellte in diesem Experiment die Verwendung des Lysatfilters in Experiment II dar. Es wurde DNA in relativ großer Menge isoliert und Zelltrümmer und Proteine effektiv von der DNA-Lösung abgetrennt. Lediglich RNA wurde mit aufgereinigt, die aber vermutlich keinen oder nur geringen Einfluss auf die Transformation in *T. thermophilus* hat. Die Verwendung des DNA-Bindefilters in Experiment III und IV lieferte sehr reine DNA, aber nur in geringen Mengen, die für eine effiziente Transformation nicht ausreichen. Zusammenfassend führte die Verwendung des Lysatfilters zu einer starken Erhöhung der Transformationseffizienz und einem geringeren zeitlichen Aufwand bei der Fosmid-Extraktion. Allerdings steigt der finanzielle Aufwand, den diese Methode mit sich bringt, stark an, so dass zur Kostenersparnis die Nutzung der herkömmlichen Methode der alkalischen Lyse sinnvoll ist.

Eine enorme Kosten- und Zeitersparnis kann auch mit dem Pool-Screening erreicht werden. Hierbei werden bis zu 48 Klone in einem Reaktionsgefäß zusammengeführt und gleichzeitig durchmustert. Mit den verschiedenen Durchführungsprotokollen beim Pool-Screening wurden drei Fragen geklärt: 1. Kann ein aktiver Klon aus einem Pool mit 48 Einzelfosmiden identifiziert werden? 2. Ist die Zugabe der Auto Induction Solutionen (AIS) für das funktionelle Screening notwendig? 3. Wie gut funktioniert die Transformation des Pools im Vergleich zu der des Einzelklons?

Die Ergebnisse zeigen, dass der Xylanase-aktive Klon B8-M12-22 aus der Kompost-Metagenom-Genbank beim Durchforsten des ganzen Pools identifiziert werden konnte. Sowohl im Einzelausstrich als auch beim Ausplattieren waren die hellen Höfe nach der Kongorot-Färbung, die auf einen Xylanabbau hinweisen, gut ersichtlich. Die Proben, die vorher mit AIS behandelt worden waren, zeigten auf den Platten mehr Kolonien und auch bei den Aktivitäts-Assays deutlichere Unterschiede zu den nicht aktiven Klonen. Für die Transformation wurde der Pool sowie der Einzelklon nach der *in vivo* Modifizierung in 5 mL

Medium mit Antibiotika-Zusatz und AIS über Nacht kultiviert und die Fosmide isoliert. Um die Diversität der gemischten Fosmide im Pool zu wahren wurde der gesamte aufgereinigte Ansatz zu *T. thermophilus* gegeben, parallel dazu wurde *T. thermophilus* mit aufgereinigter Fosmid-DNA aus dem B8-Einzelklon transformiert. In beiden Fällen wuchsen ca. 30 Klone, was bedeutet, dass bei der Pool-Transformation nicht alle Fosmide in *T. thermophilus* eingebracht wurden. Für eine bessere Aussage müssen mehrere Pools getestet werden und eine statistische Analyse der Diversität der Fosmide im Pool nach der *in vivo* Modifikation und nach der Transformation in *T. thermophilus* erstellt werden. Zusammenfassend hatte die Pool-Screening Methode keine negativen Auswirkungen auf die *in vivo* Modifizierungsprozedur und führte zu aktiven Klonen obwohl die Transformationseffizienz noch weiter verbessert werden muss.

Neben dem Pool-Screening gibt es für die Durchmusterung von Metagenom-Genbanken, hoch entwickelte Technologien und Geräte, die einen sehr hohen Durchsatz an Proben erlauben. Gong *et al.* verwendete für das Screening einer Kuhmagen-Metagenom-Bibliothek den Koloniepickenden Roboter Genetic QPix mit 22 cm x 22 cm großen Platten, wodurch bis zu 6912 Klone durchforstet werden konnten (Gong *et al.*, 2012). Pham *et al.* hat ein Protokoll entwickelt, indem es möglich ist 10.000 Klone auf eine Nylonfiltermembran mithilfe eines Roboters zu übertragen und die DNA durch mehrere Aufreinigungs- und Fixierungsschritte an die Membran zu binden. Die Membran wird mit Proben hybridisiert und später in einem Fluoreszenz-Scanner analysiert. Dadurch können im Hochdurchsatz-Verfahren große Fosmidbanken nach Genen durchforstet werden (Pham *et al.*, 2007). Trotz enormer Präzision und hohen Durchforstungsraten sind solche Technologien nur für wenige Labore rentabel, deshalb ist es weiterhin wichtig vereinfachte Protokolle zu entwickeln.

Die Trefferquote beim funktionellen Screening ist je nach Methode und gesuchtem Enzym sehr gering im Vergleich zu der Sequenz-basierter Suche. Um die Wahrscheinlichkeit zu erhöhen ein neues Enzym zu finden, wird es auch künftig wichtig sein, Methoden und Protokolle zu entwickeln, die einen hohen Durchsatz bei der Durchmusterung der Metagenom-Genbibliotheken ermöglichen. Der Paralleleinsatz verschiedener Wirte wie in dieser Arbeit bietet den Vorteil, sich nicht alleine auf den Genexpressionsapparat eines Wirtes bei der heterologen Expression metagenomischer Gene verlassen zu müssen. Im folgenden Abschnitt wird darauf noch weiter eingegangen. Damit der Mehraufwand beim Einsatz zusätzlicher Screeningwirte beherrschbar bleibt, können Methoden wie das hier beschriebene Pool-Screening hilfreich sein.

#### 5.4.5 Entwicklung und Evaluierung neuer Expressionswirte

In den meisten metagenomischen Studien wird aufgrund der einfachen Handhabung im Labor und der umfangreichen genetischen Methoden, bevorzugt mit *E. coli* gearbeitet. Jedoch werden viele Gene aufgrund der fehlenden Passung zum Expressionsapparat von *E. coli* in diesem Wirt nicht exprimiert und können somit nicht identifiziert und charakterisiert werden. Diese Einschränkung wird umgangen, indem man andere Expressionswirte bei der Durchforstung von Metagenom-Genbibliotheken zum Einsatz bringt. Diese können beispielsweise eine erweiterte Erkennung metagenomischer Promotoren erlauben, eine für bestimmte Gene passendere Codon-Usage verwenden oder Helferproteine besitzen, die der inkorrekten Proteinfaltung eines Enzyms vorbeugen können. Diese neuen Expressionswirte müssen charakterisiert und der Nutzen gegenüber *E. coli* muss validiert werden. Darüber hinaus müssen Protokoll und Systeme entwickelt werden, die es ermöglichen Metagenom-Genbanken in die neuen Wirte zu transportieren. Ein weiterer Fortschritt wäre die Erstellung einer broad-host Metagenom-Genbibliothek, die in mehrere Wirte gleichzeitig eingebracht werden kann und das parallel funktionelle Screenings ermöglicht.

In dieser Arbeit wurde jeweils ein Vektor für *B. subtilis* (pRetro-B1) und für *M. luteus* (pRetro-M) konstruiert, die nach dem gleichen Prinzip wie das früher für *T. thermophilus* konstruierte pRetro-T2 Plasmid funktionieren sollten. Dazu wurden eine Kanamycinresistenzkassette, *ori* und *loxP*-Seite und eine 1,5 kb lange chromosomale DNA-Sequenz aus dem jeweiligen Expressionswirt via Gibson assembliert. pRetro-M konnte nicht in EPI300 Zellen transformiert werden, vermutlich wegen nicht geklärter toxischer Effekte, die das Wachstum von *E. coli* unterdrückten (Warren et al., 2008). Auch die Verwendung verschiedener Abschnitte des *trpE*-Gens als alternative chromosomale DNA-Sequenzen wurden keine Kolonien erhalten. Da kein Vektor für *M. luteus* hergestellt werden konnte, wurde nur mit pRetro-B1 weiter gearbeitet. Dieses Plasmid konnte erfolgreich in EPI300 Zellen, die den Leervektor pCC1FOS tragen, übertragen und mithilfe von pJM2545 *in vivo* modifiziert werden. Jedoch konnten Fosmide mit einem metagenomischen Insert (beispielhaft erprobt mit pCC1FOS-HA-CMC-1 und pCC1FOS-TB-EST-7) nicht mit dem *Bacillus* spezifischen Vektor pRetro-B1 modifiziert werden. Die Selektion von Klonen und die Propagation des Plasmids wurde zwar mittels Kanamycin erreicht, jedoch entsprach das Restriktionsmuster des pRetro-B1-Vektors nicht den Erwartungen. Die Klone sind auf Kanamycin gewachsen, was auf eine Aufnahme des pRetro-B1 Vektors hinweist. Der Kontrollverdau zeigt jedoch keine Bande mit der passenden Größe von pRetro-B1 (Daten nicht gezeigt).

Neben diesen zwei natürlich kompetenten Bakterien und dem bereits genannten *T. thermophilus* könnten sich diverse andere bakterielle Wirte wegen ihrer besonderen Eigenschaften für funktionelle Screenings eignen. Der Vorteil bei *Pseudomonas antartica* wäre



sein niedriges Wachstum-Temperatur optimum von 22 °C (Reddy et al., 2004). Ein passender Wirt für die Expression von membran-gebundenen Dehydrogenasen ist *Gluconobacter oxydans*. Es existieren bereits multiple Gen-Deletionsstämme, die sich hervorragend für funktionelle Screenings nach metagenomischen Dehydrogenasen einsetzen ließen (Liebl et al., 2014a; Peters et al., 2013). Ein ebenfalls gut geeigneter Wirt für membran-gebundene Proteine ist *Rhodothermus capsulatus*. Ein fakultativer phototropher Wirt, welcher je nach den verfügbaren Bedingungen phototroph, chemotroph, autotroph oder heterotroph wachsen kann (Bauer and Wu, 2002; Masepohl and Hallenbeck, 2010). Martinez *et al.* (2004) haben die Wirte *Pseudomonas putida* und *Streptomyces lividans* modifiziert, so dass sie Umwelt-DNA besser aufnehmen und exprimieren können. In *S. lividans* wurden die Antibiotikaresistenzgene Aktinrhodin und Undecylprodiginine (*act* und *red*) deletiert, um die endogene Aktivität zu verringern und Interaktionen mit heterologen Komponenten zu vermeiden. Der neue *P. putida* Stamm MBD1 besitzt eine zusätzliche  $\Phi$ C31 *attB* Sequenz, über die BAC Vektoren mit unterschiedlichen Insertgrößen ins Genom integrieren können. Des Weiteren hat die Gruppe einen BAC Vektor (pMBD14) etabliert, der bis zu 85 kB große Inserts aufnehmen und über eine High-throughput Konjugation in beide Stämme transferiert werden kann (Martinez et al., 2004).

#### 5.4.6 BAC-basierte Metagenom-Genbanken

Neben der Erweiterung des Wirtsspektrums kann man die Insert-Größe der Metagenom-Genbanken expandieren. In dieser Arbeit wurde versucht eine BAC-Metagenom-Bibliothek in *T. thermophilus* zu erstellen. Dazu wurden zwei verschiedene BAC-artige Vektoren für *T. thermophilus* konstruiert, die unterschiedliche Replikationselemente aus dem Megaplasmid pTT27 vom *Thermus*-Wildtyp HB27 tragen. Auf pRetro-BAC1 wurden der *ori* und das putative *repA* Gen des Megaplasמידs eingefügt. pRetro-BAC2 ist eine Erweiterung der ersten Variante und beinhaltet zusätzlich noch den *par*-Locus mit den *parA*, *B* und *parS* Genen, welches von Haijuan Li in ihrer Doktorarbeit am Lehrstuhl für Mikrobiologie genauer untersucht wurde (Li). Die Vektoren wurden mittels Sequenzierung überprüft und es folgte die *in vivo* Modifizierung von zwei rekombinanten Fosmiden mit diesen pRetro-BAC Vektoren und dann die Übertragung der BAC-Konstrukte in *T. thermophilus*, zum einen die leeren pRetro-BAC-Vektoren und zum anderen die *in vivo* mit den Fosmiden rekombinierten Vektoren (kurz: BAC1, BAC2, rfBAC1 und rfBAC2). Southern Blot-Analysen zeigten eine Integration der BAC-Konstrukte in das Genom von HB27. Bei hohen Kanamycinkonzentrationen wuchsen nur die BAC-Konstrukte ohne integrierter metagenomischer DNA im Flüssigmedium an. Das lässt darauf schließen, dass die Aufnahme des BAC-Konstrukts mit einem ca 40 kb nicht funktioniert. Möglich wäre aber eine Integration eines kleinen Genabschnittes mithilfe des BAC-Konstrukts.

Rondon und Kollegen haben BAC Vektoren (pBeloBAC11) für die Klonierung metagenomischer DNA aus Bodenproben verwendet. Dabei entstanden zwei BAC-Bibliotheken SL1 und SL2. SL1 umfasste insgesamt 10 Mbp DNA und die Durchschnittslänge der Fremd-DNA betrug ca. 27 kb, während die zweite Bank das Zehnfache der DNA und eine durchschnittliche Größe von 44 kb aufweisen konnte. Die funktionellen Screenings auf festen Agarplatten in *E. coli* waren erfolgreich und führten zur Auffindung von Genen für DNAsen, zwei Lipasen und acht Amylasen, weiterhin wurde ein antibakterieller Klon identifiziert. Der Klon zeigte eine Wachstums-inhibierende Wirkung auf *B. subtilis* und *Staphylococcus aureus*. Neben der Möglichkeit der Identifizierung neuer Gene, können die klonierten langen kontinuierlichen Genomfragmente die phylogenetische Analyse der aufgefundenen Gene erleichtern (Rondon et al., 2000). Auch größere DNA-Insertionen als 27 - 44 kb konnten mit demselben pBeloBAC11 kloniert werden (Sabehi and Bèjà, 2013).

Prof. Liles von der Universität Auburn in Alabama hat einen Shuttle-BAC Vektor erstellt, der sich über Konjugation oder Elektroporation in diverse Gram-negative Bakterien übertragen lässt (Kakirde et al., 2012). Der pGNS-BAC enthält das F-Replikon, welches die „single-copy Replikation“ in *E. coli* stabilisiert und ein weiteres „broad-host range“ RK2 Minireplikon für die „high-copy“ Replikation in Gram-negativen Mikroorganismen. Für die Selektion kodiert der

Vektor zwei Resistenzgene für Gentamicin und Chloramphenicol und durch den *oriV* aus dem IncP Plasmid ist der Vektor mit Arabinose induzierbar (Popowska and Krawczyk-Balska, 2013). Die aus Bodenproben generierte pGNS-BAC-1 Bibliothek wurde bereits erfolgreich in *E. coli*, *Pseudomonas spp.*, *Salmonella enterica*, *Serratia marcescens*, *Vibrio vulnificus* und *Enterobacter nimipressuralis* übertragen (Kakirde et al., 2012). Eine größere broad-host BAC Bibliothek mit einer Insertgröße von 200 kb für die Stämme *Pseudomonas fluorescence* und *Xanthomonas campestris* wurde von Aakvik etabliert (Aakvik et al., 2009). Dieses Ergebnis ist ein hervorragendes Beispiel für die Kombination einer BAC-Metagenom-Bibliothek mit einem breiten Wirtsspektrum.

Trotz anfänglicher Erfolge bei der Herstellung von BAC-artigen Vektoren für *T. thermophilus* in der vorliegenden Arbeit sind weitere Anstrengungen nötig, bis eine effiziente Klonierung von sehr großen metagenomischen DNA-Fragmenten für nachfolgende funktionelles Screenings im *T. thermophilus* etabliert sein wird.

## 6 Ausblick

Das Projekt mit der Endoglukanase EngU wird am Lehrstuhl für Mikrobiologie fortgeführt und weitere pETDuet Konstrukte nur mit der CBM Domäne erstellt, um diese durch elektrophoretische Gelshiftsassays unter Verwendung von Gelen, die mit verschiedene lösliche Substrate enthalten, zu analysieren. Für die Untersuchung der Bindung an mögliche unlösliche Substrate wurden bereits Bindungsassays ohne native Gele hergestellt. Dabei wurde eine bestimmte Enzymmenge vor und nach der Inkubation mit dem unlöslichen Substrat auf eine SDS-Page übertragen (siehe Bachelorarbeit von Helena Ortiz-Canavate, 2017).

Das verkürzte Protein GH12E (Cel12E ohne dessen zwei CBM Module) wurde aufgereinigt und an die Kooperationspartner an der Swedish University of Agriculture Science weitergegeben. Die Kristallationsversuche des Proteins sind noch im Gange. Für die Gewinnung neuer Glykosynthasen wurde GH12E an der Position E171 mit jeweils drei verschiedenen Aminosäuren substituiert, woraus die Mutanten E171A, E171S und E171G entstanden. Alle drei Mutanten wurden ebenfalls an die Partner in Schweden versendet. Bis heute (Stand 11.2017) sind noch keine Ergebnisse bekannt. Es wird mit großem Interesse verfolgt, welches Potential die Mutanten für eine Anwendung als Glykosynthasen haben.

Das funktionelle Screening von Metagenombanken bedeutet einen nicht unerheblichen Zeitaufwand. Mit gezielten Screeningmethoden, den passenden Expressionswirten und potenziellen Umweltproben ist es dennoch möglich, neuartige Enzyme mit ungewöhnlichem modularen Aufbau und neuen katalytischen Eigenschaften zu entdecken, was für biotechnologische Anwendungen von Interesse ist. Daher ist es wichtig weitere Metagenombanken aus interessanten Habitaten zu erstellen und diese im Labor zu durchforsten. So sollten auch die Analysen zu den Isolaten weitergeführt werden, etwa mit „Mexico SH-Avicel“, die positive Reaktionen auf Xylan, CMC, Milch und Tributyrin gezeigt hat.

## 7 Anhang

Tabelle 38: Primerliste für die EngU Expressionskonstrukte.

| Primerbezeichnung  | Sequenz                                       | Notizen   |
|--------------------|---|---|
| pRetro-B-R         | GTTTCTGTTATACTCCCGGGAGAAGACGGCGAGC<br>TTCTTG  | pRetro-T1 +<br>repA and par<br>locus                              |
| pRetro-B1-F        | CATTATACGAAGTTATCTAGAGCGGGCCATCTCG<br>CTCACGG | pRetro-T1 (ori,<br>kat) + repA+ori                                |
| pRetro-B2-F        | CATTATACGAAGTTATCTAGAGGCAACCCCCTAG<br>AGCTACC | pRetro-T1 (ori,<br>kat) +<br>repA+ori+par<br>locus                |
| pRetro-B S1        | AAATCAGCAACTTAAATAGCC                         | Sequenzierpri<br>mer über loxP<br>von pRetro                      |
| pMKE2 Hind F       | TGCGGCCGCACTCGAGCACC                          | pMKE2 vector<br>PCR   |
| pMKE2 NcoI R       | GTATGCACCTCCGGCCCCAG                          | pMKE2 vector<br>PCR   |
| TB-EST 32-HIS<br>F | TTTAAGAAGGAGATATACATATGCAGCGAGAAAA<br>ACAGCG  | TB-EST-7 32<br>ORF mit<br>HISTaq<br>(pET21a)                      |
| TB-EST 32-HIS<br>R | TGGTGGTGGTGGTGCTCGAGTGC GTCGGCAACT<br>GTTGAGG | TB-EST-7 32<br>ORF mit<br>HISTaq<br>(pET21a)                      |
| qPCR Kan F         | CAAATCGCATCGTGGAAC                            | qPCR  |
| qPCR Kan R         | CATCTCAGGTGGAATCAG                            | qPCR  |
| qKan F1            | TGGAATCAGATTGGCCGCTT                          | qPCR  |
| qKan R1            | CGCATCGTGGAAC TTTTGGG                         | qPCR  |
| Ori lox F          | GAGTTCTTCTGAGCGGGACTCTACCGCATCAGGC<br>GCCATTC | pRetro-M/B Ori<br>und Lox aus<br>pRetro-T2,<br>Überhang zu<br>Kan |
| Ori lox R          | TGAAAGGGCGAATTCCACAG                          | pRetro-M/B Ori<br>und Lox aus<br>pRetro-T2                        |
| amyE F             | CTGTGGAATTCGCCCTTTCATCAGACATCTCCGAT<br>TAACC  | 1,5 kb<br>fragment aus  |

| Primerbezeichnung | Sequenz                                       | Notizen   |
|-------------------|---|---|
|                   |   | amyE (B. subtilis); ÜH zu Ori lox                   |
| amyE R            | CTTTGCAGGGCTTCCCAACCTTCCGGTCAATTCA<br>ATTACTC | 1,5 kb fragment aus amyE (B. subtilis); ÜH zu Kan   |
| trpE TP F         | CTGTGGAATTCGCCCTTTCCTCCCTCCCGTGCG<br>GAGTTC   | 1,5 kb fragment aus trpE (M. luteus); ÜH zu Ori lox |
| trpE TP R         | CTTTGCAGGGCTTCCCAACCTAGCTCGCCCGCG<br>GTCAGCAC | 1,5 kb fragment aus trpE (M. luteus); ÜH zu Kan     |
| trpE TP F2        | CTGTGGAATTCGCCCTTTCAGGTGCTGGCGGACG<br>GCCTCA  | 1,5 kb fragment aus trpE (M. luteus); ÜH zu Ori lox |
| trpE TP R2        | CTTTGCAGGGCTTCCCAACCTAGGTCCACGAGCA<br>TGAGGTG | 1,5 kb fragment aus trpE (M. luteus); ÜH zu Kan     |
| pMK EST HIS F1    | TTGAAGGAGGTGTGAGGCATATGCAGCGAGAAAA<br>ACAGCG  | OH to pMK18 vector PCR and pET-EST-HIS              |
| pMK EST HIS R1    | AATCGCCTTGCAGCACATCCTCAGTGGTGGTGGT<br>GGTGCT  | OH to pMK18 vector PCR and pET-EST-HIS              |
| pET Upstream      | ATGCGTCCGGCGTAGA                              | Sequenzieren von pETvector / pETDUET                |
| DuetUP2 Primer    | TTGTACACGGCCGCATAATC                          | Sequenzieren von pETDUET                            |
| DuetDOWN1 Primer  | GATTATGCGGCCGTGTACAA                          | Sequenzieren von pETDUET                            |
| T7 Terminator     | GCTAGTTATTGCTCAGCGG                           | Sequenzieren von pETvector / pETDUET                |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>                                 | <b>Notizen</b>                                    |
|--------------------------|--|---|
| HA CMC F                 | TTTAAGAAGGAGATATACAATGCAGGAGACAACA<br>GTGCTGGA | pET21a_GH1<br>2E Gibson<br>Primer                 |
| GH12E R                  | TGGTGGTGCTCGAGTGCTCAGCTTGATGGCAACA<br>ATGCCT   | pET21a_GH1<br>2E Gibson<br>Primer mit<br>TGA      |
| E171A F                  | GGGTTAAAGCAGGGGATATCGCGATAATGGTGTG<br>GCTCTA   | pET21a_GH1<br>2E Mutation<br>E171 Alanin<br>(GCG) |
| E171S F                  | TTAAAGCAGGGGATATCTCGATAATGGTGTGGCT<br>CTA      | pET21a_GH1<br>2E Mutation<br>E171 Serin<br>(UCG)  |
| E171G F                  | GTAAAGCAGGGGATATCGGCATAATGGTGTGGC<br>TCTAC     | pET21a_GH1<br>2E Mutation<br>E171 Glycin<br>(GGC) |
| D123A R                  | CGTTGGGAAACTCTTGCATCGCCATGTGATGGAA<br>TCGGAC   | pET21a_bga4<br>8full Mutation<br>D123A (GCG)      |
| E235A R                  | CCGTTTTTCGTTGTTGATCGCCACAAAGGCAATTG<br>CAGGGT  | pET21a_bga4<br>8full Mutation<br>E239A (GCG)      |
| E593A R                  | CGAGCAACAAGAAGGCCGCACTGCTGTAAGTGTT             | pET21a_bga4<br>8full Mutation<br>E593A (GCG)      |
| D605A R                  | GAAGGCGTAAATTGCGTCCCACGCCTGAAGGGC<br>TGCGTAG   | pET21a_bga4<br>8full Mutation<br>D605A (GCG)      |
| Duet bga48 F1            | TAAC TT TAAGAAGGAGATATACCATGCGTTGCGG<br>TCTGAC | Umklonierung<br>von bga48full<br>in pETDuet       |
| Duet bga48 R             | CTTAAGCATTATGCGGCCGCATCATTGCACCACA<br>ACTTCA   | Umklonierung<br>von bga48full<br>in pETDuet       |
| Bgl48 Seq 01 R           | CTCTTTCGTCAACCTGACAAC                          | Sequenzierung                                     |
| Bgl48 Seq 01 R           | CAGCCA ACTCAGCTTCCTTTTCG                       | Sequenzierung                                     |
| pRetro-B-R               | GTTTCTGTTATACTCCCGGGAGAAGACGGCGAGC<br>TTCTTG   | pRetro-T1 +<br>repA and par<br>locus              |
| bga_for                  | GATCCATATGCGTTGCGGTCTGACACAGGAA                | Bgl Konstrukte                                    |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>                                 | <b>Notizen</b> |
|--------------------------|--|----------------|
| bga_rev                  | GATCAAGCTTTCATTGCACCACAACCTTCATAC              | Bgl Konstrukte |
| cbm_rev                  | GATCAAGCTTTCACGTCTGTGCCCCCAAGTTGCTG            | Bgl Konstrukte |
| gh42_rev                 | GATCAAGCTTTCAAAAGTTGGCGTTGCGCAACAT             | Bgl Konstrukte |
| bga_for ecoRI            | GATCGAATTCATGCGTTGCGGTCTGACACA                 | Bgl Konstrukte |
| bga_for NcoI             | GATGCCATGGGTTGCGGTCTGACACA                     | Bgl Konstrukte |
| cterm_for                | GATCCATATGGCACGGGTCATATTCAG                    | Bgl Konstrukte |
| bga_rev KpnI             | GGCCGGTACCTTTCATTGCACCACAACCTTCAT              | Bgl Konstrukte |
| Duet_Bga48 F1            | TAACCTTTAAGAAGGAGATATACCATGCGTTGCGGTCTGAC      | Bgl Konstrukte |
| Duet_Bga48 R             | CTTAAGCATTATGCGGCCGCATCATTGCACCACA<br>ACTTCA   | Bgl Konstrukte |
| gh42a_for                | ACTTTAAGAAGGAGATATACCTGGACTACCTGATT<br>GCCCA   | Bgl Konstrukte |
| gh42 b rev2              | ATTATGCGGCCCGCAGTGTCACCTGACAACCACTT<br>GACGCT  | Bgl Konstrukte |
| Duet bga48 F1            | TAACCTTTAAGAAGGAGATATACCATGCGTTGCGGTCTGAC      | Bgl Konstrukte |
| Duet bga48 R             | CTTAAGCATTATGCGGCCGCATCATTGCACCACA<br>ACTTCA   | Bgl Konstrukte |
| E239A                    | CAAGCGTGAACGAGACCGTTCGCGTTGTTGATTT<br>CCACAAAG | Bgl Konstrukte |
| E581A                    | GTAAACCCTTCTCGGTTACAGCGTACAACCATTCT<br>GCTCCG  | Bgl Konstrukte |
| D151A_rev                | AGTTGGGCAATCAGGTACGCCAGTCGGTCAAGG<br>GCTT      | Bgl Konstrukte |
| E183A_for                | GCCGACGGTTTGCCCAAAGCGATTGAACAACCTTG<br>GTT     | Bgl Konstrukte |
| D123A R                  | CGTTGGGAAACTCTTGCATCGCCATGTGATGGAA<br>TCGGAC   | Bgl Konstrukte |
| E235A R                  | CCGTTTTTCGTTGTTGATCGCCACAAAGGCAATTG<br>CAGGGT  | Bgl Konstrukte |
| E593A R                  | CGAGCAACAAGAAGGCCGCACTGCTGTAAGTGTT             | Bgl Konstrukte |
| D605A R                  | GAAGGCGTAAATTGCGTCCCACGCCTGAAGGGC<br>TGCGTAG   | Bgl Konstrukte |
| pET-Duet 1 SEQ           | TGGCAAGTGTAGCGGTCACG                           | Bgl Konstrukte |



Tabelle 39: Primerliste für die Transkriptomanalyse des pCT3FK-HA-CMC-1 Fosmids

| <b>Primerbezeichnung</b> | <b>Sequenz</b>         | <b>Region</b> |
|--------------------------|------------------------|---------------|
| HA01-F                   | TCGACAAACGGGCCCTGATAG  | 649 - 760     |
| HA01-R                   | ACGTCGTGGCGTTTGAACAGC  |               |
| HA02-F                   | TCTTGCCCTTGACGGGAAC    | 1805 - 1712   |
| HA02-R                   | CCCGCCAACAGGAAGAAAC    |               |
| HA03-F                   | CTCTGCCCCGAAATCCCATC   | 2020 - 1901   |
| HA03-R                   | AGCTCGAACTCCTGAGAAG    |               |
| HA05-F                   | GCCGGCGACCATCATTATTTTC | 3180 - 3287   |
| HA05-R                   | AACCACCAAGGGCTTCACAC   |               |
| HA06-F                   | AGAGCAGCAAGGCTCCAAC    | 3876 - 3980   |
| HA06-R                   | GCGCATTGGGCAGTGAAAC    |               |
| HA07-F                   | AAGGAGCTCGGCTACAACG    | 5057 - 5117   |
| HA07-R                   | CCACTTCTCACCGAGGAAC    |               |
| HA08-F                   | CAACGGCCTCAAGAACTTCC   | 6265 - 6357   |
| HA08-R                   | ACGTCGCTGACGTGAGATG    |               |
| HA09-F                   | AGGCGACAGAGATAGGAAG    | 6963 - 7069   |
| HA09-R                   | TGCGGCCTCAAGTCCTATG    |               |
| HA10-F                   | ACTGGGCAGGACTTGAAAC    | 8149 - 8247   |
| HA10-R                   | CGCTGAGAACGGCTTTAACAC  |               |
| HA11-F                   | CGAGTACTACAAGCCCTATC   | 9321 - 9439   |
| HA11-R                   | CCATCTCGACTTCCATTC     |               |
| HA12-F                   | GGCCATTGAGCTGAAAAC     | 9570 - 9687   |
| HA12-R                   | GAAGTCCGGTAGATCCTTAG   |               |
| HA13-F                   | AGCGGTAAATGCTGGGAAG    | 9938 - 10057  |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>        | <b>Region</b> |
|--------------------------|-----------------------|---------------|
| HA13-R                   | TCCGTGGTGAGAACTATG    |               |
| HA14-F                   | CTAACCGCCACTGAGAG     | 10550 - 10645 |
| HA14-R                   | TATGACGGCGAACGTCC     |               |
| HA15-F                   | TTCAGCGAGCCGTTTATC    | 11160 - 11256 |
| HA15-R                   | CTGAGGAGAGCTTCATAAGG  |               |
| HA16-F                   | CCCTCGTATTTGCTGTACC   | 11684 - 11565 |
| HA16-R                   | GAAGCGCTACCTCAGATG    |               |
| HA18-F                   | CCGGTGCTATGAGAAGTACGG | 11982 - 12077 |
| HA18-R                   | CTCGAAACGGGTCTTGAAGGG |               |
| HA20-F                   | CCCTCGACGAGCTTGTAAC   | 13396 - 13495 |
| HA20-R                   | GTCGGCGGTGAAATCAG     |               |
| HA21-F                   | GGAGTTTCCCGGTATCATCTG | 14879 - 14984 |
| HA21-R                   | TATGTCCTCCACGCTCTTG   |               |
| HA22-F                   | TTCACGAGCGTGACATC     | 15883 - 15995 |
| HA22-R                   | GTCGTCATCTCCAGAGTAG   |               |
| HA23-F                   | ATCCGTGGAACATCCAAG    | 16658 - 16753 |
| HA23-R                   | CGATGTCCGTGAGATTAC    |               |
| HA24-F                   | TTCTCCTGACGGCTCTTC    | 18699 - 18800 |
| HA24-R                   | GATCGGTGGCGCTTATTATG  |               |
| HA25-F                   | TTGGCGAGGCCCATGAAAC   | 20108 - 20015 |
| HA25-R                   | CGCTCCGCTACATCAACTATC |               |
| HA26-F                   | GGCGAAATCGGCTCAAG     | 21941 - 21850 |
| HA26-R                   | GATGACCCGAGCATGTG     |               |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>        | <b>Region</b> |
|--------------------------|-----------------------|---------------|
| HA27-F                   | CTGGCGTTCATGCTGTTG    | 23354 - 23259 |
| HA27-R                   | GAGGGAGCTCGTTGAGATAG  |               |
| HA28-F                   | CGAAGATTGTCGCGATGTACG | 24079 - 24190 |
| HA28-R                   | ACTATCTCCACGCCATCCAC  |               |
| HA29-F                   | CCGTACTCCCACATGTTTG   | 24750 - 24844 |
| HA29-R                   | AGCTTGCGTAGGATTC      |               |
| HA30-F                   | TTCACGTGCCTCGTGTAG    | 25819 - 25713 |
| HA30-R                   | GAGCTGAGGGCGGTTATATTC |               |
| HA31-F                   | CCGAAGAGCTCAGATAGG    | 26547 - 26446 |
| HA31-R                   | TCCGGTAGTTGAGGATG     |               |
| HA32-F                   | GAACGACGAGGGCAATTAG   | 27226 - 27133 |
| HA32-R                   | ATAGGCGTCCTTGTGTACC   |               |
| HA33-F                   | CGGCCAGATTCGTCAGAAG   | 28193 - 28101 |
| HA33-R                   | TGGGCTACCTCACCTACAC   |               |
| HA34-F                   | CTCGTCTATCTCGACCATCAC | 28973 - 28872 |
| HA34-R                   | CCGAGAAGGGTTCTGGTAATC |               |
| HA35-F                   | GCCGCGCTTATAGTCATC    | 29600 - 29481 |
| HA35-R                   | CGTCATCCCAGCCCATATAG  |               |
| HA36-F                   | GTCCGAGCTCACTGTCCATC  | 30826 - 30715 |
| HA36-R                   | CCCTGAGGAGCTTCACTACG  |               |
| HA37-F                   | CAAGAGACTCGCAGAAG     | 31380 - 31273 |
| HA37-R                   | GCCTCACTGAAGAGTATG    |               |
| HA38-F                   | GAGAAGGATCATGCCGTAAG  | 32344 - 32255 |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>        | <b>Region</b> |
|--------------------------|-----------------------|---------------|
| HA38-R                   | CCGGAGGCACATTCAAAG    |               |
| HA39-F                   | GGCCGATGAAGAAGAGC     | 33246 - 33153 |
| HA39-R                   | GGGATATACGGCCTGATG    |               |
| HA40-F                   | CTCCGTCGGGAGGAAATTGG  | 33492 - 33383 |
| HA40-R                   | AGTTCAACTGGGCGAGCCTG  |               |
| HA41-F                   | AATGGCCGAGACCAGGTTG   | 33875 - 33778 |
| HA41-R                   | TGCTTGGCCTGAGCTACAC   |               |
| HA42-F                   | TATCAGGAAGGCCAGCACCAG | 34352 - 34234 |
| HA42-R                   | TCCACGCGGCAACTAAG     |               |
| HA43-F                   | CCGCTATTATGACGACAACC  | 34658 - 34558 |
| HA43-R                   | GGCGATACTGATAACCTACC  |               |
| HA44-F                   | ATGAACGCCGCTATCATC    | 35712 - 35075 |
| HA44-R                   | CTGGTTCTCATGGCATTCTG  |               |
| HA45-F                   | CCGTCAGTGAGGTTTAC     | 35690 - 35797 |
| HA45-R                   | CGAGCCAGTTCTCATAG     |               |
| HA46-F                   | TGGGACGCAAAGGTAAAGC   | 35978 - 36072 |
| HA46-R                   | GGGTTCTGTTCGAGGTAGAAG |               |
| HA47-F                   | GCGAGTTTGCTGGCTTAATC  | 37009 - 37112 |
| HA47-R                   | ATCGGCACCGAGGTAATC    |               |
| HA48-F                   | CCCACGATCTTCCTCTTG    | 38031 - 37925 |
| HA48-R                   | CCGGTTGAGTTCATAGGC    |               |

Tabelle 40: Primerliste für die Transkriptomanalyse des pCT3FK-TB-EST-7 Fosmids

| <b>Primerbezeichnung</b> | <b>Sequenz</b>        | <b>Region</b> |
|--------------------------|-----------------------|---------------|
| TB01-F                   | CGGCTATGCCGAACTAGAAAC | 230 - 338     |
| TB01-R                   | AATCTGCGTGACGAGATCC   |               |
| TB02-F                   | GCGCAACGTCTGAATCCAC   | 1101 - 1007   |
| TB02-R                   | GGGCAACGCTTGTTGAG     |               |
| TB03-F                   | TGCACCCGACCGAATAG     | 1620 - 1526   |
| TB03-R                   | GACCGGATCAACGTGCAAG   |               |
| TB04-F                   | GCCGCTACCTGTCTGTTAC   | 2521 - 2632   |
| TB04-R                   | CGATTGCGGGCAATGAACC   |               |
| TB05-F                   | GTGGCGCGGTTTAACAC     | 4745 - 4839   |
| TB05-R                   | TGCGGCAGCGAGAAATC     |               |
| TB06-F                   | AGAAGCGGTGGACACTC     | 5820 - 5934   |
| TB06-R                   | GATGCCAACCCTGTAGG     |               |
| TB07-F                   | TGCTCGAAGGACTGAAC     | 6568 - 6684   |
| TB07-R                   | GCCGTCCAATTCGAAAG     |               |
| TB08-F                   | CGTCGCCGAACTATTAC     | 8611 - 8720   |
| TB08-R                   | AAAGCGGACGAGGATTG     |               |
| TB12-F                   | TCCCTCCTCCGTGTACATC   | 13130 - 13037 |
| TB12-R                   | AGCACGGGAGCGAATTG     |               |
| TB13-F                   | CCATGGCATGCAGATTACC   | 13589 - 13493 |
| TB13-R                   | AGCTGGGAGCGTACTAAAG   |               |
| TB14-F                   | GCGCCGTAAACGCCAAATC   | 14041 - 13922 |
| TB14-R                   | GCAGGCATTGCTTGAGGAAGG |               |
| TB15-F                   | CCGAAGTGCTGGACTATC    | 14714 - 14818 |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>         | <b>Region</b> |
|--------------------------|------------------------|---------------|
| TB15-R                   | CCAGCTCGAAGAGCTTTG     |               |
| TB16-F                   | GCGCATCGTGGATCTAAAG    | 16442 - 16561 |
| TB16-R                   | ATGTCGGCACCAAACCTC     |               |
| TB17-F                   | TGGAAGAAGTGCCTCCTACTG  | 17571 - 17675 |
| TB17-R                   | GGGCTCGGTTCGATTAATAACG |               |
| TB18-F                   | CCATTGTGGCCATTCCC      | 18464 - 18580 |
| TB18-R                   | CGCCGGCTACAAATGTG      |               |
| TB19-F                   | CAGGCCGACTATGAACG      | 19331 - 19421 |
| TB19-R                   | CGGCGAACCAGTTGTTG      |               |
| TB20-F                   | AGCCGCATTCCGAATAC      | 20691 - 20594 |
| TB20-R                   | CGCGCAATGTCATCTATC     |               |
| TB21-F                   | ATGGTCTCCACCTTCAC      | 21531 - 21426 |
| TB21-R                   | GCGGATCTGGCATTCTG      |               |
| TB22-F                   | CGCAAACCGACACATAG      | 22406 - 22294 |
| TB22-R                   | GTTGGACCGGTACTTTC      |               |
| TB23-F                   | CTCTCCACGGTTGACCTTTG   | 23771 - 23659 |
| TB23-R                   | TGGGATTGGGCTATGTAGGG   |               |
| TB27-F                   | GAACCTGAGCGATACTTC     | 25066 - 25161 |
| TB27-R                   | GGTCTGTTCCCAGATTTC     |               |
| TB28-F                   | AACAGGGCCTGGATGTG      | 25780 - 25880 |
| TB28-R                   | GGTGTATCCGGGAAACGTAG   |               |
| TB29-F                   | GATGCCGTGGAAGTAAGTG    | 26091 - 26192 |
| TB29-R                   | GCGTCCTTCCTCATTGAAC    |               |

| <b>Primerbezeichnung</b> | <b>Sequenz</b>        | <b>Region</b> |
|--------------------------|-----------------------|---------------|
| TB30-F                   | GGGACTCGTCGACTTTG     | 26532 - 26627 |
| TB30-R                   | GCCGACTCCAGTTTAGG     |               |
| TB31-F                   | ACGACCAAACCAGCGGTAAC  | 27741 - 27652 |
| TB31-R                   | CTGCTCTACGGACTCATTGTG |               |
| TB32-F                   | CTACCGAATAGCCAAAGG    | 28892 - 28781 |
| TB32-R                   | CATATAGGCGTGCTGTC     |               |
| TB33-F                   | TGCTTTGCTGCTGCTCCTC   | 29273 - 29373 |
| TB33-R                   | ACTGGATCGCGCTGCACAAC  |               |
| TB32-F2                  | CACAGCAATGGTGTAACGGG  |               |
| TB32-R2                  | GCATTTTTGCGCCGGTAGAA  |               |

Tabelle 41: Gegebüberstellung von Cel12E und GH12E

|               |  |
|---------------|--|
| Name:         | <b>Cel12E</b>  |
| AS-Sequenz    | <p>MQETTVLEFPGTQGSTALIDMNGDGISDFIMEINPWNIQDAQGKQIMEY<br/>           DPNNNEIRFSSNLTDIVPKNSDRWIYGYPEVYFGSKPWNSKVAGLLK<br/>           LPKKVSDLSGFTLKFYEYSLEHDPNLPINLAMETWLTTDQLATSGVKAGD<br/> <b>I</b>EIMVWLYSSKLNPAGRKIDTVTPMIVNGRLVNESFEVWKMEGMGSG<br/>           WTYFAFRLTTPLKSAEIGIDPTLFIKKVEGYIQTNVENLYMQDWEIGT<b>E</b>F<br/>           GNPTTTSALFNWTIRNLEVNEEALLPSSTSSPEQNQTSPSNTTNLIKPG<br/>           SLDVKVNSWGSATQYSCTLYLGAQYDWSVKVELKDGSEITSYWSADA<br/>           TEENGTVVFTPKSWNRGPTASFGFIASGDVPVESITLIVDGKVWDWWP<br/>           NEAQANTTSSNQTSTQNATATNQTNTTTSTNTTSSNSNQTSTDQTNTTN<br/>           TTSDTLKPGSLSVSINDWGTGGQFNINLDLGGQYEWVVKVSVESPTTI<br/>           GNYWSAQKSEENGWIIFTPTNWNKGPTASFGFIVNGPVTGVKQIILEVN<br/>           GELWDIWTPQ</p> |
| Eigenschaften | 543 AS ; 59,96 kDa ; pl: 4,23  |
| Activität     | 4-5 U/mg auf 0,3 % CMC, 85 °C bis 96 °C  |
| Name:         | <b>GH12E 01</b>  |
| AS-Sequenz    | <p>MQETTVLEFPGTQGSTALIDMNGDGISDFIMEINPWNIQDAQGKQIMEY<br/>           DPNNNEIRFSSNLTDIVPKNSDRWIYGYPEVYFGSKPWNSKVAGLLK<br/>           LPKKVSDLSGFTLKFYEYSLEHDPNLPINLAMETWLTTDQLATSGVKAGD<br/> <b>I</b>EIMVWLYSSKLNPAGRKIDTVTPMIVNGRLVNESFEVWKMEGMGSG<br/>           WTYFAFRLTTPLKSAEIGIDPTLFIKKVEGYIQTNVENLYMQDWEIGT<b>E</b>F<br/>           GNPTTTSALFNWTIRNLEVNEEALLPSS</p>   |
| Eigenschaften | 272 AS ; 30,6 kDa ; pl: 4,33   |
| Activity      | 6-8 U/mg auf 0,2 % CMC, 80 °C  |



### Eckhardts Gel

### Pulse-field gelelectrophoration

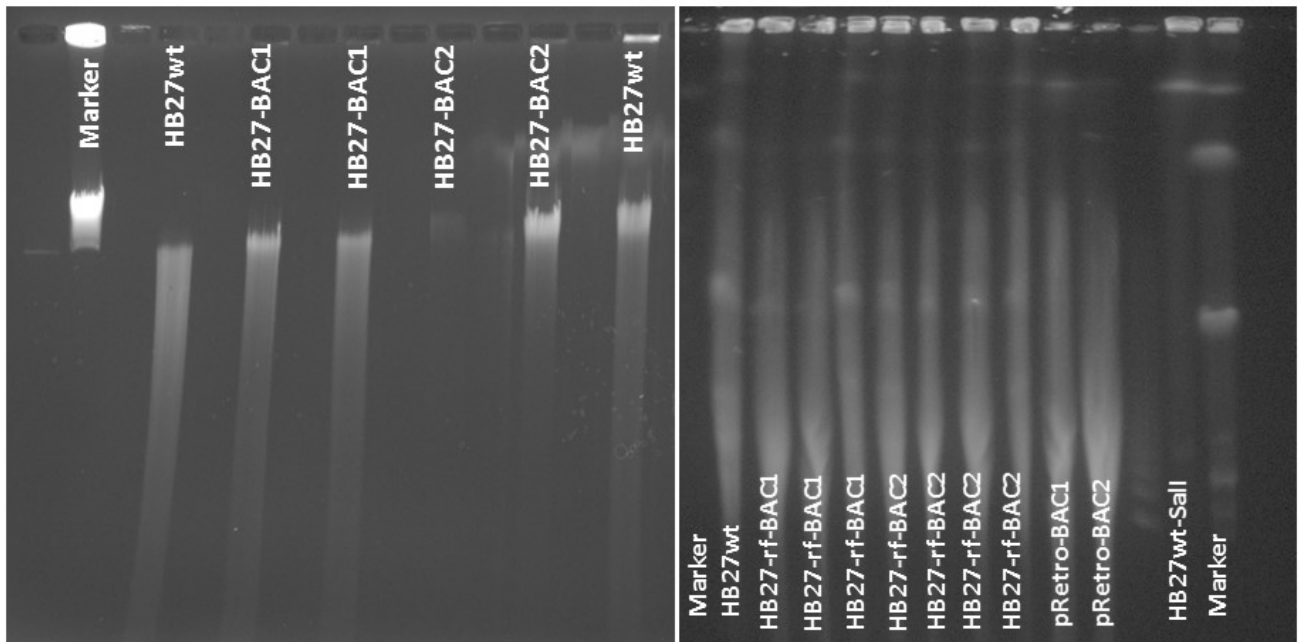


Abbildung 31: Eckhardts Gel and PFGE-Gel zur Visualisierung der pRetro-BAC1 und pRetro-BAC2 Vektoren in *T. thermophilus* HB27wt

## 8 Literaturverzeichnis

- Aakvik, T., Degnes, K. F., Dahlsrud, R., Schmidt, F., Dam, R., Yu, L., et al. (2009). A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. *FEMS Microbiol. Lett.* 296, 149–158. doi:10.1111/j.1574-6968.2009.01639.x.
- Adinarayana, K., Ellaiah, P., and Prasad, D. S. (2003). Purification and partial characterization of thermostable serine alkaline protease from a newly isolated *Bacillus subtilis* PE-11. *AAPS PharmSciTech* 4, E56. doi:10.1208/pt040456.
- Alahuhta, M., Luo, Y., Ding, S. Y., Himmel, M. E., and Lunin, V. V. (2011). Structure of CBM4 from *Clostridium thermocellum* cellulase K. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 67, 527–530. doi:10.1107/S1744309111003307.
- Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169. doi:10.1016/j.jip.2007.09.009.
- Ando, S., Ishida, H., Kosugi, Y., Ishikawa, K. (2002). Hyperthermo-stable endoglucanase from *Pyrococcus horikoshi*. *Appl. Environ. Microbiol.* 68, 430–433.
- Angelov, A., Bergen, P., Nadler, F., Hornburg, P., Lichev, A., Åœebelacker, M., et al. (2015). Novel Flp pilus biogenesis-dependent natural transformation. *Front. Microbiol.* 6, 1–11. doi:10.3389/fmicb.2015.00084.
- Angelov, A., Mientus, M., Liebl, S., and Liebl, W. (2009). A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. *Syst. Appl. Microbiol.* 32, 177–85. doi:10.1016/j.syapm.2008.01.003.
- Angelov, A., Pham, V. T. T., Übelacker, M., Brady, S., Leis, B., Pill, N., et al. (2017). A metagenome-derived thermostable  $\beta$ -glucanase with an unusual module architecture which defines the new glycoside hydrolase family GH148. *Sci. Rep.* 7, 17306. doi:10.1038/s41598-017-16839-8.
- Annamalai, N., Thavasi, R., Vijayalakshmi, S., and Balasubramanian, T. (2011). Extraction, Purification and Characterization of Thermostable, Alkaline Tolerant ??-Amylase from *Bacillus cereus*. *Indian J. Microbiol.* 51, 424–429. doi:10.1007/s12088-011-0160-z.

- Arpigny, J. L., and Jaeger, K.-E. E. (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* 343 Pt 1, 177–183. doi:10.1042/0264-6021:3430177.
- Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. a (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife*, 1–23. doi:10.7554/eLife.08504.
- Balan, A., Ibrahim, D., Abdul Rahim, R., and Ahmad Rashid, F. A. (2012). Purification and characterization of a thermostable lipase from *Geobacillus thermodenitrificans* IBRL-nra. *Enzyme Res.* 2012. doi:10.1155/2012/987523.
- Bauer, M. W., Driskill, L. E., Callen, W., Snead, M. a, Mathur, E. J., and Kelly, R. M. (1999). An Endoglucanase , EglA , from the Hyperthermophilic Archaeon *Pyrococcus furiosus* Hydrolyzes  $\beta$  -1 , 4 Bonds in Mixed-Linkage ( 1 3 3 ),( 1 3 4 ) -  $\beta$  - D -Glucans and Cellulose. 181, 284–290.
- Bauer, and Wu (2002). RegB/RegA, a global redox-responding two- component system. *Adv. Exp. Med. Biol.* 631, 131–148.
- Becker, P. (1997). Determination of the kinetic parameters during continuous cultivation of the lipase producing thermophile *Bacillus* sp. IHI-91 on olive oil. *Appl. Microbiol. Biotechnol.* 48, 184–190.
- Beloqui, a, Nechitaylo, T. Y., Lopez-Cortes, N., Ghazi, a, Guazzaroni, M. E., Polaina, J., et al. (2010). Diversity of Glycosyl Hydrolases from Cellulose-Depleting Communities Enriched from Casts of Two Earthworm Species. *Appl. Environ. Microbiol.* 76, 5934–5946. doi:10.1128/aem.00902-10.
- Belval, L., Marquette, A., Mestre, P., Piron, M. C., Demangeat, G., Merdinoglu, D., et al. (2015). A fast and simple method to eliminate Cpn60 from functional recombinant proteins produced by *E. coli* Arctic Express. *Protein Expr. Purif.* 109, 29–34. doi:10.1016/j.pep.2015.01.009.
- Berendsen, E. M., Wells-bennik, M. H. J., Krawczyk, A. O., Jong, A. De, Heel, A. Van, Holsappel, S., et al. (2016). Draft Genome Sequences of Seven Thermophilic Spore-Forming Bacteria Isolated from Foods That Produce Highly Heat-Resistant Spores , Comprising *Geobacillus* spp ., *Caldibacillus debilis* , and. 4, 4–5. doi:10.1128/genomeA.00105-16. Copyright.

- Bjornsdottir, S. H., Fridjonsson, O. H., Kristjansson, J. K., and Eggertsson, G. (2007). Cloning and expression of heterologous genes in *Rhodothermus marinus*. *Extremophiles* 11, 283–293. doi:10.1007/s00792-006-0037-y.
- Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382, 769–781. doi:10.1042/BJ20040892.
- Bosma, E. F., van de Weijer, A. H. P., Daas, M. J. a, van der Oost, J., de Vos, W. M., and van Kranenburg, R. (2015a). Isolation and screening of thermophilic bacilli from compost for electrotransformation and fermentation: Characterization of bacillus smithii ET 138 as a new biocatalyst. *Appl. Environ. Microbiol.* 81, 1874–1883. doi:10.1128/AEM.03640-14.
- Bosma, E. F., van de Weijer, A. H. P., van der Vlist, L., de Vos, W. M., van der Oost, J., and van Kranenburg, R. (2015b). Establishment of markerless gene deletion tools in thermophilic *Bacillus smithii* and construction of multiple mutant strains. *Microb. Cell Fact.* 14, 99. doi:10.1186/s12934-015-0286-5.
- Chang, Y.-Y., Hung, C.-H., Hwang, T.-S., and Hsu, C.-H. (2013). Cloning, overexpression, purification and crystallization of malate dehydrogenase from *Thermus thermophilus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 69, 1249–1251. doi:10.1107/S174430911302472X.
- Charnock, S. J., Brown, I. E., Turkenburg, J. P., Black, G. W., and Davies, G. J. (2001). Characterization of a novel pectate lyase, Pel10A, from *Pseudomonas cellulosa*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 57, 1141–1143. doi:10.1107/S0907444901007491.
- Chin, K. J., and Janssen, P. H. (2002). Propionate formation by *Opiritus terrae* in pure culture and in mixed culture with a hydrogenotrophic methanogen and implications for carbon fluxes in anoxic rice paddy soil. *Appl. Environ. Microbiol.* 68, 2089–2092. doi:10.1128/AEM.68.4.2089-2092.2002.
- Choo, Y. J., Lee, K., Song, J., and Cho, J. C. (2007). *Puniceicoccus vermicola* gen. nov., sp. nov., a novel marine bacterium, and description of Puniceicocaceae fam. nov., Puniceicoccales ord. nov., Opiritutaceae fam. nov., Opiritutales ord. nov. and Opiritutae classis nov. in the phylum “Verrucomicrobia.” *Int. J. Syst. Evol. Microbiol.* 57, 532–537. doi:10.1099/ijs.0.64616-0.

- Culligan, P., Sleator, R. D., Marchesi, J. R., and Hill, C. (2014). Metagenomics and novel gene discovery Promise and potential for novel therapeutics. 399–412.
- Daniel, R. (2005). The metagenomics of soil. *Nat. Rev. Microbiol.* 3, 470–478. doi:10.1038/nrmicro1160.
- Das, K., Doley, R., and Mukherjee, A. K. (2004). Purification and biochemical characterization of a thermostable, alkaliphilic, extracellular alpha-amylase from *Bacillus subtilis* DM-03, a strain isolated from the traditional fermented food of India. *Biotechnol. Appl. Biochem.* 40, 291–8. doi:10.1042/BA20040034.
- Davies G, Wilkson KS, H. B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochemistry* 321, 557–559.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* 3, 853–859. doi:10.1016/S0969-2126(01)00220-9.
- Dean, F. B., Nelson, J. R., Giesler, T. L., and Lasken, R. S. (2001). Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification. *Genome Res.*, 1095–1099. doi:10.1101/gr.180501.4.
- Dimarogona, M., Topakas, E., Olsson, L., and Christakopoulos, P. (2012). Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour. Technol.* 110, 480–487. doi:10.1016/j.biortech.2012.01.116.
- Ding, S. ., Lamed, R., Bayer, E. A., and Himmel, M. E. (2003). The bacterial scaffoldin: structure, function and potential applications in the nanoscience. *Genet. Eng.* 25, 209–225.
- Distel, D. L., Morrill, W., Maclaren-toussaint, N., Franks, D., and Waterbury, J. (2015). *Teredinibacter turnerae* gen . nov ., sp . nov ., a dinitrogen-fixing , cellulolytic , endosymbiotic  $\gamma$  -proteobacterium isolated from the gills of wood-boring molluscs ( *Bivalvia* : *Teredinidae* ). 2261–2269.
- Dwidar, M., Park, J.-Y., Mitchell, R. J., and Sang, B.-I. (2012). The Future of Butyric Acid in Industry. *Sci. World J.* 2012, 1–10. doi:10.1100/2012/471417.
- Ebina, T., Toh, H., and Kuroda, Y. (2009). Loop-length-dependent SVM prediction of domain linkers for high-throughput structural proteomics. *Biopolymers* 92, 1–8. doi:10.1002/bip.21105.

- Ekkers, D. M., Cretoiu, M. S., Kielak, A. M., and Van Elsas, J. D. (2012). The great screen anomaly—a new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.* 93, 1005–1020. doi:10.1007/s00253-011-3804-3.
- Eloe, E. a., Fadrosch, D. W., Novotny, M., Zeigler Allen, L., Kim, M., Lombardo, M.-J., et al. (2011). Going Deeper: Metagenome of a Hadopelagic Microbial Community. *PLoS One* 6, e20388. doi:10.1371/journal.pone.0020388.
- Eminoğlu, A., Ülker, S., and Sandallı, C. (2015). Cloning, Purification and Characterization of Acetyl Xylane Esterase from *Anoxybacillus flavithermus* DSM 2641T with Activity on Low Molecular-Weight Acetates. *Protein J.* 34, 237–242. doi:10.1007/s10930-015-9618-x.
- Ferrer, M., Ghazi, A., Beloqui, A., Vieites, J. M., López-Cortés, N., Marín-Navarro, J., et al. (2012). Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 7, e38134. doi:10.1371/journal.pone.0038134.
- Ferrer, M., Martinezabarca, F., and Golyshin, P. (2005). Mining genomes and “metagenomes” for novel catalysts. *Curr. Opin. Biotechnol.* 16, 588–593. doi:10.1016/j.copbio.2005.09.001.
- Fink, A. L. (1999). Chaperone-mediated protein folding. *Physiol. Rev.* 79, 425–449.
- Frostegård, Å., Courtois, S., Ramisse, V., Clerc, S., Bernillon, D., Le Gall, F., et al. (1999). Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* 65, 5409–5420.
- Fuerst, J. a, Gwilliam, H. G., Lindsay, M., Lichanska, A., Belcher, C., Vickers, J. E., et al. (1997). Isolation and molecular identification of planctomycete bacteria from postlarvae of the giant tiger prawn , *Penaeus monodon* . Isolation and Molecular Identification of Planctomycete Bacteria from Postlarvae of the Giant Tiger Prawn , *Penaeus monodon*. 63, 254–262.
- Gaur, R., and Tiwari, S. (2015). Isolation, production, purification and characterization of an organic-solvent-thermostable alkalophilic cellulase from *Bacillus vallismortis* RG-07. *BMC Biotechnol.* 15, 19. doi:10.1186/s12896-015-0129-9.

- Gloux, K., Berteau, O., El oumami, H., Beguet, F., Leclerc, M., and Dore, J. (2011). A metagenomic  $\alpha$ -glucuronidase uncovers a core adaptive function of the human intestinal microbiome. *Proc. Natl. Acad. Sci.* 108, 4539–4546. doi:10.1073/pnas.1000066107.
- Goldman, A. D., Beatty, J. T., and Landweber, L. F. (2016). The TIM Barrel Architecture Facilitated the Early Evolution of Protein-Mediated Metabolism. *J. Mol. Evol.* 82, 1–10. doi:10.1007/s00239-015-9722-8.
- Gomes, I., Gomes, J., and Steiner, W. (2003). Highly thermostable amylase and pullulanase of the extreme thermophilic eubacterium *Rhodothermus marinus*: Production and partial characterization. *Bioresour. Technol.* 90, 207–214. doi:10.1016/S0960-8524(03)00110-X.
- Gong, X., Gruninger, R. J., Qi, M., Paterson, L., Forster, R. J., Teather, R. M., et al. (2012). Cloning and identification of novel hydrolase genes from a dairy cow rumen metagenomic library and characterization of a cellulase gene. *BMC Res. Notes* 5, 1. doi:10.1186/1756-0500-5-566.
- Hadad, D., Geresh, S., and Sivan, A. (2005). Biodegradation of polyethylene by the thermophilic bacterium *Brevibacillus borstelensis*. *J. Appl. Microbiol.* 98, 1093–1100. doi:10.1111/j.1365-2672.2005.02553.x.
- Handelsman, J. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Metagenomics: Application of Genomics to Uncultured Microorganisms.* 68, 669–685. doi:10.1128/MBR.68.4.669.
- Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, R245–R249. doi:10.1016/S1074-5521(98)90108-9.
- Helena and Härd, T. H. M. H. N. van den B. S. B. (2002). Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.* 11, 313–21. doi:10.1110/ps.22102.
- Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P., and Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7090–7094. doi:10.1073/pnas.92.15.7090.

- Henrissat, B., Vegetales, M., and Grenoble, F. (1991). A classification of glycosyl hydrolases based sequence similarities amino acid. *Biochem. J.* 280, 309–316. doi:10.1007/s007920050009.
- Hidaka, M., Fushinobu, S., Ohtsu, N., Motoshima, H., Matsuzawa, H., Shoun, H., et al. (2002a). Trimeric crystal structure of the glycoside hydrolase family 42 ??-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose. *J. Mol. Biol.* 322, 79–91. doi:10.1016/S0022-2836(02)00746-5.
- Hidaka, M., Fushinobu, S., Ohtsu, N., Motoshima, H., Matsuzawa, H., Shoun, H., et al. (2002b). Trimeric Crystal Structure of the Glycoside Hydrolase Family 42  $\beta$ -Galactosidase from *Thermus thermophilus* A4 and the Structure of its Complex with Galactose. *J. Mol. Biol.* 322, 79–91. doi:10.1016/S0022-2836(02)00746-5.
- Hidaka, Y., Hasegawa, M., Nakahara, T., and Hoshino, T. (1994). The entire population of *Thermus thermophilus* cells is always competent at any growth phase. *Biosci Biotechnol Biochem* 58, 1338–1339. doi:10.1271/bbb.58.1338.
- Hidalgo, A., Betancor, L., Moreno, R., Zafra, O., Cava, F., Fernández-Lafuente, R., et al. (2004). *Thermus thermophilus* as a cell factory for the production of a thermophilic Mn-dependent catalase which fails to be synthesized in an active form in *Escherichia coli*. *Appl. Environ. Microbiol.* 70, 3839–3844. doi:10.1128/AEM.70.7.3839-3844.2004.
- Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, a, and Sterner, R. (2001). Dissection of a (betaalpha)8-barrel enzyme into two folded halves. *Nat. Struct. Biol.* 8, 32–36. doi:10.1038/83021.
- Hong, J., Tamaki, H., Yamamoto, K., and Kumagai, H. (2003). Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. *Appl. Microbiol. Biotechnol.* 63, 42–50. doi:10.1007/s00253-003-1379-3.
- Hosono, S., Faruqi, a. F., Dean, F. B., Du, Y., Sun, Z., Wu, X., et al. (2003). Unbiased whole-genome amplification directly from clinical samples. *Genome Res.* 13, 954–964. doi:10.1101/gr.816903.
- Hu, Y., Fu, C., Huang, Y., Yin, Y., Cheng, G., Lei, F., et al. (2010). Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. *FEMS Microbiol. Ecol.* 72, 228–237. doi:10.1111/j.1574-6941.2010.00851.x.



- Ilmberger, N., Meske, D., Juergensen, J., Schulte, M., Barthen, P., Rabausch, U., et al. (2012). Metagenomic cellulases highly tolerant towards the presence of ionic liquids - Linking thermostability and halotolerance. *Appl. Microbiol. Biotechnol.* 95, 135–146. doi:10.1007/s00253-011-3732-2.
- Imachi, H., Sakai, S., Hirayama, H., Nakagawa, S., Nunoura, T., Takai, K., et al. (2008). *Exilispira thermophila* gen. nov., sp. nov., an anaerobic, thermophilic spirochaete isolated from a deep-sea hydrothermal vent chimney. *Int. J. Syst. Evol. Microbiol.* 58, 2258–2265. doi:10.1099/ijs.0.65727-0.
- Jaeger, K. E., Dijkstra, B. W., and Reetz, M. T. (1999). Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu. Rev. Microbiol.* 53, 315–351. doi:10.1146/annurev.micro.53.1.315.
- Jaeger, K., and Reetz, M. T. (1998). Microbial lipases form versatile tools for biotechnology. 16, 396–403.
- Jeanthon, C., Haridon, L., Banta, A., Reysenbach, A., and Prieur, D. (2002). nov ., a thermophilic , chemolithoautotrophic , sulfate-reducing bacterium isolated from a deep-sea hydrothermal vent at Guaymas Basin ., *Int. J. Syst. Evol. Microbiol.*, 765–772. doi:10.1099/ijs.0.02025-0.The.
- Jia, H., Li, Y., Liu, Y., Yan, Q., Yang, S., and Jiang, Z. (2012). Engineering a thermostable  $\beta$ -1,3-1,4-glucanase from *Paecilomyces thermophila* to improve catalytic efficiency at acidic pH. *J. Biotechnol.* 159, 50–5. doi:10.1016/j.jbiotec.2012.02.007.
- Kadnikov, V. V., Mardanov, A. V., Podosokorskaya, O. a., Gavrilov, S. N., Kublanov, I. V., Beletsky, A. V., et al. (2013). Genomic Analysis of *Melioribacter roseus*, Facultatively Anaerobic Organotrophic Bacterium Representing a Novel Deep Lineage within Bacteroidetes/Chlorobi Group. *PLoS One* 8. doi:10.1371/journal.pone.0053047.
- Kakirde, K. S., Parsley, L. C., and Liles, M. R. (2010). Size does matter: Application-driven approaches for soil metagenomics. *Soil Biol. Biochem.* 42, 1911–1923. doi:10.1016/j.soilbio.2010.07.021.
- Kakirde, K. S., Wild, J., Godiska, R., Mead, D. a, Wiggins, A. G., Goodman, R. M., et al. (2012). Gram negative shuttle BAC vector for heterologous expression of metagenomic libraries. *Gene* 475, 57–62. doi:10.1016/j.gene.2010.11.004.Gram.

- Kang, C.-H., Oh, K.-H., Lee, M.-H., Oh, T.-K., Kim, B. H., and Yoon, J.-H. (2011). A novel family VII esterase with industrial potential from compost metagenomic library. *Microb. Cell Fact.* 10, 41. doi:10.1186/1475-2859-10-41.
- Karnaouri, A. C., Topakas, E., and Christakopoulos, P. (2014). Cloning, expression, and characterization of a thermostable GH7 endoglucanase from *Myceliophthora thermophila* capable of high-consistency enzymatic liquefaction. *Appl. Microbiol. Biotechnol.* 98, 231–42. doi:10.1007/s00253-013-4895-9.
- Kellner, R., Hofmann, H., Barducci, A., Wunderlich, B., Nettels, D., and Schuler, B. (2014). Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proc. Natl. Acad. Sci. U. S. A.* 111, 13355–13360. doi:10.1073/pnas.1407086111.
- Kengen, S. W. M., Luesink, E. J., Stams, A. J. M., and Zehnder, A. J. B. (1993). Purification and characterization of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* 213, 305–312. doi:10.1111/j.1432-1033.1993.tb17763.x.
- Kim, B. S., Kim, S. Y., Park, J., Park, W., Hwang, K. Y., Yoon, Y. J., et al. (2007). Sequence-based screening for self-sufficient P450 monooxygenase from a metagenome library. *J. Appl. Microbiol.* 102, 1392–1400. doi:10.1111/j.1365-2672.2006.03169.x.
- Ko, K.-C., Rim, S.-O., Han, Y., Shin, B. S., Kim, G.-J., Choi, J. H., et al. (2012). Identification and characterization of a novel cold-adapted esterase from a metagenomic library of mountain soil. *J. Ind. Microbiol. Biotechnol.* 39, 681–9. doi:10.1007/s10295-011-1080-y.
- Kowalchuk, G. a., Speksnijder, A. G. C. L., Zhang, K., Goodman, R. M., and van Veen, J. a. (2007). Finding the Needles in the Metagenome Haystack. *Microb. Ecol.* 53, 475–485. doi:10.1007/s00248-006-9201-2.
- Koyama, Y., Hoshino, T., Tomizuka, N., and Furukawa, K. (1986). Genetic Transformation of the Extreme Thermophile *Thermus thermophilus* and of Other *Thermus* spp. *J. Bacteriol.* 166, 338–340.
- Krahe, M., Antranikian, G., and Markel, H. (1996). Fermentation of extremophilic microorganisms. *FEMS Microbiol. Rev.* 18, 271–285.

- Kroon, P. a., Williamson, G., Fish, N. M., Archer, D. B., and Belshaw, N. J. (2000). A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plant cell walls and binds crystalline cellulose contains a carbohydrate binding module. *Eur. J. Biochem.* 267, 6740–6752. doi:10.1046/j.1432-1327.2000.01742.x.
- Kudla, G., Murray, A. W., Tollervey, D., and Plotkin, J. B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324, 255–8. doi:10.1126/science.1170160.
- Kumar, H. D. and, and Swati, S. (2001). *Modern Concepts of Microbiology*, second revised ed. *Vikas Publ. House Pvt. Ltd., New Delhi*.
- Kwon, S.-W., Kim, B.-Y., Kim, W.-G., Yoo, K.-H., Yoo, S.-H., Son, J. –, et al. (2008). *Paludibacterium yongneupense* gen. nov., sp. nov., isolated from a wetland, Yongneup, in Korea. *Int. J. Syst. Evol. Microbiol.* 58, 190–194. doi:10.1099/ijs.0.64831-0.
- Lachmann, K. (2013). Charakterisierung von Fosmidklonen aus der vergleichenden Durchmusterung von metagenomischen Genbibliotheken in *Escherichia coli* und *Thermus thermophilus*.
- Laderman, K. a, Asada, K., Uemori, T., Mukai, H., Taguchi, Y., Kato, I., et al. (1993). Alpha-amylase from the hyperthermophilic archaebacterium *Pyrococcus furiosus*. Cloning and sequencing of the gene and expression in *Escherichia coli*. *J. Biol. Chem.* 268, 24402–24407. Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8226990](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8226990).
- Lam, K. N., Hall, M. W., Engel, K., Vey, G., Cheng, J., Neufeld, J. D., et al. (2014). Evaluation of a pooled strategy for high-throughput sequencing of cosmid clones from metagenomic libraries. *PLoS One* 9, e98968. doi:10.1371/journal.pone.0098968.
- Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. a, Sogin, M. L., and Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. U. S. A.* 82, 6955–6959. doi:10.1073/pnas.82.20.6955.
- Lang, D., Thoma, R., Henn-Sax, M., Sterner, R., and Wilmanns, M. (2000). Structural Evidence for Evolution of the beta /alpha Barrel Scaffold by Gene Duplication and Fusion. *Science* (80-. ). 289, 1546–1550. doi:http://dx.doi.org/10.1126/science.289.5484.1546.

- Lantz, S. E., Goedegebuur, F., Hommes, R., Kaper, T., Kelemen, B. R., Mitchinson, C., et al. (2010). Hypocrea jecorina CEL6A protein engineering. *Biotechnol. Biofuels* 3, 20. doi:10.1186/1754-6834-3-20.
- Lasken, R. S. (2009). Genomic DNA amplification by the multiple displacement amplification (MDA) method: Figure 1. *Biochem. Soc. Trans.* 37, 450–453. doi:10.1042/BST0370450.
- Lee, C., Yashiro, S., Dotson, D. L., Uzdavinyas, P., Iwata, S., Sansom, M. S. P., et al. (2014). Crystal structure of the sodium-proton antiporter NhaA dimer and new mechanistic insights. *J. Gen. Physiol.* 144, 529–44. doi:10.1085/jgp.201411219.
- Lee, Y. J., Wagner, I. D., Brice, M. E., Kevbrin, V. V., Mills, G. L., Romanek, C. S., et al. (2005). *Thermosediminibacter oceani* gen. nov., sp. nov. and *Thermosediminibacter litoriperuensis* sp. nov., new anaerobic thermophilic bacteria isolated from Peru Margin. *Extremophiles* 9, 375–383. doi:10.1007/s00792-005-0453-4.
- Leis, B. (2015a). Entwicklung von *Thermus thermophilus* als Wirt für (Meta)Genomanalysen. *Lehrstuhl für Mikrobiol.*
- Leis, B. (2015b). Entwicklung von *Thermus thermophilus* als Wirt für (Meta)Genomanalysen.
- Leis, B., Angelov, A., Mientus, M., Li, H., Pham, V. T. T., Lauinger, B., et al. (2015a). Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00275.
- Leis, B., Heinze, S., Angelov, A., Pham, V. T. T., Thürmer, A., Jebbar, M., et al. (2015b). Functional Screening of Hydrolytic Activities Reveals an Extremely Thermostable Cellulase from a Deep-Sea Archaeon. *Front. Bioeng. Biotechnol.* 3. doi:10.3389/fbioe.2015.00095.
- Li, H. Chromosome and Megaplasmid partitioning in *Thermus thermophilus* HB27.
- Li, L.-L., Taghavi, S., McCorkle, S. M., Zhang, Y.-B., Blewitt, M. G., Brunecky, R., et al. (2011). Bioprospecting metagenomics of decaying wood: mining for new glycoside hydrolases. *Biotechnol. Biofuels* 4, 23. doi:10.1186/1754-6834-4-23.
- Li, S., Xu, L., Hua, H., Ren, C., and Lin, Z. (2007). A set of UV-inducible autolytic vectors for high throughput screening. *J. Biotechnol.* 127, 647–652. doi:10.1016/j.jbiotec.2006.07.030.

- Liaw, S.-T. (2010). Genetics and genomics in general practice. *Aust. Fam. Physician* 39, 689–691.
- Liebl, W. (2004). Genomics taken to the extreme. *Nat. Biotechnol.* 22, 524–525. doi:10.1038/nbt0504-524.
- Liebl, W., Angelov, A., Juergensen, J., Chow, J., Loeschcke, A., Drepper, T., et al. (2014a). Alternative hosts for functional ( meta ) genome analysis. 8099–8109. doi:10.1007/s00253-014-5961-7.
- Liebl, W., Angelov, A., Juergensen, J., Chow, J., Loeschcke, A., Drepper, T., et al. (2014b). Alternative hosts for functional (meta)genome analysis. *Appl. Microbiol. Biotechnol.* 98, 8099–8109. doi:10.1007/s00253-014-5961-7.
- Liebl, W., Kloos, W. E., and Ludwig, W. (2002). Plasmid-borne macrolide resistance in *Micrococcus luteus*. *Microbiology* 148, 2479–2487. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12177341>.
- Linko, Y., Lamsa, M., Huhatala, A., and Linko, P. (1994). Lipase catalyzed transesterification of rapeseed oil snf 2-ethyl-1-hexanol. *JAOCS*, 1411–1414.
- Littlechild, J. a. (2015). Enzymes from Extreme Environments and Their Industrial Applications. *Front. Bioeng. Biotechnol.* 3, 1–9. doi:10.3389/fbioe.2015.00161.
- Liu, Y., Dun, B., Shi, P., Ma, R., Luo, H., and Bai, Y. (2015). A Novel GH7 Endo-  $\beta$  -1 , 4- Glucanase from *Neosartorya fischeri* P1 with Good Thermostability , Broad Substrate Specificity and Potential Application in the Brewing Industry. 1–12. doi:10.1371/journal.pone.0137485.
- Lymar, E. S., Li, B., and Renganathan, V. (1995). Purification and Characterization of a Cellulose-Binding (beta)-Glucosidase from Cellulose-Degrading Cultures of *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* 61, 2976–2980.
- Ma, J.-Y., Cao, H.-L., Tan, H.-D., Hu, X.-J., Liu, W.-J., Du, Y.-G., et al. (2015). Cloning, Expression, Characterization, and Mutagenesis of a Thermostable Exoinulinase From *Kluyveromyces cicerisporus*. *Appl. Biochem. Biotechnol.* doi:10.1007/s12010-015-1864-z.
- Manco, G., Adinolfi, E., Pisani, F. M., Ottolina, G., Carrea, G., and Rossi, M. (1998). Overexpression and properties of a new thermophilic and thermostable esterase from

*Bacillus acidocaldarius* with sequence similarity to hormone-sensitive lipase subfamily. *Biochem. J.* 332, 203–212. doi:10.1042/bj3320203.

Marrero, G., Schneider, K. L., Jenkins, D. M., and Alvarez, A. M. (2013). Phylogeny and classification of *Dickeya* based on multilocus sequence analysis. *Int. J. Syst. Evol. Microbiol.* 63, 3524–3539. doi:10.1099/ijs.0.046490-0.

Martinez, A., Kolvek, S. J., Yip, C. L. T., Hopke, J., Brown, K. a., MacNeil, I. a., et al. (2004). Genetically Modified Bacterial Strains and Novel Bacterial Artificial Chromosome Shuttle Vectors for Constructing Environmental Libraries and Detecting Heterologous Natural Products in Multiple Expression Hosts. *Appl. Environ. Microbiol.* 70, 2452–2463. doi:10.1128/AEM.70.4.2452-2463.2004.

Masepohl and Hallenbeck (2010). No Title Nitrogen and molybdenum control of nitrogen fixation in the phototrophic bacterium *Rhodobacter capsulatus*. *Adv. Exp. Med. Biol.* 675, 49–70.

Matsuzawa, T., Saito, Y., and Yaoi, K. (2014). Key amino acid residues for the endo-processive activity of GH74 xyloglucanase. *FEBS Lett.* 588, 1731–8. doi:10.1016/j.febslet.2014.03.023.

Mawadza, C., Hatti-Kaul, R., Zvauya, R., and Mattiasson, B. (2000). Purification and characterization of cellulases produced by two *Bacillus* strains. *J. Biotechnol.* 83, 177–187. doi:10.1016/S0168-1656(00)00305-9.

McGrath, C. E., Vuong, T. V., and Wilson, D. B. (2009). Site-directed mutagenesis to probe catalysis by a *Thermobifida fusca*  $\alpha$ -1,3-glucanase (Lam81A). *Protein Eng. Des. Sel.* 22, 375–382. doi:10.1093/protein/gzp015.

Michelin, M., Silva, T. M., Benassi, V. M., Peixoto-Nogueira, S. C., Moraes, L. a B., Le??o, J. M., et al. (2010). Purification and characterization of a thermostable  $\alpha$ -amylase produced by the fungus *Paecilomyces variotii*. *Carbohydr. Res.* 345, 2348–2353. doi:10.1016/j.carres.2010.08.013.

Mientus, M., Brady, S., Angelov, A., Zimmermann, P., Wenheuer, B., Schuldes, J., et al. (2013). Thermostable Xylanase and  $\beta$ -Glucanase Derived from the Metagenome of the Avachinsky Crater in Kamchatka (Russia). *Curr. Biotechnol.* 2, 284–294.

- Miller, D. N., Bryant, J. E., Madsen, E. L., and Ghiorse, W. C. (1999). Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl. Environ. Microbiol.* 65, 4715–4724.
- Miyanaaga, A., Koseki, T., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2004). Crystal structure of a family 54 ??-L-arabinofuranosidase reveals a novel carbohydrate-binding module that can bind arabinose. *J. Biol. Chem.* 279, 44907–44914. doi:10.1074/jbc.M405390200.
- Miyazaki, T., Ichikawa, M., Iino, H., Nishikawa, A., and Tonozuka, T. (2015). Crystal structure and substrate-binding mode of GH63 mannosylglycerate hydrolase from *Thermus thermophilus* HB8. *J. Struct. Biol.* 190, 21–30. doi:10.1016/j.jsb.2015.02.006.
- Moran, M. A., Torsvik, V. L., Torsvik, T., and Hodson, R. E. (1993). Direct extraction and purification of rRNA for ecological studies. *Appl. Environ. Microbiol.* 59, 915–918.
- Moreno, R., Haro, A., Castellanos, A., Moreno, R., Haro, A., and Castellanos, A. (2005). High-Level Overproduction of His-Tagged Tth DNA Polymerase in *Thermus thermophilus*. High-Level Overproduction of His-Tagged Tth DNA Polymerase in *Thermus thermophilus*. 71, 1–4. doi:10.1128/AEM.71.1.591.
- Morikawa, M., Izawa, Y., Rashid, N., Hoaki, T., and Imanaka, T. (1994). Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic *Pyrococcus* sp. *Appl. Environ. Microbiol.* 60, 4559–4566.
- Mujacic, M., Bader, M. W., and Baneyx, F. (2003). *Escherichia coli* Hsp31 functions as a holding chaperone that cooperates with the DnaK-DnaJ-GrpE system in the management of protein misfolding under severe stress conditions. *Mol. Microbiol.* 51, 849–859. doi:10.1046/j.1365-2958.2003.03871.x.
- Nakamura, H., Kubota, H., Kono, T., Isogai, A., and Onabe, F. (2001). Modification of pulp properties by cellulase treatment and application of cellulase to wastepaper deinking and mechanical pulp refining. 2–5.
- Nanda, V. (2016). Getting to the bottom of the TIM barrel. *Nat. Chem. Biol.* 12, 2–3. doi:10.1038/nchembio.1987.
- Nanmori, T., Watanabe, T., Shinke, R., Kohno, a, and Kawamura, Y. (1990). Purification and properties of thermostable xylanase and beta-xylosidase produced by a newly isolated

- Bacillus stearothermophilus* strain. *J. Bacteriol.* 172, 6669–72. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=210778&tool=pmcentrez&rendertype=abstract>.
- Niederberger, T. D., Götz, D. K., McDonald, I. R., Ronimus, R. S., and Morgan, H. W. (2006). *Ignisphaera aggregans* gen. nov., sp. nov., a novel hyperthermophilic crenarchaeote isolated from hot springs in Rotorua and Tokaanu, New Zealand. *Int. J. Syst. Evol. Microbiol.* 56, 965–971. doi:10.1099/ijs.0.63899-0.
- Ohtani, N., Tomita, M., and Itaya, M. (2010). An extreme thermophile, *Thermus thermophilus*, is a polyploid bacterium. *J. Bacteriol.* 192, 5499–505. doi:10.1128/JB.00662-10.
- Oshima, T., and Imahori, K. (1974). 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. doi:10.1099/00207713-24-1-102.
- Padan, E., Venturi, M., Gerchman, Y., and Dover, N. (2001). Na<sup>+</sup>/H<sup>+</sup> antiporters. *Biochim. Biophys. Acta - Bioenerg.* 1505, 144–157. doi:10.1016/S0005-2728(00)00284-X.
- Pandey, A., Benjamin, S., Soccol, C., Nigam, P., Krieger, N., and Soccol, V. (1999). The realm of microbial lipases in biotechnology: a review. *Biotechnol. Appl. Biochem.* 29, 119–131.
- Parry, N. J., Beaver, D. E., Owen, E., Vandenberghe, I., Van Beeumen, J., and Bhat, M. K. (2001). Biochemical characterization and mechanism of action of a thermostable beta-glucosidase purified from *Thermoascus aurantiacus*. *Biochem. J.* 353, 117–127. doi:10.1042/0264-6021:3530117.
- Peters, B., Mientus, M., Kostner, D., Junker, A., Liebl, W., and Ehrenreich, A. (2013). Characterization of membrane-bound dehydrogenases from *Gluconobacter oxydans* 621H via whole-cell activity assays using multideletion strains. *Appl. Microbiol. Biotechnol.* 97, 6397–6412. doi:10.1007/s00253-013-4824-y.
- Pham, V. D., Palden, T., and DeLong, E. F. (2007). Large-scale screens of metagenomic libraries. *J. Vis. Exp.*, 201. doi:10.3791/201.
- Popowska, M., and Krawczyk-Balska, A. (2013). Broad-host-range IncP-1 plasmids and their resistance potential. *Front. Microbiol.* 4, 1–8. doi:10.3389/fmicb.2013.00044.



- R. M. Ray, H. C. Aldrich, A. C. Wilkie, L. O. Ingram, K. T. S. and, and Healy, F. G. (1995). Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic.pdf. *Appl. Microbiol. Biotechnol.* 43, 667–674.
- Ramírez, S., Moreno, R., Zafra, O., Castán, P., Vallés, C., and Berenguer, J. (2000). Two nitrate/nitrite transporters are encoded within the mobilizable plasmid for nitrate respiration of *Thermus thermophilus* HB8. *J. Bacteriol.* 182, 2179–2183. doi:10.1128/JB.182.8.2179-2183.2000.
- Reddy, G. S. N., Matsumoto, G. I., Schumann, P., Stackerbrandt, E., and Shivaji, S. (2004). Psychrophilic pseudomonads from Antarctica: *Pseudomonas antarctica* sp. nov., *Pseudomonas meridiana* sp. nov. and *Pseudomonas proteolytica* sp. nov. *Int. J. Syst. Evol. Microbiol.* 54, 713–719. doi:10.1099/ij.s.0.02827-0.
- Van Rensburg, P., Van Zyl, W. H., and Pretorius, I. S. (1997). Over-expression of the *Saccharomyces cerevisiae* exo- $\beta$ -1,3-glucanase gene together with the *Bacillus subtilis* endo- $\beta$ -1,3-1,4-glucanase gene and the *Butyrivibrio fibrisolvens* endo- $\beta$ -1,4-glucanase gene in yeast. *J. Biotechnol.* 55, 43–53. doi:10.1016/S0168-1656(97)00059-X.
- Rhee, J.-K., Ahn, D.-G., Kim, Y.-G., and Oh, J.-W. (2005). New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library. *Appl. Environ. Microbiol.* 71, 817–25. doi:10.1128/AEM.71.2.817-825.2005.
- Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., et al. (2000). Cloning the soil metagenome: A strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66, 2541–2547. doi:10.1128/AEM.66.6.2541-2547.2000.
- Rouvinen, J., Bergfors, T., Teeri, T., and Knowles, J. K. C. (1967). Three-dimensional Structure of Tosyl-a-chymotrypsin. *Nature*. doi:10.1038/214652a0.
- Ruttersmith, L. D., and Daniel, R. M. (1991). Thermostable cellobiohydrolase from the thermophilic eubacterium *Thermotoga* sp. strain FjSS3-B.1. Purification and properties. *Biochem. J.* 277 ( Pt 3, 887–890.
- Sabehi, G., and Béjà, O. (2013). *Preparation of BAC Libraries from Marine Microbial Populations*. 1st ed. Elsevier Inc. doi:10.1016/B978-0-12-407863-5.00006-X.

- Sangwan, P., Chen, X., Hugenholtz, P., and Janssen, P. H. (2004). *Chthoniobacter flavus* gen. nov., sp. nov., the first pure-culture representative of subdivision two, Spartobacteria classis nov., of the phylum Verrucomicrobia. *Appl. Environ. Microbiol.* 70, 5875–5881. doi:10.1128/AEM.70.10.5875-5881.2004.
- Sastry, M. S. R., Quigley, P. M., Hol, W. G. J., and Baneyx, F. (2004). The linker-loop region of *Escherichia coli* chaperone Hsp31 functions as a gate that modulates high-affinity substrate binding at elevated temperatures. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8587–8592. doi:10.1073/pnas.0403033101.
- Schipper, C., Hornung, C., Bijtenhoorn, P., Quitschau, M., Grond, S., and Streit, W. R. (2009). Metagenome-derived clones encoding two novel lactonase family proteins involved in biofilm inhibition in *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* 75, 224–233. doi:10.1128/AEM.01389-08.
- Schlegel, A., Danot, O., Richet, E., Ferenci, T., and Boos, W. (2006). The N Terminus of the *Escherichia coli* Transcription Activator MalT Is the Domain of Interaction with MalY. 184, 3069–3077. doi:10.1128/JB.184.11.3069.
- Schmeisser, C., Steele, H., and Streit, W. R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biotechnol.* 75, 955–962. doi:10.1007/s00253-007-0945-5.
- Schmidt, T. M., DeLong, E. F., and Pace, N. R. (1991). Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* 173, 4371–4378.
- Schuldes, J. (2008). Charakterisierung und Nutzung der mikrobiellen Diversität extremer Habitats der Dissertation.
- Schwarzenlander, C., and Averhoff, B. (2006). Characterization of DNA transport in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS J.* 273, 4210–8. doi:10.1111/j.1742-4658.2006.05416.x.
- SD. Baird, M.A. Hefford, D. A. Johnson, W. L. S. M. Y. and V. L. S. (1990). The Glu residue in the conserved Asn-Glu-Pro sequence of two highly divergent endo  $\beta$ -1,4-glucanases is essential for enzymatic activity. 169, 1028–1034.

- Sekhar, A., Lam, H. N., and Cavagnero, S. (2012). Protein folding rates and thermodynamic stability are key determinants for interaction with the Hsp70 chaperone system. *Protein Sci.* 21, 1489–1502. doi:10.1002/pro.2139.
- Shariff, F. M., Rahman, R. N. Z. R. A., Basri, M., and Salleh, A. B. (2011). A newly isolated thermostable lipase from bacillus sp. *Int. J. Mol. Sci.* 12, 2917–2934. doi:10.3390/ijms12052917.
- Shoseyov, O., Shani, Z., and Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* 70, 283–95. doi:10.1128/MMBR.00028-05.
- Simon, C., and Daniel, R. (2009a). Achievements and new knowledge unraveled by metagenomic approaches. *Appl. Microbiol. Biotechnol.* 85, 265–276. doi:10.1007/s00253-009-2233-z.
- Simon, C., and Daniel, R. (2009b). Achievements and new knowledge unraveled by metagenomic approaches. 265–276. doi:10.1007/s00253-009-2233-z.
- Simon, C., and Daniel, R. (2011). Metagenomic Analyses: Past and Future Trends. *Appl. Environ. Microbiol.* 77, 1153–1161. doi:10.1128/AEM.02345-10.
- Sonne-hansen, J., and Westermann, P. (1999). Kinetics of Sulfate and Hydrogen Uptake by the Thermophilic Sulfate-Reducing Bacteria *Thermodesulfobacterium* sp . Strain. *Appl. Environ. Microbiol.* 65, 1304–1307.
- Sookkheo, B., Sinchaikul, S., Phutrakul, S., and Chen, S. T. (2000). Purification and characterization of the highly thermostable proteases from *Bacillus stearothermophilus* TLS33. *Protein Expr. Purif.* 20, 142–51. doi:10.1006/prev.2000.1282.
- Stålbrand, H., Saloheimo, a, Vehmaanperä, J., Henrissat, B., and Penttilä, M. (1995). Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei* beta-mannanase gene containing a cellulose binding domain. *Appl. Environ. Microbiol.* 61, 1090–7. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=167364&tool=pmcentrez&rendertype=abstract>.
- Sterner, R., and Liebl, W. (2001). Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 36, 39–106. doi:10.1080/20014091074174.

- Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., and Davies, G. J. (1996). Structure of the *Fusarium oxysporum* endoglucanase I with a nonhydrolyzable substrate analogue: Substrate distortion gives rise to the preferred axial orientation for the leaving group. *Biochemistry* 35, 15280–15287. doi:10.1021/bi961946h.
- Taupp, M., Hallam, S. J., and Mewis, K. (2011). A high throughput screen for biomining cellulase activity from metagenomic libraries. *J. Vis. Exp.*, 2–5. doi:10.3791/2461.
- Tegel, H., Tourle, S., Ottosson, J., and Persson, A. (2010). Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expr. Purif.* 69, 159–167. doi:10.1016/j.pep.2009.08.017.
- Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 60, 523–533. doi:10.1007/s00253-002-1158-6.
- Thebti, W., Riahi, Y., and Belhadj, O. (2016). Purification and Characterization of a New Thermostable , Haloalkaline , Solvent Stable , and Detergent Compatible Serine Protease from *Geobacillus toebii* Strain LBT 77. 2016. doi:10.1155/2016/9178962.
- Thoma, J. a., and Koshland, D. E. (1960). Stereochemistry of enzyme, substrate, and products during beta-amylase action. *J. Biol. Chem.* 235, 2511–2517.
- Tirawongsaroj, P., Sriprang, R., Harnpicharnchai, P., Thongaram, T., Champreda, V., Tanapongpipat, S., et al. (2008). Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. *J. Biotechnol.* 133, 42–49. doi:10.1016/j.jbiotec.2007.08.046.
- Tishkov, V. I., Gusakov, A. V., Cherkashina, A. S., and Sinitsyn, A. P. (2013). Engineering the pH-optimum of activity of the GH12 family endoglucanase by site-directed mutagenesis. *Biochimie* 95, 1704–1710. doi:10.1016/j.biochi.2013.05.018.
- Tomme P., Warren R.A., Miller R.C., Jr. Kilburn D.G., Gilkes, N. . (1995). Cellulose-binding domains: classification and properties. *Enzym. Degrad. Insoluble Polysacch.*, 142–163.
- Torsvik, V., Goksoyr, J., and Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* 56, 782–787.
- Traore, M.K. and Buschle-Diller, G. (2000). Environmentally freindly scouring processes. *Text. Chem.* 32, 40–43.

- Tripathy, S., Padhi, S. K., Sen, R., Maji, U., Samanta, M., Mohanty, S., et al. (2016). Draft Genome Sequence of *Brevibacillus borstelensis* cifa\_chp40, a Thermophilic Strain Having Biotechnological Importance. *J. Genomics* 4, 4–6. doi:10.7150/jgen.14036.
- Tuohy, M. G., Walsh, D. J., Murray, P. G., Claeysens, M., Cuffe, M. M., Savage, A. V., et al. (2002). Kinetic parameters and mode of action of the cellobiohydrolases produced by *Talaromyces emersonii*. *Biochim. Biophys. Acta* 1596, 366–80. doi:10.1016/S0167-4838(01)00308-9.
- Uchiyama, T., and Miyazaki, K. (2010). Product-Induced Gene Expression, a Product-Responsive Reporter Assay Used To Screen Metagenomic Libraries for Enzyme-Encoding Genes. *Appl. Environ. Microbiol.* 76, 7029–7035. doi:10.1128/AEM.00464-10.
- Ullmann, Jacob, and Monod (1967). Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli*. *J Mol Biol* 24, 339–343. doi:10.1016/0022-2836(67)90341-5.
- Venter, J. C. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science (80-. )*. 304, 66–74. doi:10.1126/science.1093857.
- Walsh, G. a., Power, R. F., and Headon, D. R. (1994). Enzymes in the animal-feed industry. *Trends Food Sci. Technol.* 5, 81–87. doi:10.1016/0924-2244(94)90242-9.
- Wang, B., Wang, A., Cao, Z., and Zhu, G. (2015). Characterization of a novel highly thermostable esterase from the gram-positive soil bacterium *Streptomyces lividans* TK64. *Biotechnol. Appl. Biochem.*, 6–8. doi:10.1002/bab.1465.
- Wang, T., and Leyh, T. S. (2012). Three-stage assembly of the cysteine synthase complex from *Escherichia coli*. *J. Biol. Chem.* 287, 4360–4367. doi:10.1074/jbc.M111.288423.
- Warren, R. L., Freeman, J. D., Levesque, R. C., Smailus, D. E., Flibotte, S., and Holt, R. a. (2008). Transcription of foreign DNA in *Escherichia coli*. *Genome Res.* 18, 1798–1805. doi:10.1101/gr.080358.108.
- Welply, J. K., Fowler, a. V., and Zabin, I. (1981). beta-Galactosidase alpha-complementation. Overlapping sequences. *J. Biol. Chem.* 256, 6804–6810.
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* 95, 6578–6583. doi:10.1073/pnas.95.12.6578.

- Williamson, L. L., Borlee, B. R., Schloss, P. D., Guan, C., Allen, H. K., and Handelsman, J. (2005). Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl. Environ. Microbiol.* 71, 6335–44. doi:10.1128/AEM.71.10.6335-6344.2005.
- Winterhalter, C., and Liebl, W. (1995). Two Extremely Thermostable Xylanases of the Hyperthermophilic Bacterium *Thermotoga maritima* MSB8. 61, 1810–1815.
- Wittenberg, S. (2009). Metagenomanalysen von zwei Habitaten mit (heim-)cellulolytischen mikrobiellen Gemeinschaften.pdf.
- Xu, B., Xu, W., Li, J., Dai, L., Xiong, C., Tang, X., et al. (2015a). Metagenomic analysis of the *Rhinopithecus bieti* fecal microbiome reveals a broad diversity of bacterial and glycoside hydrolase profiles related to lignocellulose degradation. *BMC Genomics* 16, 174. doi:10.1186/s12864-015-1378-7.
- Xu, X., Li, J., Zhang, W., Huang, H., Shi, P., Luo, H., et al. (2015b). A neutral thermostable 1,4-glucanase from *Humicola insolens* Y1 with potential for applications in various industries. *PLoS One* 10, 1–13. doi:10.1371/journal.pone.0124925.
- Xue, Y., Zhang, X., Zhou, C., Zhao, Y., Cowan, D. a., Heaphy, S., et al. (2006). *Caldalkalibacillus thermarum* gen. nov., sp. nov., a novel alkalithermophilic bacterium from a hot spring in China. *Int. J. Syst. Evol. Microbiol.* 56, 1217–1221. doi:10.1099/ijs.0.64105-0.
- Yamamura, A., Okada, A., Kameda, Y., Ohtsuka, J., Nakagawa, N., Ebihara, A., et al. (2009). Structure of TTHA1623, a novel metallo-lactamase superfamily protein from *Thermus thermophilus* HB8. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 455–459. doi:10.1107/S174430910901361X.
- Yang, W., Cao, H., Xu, L., Zhang, H., and Yan, Y. (2015). A novel eurythermic and thermostable lipase LipM from *Pseudomonas moraviensis* M9 and its application in the partial hydrolysis of algal oil. *BMC Biotechnol.* 15, 94. doi:10.1186/s12896-015-0214-0.
- Yeoman, C. J. (2015). Thermostable Enzymes as Biocatalysts in the Biofuel Industry. 33, 395–401. doi:10.1038/nbt.3121.ChIP-nexus.
- Yin, Y. R., Zhang, F., Hu, Q. W., Xian, W. D., Hozzein, W. N., Zhou, E. M., et al. (2015). Heterologous expression and characterization of a novel halotolerant, thermostable, and

- alkali-stable GH6 endoglucanase from *Thermobifida halotolerans*. *Biotechnol. Lett.* 37, 857–862. doi:10.1007/s10529-014-1742-8.
- Yoshida, K., Imaizumi, N., Kaneko, S., Kawagoe, Y., Tagiri, A., Tanaka, H., et al. (2006). Carbohydrate-binding module of a rice endo- $\alpha$ -1,4-glycanase, OsCel9A, expressed in auxin-induced lateral root primordia, is post-translationally truncated. *Plant Cell Physiol.* 47, 1555–1571. doi:10.1093/pcp/pcl021.
- Yun, J., Kang, S., Park, S., Yoon, H., Kim, M., Heu, S., et al. (2005). Yun J , Kang S , Park S , Yoon H , Kim MJ , Hew S et al .. Characterization of a novel amyolytic enzyme encoded by a gene from a soil derived metagenomic library . *Appl Environ Microbiol ...* Characterization of a Novel Amyolytic Enzyme Encoded by a Gene. 70, 7229–7235. doi:10.1128/AEM.70.12.7229.
- Zafra, O., Ramírez, S., Castán, P., Moreno, R., Cava, F., Vallés, C., et al. (2002). A cytochrome c encoded by the nar operon is required for the synthesis of active respiratory nitrate reductase in *Thermus thermophilus*. *FEBS Lett.* 523, 99–102. doi:S0014579302029538 [pii].
- Zare, H., Sangurdekar, D., Srivastava, P., Kaveh, M., and Khodursky, a B. (2009). Reconstruction of *Escherichia coli* transcriptional regulatory networks via regulon-based associations. *BMC Syst. Biol.* 3, 39. doi:10.1186/1752-0509-3-39.
- Zerbs, S., Giuliani, S., and Collart, F. (2014). *Small-Scale Expression of Proteins in E. coli*. 1st ed. Elsevier Inc. doi:10.1016/B978-0-12-420070-8.00011-8.
- Zhou, C., Xue, Y., and Ma, Y. (2015). Evaluation and directed evolution for thermostability improvement of a GH 13 thermostable  $\alpha$ -glucosidase from *Thermus thermophilus* TC11. *BMC Biotechnol.* 15, 97. doi:10.1186/s12896-015-0197-x.
- Zhou, L., Yeung, K, Yuen, C. (2001). Combined cellulase ans wrinkle-free treatment on cotton fabric. 18, 11–15.
- Zverlov, V., Mahr, S., Riedel, K., and Bronnenmeier, K. (1998a). Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile “*Anaerocellum thermophilum*” with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology* 144, 457–465. doi:10.1099/00221287-144-2-457.

Zverlov, V., Piotukh, K., Dakhova, O., Velikodvorskaya, G., and Borriss, R. (1996). The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. *Appl. Microbiol. Biotechnol.* 45, 245–247. doi:10.1007/s002530050678.

Zverlov, V. V., Velikodvorskaya, G. V., Schwarz, W. H., Bronnenmeier, K., Kellermann, J., and Staudenbauer, W. L. (1998b). Multidomain structure and cellulosomal localization of the *Clostridium thermocellum* cellobiohydrolase CbhA. *J. Bacteriol.* 180, 3091–3099.



## 9 PUBLIKATIONEN UND KONFERENZEN

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung des Lehrstuhls für Mikrobiologie vorab veröffentlicht.

12.2017 Angelov\* A., Pham\* V. T. T., Übelacker M., Brady S., Leis B., Pill N., Brolle J., Mechelke M., Mörch M., Henrissat B., Liebl W. (\* equally contributing authors). A metagenome-derived thermostable  $\beta$ -glucanase with an unusual module architecture which defines the new glycoside hydrolase family GH148. *Sci Rep*, 2017.

Sonstige Veröffentlichungen während der Promotion

09.2017 Zhou J., Lewin A., Pham V. T., Haugen T., El Zeiny M., Wentzel A., Liles M.. Novel Archaeal Thermostable Cellulases from an Oil Reservoir Metagenome. *AMB Express*, 2017 Sep 29; 7(1)

04.2015 Leis B., Angelov A., Mientus M., Li H., Pham V. T., Launinger B., Liebl W. Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. *Front Microbiol.* 2015; 6:275

04.2015 Li H., Angelov A., Pham V. T., Leis B., Liebl W. Characterization of chromosomal and megaplasmid partitioning loci in *Thermus thermophilus* HB27. *BMC Genomics*, 2015 16:37

02.2014 Panitz J.C.; Zverlov V.V., Pham V. T., Stürzl S., Schieder D., Schwarz W.H. Isolation of a solventogenic *Clostridium sp. strain*: fermentation of glycerol to n-butanol, analysis of the bcs operon region and its potential regulatory elements. *Syst Appl Microbiol.*, 2014, Vol. 34, Issue 1

Konferenz:

99th Annual Meeting of the Southeastern Branch of the (ASM)  
Auburn, USA im November.2013

## **10 Eigenständigkeitserklärung**

Ich erkläre hiermit, dass ich meine Dissertation mit dem Titel „Charakterisierung thermostabiler Enzyme aus funktionell durchforsteten metagenomischen Genbanken“ (englischer Titel: „Characterization of thermostable enzymes obtained from functional screenings of metagenomic libraries“) selbstständig und nur mit den angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, wurden durch Angaben von Quellen als Entlehung kenntlich gemacht.

Freiburg, den 11.12.2017

Pham, Vu Thuy Trang

## 11 Danksagung

Als erstes möchte ich mich bei **Herrn Prof. Dr. Wolfgang Liebl** für die Themenstellung und Begutachtung meiner Disseratation bedanken.

Gleich im Anschluss geht ein großer Dank an meine gesamte Arbeitsgruppe, **Angel, Maria, Bene, Li, Tony, Matthias, Max und die Studenten**, für die produktive und unvergessliche Zeit im Labor. Es waren die Rituale wie Punkt-11 Uhr-Mittagessen, Proteinshakes und Kaffee trinken, Labordiskussionen, excessive Putztage und die gute Musik, die das Team und die Stimmung so einzigartig gemacht haben.

Ein besonderer Dank gebührt meinen Praktikanten, Bacheloranden und Masteranden **Nicki, Sarah, Katharina, Judith, Alexandra und Bettina** für ihre fleißige Arbeit und die erfrischende Abwechslung im Doktoranden-Alltag.

Der lieben **Konny** danke ich für ihre Arbeit, ihr stets offenes Ohr und ihre guten Ratschläge über die vielen Jahre.

Ich möchte **allen Kollegen und Kolleginnen am Lehrstuhl für Mikrobiologie** für die gute Arbeitsathmosphäre und Hilfsbereitschaft danken.

Bei **Claudia, Daniela, Martin, Bene, Maria und Raena** bedanke ich mich die Korrektur meiner Dissertation und meines Papers.

Aus tiefsten Herzen möchte ich meiner besten Freundin **Maria** danken. Du bist die gute Seele des Lehrstuhls und ich danke dir für deine enorme Unterstützung im Laboralltag und deine Hilfestellung in allen Lebensphasen.

Meine lieben Freunde **Sabrina, Steffi, Martin, Angi, Mali, Susi, Gitti, Beate, Ingi, Raena und Markus** haben mich durch das Auf und Ab der Doktorzeit getragen und mich mit aufbauenden Worten und Taten unterstützt.

Besten Dank an meine beiden Mentoren **Christian** und **Claudia**, die immer nur einen Anruf (früher nur eine Tür ☺) weit entfernt sind.

Lieber **Julian**, ich danke dir für die schönen bewussten Augenblicke und den kompletten Ausgleich in meinem turbulenten Leben.

Mein größter Dank gebührt meiner **Phamily Vu**. Ich bin sehr dankbar für all eure Weisheiten und den unendlichen Support in meinem Leben. Cảm ơn cha và mẹ cho mọi thứ trong cuộc sống của con. **Papa, Mama, Jenny und Josefine** ich liebe euch sehr und widme euch diese Arbeit.