

**Analyse von Regulationsnetzwerken der
Extracytoplasmic Function (ECF)-Sigmafaktoren in
*Corynebacterium glutamicum***

DISSERTATION

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)
der Fakultät für Biologie der Universität Bielefeld

vorgelegt von
Tobias Busche
aus Höxter

Bielefeld, im Juli 2013

für Jessie, Joline und Raphael

Inhaltsverzeichnis

Abbildungsverzeichnis	- 8 -
Tabellenverzeichnis.....	- 10 -
Abkürzungsverzeichnis	- 11 -
I. Zusammenfassung	1
II. Einleitung	3
1. Die Gattung Corynebakterium	3
1.1. Taxonomie der Corynebakterien.....	3
1.2. Biotechnologische und wirtschaftliche Relevanz von <i>Corynebacterium glutamicum</i> ..	4
2. Regulation bakterieller Genexpression.....	6
2.1. Bakterielle Transkription	6
2.2. Regulation der Transkription	8
2.3. Klassifizierung und Aufbau von Sigmafaktoren	10
2.4. ECF-Sigmafaktoren	12
2.5. Anti-Sigmafaktoren	15
3. Die Sigmafaktoren von <i>Corynebacterium glutamicum</i>	17
3.1. Der Hauptsigmafaktor SigA.....	17
3.2. Der nicht-essentielle Sigmafaktor SigB	18
3.3. ECF-Sigmafaktoren in <i>Corynebacterium glutamicum</i>	20
3.4. Der ECF-Sigmafaktor SigH.....	21
3.5. Der ECF-Sigmafaktor SigM.....	22
3.6. Der ECF-Sigmafaktoren SigE	23
3.7. Die ECF-Sigmafaktoren SigC und SigD	25
4. Klassische Methoden zur Analyse von Sigmafaktornetzwerken	26
5. RNA-Sequenzierung	27
5.1. Bioinformatische Analyse der RNA-Seq-Daten	30
5.1.1. <i>Mapping</i> sequenzierter <i>Reads</i> durch SARUMAN	30
5.1.2. RNA-Seq-Datenvisualisierung durch VAMP	30

6. Run-Off-Transcription/Microarray-Analysis (ROMA)	31
7. Zielsetzung der Arbeit	34
8. Bisherige Publikationen	35
III. Material und Methoden	36
1. Bakterienstämme und Plasmide	36
1.1. Bakterienstämme	36
1.2. Plasmide	37
2. Primer	38
3. Enzyme, Chemikalien und andere Materialien	41
3.1. Enzyme, Marker	41
3.2. Chemikalien	42
3.3. Kits	43
3.4. Materialien	44
3.5. Geräte und Apparaturen	45
3.6. Software	46
4. Medien und Zusätze	47
4.1. Nährmedien	47
4.2. Zusätze zu den Nährmedien	48
5. Verwendete Puffer und Lösungen	49
5.1. Elektrophorese	49
5.2. PCR	50
5.3. DNA-Transfertechniken	50
5.4. Proteinaufreinigung IMPACT-CN-System	51
5.5. 1D SDS-Gelelektrophorese und Proteinfärbung	52
5.6. Tryptischer Verdau	53
5.7. cDNA-Synthese und -Markierung für <i>Microarrays</i>	54
5.8. Lösungen und Puffer für <i>Microarray</i> -Hybridisierung	54
5.9. Northern Blot Analyse (Gel, Blot und Hybridisierung)	55
5.10. RNA-Seq	57
5.11. <i>Run-Off</i> -Transkription	57
5.12. <i>Run-Off</i> -Puffer	57
5.13. Gibson- <i>Assembly</i>	58

6. Kultivierung von Bakterien	59
6.1. Anzucht und Lagerung von <i>E. coli</i> Zellen	59
6.2. Anzucht und Lagerung von <i>C. glutamicum</i> Zellen	59
6.3. Dauerhafte Stammkonservierung - Herstellung von Glycerinkulturen	59
6.4. Bestimmung des Bakterientiters	60
6.5. Stochertest	60
7. Allgemeine DNA-Arbeiten	60
7.1. Isolierung von Plasmid-DNA	60
7.2. Aufreinigung von PCR-Produkten	60
7.3. Agarosegelelektrophorese	61
7.4. Isolierung von DNA-Fragmenten aus einem Agarosegel.....	61
7.5. Isolierung genomischer DNA aus <i>C. glutamicum</i>	62
7.6. Quantifizierung genomischer DNA aus <i>C. glutamicum</i>	62
7.7. Restriktion genomischer DNA aus <i>C. glutamicum</i>	62
7.7.1. Analyse geschnittener genomischer DNA aus <i>C. glutamicum</i>	63
8. Polymerasekettenreaktion (PCR)	63
8.1. Primer-Design.....	63
8.2. PCR-Reaktionsansätze und -Programme	64
8.2.1. PCR-Reaktionsansatz und -Programm für die KOD-Polymerase	65
8.2.2. PCR-Reaktionsansatz und -Programm für die Phusion-Polymerase	65
9. Klonierungsexperimente	66
9.1. DNA-Restriktionsspaltung.....	66
9.2. DNA-Ligation	67
10. DNA-Transfertechniken	67
10.1. Elektroporation	67
10.1.1. Herstellung elektrokompenter <i>C. glutamicum</i> Zellen.....	67
10.1.2. Elektroporation nach <i>C. glutamicum</i>	67
10.2. Hitzeschocktransformation von ultrakompetenten <i>E. coli</i> Zellen	68
10.2.1. Herstellung von ultrakompetenten <i>E. coli</i> Zellen (Morrison, 1977).....	68
10.2.2. Hitzeschocktransformation von ultrakompetenten <i>E. coli</i> Zellen.....	68
11. Konstruktion definierter Deletionen in <i>C. glutamicum</i>	69
11.1. Erstellung von Deletionskonstrukten mittels <i>Gene Splicing by Overlap Extension</i> (GeneSOEing)	69
11.2. Erstellung von Deletionskonstrukten mittels <i>Gibson-Assembly</i>	71

11.3. Etablierung einer Deletion im Genom	72
11.4. Deletionsnachweis mittels PCR	73
12. Transkriptomik	74
12.1. Kultivierung und Lagerung der <i>C. glutamicum</i> Zellen	74
12.2. Gesamt-RNA-Isolierung aus <i>C. glutamicum</i>	74
12.3. Test-PCR auf DNA-Reste	76
12.4. RNA Konzentrationsbestimmung	76
12.5. Transkriptionsanalyse mittels RT-qPCR	76
12.6. Globale Transkriptionsanalyse mittels DNA- <i>Microarray</i> -Technologie	77
12.7. Northern Blot Analyse	78
12.8. RNA-Sequenzierung (RNA-Seq)	79
12.8.1. Native 5'-Enden RNA-Seq-Protokoll	80
12.8.2. Entfernung ribosomaler RNA	80
12.8.3. Phenol-Chloroform-Isoamyl-Alkohol-Extraktion	81
12.8.4. RNA Fragmentierung	81
12.8.5. Terminator 5'-Phosphate-dependent Exonuklease Behandlung	81
12.8.6. 5'-Enden Vorbereitung auf Ligation	81
12.8.7. RNA-Adapterligation	82
12.8.8. Vorbereitung des <i>Loop</i> -DNA-Adapters	82
12.8.9. cDNA-Synthese und <i>Tagging</i> des 3'-Endes der cDNA	82
12.8.10. Amplifikation cDNA und Aufreinigung der 5'- cDNA-Bibliothek	83
12.8.11. Sequenzierung der 5'-cDNA-Bibliotheken	84
12.9. Bioinformatische Analyse der RNA-Seq-Daten	84
12.9.1. Mapping der sequenzierten <i>Reads</i> durch SARUMAN	84
12.9.2. RNA-Seq-Datenvisualisierung durch VAMP	84
12.9.3. Identifikation von Transkriptionsstartpunkten	84
12.9.4. <i>In vitro Run-Off</i> -Transkription	85
12.10. <i>Run-Off-Transcription/RNA-Sequencing</i> (ROSE)	87
12.11. DNA-Sequenzierung	87
13. Proteinherstellung und -analytik	88
13.1. Proteinaufreinigung mittels IMPACT™-CN-System	88
13.1.1. Konstruktion von Protein-CBD-Inteinfusionen	89
13.1.2. Expression der Intein-CBD-Fusionsproteine	89
13.1.3. Proteinaufreinigung mittels IMPACT-CN-System	89
13.2. Quantitative Bestimmung des Proteingehaltes	90
13.3. Diskontinuierliche SDS-PAGE	90

13.4. Massenspektrometrische Analyse zur Identifizierung von Proteinen.....	92
13.4.1. Tryptischer Verdau.....	92
13.4.2.MALDI-TOF Analyse.....	93
14. Sequenzanalyse	94
14.1.Motivsuche und Bewertung der Motive in potentiellen Promotorregionen	94
IV. Ergebnisse	95
1.Transkriptionelle Regulation des <i>sigH-rshA</i>-Operons und die Kontrolle des SigH-Regulationsnetzwerkes in <i>Corynebacterium glutamicum</i>.....	95
1.1. Das <i>sigH</i> -Gen und das <i>rshA</i> -Gen, welches für den Anti-Sigmafaktor kodiert, sind als Operon angeordnet.....	95
1.2. Die Gene des <i>sigH-rshA</i> -Operons werden von mehreren unterschiedlichen Promotoren aus transkribiert.....	98
1.3. Die Analyse des Transkriptionsprofil der Deletionsmutante RES167 $\Delta rshA$ identifiziert die meisten bekannten SigH-abhängigen Gene und neue Targetgene	101
1.4. Die differentielle Transkription ausgewählter SigH-abhängiger Gene konnte durch quantitative RT-PCR bestätigt werden	106
1.5. Experimentelle Bestimmung SigH-abhängiger Promotoren und Ableitung einer Promotorkonsensussequenz.....	107
2.Tiefenanalyse des <i>Corynebacterium glutamicum</i> SigH-Regulationsnetzwerkes mittels RNA-Sequenzierung der Deletionsmutanten RES167 $\Delta sigH$ und $\Delta rshA$... 111	
2.1. Erstellung der cDNA-Bibliotheken von mRNA des Referenzstammes <i>C. glutamicum</i> RES167 und der Deletionsstämmen <i>C. glutamicum</i> RES167 $\Delta sigH$ und $\Delta rshA$	111
2.2. Mapping-Statistik der RNA-Seq Läufe der Stämme <i>C. glutamicum</i> RES167, <i>C. glutamicum</i> RES167 $\Delta sigH$ und <i>C. glutamicum</i> RES167 $\Delta rshA$	112
2.3. Identifizierung der Transkriptionsstartpunkte durch Analyse der RNA-Seq Datensätze von <i>C. glutamicum</i> RES167, <i>C. glutamicum</i> RES167 $\Delta sigH$ und <i>C. glutamicum</i> RES167 $\Delta rshA$	112
2.4. Bestätigung der transkriptionellen Organisation des <i>sigH-rshA</i> -Operons	113
2.5. Experimentelle Bestätigung der vorhergesagten SigH-abhängigen Promotoren und Korrektur einiger experimentell bestimmter TSP durch die TSP-Identifizierung per RNA-Seq des Deletionsstamms RES167 $\Delta rshA$	116
2.6. Einige SigH-abhängige Promotoren sind in dem Deletionsstamm RES167 $\Delta sigH$ weiterhin aktiv.....	122

2.7. Die aktiven SigH-abhängigen Promotoren in der Deletionsmutante $\Delta sigH$ unterscheiden sich am 3'-Ende -35-Region von den übrigen SigH-abhängigen Promotoren.....	125
3. Entwicklung eines genomweiten <i>in vitro</i> Transkriptions/ Sequenzierungsansatzes (ROSE) und Anwendung auf die Regulationsnetzwerke von SigH und SigE in <i>Corynebacterium glutamicum</i>.....	126
3.1. Die ECF-Sigmafaktoren SigC, SigD, SigE, SigH und der <i>Housekeeping</i> -Sigmafaktor SigA konnten als native Proteine für die Nutzung in ROSE aufgereinigt werden ...	127
3.2. Direkte Identifizierung von exklusiv und partiell SigH-abhängigen Promotoren und Nachweis der direkten Kontrolle durch SigH.....	129
3.3. Nachweis der zusätzlichen und direkten Kontrolle partiell SigH-abhängiger Promotoren durch den ECF-Sigmafaktor SigE und die Identifizierung putativ exklusiv SigE-abhängiger Promotoren	140
3.4. Das globale Transkriptionsprofil der Deletionsmutante $\Delta cseE$ liefert eine Bestätigung für die Kontrolle <i>in vitro</i> identifizierter SigE-abhängiger Gene durch SigE <i>in vivo</i> und validiert CseE als Anti-Sigmafaktor von SigE	147
3.5. Die differentielle SigE-abhängige Transkription der Gene <i>amtR</i> und <i>cseE</i> konnte durch quantitative <i>real-time</i> RT-PCR in den Deletionsderivaten $\Delta cseE$ und $\Delta sigE$ im Vergleich zum Referenzstamm bestätigt werden	150
3.6. Ableitung der Promotorkonsensussequenzen exklusiv SigH-abhängiger und SigH/SigE-abhängiger Promotoren.....	152
V. Diskussion.....	155
1. Die Kombination von <i>Bottom-Up</i>- und <i>Top-Down</i>-Methoden zur Analyse von ECF-Sigmafaktornetzwerken ermöglicht zusammen mit RNA-Sequenzierung eine neue Betrachtungsweise von transkriptionellen Regulationsnetzwerken.....	156
1.1. RNA-Sequenzierung ermöglicht die genomweite qualitative und quantitative Betrachtungsweise von Promotoren	158
1.2. <i>Run-off-Transcription/RNA-Sequencing</i> (ROSE) ermöglicht die Genom-weite Identifizierung von Promotoren unter Ausblendung aller sekundären Effekte.....	160
1.3. Vergleich der Methoden zur Identifizierung und zum Vergleich von ECF-Sigmafaktornetzwerken	161
2. Das SigH-Regulationsnetzwerk in <i>Corynebacterium glutamicum</i>	165
2.1. Das <i>sigH-rshA</i> -Operon in <i>Corynebacterium glutamicum</i> weist eine komplexe transkriptionelle Organisation und eine Autoregulation auf.....	166

2.2. Das bekannte SigH-Regulationsnetzwerk konnte durch die globale Expressionsanalyse der <i>rshA</i> -Deletionsmutante und dem Einsatz von <i>Run-Off-Transcription/RNA-Sequencing</i> validiert und erweitert werden.....	168
3. Das SigE-Regulationsnetzwerk in <i>Corynebacterium glutamicum</i>	174
3.1. Das <i>sigE-cseE</i> Operon in <i>Corynebacterium glutamicum</i> weist eine ähnliche transkriptionale Organisation wie das <i>sigH-rshA</i> Operon und eine Autoregulation auf	175
3.2. Das SigE-Regulationsnetzwerk in <i>Corynebacterium glutamicum</i> umfasst Gene für Membranproteine, Chaperone, Proteasen und den Regulator des Stickstoffmetabolismus AmtR	178
4. Überschneidungen der SigH- und SigE-Regulationetzwerke in <i>Corynebacterium glutamicum</i>	181
5. Das 3'-Ende des -35 Promotormotivs bewirkt die Spezifität der Erkennung durch die <i>Corynebacterium glutamicum</i> ECF-Sigmafaktoren SigH und SigE	183
6. Das Sigmafaktornetzwerk in <i>Corynebacterium glutamicum</i>	186
7. <i>Corynebacterium glutamicum</i> als Modell-Organismus.....	190
8. Ausblick	191
VI. Literaturverzeichnis.....	194
VII. Anhang	204
1. Publikationsliste	204
2. Ausführliche Darstellung differentiell transkribierter Gene aus den <i>Microarray</i> -Experimenten.....	205
3. Zusammenfassung der RNA-Seq- (<i>Top-Down</i>) und ROSE-Ergebnisse (<i>Bottom-Up</i>)	214
Danksagung	221
Erklärung.....	222

Abbildungsverzeichnis

Abb. II-1 Zirkuläre Darstellung des <i>C. glutamicum</i> Chromosoms	5
Abb. II-2 DNA-Elemente und RNAP-Module die zu der Promotorerkennung beitragen.....	10
Abb. II-3 Strukturelle Organisation und schematische Übersicht der konservierten Regionen der Sigmafaktorgruppen 1, 2, 3 und 4	11
Abb. II-4 Phylogenetischer Baum der ECF-Sigmafaktoren	14
Abb. II-5 Modell der Aktivierung von SigE in <i>E. coli</i>	16
Abb. II-6 Verteilung der Nukleotide in der –35- und erweiterten –10-Region der <i>C. glutamicum</i> Promotoren von <i>Housekeeping</i> -Genen.....	17
Abb. II-7 Erstellung einer nativen 5'-Enden-angereicherten RNA-Seq-Bibliothek inklusive Ablauf der Sequenzierung und bioinformatischen Analyse der cDNA-Bibliothek.....	29
Abb. III-1 <i>Gene Splicing by Overlap Extension</i> – Erstellung eines Deletionskonstruktes.....	70
Abb. III-2 Gibson-Assembly - Isothermale <i>in vitro</i> Rekombination von DNA-Fragmenten in einem einzigen Schritt	72
Abb. III-3 Schematische Darstellung des IMPACT™-CN-System.....	88
Abb. IV-1 Alignment der Aminosäuresequenz zwischen vier corynebakteriellen SigH-Anti- Sigmafaktoren und deren Orthologen aus <i>M. tuberculosis</i> und aus <i>S. coelicolor</i>	96
Abb. IV-2 Genkarte des <i>sigH-rshA</i> -Operons und dessen Analyse durch Northern-Hybridisierung in <i>C. glutamicum</i> RES167 und Deletionsmutanten.....	97
Abb. IV-3 Identifikation der Transkriptionsstarts von <i>sigH</i> und <i>rshA</i> und Sequenz der Promotoren.....	99
Abb. IV-4 Genregion <i>sigH-rshA</i> , inklusive Kartierung der $\Delta rshA$ -Deletion und entsprechendem Deletionsnachweis per Test-PCR	101
Abb. IV-5 <i>Microarray</i> -Analyse von <i>C. glutamicum</i> RES167 im Vergleich mit der $\Delta rshA$ - Deletionsmutante	103
Abb. IV-6 Relative Transkriptmengen ausgewählter potentiell SigH-abhängiger Gene in <i>C. glutamicum</i> $\Delta rshA$	107
Abb. IV-7 Identifikation der Transkriptionsstartpunkte der Gene <i>uvrA</i> und <i>dnaJ2</i> durch PEX .	108
Abb. IV-8 Sequenzen der mutmaßlichen SigH-abhängigen Promotoren von <i>C. glutamicum</i>	109
Abb. IV-9 Nukleotidverteilung innerhalb der –35- und –10-Kernregionen SigH-abhängiger Promotoren in <i>C. glutamicum</i>	110
Abb. IV-10 Visualisierte native 5'-Enden RNA-Seq-Daten des genomischen Bereichs <i>sigH-rshA</i> von <i>C. glutamicum</i> RES167, RES167 $\Delta sigH$ und RES167 $\Delta rshA$	113
Abb. IV-11 Visualisierte native 5'-Enden RNA-Seq-Daten des genomische Bereichs stromaufwärts von <i>sigH</i> von <i>C. glutamicum</i> RES167 $\Delta rshA$	114
Abb. IV-12 Visualisierte native 5'-Enden RNA-Seq-Daten des genomischen Bereichs stromaufwärts von <i>rshA</i> von <i>C. glutamicum</i>	115
Abb. IV-13 Profile der <i>mappenden Reads</i> im Bereich der zwei TSP des <i>C. glutamicum</i> Gens <i>dnaK</i>	124

Abb. IV-14 Isolierte <i>C. glutamicum</i> Sigmafaktoren (Coomassie-Blau gefärbte SDS-PAGE).....	128
Abb. IV-15 Visualisierung der nativen 5´-Enden RNA-Seq Datensätze der <i>dnak</i> -Genregion der <i>C. glutamicum</i> RES167 Deletionsderivate $\Delta rshA$, $\Delta sigH$ und der <i>in vitro</i> Datensätze von ROSE-SigH und ROSE-Core (Kontrolle).....	131
Abb. IV-16 Visualisierung der nativen 5´-Enden RNA-Seq-Datensätze im Bereich des Gens <i>uvrD3</i> der Stämme <i>C. glutamicum</i> RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der <i>in vitro</i> Datensätze von ROSE-SigH, und ROSE RNAP-Core (Kontrolle)	134
Abb. IV-17 Visualisierung der nativen 5´-Enden RNA-Seq-Datensätzen im Bereich des Gens <i>trxB1</i> von <i>in vivo</i> -Transkriptomen der Stämme <i>C. glutamicum</i> RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der <i>in vitro</i> Transkriptome von ROSE-SigH, and ROSE-Core (Kontrolle).....	136
Abb. IV-18 Visualisierung der nativen 5´-Enden RNA-Seq-Datensätze im Bereich des Gens <i>dnak</i> der <i>C. glutamicum</i> RES167 Deletionsderivate $\Delta rshA$, $\Delta sigH$ und der <i>in vitro</i> Datensätze von ROSE-SigH, ROSE-SigE und ROSE-Core (Kontrolle).....	141
Abb. IV-19 Visualisierung der RNA-Seq-Datensätze der nativen 5´-Enden im Bereich der genomischen Region <i>sigE-cseE</i> der Stämme <i>C. glutamicum</i> RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der <i>in vitro</i> Datensätze von ROSE-SigH, ROSE-SigE und ROSE-Core-RNAP (Kontrolle)	146
Abb. IV-20 Genregion <i>sigE-cseE</i> inklusive Kartierung der Deletion $\Delta cseE$ und entsprechendem Deletionsnachweis per Test-PCR	147
Abb. IV-21 <i>Microarray</i> -Analyse von <i>C. glutamicum</i> RES167 im Vergleich mit der Deletionsmutante $\Delta cseE$	149
Abb. IV-22 Relative Transkriptmengen der Gene <i>amtR</i> und <i>cseE</i> in <i>C. glutamicum</i> $\Delta cseE/C. glutamicum$ RES167 und <i>C. glutamicum</i> $\Delta sigE/C. glutamicum$ RES167	151
Abb. IV-23 DNA-Sequenzlogo exklusiv SigH-abhängiger Promotoren.....	152
Abb. IV-24 DNA-Sequenzlogo SigH/SigE-abhängiger Promotoren.....	153
Abb. V-1 Konzept der Kombination aus <i>Top-Down</i> - und <i>Bottom-Up</i> -Ansatz zur Analyse von ECF-Sigmafaktornetzwerken.	156
Abb. V-2 Grafische Darstellung der Überschneidungen zwischen den drei angewandten Methoden zur Bestimmung des SigH-Regulon in <i>C. glutamicum</i>	162
Abb. V-3 Erweitertes Modell des SigH Regulationsnetzwerkes in <i>C. glutamicum</i>	170
Abb. V-4 Transkriptionelle Organisation des <i>sigE-cseE</i> Operons.....	176
Abb. V-5 Übersicht über die SigH und SigE Regulationsnetzwerke und deren Überschneidung in <i>C. glutamicum</i>	182
Abb. V-6 Konsensusmotive SigH und SigE-abhängiger Promotoren in <i>C. glutamicum</i> und <i>M. tuberculosis</i>	184
Abb. V-7 Interaktionsnetzwerk der Sigmafaktoren in <i>C. glutamicum</i>	187

Tabellenverzeichnis

Tab. II.1 Vorkommen von ECF-Sigmafaktoren bei Bakterien (Staroń <i>et al.</i> , 2009).....	12
Tab. III.1 Verwendete Bakterienstämme.....	36
Tab. III.2 Verwendete Plasmide	37
Tab. III.3 Verwendete Primer (Hersteller Metabion)	38
Tab. III.4 Verwendete Enzyme und Marker.....	41
Tab. III.5 Verwendete Chemikalien	42
Tab. III.6 Verwendete Kits	43
Tab. III.7 Verwendete Materialien.....	44
Tab. III.8 Verwendete Geräte und Apparaturen.....	45
Tab. III.9 Verwendete Software	46
Tab. III.10 Antibiotika	48
Tab. III.11 Zusammensetzung des Trenngels für die diskontinuierliche SDS-PAGE. Das Gel trennt Proteine in einem Bereich zwischen 15 und 60 kDa.....	91
Tab. III.12 Zusammensetzung des Sammelgels für die diskontinuierliche SDS-PAGE	91
Tab. IV.1 Gene mit erhöhtem mRNA-Pool in <i>C. glutamicum</i> $\Delta rshA$ verglichen mit <i>C.</i> <i>glutamicum</i> RES167 (Referenz), nach Funktion sortiert.*	104
Tab. IV.2 Mapping-Statistik der RNA-Seq Läufe der Stämme <i>C. glutamicum</i> RES167, $\Delta sigH$ und $\Delta rshA$	112
Tab. IV.3 Promotoren und TSP SigH-abhängiger Gene per RNA-Seq von dem Deletionsstamm RES167 $\Delta rshA$ bestimmt – Experimentelle Bestätigung der vorhergesagten SigH- abhängigen Promotoren und Korrektur einiger experimentell bestimmter TSP	118
Tab. IV.4 Promotoren SigH-abhängiger Gene die im Deletionsstamm RES167 $\Delta sigH$ aktiv sind.	123
Tab. IV.5 Konzentrationen der aufgereinigten Sigmafaktoren (N-terminal) per IMPACT-System	127
Tab. IV.6 Mapping-Statistik der ROSE-Experimente.....	130
Tab. IV.7 Exklusiv SigH-abhängige Promotoren (Identifikation per ROSE-SigH).....	133
Tab. IV.8 Partiell SigH-abhängige Promotoren (Identifikation per ROSE-SigH).....	139
Tab. IV.9 SigH- und zusätzlich SigE-abhängige Promotoren (Identifikation per ROSE-SigH und ROSE-SigE).....	143
Tab. IV.10 Putative exklusiv SigE-abhängige Promotoren (Identifikation per ROSE-SigE).....	146
Anhang Tab. 1 Differentiell transkribierte Gene in <i>C. glutamicum</i> RES167 $\Delta rshA$ im Vergleich zu <i>C. glutamicum</i> RES167 (Referenz) sortiert nach Änderung.....	205
Anhang Tab. 2 Differentiell transkribierte Gene in <i>C. glutamicum</i> RES167 $\Delta cseE$ im Vergleich zu <i>C. glutamicum</i> RES167 (Referenz) sortiert nach Änderung.....	209
Anhang Tab. 3 Partiell SigH-abhängige Promotoren.....	214
Anhang Tab. 4 SigH- und SigE- abhängigen Promotoren.....	216
Anhang Tab. 5 SigH-exklusive Promotoren.....	217
Anhang Tab. 6 Potentiell exklusiv SigE-abhängige Promotoren	220

Abkürzungsverzeichnis

Abkürzung	Bedeutung
1-CS	1-Komponenten-Systemen
2-CS	2-Komponenten-Systeme
A	Adenin
Abb.	Abbildung
Amp/ Ap	Ampicillin
APS	Ammoniumpersulfat
ASD	Antisigmadomäne
asRNA	antisense RNA
ATCC	<i>American Type Culture Collection</i>
ATP	Adenosintriphosphat
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
BHI	Brain Heart Infusion
bp	Basenpaar(e)
BPB	Bromphenolblau
C	Cytosin
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
CAS	<i>Candidate Anti-Sigma Sequence</i>
CASO	Caseinpepton-Sojamehlpepton
cat	Chloramphenicol Acetyltransferase
cDNA	<i>complementary DNA</i>
CoA	<i>Coenzym A</i>
<i>Corynebacterium</i>	<i>C.</i>
CP	<i>Crossing Point</i>
C-Quelle	Kohlenstoffquelle
CT	<i>Crossing Point Threshold</i>
C-Terminus	Carboxyl-Terminus
D	A oder G
DEPC	Diethyldicarbonat
dH ₂ O	Destilliertes Wasser
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTPs	Desoxyribonukleosidtriphosphate
dRNA-Seq	differentielle RNA-Seq
ds	doppelsträngig
DTT	1,4-Dithiothreitol
EC	Elongationskomplex
ECF	<i>Extracytoplasmic Function</i>
<i>Eschechiria coli</i>	<i>E. coli</i>
<i>et al.</i>	<i>et alii</i> (und andere)
EtBr	Ethidiumbromid
EtOH	Ethanol
G	Guanidin
GC-Gehalt	Guanin-Cytosin-Gehalt
gDNA	genomische DANN
GeneSOEing	<i>Gene-Splicing by Overlap Extension</i>
GRAS	<i>generally regarded as safe</i>
HCl	Salzsäure
HTH	<i>Helix-Turn-Helix</i>
IPTG	Isopropyl- β -D-thiogalactopyranosid

Abkürzung	Bedeutung
ITC	initiale Transkriptionskomplex
K	G oder T
kb	<i>kilo bases</i> (Kilobasen; 1000 Nukleotide)
kDa	Kilo Dalton; Molmasse
Km	Kanamycin
LB	Luria Bertani
M	C oder A
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
m/z	Verhältnis von Molekülmasse zu Ionenladung
mA	Milli-Ampere (Einheit der elektrischen Stromstärke)
max.	maximal
MCS	<i>Multiple Cloning Site</i>
min.	minimal
MOPS	3-Morpholinopropansulfonsäure
mRNA	<i>messenger RNA</i>
MS	Massenspektrometrie
nt	Nukleotide
N-Terminus	Amino-Terminus
NTP	Ribonukleosidtriphosphate
Nx	Nalidixinsäure
OD600	Optische Dichte bei einer Wellenlänge von 600 nm
<i>p.a.</i>	<i>per analysis</i> (zur Analyse)
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase chain reaction
PEX	<i>Primer-Extension-Analyse</i>
PMSF	Phenylmethylsulfonylfluorid
r	Resistent
RACE	Rapid Amplification of cDNA-Ends
RBS	Ribosomen-Bindestelle
RNA	Ribonukleinsäure
RNAP	RNA-Polymerase
RNAP-Core Enzym	RNA-Polymerase-Core-Enzym
RNAP-Holo Enzym	RNA-Polymerase-Holoenzym
RNA-Seq	RNA-Sequenzierung
ROSE	<i>Run-Off Transcription/RNA-Sequencing</i>
rRNA	ribosomale RNA
RT	Raumtemperatur
RT-qPCR	<i>Reverse transcription-quantitative PCR</i>
<i>S. coelicolor</i>	<i>Streptomyces coelicolor</i>
SDS	sodium dodecylsulfate, Natriumdodecylsulfat
Sig	Sigmafaktor
sRNA	<i>small RNA</i>
T	Thymin
Tab.	Tabelle
TAE	Tris-Acetat-EDTA(-Puffer)
TEMED	Tetramethylethylendiamin
TEN	<i>Terminator Exo-Nuclease</i>
TF	Transkriptionsfaktoren
TLS	Translationsstart
TOF	<i>Time of flight</i>
Tris	Tris(hydroxymethyl)-aminomethan
tRNA	Transfer-RNA
U	<i>Unit</i>
ün	über Nacht

Abkürzung	Bedeutung
UTR	untranslatierte Region
UV	Ultraviolett
V	Volt (Einheit der elektrischen Spannung)
vs.	<i>versus</i> , gegen
W	A oder T
WT	Wildtyp
X-Gal	5-Brom-4-chlor-3-indoxyl- β -D-galactopyranosid
Y	C oder T
z.B.	zum Beispiel
ZAS-Anti-Sigmafaktoren	Zink-bindende Anti-Sigmafaktoren
ZAS-Domäne	Zink bindende Anti-Sigma Domäne
σ	Sigmafaktors

I. Zusammenfassung

Die Anpassung an sich ändernde Umwelteinflüsse ist eine wesentliche Voraussetzung für das Leben allgemein, und im speziellen für das Überleben von Bakterien. Dies geschieht unter anderem über die Regulation der Genexpression. Der erste Schritt der Genexpression, die Transkriptionsinitiation, ist der am häufigsten und am stärksten regulierte Schritt in Bakterien. Hierfür verwenden Bakterien unter anderem verschiedenen Sigmafaktoren.

Sigmafaktoren sind Untereinheiten der bakteriellen RNA-Polymerase und sorgen zusammen mit dem RNA-Polymerase *Core*-Enzym als *Holo*-Enzyme für die Erkennung spezifischer Promotorsequenzen. Sie kontrollieren oft große Netzwerke von Genen. Gerade die ECF-Sigmafaktoren sind in Bakterien weit verbreitet und erfüllen wichtige Funktionen in der Stressantwort. Der Gram-positive Modellorganismus *Corynebacterium glutamicum* besitzt sieben Sigmafaktoren, darunter den *Housekeeping* Sigmafaktor SigA, den alternativen SigA-ähnlichen Sigmafaktor SigB und fünf ECF (*Extracytoplasmic Function*)-Sigmafaktoren (SigC, SigD, SigE, SigH und SigM).

Im Rahmen dieser Arbeit wurden die ECF-Sigmafaktoren SigH und SigE und deren Anti-Sigmafaktoren RshA bzw. CseE näher untersucht. Für SigH war bekannt, dass es in *C. glutamicum* und in anderen Actinomyzeten die Antwort auf Hitze- und oxidativen Stress reguliert. Für SigE war eine Beteiligung an der Antwort auf Oberflächen-Stress vermutet worden.

Durch die *in vivo* Transkriptomanalyse von Sigmafaktor- und Antisigmafaktorgen-Mutanten mittels *Microarray*-Analyse und RNA-Sequenzierung in Kombination mit der Analyse spezifischer Sigmafaktor-RNAP-Komplexe *in vitro* durch ROSE (**R**un-**O**ff **T**ranscription/**R**NA-**S**equencing) konnten die ECF-Sigmafaktor-Netzwerke von SigH und SigE und ihre Interaktionen, die Funktion der Anti-Sigmafaktoren und die transkriptionelle Organisation der beiden Sigmafaktor/Anti-Sigmafaktor Operons erstmalig und detailliert beschrieben werden. Die ROSE-Methodik stellt dabei eine im Rahmen dieser Arbeit erfolgte Neuentwicklung zur genomweiten *in-vitro* Transkription mit Einzelnukleotidauflösung dar.

Das SigH-Regulon war bereits durch frühere Arbeiten untersucht worden, konnte aber um Gene der SOS-Antwort (*uvrA*, *uvrC* und *uvrD3*), der Proteinqualitätskontrolle (*pup* und diverse Proteasen) und einer Reihe von Genen, die Membranproteine kodieren, erweitert werden. Daneben wurde das Spektrum der Proteine des SigH-Regulons erweitert, die der Aufrechterhaltung der Redox-Homöostase dienen. Es konnte weiter gezeigt werden, dass

RshA in *C. glutamicum* die Aufgabe als Anti-Sigmafaktor besitzt und das das *rshA*-Gen SigH-abhängig transkribiert wird.

Durch diese Arbeit konnte das SigE-Regulon erstmalig umfassend beschrieben werden. Bei den Genfunktionen des SigE-Netzwerkes handelt es sich hauptsächlich um Membranproteine, was im Einklang mit dem Einfluß des SigE-Netzwerkes auf die Antwort von Membran-Stress steht. Daneben konnte interessanterweise das Gen für den globalen Regulator des Stickstoffmetabolismus AmtR als Mitglied des SigE-Netzwerkes identifiziert werden. Es konnte weiter gezeigt werden, dass der Großteil des SigE-Regulons bis auf AmtR ebenfalls Bestandteil des SigH-Regulon ist. Das *sigE*-Gen liegt, wie *sigH-rshA*, in einem Operon mit dem Gen für den Anti-Sigmafaktor, *rseA*. Auch hier wird das Anti-Sigmafaktorgen in Abhängigkeit des entsprechenden Sigmafaktors transkribiert. Der *rseA*-Promotor wird allerdings zusätzlich von SigH erkannt.

Es konnte, wie für *Mycobacterium tuberculosis* beschrieben, gezeigt werden, dass die ECF-Sigmafaktoren SigH und SigE in *C. glutamicum* ähnliche Promotoren erkennen. Dabei ist der 3'-Bereich des -35 Promotormotivs für die Unterscheidung von SigH und zusätzlich SigE-abhängigen Promotoren verantwortlich.

Somit konnte klar gestellt werden, dass die Regulationsnetzwerke von SigH und SigE auf verschiedenen Ebenen hierarchisch verflochten sind. Während das SigH-Netzwerk vermutlich der Bekämpfung der unter vielen Stressbedingungen auftretenden Auslenkung der Redox-Homöostase dient, integriert das partiell untergeordnete SigE-Netzwerk vermutlich die spezielle Reaktion auf Oberflächenstress und ein Herunterfahren des Stickstoffmetabolismus unter den spezifischen Induktionsbedingungen für SigE. Da Oberflächenstress sicher auch ohne massive Störung der Redox-Homöostase vorkommen kann, erlaubt die Verknüpfung der Netzwerke so eine flexiblere Antwort.

Zusammenfassend konnten die erzielten Ergebnisse dazu genutzt werden das Sigmafaktor-Netzwerk in *C. glutamicum* näher zu charakterisieren. Es zeigen sich bei der Verknüpfung viele Parallelen zu orthologen Sigmafaktoren in den verwandten Actinobakterien *M. tuberculosis* und *Streptomyces coelicolor*. Die entwickelten komplementären *in vivo* und *in vitro*-Methoden haben ein großes Potential für zukünftige Studien von bakteriellen Regulationsnetzwerken.

II. Einleitung

1. Die Gattung Corynebakterium

1.1. Taxonomie der Corynebakterien

Die Gattung der Corynebakterien beinhaltet Gram-positive, nicht-motile, nicht-sporulierende und fakultativ anaerob bis aerob lebende Bakterien mit stäbchen- bis keulenförmiger Gestalt (griech. *coryne* = Keule). Weiter zeichnen sich Corynebakterien durch einen hohen G+C-Gehalt der DNA aus. Erstmals wurde diese Gattung durch Lehmann und Neumann 1896 zur taxonomischen Einordnung des humanpathogenen Vertreters *Corynebacterium diphtheriae* in *Bergey's Manual of Systematic Bacteriology* beschrieben (Goodfellow *et al.*, 2012). Phylogenetisch gesehen gehören die Corynebakterien zu den mycolsäurehaltigen Actinomyceten und sind eng verwandt mit den Mycobakterien (Nolden *et al.*, 2000), zu denen unter anderem der pathogene Vertreter *Mycobacterium tuberculosis* zählt. Aufgrund von Gemeinsamkeiten in der Lipidzusammensetzung und Peptidoglykanstruktur der mycolsäurehaltigen Zellwand werden Corynebakterien und Mycobakterien mit der Gattung *Nocardia* in der CMN-Gruppe zusammengefasst (Bansal-Mutalik und Nikaido, 2011).

Die bisher bekannten Corynebakterien wurden sowohl aus human- und veterinärklinischen Proben isoliert, als auch aus dem Boden und aus Fäkalien, aus Käse und anderen Milchprodukten, von Gemüse oder Früchten, sowie aus Tierfutter (Liebl, 2005). Aufgrund von 16S rDNA-Analysen lassen sich Corynebakterien allerdings eindeutig als monophyletische Gruppe beschreiben. Viele der zur Zeit beschriebenen 81 verschiedenen Corynebakterien Spezies sind von großem medizinischen und wirtschaftlichen Interesse (Eikmanns *et al.*, 1991; Ventura *et al.*, 2007).

Corynebakterien weisen eine hohe Diversität auf, die sich durch das Vorkommen von Mycolsäuren, dem stark schwankenden G+C-Gehalt der DNA (46 % bis 74 %) und der weiten Verbreitung unterschiedlicher Corynebakterien in verschiedenen Lebensräumen widerspiegelt (Funke *et al.*, 1995; Liebl, 2005)

Ein weiteres Merkmal der Corynebakterien ist die *Snapping Division*, die während der Zellteilung auftritt und ein V-förmiges Abwinkeln der Zellen beschreibt, welches durch das asymmetrische Aufreißen der Zellwand verursacht wird. Eine Besonderheit der Corynebakterien ist die hohe Stabilität ihrer Genome (stabile Genanordnung). Diese Eigenschaft wird nicht von den nahe verwandten Mycobakterien geteilt, deren Genome teilweise drastischen Veränderungen unterworfen sind (Nakamura *et al.*, 2003). Einige der

corynebakterielle Genome haben sich seit der evolutionären Trennung von den Mycobakterien kaum verändert. Das Fehlen der *recBCD*-Gene könnte eine Erklärung für diese Syntenie sein, da diese Gene an der Entstehung chromosomaler Inversionen beteiligt sind (Nakamura *et al.*, 2003). Dass es trotz der starken Syntenie corynebakterieller Genome zu einer diversen Artenbildung gekommen ist, ist deshalb wahrscheinlich auf den Verlust und den Zugewinn von Genen zurückzuführen, sowie auf Substitutionen einzelner Nukleotide in ihren Genomen (Ventura *et al.*, 2007).

1.2. **Biotechnologische und wirtschaftliche Relevanz von *Corynebacterium glutamicum***

Neben den medizinisch relevanten Corynebakterien, wie *C. diphtheriae* und *Corynebacterium pseudotuberculosis*, zeichnen sich andere Vertreter dieser Gattung durch ihr breites Spektrum an StoffwechsellLeistungen aus und sind dadurch von großem wirtschaftlichen Interesse. Die hier relevanten Spezies sind Boden- oder Wasserbewohner und kommen aufgrund ihrer Einordnung in die Gruppe der GRAS (*generally regarded as safe*) Organismen überwiegend in der fermentativen Produktion von Aminosäuren zum Einsatz, die als Futter- und Nahrungs- sowie als Arzneimittelzusätze verwendet werden (Schaffer *et al.*, 2001). Weitere Einsatzgebiete der Corynebakterien sind die Herstellung und die Konversion von Steroiden in biotechnologischen Verfahren (Constantinides, 1980), sowie der Abbau halogenierter Kohlenwasserstoffe (Yokota *et al.*, 1987), die Klärung von Abwässern (Calabrese *et al.*, 2007) und die Käsereifung (Schröder *et al.*, 2011).

Bereits 1957 wurde das wohl bekannteste produktionsrelevante Mitglied der Corynebakterien, *Corynebacterium glutamicum*, in einem *Screening*-Verfahren für Aminosäure-produzierende Mikroorganismen identifiziert; allerdings wurde es damals unter der Bezeichnung „*Micrococcus glutamicus*“ geführt (Kinoshita *et al.*, 2004). Die Eigenschaft zur natürlichen Sekretion von Aminosäuren und die daraus resultierende Eignung von *C. glutamicum* als Aminosäureproduzent beruht auf einem *Overflow*-Metabolismus, sowie begrenzten katabolischen Stoffwechselwegen zum Recycling überflüssiger Aminosäuren (Wohlleben *et al.*, 1993). Aufgrund der großen industriellen Bedeutung von *C. glutamicum* wurde die Forschung an diesem Bakterium in den letzten Jahren stark vorangetrieben, sodass eine kontinuierliche Entwicklung und Optimierung von gentechnischen Methoden für *C. glutamicum* stattfinden konnte, welche sowohl für die weitere Grundlagenforschung als auch für gezieltes *Metabolic Engineering* die Grundlagen schafften (Bardonnet und Blanco, 1991; Cadenas *et al.*, 1991; Eikmanns *et al.*, 1991; Kirchner und Tauch, 2003; Schäfer *et al.*, 1994; Tauch *et al.*,

1995a). Im Jahr 2003 wurde die Genomsequenz von *C. glutamicum* entschlüsselt (Abb. II-1) (Ikeda und Nakagawa, 2003; Kalinowski *et al.*, 2003), welche, in Kombination mit rekombinanten DNA-Transfertechniken zum einen die Grundlage für die Konstruktion verbesserter Produktionsstämme von *C. glutamicum* bildete und zum anderen die Bedeutung von *C. glutamicum* als Modellorganismus der *Actinomycetales*, insbesondere für die pathogenen Vertreter, wie *C. diphtheriae*, *Mycobacterium tuberculosis*, stärkte. Das sequenzierte Genom ermöglichte weiter die Etablierung der DNA-Microarray-Technologie zur globalen Analyse des Transkriptom von *C. glutamicum* (Hüser *et al.*, 2003) und eröffnete damit die postgenomische Ära für diesen Modellorganismus.

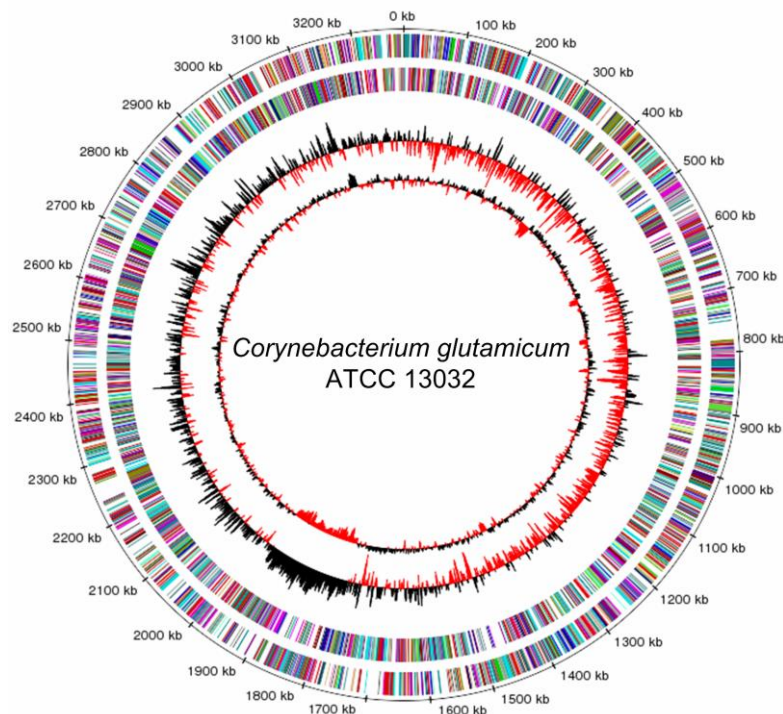


Abb. II-1 Zirkuläre Darstellung des *C. glutamicum* Chromosoms. Die Darstellung zeigt das 3,28 Mbp große Genom von *C. glutamicum* ATCC 13032 mit drei charakteristischen Merkmalen. Die Kreise beschreiben von außen nach innen: (i) die Kodierregionen auf dem positiven DNA-Strang, (ii) die Kodierregionen auf dem negativen DNA-Strang, (iii) den GC-Skew, der zur Identifizierung des *leading* und des *lagging Strand* der DNA-Replikation genutzt wird und (iv) den durchschnittlichen G+C-Gehalt der DNA, wobei rote Werte negativ und schwarze Werte positiv vom Durchschnittswert abweichen.

Gegenwärtig werden mit genetisch veränderten *C. glutamicum* Stämmen überwiegend die Aminosäuren L-Glutamat und L-Lysin mit einem jährlichen Gesamtvolumen von über 2 Mio t für die Tierzucht produziert (Ajinomoto, 2012)

2. Regulation bakterieller Genexpression

Die Anpassung an sich ändernde Umwelteinflüsse ist eine wesentliche Voraussetzung für das Leben allgemein, und im speziellen für das Überleben von Bakterien. Dies ist besonders in Umgebungen wichtig, in denen Umwelteinflüsse schnell oder oft wechseln. Das natürliche Habitat des Actinobakteriums *C. glutamicum*, der Boden, ist ein Beispiel hierfür. Das Bakterium muss sich an die ständig wechselnden Umweltbedingungen in kürzester Zeit anpassen, um das Überleben zu sichern. Dabei muss zum einen der Metabolismus angepasst werden, und zum anderen müssen Proteine gebildet werden, die der Zelle das Überleben ermöglichen. Dies geschieht unter anderem über die Regulation der Genexpression. Für die Anpassung an veränderte Umwelteinflüsse müssen entsprechende Signale zuerst sensiert und dann weitergeleitet werden. Dies geschieht in der bakteriellen Zelle durch ein großes Repertoire an 1- & 2-Komponenten-Systemen (Staroń *et al.*, 2009). Bei 1-Komponenten-Systemen (1-CS) befinden sich die *Input*-Domäne und die *Output*-Domäne innerhalb desselben Proteins (Ulrich *et al.*, 2005). Die 2-Komponenten-Systeme (2-CS) bestehen hingegen aus zwei unterschiedlichen Proteinen, einem Sensorprotein und einem Antwortregulator (Stock and Da Re, 2000). Über eine Signaltransduktionskette wird ein sensiertes Signal auf das Regulatorprotein durch Phosphorylierung übertragen und damit zu einer veränderten Bindeaffinität zur Ziel-DNA des Regulatorproteins führt, wodurch die Genexpression spezifischer Gene reguliert werden kann.

Die Regulation der Genexpression kann an jedem Einzelschritt der Genexpression von der Transkription bis zur Translation erfolgen. Eine posttranslationale Modifizierung der Proteinausstattung ist zusätzlich möglich. Die Transkription oder genauer gesagt die Transkriptionsinitiation ist aber in Bakterien der am häufigsten und am stärksten regulierte Schritt (Browning und Busby, 2004).

2.1. Bakterielle Transkription

Die Transkription umfasst drei Phasen: Initiation, Elongation und Termination. Die Transkriptionsinitiation, als entscheidender Schritt der Regulation der Genexpression, lässt sich wiederum in drei weitere Schritte unterteilen: Promotorerkennung, Initiation der Transkription und Promotorfreigabe (Helmann, 2009). Der Prozess der Transkription beginnt mit der Bindung der RNA-Polymerase (RNAP) an die Promotorregion. Als Promotor kann man den Bereich der DNA bezeichnen, der während der Transkriptionsinitiation mit der RNAP in Wechselwirkung steht. Im Gegensatz zu Archaea und Eukaryoten besitzen Bakterien nur eine einzige RNAP, welche für die Synthese der mRNA, rRNA und auch der tRNA verantwortlich

ist. Man unterscheidet grundsätzlich zwei Formen der RNAP (Burgess et al., 1969). Das RNA-Polymerase-Core-Enzym (RNAP-Core Enzym) besteht aus vier unterschiedlichen Untereinheiten (α_2 , β , β' , ω) und bildet das katalytische Grundgerüst der RNAP. Die β -Untereinheit und die β' -Untereinheit bilden die Hauptstruktur des RNAP-Core Enzyms. Diese Untereinheiten sind für die Bindung an die DNA notwendig und tragen das aktive Zentrum, an dem neue Nukleotide an das 3'-OH-Ende der RNA gehängt werden. Die zwei α -Untereinheiten sind zum einen für den Zusammenbau und die Stabilität des RNAP-Core Enzyms verantwortlich, und zum anderen unterstützen sie die Bindung an die Promotorregion durch Erkennung einer stromaufwärts gelegenen alternierenden Serie aus den Nukleotiden A und T (UP-Element) einiger Promotoren. Die kleinste Untereinheit der RNAP ist ω . Ihre Funktion ist es, die dreidimensionale Struktur der RNAP zu stabilisieren und aufrechtzuerhalten. Allgemein besitzt das RNAP-Core Enzym nur basale unspezifische Aktivitäten. Die spezifischen Aktivitäten der RNAP, d. h. Promotorerkennung und Transkriptionsinitiation, werden erst durch die reversible Bindung des Sigmafaktors (σ) ermöglicht. Bindet ein Sigmafaktor an das RNAP-Core Enzym, spricht man von dem RNA-Polymerase-Holoenzym (RNAP-Holo Enzym), dass aus den Einheiten α_2 , β , β' , ω und σ aufgebaut ist.

Vor der Erkennung eines Promotors ist die RNAP unspezifisch an die DNA gebunden und diffundiert eindimensional auf dem DNA-Strang entlang (*Sliding*), bis ein Promotor erkannt wird (Bustamante et al., 1999). Die Promotorerkennung beginnt mit der Bildung eines Erkennungskomplexes zwischen dem RNAP-Holo Enzym und der Promotorregion (Borukhov und Nudler, 2008; Haugen et al., 2008). Diese Erkennung beruht auf der hohen Affinität des gebundenen Sigmafaktors zu den beiden Promotorkernmotiven, die sich ca. 10 bzw. 35 Nukleotide stromaufwärts des Startpunts der Transkription (+1) befinden und daher als -10 und -35 Motive bezeichnet werden. In Abhängigkeit vom Sigmafaktor und vom Promotor kann die Lage jedoch abweichen oder angrenzende Sequenzen, wie die erweiterte -10 Region, umfassen (Helmann, 2009). Weiterhin ist die α -Untereinheit der RNAP zum Teil auch an der Erkennung des Promotors durch Bindung an UP-Elemente beteiligt (Abb. II-2). Die verantwortlichen Regionen des Sigmafaktors für die Erkennung des Promotors werden in Kapitel 0 im Detail dargestellt.

Im initialen „geschlossenen Komplex“ liegt die DNA noch als Doppelstrang vor (de Haseth et al., 1998). In den nachfolgenden Schritten wird der DNA-Doppelstrang aufgeschmolzen. Dieser Vorgang ist von der Sequenz-spezifischen Erkennung des nicht-*Template* Stranges durch den Sigmafaktor und die Positionierung des *Template*-Stranges im aktiven Zentrum der RNAP abhängig (Borukhov und Nudler, 2003; Helmann und de Haseth, 1999; Young et al., 2004). Ist der „offene Komplex“ gebildet, kann die Transkriptionsinitiation beginnen.

In dieser Reaktion bindet die RNAP zwei NTPs, die zu den Positionen +1 und +2 des *Template-Stranges* komplementär sind. Das +1 NTP an der *i*-Stelle (NTP/RNA-Bindestelle) der RNAP ist vorwiegend ein Purin, d. h. die meisten bakteriellen Transkripte beginnen mit einem A oder G (Hermann, 2009). Das +2 NTP liegt an der *i*+1-Stelle der RNAP. Die NTPs werden durch eine Phosphodiesterbindung verbunden und das entstehende terminale NMP von der RNAP-Stelle *i*+1 auf die Position *i* verschoben, da diese eine sehr geringe Affinität für NTPs besitzt. Dieser Zyklus wird für jedes folgende NTP wiederholt. Generell wird die Transkription in einem Bereich von sechs bis neun Nukleotiden stromabwärts des -10-Promotormotivs initiiert (Hermann, 2009). Der initiale Transkriptionskomplex (ITC) beinhaltet weiterhin den Sigmafaktor und hat Kontakt zu den Promotorelementen. Erst nach der mehrfachen Bildung kurzer Transkripte findet der erfolgreiche Übergang zu dem Elongationskomplex (EC) statt (Hsu, 2009). Der EC besitzt eine wesentlich geringere Affinität zu dem Sigmafaktor als der ITC und verliert diesen. Dieser Übergang vom ITC zum EC wird auch Promotorfreigabe genannt.

Einmal geformt, synthetisiert der EC RNA bis zu einer Länge von mehreren Tausend Nukleotiden. Der finale Schritt der Transkription ist die Termination, die durch spezifische G+C-reiche *Loop*-Terminatorsequenzen oder durch den Terminationsfaktor Rho hervorgerufen wird, und mit der Freisetzung des Transkripts endet (Ciampi, 2006).

2.2. Regulation der Transkription

Wie bereits geschildert, ist der erste Schritt der Genexpression, die Transkriptionsinitiation, der am häufigsten und am stärksten regulierte Schritt. Dies ist schon allein aus ökonomischer Sicht sinnvoll (Browning und Busby, 2004). Die Transkription kann dabei durch DNA-bindende Transkriptionsfaktoren (TF) entweder reprimiert werden, wie z. B. durch den Lac-Repressor aus *Escherichia coli* (Lewis, 2005) oder den Hauptregulator des Stickstoffmetabolismus AmtR aus *C. glutamicum* (Hasselt *et al.*, 2011), oder sie wird aktiviert, wie z. B. durch MalT aus *E. coli* (Boos and Böhm, 2000). Es gibt aber auch eine Reihe von dualen Regulatoren, die die Transkription einzelner Gene reprimieren, während sie die Transkription anderer Gene aktivieren. TF binden an eine spezifische Erkennungssequenz stromaufwärts eines Zielgens, den Operator, und können somit die Transkription verhindern oder aktivieren. Für die Bindung an die DNA ist die DNA-Bindedomäne eines TF verantwortlich, die in Bakterien in den allermeisten Fällen aus einem Helix-Turn-Helix-Motiv (HTH-Motiv) besteht (Pabo und Sauer, 1992; Pérez-Rueda und Collado-Vides, 2001). Häufig ist die regulatorische Funktion eines

DNA-bindenden TF für ein Zielgen von der Position der entsprechenden Bindesequenz abhängig. Liegt diese Sequenz stromaufwärts der –35-Region des Promoters, so wird die Transkription des Gens normalerweise aktiviert. Die Bindung des TF in oder stromabwärts der Promotorregion führt hingegen zu einer Inhibierung der Transkription, da die RNAP entweder nicht mehr an die Promotorregion binden kann, oder aber die initiierte Transkription dadurch abgebrochen wird, dass ein TF den folgenden DNA-Abschnitt in Form eines sogenannten *Road Block* blockiert (Babu und Teichmann, 2003). Neben den beschriebenen TF besitzen Bakterien noch eine weitere grundlegende Möglichkeit die Genexpression auf Basis der Transkriptionsinitiation zu verändern. Dies geschieht durch die Verwendung unterschiedlicher RNAP-*Holo* Enzyme.

In Bakterien gibt es unterschiedliche Sigmafaktoren, die in Abhängigkeit der bestehenden Bedingungen aktiv sind und zur Bildung unterschiedlicher RNAP-*Holo* Enzyme führen. Die verschiedenen RNAP-Holo Enzyme besitzen unterschiedliche Spezifitäten für Promotoren, die allein auf dem gebundenen Sigmafaktor beruhen. Die Verwendung unterschiedlicher RNAP-*Holo* Enzyme mit deren unterschiedlicher Promotorspezifität ermöglicht es den Bakterien die Expression größerer Netzwerke abhängig von äußeren oder inneren Einflüssen umzuprogrammieren.

2.3. Klassifizierung und Aufbau von Sigmafaktoren

Bakterielle Sigmafaktoren sind in zwei strukturell und funktional unterschiedliche Familien eingeteilt. Diese beiden Familien werden als Sigma⁷⁰ und Sigma⁵⁴ bezeichnet (T M Gruber and C A Gross, 2003). Vertreter der Familie Sigma⁷⁰ sind in allen bekannten bakteriellen Spezies vorhanden, während Sigmafaktoren der Familie Sigma⁵⁴ eher selten vorkommen. Diese selten vorkommenden Sigmafaktoren benötigen für ihre Aktivität ATP und die zusätzliche Beteiligung von *Enhancer*-Proteinen. Die Sigmafaktoren beider Familien sind modular aufgebaut, besitzen aber keinerlei Sequenzähnlichkeit. Die Sigma⁷⁰ Sigmafaktoren bestehen aus bis zu vier konservierten Regionen (T M Gruber and C A Gross, 2003).

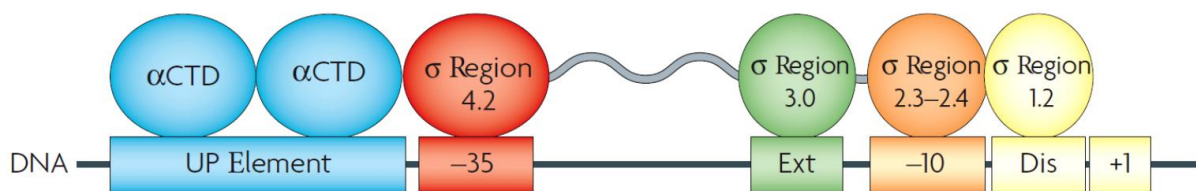


Abb. II-2 DNA-Elemente und RNAP-Module die zu der Promotorerkennung beitragen.

UP Element, alternierende Serie aus A und T; -35; -35-Promotormotiv; -10, -10-Promotormotiv; Ext, erweiterte -10-Region; Dis, *Discriminator Region*; +1, Transkriptionsstartpunkt (TSP). Die Region 2.3 - 2.4 des Sigmafaktors erkennt und bindet die -10-Region des Promotors. Die Region 4.2 erkennt und bindet die -35-Promotorregion. Einige Promotoren besitzen eine erweiterte -10-Region (Ext), die durch die Region 3.0 des Sigmafaktors erkannt wird. Die Region 1.2 ist für die Erkennung einer G+C-reichen Region zwischen -10 und +1 in Promotoren ribosomaler RNA und Transfer-RNA verantwortlich (Dis). Das UP-Element wird durch die carboxyterminalen Enden der RNAP- α -Untereinheiten erkannt (α CTD). Abbildung nach (Haugen *et al.*, 2008).

Die DNA-Bindedomänen der Sigmafaktoren sind für die Erkennung der -10- und -35-Motive der Promotoren verantwortlich. Die Regionen 2.3 und 2.4 erkennen und binden die -10-Region des Promotors und die Region 4.2 die -35-Region.

Einige Promotoren besitzen eine erweiterte -10-Region (*extended -10*), die durch die Region 3.0 des Sigmafaktors erkannt wird (Gruber und Gross, 2003). Die Region 1.2 ist für die Erkennung einer G+C-reichen Region (*Discriminator Region*) zwischen -10 und +1 in Promotoren ribosomaler RNA (rRNA) und Transfer-RNA (tRNA) verantwortlich (Haugen *et al.*, 2008). Neben der Interaktion des Sigmafaktors mit dem Promotor interagieren zusätzlich die carboxyterminalen Enden der α -Untereinheiten der RNAP mit UP-Elementen einiger Promotoren (Abb. II-2).

Die Sigmafaktoren der Sigma⁷⁰ Familie sind in vier Gruppen unterteilt. Diese Unterteilung wurde auf Basis der Struktur und der physiologischen Rolle der Sigmafaktoren durchgeführt

(Abb. II-3). Die Gruppe 1 Sigmafaktoren besitzen alle vier konservierten Regionen und werden als essentielle primäre Sigmafaktoren bezeichnet. Die primären Sigmafaktoren sind für die Transkription von *Housekeeping*-Genen verantwortlich und werden zum Teil auch als *Housekeeping*-Sigmafaktoren bezeichnet. Alle Bakterien besitzen einen solchen primären (oder *Housekeeping*-) Sigmafaktor.

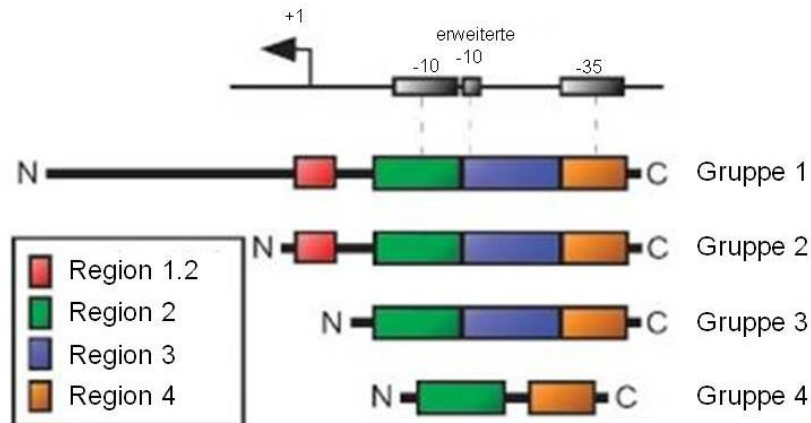


Abb. II-3 Strukturelle Organisation und schematische Übersicht der konservierten Regionen der Sigmafaktorgruppen 1, 2, 3 und 4. Die unterbrochenen Linien verbinden die Sigmafaktorregionen mit den Promotorelementen, die diese erkennen. Der Transkriptionsstartpunkt ist durch einen geknickten Pfeil und +1 gekennzeichnet. Gruppe 1: primäre (*Housekeeping*) Sigmafaktoren; Gruppe 2: *primary-like* Sigmafaktoren; Gruppe 3: Flagella-, Sporulations-, Hitzeschock-Sigmafaktoren; Gruppe 4: ECF-Sigmafaktoren. Abbildung nach (Rodrigue et al., 2006).

Die meisten Bakterien besitzen über diesen hinaus noch alternative Sigmafaktoren. Besonders Bakterien mit einem komplexen Lebensstil besitzen gewöhnlich eine umfangreiche Ausstattung an Sigmafaktoren. Wie in Kapitel II.2.2 bereits erwähnt, kann das Bakterium so die Transkription der RNAP durch die Verwendung alternativer Sigmafaktoren unter Verdrängung des primären Sigmafaktors grundlegend umlenken. Zu der Gruppe 2 zählen die *primary-like* Sigmafaktoren, die bis auf die verkürzte N-terminale Region eine ähnliche Struktur wie die Sigmafaktoren der Gruppe 1 aufweisen, aber nicht essentiell sind. Sie dienen als eine Art Back-up zu den primären Sigmafaktoren, da sie ähnliche Promotoren erkennen. Gruppe 2 Sigmafaktoren ermöglichen die Transkription von *Housekeeping*-Genen unter bestimmten Stressbedingungen (Patek und J Nesvera, 2011). Die Gruppe 3 Sigmafaktoren kontrollieren in einigen Bakterien Promotoren von Genen, die an spezifischen Funktionen wie Sporulation, Flagellensynthese oder Hitzeschockantwort beteiligt sind. Ihnen fehlt die Region 1.

Die Gruppe 4 ist die größte und diverseste Gruppe der Sigma⁷⁰ Sigmafaktoren. Diese Gruppe wird oft als *Extracytoplasmic Function* (ECF)-Sigmafaktorfamilie bezeichnet (Helmann, 2002). ECF-Sigmafaktoren besitzen nur die konservierten Regionen 2 und 4, die für die Bindung an die RNAP und die Erkennung der Promotormotive -10 und -35 verantwortlich sind.

In Abwesenheit eines Signals sind die meisten alternativen Sigmafaktoren durch einen entsprechenden Anti-Sigmafaktor mittels Protein-Protein-Interaktion inaktiviert (Hermann, 1999).

2.4. ECF-Sigmafaktoren

Lonetto und Mitarbeiter haben 1994 den Begriff ECF (*Extracytoplasmic Function*) definiert. In ihrer Arbeit konnten sie einer Gruppe nur schwach verwandter kleiner Proteine, die an der Regulation von Zelloberflächenprozessen bzw. Transportvorgängen beteiligt waren, die Funktion alternativer Sigmafaktoren zuweisen und vergaben den Namen entsprechend ihrer Funktion (Lonetto *et al.*, 1994). Der überwiegende Anteil dieser Sigmafaktoren ist an der Antwort auf verschiedene Stresse beteiligt. Die Promotormotive, die von ECF-Sigmafaktoren erkannt werden, unterscheiden sich teils deutlich von denen, die durch primäre Sigmafaktoren erkannt werden.

Die Anzahl und Art der ECF-Sigmafaktoren ist in unterschiedlichen Bakterien sehr variabel (Tab. II.1). Die Vertreter der bakteriellen Taxa *Aquificae*, *Chlamydiae*, *Rickettsiales* und *Borrelia* besitzen beispielsweise keine ECF-Sigmafaktoren, während *Streptomyces coelicolor* A3(2) mit 51 ECF-Sigmafaktoren eine umfassende Ausstattung besitzt (Ulrich und Zhulin, 2010). Die Anzahl der ECF-Sigmafaktoren spiegelt oft den Lebensstil eines Bakterium wider, oder die Umgebung in der dieses lebt. Je komplexer der Lebensstil ist, desto größer ist die Ausstattung an ECF-Sigmafaktoren. Die durchschnittliche Anzahl an ECF-Sigmafaktoren beträgt dabei sechs pro Genom (Staroń *et al.*, 2009).

Tab. II.1 Vorkommen von ECF-Sigmafaktoren bei Bakterien (Staroń *et al.*, 2009).

Anzahl ECF	Anzahl Genome	Representatives Beispiel ^a
0	239	<i>Aquificae</i> , <i>Chlamydiae</i> , <i>Rickettsiales</i> , <i>Borrelia</i> , <i>Mycoplasma spp.</i>
1	98	<i>Staphylococcus spp.</i> , <i>Anabaena variabilis</i>
2	62	<i>Escherichia coli</i>
3 – 5	85	<i>Corynebacterium glutamicum</i> , <i>Enterococcus faecalis</i>
6 – 10	95	<i>Bacillus subtilis</i> , <i>Rhodobacter sphaeroides</i>
11 – 20	89	<i>Bacillus anthracis</i> , <i>Caulobacter crescentus</i> , <i>Pseudomonas aeruginosa</i>
21 – 30	18	<i>Pseudomonas fluorescens</i> , <i>Pseudomonas putida</i> , <i>Frankia sp.</i>
> 30	14	<i>Streptomyces spp.</i> , <i>Bacteroides spp.</i> , <i>Myxococcus xanthus</i>

^a Representative Bakterien oder Bakteriengruppen sind als Beispiel aufgelistet

Staroń beschreibt die ECF-Sigmafaktoren als den dritten fundamentalen Mechanismus der bakteriellen Signaltransduktion neben 1-CS und 2-CS. In einer umfassenden Analyse der ECF-Sigmafaktorfamilie wurden die ECF-Sigmafaktoren dabei in über 43 Hauptgruppen

unterteilt (Staroń *et al.*, 2009). Die Analyse basierte auf einem Datensatz aus über 2.700 vorhergesagten ECF-Sigmafaktorproteinsequenzen aus 369 mikrobiellen Genomen. Insgesamt 1.735 der 2.708 ECF-Sigmafaktoren konnten durch die *HMM*-Analyse einer der 43 ECF-Hauptgruppen (ECF01 bis ECF43) zugeordnet werden, von denen jede mehr als zehn Proteine umfasst. Innerhalb der Klassifizierung sind die sehr gut charakterisierten RpoE-ähnlichen Sigmafaktoren (z. B. SigE aus *E. coli*; SigW aus *Bacillus subtilis*) und die Fecl-ähnlichen Sigmafaktoren die dominierenden Gruppen und konnten weiter unterteilt werden (RpoE-*like*: ECF01–04 mit 272 Proteinen; Fecl-*like*: ECF05–ECF10 mit 254 Proteinen). Im Vergleich zu den RpoE-*like* und Fecl-*like* Proteinen zeigen die fünf Gruppen, die die *cytoplasmic-sensing* ECF-Sigmafaktoren beinhalten (ECF11–ECF15 mit 213 Proteinen), eine nur geringe Sequenzähnlichkeit (Abb. II-4). Die Gemeinsamkeit dieser ECF-Sigmafaktoren besteht darin, dass sie einen cytoplasmatisch lokalisierten Anti-Sigmafaktor besitzen und damit wahrscheinlich cytoplasmatische Veränderungen sensieren können. Die vorwiegende Anzahl an Anti-Sigmafaktoren ist dagegen Membran-gebunden.

Die Daten der erfolgten Klassifizierung der ECF-Sigmafaktoren durch Staroń sind in einer umfassenden Datenbank *Microbial Signal Transduction 2* (MiST2 Database) hinterlegt (Ulrich und Zhulin, 2010).

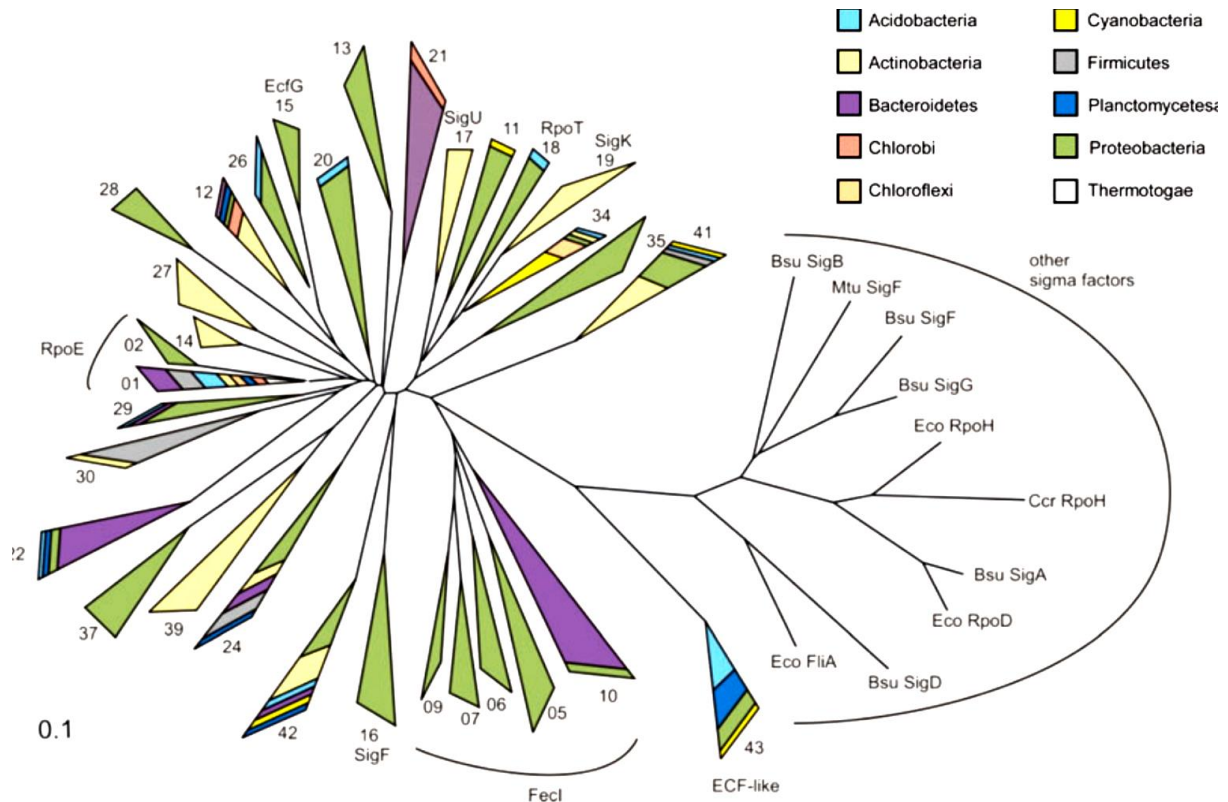


Abb. II-4 Phylogenetischer Baum der ECF-Sigmafaktoren. Die taxonomische Diversität, die phylogenetischen Beziehungen und die Sequenzdiversität der einzelnen ECF-Gruppen sind hier visualisiert. Der phylogenetische Baum basiert auf einem multiplen Sequenz-*Alignment* der Regionen 2 und 4 der beiden unterschiedlichsten Vertreter einer Gruppe. Jedes Dreieck repräsentiert eine der Hauptgruppen, die unterschiedlichen Farben kennzeichnen die phylogentischen Zugehörigkeiten. Die Länge der Dreieckseiten stellt die Sequenzdiversität innerhalb einer ECF-Gruppe dar. Je länger die Seiten, desto größer ist die Diversität. Abbildung aus (Staroń *et al.*, 2009)).

Die am besten verstandenen ECF-Sigmafaktoren sind SigE aus *E. coli*, SigW aus *B. subtilis*, SigE aus *Rhodobacter sphaeroides* und SigR aus *Streptomyces coelicolor* (Staroń *et al.*, 2009). In der Arbeit von Staroń konnte gezeigt werden, dass diese Proteine beispielhaft für andere ECF-Sigmafaktoren die Mechanismen und Charakteristika der ECF-vermittelten Signaltransduktion und Genregulation aufzeigen. Erstens autoregulieren die meisten ECF-Sigmafaktoren ihre eigene Expression; zweitens werden die Gene der ECF-Sigmafaktoren für gewöhnlich mit ihren Anti-Sigmafaktorgenen kotranskribiert; drittens werden die ECF-Sigmafaktoren in Abwesenheit eines Stimulus (z. B. Stress) durch den Anti-Sigmafaktor inaktiv gehalten und viertens werden die Anti-Sigmafaktoren durch ein spezifisches Signal inaktiviert und führen zur Freisetzung und damit zur Aktivierung des Sigmafaktors der nun an ein RNAP-*Core* Enzym binden kann und dadurch zu Transkription spezifischer Gene führt.

Für die ECF-Sigmafaktoren, die keinen Anti-Sigmafaktor besitzen, erfolgt die Regulation der Sigmafaktoraktivität zum Beispiel über die Transkriptmenge des ECF-Sigmafaktors. Dies wurde für SigE in *S. coelicolor* beschrieben (Staroń *et al.*, 2009).

Durch die Ausstattung mit unterschiedlichen ECF-Sigmafaktoren in Zusammenhang mit den sensierenden und inhibierenden Funktion der Anti-Sigmafaktoren ergibt sich ein hochmodulares System zur Regulation der Genexpression, der dritte fundamentale Mechanismus der Signaltransduktion in Bakterien (Staroń *et al.*, 2009).

2.5. Anti-Sigmafaktoren

Wie in den vorangehenden Abschnitten erwähnt, werden die meisten ECF-Sigmafaktoren durch einen entsprechenden Anti-Sigmafaktor durch Protein-Protein-Interaktion inaktiv gehalten (Helmann, 1999). Die zelluläre Konzentration an freien ECF-Sigmafaktor ist abhängig von der Aktivität und Konzentration des entsprechenden Anti-Sigmafaktors. Diese kommen in den meisten Fällen Membran-assoziiert vor. Es gibt aber auch Vertreter die gelöst im Cytoplasma vorliegen, wie die Anti-Sigmafaktoren der ECF-Gruppen ECF11 - ECF15). Gene, die ECF-Sigmafaktoren kodieren, sind normalerweise mit den Genen der Anti-Sigmafaktoren kotranskribiert. Viele Anti-Sigmafaktoren besitzen ein strukturelles Motiv, die sogenannte Anti-Sigmadomäne (ASD). Die ASD spielen eine wichtige Rolle bei der Reaktion auf spezifische Signale, die die Aktivität des Anti-Sigmafaktors kontrollieren (Campbell *et al.*, 2007). Ungefähr 30 % der Proteine, die eine ASD enthalten, werden als Zink-bindende Anti-Sigmafaktoren (ZAS-Proteine) bezeichnet (Staroń *et al.*, 2009).

Die Inaktivierung der Anti-Sigmafaktoren, und dadurch die Reaktivierung der Sigmafaktoren, kann über verschiedene Mechanismen erfolgen. Zum einen kann das über die Degradation des Anti-Sigmafaktors durch Proteolyse geschehen, wie bei dem Anti-Sigmafaktor RseA von SigE aus *E. coli* (Abb. II-5; Ho und Ellermeier, 2012) oder durch eine Konformationsänderung des Anti-Sigmafaktors. Ein Beispiel für eine solche Konformationsänderung ist der Redox-sensitive Anti-Sigmafaktor RsrA aus *S. coelicolor*. Es handelt sich um einen Anti-Sigmafaktor der ZAS-Familie. Unter reduzierenden Bedingungen wird der ECF-Sigmafaktor SigR durch die Bindung des Anti-Sigmafaktors RsrA inaktiv gehalten. In Antwort auf Thiol-oxidativen Stress, verliert RsrA ein Zinkion und es bildet sich eine Disulfidbindung zwischen zwei Cysteinen, die zuvor das Zinkion koordinierten. Dies führt zu einer Konformationsänderung des Anti-Sigmafaktors und damit zur Freigabe des ECF-Sigmafaktors SigR (Campbell *et al.*, 2008).

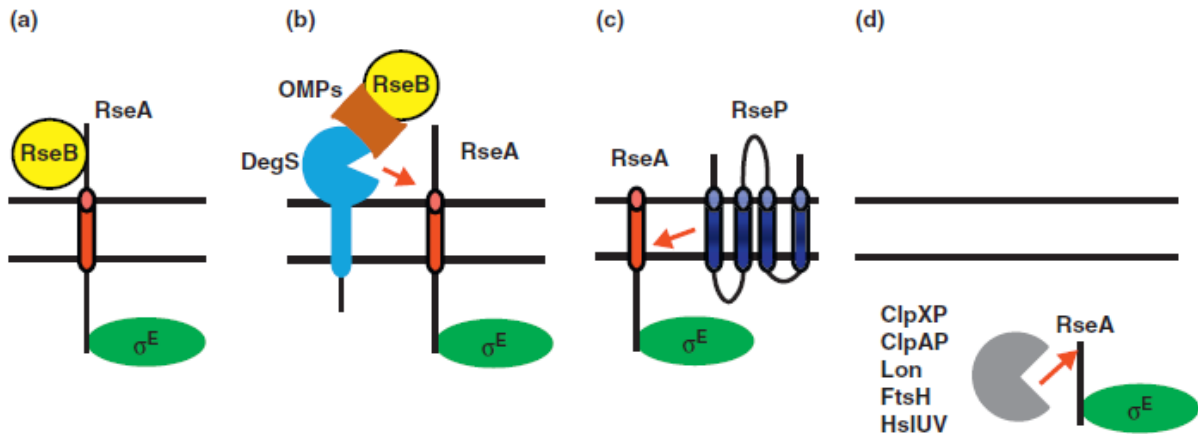


Abb. II-5 Modell der Aktivierung von SigE in *E. coli*. (a) In Abwesenheit von Stress bindet der Anti-Sigmafaktor RseA mit dem N-terminalen cytoplasmatischen Bereich an SigE und unterbindet dessen Aktivität. Das periplasmatische Protein RseB moduliert die Aktivität von SigE über RseA. (b) Ungefaltete Proteine der äußeren Membran (*Outer Membran Protein, OMP*) binden an DegS (periplasmatische Protease) und RseB. Dies führt zu einer Abspaltung der C-terminalen Region des Anti-Sigmafaktor RseA durch DegS. (c) Das verkürzte RseA wird in einem zweiten proteolytischen Schritt durch die Protease RseP nahe der Transmembrandomäne gespalten. (d) Der Komplex aus abgespaltenem N-terminalen Rest des Anti-Sigmafaktor RseA und SigE liegt nun frei im Cytoplasma vor. Die finale proteolytische Inaktivierung des nun verkürzten cytoplasmatischen Anti-Sigmafaktors und damit Aktivierung von SigE erfolgt durch Degradierung durch eine der cytosolischen Proteasen (ClpXP, ClpAP, Lon, FtsH oder HslUV). Rote Pfeile deuten die Position der Spaltung durch die Proteasen an. Abbildung aus (Ho und Ellermeier, 2012).

In manchen *Bacillus*-Arten wird die Aktivität von Anti-Sigmafaktoren noch weit komplexer reguliert. Hier erfolgt die Regulation über Anti-Anti-Sigmafaktoren, die den Anti-Sigmafaktor inaktivieren können und somit zur Reaktivierung des Sigmafaktors führen und Ko-Anti-Sigmafaktoren, die zusammen mit ihrem Anti-Sigmafaktor die Aktivität von einem entsprechenden Sigmafaktor inhibieren oder umlenken können. In Corynebakterien ist eine solche Regulation allerdings nicht bekannt (Campbell *et al.*, 2008).

Die genannten Mechanismen der Inaktivierung der Anti-Sigmafaktoren führen zu der Freisetzung eines ECF-Sigmafaktors, der dann zu der Umprogrammierung der Genexpression führt, in dem er an das RNAP-Core Enzym bindet und die Spezifität für spezifische Promotoren vermittelt.

3. Die Sigmafaktoren von *Corynebacterium glutamicum*

In Corynebakterien gibt es generell keine Sigmafaktoren der Familie Sigma⁵⁴ und keine der Gruppe 3 der Sigma⁷⁰-Familie. Neben dem Hauptsigmafaktor SigA (Gruppe 1) und dem alternativen SigA-ähnlichen Sigmafaktor SigB (Gruppe 2) besitzen Corynebakterien einen gemeinsamen Kernsatz aus ECF-Sigmafaktoren (SigC, SigE, SigH und SigM) der in allen sequenzierten Corynebakterien vorkommt (Patek und Nesvera, 2011). Daneben existieren eine Reihe weiterer ECF-Sigmafaktoren, die in vereinzelt Corynebakterien vorkommen. *C. glutamicum* besitzt neben dem Kernsatz an corynebakteriellen ECF-Sigmafaktoren (SigC, SigE, SigH und SigM) zusätzlich den ECF-Sigmafaktor sigD.

3.1. Der Hauptsigmafaktor SigA

Der Sigmafaktor SigA ist der Hauptsigmafaktor in *C. glutamicum* und ist verantwortlich für die Transkription von *Housekeeping*-Genen (Oguiza *et al.*, 1997). Die Transkription des *sigA*-Gens ist während des exponentiellen Wachstums gleichbleibend hoch und schwächt sich in der Transitionsphase, dem Übergang vom exponentiellen Wachstum zur stationären Phase, ab (Larisch *et al.*, 2007). Patek und Mitarbeiter untersuchten 159 SigA-abhängige Promotoren und konnten die Nukleotidverteilung an jeder Position des -35- und -10-Motivs für diese Promotoren darstellen (Abb. II-6). Die daraus resultierende statistische Konsensussequenz für die -35-Region und die erweiterte -10-Region lautet TTGNC bzw. GNTANANTNG. Dabei sind fett gedruckte Nukleotide in mehr als 80 % der Sequenzen vorhanden, die anderen in mehr als 35 %, wobei die Kernmotive unterstrichen sind.

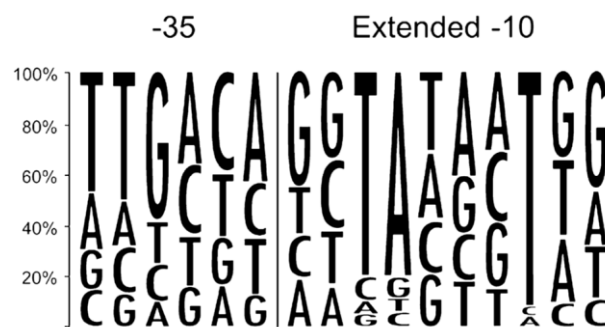


Abb. II-6 Verteilung der Nukleotide in der -35- und erweiterten -10-Region der *C. glutamicum* Promotoren von *Housekeeping*-Genen. Die Verteilung der Nukleotide an den angegebenen Positionen ist in Prozent durch die Höhe der Nukleotidsymbole (A, C, G, T) dargestellt. Das -10-Kernmotiv ist unterstrichen. Die Analyse basiert auf 159 SigA-abhängigen Promotoren. Abbildung aus (Patek und Nesvera, 2011).

In den meisten Sequenzen ist das -35-Motiv nur schwer zu erkennen und in vielen Promotorsequenzen fehlt dieses Motiv ganz. Es spielt somit keine entscheidende Rolle bei der Identifizierung durch SigA (Patek und Nesvera, 2011). Die erweiterte -10-Region jedoch stellt ein stärker konserviertes Motiv der SigA-abhängigen Promotoren in *C. glutamicum* dar (Pátek *et al.*, 2003).

3.2. Der nicht-essentielle Sigmafaktor SigB

SigB ist ein nicht-essentieller *primary-like* Sigmafaktor der Gruppe 2. Die Transkription des *sigB*-Gens ist während des exponentiellen Wachstums geringer als die des primären Sigmafaktorgens *sigA* (Oguiza *et al.*, 1997). Die Transkriptmenge von *sigB* erhöht sich erst in der Transitionsphase (Larisch *et al.*, 2007), was mit einer gleichzeitigen Verringerung der Transkriptmenge des Gens *sigA* einhergeht. Vermutlich findet eine Verdrängung des Sigmafaktors SigA durch SigB an der RNAP statt und dies erzeugt die vermehrte Transkription SigB-abhängiger Gene. Die Transkription von *sigB* ist in Antwort auf Säure-, Salz-, Kälte-, Hitze- und Ethanolstress (Halgasova *et al.*, 2001), Sauerstoffmangel (Ehira *et al.*, 2008), Adaption an Milchsäure (Jakob *et al.*, 2007) und bei der stringenten Antwort (Brockmann-Gretza und Kalinowski, 2006) zusätzlich erhöht. Die erhöhte Transkription vieler Gene in der Transitionsphase ist von SigB abhängig (Larisch *et al.*, 2007). Die Gene, die in der Transitionsphase nur in Anwesenheit von SigB verstärkt transkribiert werden, sind am Aminosäuremetabolismus, an Stressbekämpfungsmechanismen, an Membranprozessen und allgemein an regulatorischen Prozessen beteiligt. Durch die verstärkte Transkription der Gene, die an Stressabwehrmechanismen beteiligt sind, passen sich die Zellen an Bedingungen wie Nährstoffmangel, oxidativem Stress und andere Arten von Umgebungs-Stress an, die im Verlaufe der Stationärphase zunehmen (Patek und Nesvera, 2011). SigB aus *C. glutamicum* könnte somit eine ähnliche Funktion haben wie Sig³⁸, dem Hauptregulator der allgemeinen Stressantwort in *E. coli* (Becker und Hengge-Aronis, 2001) und SigB aus *M. tuberculosis*, der an der Adaption an die stationäre Phase und an Nährstoffmangelbedingungen beteiligt ist (Sachdeva *et al.*, 2010). Zusätzlich zu den Genen, die in an der Stressantwort beteiligt sind, wurde in *C. glutamicum* eine SigB-abhängige Transkription von Genen der Glukoseverwertung unter Sauerstoffmangel beobachtet (Ehira *et al.*, 2008).

Für das *sigB*-Gen aus *C. glutamicum* wurde der TSP bestimmt (Halgasova *et al.*, 2001). Der dazugehörige Promotor ist SigH-abhängig (Ehira *et al.*, 2009), was in Übereinstimmung mit dem Einfluss von SigB auf verschiedene Stressbedingungen steht. In *M. tuberculosis* ist die Regulation der Transkription des *sigB*-Gens komplex (Sachdeva *et al.*, 2010). Ein Promotor, der Ähnlichkeiten zu dem Promotor des *sigB*-Gens in *C. glutamicum* aufweist, wird von drei

unterschiedlichen Sigmafaktoren erkannt (SigB, SigE und SigL) und ein zweiter Promotor ist SigF-abhängig.

Ein *Alignment* der Promotorregionen von Genen die in der Transitionphase verstärkt transkribiert werden und einiger Gene des Glucosemetabolismus die SigB-abhängig sind ergab die Promotorkonsensussequenz CGGCAA-N(17-18)-TANAAT. Die verwendeten Promotoren enthalten ein -10-Promotormotiv, dass nicht von dem SigA-abhängiger Promotoren zu unterscheiden ist (Patek und Nesvera, 2011). Weiter sind nur geringe Unterschiede in den -35-Motiven SigB- und SigA-abhängiger Promotoren vorhanden (Ehira *et al.*, 2008). Dieses Ergebnis unterstützt eine frühere Hypothese (Oguiza *et al.*, 1997), dass die Sigmafaktoren SigA und SigB beide ähnliche Promotoren in *C. glutamicum* erkennen (Schröder und Tauch, 2010).

Die Beobachtungen für SigB-abhängige Promotoren in *C. glutamicum* sind ähnlich zu den Erkenntnissen zu Sig³⁸-abhängigen Promotoren in *E. coli*, wo kein klares Motiv der -35-Region vorhanden ist, und sich die -10-Region von Sig³⁸- und Sig⁷⁰-abhängigen Promotoren nicht unterscheiden lässt (Weber *et al.*, 2005). Für einige Promotoren konnte sowohl eine Sig³⁸- als auch Sig⁷⁰-abhängige Transkription nachgewiesen werden (Olvera *et al.*, 2009).

Die Transkriptionsinitiation von diesen Promotoren aus erfolgt während des exponentiellen Wachstums vorwiegend in Abhängigkeit von Sig⁷⁰. Sind die Wachstumsbedingungen nicht optimal, erfolgt die Transkriptionsinitiation in Abhängigkeit von Sig³⁸. Man nimmt an, dass dieser Wechsel zwischen Sig⁷⁰- und Sig³⁸-abhängiger Transkription in *E. coli* auf der Menge der Sigmafaktoren und ihrer Affinität zu dem RNAP-Core Enzym basiert (Sharma und Chatterji, 2010). Eine ähnliche Situation ist auch für *C. glutamicum* SigA und SigB vorstellbar (Pátek *et al.*, 2003). Zusammenfassend lässt die Aufgabe von SigB als eine Art *Back-up* zu SigA beschreiben, der die Initiierung der Transkription an (bestimmten) SigA-ähnlichen Promotoren während der Stationärphase oder unter anderen nicht optimalen Bedingungen gewährleistet.

3.3. ECF-Sigmafaktoren in *Corynebacterium glutamicum*

Die ECF-Sigmafaktoren von *C. glutamicum* sind alle in unterschiedliche Hauptgruppen der ECF-Sigmafaktoren eingeteilt (Staroń *et al.*, 2009). SigC gehört der Gruppe ECF36 und SigD der Gruppe ECF40 an. Die Mitglieder der ECF-Gruppen 36 bis 40 kommen ausschließlich in Actinobakterien vor mit Ausnahme der ECF-Gruppe 37, die auf Proteobakterien beschränkt ist. Nur drei Vertreter dieser Gruppen wurden bis jetzt analysiert. SigC (ECF36) aus *M. tuberculosis* wurde als wichtiger Regulator der Pathogenität und der Adaption des Überlebens in Lunge und Milz beschrieben. SigD (ECF40), ebenfalls aus *M. tuberculosis*, wird auch eine Bedeutung in Zusammenhang mit der Pathogenität zugeschrieben sowie die Regulation ribosomaler Gene und ATP-Transporter.

Mit Ausnahme von ECF36, besitzen vermutlich alle Sigmafaktoren dieser Gruppen einen Anti-Sigmafaktor, wenn auch teilweise ohne die normalerweise konservierten charakteristischen Domänen (Staroń *et al.*, 2009).

Der *C. glutamicum* ECF-Sigmafaktor SigH wird zur ECF12-Gruppe und SigE zur ECF14-Gruppe gezählt. Die Proteine der Untergruppe ECF11 bis ECF15 sind untereinander nicht nahe verwandt (Staroń *et al.*, 2009). Die Mitglieder dieser Gruppen vereint das Vorhandensein eines nicht in der Membran verankerten Anti-Sigmafaktors, was die Vermutung nahe legt, dass diese Sigmafaktoren cytoplasmische Stressbedingungen sensieren. Die Anti-Sigmafaktorproteine der ECF-Gruppen 11 bis 14 besitzen alle eine ZAS-Domäne (Zink bindende Anti-Sigma Domäne) und alle bisher untersuchten Mitglieder haben ihre Funktion in der Sensierung des Redox-Gleichgewichts der Zelle und damit verbunden bei der Antwort auf oxidativen Stress.

Der ECF-Sigmafaktor SigM aus *C. glutamicum* gehört der ECF-Gruppe 27 an. Diese Gruppe ist auf ECF-Sigmafaktoren aus Actinobakterien beschränkt (Staroń *et al.*, 2009) Für diese Gruppe ist nur die Funktion von SigM aus *C. glutamicum* näher untersucht worden (Nakunst *et al.*, 2007). Für einige Mitglieder dieser Gruppe konnten Anti-Sigmafaktoren identifiziert werden. Allerdings konnte bei *C. glutamicum* SigM bisher ein solcher Anti-Sigmafaktor nicht identifiziert werden.

3.4. Der ECF-Sigmafaktor SigH

SigH reguliert in *C. glutamicum* und in anderen Actinobakterien die Antwort auf Hitze- und oxidativen Stress. Die Aktivität des Promotors von *sigH* ist laut den Ergebnissen einer Reportengenfusionsanalyse während der stationären Phase und nach oxidativem Stress verstärkt (Kim *et al.*, 2005). In *Microarray*-Analysen konnte allerdings keine erhöhte Transkriptmenge des *sigH*-Gens nach Hitzestress oder der Transitionsphase beobachtet werden (Larisch *et al.*, 2007). Diese Diskrepanz kann z. B. mit dem Einfluss negativer TF auf die Transkription von *sigH*, die durch die Verwendung der Operatorregion in *multicopy* Plasmiden, die für die Reportegenfusionsanalyse genutzt wurden, abgeschwächt ist (Patek und J Nesvera, 2011).

SigH ist in die Transkription des *primary-like* Sigmafaktors SigB (Ehira *et al.*, 2008) und des ECF-Sigmafaktors SigM (Nakunst *et al.*, 2007) und seiner eigenen Regulation involviert (Kim *et al.*, 2005). Die von Kim und Mitarbeitern berichtete Autoregulation konnte in Folgearbeiten allerdings nicht bestätigt werden (Ehira *et al.*, 2009). Das SigH-Regulon umfasst eine große Anzahl Regulatoren, die an der Stressantwort beteiligt sind wie HspR (Barreiro *et al.*, 2009), ClgR (Engels *et al.*, 2004) und SufR (Ehira *et al.*, 2009). SigH nimmt somit eine zentrale Position in den vernetzten Regulationsnetzwerken von Sigmafaktoren und weiteren Regulatoren ein, die an der Antwort auf verschiedene Stressbedingungen beteiligt sind. Es wird angenommen, dass SigH als globaler Regulator die Antwort auf Hitze- und oxidativen Stress und eine Reihe anderer Stressbedingungen die während der stationären Phase immanent werden, kontrolliert (Schröder und Tauch, 2010). Die Präsenz des *sigH*-Gens in allen Corynebakterien und die hohe Ähnlichkeit in den Corynebakterien, aber auch in Mycobakterien unterstützt diese Annahme. Auch die bekannten Funktionen ähneln sich stark. Der ECF-Sigmafaktor SigH aus *M. tuberculosis* kontrolliert ebenfalls die Transkription von weiteren Sigmafaktoren (SigB, SigE) und weiteren TF (Sachdeva *et al.*, 2010).

Die Aktivität des SigH-Orthologs von *M. tuberculosis* bzw. *S. coelicolor* wird durch den Anti-Sigmafaktor RshA bzw. RsrA kontrolliert (Song *et al.*, 2003; Kang *et al.*, 1999). Der Anti-Sigmafaktor und gleichzeitige Stresssensor bindet in *M. tuberculosis* unter ungestressten Bedingungen den ECF-Sigmafaktor SigH und unterbindet so die SigH-abhängige Transkription von Genen. Durch oxidativen Stress oder Hitzestress wird der Sigmafaktor-Anti-Sigmafaktor-Komplex aufgelöst, SigH kann an das RNAP-Core Enzym binden und die SigH-abhängige Transkription von Genen ist möglich (Song *et al.*, 2003). In *C. glutamicum* ist das zu *M. tuberculosis* homologe Anti-Sigmafaktorgen *rshA* nur zwei bp stromabwärts des *sigH*-

Gens angeordnet. Ob RshA in *C. glutamicum* die gleiche Funktion, wie in *M. tuberculosis* erfüllt ist bisher nicht gezeigt.

Eine größere Anzahl SigH-abhängiger Gene wurde in einer Arbeit zu der Hitzeschockantwort in *C. glutamicum* durch Nutzung von DNA-Microarrays und *sigH*-Deletionsmutanten bzw. Überproduzenten identifiziert (Ehira *et al.*, 2009). In dieser Arbeit von Ehira und Mitarbeitern konnten 45 Gene, organisiert in 29 Transkriptionseinheiten, als SigH-abhängig transkribiert identifiziert werden. Es handelt sich um Gene, die an der Bekämpfung des oxidativen Stresses beteiligt sind und um Hitzeschock-relevante Gene. Einige dieser Gene, wie *trxB* und *trxC* wurden allerdings in einer früheren Arbeit auch als SigM-abhängig beschrieben (Nakunst *et al.*, 2007). Die Kernmotive –35 und –10 dieser SigH-abhängigen Promotoren (G/TGGAA–17-20nt–T/CGTT) sind hochkonserviert (Patek und Nesvera, 2012) und entsprechen den Konsensussequenzen von SigH- bzw. SigR-abhängigen Promotoren in *M. tuberculosis* bzw. *S. coelicolor* (Raman *et al.*, 2001; Paget *et al.*, 2001). Das Konsensuspromotormotiv mycobakterieller SigE-abhängiger Gene zeigt hohe Ähnlichkeit zu dem SigH-abhängiger Promotoren. In *M. tuberculosis* lautet die Konsensussequenz SigE-abhängiger Promotoren gGGAACYa–N15-16–cGTT und wird teilweise auch von SigH erkannt (Song *et al.*, 2008). Eine solche duale Kontrolle SigH-abhängiger Promotoren ist für *C. glutamicum* bisher nicht bekannt.

3.5. Der ECF-Sigmafaktor SigM

SigM ist an der Regulation der Antwort auf oxidativen Stress und Hitzestress beteiligt (Nakunst *et al.*, 2007). Entgegen der anderen ECF-Sigmafaktoren aus *C. glutamicum* teilt SigM eine nur 66 %ige Identität mit SigM aus *C. efficiens*. Die Ähnlichkeit zu SigM aus *M. tuberculosis* beträgt nur 35 %. Die Transkription des *sigM*-Gens ist nach Hitze-, Kälte-, Disulfid- und Peroxidstress erhöht. Dies ist in Übereinstimmung mit der Beobachtung, dass die Deletionsmutante $\Delta sigM$ sensitiv gegenüber eben diesen Stressbedingungen ist. Das *sigM*-Gen wird durch einen SigH-abhängigen Promotor transkribiert (Nakunst *et al.*, 2007). Im Gegensatz zu den Regulationsnetzwerken von SigH und SigB besteht das Netzwerk von SigM aus nur einer kleinen Gruppe von Genen. Diese Gene wurden in einem *Microarray*-Experiment identifiziert, in dem eine *sigM*-Deletionsmutanten mit dem Vorläuferstamm RES167, ein Restriktions-defektes Derivat des Typstammes ATCC 13032, verglichen wurde. Es handelt sich um die Thioredoxinmetabolismogene *trxB*, *trxB1* und *trxC*, sowie das *suf*-Operon, das eine Funktion bei der Reparatur von Eisenschwefelclustern in Proteinen spielt. Alle diese Genfunktionen sind an der Antwort auf Disulfidstress beteiligt. Zusätzlich wurden einige Gene identifiziert, die an

der Antwort auf Hitzestress beteiligt sind (Nakunst *et al.*, 2007). Durch die Bestimmung der TSP konnten vier Promotoren identifiziert werden, die das zu *sigH*-Promotoren identische –35-Motiv GGAA und das –10-Motiv GTT tragen. Die Promotoren der Gene *trxB*, *trxC* und *sufR*, die in der Arbeit mit *C. glutamicum* RES167 als SigM-abhängig beschrieben worden sind, wurden in einer Arbeit mit *C. glutamicum* R als SigH-abhängig beschrieben (Ehira *et al.*, 2009). Diese Gene waren in einer *sigH*-Deletionsmutante in ihrer Transkriptmenge im Vergleich zum Referenzstamm verringert und in einem *sigH*-Überexpressionsstamm erhöht. Die erhöhte Transkriptmenge dieser Gene und weiterer Hitzeschock relevanter Gene konnte durch Hitzestress auch in einer *sigM*-Deletionsmutante beobachtet werden (Ehira *et al.*, 2009).

Diese Ergebnisse deuten auf eine Erkennung durch beide Sigmafaktoren hin (Patek und Nesvera, 2011). In *M. tuberculosis* ist dies unter anderem für den Promotor des Gens *sigB* bekannt. Das *sigB*-Gen wird unter Hitzestress SigH-abhängig und unter Oberflächenstress SigE-abhängig transkribiert (Manganelli *et al.*, 2002). Die Erkennung eines Promotors durch mehrere Sigmafaktoren ist für *B. subtilis* ebenfalls bereits beschrieben (Qiu und Helmann, 2001).

3.6. Der ECF-Sigmafaktoren SigE

SigE ist an der Antwort auf Zelloberflächenstress und der Hitzeschockantwort in *C. glutamicum* beteiligt. Es wurde beobachtet, dass eine $\Delta sigE$ -Mutante sensitiv gegenüber Hitze, SDS, EDTA, Lysozym und verschiedenen Antibiotika ist (Park *et al.*, 2008). Die Deletionsmutante zeigte eine erhöhte Sensitivität gegenüber den Zellwandantibiotika Penicillin bzw. Vancomycin, was mit der Sensitivität gegenüber Detergenzien in Übereinstimmung steht, die ebenfalls an der Zellhülle wirken. Da neben Antibiotika, die auf die Stabilität der Zellwand einwirken, auch eine Sensitivität gegen Antibiotika mit anderen Wirkmechanismen besteht (Nalidixinsäure und Tetracyclin), wurde davon ausgegangen, dass dieser Umstand eventuell auf eine instabilere Zellhülle der Deletionsmutante $\Delta sigE$ zurückzuführen ist. Park und Mitarbeiter begründen die potentiell instabilere Membran der Deletionsmutante mit dem Fehlen einer putativen Methyltransferase, deren Gen (*cg1266*) SigE-abhängig transkribiert zu werden scheint (Park *et al.*, 2008). *Cg1266* stellt das einzige bis jetzt als SigE-abhängig beschriebenes Gen in *C. glutamicum* dar. Die SigE-abhängige Expression des Gens wurde allerdings nur indirekt durch einen komparativen Proteomikansatz postuliert. Hier konnte das Cg1266-Protein durch 2D-Proteingelelektrophorese in einer $\Delta sigE$ -Deletionsmutante nicht nachgewiesen werden. Ob *cg1266* unter der primären Kontrolle oder aber sekundär von SigE abhängig ist wurde in der Arbeit von Park *et al.* nicht geklärt.

Die Transkription des Gens *sigE* ist während des exponentiellen Wachstums schwach und bei Zelloberflächenstress oder Hitzeschock erhöht (Park *et al.*, 2008). Für *C. glutamicum* sind weder die TSP des Gens *sigE* bekannt, noch die Regulation der *sigE*-Transkription verstanden. Die *sigE*-Transkriptmenge ist unter weiteren Bedingungen im Vergleich zum Wildtyp erhöht, so unter Milchsäurestress (Jakob *et al.*, 2007), Stickstoffmangel (Silberbach *et al.*, 2005) oder in einem *C. glutamicum rel*-Deletionsstamm, der keine Stringente Kontrolle mehr ausüben kann (Brockmann-Gretza und Kalinowski, 2006). Dies legt die Vermutung nahe, dass SigE neben einer Hauptaufgabe, der Bekämpfung von Zelloberflächenstress, zusätzlich noch in die Regulation weiterer zellulärer Prozesse involviert ist.

Inwieweit SigE unter den verschiedenen Bedingungen tatsächlich aktiv ist hängt von der Aktivität des dazugehörigen Anti-Sigmafaktors (CseE) ab. Das Anti-Sigmafaktorgen *cseE* liegt unmittelbar stromabwärts des *sigE*-Gens nur durch eine kurze intergenische Region getrennt. Beide Gene werden als Operon transkribiert (Park *et al.*, 2008). Die Funktion von CseE als Anti-Sigmafaktor von SigE wurde durch den Nachweis der Interaktion der beiden Proteine *in vitro* und die phänotypische Komplementation der $\Delta cseE$ -Mutante durch die zusätzliche Deletion des *sigE*-Gens abgeleitet (Park *et al.*, 2008). Das CseE-Protein besitzt vier konservierte Cys-Reste, die für die Interaktion mit SigE verantwortlich sein könnten, ähnlich wie bei der Interaktion von RshA mit SigH (Park *et al.*, 2008). Ein Mechanismus für die Inaktivierung der CseE-SigE-Interaktion und damit die Aktivierung von SigE ist nicht bekannt.

Der ECF-Sigmafaktor SigE von *C. glutamicum* scheint eine ähnliche Funktion wie SigE in *M. tuberculosis* zu besitzen, der eine wichtige Rolle bei dem Überleben der Zellen unter Zelloberflächen- und Hitzestress spielt und auch durch einen Anti-Sigmafaktor reguliert wird. In *M. tuberculosis* wird *sigE* von drei unterschiedliche Promotoren aus transkribiert: Der Promotor P1 wird unter normalen physiologischen Bedingungen genutzt (Donà *et al.*, 2008), der Promotor P2 ist durch MprAB (einem 2-CS) reguliert, welches durch Oberflächenstress oder alkalischen pH induziert wird (He *et al.*, 2006), und der dritte Promotor P3 ist SigH-abhängig und unter oxidativem Stress oder Hitzeschock aktiv. Die Konsensussequenz SigE-abhängiger Promotoren lautet in *M. tuberculosis* gGGAACYa-N15-16-cGTT und wird teilweise auch von SigH erkannt (Song *et al.*, 2008). Bis jetzt ist für *C. glutamicum* kein direkt in Abhängigkeit von SigE transkribiertes Gen oder die Promotorkonsensussequenz bekannt.

3.7. Die ECF-Sigmafaktoren SigC und SigD

Die ECF-Sigmafaktoren SigC und SigD wurden in *C. glutamicum* bis jetzt nicht intensiv untersucht. Somit liegen zum jetzigen Zeitpunkt keine Informationen zu der physiologischen Funktion von SigC vor. Für SigD ist ein putativer Anti-Sigmafaktor (*cg0697*) direkt stromabwärts des Sigmafaktorgens *sigD* (*cg0696*) identifiziert worden, der anstelle einer eindeutigen Anti-Sigmadomäne (ASD) eine *Candidate Anti-Sigma Sequence* (CAS) besitzt und damit eine eher ungewöhnliche Struktur aufweist. Diese ungewöhnliche Struktur des Anti-Sigmafaktors ist unter den Anti-Sigmafaktoren der ECF-Gruppe 40 konserviert (Staroń *et al.*, 2009).

Im Gegensatz zu der Funktion von SigC gibt es für die Funktion von SigD Vermutungen. Komplementationsanalysen von *C. glutamicum*-Mutanten, die für das Wachstum höhere Konzentrationen an Sauerstoff im Vergleich zum Wildtyp benötigten, lassen vermuten, dass SigD an der Transkription von Genen beteiligt ist, die für Wachstum bei niedriger Sauerstoffkonzentration essentiell sind (Ikeda *et al.*, 2009). In einer weiteren Studie war die Transkription des *sigD*-Gens in einer *rel*-Deletionsmutante von *C. glutamicum* im Vergleich zum Wildtyp verringert (Brockmann-Gretza und Kalinowski, 2006). *C. glutamicum* ist durch die Deletion des Gens *rel* nicht fähig das Signalmolekül (p)ppGpp unter Aminosäureverknappung zu synthetisieren (Brockmann-Gretza und Kalinowski, 2006). Für das orthologe Gen aus *M. tuberculosis* ist eine Verbindung von SigD und der Stringenten Kontrolle bereits bekannt. Eine *sigD*-Deletionsmutante (*M. tuberculosis* $\Delta sigD$) hatte im Vergleich zum Wildtyp geringere Transkriptmengen von Genen, die für Bestandteile der Ribosomen, Elongationsfaktoren, DNA-bindende Proteine und Enzymen der ATP-Synthese kodieren (Calamita *et al.*, 2005). Weiter ist *sigD* in *M. tuberculosis* unter Nährstoffmangel verstärkt exprimiert (Betts *et al.*, 2002), einer Bedingung, die zumindest teilweise Rel-abhängig reguliert wird.

4. Klassische Methoden zur Analyse von Sigmafaktornetzwerken

Klassisch werden Regulationsnetzwerke durch *in vivo* basierte Methoden der Transkriptomanalyse untersucht. Um die Aktion eines Regulators (z. B. TF oder Sigmafaktor) zu untersuchen, kann das entsprechende Regulatorgen deletiert oder überexprimiert werden. Das Transkriptionsprofil einer solchen Mutante wird dann mit dem Transkriptionsprofil des Wildtyps verglichen. Unterscheiden sich die relativen Transkriptmengen einzelner Gene in den Datensätzen signifikant voneinander, kann es sich bei diesen Genen um potentiell durch den Regulator kontrollierte Gene handeln. Falls es sich allerdings um einen Regulator handelt, der unter den gewählten Anzuchtbedingungen nicht aktiv ist, werden keine Unterschiede zwischen den Transkriptionsprofilen des Wildtyps und der Mutante erkennbar sein.

Um das Regulon eines solchen Regulators aufzudecken, müssen dementsprechend Bedingungen, z.B. Stressbedingungen oder Bereitstellung bestimmter Nährstoffe, gewählt werden, bei denen der zu untersuchende Regulator aktiv ist.

In *in vivo* Experimenten werden durch die gewählten induzierenden Bedingungen neben der primären Antwort der Zelle fast immer komplexe sekundäre Reaktionen durch den Einfluss auf den Metabolismus oder der Induzierung hierarchisch angeordneter und quervernetzter Regulationsnetzwerke hervorrufen, die der primären Antwort nachgeschaltet sind. Diese Tatsache erschwert die Identifizierung der primären Aktion eines Regulators bzw. Sigmafaktors. Weiter müssen unter der gewählten Bedingung zur Aktivierung des Regulationsnetzwerkes nicht alle Mitglieder des Netzwerkes in Deletionsmutante und Wildtyp unterschiedlich transkribiert werden. Es kann nicht ausgeschlossen werden, dass unter unterschiedlichen Bedingungen andere Gene in Abhängigkeit des Regulators differentiell transkribiert werden. Eine solche Situation ist denkbar, wenn verschiedene Aktivatoren und Repressoren die Transkription zusätzlich beeinflussen und nur unter bestimmten Bedingungen aktiv oder inaktiv sind. Sind potentielle Zielgene eines Regulators oder Sigmafaktors z. B. durch *Microarray*-Analyse identifiziert worden, ist es für ein genaues Verständnis der Wirkungsweise meist notwendig, die TSP zu bestimmen. Für die Analyse kleiner Regulons kann dies durch Methoden wie *Primer-Extension-Analyse* (PEX) oder *5'-rapid Amplification of 5'-Ends* (5'-RACE) einzeln für jedes Gen geschehen. Ausgehend von den TSP kann die stromaufwärts liegende Sequenz der entsprechenden Transkripte analysiert werden, um Bindemotive der Regulatoren bzw. Promotoren zu identifizieren.

5. RNA-Sequenzierung

Für größere Regulons, zum Beispiel solche, die von Sigmafaktoren oder globalen Regulatoren kontrolliert werden, ist die genomweite Bestimmung von TSP von Nutzen. Hierzu kommt verstärkt die Methode der RNA-Sequenzierung (RNA-Seq), d. h. die Hochdurchsatz-Sequenzierung von cDNA-Bibliotheken, zum Einsatz. Neben der Bestimmung des kompletten Transkriptoms eines Bakteriums für qualitative und quantitative Analysen können auch gezielt 5'-Enden angereichert, und damit alle unter den gewählten Bedingungen transkribierten TSP identifiziert, werden. Die Lokalisierung der entsprechenden Promotoren auf genomweiter Ebene sowie die Identifizierung noch nicht bekannter Transkripte (Sorek und Cossart, 2010) sind damit entscheidende Vorteile gegenüber klassischen Methoden wie der *Microarray*-Analyse (Hoen *et al.*, 2008).

In der Vergangenheit war diese Methode aufgrund der leichteren Zugänglichkeit der mit PolyA-Schwänzen versehenen mRNA auf Eukaryoten beschränkt (Van Vliet, 2010). Es konnten aber verschiedene Methoden entwickelt werden, die es ermöglichen, bakterielle Transkripte genomweit mit Einzelbasenauflösung zu analysieren (Van Vliet, 2010). Es handelt sich hierbei um alternative *Priming*-Methoden. Grundsätzlich werden in den meisten Protokollen zufällige Hexamer-Primer (*Random Priming*) (Passalacqua *et al.*, 2009; Perkins *et al.*, 2009; Yoder-Himes *et al.*, 2009) und, falls künstlich poly-adenylierte mRNA eingesetzt wird, Oligo(dT)-Primer (Frias-Lopez *et al.*, 2008), genutzt. Eine weitere Möglichkeit ist die Ligation spezifischer RNA-Sonden an die mRNA (Sittka *et al.*, 2008; Wurtzel *et al.*, 2010). In Verbindung mit Hochdurchsatz-Sequenzierungstechnologien (Roche 454, Illumina Genome Analyser und Applied Biosystems SOLiD) kann RNA-Seq damit auch bei Prokaryoten eingesetzt werden (Sorek und Cossart, 2010).

Wie bei klassischen Transkriptomanalysen wird auch bei RNA-Seq mit isolierter Gesamt-RNA gestartet. Da in Mikroorganismen in der Regel bis zu 80 % der isolierten RNA auf stabile RNA-Spezies (rRNA und tRNA) entfallen (Condon, 2007) ist es unabdingbar, diese vor der RNA-Seq abzureichern. In Extremfällen kann der Anteil der stabilen RNA-Spezies bis zu 99,5 % betragen (Rückert, persönliche Mitteilung). Eine erfolgreiche Reduktion der rRNA/tRNA-Fractionen von beispielsweise 95 % auf 50 % der RNA-Isolierung würde in einer 10-fachen Anreicherung der mRNA-Menge resultieren und ist somit ein Schlüsselfaktor einer erfolgreichen RNA-Seq in Prokaryoten (Sorek und Cossart, 2010). Für die Abreicherung stehen mittlerweile unterschiedliche Methoden zur Verfügung, wie rRNA *Capture* mit spezifischen Sonden, Degradation prozessierter RNA, selektive Poly-Adenylierung von mRNA oder rRNA *Capture* mittels spezifischer Antikörper (Sorek und Cossart, 2010).

Nach der Abreicherung der stabilen RNA muss in mehreren Schritten eine cDNA-Bibliothek erstellt werden. Da die eingesetzten Sequenziertechnologien (z. B. der Firma Illumina) eine bestimmte Länge der cDNAs voraussetzen, müssen entweder die Transkripte oder die cDNA fragmentiert werden (Abb. II-7). Zudem müssen Plattform-spezifische Adaptoren eingebracht werden. Das weitere Vorgehen nach der Abreicherung der stabilen RNA-Spezies ist daher extrem von der gewählten Art der Bibliothekserstellung abhängig (Van Vliet, 2010).

Mit den heute zur Verfügung stehenden *Next-Generation*-Sequenziertechnologien resultiert eine solche Sequenzierung in mehreren Millionen DNA-Sequenzen (*Reads*). In dieser Arbeit wurde der Illumina Genome Analyzer IIx verwendet. Die *Read*-Länge, die man mit dieser Plattform erhält, ist mit 35 - 150 Basen geringer als für andere *Next-Generation*-Sequenziertechnologien, wie z. B. Roche 454, aber vollkommen ausreichen für RNA-Seq von bekannten bakteriellen Transkriptomen. Die Anzahl an *Reads* ist mit bis zu 30 Millionen pro *Lane* einer acht *Lanes*-umfassenden *Flow Cell* höher als die anderer Technologien (aktueller Stand, Rückert, persönliche Mitteilung). Zudem erlaubt die Plattform die Sequenzierung von beiden Seiten (*Paired-End* Sequenzierung), wodurch das 3'-Ende von Transkripten bei der Bestimmung des vollständigen Transkriptoms deutlich besser erfasst werden kann.

Die *Reads* müssen zur bioinformatischen Analyse, wie TSP-Identifizierung, unter Verwendung von *Mapping*-Algorithmen auf einem Referenzgenom kartiert werden (Abb. II-7). Die in dieser Arbeit verwendeten bioinformatischen Methoden sind in den Folgekapiteln erläutert.

Eine Modifikation der RNA-Seq-Methode ist die sogenannte differentielle RNA-Seq (dRNA-Seq) (Sharma *et al.*, 2010). Bei dieser Methode werden prozessierte Transkripte, die ein 5'-Monophosphat enthalten, durch den Einsatz einer 5'-Monophosphat-abhängigen Exonuklease (*Terminator Exo-Nuclease*, TEN) abgebaut. Durch den Einsatz der TEN bleiben nur primäre/native Transkripte erhalten, die ein 5'-Triphosphat besitzen. Die auf Basis eines nativen 5'-Enden-angereicherten RNA-*Pools* erstellten ds-cDNA-Bibliotheken werden nur vom 5'-Ende aus sequenziert (Abb. II-7). Die Methode der dRNA-Seq ist besonders bei der eindeutigen Identifizierung von TSP und damit Promotorsequenzen vorteilhaft, da prozessierte RNA von der Analyse ausgeschlossen ist. Hiermit steht eine Methode zur Verfügung die gezielt zur Identifizierung von Promotoren auf einem genomweiten Level eingesetzt werden kann.

Diese nun mögliche Identifizierung eines genomweiten „Promoteroms“ ist ein großer Schritt in der qualitativen und auch quantitativen Betrachtungsweise von Promotoren, da die Methode zum einen eine Einzelnukleotidauflösung besitzt und zum anderen Auskunft über die relative Menge an mRNA mit intaktem (primärem) 5'-Ende gibt (Pátek *et al.*, 2013).

Um die TSP von *C. glutamicum* auf diese Weise genomweit untersuchen zu können, wurde in dieser Arbeit die Methode der dRNA-Seq genutzt. Hierfür wurde das durch Pfeifer-Sancar und Rückert 2013 entwickelte native 5'-Enden RNA-Seq-Protokoll verwendet (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung). Hierbei handelt es sich um eine Modifikation der dRNA-Seq, die sich in der Erstellung der cDNA-Bibliotheken von anderen Methoden unterscheidet. Um die Informationen über das 3'- und 5'-Ende eines Transkripts zu erhalten, wird zunächst ein Voll-RNA-Adapter an das 5'-Ende der nativen mRNA-Transkripte ligiert, und dann unter Verwendung eines *Loop*-DNA-Adapters die Erststrangsynthese durchgeführt. Die cDNA-Bibliothek wird final durch die Verwendung von Primern spezifisch für den 5'- und 3'-Adapter in einer *Low-Cycle* PCR durchgeführt (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung). Der Ablauf des nativen 5'-Enden RNA-Seq Protokolls inklusive der Sequenzierung und bioinformatischen Auswertung ist in Abb. II-7 grafisch zusammengefasst.

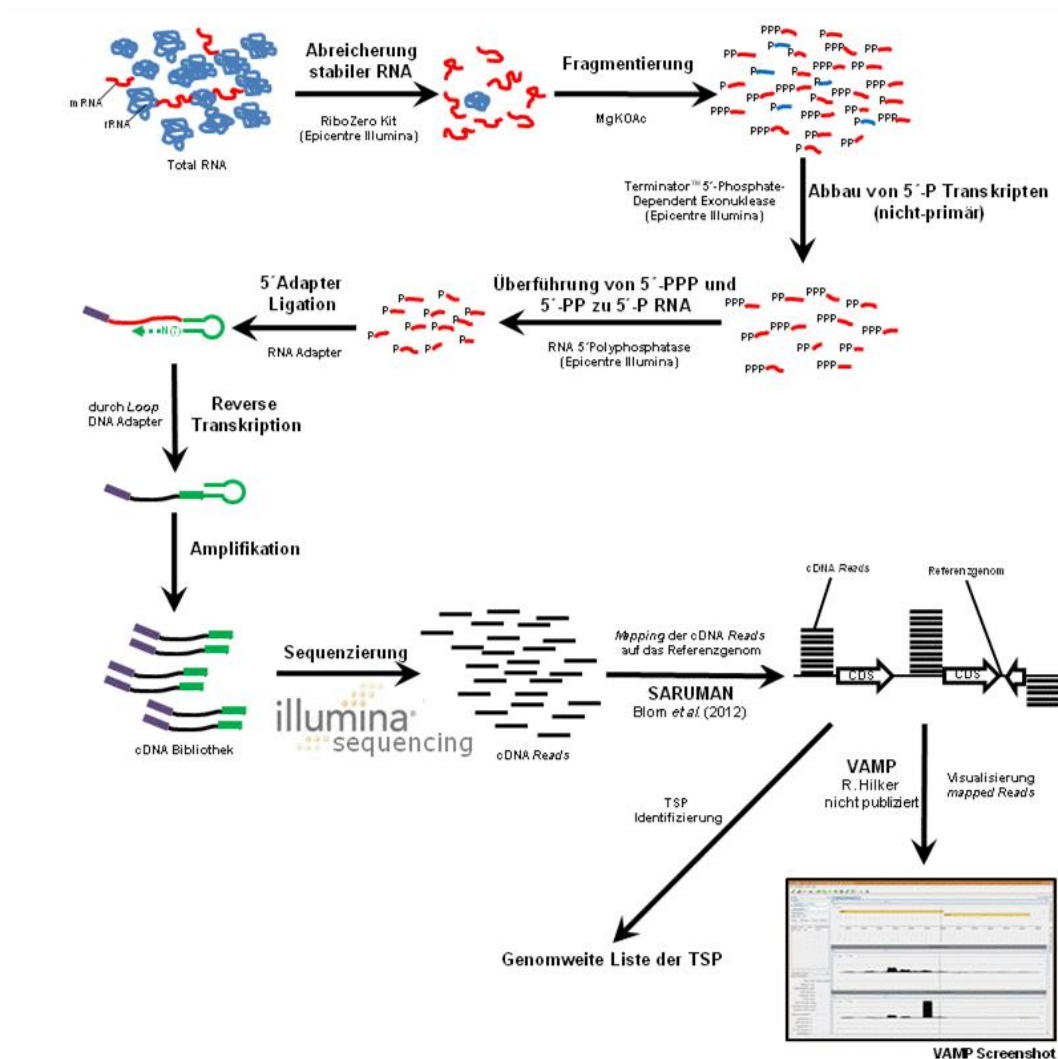


Abb. II-7 Erstellung einer nativen 5'-Enden-angereicherten RNA-Seq-Bibliothek inklusive Ablauf der Sequenzierung und bioinformatischen Analyse der cDNA-Bibliothek.

5.1. Bioinformatische Analyse der RNA-Seq-Daten

Die enorme Menge an kurzen cDNA-Reads, die durch die Sequenzierung mit den Hochdurchsatz-Sequenzieretechnologien entstehen, muss auf ein Referenzgenom *gemappt* werden. Mit einem Lauf einer *Flow Cell* im *Single-Read*-Modus wird eine Datenmenge von 10 - 12 Gb erzeugt. Theoretisch passieren 320 Millionen Reads den Filter der *Illumina Instrument* Software und müssen auf das Referenzgenom *gemappt* werden. Hierzu sind viele Rechenschritte nötig, was eine enorme Rechenleistung erfordert.

5.1.1. Mapping sequenzierter Reads durch SARUMAN

Die sequenzierten Reads werden unter Verwendung von SARUMAN (**S**emiglobal **A**lignment of short **R**eads **U**sing **C**UDA and **N**eedle**MAN**-Wunsch) (Jochen Blom et al., 2011) auf einem Referenzgenom kartiert. Die Nutzung von SARUMAN bietet mehrere Vorteile. Zum einen arbeitet es im Vergleich zu anderen *Mapping*-Methoden nicht mit einem heuristischen Modell, bei dem nicht alle möglichen *Alignment*-Positionen erkannt werden, sondern mit einem mathematisch exakten Ansatz. Zum anderen gibt SARUMAN für einen cDNA-Read alle möglichen Positionen an, wodurch sich diese Anwendung deutlich von anderen Anwendungen unterscheidet. Bei anderen Programmen werden meist nur die n besten Treffer für die gesuchte Sequenz ausgegeben und somit gehen wichtige Informationen für alternative *Alignment*-Positionen mit identischen Werten verloren. Zusammenfassend erklärt, gibt SARUMAN alle möglichen Positionen eines cDNA-Reads in der Referenzsequenz zusammen mit einem optimalen *Alignment* für jede dieser Positionen unter Berücksichtigung einer Fehlerrate (in dieser Arbeit 6 %, d. h. 1 Nukleotid-*missmatch* pro *Read*, bei einer *Read*-Länge von 26 bp) an. Ein weiterer Vorteil bei der Nutzung von SARUMAN ist die Geschwindigkeit, in der die Berechnungen laufen. Durch die Nutzung von CUDA (*Compute Unified Device Architecture*) laufen die Berechnungen der *Alignments* auf Grafikkarten, die in der Lage sind parallelisierte Programmabläufe (Datenparallelität) signifikant schneller als herkömmliche CPUs zu bearbeiten (Jochen Blom et al., 2011).

5.1.2. RNA-Seq-Datenvisualisierung durch VAMP

Für die Visualisierung der *gemappten Reads* wird die Software VAMP (*Visualization and Analysis of Mapped Sequences*) (Hilker et al., 2013, in Vorbereitung) genutzt. VAMP ermöglicht den Import und die Visualisierung einer Referenzsequenz und der dazugehörigen, durch SARUMAN *gemappten Reads* (*Read Mapping Data*) einer Sequenzierung. Die Referenzsequenz wird mit allen bekannten kodierenden Regionen (CDS) auf dem Vorwärts- oder Rückwärtsstrang im entsprechenden Leseraster dargestellt. Die RNA-Seq *Mapping*-

Daten werden in Form von *Tracks* dargestellt, in denen die Anzahl an *gemappten Reads* für eine Position des Referenzgenoms als *Peaks* dargestellt sind. Innerhalb der Software ist es möglich durch das Referenzgenom und der dazugehörigen *Tracks* zu navigieren und sich die *gemappten Reads* an unterschiedlichen Positionen des Genoms darstellen zu lassen. Dies erlaubt dem Nutzer zu ermitteln wie viele *Reads* sich an welcher Stelle des Referenzgenoms befinden. Es kann bis auf Einzelnukleotid-Auflösung vergrößert werden, sodass man die Einzelbasenpaar-genaue Position der gemappten *Reads* innerhalb der *Tracks* begutachten kann. Weiter ermöglicht VAMP die Visualisierung mehrerer *Tracks*, sodass dem Nutzer ermöglicht wird, verschiedene RNA-Seq-Datensätze (Transkriptome) miteinander zu vergleichen.

6. **Run-Off-Transcription/Microarray-Analysis (ROMA)**

Das Hauptproblem konventioneller oder klassischer Methoden der Analyse transkriptioneller Regulation durch einen einzigen Regulator *in vivo* (*Top-Down*) ist die Schwierigkeit, direkte primäre Effekte einer genetischen Veränderung oder chemischen bzw. physikalischen Behandlung von Zellen von indirekten sekundären Effekten zu unterscheiden. Ein weiteres Problem ist die zusätzliche Regulation durch weitere TF neben dem zu untersuchenden Regulator, die dessen Wirkung in Abhängigkeit von Versuchsbedingungen maskieren können. Die Problematik ist in Kapitel 0 eingehend dargestellt.

Die Nutzung der *in vitro* Methode *Run-Off-Transcription/Microarray-Analysis* (ROMA) bietet hierfür eine Lösung. ROMA erlaubt die Analyse der Aktivität eines Sigmafaktors oder TF in Abwesenheit aller *in vivo* auftretenden sekundären Effekte (Maclellan *et al.*, 2009). Während die Analyse der *in vivo* Transkriptome von Wildtyp und Mutanten, also von komplexen Systemen, als *Top-Down*-Ansatz bezeichnet werden kann, handelt es sich hier um die Rekonstruktion eines Systems aus einfachen und bekannten Bauteilen. Entsprechend handelt es sich um einen komplementären *Bottom-Up*-Ansatz.

ROMA nutzt aufgereinigte RNAP mit oder ohne regulatorischem Protein, um die spezifischen Transkriptionsprodukte auf Basis fragmentierter, linearer genomischer *Template*-DNA zu untersuchen. Im Fall der Untersuchung von Sigmafaktoren wird das RNAP-Core Enzym genutzt und als „*regulatorisches Protein*“ ein Sigmafaktor. Die *in vitro* transkribierte RNA wird durch eine *reverse* Transkriptase in cDNA überführt und durch konventionelle komparative *Microarray*-Analysen analysiert (z. B. Vergleich von mRNA, die in Anwesenheit eines Sigmafaktors transkribiert wurde, mit mRNA, die nur in Anwesenheit des RNAP-Core Enzym

transkribiert wurde). Dies ermöglicht die Identifizierung von Genen, die spezifisch von dem eingesetzten Regulator abhängig sind.

Neben den Vorteilen, die die *in vitro* Methode ROMA dank ihrer „Einfachheit“ gegenüber dem komplexen *in vivo* System bietet, unterliegt diese Methode auch einigen Beschränkungen, die nicht unberücksichtigt bleiben dürfen (Maclellan *et al.*, 2009). Transkription, die unter dem Einfluss unbekannter Bedingungen oder Chemikalien steht, oder durch zusätzliche nicht bekannte Aktivatoren initiiert wird, kann durch ROMA nicht untersucht werden. In einigen Fällen ist die Transkription abhängig vom *Supercoiling* der DNA, was zur Folge hat, dass die entsprechenden Transkripte in einem ROMA-Ansatz, bei dem lineare DNA als *Template* genutzt wird, deutlich unterrepräsentiert sind. Ein Beispiel für die Abhängigkeit eines Sigmafaktors von der Konformation der DNA ist die Promotor-Erkennung durch SigS in *E. coli*, die zum Teil vom Grad des *Supercoiling* der DNA abhängt (Kusano *et al.*, 1996). Weiter können durch ROMA falsch positive Ergebnisse durch eine ineffiziente Transkriptions-Termination *in vitro* generiert werden, was auch als *Read-Through* bezeichnet wird. Die Länge der *in vitro* Transkripte kann durch die fehlende oder unzureichende Termination bis zu 10 kb betragen (MacLellan *et al.*, 2008). Dadurch können Transkripte nach der *Microarray*-Analyse fälschlicherweise einem Promotor zugeordnet werden, von dem *in vivo* keine Transkripte einer solchen Länge aus transkribiert werden. Insbesondere der *Read-Through* in konvergente Gene ist *in vitro* häufiger zu beobachten als *in vivo* (Maclellan *et al.*, 2009). Unter Berücksichtigung dieser Limitierungen wurde die Methode ROMA in mehreren verschiedenen Arbeiten erfolgreich zur Analyse von Sigmafaktor- und Transkriptionsfaktor-Regulationsnetzwerken eingesetzt.

Eine frühe Version von ROMA wurde bereits 1984 im Chamberlin-Labor zur Identifizierung von Mitgliedern des SigD-Regulons in *B. subtilis* genutzt. Anstelle der *Microarray*-Analyse wurde 1984 die *in vitro* transkribierte RNA durch die Nutzung genomischer λ -Bibliotheken analysiert (Maclellan *et al.*, 2009). Der Begriff ROMA und die Einführung der Methode erfolgte bei der Untersuchung des SigW-Regulationsnetzwerkes von *B. subtilis* (Cao *et al.*, 2002), wobei hier noch *Macroarrays* eingesetzt wurden, wie auch in einer ähnlichen Studie zu SigX aus *B. subtilis* (Maclellan *et al.*, 2009). *Microarrays* rückten erst in den Folgejahren an die Position der *Macroarrays*. Neben den Sigmafaktoren SigW und SigX aus *B. subtilis* wurde auch das Regulon von SigM erfolgreich durch die Nutzung dieser Methode als Teil einer umfassenden Arbeit aufgedeckt (Maclellan *et al.*, 2009). In einer aktuellen Studie in *E. coli* konnte primär durch ROMA aufgeklärt werden, wie sich die kaum unterscheidbaren Sigma⁷⁰-abhängigen Promotoren (*Housekeeping*-Sigmafaktor) von SigS-abhängigen Promotoren (alternativer Sigmafaktor) differenzieren lassen, und somit das SigS-Regulon neu definiert werden (Maciag

et al., 2011). Da SigS aus *E. coli* wie oben beschrieben zum Teil von der Konformation der DNA abhängt (Kusano *et al.*, 1996), wurde in dieser Studie zusätzlich zu dem ROMA-Ansatz eine *in vitro* Transkription mit ausgewählten Promotoren auf *Supercoiled* DNA ausgeführt. Für die ausgewählten SigS-abhängigen Promotoren konnte hierbei allerdings kein Unterschied bei der Verwendung von linearer und *supercoiled* DNA in den beiden *in vitro* Ansätzen beobachtet werden (Maciag *et al.*, 2011). Neben der Analyse der regulatorischen Funktion von Sigmafaktoren wurde ROMA ebenfalls bei der Analyse von klassischen TF eingesetzt wie z. B. dem cAMP-Rezeptorprotein (Crp) aus *E. coli* (Zheng *et al.*, 2004).

7. Zielsetzung der Arbeit

Corynebacterium glutamicum ist aufgrund seines einfachen Sets von nur fünf ECF-Sigmafaktoren, die alle unterschiedlichen Familien zugeordnet werden können, besonders für die Untersuchung der ECF-Regulationsnetzwerke geeignet, und kann auch in dieser Hinsicht als Modellorganismus der *Actinomycetales* betrachtet werden. Ziel dieser Arbeit war die Aufklärung von ECF-Regulationsnetzwerken und ihren Interaktionen mit aktuellen Methoden der Transkriptomanalyse. Da im Lauf der Arbeit in der Arbeitsgruppe die Transkriptomanalyse durch Sequenzierung von cDNA (RNA-Seq) etabliert werden konnte, sollte diese Technologie aufgrund ihrer hohen Auflösung in Bezug auf Transkriptionsstarts und Transkriptmengen ebenfalls eingesetzt werden. Die RNA-Seq Methodik erlaubte dann auch die hier durchgeführte Entwicklung einer neuen Analysemethode von Sigmafaktor-Regulationsnetzwerken durch genomweite *in vitro* Transkription und die Bestimmung der Sigmafaktor-spezifischen Transkripte durch RNA-Seq.

Damit können Sigmafaktornetzwerke nun in komplementären *in vivo Top-Down*-Ansätzen durch Transkriptomanalysen in Sigmafaktor- und Antisigmafaktorgen-Mutanten und *in vitro Bottom-Up*-Ansätzen mittels Assemblierung spezifischer Sigmafaktor-RNAP-Komplexe und genomweiter *in vitro* Transkription analysiert werden. Beispielhaft sollten mit diesen experimentellen Ansätzen die Netzwerke der Sigmafaktoren SigH und SigE entschlüsselt und ihre Interaktion an gemeinsam genutzten Promotoren aufgeklärt werden.

8. Bisherige Publikationen

Teile dieser Arbeit wurden bereits in folgenden Publikationen veröffentlicht:

Busche, T., Silar, R., Pičmanová, M., Pátek, M., Kalinowski, J., 2012. Transcriptional regulation of the operon encoding stress-responsive ECF sigma factor SigH and its anti-sigma factor RshA, and control of its regulatory network in *Corynebacterium glutamicum*. BMC genomics **13**, 445.

Pátek, M., Holátko, J., Busche, T., Kalinowski, J., Nešvera, J., 2013. *Corynebacterium glutamicum* promoters: a practical approach. Microbial biotechnology (2):103-17

III. Material und Methoden

1. Bakterienstämme und Plasmide

1.1. Bakterienstämme

Tab. III.1 Verwendete Bakterienstämme

Bakterium	Relevanter Genotyp/Information	Quelle/Referenz
<i>E. coli</i>		
<i>E. coli</i> JM109	<i>endA1, recA1, gyrA96, thi, hsdR17</i> (r_k^- , m_k^+), <i>rθIA1, supE44, Δ(lac-proAB), F' traD36, proAB, laq^F lacZΔM15</i>	Yanisch-Perron <i>et al.</i> , 1985
<i>E. coli</i> DH5 α MCR	<i>F-endA1 supE44 thi-l λ-re-ca1 deoR Δ(lacZYA – argF) U169 Φ80dlacZΔM15 mcrA Δ(mrr-hsdRMS-mcrBC)</i>	Grant <i>et al.</i> , 1990
<i>E. coli</i> ER2566	IPTG induzierbarer T7 Expressionstamm zur Expression von Intein-CBD-Protein-Fusionen des IMPACT-Systems (F- λ - fhuA2 [lon] omplacZ::T7 gene1galsul A1 Δ (<i>mcrC-mrr</i>)144::IS10 R(<i>mcr-73::miniTn 10TetS</i>)2R (<i>zgb210::Tn10</i>)(<i>Tn10</i>) <i>endA1[dcml]</i>)	New England Biolabs
<i>E. coli</i> ER2566_pTXB1::sigA_Cterm	wie <i>E. coli</i> ER2566 + pTXB1::sigA_Cterm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigA_NTerm	wie <i>E. coli</i> ER2566 + pTYB11::sigA_Cterm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTXB11::sigC_CTerm	wie <i>E. coli</i> ER2566 + pTXB11::sigC_Cterm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigC_Nterm	wie <i>E. coli</i> ER2566 + pTYB11::sigC_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTXB1::sigD_CTerm	wie <i>E. coli</i> ER2566 + pTXB11::sigD_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigD_NTerm	wie <i>E. coli</i> ER2566 + pTYB11::sigD_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB2::sigE_CTerm	wie <i>E. coli</i> ER2566 + pTYB2::sigE_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigE_Nterm	wie <i>E. coli</i> ER2566 + pTYB11::sigE_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTXB1::sigH_CTerm	wie <i>E. coli</i> ER2566 + pTXB11::sigH_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigH_NTerm	wie <i>E. coli</i> ER2566 + pTYB11::sigH_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTXB1::sigM_CTerm	wie <i>E. coli</i> ER2566 + pTXB11::sigM_CTerm; (Amp ^R)	diese Arbeit
<i>E. coli</i> ER2566_pTYB11::sigM_NTerm	wie <i>E. coli</i> ER2566 + pTYB11::sigM_Nterm; (Amp ^R)	diese Arbeit
<i>C. glutamicum</i>		
RES167	Restriktionsdefekter <i>C. glutamicum</i> Stamm(Δ <i>cgIIIM-cgIIIR-cgIIIR</i>)	Tauch <i>et al.</i> , 2002
DN2	RES167 Deletion von <i>sigH</i>	Nakunst <i>et al.</i> , 2007
AS1	RES167 Deletion von <i>sigH-rshA</i>	Soshi, 2008
AH1	RES167 Deletion von <i>sigE</i>	Hüttges, 2005
RES167 Δ <i>rshA</i>	RES167 Deletion von <i>rshA</i>	Busche <i>et al.</i> , 2012
RES167 Δ <i>cseE</i>	RES167 Deletion von <i>cseE</i>	diese Arbeit

1.2. Plasmide

Folgende Plasmide kamen im Rahmen dieser Arbeit zum Einsatz.

Tab. III.2 Verwendete Plasmide

Plasmid	Relevanter Genotyp/Information	Quelle/Referenz
pK18 <i>mobsacB</i>	<i>sacB</i> , <i>lacZa</i> , <i>mcs</i> , (5,7 kb, Km ^R)	Schäfer <i>et al.</i> , 1994
pK18 <i>mobsacB</i> :: Δ <i>rshA</i>	pK18 <i>mobsacB</i> Derivat; <i>rshA</i> Deletionskonstrukt	Busche <i>et al.</i> , 2012
pK18 <i>mobsacB</i> :: Δ <i>cseE</i>	pK18 <i>mobsacB</i> Derivat; <i>cseE</i> Deletionskonstrukt	diese Arbeit
pET2	<i>E. coli</i> - <i>C. glutamicum</i> promoter-probe vector (7.5 kb, Km ^r , promotoloses <i>cat</i> Gen)	Vasicova <i>et al.</i> , 1998
pET2:: <i>sigH</i>	Chromosomales Fragment (550 bp) in pET2	Busche <i>et al.</i> , 2012
pET2:: <i>rshA</i>	Chromosomales Fragment (301 bp) in pET2	Busche <i>et al.</i> , 2012
pET2:: <i>sigH4</i>	P4 <i>sigH</i> Promotorregion (348 bp) in pET2	Busche <i>et al.</i> , 2012
pTXB1	C-terminaler Intein-CBD-Fusionsvektor; trägt eine Intein gekoppelte Chitin bindende Domäne (6,7 kb; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i>)	New England Biolabs
pTYB2	C-terminaler Intein-CBD-Fusionsvektor; trägt eine Intein gekoppelte Chitin bindende Domäne (7,5 kb; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i>)	New England Biolabs
pTYB11	N-terminaler Intein-CBD-Fusionsvektor; trägt eine Intein gekoppelte Chitin bindende Domäne (7,4 kb; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i>)	New England Biolabs
pTYB12	N-terminaler Intein-CBD-Fusionsvektor; trägt eine Intein gekoppelte Chitin bindende Domäne (7,4 kb; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i>)	New England Biolabs
pTXB1:: <i>sigA</i> _CTerm	C-terminale Intein-CBD-Fusion mit SigA aus <i>C. glutamicum</i> ; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigA</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB11:: <i>sigA</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigA aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigA</i> aus <i>C. glutamicum</i>	Esken, 2011
pTXB1:: <i>sigC</i> _CTerm	C-terminale Intein-CBD-Fusion mit SigC aus <i>C. glutamicum</i> ; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigC</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB11:: <i>sigC</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigC aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigC</i> aus <i>C. glutamicum</i>	Esken, 2011
pTXB1:: <i>sigD</i> _CTerm	C-terminale Intein-CBD-Fusion mit SigD aus <i>C. glutamicum</i> ; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigD</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB11:: <i>sigD</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigD aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigD</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB2:: <i>sigE</i> _CTerm	C-terminale Intein-CBD-Fusion mit SigE aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigE</i> aus <i>C. glutamicum</i>	(Esken, 2011)
pTYB11:: <i>sigE</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigE aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigE</i> aus <i>C. glutamicum</i>	Esken, 2011
pTXB1:: <i>sigH</i> _CTerm	C-terminale Intein-CBD-Fusion mit sigH aus <i>C. glutamicum</i> ; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigH</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB11:: <i>sigH</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigH aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigH</i> aus <i>C. glutamicum</i>	Esken, 2011
pTXB1:: <i>sigM</i> _CTerm	C-terminale Intein-CBD-Fusion mit SigM aus <i>C. glutamicum</i> ; Mxe GyrA intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigM</i> aus <i>C. glutamicum</i>	Esken, 2011
pTYB11:: <i>sigM</i> _NTerm	N-terminale Intein-CBD-Fusion mit SigM aus <i>C. glutamicum</i> ; Sce VMA1 intein, CBD, T7/lac Promotor, Amp ^R , ori, M13 ori, <i>mcs</i> , <i>sigM</i> aus <i>C. glutamicum</i>	Esken, 2011

2. Primer

Tab. III.3 Verwendete Primer (Hersteller Metabion)

Primer	DNA-Sequenz	Eigenschaften
Deletion/Plasmid Konstruktions Primer		
pk18mobsacb::ΔrshA		
<i>rshA_del1</i>	GATCTAGAATTC-AGGCACTGCCACTCCTTGAC	<i>EcoRI</i>
<i>rshA_del2</i>	CAGATGATGCTCGACACGAA-GAGTCGCTGCGGTTGAGATT	
<i>rshA_del3</i>	CGATGACGAATCTCAACCGC-TTCGTGTCGAGCATCATCTG	
<i>rshA_del4</i>	GATCTATCTAGA-CAGGCCACTAGATGGTCTCC	<i>XbaI</i>
pk18mobsacb::ΔcseE		
<i>V_ΔcseE_for</i>	CTTCAGGCAGCGCCTCCATGCGGGAATTCGTAATCATGGT	
<i>V_ΔcseE_rev</i>	CTCTTGGACCGATGAGCTCGGTACCCGGGGATCCTCTAGA	
<i>R_ΔcseE_for</i>	ACACAGGAGGTCATCGGAAGATTTTTTAATCCTTATTTAC	
<i>R_ΔcseE_rev</i>	GGTACCGAGCTCATCGGTCCAAGAGACGCCACCTGAGTAA	
<i>L_ΔcseE_for</i>	ACCATGATTACGAATTCCTCGCATGGAGGCGCTGCCTGAAG	
<i>L_ΔcseE_rev</i>	TGTAAATAAGGATTAATAAATCTCCGATGACCTCCTGTG	
Klonierung der Promotoren in pET2		
		Durchführung Patek
PSIGHF	GAACTTTTCTGCAGCTTTAGTGAT	<i>PstI</i>
PSIGHR	CAACGCAGATCTATCGACTGTGCC	<i>BglII</i>
PSIGH4R	CCAGATCTCCACCACAGCAGTTAGTACA	<i>BglII</i>
PRSHAF	AACTGCAGATGCAATGAATCAACTC	<i>PstI</i>
PRSHAR	GAGGATCCCAGGGCAGCCACAATC	<i>BamHI</i>
PDNAJ2F	ACCTGCAGTTACGGCTCCCAG	<i>PstI</i>
PDNAJ2R	GCCTTTTGGATCCCTGATTCGGTT	<i>BamHI</i>
PUVRAF	ATCTGCAGCATCAATGAGTA	<i>PstI</i>
PUVRAR	TCGGATCCTTTTAGGTTATGTTACG	<i>BamHI</i>
PMSHCF	TGATGCATAGCGTGAAGTAGCCACCA	<i>NsiI</i>
PMSHCR	GGGGATCCAGCGAGAGCAGG	<i>BamHI</i>
PMCAF	CAACTTCATGCATGAAGGTGTCGTCG	<i>NsiI</i>
PMCAR	TATTGGATCCATGATAGTCGTTTC	<i>BamHI</i>
PUVRDF	GTCTGCAGAAGAAGCTGAAGGGC	<i>PstI</i>
PUVRDR	CTGGATCCGGCTAATTGCTGTGGTGA	<i>BamHI</i>

Fortsetzung Tab. III.3

Primer	DNA-Sequenz	Eigenschaften
Primer-Extension (PEX)		
CM4	GAAAATCTCGTCGAAGCGTCG	Cy5-labeled
CM5	aagctcggcggattgtc	Cy5-labeled
XMSHC	CGACAAGCTATTTGCAAGCTAT	Cy5-labeled
XMCA	CTGAGGTGACTAAGATGTA	Cy5-labeled
XUVRD	GAGACCTCCTGTTTCCTGAT	Cy5-labeled
Deletionsnachweisprimer		
Nachweisprimer $\Delta rshA$		
<i>sigH</i> delN1	GGCATCGACGGTGATACTGA	
<i>sigH</i> delN2	CCGCTGCGTAACCTCATCTT	
Nachweisprimer $\Delta cseE$		
TEST_Δ <i>cseE</i> _for	CTTACCGCTGTGACAGGTGA	
TEST_Δ <i>cseE</i> _rev	CTGCAGCATTGATGCTGGA	
Northernprimer		
<i>sigH</i> Sonde		
<i>sigH</i> _Northern1	GCGGTGCTCTGCGCATGACTAGAA	
<i>sigH</i> _Northern2	GAAATTAATACGACTCACTATAGG- GAGCCTCAACCTCGGCGGATTCCA	T7-Promotersequenz
<i>rshA</i> Sonde		
<i>rshA</i> _Northern1	GCGAAGGAACAAGGCATTGGTCTTGA	
<i>rshA</i> _Northern2	GAAATTAATACGACTCACTATAGG- GTGCCGATTGCGCAGCAGCACTT	T7-Promotersequenz
RNA-Seq		
5'-RNA-Adapter	CCCUACACGACGCUCUUCGGAUCGAG	Voll-RNA
Loop Primer	AGATCGGAAGAGAGACGTGTGCTCTTCCGATCTNNNNNNN	7 wobble Basen
PCR Primer 1	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACA CGACGCTCTCCGATCGAG	
PCR Primer 2.1	CAAGCAGAAGACGGCATACGAGATcgatGTGACTGGAGTTC AGACGTGTGCTCTTCCGATCT	Barkode cgatgat
RNA-Seq Seq Primer 1	ACTCTTCCCTACACGACGCTCTTCCGATCGAG	
Multiplexing Read 2 Sequencing Primer	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	

Fortsetzung Tab. III.3

Primer	DNA-Sequenz	Eigenschaften
q-RT-PCR		
<i>arnA</i> _LC1	GCTGCGGACATAGTATGTTT	
<i>arnA</i> _LC2	ACATCGTTACCCCGAACAG	
<i>cg2838</i> _LC1	CAGCACATCGAGGTCATCAA	
<i>cg2838</i> _LC2	AATGAACGGCCAAGTCCAAG	
<i>cg1560_uvrA</i> _LC1	AACCAGCGCTTGATCACTAC	
<i>cg1560_uvrA</i> _LC2	GACAAGTAAGCACCTGTGAG	
<i>cg2194_mtr</i> _LC1	TCACCGCAGTACACAACAAG	
<i>cg2194_mtr</i> _LC2	ACTGGTGGACCGAATTCAT	
<i>cg1709_mshC</i> _LC1	TGGCTACGAGTCCAATTACG	
<i>cg1709_mshC</i> _LC2	GTTGGTAGCGATTGCTGAAC	
<i>cg1555_uvrD</i> _LC1	AATCATCGTGCAGCTGATCG	
<i>cg1555_uvrD</i> _LC2	ATCGTGGTCTCTGTTGATGG	
<i>cg3405</i> _LC1	CGCCATCAGCCTTCTTGATT	
<i>cg3405</i> _LC2	TTGAGCAGCGAGTTCTTGGA	
<i>cg1127_mca</i> _LC1	CGGATCACCTCAAGGTTTAT	
<i>cg1127_mca</i> _LC2	AGCCATCACATCAGCCTCAT	
<i>cg0876_sigH</i> _LC1	TCAAGGCGTACCAGGCGTTC	
<i>cg0876_sigH</i> _LC2	TCAACCTCGGCGGATTCCAG	
<i>cg1689_pup</i> _LC1	CGCCACCCTTTTGTACATAG	
<i>cg1689_pup</i> _LC2	AGTGAACGCAAAGCAAACCC	
<i>cseE</i> in $\Delta cseE$ _LC1	CACGTACCGCTATCGACAAT	
<i>cseE</i> in $\Delta cseE$ _LC2	TTGTCAGCTGCCTGCTTGTT	
<i>cg0986_amtR</i> _LC1	CCAAGCCTCGCTGTATTATC	
<i>cg0986_amtR</i> _LC2	GGCGAACTCTTCAGAACCAA	

3. Enzyme, Chemikalien und andere Materialien

3.1. Enzyme, Marker

Tab. III.4 Verwendete Enzyme und Marker

Produkt	Bezugsquelle
Agencourt AMPure XP - PCR Purification System	Beckman Coulter
Anti-Digoxigenin-AP, Fab fragments	Roche
BioScript Reverse Transcriptase	Bioline
CDP- <i>Star</i>	Roche
DIG RNA Labeling Kit (SP6/T7)	Roche
DNA-Marker 100 bp-Leiter	MBBL
DNA-Marker 250 bp-Leiter	MBBL
DNase I (RNase-frei)	Roche
FastDigest <i>EcoRI</i>	Thermo Scientific
High Range Marker	MBI Fermentas- Thermo Scientific
KOD-Polymerase	Merck Millipore
Lysozym	Serva
PageRuler™ Prestained Protein Ladder	MBI Fermentas - Thermo Scientific
PageRuler™ Unstained Protein Ladder	MBI-Fermentas - Thermo Scientific
Phusion-Polymerase	MBI Fermentas - Thermo Scientific
Phusion® High-Fidelity DNA Polymerase	New England Biolabs
Pronase E	Serva
Restriktionsendonukleasen	MBI Fermentas - Thermo Scientific
RNA Polymerase Core Enzym	Epicenter
RNA 5´-Polyphosphatase	Epicenter
RNase	Roche
Shrimp Alkalische Phosphatase	MBI Fermentas - Thermo Scientific
Superscript II RNase H Reverse Transkriptase	Invitrogen
<i>Taq</i> -DNA-Polymerase	MBI Fermentas - Thermo Scientific
T4-DNA-Ligase	MBI Fermentas - Thermo Scientific
T4 RNA Ligase	Epicenter
T4 RNA Ligase	New England Biolabs
Terminator™ 5´-Phosphate-Dependent Exonuclease	Epicenter
Trypsin	Promega

3.2. Chemikalien

Alle verwendeten Chemikalien entsprechen dem Reinheitsgrad p.a..

Tab. III.5 Verwendete Chemikalien

Produkt	Bezugsquelle
Acetonitril	Roth
Acrylamid / Bisacrylamid-Lösung	Roth
Agar	Invitrogen
Agarose	PeqLab
Ammoniumhydrogencarbonat	Merck
Antibiotic Medium No.3 (PA-Medium)	Oxoid
Antibiotika	Sigma
APS	Sigma
Arginin	Sigma
Biotin	Sigma
Blocking-Pulver	Roche
Blocking-Reagenz	Roche
Brain-Heart-Infusion (BHI)	Roth
Bromphenolblau (BPB)	Serva
5-Brom-4-Chlor-Indolylphosphat-Toluidinsalz (X-Gal)	GIBCO-BRL
CASO-Boullion	Merck
Chitinkügelchen	NEB
Coomassie®Brilliant Blue R250/G250	Serva
Cy3 und Cy5	Amersham Biosciences
DEPC-Wasser	eigene Herstellung / Roth
Dikaliumhydrogenphosphat	Merck
Dimethylsulfoxid (DMSO)	Fluka
Dinatriumhydrogenphosphat	Merck
Dithiothreitol (DTT)	Sigma
dNTPs	MBI Fermentas
Eisessig	Roth
Ethanol	Roth
Ethidiumbromid	Serva
Ethylendiamintetraessigsäure (EDTA)	Amersham
FluoroLink Cy3 monofunctional dye 3-pack	Amersham
FluoroLink Cy5 monofunctional dye 5-pack	Amersham
Glucose	Merck
Glycerin	Roth
Glykogen	Fermentas
Glykogen RNase-frei	Invitrogen
Hefeextrakt	Oxoid
Isopropyl- β -DThiogalactopyranosid (IPTG)	Fermentas
Kaliumchlorid	Merck
Kaliumdihydrogenphosphat	Merck
Kaliumglutamat	Fermentas

Fortsetzung Tab. III.5

Produkt	Bezugsquelle
LB Broth Base	Invitrogen
Magnesiumchlorid	Merck
Magnesiumsulfat	Merck
Methanol	Roth
MOPS	Serva
Natriumacetat	Merck
Natriumchlorid	Merck
Natriumdodecylphosphat (SDS)	Serva
Natronlauge	Merck
Phenol-Chloroform-Isoamylalkohol	Sigma
PCR-Puffer für <i>Taq</i> -, Phusion-Polymerase	MBI Fermentas
PCR-Puffer für KOD-Polymerase	Merk Millipore
Restriktionspuffer	MBI Fermentas
Rnase Zap	Sigma
SDS	Sigma
TEMED	Sigma
Sorbitol	Sigma
Sucrose	Merck
T4-DNA-Ligasepuffer	MBI Fermentas
Thiamin	Sigma
Tris-HCl	Serva
Triton X-100	Merck
Trypton	Oxoid
Zitronensäure	Merck

3.3. Kits

Tab. III.6 Verwendete Kits

Produkt	Bezugsquelle
Agilent High Sensitivity DNA Kit	Agilent Technologies
Agilent DNA 12000 Kit	Agilent Technologies
Agilent RNA 6000 Nano Kit	Agilent Technologies
Agilent RNA 6000 Pico Kit	Agilent Technologies
5`/3` RACE-Kit, 2nd Generation	Roche
CYScribe GFX™ Purification Kit	GE Healthcare
CloneJET™ PCR Cloning Kit	MBI Fermentas
GeneJET™ Plasmid Miniprep Kit	MBI Fermentas
High Pure PCR Product Purification Kit	Roche
IMPACT™ Kit	New England BioLabs
innuPrep RNA Mini Kit	Analytik Jena
NucleoSpin® Extract II	Macherey-Nagel
Qiagen QIAquick Gel Extraction Kit	Qiagen
Quant-iT Pico Green	Invitrogen

Fortsetzung Tab. III.6

Produkt	Bezugsquelle
Ribo-Zero™ rRNA Removal Kit	Epicenter
RNase-Free Dnase Set	QIAGEN
Roti® Nanoquant Kit	Carl Roth
Rneasy Mini Kit	QIAGEN
SensiMix ^{MT} SYBR & Fluorescein One-Step Kit	Bioline
ThermoScript™ RT-PCR System	Invitrogen
TruSeq DNA Sample Preparation Kit LT Set A	Illumina

3.4. Materialien

Tab. III.7 Verwendete Materialien

Produkt	Bezugsquelle
2 mL-Schraubgefäße	Starlab
Amicon Ultra-4 Zentrifugationseinheiten	Millipore
50 mL Schraubdeckelgefäße	Greiner
Cellulose Acetat Filter (Porengröße: 0,45 µm, Ø = 50 mm)	Sartorius
Elektroporationsküvetten	Peqlab
Eli-CAP Polycarbonatkapillaren für quantitative PCR	BioBudget
Reaktionsgefäße 1,5 mL (Eppis)	Greiner
Filterspitzen	Biozym
Glasperlen (Ø = 0,1 mm)	Retsch
Glaswaren	Schott
Hybond-N Membran	GE Healthcare
13 mL Röhrchen	Greiner
Küvetten	Brand
Mikrotiterplatten	Greiner
Microcon YM-30	Fisher Scientific
Nanodrop 2000	Peqlab
Parafilm	American
PCR-Reaktionsgefäße	Biozym
PE-Röhrchen	Greiner
Petrischalen	Greiner
Pipettenspitzen	Greiner, Sarsted, Starlab
Proteinausschlussäulen (10 kDa)	Millipore
Proteinofilter® Columns (35 mL)	Macherey-Nagel
Ribotubes	Q-Biogene
Sterilfilter	Schleicher&Schuell
UV-Küvetten	Eppendorf
Vinylhandschuhe	Roth

3.5. Geräte und Apparaturen

Tab. III.8 Verwendete Geräte und Apparaturen

Produkt	Hersteller
Agilent Bioanalyzer 2100	Agilent Technologies
Automated Slide Processor HS 4800	Tecan
BiflexIII-MALDI-TOF-MS	Bruker
Brutschränke	Memmert
Derivatisierungsblock	Pierce
Elektrophoresekammern für Agarosegele	Eigenbau
Gene Pulser, Pulse Control	BioRad
Feinwaagen	Sartorius
French® Press	Thermo Fisher Scientific
Gefrierschränke (-80 °C)	Sanyo
Genome Analyzer Iix	Illumina
HPLC-Anlage OPA Precolumn Analyzer A0426-1	Knauer
Hybridisierungsöfen (Roller)	Bachofer
Kühlfalle RVT400-230	Thermo Savant
Kühlschrank	Privileg/Siemens
Kühltischzentrifuge 5417R	Eppendorf
Kühlzentrifuge groß (4K 15)	Sigma
LightCycler	Roche
Luftschüttler	Gerhardt
Luftschüttler Innova 4430	New Brunswick
Luminescent Image Analyzer LAS-3000	Fujifilm Europe
Magnetrührer mit Heizplatte	Heidolph, IKA
Mastercycler Pro S Thermocycler	Eppendorf
Multifuge 3 L-R	Heraeus
MTP-Reader TECAN Infinite M200	Tecan
Nanodrop 2000 Spectrometer	Peqlab
ND-1000 Photospectrometer	NanoDrop
PCR-Thermal-Cycler PTC-100	MJ Research
pH-Meter	Knick
Spectralphotometer	Eppendorf
Pipetman	Gilson, Eppendorf
Reinraumwerkbank	Heraus HeraSafe
Reinstwasseranlage	Millipore
Ribolyser	Hybaid
ScanArray 4000	Gsi Lumonics
Schnellschüttler	New Brunswick Scientific
Spannungsgeber	BioRad
Speed-Vac SPD 111V	Thermo Savant
Thermomixer	Eppendorf
Mastercycler® pro S	Eppendorf
Typhoon 8600-Scanner	Amersham Biosciences
UV-Transilluminator	UVP

Fortsetzung Tab. III.8

Produkt	Hersteller
Videoprinter P67E	Mitsubishi
Vortex-Genie 1 und 2	Scientific Industries
Wasserbäder	GFL
Zentrifugen	Eppendorf

3.6. Software

Tab. III.9 Verwendete Software

Programm	Verwendungszweck	Version	Referenz
Bioprospector	Motiv-Suche	2013	Liu <i>et al.</i> , 2001
BLAST	Sequenzvergleich	2.2	Altschul <i>et al.</i> , 1997
CloneManager	Restriktionskarten und Primerdesign	5.0	Scientific & Education Software
ClustalX2	Sequenz-Alignment	2.0	Larkin <i>et al.</i> , 2007
Corel Draw	Erstellung und Bearbeitung von Graphiken	11.0 & 12.0	Corel Corp.
EMMA	Normalisierung von <i>Microarray</i> -Daten	2.2	Dondrup <i>et al.</i> , 2009
Excel™	Tabellenkalkulation	2010	Microsoft Corp.
FlexAnalysis	Kalibrierung der MALDI-Spektren	2.4	Bruker Daltonik
FlexControl	Aufnahme von MALDI-Spektren	2.4	Bruker Daltonik
GenDB	Datenbank von <i>C. glutamicum</i>	1.0	Meyer <i>et al.</i> , 2003
HMMER Paket	Sequenzanalyse	2.0	Eddy, 1998
ImaGene & GeneSight	Spoterkennung und Quantifizierung	5.0	Bio Discovery
Improbizer	Motiv-Suche	1.0	Ao <i>et al.</i> , 2004; Kent, USCS
Mascot	Datenbanksuche für MALDI-Auswertungen		Perkins <i>et al.</i> , 1999
Microsoft Acces	Tabellenkalkulation	2010	Microsoft Corp.
Mendeley Desktop	Referenz-Manager	1.8.4	Mendeley
Nanodrop 2000	Nanodropmessung	1.4.2	Peqlab
Notepad++	Textbearbeitung	6.3.3	Notepad++Team
LightCycler Software	Aufnahme und Analyse RT-qPCR (Bestimmung <i>Crossing Points</i> , Cp)	3.5	Roche Diagnostics
Scan & Quant Array	Scannen der <i>Microarrays</i>		Gsi Lumonics
T-coffee	Sequenz-Alignment	9.03	Notredame <i>et al.</i> , 2000
Word™	Textverarbeitung	2010	Microsoft Corp.

4. Medien und Zusätze

4.1. Nährmedien

Soweit nicht anders vermerkt, beziehen sich die Mengenangaben auf 1 Liter deionisiertes Wasser (Endvolumen). Die Nährmedien wurden vor der Nutzung autoklaviert.

CASO-Bouillon

- CASO-Bouillon 30,0 g

PA-Medium (Penassay Broth)

- Antibiotic Medium No.3 17,5 g

SOC (Regenerationsmedium) (Hanahan, 1983)

- Trypton 20 g
- Hefeextrakt 5,0 g
- D(+)-Glucose-Monohydrat 4,0 g
- NaCl 580 mg
- KCl 180 mg
- MgCl₂ x 6 H₂O 2,0 g
- MgSO₄ x 7 H₂O 2,4 g

BHIS-Medium (Liebl *et al.*, 1989)

- Brain-Heart-Infusion 37,0 g
- Sorbitol 91,0 g

LB-Vollmedium (Luria-Bertani, (Sambrook *et al.*, 1989)

- LB Broth Base 20,0 g

CGXII (Keilhauer *et al.*, 1993)

- (NH₄)₂SO₄ 20,0 g
- K₂HPO₄ 1,0 g
- KH₂PO₄ 1,0 g
- MgSO₄ x 7 H₂O 0,25 g
- MOPS 42,0 g

→ in 800 mL H₂O lösen

→ Mit KOH auf pH 7 einstellen und mit H₂O auf 876 mL auffüllen und autoklavieren

nach dem Autoklavieren hinzufügen:

- D-Glucose (Stammlösung: 500 g/L), autoklaviert 100 mL
- Thiamin (Stammlösung 500 µg/mL), sterilfiltriert 1 mL
- Protokatechusäure (Stammlösung: 30 mg/mL in EtOH) 1 mL
- CGXII-Spurenelemente, sterilfiltriert 1 mL
- Biotin (Stammlösung: 0,2 mg/mL) sterilfiltriert 1 mL
- Harnstoff (0,25 g/mL) sterilfiltriert 20 mL

4.2. Zusätze zu den Nährmedien

Agar

- 1,6 % (w/v) für Festmedium

Antibiotika

Antibiotika werden als sterile Stammlösungen angesetzt (Sterilfiltration) und nach dem Autoklavieren und Abkühlen der Medien diesen zugesetzt.

Tab. III.10 Antibiotika

Antibiotikum	Abkürzung	Lösungsmittel	Stammlösung [mg/mL]	<i>C. glutamicum</i> [µg/mL]	<i>E. coli</i> [µg/mL]	Lagerung
Ampicillin	Ap	H ₂ O bidest.	200	-	200	-80 °C
Kanamycin	Km	H ₂ O bidest.	50	25	50	-4 °C
Nalidixin	Nx	0,5 N NaOH	50	50	-	-4 °C

Vitamine

Vitaminstammlösungen werden in H₂O bidest. angesetzt und steril filtriert.

- Biotin 50 µg/mL
- Thiamin 500 µg/mL

Sucrose

- Sucrose 9,1 % (w/v)

CGXII-Spurenelemente (Keilhauer *et al.*, 1993)

- FeSO₄ x 7 H₂O 10,0 g/L
- MnSO₄ x 7 H₂O 10,0 g/L
- ZnSO₄ x 7 H₂O 1,0 g/L
- CuSO₄ 0,2 g/L
- NiCl₂ x 6 H₂O 0,02 g/L

→ in 100 mL deionisiertem H₂O lösen und autoklavieren.

5. Verwendete Puffer und Lösungen

Alle Puffer und Lösung werden in deionisiertem Wasser angesetzt und wenn nötig autklaviert.

5.1. Elektrophorese

Agarose

- 0,8 bis 2,0 % (w/v) in TA-Elektrophoresepuffer je nach gewünschter Auftrennung lösen
- Aufkochen (Mikrowelle oder Heizblock) und bei 60 °C lagern

Bromphenolblau-Xylencyanol-Ladepuffer (BPB-XC-Ladepuffer)

- EDTA 1 mM
- Glycerin 50 % (w/v)
- Bromphenolblau 0,1 % (w/v)
- Xylencyanol FF 0,1 % (w/v)

→ mit HCl bis zum Umschlagpunkt (dunkelgrün) titrieren

Ethidiumbromid Färbelösung

- 10 mg/mL EtBr in H₂O bidest. (wöchentlich frisch ansetzen)

TA-Elektrophoresepuffer (50x)

- Tris 2 M
- Natriumacetat 500 mM
- EDTA 50 mM

→ mit Eisessig auf pH 7,8 einstellen

Agaroseformaldehydgel (Northern-Blot)

- Agarose 0,8 g
- RNase-freies Wasser 53 mL
- 10x MOPS 6,5 mL
- Formaldehyd 11 mL

→ mit NaOH auf pH 7,0 einstellen

10-fach MOPS-Laufpuffer (Northern-Blot) (2x autoklavieren)

- MOPS (200 mM) 41,854 g/L
- Na-Acetat (50 mM) 4,1015 g/L
- EDTA (10 mM) 3,7224 g/L

→ 2x autoklavieren

5.2. PCR

dNTP-Mix

- dATP 2,5 mM
- dCTP 2,5 mM
- dGTP 2,5 mM
- dTTP 2,5 mM

5.3. DNA-Transfertechniken

15 % Glycerin

- Glycerin 87 % (w/v) 172 mL
- mit deionisiertem H₂O auf 1 L auffüllen

10 % Glycerin

- Glycerin 87 % (w/v) 115 mL
- mit deionisiertem H₂O auf 1 L auffüllen

TG (Tris Glycerin)-Puffer (pH 7,5)

- Tris 1 mM
- Glycerin 10 % (w/v)

TB-Puffer

- PIPES 10 mM
- CaCl₂ 15 mM
- KCl 250 mM

→ in ca. 800 mL H₂O lösen und mit KOH oder HCL auf pH 6,7 einstellen

- MnCl₂ 55 mM
- CoCl₂ 5 mM

→ auf 1 L Endvolumen auffüllen und sterilfiltrieren (Filter: 0,2 µm); Lagerung bei 4 °C

DNA-Isolierung und -Reinigung

HB1

- Glukose 50 mM
- Tris 25 mM
- EDTA 10 mM

→ in H₂O lösen; frisch 10 µL RNase-Lösung auf 1 mL HB1 geben; auf pH 7 einstellen

RNAse-Lösung

- RNase 10 mg
- in 1 mL H₂O lösen und 30 min kochen; Lagerung bei 4°C

SDS-Lösung

- SDS 20 % (w/v)

5.4. Proteinaufreinigung IMPACT-CN-System**Säulenpuffer**

- Na₂HPO₄ 20 mM
- NaCl 500 mM
- EDTA (0,1 M Stammlösung) 1 mM

→ alle Substanzen in H₂O lösen, pH 6,0-9,0 mit Phosphorsäure einstellen (Vorsicht! Nur tropfenweise!)

Aufschlusspuffer

- Triton X-100 0,1 %
- PMSF 20 µM
- TCEP 0,1 -1 mM

→ PMSF (Protease-Inhibitor): 100 mM Stammlösung ansetzen, in DMSO lösen (instabil! einfrieren)

→ TCEP (Stabilisator): 1 M Stammlösung ansetzen, in H₂O lösen (instabil!)

→ alle Substanzen in Säulenpuffer lösen

Spaltungspuffer

- Na₂HPO₄ 20 mM
- NaCl 500 mM
- EDTA 1 mM
- DTT 50 mM

→ alle Substanzen in H₂O lösen, pH 7-9 mit Phosphorsäure einstellen (Vorsicht! Nur tropfenweise!)

Proteinwaschpuffer

- Na₂HPO₄ 20 mM
- NaCl 10 mM
- EDTA 1 mM

5.5. 1D SDS-Gelelektrophorese und Proteinfärbung

SDS-Lösung

- SDS 0,5 % (w/v)

Trenngeltrispuﬀer

- Tris (pH 8,8) 1,5 M

Sammelgeltrispuﬀer

- Tris (pH 6,8) 0,5 M

SDS-PAGE-Laufpuﬀer

- Tris 25 mM
- Glycin 192 mM
- SDS 0,1 % (w/v)

Probenpuﬀer (PBJR)

- Tris-HCl (pH 6,8) 3,5 mL
- SDS (0,5 M) 3,5 mL
- Glycerin (87 %) 3,5 mL
- SDS 1,0 g
- DTT 0,9 g
- BPB 1,2 mg

→ mit deionisiertem H₂O auf 10 mL auffüllen

Coomassie-Brillant-Blau-Färbung

- Coomassie Brilliant Blue R250 2,0 g
- Coomassie Brilliant Blue G250 0,5 g
- Ethanol 425 mL
- Methanol 50 mL
- H₂O bidest 425 mL

→ über Nacht unter Rühren lösen und vor Gebrauch filtrieren

Schnellentfärber

- Ethanol 450 mL
- Eisessig 100 mL
- H₂O bidest. 450 mL

Langsamentfärber

- Essigsäure 7 % (w/v)
- in deionisiertem H₂O lösen

5.6. Tryptischer Verdau**TDL1**

- CH₃CN 60 % (v/v)
- TFA 0,1 % (v/v)

TDL2

- CH₃CN 50 % (v/v)

TDL3

- CH₃CN 50 % (v/v)
- NH₄HCO₃ 50 mM

TDL4

- CH₃CN 50 % (v/v)
- NH₄HCO₃ 10 mM

TDL5

- NH₄HCO₃ 10 mM

Trypsinlösung

- Trypsin 1 µL
- NH₄HCO₃ (10 mM) 14 µL

5.7. cDNA-Synthese und -Markierung für *Microarrays*

Aminomodifizierte *Random*-Hexanukleotide

- AmC6NNNNNN 5 µg/µL

50x dNTP-Stammlösung

- Aminoallyl-dUTPs (aa-dUTPs) 200 mM
- dTTP 250 mM
- dATP 250 mM
- dGTP 250 mM
- dCTP 250 mM

Fluoreszenzfarbstofflösung

- Cy3TM/Cy5TM Fluoreszenzfarbstoff in 16,5 µL DMSO lösen
- 2 µL Aliquots in braune Eppis füllen
- 1 h in der SpeedVac trocknen

TE-Puffer (pH 7,5)

- Tris-HCl 10 mM
- EDTA 1 mM

5.8. Lösungen und Puffer für *Microarray*-Hybridisierung

20x SSC-Puffer

- NaCl 3 mM
- Natriumcitrat 0,3 mM
- in deionisiertem H₂O gelöst

Waschlösung 1

- SSC 2 fach
- SDS 0,2 % (w/v)

Waschlösung 2

- SSC 0,2 fach
- SDS 0,2 % (w/v)

Waschlösung 3

- SSC 0,2 fach

Waschlösung 4

- SSC 0,1 fach

5.9. Northern Blot Analyse (Gel, Blot und Hybridisierung)

RNase-freies Wasser

Millipore Wasser 2x autoklavieren

10-fach MOPS (2x autoklavieren)

- MOPS (200 mM) 41,854 g/L
 - Na-Acetat (50 mM) 4,1015 g/L
 - EDTA (10 mM) 3,7224 g/L
- 2x autoklavieren

Agaroseformaldehydgel

- Agarose 0,8 g
 - RNase-freies Wasser 53 mL
 - 10x MOPS 6,5 mL
 - Formaldehyd 11 mL
- mit NaOH auf pH 7,0 einstellen

Blotten des Gels und Hybridisierung

20x SSC

- Natriumcitratdihydrat 88,2 g/L
 - NaCl 175,3 g/L
- 2x autoklavieren

Denaturierungslösung

- NaOH (50 mM) 1,9998 g/L
 - NaCl (10 mM) 0,5844 g/L
- 2x autoklavieren

Neutralisierungslösung

- Tris-HCl (100 mM; pH 7,4) 18,760 g/L
- 2x autoklavieren

Blockingreagenz

- Blocking-Pulver 20 g
 - 1x Puffer I 200 mL
- bei 4 °C lagern

Puffer I

- Maleinsäure 116 g/L
 - NaCl 87,7 g/L
 - NaOH 72 g/L
- mit NaOH auf pH 7,5 einstellen

Prähybridisierungslösung (40mL)

- RNase-freies Wasser 1,5 mL
- SDS (10 %) 80 µL
- N-Laurolylsarcosin Natriumsalz (10 %) 400 µL
- Blocking-Reagenz 8 mL
- 20x SSC 10 mL
- Formamid 20 mL

Waschen und Detektion der Hybridisierungssignale**Waschlösung I**

- SSC (20x) 10 mL
- SDS (10 %) 1 mL
- RNase-freies Wasser 89 mL

Waschlösung II

- SSC (20x) 1,5 mL
- SDS (10 %) 3 mL
- RNase-freies Wasser 295,5 mL

Puffer II

- Puffer I (10x) 5 mL
 - Blocking-Reagenz 5 mL
 - RNase-freies Wasser 40 mL
- immer frisch ansetzen

Puffer III

- Diethanolamin 9,6 mL/L

Waschpuffer

- Puffer I (10x) 100 mL
- Tween20 3 mL
- RNase-freies Wasser 897 mL

5.10. RNA-Seq

Ethanolpräzipitation

- Natriumacetat (pH 5.2) 3 M
- Glykogen RNase-frei 20 µg/µL

Fragmentierungspuffer

- Kaliumacetat 100 mM
- Magnesiumacetat 30 mM
- Tris-HCl (pH 8,1) 200 mM

Fragmentierungsstopppuffer

- Tris-HCl (pH 8) 10 mM
- EDTA 1 mM

5.11. Run-Off-Transkription

5.12. Run-Off-Puffer (1,5 x)

- HEPES (1 M; pH 8,0) 2,4 µL
 - MgCl₂ (25 mM) 24 µL
 - Kaliumglutamat (1M) 9 µL
 - Dithiothreitol (DTT) (500 mM) 0,24 µL
 - BSA (10 mg/mL) 0,6 µL
 - Triton X (0,5 %) 1,2 µL
 - Ribolock (40 U/µl) 1,5 µL
 - RNase-freies Wasser 1,1 µL
- Immer frisch ansetzen!

NTP-Mix für Run-Off-Transkription

- ATP 10 mM
- CTP 10 mM
- GTP 10 mM
- TTP 10 mM

5.13. Gibson-Assembly

Isothermaler Reaktionspuffer (5x; 6 mL)

- Tris-HCl (1 M; pH 7,5) 3 mL
- MgCl² (2 M) 150 μ L
- dGTP (100 mM) 60 μ L
- dATP (100 mM) 60 μ L
- dTTP (100 mM) 60 μ L
- dCTP (100 mM) 60 μ L
- DTT (1 M) 300 μ L
- PEG-8000 1,5 g
- NAD (100 mM) 300 μ L

→ Dieser Puffer kann aliquotiert und bei $-20\text{ }^{\circ}\text{C}$ gelagert werden.

Gibson-Assembly-Mastermix

- 5x Isothermaler Reaktionspuffer 320 μ L
- T5 Exonuclease (10 U/ μ L) 0,64 μ L
- Phusion-Polymerase (2 U/ μ L) 20 μ L
- Taq-DNA-Ligase (40 U/ μ L) 160 μ L

→ Mit H₂O auf 1,2 mL auffüllen.

→ 15 μ L Aliquots erstellen und bis zur Verwendung bei $-20\text{ }^{\circ}\text{C}$ lagern.

6. Kultivierung von Bakterien

Die in dieser Arbeit verwendeten und konstruierten Bakterienstämme und Plasmide sind in den Tabellen III.1 und III.2 aufgelistet. Alle Festmedien enthalten 1,6 % (w/v) Agar. Zur Selektion von Plasmid-tragenden Stämmen werden den sterilen Medien Antibiotika aus sterilfiltrierten Stammlösungen (s. Tab. III.10) nach dem Abkühlen zugegeben. Flüssigkulturen werden in Erlenmeyerkolben im Luftschüttler Innova 4430 bei 300 U/min angezogen.

6.1. Anzucht und Lagerung von *E. coli* Zellen

Die *E. coli* Stämme JM109, DH5 α MCR und ER 2566 werden bei 37 °C in Luria-Bertani-Medium (LB) (Sambrook et al., 1989) oder auf LB-Agarplatten angezogen und für Standardklonierungen und Proteinexpressionen verwendet. Auf Festmedien werden die Bakterien über Nacht angezogen. Plasmid-tragende *E. coli* Stämme werden mit Kanamycin oder Ampicillin selektiv angezogen (Tab. III.10). *E. coli* Stämme können, wenn nötig, wenige Wochen bei 4 °C gelagert werden. Hierfür werden die Platten mit Parafilm verschlossen. Für die Induktion der IMPACT-Plasmide wurde IPTG (0,5 mM) verwendet.

6.2. Anzucht und Lagerung von *C. glutamicum* Zellen

C. glutamicum Stämme werden bei 30 °C in CGXII Minimalmedium (Keilhauer et al., 1993) oder in CASO-Bullion angezogen. Die Anzucht auf Festmedium erfolgt auf CASO-Nx-Platten (Tab. III.10). Plasmid-tragende *C. glutamicum* Stämme werden selektiv mit Kanamycin angezogen (Tab. III.10). *C. glutamicum* Stämme können, wenn nötig, wenige Wochen bei 4 °C gelagert werden. Hierfür werden die Platten mit Parafilm verschlossen.

6.3. Dauerhafte Stammkonservierung - Herstellung von Glycerinkulturen

Zur langfristigen Stammkonservierung werden frisch gewachsene Bakterien einer Festmediumsplatte in 400 μ L LB-Medium resuspendiert und mit 600 μ L 87 %igem Glycerin vermischt und sofort auf Eis gelagert. Der Einsatz von Antibiotika ist hier nicht nötig. Die erstellten Glycerinkulturen werden nachfolgend bei –80 °C gelagert. Die Lagerung ist für mehrere Jahre bedenkenlos möglich. Für eine erneute Verwendung der Bakterien werden 30 - 100 μ L der Glycerinkultur auf Festmedienplatten ausgestrichen oder für das Animpfen einer Flüssigkultur verwendet.

6.4. Bestimmung des Bakterientiters

Das Wachstum von Flüssigkulturen im Kulturkolben oder Fermenter wird im Spektralphotometer (Eppendorf) verfolgt. Dazu wird die optische Dichte der Probe bei einer Wellenlänge von $\lambda = 600 \text{ nm}$ ermittelt. Als Referenz dient die jeweilige sterile Nährlösung. Eine optische Dichte von 1 entspricht etwa 2×10^8 *E. coli* Zellen/mL und 1×10^8 *C. glutamicum* Zellen/mL.

6.5. Stochertest

Bei dem Test werden Einzelkolonien mit einer sterilen Spitze oder einem sterilen Zahnstocher von einer Festmediumplatte abgenommen. Hierbei muss darauf geachtet werden, dass es sich tatsächlich nur um eine Einzelkolonie handelt. Diese Einzelkolonien werden dann mit Hilfe eines Rasters strichförmig auf Festmedien und verschiedenen selektiven Bedingungen ausgestrichen. Nach 1 - 2 Tagen Inkubation bei 30 °C bzw. 37 °C können die Platten ausgewertet werden.

7. Allgemeine DNA-Arbeiten

7.1. Isolierung von Plasmid-DNA

Die Isolierung hochreiner Plasmid-DNA für eine Sequenzierung oder eine Standardklonierungen erfolgt mittels dem Prinzip der alkalischen Lyse (Birnboim and Doly, 1979). Hierfür wird der entsprechende Plasmid-tragende Stamm auf Festmedium angezogen. Die Isolierung erfolgt mit dem GeneJET™ Plasmid Miniprep Kit nach Anleitung des Herstellers mit folgenden Abänderungen. Die Zellen einer Platte werden auf zwei Ansätze aufgeteilt und resuspendiert. Weiter wird die nachfolgende alkalische Lyse für 2 min bei Raumtemperatur durchgeführt. Nach der Neutralisierung und Zentrifugation werden die zwei Ansätze auf einer Säule vereint. Die Elution der Plasmid-DNA erfolgt in zwei Elutionsschritten mit jeweils 2 min Inkubation der Elutionspuffers auf der Säule.

7.2. Aufreinigung von PCR-Produkten

Es besteht die Notwendigkeit PCR-Produkte vor der weiteren Verwendung wie z. B. für eine Restriktionsspaltung, eine Ligation, weiterer Amplifikation oder für das Gibson-Assembly aufzureinigen. Für die Aufreinigung der PCR-Produkte wurde in dieser Arbeit das Kit NucleoSpin® Extract II genutzt, wobei nach Vorschrift des Herstellers gearbeitet wurde. Im NucleoSpin® Extract II Kit wird die DNA (> 80 bp) an eine Säulenmatrix gebunden und Salze,

Nukleotide, Primer und Proteine durch mehrere Waschschrirte entfernt. Final kann in dem mitgelieferten Elutionspuffer oder mit H₂O eluiert werden.

7.3. Agarosegelelektrophorese

Die Agarosegelelektrophorese ist das Standardverfahren zur Auftrennung unterschiedlich großer DNA-Fragmente (Sambrook et al., 1989). Bei diesem Verfahren werden die DNA-Fragmente durch das Anlegen eines elektrischen Feldes abhängig von ihrer Größe und Konformation unterschiedlich schnell durch das Agarosegel zur Anode bewegt. Die Agarose wird dazu zunächst mit TA-Puffer vermischt, aufgekocht und in eine Gelkammer gegossen. Ein Kamm wird zur Formung der Taschen eingesetzt. Das Gel wird nach dem Erkalten mit TA-Puffer überschichtet. Die Proben werden nach der Entfernung des Gelkammes mit einem BPB-haltigen Ladepuffer in die Taschen pipettiert. Die Trennung der Fragmente in einem 0,5 - 2 %igen Agarosegel erfolgt in einem elektrischen Feld bei 90 V für 30 - 45 min. Anschließend wird das Gel in einem Ethidiumbromidbad (1 µg/mL EtBr) gefärbt, kurz mit H₂O entfärbt und unter UV-Licht fotografiert. Für die Größenanalyse von DNA-Fragmenten werden die Proben mit einem geeigneten Marker verglichen (Tab. III.4).

7.4. Isolierung von DNA-Fragmenten aus einem Agarosegel

Um ein DNA-Fragment definierter Größe aus einem Mix unterschiedlich großer Fragmente zu isolieren, werden die Fragmente zunächst durch Agarosegelelektrophorese ihrer Größe nach aufgetrennt. Die Bande mit der gewünschten Größe wird anschließend ausgeschnitten. Hierzu werden pro Probe 2 Ladespuren (15 - 20 µl Probe plus BPB-Ladepuffer) mit 2 Referenzspuren (5 µl Probe plus BPB-Ladepuffer) flankiert, welche wiederum jeweils außen mit einem geeigneten Längenmarker flankiert werden. Nur die Referenzspuren mit dem Marker werden im Ethidiumbromidbad gefärbt. Die Banden mit der gewünschten Größe werden auf den gefärbten Agarosegel-Abschnitten unter UV-Licht ausgeschnitten und somit das Gel markiert. Diese jetzt markierten flankierenden Agarosegelabschnitte dienen bei dem Ausschneiden des aufzureinigenden DNA-Fragments (Banden) als Schablone. Durch diese Vorgehensweise wird die zu isolierende DNA nicht dem interkalierenden Farbstoff Ethidiumbromid ausgesetzt. Die Aufreinigung der DNA aus diesem Agarosestück erfolgt mit dem Kit NucleoSpin® Extract II nach den Angaben des Herstellers.

7.5. Isolierung genomischer DNA aus *C. glutamicum*

Die Isolierung genomischer DNA (gDNA) aus *C. glutamicum* Zellen erfolgt durch eine kombinierte Behandlung der Zellen mit Detergens (SDS [20 % w/v]) und dem Enzymen Lysozym, das zur Lysierung der Zellen führt. Durch die Zugabe von Pronase E erfolgt der proteolytische Abbau der Zellproteine, die in den nachfolgenden Schritten mittels einer Phenolextraktion entfernt werden. Final wird die DNA durch Ethanolfällung extrahiert (A Tauch et al., 1995b).

Im Detail:

- ü/N auf Festmediumplatten angewachsene Bakterien in 3 mL HB1 (20 mg/mL Lysozym) vollständig resuspendieren und 2 h bei 37 °C im Roller inkubieren
- Zugabe von 500 µL einer 20 %igen SDS-Lösung; vorsichtig invertieren
- Zugabe einer Spatelspitze Pronase E gefolgt von einer 2 h Inkubation im 37 °C im Roller
- 2–3 ml Phenol-Chlorophorm-Mix (1:1) zugeben und vorsichtig invertieren
- Zentrifugation bei 4 °C für 20 min mit 8.000 RZB
- Überführung der wässrigen Phase mit einer abgeschnittenen blauen Spitze in ein neues PE-Röhrchen; Phenolisierung wiederholen bis kein Protein mehr ausfällt
- Überstand mit 98 %igem (v/v) Ethanol (–20 °C) auf 10 mL auffüllen und invertieren bis die DNA ausfällt; 20 min bei 12.000 U/min und 4 °C zentrifugieren
- Pellet mit 5 mL 70 %igem Ethanol (–20 °C) waschen und erneut zentrifugieren
- Überstand verwerfen und Pellet bei 60 °C trocknen und in 100 µL H₂O resuspendieren; Lagerung bei 4 °C

7.6. Quantifizierung genomischer DNA aus *C. glutamicum*

Für die Quantifizierung von gDNA aus *C. glutamicum* wird das Quant-iT™ PicoGreen® dsDNA Assay Kit verwendet, wobei der verwendete Fluoreszenzfarbstoff spezifisch an doppelsträngige DNA bindet. Die Fluoreszenz wird mit dem MTP-Reader TECAN Infinite M200 ausgelesen. Anhand einer Eichreihe wird anschließend die gDNA-Konzentration bestimmt.

7.7. Restriktion genomischer DNA aus *C. glutamicum*

Die Restriktion genomischer DNA für den Einsatz in *run-off*-Transkriptionen wird mittels FastDigest EcoRI nach Herstellerangaben durchgeführt. 5 µg genomischer DNA werden bei 37 °C mit 1 µl FastDigest EcoRI geschnitten. Anschließend erfolgt eine thermale Inaktivierung für 5 min bei 80 °C.

7.7.1. Analyse geschnittener genomischer DNA aus *C. glutamicum*

Die durchschnittlichen Längen der geschnittenen gDNA-Fragmente werden mittels Kapilargelelektrophorese an einem Agilent Bioanalyser 2100 mit dem Agilent DNA 12000 Kit (Agilent Technologies, Böblingen, Deutschland), welches eine Auftrennung von dsDNA-Fragmente von 100 bis 12.000 bp erlaubt, bestimmt. Hierbei wird nach Angaben des Herstellers verfahren.

8. Polymerasekettenreaktion (PCR)

Die Vervielfältigung von DNA-Abschnitten erfolgt mittels PCR (Saiki et al., 1988). Grundvoraussetzung ist die Kenntnis über die flankierenden Sequenzen des zu amplifizierenden DNA-Abschnittes. Ein PCR-Zyklus besteht aus drei Schritten. Durch Denaturierung bei 95–98 °C wird die ds-DNA (*Template*) im ersten Schritt in Einzelstränge aufgeschmolzen. Im zweiten Schritt, dem *Annealing*, binden die Primer über Wasserstoffbrückenbindungen an ihre komplementären Zielsequenzen. Die *Annealing*-Temperatur der Primer wird von dem Programm Clone Manager berechnet und liegt meist einige Grad Celsius unter der Schmelztemperatur der Primer. Im dritten Schritt, der *Elongation*, katalysiert eine thermostabile Polymerase bei 72 °C die Verlängerung des neuen DNA-Stranges. Hier dienen die 3'-OH-Enden der Primer als Ausgangspunkt für die Synthese. Die PCR wird mit dem Mastercycler pro S Thermocycler (Eppendorf) auf chromosomaler oder Plasmid-DNA mit den Polymerasen *Taq*, KOD oder Phusion durchgeführt. Im Unterschied zur *Taq*-Polymerase besitzen die Phusion- und KOD- Polymerase eine *Proof-Reading*-Funktion und erzeugen glatte Enden (*blunt ends*). Aufgrund dieser unterschiedlichen Eigenschaften wird die *Taq*-Polymerase nur für Kolonie-PCRs oder andere Test-PCRs genutzt, während die Phusion- und die KOD-Polymerase für die Synthese aller DNA-Fragmente eingesetzt werden, die für Klonierungen vorgesehen sind.

8.1. Primer-Design

Das Design von Primern ist ein essentieller Faktor für eine erfolgreiche PCR. Hierbei gibt es mehrere Regeln die befolgt werden müssen. Um Fehl-Amplifikate zu vermeiden muss auf eine ausreichende Spezifität der Primer geachtet werden. Dafür kann neben der Auswahl der Sequenz auch die Länge des Primers (20 - 40 bp) eine wichtige Rolle spielen. Im Idealfall sollte die Länge des Primers etwa 20 bp betragen und der G+C-Gehalt zwischen 50 - 60 % liegen. Die *Annealing*-Temperaturen der Primer sollten sich, wenn möglich, um nicht mehr als

2 °C unterscheiden. Bei dem Design der Primer muss desweiteren darauf geachtet werden, dass das Risiko für die Bildung von Haarnadelstrukturen oder von Primerdimeren minimiert wird. Weiter sollten Wiederholungen von Motiven oder Folgen von gleichen Nukleotiden vermieden werden. Erweiterungen wie Restriktionsschnittstellen inklusive GGTGGT-Kappe (diese ermöglicht die optimale Funktionalität von Restriktionsendonukleasen) oder Überlappungsbereiche für Gene-SOEing und Gibson-*Assembly* können an das 5´-Ende des Primers angehängt werden. Für die Berücksichtigung der genannten Anforderungen an den optimalen Primer wird das Programm Clone Manager verwendet. Bezugsquellen für Primer waren Metabion oder Invitrogen. Die in dieser Arbeit verwendeten Primer sind in Tabelle Tab. III.3 gelistet.

8.2. PCR-Reaktionsansätze und -Programme

PCR-Reaktionsansatz und -Programm für die *Taq*-Polymerase

20 µL *Taq*-Reaktionsansatz:

H ₂ O	13 µL
NH ₄	2 µL
MgCl ₂ (50 mM)	1,5 µL
Primer vor (100 µM)	0,5 µL
Primer rück (100 µM)	0,5 µL
dNTPs (je 2,5 mM)	2 µL
<i>Taq</i> -Polymerase	0,5 µL
<i>Template</i>	1 µL

3-Step-PCR-Programm:

Zyklen	Temperatur	Zeit	Schritt
	96 °C	bis Eingabe	Block temperieren
1	96 °C	3 min	Initiale Denaturierung
	96 °C	15 s	Denaturierung
36	55 °C(*)	15 s	<i>Annealing</i>
	72 °C	30 s (**)	<i>Elongation</i>
1	72 °C	5 min	Finale <i>Elongation</i>
	4 °C	∞	Lagerung

* abhängig von den verwendeten Primern

** abhängig von der Größe des zu amplifizierenden DNA-Fragments und der Reaktionsgeschwindigkeit der verwendeten Polymerase (*Taq*-Polymerase: 1-3 kbp/min)

8.2.1. PCR-Reaktionsansatz und -Programm für die KOD-Polymerase

50 µL KOD-Reaktionsansatz:

H ₂ O	32,25 µL
KOD-Puffer	5 µL
MgSO ₄ (25 mM)	3 µL
Primer vor (100 µM)	1,5 µL
Primer rück (100 µM)	1,5 µL
dNTPs (je 2,5 mM)	5 µL
KOD-Polymerase	0,5 µL
DMSO (optional; max 5 %)	0,75 µL
<i>Template</i> (min. 100 ng)	1 µL

3-Step-PCR-Programm der KOD-Polymerase:

Zyklen	Temperatur	Zeit	Schritt
	95 °C	bis Eingabe	Block temperieren
1	95 °C	2 min	Initiale Denaturierung
	95 °C	20 s	Denaturierung
36	55 °C(*)	10 s	<i>Annealing</i>
	70 °C	30 s (**)	<i>Elongation</i>
1	70 °C	5 min	Finale <i>Elongation</i>
	4 °C	∞	Lagerung

* abhängig von den verwendeten Primern

** abhängig von der Größe des zu amplifizierenden DNA-Fragments und der Reaktionsgeschwindigkeit der verwendeten Polymerase (KOD-Polymerase: 1-1,5 kbp/30s)

8.2.2. PCR-Reaktionsansatz und -Programm für die Phusion-Polymerase

50 µL Phusion-Reaktionsansatz:

H ₂ O	28 µL
HF- oder GC-Puffer	10 µL
Primer vor (100 µM)	2,5 µL
Primer rück (100 µM)	2,5 µL
dNTPs (je 2,5 mM)	4 µL
Phusion-Polymerase	0,5 µL
DMSO (optional)	1,5 µL
<i>Template</i>	1 µL

3-Step-PCR-Programm der Phusion-Polymerase:

Zyklen	Temperatur	Zeit	Schritt
	98 °C	bis Eingabe	Block temperieren
1	98 °C	5 min	Initiale Denaturierung
	98 °C	10 s	Denaturierung
36	55 °C(*)	30 s	<i>Annealing</i>
	72 °C	30 s (**)	<i>Elongation</i>
1	72 °C	5 min	Finale <i>Elongation</i>
	4 °C	∞	Lagerung

* abhängig von den verwendeten Primern

** abhängig von der Größe des zu amplifizierenden DNA-Fragments und der Reaktionsgeschwindigkeit der verwendeten Polymerase (Phusion-Polymerase: 1kbp/15s)

9. Klonierungsexperimente

9.1. DNA-Restriktionsspaltung

Durch die DNA-Restriktionsspaltung wird DNA durch Restriktionsendonukleasen an einer spezifischen palindromischen Sequenz geschnitten. Das Pentosephosphatrückgrat wird an diesem Sequenzabschnitt hydrolysiert. Hier gibt es zwei verschiedene Arten der Restriktion. Zum einen gibt es Restriktionsenzyme die einzelsträngige Überhänge erzeugen (*sticky ends*) und zum anderen Enzyme, die glatte Enden erzeugen (*blunt ends*). Bei der Spaltung von PCR-Fragmenten die mit Schnittstellen an deren Enden versehen sind, muss darauf geachtet werden, dass diese Restriktionsschnittstellen nicht „nackt“ an das PCR-Fragment angehängt sind. Dies bedeutet, dass eine Kappensequenz angehängt wird, wodurch die Restriktionsendonukleasen die Schnittstelle effektiv erkennen und schneiden können (s. III.8.1). Bei der Spaltung von gDNA oder Plasmid-DNA sollte auf die etwaige Methylierung der DNA geachtet werden, da einige Restriktionsendonukleasen methylierte Schnittstellen nicht erkennen oder nur mit einer geringen Effektivität spalten können. Die Restriktion wird nach Angaben des Herstellers durchgeführt. Nachfolgend werden die geschnittenen Fragmente mit dem Kit NucleoSpin® Extract II nach Angaben des Herstellers aufgereinigt (s. III.7.2). Die Restriktion kann zur Restriktionsanalyse von DNA oder für die Vorbereitung weiterer Klonierungsschritte, wie z. B. der Erstellung von pk18*mobsacB*-Derivaten, genutzt werden.

9.2. DNA-Ligation

Geschnittene DNA-Fragmente die kompatible Überhänge oder glatte Enden besitzen, können durch die T4-DNA-Ligase miteinander verknüpft werden. Hierbei werden die endständigen 5'-Phosphatgruppen und 3'-Hydroxylgruppen des Pentosephosphatrückgrats der DNA unter Ausbildung von Phosphodiesterbindungen miteinander verknüpft. Ein Beispiel ist die Erstellung von Deletionskonstrukten (*pk18mobsacB*). Hier findet die Ligation eines geschnittenen DNA-Fragments (*Insert*) in den entsprechend gespaltenen Vektor *pk18mobsacB* statt. Die T4-DNA-Ligase wird bis auf die Inkubationstemperatur und Inkubationszeit nach Angaben des Herstellers verwendet. Die Inkubation findet bei Raumtemperatur für 2 h statt.

10. DNA-Transfertechniken

10.1. Elektroporation

Die Elektroporation wird als Methode verwendet um freie DNA in Bakteriezellen zu transformieren. Durch einen elektrischen Puls wird die Membran kompetenter Zellen kurzfristig permeabel und ermöglicht die Aufnahme freier DNA. Um die Zellen für diese Methode kompetent zu machen müssen diese erst behandelt werden. Die Behandlung der Zellen und die Durchführung der Elektroporation ist in den nächsten beiden Abschnitten für *C. glutamicum* beschrieben (Tauch *et al.*, 2002, 1994).

10.1.1. Herstellung elektrokompetenter *C. glutamicum* Zellen

C. glutamicum wird in BHIS bis zu einer o.D.₆₀₀ von zwei angezogen und auf Eis abgekühlt. Alle folgenden Schritte müssen auf Eis in vorgekühlten Zentrifugen und mit vorgekühlten Lösungen durchgeführt werden. Nach der Zentrifugation (10 min; 4 °C; 3.500 RZB) werden die Zellen zweimal mit TG-Puffer (50 mL) und dreimal mit eiskaltem 10 %igen Glycerin (20 mL) gewaschen. Nach der letzten Zentrifugation werden die Zellen in 3 mL eiskaltem 10 %igen Glycerin aufgenommen. Die nun kompetenten Zellen werden in 150 µL Aliquots sofort in flüssigem Stickstoff schockgefroren. Die Lagerung erfolgt bei -80 °C.

10.1.2. Elektroporation nach *C. glutamicum*

Die kompetenten *C. glutamicum* Zellen werden auf Eiswasser für ca. 5 min aufgetaut. Die Elektroporationsküvetten und die DNA-Lösung (aufgereinigtes und mit deionisiertem H₂O eluiertes Plasmid) müssen vorgekühlt sein. Das Plasmid wird zu den kompetenten Zellen hinzugegeben und für 1 min in der vorgekühlten Elektroporationsküvette inkubiert. Der Puls

findet im Gene-Pulser statt. Folgende Pulsbedingungen werden angewendet: 25 μF Kapazität; 200 Ω Parallelwiderstand; 2,5 kV Spannung (bei 0,2 cm Küvetten). Nach erfolgter Elektroporation werden die Zellen sofort in 4 mL auf 46 °C vorgewärmtem BHIS-Medium aufgenommen und für 6 min bei 46 °C inkubiert. Die Regeneration der Zellen findet anschließend für 1 h bei 30 °C im Roller statt. Hiernach werden die Zellen auf entsprechendem Selektionsmedium ausplattiert.

10.2. Hitzeschocktransformation von ultrakompetenten *E. coli* Zellen

Der Hitzeschock ist eine weitere Methode um Bakterienzellen mit freier DNA zu transformieren und findet bei *E. coli* Zellen Anwendung. Die Bakterienzellen werden hierbei mit Calciumchlorid (1 M) behandelt, so dass zwischen der negativ geladenen DNA und der negativ geladenen Zellmembran weniger abstoßende Kräfte herrschen. Durch einen Hitzeschock (42 °C für 3 min) entstehen Poren in der Membran, so dass die DNA in die Zellen gelangen kann.

10.2.1. Herstellung von ultrakompetenten *E. coli* Zellen (Morrison, 1977)

- 500 μL einer $\bar{\text{u}}\text{N}$ -Kultur *E.coli* in 50 mL LB-Medium bis zu einer o.D.₆₀₀ von 0,5 - 0,6 anziehen
- 5 min zentrifugieren, 6.000 U/min, 0 °C
- Überstand verwerfen
- Pellet in 25 mL 100 mM CaCl_2 -Lösung (0 °C) resuspendieren
- 30 min bei 0 °C inkubieren
- 5 min zentrifugieren, 6.000 U/min, 0 °C
- Überstand verwerfen
- Pellet in 5 mL 100 mM CaCl_2 -Lösung (0 °C) resuspendieren
- 1.000 μl Glycerin zugeben
- Ansatz in 200 μL Aliquots teilen
- bei -80 °C lagern

10.2.2. Hitzeschocktransformation von ultrakompetenten *E. coli* Zellen

- Ligationsansatz oder Plasmid-DNA zu den kompetenten Zellen geben
- 30 min auf Eis (4 °C) inkubieren
- 3 min Hitzeschock bei 42 °C
- 1 mL LB-Medium zugeben
- 30 - 60 min bei 37 °C inkubieren
- 100 μL ausplattieren
- Restlichen Ansatz abzentrifugieren, Überstand verwerfen und das Pellet ebenfalls ausplattieren.

11. Konstruktion definierter Deletionen in *C. glutamicum*

11.1. Erstellung von Deletionskonstrukten mittels *Gene Splicing by Overlap Extension* (GeneSOEing)

Gene Splicing by Overlap Extension (GeneSOEing) ist eine Methode für die Rekombination von DNA-Fragmenten an einer präzisen Verbindungsstelle, unabhängig von der Sequenz der Rekombinationsstelle (Horton *et al.*, 1989). Für diese Methode werden keine Restriktionsendonukleasen oder Ligase benötigt. Die Fragmente von Genen, die durch eine PCR erstellt worden sind, können mit dieser Methode in einer separaten PCR neu kombiniert werden. Hierfür müssen die Primer so designt sein, dass die Enden der zu kombinierenden Produkte komplementär sind. Wenn die PCR-Produkte vermischt, denaturiert und wieder angelagert werden, stehen die Stränge mit der passenden Sequenz an ihrem 3' überlappenden Ende gegenseitig als Primer zur Verfügung. Durch die *Elongation* dieser Überlappung durch eine DNA-Polymerase in der separaten PCR, entsteht ein DNA-Molekül in dem die beiden ursprünglichen DNA-Fragmente zusammen gefügt sind. Eine mögliche Anwendung ist das Erzeugen von Deletionskonstrukten (pK18*mobsacB*) (Abb. III-1). Hierbei werden zunächst zwei PCR-Produkte mit einer Größe von ca. 500 bp erzeugt, die den zu deletierenden Bereich flankieren. Bei den dafür verwendeten Primern (Primer d1-d4) ist darauf zu achten, dass Primer d2 am 5'-Ende einen Überhang trägt, der komplementär zur Sequenz von Primer d3 ist. In einer nachfolgenden PCR-Reaktion werden die beiden Produkte im gleichen Verhältnis zur Erstellung des Fusionsprodukts eingesetzt. Durch den Überhang am zweiten Produkt und mit Hilfe der Primer d1 und d4 können die beiden PCR-Produkte fusioniert werden. Das entstandene Fusionsprodukt kann dann über Restriktion und Ligation in den Vektor pK18*mobsacB* kloniert werden.

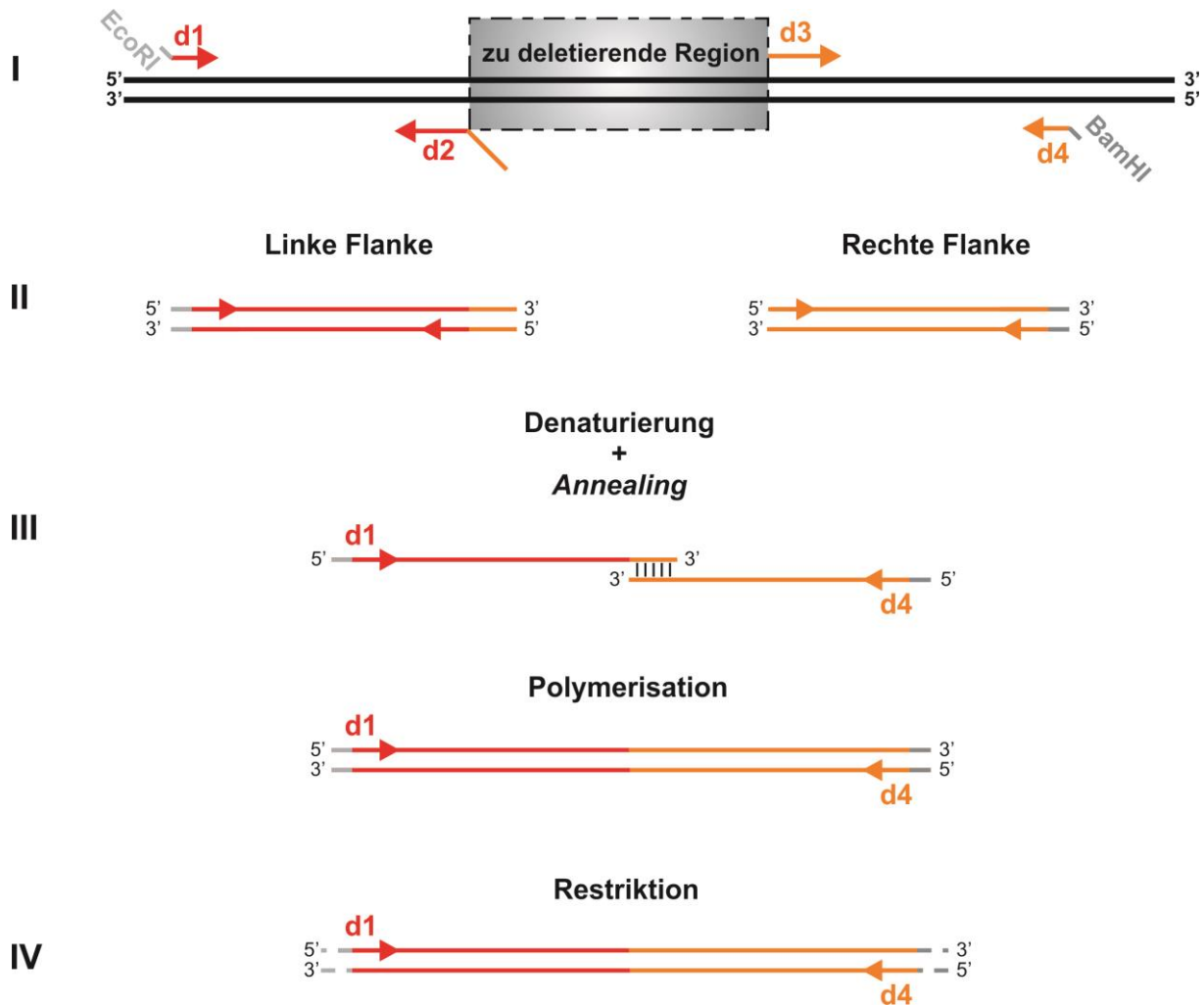


Abb. III-1 Gene Splicing by Overlap Extension – Erstellung eines Deletionskonstruktes.

I) Genregion mit eingezeichnetem zu deletierenden Bereich (graues Rechteck) inklusive Lage und Richtung der Primer d1 bis d4 (rote und orange Pfeile); die Primer d1 und d2 sind am 3'-Ende mit Restriktionsschnittstellen (graue Bereiche) und endständiger Kappe versehen. Die gewählten Restriktionsschnittstellen *Bam*HI und *Eco*RI sind nur beispielhaft. Der Primer d2 trägt einen komplementären Überhang zum d3 Primer. II) In 2 PCRs werden die flankierenden Regionen (Linke und Rechte Flanke) des zu deletierenden Bereichs amplifiziert. III) Nach der Denaturierung der Überhang des d2 Primers ein *annealing* der beiden Flanken. Die Überlappung der beiden Flanken dient somit als Primer für beide PCR-Fragmente. Durch die Polymerisation mit Hilfe der Primer d1 und d4 werden die Flanken zum neu kombinierten DNA-Produkt fusioniert. IV) Das Fusionsprodukt kann nun durch die angehängten Restriktionsschnittstellen geschnitten werden und mit einem entsprechend geschnittenen pK18*mobsacB* ligiert werden, um ein Deletionskonstrukt zu erstellen.

11.2. Erstellung von Deletionskonstrukten mittels Gibson-Assembly

Das Gibson-Assembly stellt eine Alternative zum *GeneSOEing* zur Erstellung von Deletionskonstrukten dar. Der Vorteil dieser Methode ist, dass bei der Erstellung der *pk18mobsacB*-Derivate keine weiteren Schritte zwischen der PCR der zu fusionierenden Fragmente und dem fertigen *pk18mobsacB*-Derivat liegen. Die Erstellung des Deletionskonstruktes aus den entsprechenden Vorprodukten und Insertion in den Vektor geschieht hier in einem einzigen Schritt. Das Gibson-Assembly ermöglicht die Fusion mehrere DNA-Fragmente unabhängig von der Fragmentlänge oder Sequenz in einer isothermalen Reaktion in einem einzigen Reaktionsgefäß (Gibson *et al.*, 2010, 2009), (Abb. III-2). Das Gibson-Assembly beinhaltet drei verschiedene enzymatische Schritte, die in einem Puffer ausgeführt werden:

- Die T5-Exonuklease führt zu einzelsträngigen 3'-Überhängen. Dies ermöglicht das Annealing von Fragmenten, die komplementäre Bereiche an einem Ende besitzen (Überlappung).
- Die Phusion Polymerase füllt die Lücken zwischen den zusammengelagerten Fragmenten auf.
- Die Taq-DNA-Ligase verbindet die freien 5'-Phosphatgruppen und 3'-Hydroxylgruppen.

Die zu fusionierenden DNA-Fragmente sollten mindestens eine Überlappung von 20 - 80 bp besitzen. Dies wird durch entsprechendes Primerdesign und folgender PCR, zur Produktion der zu fusionierenden DNA-Abschnitte gewährleistet. Des Weiteren wird der Vektor selbst mit entsprechenden Überlappungen zu den DNA-Fragmenten amplifiziert und im Gibson-Ansatz verwendet. Die entsprechenden Produkte werden in equimolaren Mengen in 5 µL angesetzt und zu 15 µL auf Eis aufgetautem Gibson-Mastermix (Kapitel 0) hinzugegeben. Hierbei sollten zwischen 50 - 100 ng pro Fragment eingesetzt werden. Eine weitere Möglichkeit ist, Inserts im 3-fachen Überschuss zum verwendeten Vektor einzusetzen. Der Ansatz wird bei 50 °C für 60 min inkubiert und kann direkt für Klonierungen, PCR oder die direkte Transformation eingesetzt werden.

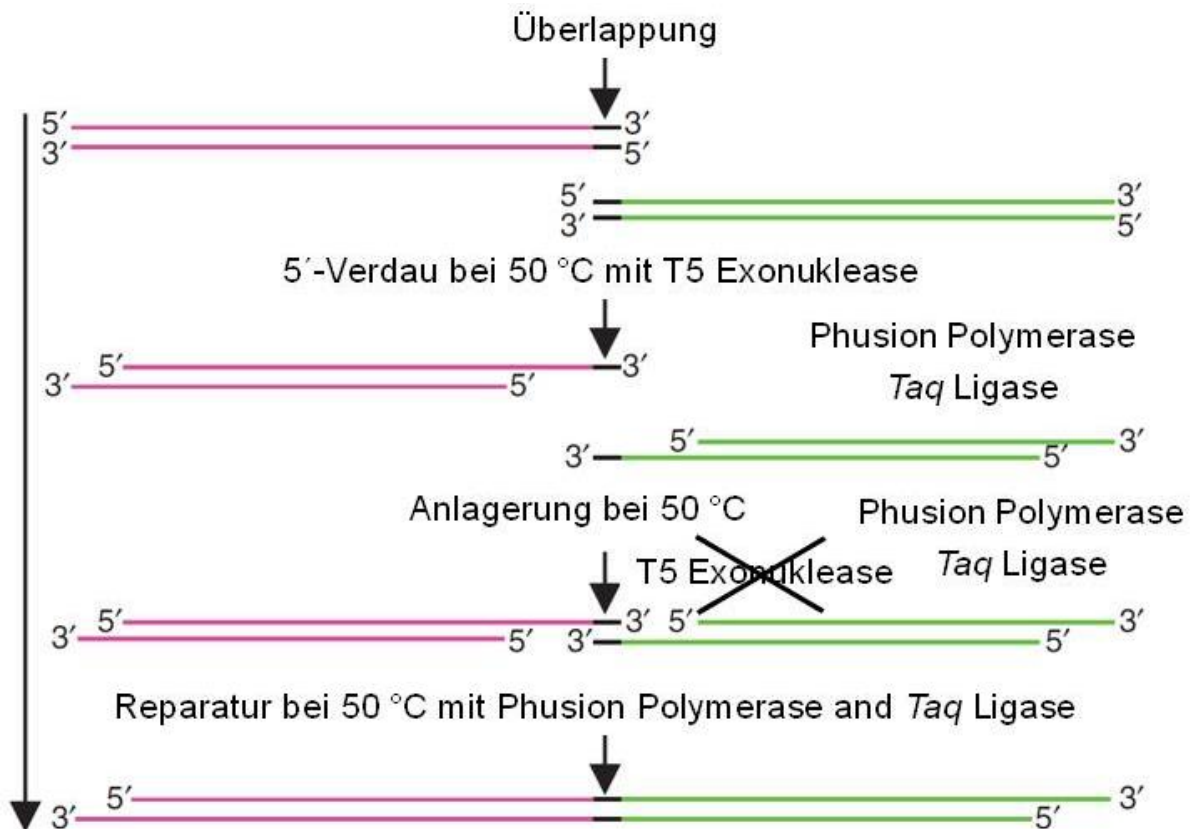


Abb. III-2 Gibson-Assembly - Isothermale *in vitro* Rekombination von DNA-Fragmenten in einem einzigen Schritt. Zwei aneinander angrenzende DNA-Fragmente (magenta und grün), die eine Überlappung an einem ihrer Enden besitzen (schwarz), werden zu einem DNA-Fragment in einem einzigen isothermalen Schritt fusioniert. Die T5-Exonuklease entfernt Nukleotide vom 5'-Ende der doppelsträngigen DNA. Die so entstehenden einzelsträngigen 3'-Überhänge der beiden DNA-Fragmente können sich im Bereich der Überlappung (schwarz) zusammenlagern, da sie komplementär sind. Die Phusion-Polymerase füllt nun die bestehenden Lücken auf. Final verbindet die Taq-Ligase die freien 5'-Phosphatgruppen und 3'-Hydroxylgruppen. Die T5-Exonuklease ist Hitze-instabil und wird im Verlauf der Inkubation bei 50 °C für 1 h inaktiviert. Abbildung nach (Gibson *et al.*, 2009).

11.3. Etablierung einer Deletion im Genom

Zur Etablierung einer Deletion, beziehungsweise eines DNA Austauschs im Genom von *C. glutamicum* wird ein entsprechendes pK18*mobsacB*-Derivat mittels Elektroporation in elektrokompente *C. glutamicum* Zellen transferiert. Das pK18*mobsacB*-Derivat trägt als Insert ein Deletionskonstrukt (fusionierter Bereich vor und nach dem zu deletierenden Bereich) welches durch *GeneSOEing* und Restriktion in den *shuttle*-Vektor kloniert wurde oder aber das Deletionskonstrukt-tragende pK18*mobsacB*-Derivat wurde mittels Gibson erstellt. Das Deletionskonstrukt wird mittels des pK18*mobsacB*-Systems über ein Doppel-Crossover gegen das native Gen ausgetauscht (Schäfer *et al.*, 1994).

Dabei wird durch Selektion auf Kanamycin zunächst auf die Integration des Vektors selektioniert, die über eine der beiden Flanken erfolgt. Für die Selektion auf das zweite Crossover-Ereignis wird das *sacB*-System des Vektors verwendet, welches eine Sensitivität

gegenüber Sucrose vermittelt. Sucrose-resistente Zellen haben diesen Vektor verloren, da der pK18*mobsacB*-Vektor in *C. glutamicum* nicht replizieren kann. Dieser Verlust kann auf zwei Arten erfolgen. Im ersten Fall erfolgt eine homologe Rekombination über dieselbe Flanke wie bei der Integration, so dass erneut der Wildtyp-Genotyp entsteht. Im zweiten Fall erfolgt die Rekombination über die zweite Flanke. Auf diese Weise verbleibt das Deletionskonstrukt im Chromosom, während das native Gen mit dem Vektor verloren geht.

11.4. Deletionsnachweis mittels PCR

Der Nachweis der Deletion erfolgt über eine PCR-Reaktion mit der *Taq*-Polymerase Testprimern, die außerhalb der linken und rechten Flanke des Deletionskonstruktes im Chromosom von *C. glutamicum* binden. Damit soll ausgeschlossen werden, dass der Vektor noch in der Genomsequenz integriert vorliegt. Dazu wird eine sehr kleine Menge Zellen für 5 min aufgekocht und die freigesetzte DNA als *Template* verwendet. Als Kontrolle dient Wildtyp-DNA.

12. Transkriptomik

12.1. Kultivierung und Lagerung der *C. glutamicum* Zellen

Für die RNA-Untersuchungen werden *C. glutamicum*-Zellen in Minimal- oder Vollmedium bis zur exponentiellen Phase (o.D.₆₀₀ = 6) angezogen, ca. $7,5 \times 10^9$ Zellen durch 15 s Zentrifugation pelletiert und nach Entfernung des Überstandes in flüssigem Stickstoff inaktiviert. Die Zellpellets können anschließend bei -80 °C gelagert werden.

12.2. Gesamt-RNA-Isolierung aus *C. glutamicum*

Für die Isolierung der Gesamt-RNA aus *C. glutamicum* wurde das RNeasy Mini Kit wie folgt verwendet:

- Vor dem Start β -Mercaptoethanol (14,3 M) (10 $\mu\text{L}/\text{mL}$ RLT-Puffer) zu der später verwendeten Menge RLT-Puffer (800 $\mu\text{L}/\text{Probe}$) hinzufügen
- Zentrifuge auf 4 °C runterkühlen
- Zellen (0,75 - 2 mL der Kultur; abhängig von der Zellzahl) in 800 μL RLT-Puffer (+ β -ME (10 $\mu\text{L}/\text{mL}$ RLT)) resuspendieren
- Lösung in vorgekühlte Ribotubes überführen
- 30 s bei 6.500 U/min ribolysieren,
- 3 min auf Eis
- 5 min bei 4 °C und 14.000 RZB zentrifugieren
- 500 μL Überstand in ein 1,5 mL Reaktionsgefäß mit 280 μL Ethanol überführen und durchmischen

Folgende Schritte bei RT:

- Ansatz auf die Säulen geben und für 30 s bei 14.000 RZB zentrifugieren
- Überstand verwerfen und neues Sammelgefäß nutzen
- 350 μL RW1-Puffer auf die Säule geben und für 30 s bei 14.000 RZB zentrifugieren

DNase-Behandlung auf der Säule (Qiagen DNaseI):

- 10 μL DNaseI, 0,5 μL RNase Inhibitor und 70 μL RDD-Puffer werden für jede Säule als Mastermix vorbereitet
- 80,5 μL Mastermix auf Säule geben; Inkubation für 60 min bei 25 °C
- 350 μL RW1-Puffer auf Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 500 μL RPE-Puffer auf die Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 500 μL RPE-Puffer auf die Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 2 min und 14.000 RZB trocken zentrifugieren
- Säule auf ein 1,5 mL Reaktionsgefäß setzen

- 50 µL RNase-freies Wasser auf die Säule geben und 1 min inkubieren
- 1 min bei 14.000 RZB zentrifugieren
- Elutionsschritt wiederholen

DNase-Behandlung im Reaktionsgefäß (Roche DNaseI):

- Mastermix erstellen: 10 µL DNaseI; 0,5 µL RNase Inhibitor und 12 µL DNaseI-Puffer mischen
- 22,5 µL zur eluierten RNA hinzugeben und 60 min bei 37 °C inkubieren
- 120 µL RNA-Lösung mit 420 µL RLT-Puffer und 350 µL EtOH mischen und auf die Säule geben
- 30 s bei 14000 RZB zentrifugieren, Durchfluss verwerfen und neues Sammelgefäß verwenden
- 350 µL RW1-Puffer auf die Säule geben und 30 s bei 14000 RZB zentrifugieren

DNase-Behandlung (Qiagen, Hilden, Deutschland) auf der Säule:

- 10 µL DNase, 0,5 µL RNase Inhibitor und 70 µL RDD-Puffer werden für jede Säule als Mastermix vorbereitet
- 80,5 µL Mastermix auf die Säule geben; Inkubation für 60 min bei 25 °C
- 350 µL RW1-Puffer auf die Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 500 µL RPE-Puffer auf die Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 500 µL RPE-Puffer auf die Säule geben, 30 s bei 14.000 RZB zentrifugieren, Durchfluss verwerfen
- 2 min und 14.000 RZB trocken zentrifugieren
- Säule auf ein 1,5 mL Reaktionsgefäß setzen
- 50 µL RNase-freies Wasser auf die Säule geben und 1 min inkubieren
- 1 min bei 14.000 RZB zentrifugieren
- Elutionsschritt wiederholen

Je 1 µL der eluierten RNA für die Test-PCR entnehmen

Falls negativ auf DNA-Rückstände getestet wurde, erfolgt in einem letzten Schritt die Aufkonzentrierung der RNA und weitere Reinigung mittels Microcon YM-30 nach Vorgabe des Herstellers.

12.3. Test-PCR auf DNA-Reste

Im Anschluss an die RNA-Isolierung wird eine Test-PCR auf DNA Reste mit der *Taq*-Polymerase durchgeführt. Dafür werden Testprimer eingesetzt, die ein etwa 250 bp großes und ein 500 bp großes Amplifikat der genomischen DNA von *C. glutamicum* erzeugen. Als Positivkontrolle dient Wildtyp-DNA. Wird mit den Testprimern ein Amplifikat in den RNA-Proben erhalten, erfolgt ein weiterer Verdau und eine weitere Aufreinigung.

12.4. RNA Konzentrationsbestimmung

Die RNA-Konzentration der Proben wird mit dem Nanodrop 2000 Spectrometer bei einer Wellenlänge von 260 nm bestimmt. Einen Hinweis auf mögliche Verunreinigungen der RNA-Proben gibt der Quotient der Absorptionen bei 260 und 280 nm (A_{260}/A_{280}). Ein Wert von $> 0,15$ deutet auf eine reine Probe hin. Für die Messung werden 2 μ L Probe eingesetzt.

12.5. Transkriptionsanalyse mittels RT-qPCR

Die quantitative reverse Transkriptions-PCR (RT-qPCR) wird auf Gesamt-RNA unterschiedlicher *C. glutamicum* Stämme mit dem LightCycler und dem SensiMix™ SYBR & Fluorescein One-Step Kit durchgeführt. Hierbei wird die differentielle Transkription einzelner Gene unterschiedlicher Proben detektiert. Für die Analyse im LightCycler werden 250 ng Gesamt-RNA eingesetzt. Die verwendeten Primer (Tab. III.3) amplifizieren jeweils etwa 200 bp große Fragmente, die in einer Schmelzkurvenanalyse überprüft werden. Unterschiede in der Transkriptmenge der Gene werden durch Vergleiche der *Crossing Points* (Cp) unterschiedlicher Proben von drei bzw. vier biologischen Replikaten, jeweils als zwei technische Replikate gemessen, ermittelt. Die Cp werden durch die LightCycler Software 3.5 ermittelt. Die Berechnung der durchschnittlichen Cp erfolgt durch die Bestimmung des Querschnittes der Sets an technischen Replikaten und der erst nachfolgenden Berechnung des Querschnittes der drei bzw. vier biologischen Replikate. Für jedes Set von drei bzw. vier biologischen Replikate wird die Standardabweichung, unter Annahme Normalverteilung der Cp, berechnet. Die kombinierte Standardabweichung für Δ Cp wurde durch die Standard-Berechnung der Fehlerfortpflanzung (unter der Annahme nicht-korrelativer Fehler) approximiert.

Über die Cp-Werte lassen sich anschließend die relativen Transkriptmengen durch einen Vergleich zu einer Referenz berechnen.

12.6. Globale Transkriptionsanalyse mittels DNA-Microarray-Technologie

Um die Transkriptionsprofile unterschiedlicher Stämme miteinander zu vergleichen wird die *Microarray*-Technologie verwendet. Dabei wird die Eigenschaft einzelsträngiger Nukleinsäuren ausgenutzt, bei vorliegenden komplementären Sequenzen spezifische Basenpaarungen einzugehen. Die für die Hybridisierung verwendeten *Slides* enthalten kovalent gebundene *Targets*, die einen Großteil der im Genom von *C. glutamicum* vorausgesagten *Open Reading Frames* (ORFs) repräsentieren. Als Sonde wird Fluoreszenzmarkierte cDNA zugegeben, die bei einer reversen Transkription der zu untersuchenden RNA hergestellt wird. Hierfür werden jeweils 15 µg RNA-Probe der beiden miteinander zu vergleichenden Bakterienkulturen in einer reversen Transkription mit BioScript Reverse Transcriptase mit Aminoallyl-dUTPs (aa-dUTPs) in cDNA umgeschrieben. Anschließend wird die cDNA mit zwei verschiedenen Fluoreszenzfarbstoffen, Cy3 und Cy5, gekoppelt. Es folgt eine Aufreinigung mit dem CyScribe GFX™ Purification Kit. Bei der folgenden Hybridisierung bindet nun die Sonde über Wasserstoffbrückenbindungen an komplementäre Bereiche auf dem *Slide*. Auf einem *Array* sind jeweils vier Replikate eines *C. glutamicum* Gens vertreten. Der detaillierte Ablauf des *Labeling* und der Hybridisierung des *C. glutamicum* *Microarrays* ist beschrieben in Rückert *et al.* (2008).

Durch den Einsatz zweier unterschiedlicher Farbstoffe (Cy3 und Cy5) können gleichzeitig zwei verschiedene Proben/Sonden miteinander verglichen werden. So kann darauf geschlossen werden, in welcher Probe ein bestimmtes Gen eine erhöhte mRNA-Menge aufweist (Hüser *et al.*, 2003). Um mögliche Unterschiede in den Eigenschaften der beiden Farbstoffe auszugleichen, empfiehlt es sich, die biologischen Replikate jeweils mit einer „*Label-Umkehr*“ zu markieren. Durch die anschließende Zusammenfassung der gewonnenen Daten werden Abweichungen durch Farbstoffunterschiede beseitigt. Nach der Hybridisierung am Automated Slide Processor HS 4800 wird der *Slide* am ScanArray 4000 gescannt. Die Spoterkennung erfolgt mit Hilfe des Programms ImaGene. Die Normalisierung und Auswertung der Daten wird mit EMMA (Dondrup *et al.*, 2009) durchgeführt. Hier erfolgt die Normalisierung mittels der LOWESS Funktion (*locally weighted scatterplot smoothing*). Die Evaluierung der Hybridisierungsdaten wird mit einem *cut off* > 7 für die mittlere Signalintensität (α -Wert) und einem *cut off* $\pm 0,6$ für den m -Wert (\log_2 -Verhältniss der mittleren Signalintensität der verglichenen Proben) durchgeführt. Ein m -Wert von 0,6 entspricht der technischen Varianz und einer Änderung der relativen Transkriptionsstärke um das 1,5-fache ($2^{(m\text{-Wert})}$) (Hüser *et al.*, 2003). Dabei wird eine Fehlerwahrscheinlichkeit von weniger als 5 % im *students t test* zugrunde gelegt. Für eine stringenterer Auswertung kann auch ein m -Wert von 1 als Signifikanzgrenze gewählt werden, was einer relativen Änderung der Transkriptmenge von 2 bzw. 0,5 entspricht.

12.7. Northern Blot Analyse

Im Northern Blot wird in einem Gel aufgetrennte RNA auf eine Membran übertragen und kann durch Hybridisierung mit spezifischen RNA-Sonden nachgewiesen werden. Die Namensgebung Northern ist eine Anspielung auf den Southern Blot, der von Edwin Southern entwickelt wurde und mit DNA, statt RNA arbeitet. Da der wie zuvor beschriebene Northern Blot (Homuth et al., 1996) mit einigen Modifikationen durchgeführt wurde, wird die Durchführung hier im Detail beschrieben.

Die mit Digoxigenin (DIG) gelabelten RNA Sonden für spezifische Transkripte von Genen (in dieser Arbeit für die Gene *sigH* und *rshA*) werden per *in vitro* Transkription mit T7-RNA-Polymerase, NTP-DIG-label Mix (DIG RNA Labeling Kit) und Gen-spezifischen Primern, wobei der *reverse* Primer eine angehängte T7-Promotersequenz besitzt, hergestellt. Vor der Hybridisierung werden die so erstellten Sonden für 10 min bei 95 °C inkubiert und dadurch denaturiert. 5 µg Gesamt-RNA unterschiedlicher Stämme, isoliert wie unter III.12.1 und III.12.2 beschrieben, werden unter denaturierenden Bedingungen in 1 %igen Agaroseformaldehydgelen mit 1x MOPS-Laufpuffer aufgetrennt und mittels Ethidiumbromid gefärbt. Die nun aufgetrennte RNA wird durch Vakkum-*Blotting* auf eine Hybond-N Membran übertragen. Hierfür wird das Gel während des Vakkum-*Blotting* zuerst 5 min mit Denaturierungslösung und nachfolgend 5 min mit Neutralisierungslösung behandelt. Das *Blotting* findet für 5 h unter ständiger Zugabe von 20x SSC statt.

Die Hybridisierung mit den erstellten DIG-gelabelten RNA-Sonden und die Detektion der Hybridisierungssignale finden wie folgt statt. Die Membran wird nach dem Backen (0,5 h bei 120 °C) mit der Prähybridisierungslösung ohne Sonde vorbehandelt, um unspezifische Bindestellen der Membran zu blocken. In einem zweiten Schritt findet die Hybridisierung mit den Digoxigenin (DIG)-gelabelten RNA-Sonden (50 ng/mL) bei 68 °C über Nacht statt. Die nun hybridisierte Membran wird durch zweimaliges Waschen mit Waschlösung I (10 min bei RT) und dreifachem Waschen mit Waschlösung II (15 min bei 68 °C) von überschüssiger Hybridisierungslösung befreit. Um die Hybridisierungssignale immunologisch zu detektieren wird Anti-Digoxigenin-AP, Fab fragments und CDP-*Star* nach Angaben des Herstellers verwendet. Die Blots werden zunächst mit dem Anti-Digoxigenin Antikörper, der mit alkalischer Phosphatase (AP) konjugiert ist, behandelt. Hierfür werden 2 µL DIG-Antikörper in 20 mL Puffer II gelöst und die Membran für 30 min inkubiert. Die Membran wird anschließend zweimal für 30 min mit Waschpuffer gewaschen um überschüssigen Antikörper zu entfernen. Das folgende Waschen mit Puffer III dient der Vorbereitung auf die enzymatische Reaktion der AP mit CDP-*Star*. Jetzt werden 7,5 µL CDP-*Star* in 2,5 mL Puffer III verdünnt und auf die Membran gegeben. Hier wird CDP-*Star* durch die AP dephosphoryliert, was zu einer Lichtemission bei

466 nm führt. Diese Hybridisierungssignale werden mit dem Luminescent Image Analyzer LAS-3000 detektiert. Die Größe der über die Hybridisierungssignale detektierten Transkripte wird durch den Vergleich mit den Markierungen des im Verlauf des Northern Blots mit Ethidiumbromid gefärbten High Range Markers auf der Membran bestimmt.

12.8. RNA-Sequenzierung (RNA-Seq)

Da die Hochdurchsatz-Sequenzieretechnologien nicht nur zur Sequenzierung von DNA verwendet werden können, sondern auch für die RNA-Sequenzierung (RNA-Seq) zur Verfügung stehen, konnten Methoden entwickelt werden, die es ermöglichen Transkripte genomweit zu analysieren (Van Vliet, 2010). Diese Methoden führen unter anderem auch zur Identifizierung einer enormen Anzahl an Transkriptionsstartpunkten (TSP) und damit auch zur Lokalisierung der entsprechenden Promotoren. Die Identifizierung eines genomweiten „Promoteroms“ ist ein großer Schritt in der qualitativen und auch quantitativen Betrachtungsweise von Promotoren, da die Methode zum einen eine Einzelnukleotidauflösung besitzt und zum anderen Auskunft über die relative Menge an mRNA mit intaktem (primären) 5'-Ende gibt (Pátek *et al.*, 2013) .

Um das Transkriptom durch RNA-Seq zu untersuchen, wird zuerst eine RNA-Seq-Bibliothek erstellt. Dieser Schritt ist entscheidend für eine erfolgreiche Sequenzierung. Bei der Erstellung einer solchen RNA-Seq-Bibliothek gilt es mehrere Dinge zu beachten. Zum einen ist es entscheidend die stabilen RNA-Spezies (rRNAs und tRNAs) aus einer Gesamt-RNA-Isolierung effektiv zu entfernen, da diese repetitiven Spezies mehr als 95 % der isolierten RNA bzw. der RNA in einer Bakterienzelle ausmachen. Zum anderen soll die Information, von welchem Strang ein Transkript stammt, erhalten bleiben. Diese und weitere Punkte wurden bei der Erstellung eines Protokolls zur Erstellung einer nativen 5'- RNA-Seq-Bibliothek (nur 5'-Enden von Primärtranskripten) berücksichtigt. Zum Zeitpunkt dieser Arbeit ist das Protokoll noch nicht publiziert (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung). Im Folgenden wird das Protokoll zur Erstellung einer nativen 5' -RNA-Seq-Bibliothek daher im Detail vorgestellt.

12.8.1. Native 5'-Enden RNA-Seq-Protokoll

Um die nativen TSP von *C. glutamicum* genomweit zu untersuchen, deren Vorhandensein im Genom in erster Linie von der Aktivität der RNAP abhängt, wird das native 5'-Enden-RNA-Seq-Protokoll (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung) verwendet. Dies setzt sich wie folgt zusammen.

Das native 5'-Enden-RNA-Seq-Protokoll startet mit der Verwendung von 10 µg Gesamt-RNA. Die stabilen RNA-Spezies werden unter Verwendung des Ribo-Zero™ rRNA Removal Kit effektiv entfernt. Hiernach wird die nun hauptsächlich aus mRNAs bestehende RNA auf eine Größe zwischen 200 - 500 bp fragmentiert. Die abgereichert und fragmentierte RNA wird einer Behandlung mit Terminator™ 5'-Phosphate-Dependent Exonuclease (TEN) unterzogen, wodurch die primären Transkripte angereichert werden. Die Exonuklease TEN erkennt und spaltet nur degradierte, gebrochene oder prozessierte RNA, die charakteristischerweise entgegen den Primärtranskripten ein Monophosphat an ihrem 5'-Ende besitzen. Somit verbleiben nur die bakteriellen Primärtranskripte mit ihrem Triphosphat in der RNA-Lösung. Für die Ligation des RNA-Adapters werden die Triphosphate der Primärtranskripte in 5'-Monophosphate durch eine RNA 5'-Polyphosphatase umgewandelt. Die für die Ligation vorbereiteten RNA-Moleküle werden durch die T4 RNA Ligase (Epicenter) mit einem Voll-RNA-Adapter versehen. Das *Tagging* des 3'-Endes der RNA-Fragmente und die Konvertierung in cDNA durch reverse Transkription erfolgt in diesem Protokoll in einem Schritt und ohne Ligation. Hierfür wird ein *Stem-Loop (Loop)* DNA-Adapter verwendet, der an seinem 3'-Ende mit 7 *Wobble*-Basen versehen ist. Diese *Wobble*-Basen am 3'-Ende des DNA-Adapters hybridisieren mit den 3'-Enden der RNA-Fragmente und fungieren somit als Primer für die Reverse Transkriptase. Nachfolgend werden die cDNA-Fragmente mit entsprechenden Sequenzierprimern amplifiziert. Die so erstellte cDNA-Bibliothek steht für die Sequenzierung durch den Genome Analyzer IIx und dem entsprechenden TruSeq DNA Sample Preparation Kit (LT, Set A) zur Verfügung. Der Ablauf des nativen 5'-Enden-RNA-Seq-Protokolls inklusive der Sequenzierung und bioinformatischen Auswertung ist in Abb. II-7 grafisch zusammengefasst. Die einzelnen Schritte der Bibliotheks-Erstellung werden im weiteren Verlauf erläutert.

12.8.2. Entfernung ribosomaler RNA

10 µg Gesamt-RNA werden nach Vorgaben des Herstellers mit dem Ribo-Zero rRNA Removal Kit von rRNA befreit. Anschließend erfolgt eine Aufreinigung der RNA mittels Ethanolpräzipitation.

12.8.3. Phenol-Chloroform-Isoamyl-Alkohol-Extraktion

Die RNA wird nach jedem enzymatischen Schritt einer Phenol-Chloroform-Isoamylalkohol-Extraktion (PCI) unterzogen. Hierzu wird 1 Volumen PCI-Lösung zu der Probe hinzugefügt, 30 s gründlich gemischt und dann bei 16.000 RZB für 15 min zentrifugiert. Der wässrige Überstand wird restlos in ein neues Reaktionsgefäß überführt. Die Ethanolpräzipitation erfolgt durch Zugabe von 0,3 Volumen Natriumacetat (3 M; pH 5,2), 20 µg Glykogen und 2,7 Volumen eiskaltem Ethanol (~99 %). Die Lösung wird gut durchmischt und über Nacht bei -20 °C inkubiert. Dann wird die Probe bei 16.000 RZB für 20 min bei 4 °C zentrifugiert. Das abzentrifugierte RNA Pellet wird dreimal vorsichtig mit Ethanol (75 %) gewaschen und bei 37 °C im Heizblock für 5 min an der Luft getrocknet und final in RNase-freiem Wasser gelöst.

12.8.4. RNA Fragmentierung

Die RNA wird durch Zugabe von 0,25 Volumen Fragmentierungspuffer mit anschließender sehr guter Durchmischung und nachfolgender Inkubation bei 94 °C für genau 2,5 min zu einer Größe von 200 - 500 bp fragmentiert. Um die Fragmentierung nach genau 2,5 min zu stoppen, wird 1 Volumen eiskalter Fragmentierungsstoppuffer hinzugegeben und die Lösung für 5 min auf Eis gekühlt. Die RNA-Probe wird wie unter 0 beschrieben durch Ethanolpräzipitation gefällt. Die Fragmentierung der RNA wird mittels Kapilargelelektrophorese an einem Agilent Bioanalyser 2100 mit dem Agilent RNA 6000 Pico Kit und Agilent RNA 6000 Nano Kit bestimmt. Hierbei wird nach Angaben des Herstellers verfahren. Wenn die Fragmente nicht im Bereich von 200 - 500 bp liegen, wird die Fragmentierung mit verkürzter Zeit so oft wiederholt, bis die notwendige Fragmentgröße für die weitere Bearbeitung erreicht ist. Die korrekte Fragmentgröße ist entscheidend für die erfolgreiche Sequenzierung der Bibliothek.

12.8.5. Terminator 5'-Phosphate-dependent Exonuklease Behandlung

Um die nicht primären Transkripte abzubauen wird die fragmentierte RNA mit Terminator 5'-Phosphate-Dependent Exonuklease (TEN) des Herstellers behandelt.

12.8.6. 5'-Enden Vorbereitung auf Ligation

Da zur Ligation des Adapter die T4 RNA Ligase 1 (New England BioLabs) zum Einsatz kommt und diese nur 5'-Monophosphate akzeptiert, ist es vorher notwendig die 5'-Triphosphate der Primärtranskripte in 5'-Monophosphate zu konvertieren. Hierfür wird die RNA 5'-Polyphosphatase nach Angaben des Herstellers verwendet. Das Enzym entfernt sequentiell die γ - und β -Phosphate von 5'-Triphosphat RNA (primäre RNA Transkripte): $5'\text{-pppN-OH } 3' \rightarrow 5'\text{-pN-OH } 3' + 2 \text{ P}_i$. Nach der Behandlung der RNA mit der RNA 5'-Polyphosphatase wird diese einer PCI-Extraktion unterzogen (s. 0). Die so vorbereitete RNA wird im Folgeschritt mit einem Voll-RNA-Adapter versehen.

12.8.7. RNA-Adapterligation

Für die Ligation des Voll-RNA-Adapters an das 5'-Ende der vorbereiteten RNA werden 100 μM RNA-Adapters mit der T4 RNA Ligase 1 (New England BioLabs) nach Vorgaben des Herstellers eingesetzt.

12.8.8. Vorbereitung des *Loop*-DNA-Adapters

Bevor der *Stem-Loop* (*Loop*) DNA-Adapter in der nachfolgenden reversen Transkription eingesetzt werden kann, wird die korrekte Faltung des Adapters sicher gestellt. Dies geschieht durch eine Hitzedenaturierung des Adapters und kontrolliertem Abkühlen auf 25 °C, wodurch sich der Adapter in die gewünschte *Loop*-Struktur faltet. Es werden 20 μL des *Loop*-Adapters (100 μM) für 3 min bei 98 °C inkubiert und anschließend bei einer Rate von 1 °C/10 s auf 25 °C im Mastercycler® pro S abgekühlt.

12.8.9. cDNA-Synthese und *Tagging* des 3'-Endes der cDNA

Die cDNA-Synthese durch reverse Transkription geht einher mit dem Anhängen einer flankierenden Sequenz an das 3'-Ende der cDNA durch die Nutzung des DNA-*Loop*-Adapters. Die mit einem 5'-Voll-RNA-Adapter versehene RNA wird durch das ThermoScript™ RT-PCR System nach Angaben des Herstellers per reverser Transkription in cDNA konvertiert. Es wird 1 μL des korrekt gefalteten DNA-*Loop*-Adapters (100 μM) als cDNA-Syntheseprimer verwendet. Dies ist möglich, da der DNA-*Loop*-Adapter 7 *Wobble*-Basen am 3'-Ende trägt, die mit den 3'-Enden der RNA-Fragmente hybridisieren und somit als Primer für die reverse Transkriptase fungieren können. Der restliche Anteil des DNA-*Loop*-Adapters ist nun als flankierende Sequenz an das 3'-Ende der cDNA angefügt. Die cDNA wird direkt für die Amplifikation genutzt.

12.8.10. Amplifikation cDNA und Aufreinigung der 5'- cDNA-Bibliothek

Die cDNA wird mit Phusion® High-Fidelity DNA Polymerase (New England BioLabs) erstellt.

100 µL Phusion-Reaktionsansatz:

Betain (5 M)	30 µL
HF-Puffer	20 µL
Primer 1 (10 µM)	2 µL
Primer 2 (10 µM)	2 µL
dNTPs (je 2,5 mM)	10 µL
Phusion-Polymerase (0,5 U/µL)	0,5 µL
DMSO (100 %)	1,5 µL
Template cDNA aus Loop-RT	34 µL

Der 100 µL Phusion-PCR-Ansatz für die cDNA Amplifikation wird für die PCR auf zwei PCR-Tubes aufgeteilt. Es wird eine low-cycle Amplifikation im Thermocycler Mastercycler pro S (Eppendorf) durchgeführt.

Low-Cycle 3-Step-PCR-Programm der Phusion-Polymerase:

Zyklen	Temperatur	Zeit	Schritt
	98 °C	bis Eingabe	Block temperieren
1	98 °C	30 s	Initiale Denaturierung
	98 °C	10 s	Denaturierung
18	60 °C	30 s	<i>Annealing</i>
	72 °C	30 s	<i>Elongation</i>
1	72 °C	2 min	Finale <i>Elongation</i>
	4 °C	∞	Lagerung

Nach der Amplifikation der cDNA-Bibliotheken werden die Proben nach Vorgaben des Herstellers mit dem Agencourt AMPure XP - PCR Purification System aufgereinigt und von DNA > 150 bp (Primer- bzw. Adapterdimere) befreit.

Die Qualitätskontrolle und Quantifizierung findet mit dem Agilent High Sensitivity DNA Kit am Agilent 2100 Bioanalyser statt (III.7.7.1). Falls weitere DNA-Moleküle > 150 bp vorhanden sind, wird die Aufreinigung mit Agencourt AMPure XP - PCR Purification System wiederholt oder die Methode der Gelextraktion angewandt. Hierbei finden hochreine Agarosegele Anwendung, um die Sequenzierung nicht negativ zu beeinflussen. Für die Gelextraktion von cDNA-Bibliotheken stehen Kits verschiedener Hersteller zur Verfügung, z. B. das Qiagen QIAquick Gel Extraction Kit.

12.8.11. Sequenzierung der 5'-cDNA-Bibliotheken

Die Sequenzierung der cDNA-Bibliotheken wird am Centrum für Biotechnologie der Universität Bielefeld unter der Verwendung der Cluster Station, dem Genome Analyzer Iix durchgeführt. Die Bibliotheken werden *Single-Read* sequenziert (1x35 bp). Für jede Probe wird hierbei eine *Lane* verwendet. Die primäre Datenanalyse und das *Base Calling* wird von der Illumina instrument Software durchgeführt.

12.9. Bioinformatische Analyse der RNA-Seq-Daten

Die enorme Menge an kurzen cDNA-*Reads*, die durch die Sequenzierung mit den Hochdurchsatz-Sequenzieretechnologien entstehen, muss auf ein Referenzgenom gemappt werden.

12.9.1. Mapping der sequenzierten *Reads* durch SARUMAN

Vor dem *Mapping* werden die *Reads* getrimmt. Die getrimmten *Reads* (26 bp) werden unter Verwendung von SARUMAN (semiglobal Alignment of short Reads using CUDA and NeedleMAN-Wunsch) (Jochen Blom et al., 2011) auf das Referenzgenom von *C. glutamicum* ATCC 13032 (Kalinowski et al., 2003) mit einer zulässigen Fehlerrate von 6 % der *Read*-Länge gemappt.

12.9.2. RNA-Seq-Datenvisualisierung durch VAMP

Für die Visualisierung der gemappten *Reads* wird die Software VAMP (Visualization and Analysis of Mapped Sequences) genutzt (Hilker et al., 2013, in Vorbereitung). Im Detail werden die durch SARUMAN auf das Referenzgenom *C. glutamicum* ATCC 13032 (Kalinowski et al., 2003) gemappten *Reads* auf das entsprechende Genom mit der *Genome Accession ID* BX927147 visualisiert.

12.9.3. Identifikation von Transkriptionsstartpunkten

Um in den Datensätzen aus den SARUMAN-*Mappings* mögliche TSP zu identifizieren, wurden die Daten mittels *ad hoc* Perl-Skripten analysiert. Hierzu wurden die Startpunkte der gemappten *Reads* für jede Position des Genoms strangspezifisch aufsummiert. Als TSP wurden alle Positionen angenommen, an denen mindestens 15 Startpunkte kumuliert sind und an der Position direkt stromaufwärts 5 fach weniger Startpunkte auftreten (Rückert, pers. Mitteilung).

12.9.4. *In vitro* Run-Off-Transkription

Die *in vitro* Run-Off-Transkription (RO) wird wie für *E. coli* beschrieben (Maciag *et al.*, 2011), durchgeführt, jedoch mit einigen Anpassungen für *C. glutamicum*. Bei dem hier beschriebenen Ansatz wird an Stelle der *C. glutamicum* Core-RNA-Polymerase (RNAP) die *E. coli* Core-RNAP als heterologes Enzym genutzt, wie bereits für *M. tuberculosis* (Rodrigue *et al.*, 2007) beschrieben. Es wird das *E. coli* RNAP-Core-Enzym mit nativen *C. glutamicum* ECF-Sigmafaktoren zu einem heterologen aktiven RNAP-Holoenzym rekonstituiert. Die ECF-Sigmafaktoren werden hierfür mit dem IMPACT™ Kit (*Intein* Mediated Purification with an Affinity Chitin-binding Tag) aufgereinigt (III.13.1). Als *Template* für die *run-off*-Transkription wird gDNA von *C. glutamicum* ATCC 13032 genutzt. Diese wurde, wie in III.7.5 beschrieben, aufgereinigt und mit FastDigest EcoRI fragmentiert (III.7.7). Die durchschnittlichen Längen der geschnittenen genomischen DNA-Fragmente werden mittels Kapillargelelektrophorese am Agilent Bioanalyzer 2100 bestimmt (III.7.7.1.). Die resultierenden Fragmentgrößen reichen von 0,6 kb bis 30 kb, wobei 1 µg geschnittener genomischer *C. glutamicum* DNA für eine einzelne *run-off*-Transkription eingesetzt wird.

Für die Rekonstitution des RNAP-Holoenzym wird das *E. coli* RNA Polymerase Core Enzym in einem molaren Verhältniss von 1:10 mit einem *C. glutamicum* ECF-Sigmafaktor im Thermocycler für 10 min bei 37 °C inkubiert. Für den Einsatz des RNAP-Holoenzym in der *run-off*-Transkription wird davon ausgegangen, dass die RNAP voll mit dem Sigmafaktor durch dessen hohen Überschuss, gesättigt ist. Die *in vitro* Transkription wird in RNase-freien PCR-Reaktionsgefäßen im Thermocycler in *run-off*-Puffer durchgeführt.

Es wird wie folgt für die *in vitro* *run-off* Transkription im Detail verfahren:

run-off-Puffer (1,5x) inklusive Ribolock frisch ansetzen:

HEPES (1 M, mit Tris-HCl auf pH 8,0)	2,4 µL	40 mM (finale Konzentration)
MgCl ₂ (25 mM)	24 µL	10 mM (finale Konzentration)
Kaliumglutamat (1 M)	9 µL	150 mM (finale Konzentration)
DTT (500 mM)	0,24 µL	2 mM (finale Konzentration)
BSA RNase-frei (10 mg/mL)	0,6 µL	1 µg/10 µL (finale Konzentration)
Triton (0,5 %)	1,2 µL	0.01 % (finale Konzentration)
Ribolock (40 U/µL)	1,5 µL	1 U/µL (finale Konzentration)
H ₂ O	1,1µL	

60 µL run-off-Reaktionsansatz:

run-off-Puffer (1,5x)	40 µL	1x (finale Konzentration)
H ₂ O	x µL *	
In Schritt (ii) hinzufügen		
RNAP Core- oder Holoenzym	y µL **	100 nM (finale Konzentration)
In Schritt (iii) hinzufügen		
fragmentierte gDNA (100 ng/µL)	10 µL	1 µg (finale Konzentration)
In Schritt (iv) hinzufügen		
NTPs (je 10 mM)	3 µL	500 µM (finale Konzentration)

*abhängig vom Volumen der RNAP (y µL)

**abhängig von der Konzentration der RNAP

- i. Vorlegen von 40 µL *run-off*-Puffer (1,5x) (s. III.5.10) in PCR-Reaktionsgefäßen; x µL H₂O hinzufügen – Menge hängt von der Konzentration der RNAP ab (7 µL - y µL RNAP (Final: 100 nM in 60 µL) = x µL H₂O)
- ii. 100 nM Holo-RNAP (rekonstituiert) oder *Core*-RNAP (als Negativkontrolle), bezogen auf 60 µL Endvolumen, hinzufügen (y µL) und für 10 min bei 37 °C inkubieren
- iii. *Template* hinzufügen: 10 µL 100 ng/µL gDNA von *C. glutamicum* hinzufügen (Endvolumen 57 µL) und für 15 min bei 37 °C im Thermocycler inkubieren. Dies erlaubt das binden der RNAP an die DNA
- iv. Starten der *in vitro* Transkription durch Zugabe von 3 µL of NTP-Mix (je 10 mM)
- v. Inkubation 30 min bei 37 °C
- vi. Stoppen der *in vitro run-off*-Transkription durch Inkubation bei 65 °C für 5 min
- vii. Direkte DNase-Behandlung mit 10 U (1 µL) DNaseI (Roche) bei 37 °C für 30 min, um genomische DNA anzubauen
- viii. Lagerung der RNA bei –80 °C!

Für die Aufreinigung der RNA aus wird das RNeasy Mini Kit genutzt (s. III.12.2). Zehn unabhängige *in vitro run-off*-Transkriptionen werden hierfür gepoolt. Die resultierende Gesamtmenge liegt meist bei 1 µg RNA und wurde mittels Nanodrop bestimmt. Die durchschnittliche Länge der Transkripte, wie auch die daraus resultierende cDNA wird mittels Kapillargelelektrophorese an einem Agilent Bioanalyzer 2100 mit dem Agilent RNA 6000 Pico Kit und dem Agilent RNA 6000 Nano Kit bestimmt. Hierbei wird nach Angaben des Herstellers verfahren. Die RNA steht nun für die weitere Untersuchung durch RNA-Seq (III.12.8) bereit.

12.10. **Run-Off-Transcription/RNA-Sequencing (ROSE)**

Run-Off-Transcription/RNA-Sequencing (ROSE) erlaubt die direkte genomweite Identifikation von TSP in Abhängigkeit eines Regulators. ROSE ist eine in dieser Arbeit durchgeführte Weiterentwicklung der Methode ROMA (*Run-Off-Transcription/Microarray-Analysis*) (MacLellan et al., 2009). Die ursprüngliche Methode ROMA nutzt die RNAP mit und ohne Regulator (z. B. Sigmafaktoren), um Transkripte abhängig vom gewählten Regulator *in vitro* auf Basis von gDNA zu transkribieren (*Run-Off-Transkription*) um Regulator-abhängige Transkription zu identifizieren. Während bei ROMA die *in vitro* Transkripte durch *Microarray*-Analyse untersucht werden, wird bei ROSE die RNA durch RNA-Seq analysiert. Für die RNA-Seq werden 5'-angereicherte Bibliotheken erstellt (III.12.8). Im Vergleich zur RNA-Seq-Analyse von Gesamt-Transkripten, wird bei ROSE die *in vitro* transkribierte RNA keinem rRNA-Abbau unterzogen. Mit der Wahl von RNA-Seq von 5'-angereicherten Bibliotheken als Analyseverfahren wird ein Problem des ROMA Ansatzes, das *Read-Through* von Genen bei der *in vitro* Transkription, was bei der Analyse durch *Microarrays* zu falsch-positiven Signalen führt, eliminiert. Mit ROSE werden nur die tatsächlichen TSP mit einer Einzelbasenauflösung identifiziert. Weiter ist diese Methode nicht vom Design von *Microarrays* abhängig und kann zur Identifizierung vorher nicht bekannter Transkripte beitragen. ROSE erlaubt somit die direkte Identifizierung von TSP die primär von einem Regulator abhängig sind. Dies geschieht unter Ausblendung aller reprimierenden Transkriptionsfaktoren oder sekundären Effekten des Regulators.

12.11. **DNA-Sequenzierung**

DNA-Sequenzierungen für PCR-Produkte oder Plasmide werden beim IIT Biotech (Bielefeld) in Auftrag gegeben. Die Proben werden per SANGER-Sequenzierung sequenziert.

13. Proteinherstellung und -analytik

13.1. Proteinaufreinigung mittels IMPACT™-CN-System

Die Aufreinigung der ECF-Sigmafaktoren und des Sigmafaktors SigA erfolgt mit dem IMPACT™ System nach den Vorgaben des Herstellers. Das Prinzip von IMPACT™ (*Intein Mediated Purification with an Affinity Chitin-binding Tag*) beruht auf der Erzeugung eines Fusionsproteins bestehend aus dem Zielprotein, einem Intein sowie einem Chitinbindepotein (CBD). Die CBD dient hierbei als Affinitäts-Tag. Dieser erlaubt die Aufreinigung über eine Chitinaffinitätssäule, von der das Zielprotein durch die induzierbare Selbstspaltungsaktivität des Inteins (*Protein-Splicing-Element*) nach stringenter Waschung in reiner Form ohne den Einsatz von Proteasen eluiert werden kann. Die Induktion der *On-Column*-Spaltung erfolgt durch die Zugabe von Thiolreagenzien wie Dithiothreitol (DTT). Der Ablauf einer solchen Aufreinigung ist schematisch in Abb. III-3 dargestellt.

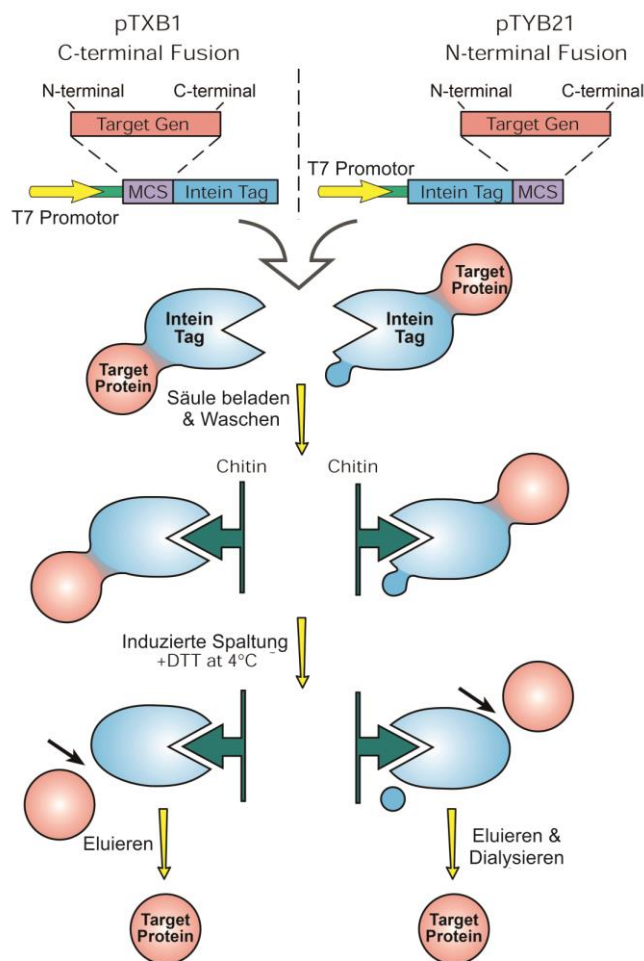


Abb. III-3 Schematische Darstellung des IMPACT™-CN-System (NEB, New England Biolabs, USA). Die Reinigung des Intein-CBD-Fusionsproteins erfolgt mit einer Chitinsäule. Zur Entfernung des Intein-Tags wird die selbstspaltende Aktivität der Inteine genutzt, die durch Thiole induziert wird. Abb. aus (Manuel IMPACT™-CN System; NEB, New England Biolabs, USA)

13.1.1. Konstruktion von Protein-CBD-Inteinfusionen

Um ein Intein-CBD-Fusionsprotein zu erzeugen wird das entsprechende Gen in einen der „IMPACT“-Vektoren kloniert. Hierfür stehen Vektoren für eine C-terminale oder N-terminale Fusion mit unterschiedlicher *Multiple Cloning Sites* (MCS) zur Verfügung. Die Klonierarbeiten werden in dem Stamm *E. coli* JM109 durchgeführt. Das fertige Konstrukt wird sequenziert und anschließend in den Expressionstamm *E. coli* ER2566 transformiert. Dieser Stamm trägt eine Kopie der T7-RNA-Polymerase innerhalb des vom *lac*-Promotor kontrollierten *lacZ*-Gens. In Abwesenheit von IPTG wird die Expression der T7-RNA-Polymerase durch die Bindung des *lacI*-Repressors an den *lac*-Promotor unterdrückt. Zudem enthält der Stamm weder *lon*- noch *ompT*-Proteasen.

13.1.2. Expression der Intein-CBD-Fusionsproteine

Für die Expression der rekombinanten Proteine werden die Plasmid-tragenden *E. coli* ER 2566 Zellen bei 37 °C selektiv (200 µg/mL Amp) bis zu einer o.D.₆₀₀ zwischen 0,6 und 0,8 angezogen. Die Expression des klonierten Gens wird mit 0,5 mM IPTG induziert. Die Expression erfolgt bei 19 °C über Nacht. Die Dauer der Expression und auch die Temperatur sind variabel und müssen ggf. angepasst werden um *Inclusion Bodies* zu vermeiden. Hierzu gibt es eine Vielzahl von Vorschlägen seitens der Hersteller NEB. Nach erfolgter Expression des Fusionsproteins werden die *E. coli* ER2566 Zellen geerntet und in 25 mL Aufschlusspuffer resuspendiert. Für den Aufschluss des Zellextrakts wird die French Pressure Cell Press bei 1.200 psi Druck verwendet. Der Zellextrakt wird im Anschluss bei 4 °C für 20 min zentrifugiert. Im Überstand befindet sich das Fusionsprotein in Lösung.

13.1.3. Proteinaufreinigung mittels IMPACT-CN-System

Die Proteinaufreinigung mittels IMPACT-CN-System wird, wie oben erwähnt, nach den Vorgaben des Herstellers durchgeführt und ist hier kurz dargestellt:

- 14 mL Proteinofilter® Columns gleichmäßig mit 5 mL Chitinkügelchen befüllen; mit Polycarbonat-Filter abdichten
- Äquilibration der Säulen mit 10 Volumen Säulenpuffer (4 °C)
- Proteinlösung auf die Säule geben (Bindung der CBD an die Chitinmatrix)
- Säulen mit 10 Volumen Säulenpuffer waschen
- Induktion der Inteinspaltung: 3x Waschen der Säule mit Spaltungspuffer (DTT-haltig)
- Säule mit Parafilm abdichten und Spaltung üN bei 4 °C inkubieren (optimale Dauer und Temperatur der Spaltung hängt von der endständigen Aminosäure des fusionierten Proteins zum Intein-Tag ab und kann der Anleitung des IMPACT-CN Kits entnommen werden)

-
- Elution des nativen Proteins durch Waschen der Säule mit 2 Volumen Säulenpuffer
 - Aufkonzentrierung des Eluats mit Hilfe eines Proteinfilters (in dieser Arbeit *cut-off*: 10 kDa, Amicon Ultra-4)
 - Waschen der aufkonzentrierten Proteinlösung zweimal mit 4 mL Waschpuffer (ggf. Austausch des Proteinpuffers)
 - Aufgereinigtes Protein aliquotieren mit flüssigem Stickstoff schockgefrieren und bei $-80\text{ }^{\circ}\text{C}$ lagern

13.2. Quantitative Bestimmung des Proteingehaltes

Zur quantitativen Bestimmung des Proteingehaltes wird das Roti[®]-Nanoquant Kit eingesetzt. Diese Methode beruht auf einer Modifikation der Proteinbestimmung nach Bradford (Bradford, 1976). Mit der verwendeten Coomassie-Färbelösung können in wässrigen Lösungen reproduzierbar Proteinmengen ab 200 ng ($c = 1\text{ ng}/\mu\text{L}$, $V = 200\text{ }\mu\text{L}$) bestimmt werden. Hierzu wird jede Probe in einer Verdünnungsreihe bei 590 nm und 450 nm vermessen. Die Linearität ergibt sich aus dem Quotient o.D. 590/450. Die Proteinmenge der Probe wird anhand einer Eichreihe bestimmt.

13.3. Diskontinuierliche SDS-PAGE

Die diskontinuierliche Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) wird verwendet, um Proteine unter denaturierenden Bedingungen nach ihrer molaren Masse aufzutrennen (Laemmli, 1970). Bei einer diskontinuierlichen SDS-PAGE werden die Proben zunächst auf ein großporiges Sammelgel (in dieser Arbeit 5 % Acrylamid) aufgetragen und anschließend im Trenngel (in dieser Arbeit 15 % Acrylamid) augetrennt. Die Durchführung der diskontinuierlichen SDS-PAGE erfolgt gemäß Beschreibung von Laemmli.

Das Gießen eines Gels für die diskontinuierliche SDS-PAGE erfolgt in 2 Schritten. Zuerst wird die Lösung für das Trenngel erstellt. Nach Zugabe des APS muss dieses sofort in den Gießstand überführt werden, da die Auspolymerisierung rasch einsetzt. Um einen glatten Übergang zum Sammelgel zu erhalten und Luftblasen zu vermeiden, wird das Trenngel mit Wasser überschichtet. Das Trenngel sollte ca. 80 % der Laufstrecke der SDS-PAGE in Anspruch nehmen.

Tab. III.11 Zusammensetzung des Trenngels für die diskontinuierliche SDS-PAGE. Das Gel trennt Proteine in einem Bereich zwischen 15 und 60 kDa.

Trenngel (12,5 %)	Volumen
H ₂ O	1,1 mL
Acrylamid (Rotiphorese Gel 30)	2,5 mL
SDS (0,5 %)	1,2 mL
Tris-HCl (1,88 M; pH 8,8)	1,2 mL
TEMED (10 %)	50 µL
APS (10 %)	30 µL

APS immer frisch ansetzen!

Nach der Auspolymerisation des Trenngels (ca. 1 h) kann das Sammelgel erstellt und nach Entfernung des Wassers auf das Trenngel gegossen werden. Sofort hiernach wird der Kamm zur Formung der Ladetaschen eingesetzt.

Tab. III.12 Zusammensetzung des Sammelgels für die diskontinuierliche SDS-PAGE

Sammelgel (5 %)	Volumen
H ₂ O	870 µL
Acrylamid (Rotiphorese Gel 30)	330 µL
SDS (0,5 %)	400 µL
Tris-HCl (0,625 M; pH 6,8)	400 µL
TEMED (10 %)	20 µL
APS (10 %)	10 µL

APS immer frisch ansetzen!

Die Gele können entweder sofort verwendet oder vor Austrocknung geschützt bei 4 °C für ca. 14 Tage gelagert werden. Für die Auftrennung per SDS-PAGE werden die Proben (2 – 8 µg Protein) mit Wasser auf 20 µL aufgefüllt und mit 4 µL Probenpuffer (PBJR) (6x) für 10 min im Thermoblock bei 99 °C gekocht. Die komplette Lösung wird auf das Gel aufgetragen. Als Längenstandard werden 5 µl Proteinmarker PageRule™ Prestained Protein Ladder und PageRuler™ Unstained Protein Ladder aufgetragen. Die Proben werden zunächst bei 15 mA pro Gel für 1 h im Sammelgel gesammelt und danach bei 30 mA pro Gel im Trenngel aufgetrennt. Fertige SDS-Gele werden in Coomassie-Blue R250/G250 über Nacht gefärbt. Überschüssiges Coomassie-Blue bzw. eine zu starke Färbung wird durch 2mal 15 min Schnellentfärbbehandlung entfernt. Verbleibende unerwünschte Hintergrundfärbung der SDS-PAGE wird durch Langsamentfärb abgeschwächt. Die SDS-PAGE kann nun mit einem Durchlichtscanner dokumentiert werden.

13.4. Massenspektrometrische Analyse zur Identifizierung von Proteinen

Zur Identifizierung der Proteine, die in einer SDS-PAGE aufgetrennt wurden, werden die Banden aus dem Gel ausgeschnitten, isoliert und durch Proteasen (in dieser Arbeit Trypsin) in Fragmente unterschiedlicher Größe gespalten. Diese Fragmentgrößen unterscheiden sich von Protein zu Protein und stellen einen charakteristischen Fingerabdruck des Proteins dar. Die Massen der einzelnen Fragmente werden mit Hilfe der Matrix-unterstützten Laser-Desorption/Ionisation-Massenspektrometrie (MALDI-MS) ermittelt. Hierfür werden die Proben in eine spezifische Matrix eingebettet. Der Einbau der Fragmente in das Kristallgitter der Matrix ist eine Voraussetzung für die erfolgreiche Laser-Desorption/Ionisation. Mit einem gepulsten Laser werden Matrix und Peptide ionisiert. Anschließend werden die Ionen in einem elektrischen Feld mit sehr hoher Spannung (einige kV) beschleunigt, woraufhin sie durch den Flugzeitanalysator fliegen. Kleinere Peptide werden dabei auf eine höhere Geschwindigkeit beschleunigt als Größere, d. h. sie durchqueren den Flugzeitanalysator schneller und treffen eher auf den Detektor. Aus der Flugdauer (TOF = *time of flight*) kann nach Abgleich mit einem Standard das Molekulargewicht der Peptide ermittelt werden. Die Gesamtheit der Massen der tryptischen Fragmente eines Proteins ergibt den MS-Fingerabdruck, mit dem das Protein durch Vergleich mit geeigneten Datenbanken identifiziert werden kann.

13.4.1. Tryptischer Verdau

Um die aufgereinigten Proteine (in dieser Arbeit Sigmafaktoren) eindeutig zu identifizieren, werden Proteine im SDS-Gel aufgetrennt (III.13.3) und die entsprechenden Banden aus dem SDS-Gel ausgeschnitten. Als ersten Schritt auf dem Weg zur Identifizierung werden die im Gel enthaltenen Proteine einem tryptischen Verdau unterzogen. Im Rahmen des tryptischen Verdau wird das in der Gelbande enthaltene Protein durch Einwirkung der Protease Trypsin in Fragmente charakteristischer Größe gespalten:

- TFA-gewaschene Reaktionsgefäße mit je 500 µL TDL1 waschen; Lösung vollständig entfernen; unter dem Abzug üN abdampfen lassen
- Coomassie-gefärbte Gelbande mit sauberen Skalpell so eng wie möglich ausschneiden, in gewaschenes Reaktionsgefäß überführen und zerkleinern
- 100 µL TDL2 zugeben; 5 min bei RT schütteln, Waschlösung entfernen
- 100 µL TDL3 zugeben; 30 min bei RT schütteln; Waschlösung entfernen
- 100 µL TDL4 zugeben; 30 min bei RT schütteln; Waschlösung entfernen
- Gelstückchen bei RT für ca. 30 - 60 min in der Vakuumtrocknungsanlage SpeedVac vollständig trocknen
- 15 µL Trypsinlösung zugeben; für 10 min bei RT inkubieren
- 20 µL TDL5 zugeben; 24 h bei 37 °C inkubieren

Tryptisch verdaute Proben können sofort zur MALDI-TOF-Analyse einsetzen werden oder bei -20 °C gelagert werden.

13.4.2. MALDI-TOF Analyse

Die durch Trypsinverdau entstehenden charakteristischen Fragmente des Proteins werden anschließend mittels MALDI-TOF-Massenspektrometrie im Ultraflex Massenspektrometer und der MASCOT-Software (Perkins *et al.*, 1999) analysiert (Hansmeier *et al.*, 2006). Dabei werden die ermittelten Massen der Peptidfragmente des zu analysierenden Proteins mit den theoretisch errechneten Fragmenten von Proteinen aus verschiedenen Datenbanken (z. B. *C. glutamicum*-Genomprojekt oder dem öffentlich zugänglichen TREMBL) verglichen. Aus dem Grad der Übereinstimmung wird die Wahrscheinlichkeit bestimmt, mit der das Analytprotein identifiziert wurde. Ein Suchergebnis gilt als signifikant, wenn die Wahrscheinlichkeit einer zufälligen Übereinstimmung unter 0,05 % liegt. Dies entspricht einem Treffer von 50 oder mehr.

Zur Auswertung der Massenspektren mit *Mascot* werden folgende Parameter verwendet:

- *type of digest*: Trypsin
- *missed cleavages*: 0 - 2
- *peptide tolerance*: 70 - 500 ppm
- *variable modifications*: Oxidation (M)

14. Sequenzanalyse

14.1. Motivsuche und Bewertung der Motive in potentiellen Promotorregionen

Für die Identifizierung von Motiven in DNA-Sequenzen mittels PSSMs (*position specific scoring matrix*) gibt es zwei Basismethoden. Zum einen den EM-Algorithmus (expectation maximization (Lawrence and Reilly, 1990)) und zum anderen das Gibbs *sampling* (Lawrence et al., 1993). Beide Methoden starten mit einem initialen *Alignment* von Subsequenzen, welche iterativ auf Basis einer spezifischen Bewertungsfunktion bis ein lokales Optimum erreicht ist verfeinert wird. In dieser Arbeit wurden beide Methoden in Form der Programme BioProspector (Liu et al., 2001) und Improbizer (Ao et al., 2004); www.soe.ucsc.edu/~kent/improbizer/improbizer.html) zur Motivsuche und Bewertung identifizierter Motive in stromaufwärts liegenden Sequenzabschnitten potentiell ECF-Sigmafaktor-regulierter Gene genutzt.

Das C-Programm BioProspector verwendet Gibbs *sampling*. BioProspector nutzt hierbei nul bis dritte Ordnung Markov-Hintergrund-Modelle, die durch den initiale Datensatz an Sequenzen bestimmt werden. Die Signifikanz jedes identifizierten Motivs wird auf Basis der Werteverteilung durch die Monte Carlo Methode abgeschätzt.

Das Programm Improbizer sucht in einem Set von Sequenzen auch nach Konsensusmotiven, die mit einer unwahrscheinlichen Regelmäßigkeit auftreten, nutzt hierbei aber eine Abwandlung des EM-Algorithmus. Der Hintergrundsignifikanzwert für die Bestimmung validierter Promotorelemente (Motive) wurde durch die 5-fache Verwendung eines Kontrolllaufs ermittelt. Hier wird eine zufällige Verteilung zweier Motive auf Basis des eingegebenen Datensatzes simuliert. Ein Motiv gilt als validiert, wenn die in den übrigen Sequenzen enthaltenen Motive in dieser Sequenz enthalten waren und die entsprechende Bewertung über den ermittelten Hintergrundsignifikanzwert lag.

IV. Ergebnisse

1. Transkriptionelle Regulation des *sigH-rshA*-Operons und die Kontrolle des SigH-Regulationsnetzwerkes in *Corynebacterium glutamicum*

1.1. Das *sigH*-Gen und das *rshA*-Gen, welches für den Anti-Sigmafaktor kodiert, sind als Operon angeordnet

Die Gene, die für SigR (ortholog zu *C. glutamicum* SigH) in *S. coelicolor* und für SigH in einigen Mycobakterien (zum Beispiel *M. avium*, *M. smegmatis* und *M. tuberculosis*) kodieren, sind in unmittelbarer Nähe zu den Genen lokalisiert, die für die entsprechenden Anti-Sigmafaktoren RsrA bzw. RshA kodieren. Die Anti-Sigmafaktoren konnten direkt *downstream* der Sigmafaktoren im Genom identifiziert werden (Kim *et al.*, 2009; Song *et al.*, 2003). Dieselbe Anordnung wurde auch für die Gene *sigH* (*cg0876*) und *rshA* (*cg0877*) in den Genomen von *C. glutamicum* ATCC 13032 (Kalinowski *et al.*, 2003) und *C. jeikeium* (Tauch *et al.*, 2005) beschrieben (Abb. IV-2 A). Aufgrund seiner geringen Größe von 267 Nukleotiden (89 Aminosäuren), wurde das Gen *rshA* in zwei *C. glutamicum* Genomsequenzen, unabhängig publiziert von Ikeda *et al.* und Yukawa *et al.* (Stamm R) (Ikeda und Nakagawa, 2003; Yukawa *et al.*, 2007), nicht automatisch erkannt. Durch eine BLASTX-Suche konnte *rshA* aber in diesen Genomen identifiziert werden (Daten nicht gezeigt). In diesem Zusammenhang ist es interessant anzumerken, dass das Fehlen des *rshA*-Gens in der Annotation einer der *C. glutamicum* Genomsequenzen (Ikeda und Nakagawa, 2003) die Autoren einer aktuellen Studie fehlgeleitet hat (Jung *et al.*, 2011). In dieser Studie haben die Autoren das falsche, nicht orthologe Gen von *C. glutamicum* verwendet, um eine funktionale Komplementation von *rsrA* in *Streptomyces* zu überprüfen.

In allen *C. glutamicum* Genomen ist das Stopcodon des *sigH*-Gens vom Startcodon des *rshA*-Gens nur durch zwei Basenpaare getrennt. Diese Anordnung hat Ähnlichkeit mit einer klassischen Operonstruktur. Die abgeleitete RshA-Proteinsequenz besitzt nur mäßige Ähnlichkeit zu RshA aus *M. tuberculosis* (35 %) bzw. zu RsrA aus *S. coelicolor* (28 %). Ein Aminosäuresequenzabgleich zwischen den Proteinen der vier Corynebakterienspezies *C. glutamicum*, *Corynebacterium efficiens*, *Corynebacterium diphtheriae* und *Corynebacterium jeikeium* und deren Orthologen aus *M. tuberculosis* und aus *S. coelicolor* (Abb. IV-1) zeigt aber,

dass RshA aus *C. glutamicum* und anderen Corynebakterien die konservierten Cysteine, die die Interaktion von SigH und RshA in der ZAS-Domäne vermitteln (Jung *et al.*, 2011), besitzt.

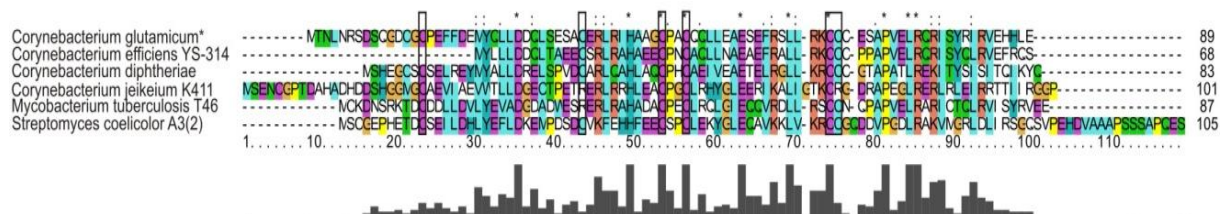


Abb. IV-1 Alignment der Aminosäuresequenz zwischen vier corynebakteriellen SigH-Anti-Sigmafaktoren und deren Orthologen aus *M. tuberculosis* und aus *S. coelicolor*. Im Detail: Alignment von RshA von *C. glutamicum*, *C. efficiens*, *C. diphtheriae* und *C. jeikeium*, wie auch *M. tuberculosis* und RsrA aus *S. coelicolor* durch CLUSTALX (Larkin *et al.*, 2007). Die konservierten Cysteine sind mit einer Box umgeben. Identische Aminosäurereste sind mit einem Stern markiert; "." zeigt einen hohen Grad der Konservierung an, während "." einen schwächeren Grad der Konservierung anzeigt.

Wie oben beschrieben sind die Gene *sigH* und *rshA* in *C. glutamicum* aufgrund des geringen Abstands im Genom vermutlich in einem Operon arrangiert. Um deren transkriptionelle Organisation aufzuklären, wurden Northern-Hybridisierungen durchgeführt. Die Gelaufftrennung wurde mit Gesamt-RNA aus *C. glutamicum* RES167 (restriktionsdefiziente Variante des ATCC 13032 Typstammes) und den darauf basierenden Deletionsmutanten DN2 (Deletion innerhalb des *sigH*-Gens) und AS1 (komplette Deletion der Genregion *sigH-rshA*) durchgeführt. Nach dem *Blotting* wurde die Membran mit DIG-markierten RNA-Sonden spezifisch für das Gen *sigH* bzw. das Gen *rshA* hybridisiert. Für die RES167 Gesamt-RNA konnte ein einziges, ca. 1 kb großes, Transkript mit der *sigH*-spezifischen RNA-Sonde nachgewiesen werden (Abb. IV-2 B). Ein Transkript mit identischer Länge hybridisierte auch mit der *rshA*-Sonde. Dies zeigt, dass beide Gene in Form einer einzelnen mRNA von einem Promoter, stromaufwärts von *sigH* gelegen, transkribiert werden. Bei der Nutzung der *rshA*-Sonde wurde zusätzlich ein weiteres Transkript detektiert. Dieses hat eine Größe von ungefähr 370 bp, was darauf hindeutet, dass es sich hier um eine mRNA handelt, die nur das *rshA*-Gen beinhaltet und von einem Promotor (P_{rshA}), innerhalb der Kodierregion von *sigH* gelegen, transkribiert wird.

Um die auftretende Frage zu beantworten, ob die Promotoren der Gene *sigH* und *rshA* von dem Sigmafaktor SigH kontrolliert werden, wurde für weitere Northern-Hybridisierungen die Gesamt-RNA des *sigH*-Deletionsstammes DN2 verwendet. Die Annahme war, dass SigH-abhängige Transkripte in der Gesamt-RNA von DN2 nicht vorkommen. Entsprechend wurde kein Signal bei der Verwendung der *sigH*-Sonde detektiert, da der komplementäre Bereich in DN2 deletiert ist. Dies entspricht folglich der Negativkontrolle. Allerdings wurde durch den

Einsatz der *rshA*-spezifischen Sonde ein ca. 550 bp großes Transkript detektiert (Abb. IV-2 B). Die Transkription dieser mRNA wird am wahrscheinlichsten stromaufwärts von *sigH* durch einen *sigH*-Promoter initiiert, da die Länge des Transkriptes der Länge des *sigH-rshA*-Transkriptes, abzüglich der Länge der Deletion im *sigH*-Gen des Stammes DN2, entspricht (Abb. IV-2 A). Dieses Ergebniss zeigt, dass das bi-cistronische *sigH-rshA*-Transkript SigH-unabhängig transkribiert wird. Im Gegensatz dazu konnte das separat transkribierte reine *rshA*-Transkript mit der dazugehörigen RNA-Sonde nicht identifiziert werden, obwohl die Deletion innerhalb der Kodierregion von *sigH* nicht den mutmaßlichen *rshA*-Promotor umfasst.

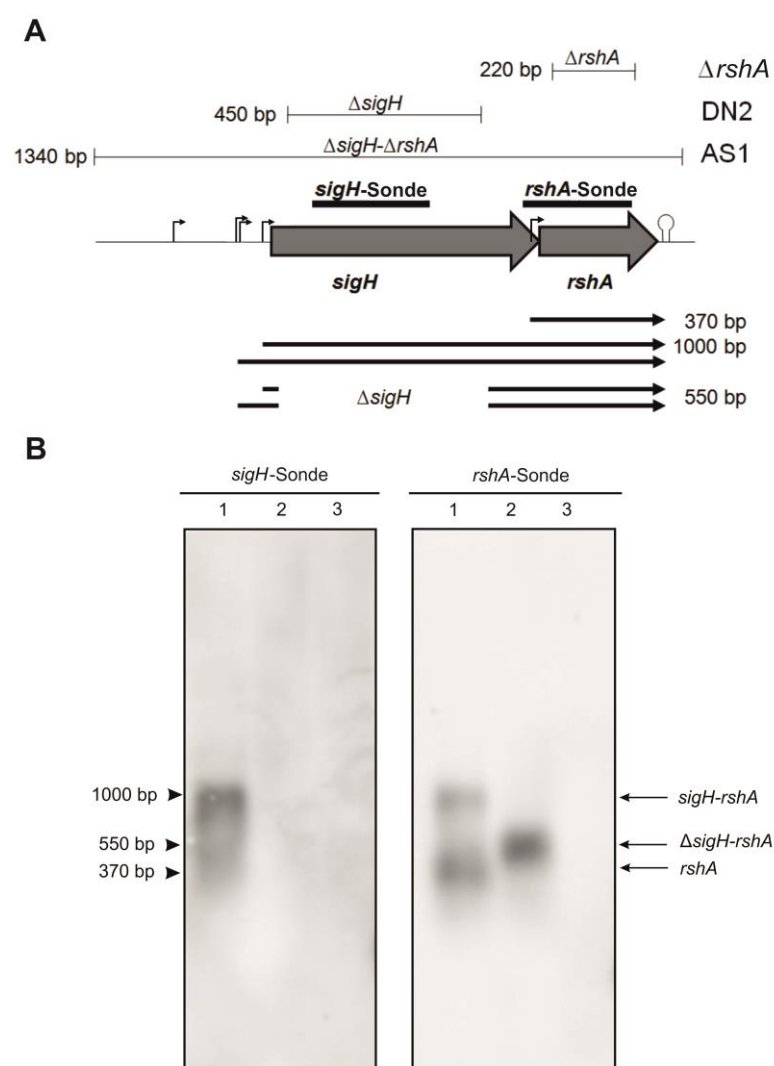


Abb. IV-2 Genkarte des *sigH-rshA*-Operons und dessen Analyse durch Northern-Hybridisierung in *C. glutamicum* RES167 und Deletionsmutanten. a. Genkarte der *sigH-rshA*-Region mit Kennzeichnung der genauen Position und Größe der Deletionen im Chromosom der Stämme *C. glutamicum* $\Delta rshA$, DN2 ($\Delta sigH$) and AS1 ($\Delta sigH-rshA$), vorhergesagte Größen der entsprechenden *sigH-rshA*- und *rshA*-Transkripte (Pfeile) und Position der Sonden für die Northern-Hybridisierung. Promotoren sind mit geneigten Pfeilen und der Terminator als Haarnadelstruktur-Symbol gekennzeichnet. b. Northern Blot: *sigH*-Sonde (linker Blot) und *rshA*-Sonde (rechter Blot) hybridisiert mit Gesamt-RNA isoliert von: RES167 Zellen (Spur 1); DN2 Zellen ($\Delta sigH$ Deletion; Spur 2); AS1 Zellen ($\Delta sigH-rshA$ Deletion; Spur 3). Die abgeschätzten Längen der detektierten Transkripte (links) und deren Bezeichnung (rechts) sind dargestellt.

1.2. Die Gene des *sigH-rshA*-Operons werden von mehreren unterschiedlichen Promotoren aus transkribiert

Um die Promotorregionen des *sigH-rshA*-Operons und des einzelnen *rshA*-Gens zu untersuchen, wurden durch die Gruppe um Miroslav Patek folgende Untersuchungen durchgeführt. Die potentiell die Promotoren beinhaltenden DNA-Fragmente (504 bp stromaufwärts von *sigH* und 301 bp stromaufwärts von *rshA*) wurden in den Promotor-Test-Vektor pET2 kloniert und so eine transkriptionelle Fusion zwischen den Promotor-tragenden Fragmenten und dem Reporter gen *cat* erstellt. Das Gen *cat* kodiert für die Chloramphenicol-acyltransferase (CAT). Die Aktivität der Promotoren wurde durch die Bestimmung der CAT-Enzymaktivität in zellfreien Extrakten von *C. glutamicum* pET2::*sigH* und *C. glutamicum* pET2::*rshA* gemessen. Die Aktivität von P_{sigH} belief sich während der exponentiellen Wachstumsphase auf $0,1 \pm 0,015$ U (mg Protein)⁻¹, wohingegen die Aktivität von P_{rshA} nur bei $0,03 \pm 0,005$ U (mg Protein)⁻¹ lag. Die Negativkontrolle mit dem Leervektor pET2 wies einen zu vernachlässigenden Grundumsatz von $\leq 0,003$ U (mg Protein)⁻¹ auf. Diese Messungen bestätigten, dass *rshA* von einem separaten Promotor (P_{rshA}) erkannt und transkribiert wird, der während des exponentiellen Wachstums eine geringere Aktivität besitzt als der Promotor, der das gesamte Operon transkribiert.

Um die Transkriptionsstartpunkte (TSP) der *sigH-rshA*- und *rshA*-Transkripte zu bestimmen, wurden von der Gruppe um Miroslav Patek *Primer-Extension-Analysen* (PEX) durchgeführt. Hierfür wurde der Primer CM4 und Gesamt-RNA aus *C. glutamicum* pET2::*sigH* bzw. *C. glutamicum* pET2::*rshA* verwendet. Durch diese Analysen konnten drei TSP stromaufwärts von *sigH* identifiziert werden. TSP1, TSP2 und TSP3 wurden jeweils auf das Nukleotid A gemappt und befinden sich 22 bp, 89 bp und 93 bp stromaufwärts des Startcodons von *sigH* (Abb. IV-3 A). Dieses Ergebnis konnte unter Verwendung des Primers CM5 bestätigt werden (Daten nicht gezeigt). Die putativen -10-Regionen der entsprechenden Promotoren, TAGAAT (P1), TAAAGT (P2) und TAGAGT (P3), sind zum einen untereinander ähnlich und besitzen zum anderen eine hohe Übereinstimmung mit der Konsensussequenz der -10-Region (TANANT) SigA-abhängiger Promotoren (Patek und Nesvera, 2011). SigA-abhängige Promotoren sind für die Expression von *Housekeeping*-Genen in *C. glutamicum* verantwortlich. Die putativen -35-Regionen der Promotoren P1, P2 und P3 besitzen weniger Ähnlichkeit zu der Konsensussequenz SigA-abhängiger Promotoren. Diese Tatsache ist aber eine allgemeine Eigenschaft der *Housekeeping*-Promotoren in *C. glutamicum*. Zusammenfassend konnten diese drei Promotoren als SigA-abhängig klassifiziert werden.

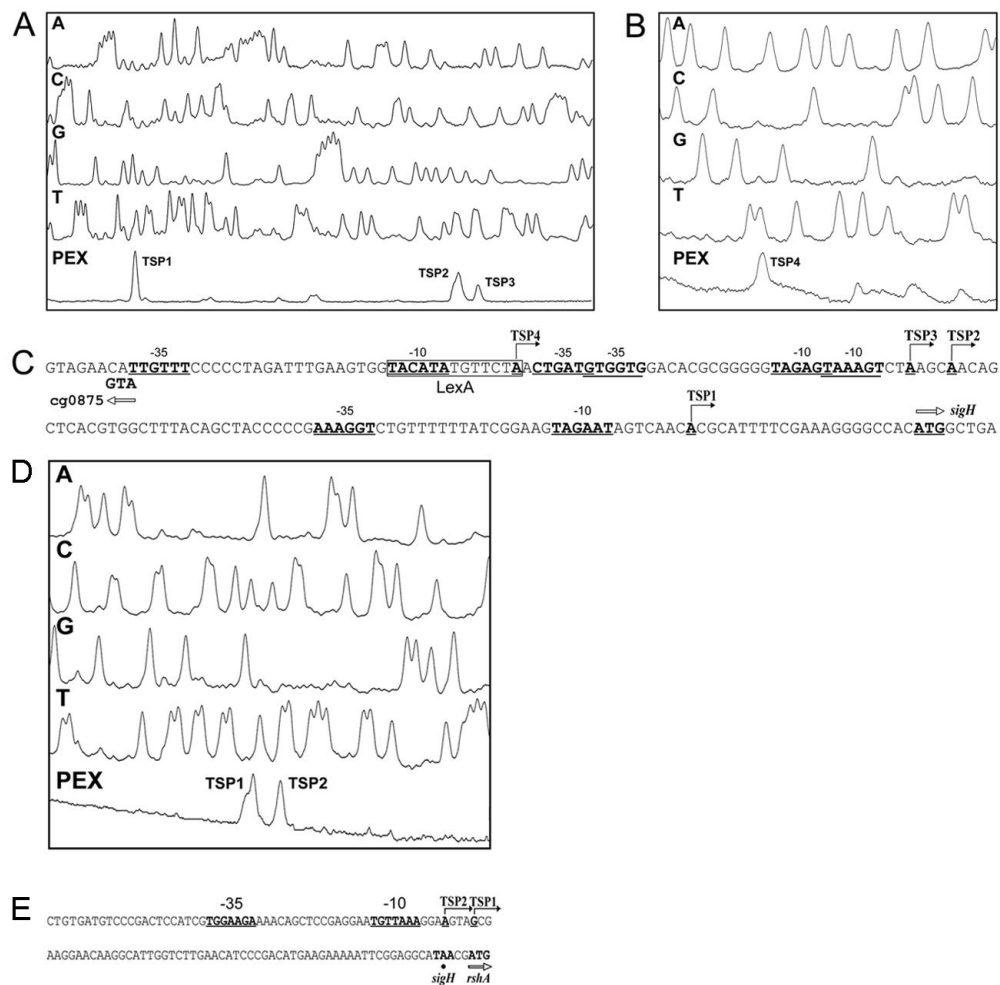


Abb. IV-3 Identifikation der Transkriptionsstarts von *sigH* und *rshA* und Sequenz der Promotoren. (a) und (b) Identifikation des *sigH*-Transkriptionsstarts (TSP) durch *Primer-Extension-Analyse* (PEX). Die unteren *Peaks* (PEX) repräsentieren cDNA die in einer reversen Transkription aus Gesamt-RNA von *C. glutamicum* (pET2::*sigH*) bzw. *C. glutamicum* (pET2::*sigH4*) hergestellt wurde. Die kleineren *Peaks* waren in wiederholten Experimenten nicht reproduzierbar. Die *Peaks* (A, C, G, T) repräsentieren die Produkte der Sequenzierreaktion, die mit demselben Fluoreszenz-markierten Primer durchgeführt wurde, der für die reverse Transkription eingesetzt wurde. (c) Nukleotidsequenz der stromaufwärts liegenden Region von *sigH*. TSP und die entsprechenden -35 - und -10 -Promotorelemente sind Fett geschrieben und unterstrichen. Die Transkriptionsinitiation ist mit einem gekrümmten Pfeil markiert. Die Bindestelle des LexA-Regulators (SOS-Box) ist mit einer Box umgeben. Die Startcodons der Gene *sigH* und *cg0875* sind Fett geschrieben und mit einem nicht gefüllten Pfeil gekennzeichnet. (d) Identifikation des *rshA*-TSP. (e) Nukleotidsequenz der stromaufwärts liegenden Region von *rshA*. Das Stoppcodon (Fett) des *sigH*-Gens ist mit einem schwarzen Punkt markiert. Wichtig: die gezeigten Sequenzen in (c) und (e) sind komplementär und *reverse* zu den Sequenzen, die von den *Peaks* der Sequenzierung abgeleitet wurden.

Da in einigen PEX-Analysen von *C. glutamicum* pET2::*sigH* ein weiterer TSP stromaufwärts des TSP3 identifiziert werden konnte, wurden erneut CAT-Enzymaktivitäten und PEX-Analysen durchgeführt. Hierfür wurde ein 348 bp großes stromaufwärts des *sigH* Startcodons

(Position 462 bp bis 115 bp), außerhalb der Promotoren P1, P2 und P3 separat in den Vektor pET2 kloniert (pET2::*sigH4*). Die Analyse dieser transkriptionellen Fusion ergab eine CAT-Aktivität von $0,009 \pm 0,002$ U (mg Protein)⁻¹. Dieser geringe Wert deutet auf eine sehr schwache Promoteraktivität während der exponentiellen Wachstumsphase hin. Durch die PEX-Analyse mit Gesamt-RNA von *C. glutamicum* pET2::*sigH4*, unter Verwendung der Primer CM4 und CM5, konnte ein TSP 131 bp stromaufwärts des *sigH*-Startcodons identifiziert werden (TSP4; Abb. IV-3 B). Dieser TSP konnte zusätzlich durch eine RACE (*Rapid Amplification of C-terminal Ends*)-Analyse bestätigt werden (Daten nicht gezeigt). Die Hexamere TACATA und TTGTTT (mit einem Abstand von 19 bp) im zu erwartenden Abstand zum TSP4 könnten als weiterer SigA-abhängiger Promotor (P4) fungieren (Abb. IV-3 C). Die Sequenz TGGTACATATGTTCTA überlappt mit der -10-Region des Promotors P4. Sie entspricht der Konsensussequenz der SOS-Box, die als Bindestelle für LexA in *C. glutamicum* beschrieben ist (Jochmann *et al.*, 2009).

Um den oder die TSP von *rshA* zu bestimmen, wurde die Gesamt-RNA von *C. glutamicum* pET2::*rshA* zusammen mit dem Primer CM4 in einer PEX-Analyse verwendet. Es konnten zwei TSP bestimmt werden. TSP1 wurde mit dem Nukleotid G 62 bp und TSP2 mit dem Nukleotid A 66 bp stromaufwärts des *rshA*-Startcodons identifiziert (Abb. IV-3 D). TSP1 konnte unter Verwendung des Primers CM5 bestätigt werden, aber weist ein deutlich schwächeres Signal auf (Daten nicht gezeigt). Die Motife TGGGAAGA in der -35-Region und TGTTAAA in der -10-Region relativ zum TSP1 passen sehr gut zu den -35- und -10-Regionen der SigH-abhängigen Promotoren von *M. tuberculosis* (^G/TGGAA^C/TA -16 nt -^C/GGTT) (Raman *et al.*, 2001b), zu den SigR-abhängigen Promotoren von *S. coelicolor* (GGGAAT^G/C - 16 nt -^C/GGTTG) (Paget *et al.*, 2001) und auch zu der publizierten *C. glutamicum* Konsensussequenz von SigH-abhängigen Promotoren gGGAA^ata - 16–19 nt -^C/TGTTgaa (Ehira *et al.*, 2009) bzw. ^G/TGGAA^aTA - 16–19 nt -^C/TGTTGAA (Patek und Nesvera, 2011). Diese Ergebnisse lassen vermuten, dass der P_{*rshA*} Promotor der Kontrolle von SigH unterliegt, was wiederum die Ergebnisse der Northern-Analysen bestätigt.

1.3. Die Analyse des Transkriptionsprofil der Deletionsmutante RES167 $\Delta rshA$ identifiziert die meisten bekannten SigH-abhängigen Gene und neue Targetgene

Um die Gene aufzudecken, die unter der Kontrolle von SigH stehen, wurde der hierfür konstruierte Deletionsstamm *C. glutamicum* RES167 $\Delta rshA$ verwendet (Abb. IV-4 A). Die Deletion $\Delta rshA$ im Genom von *C. glutamicum* RES167 wurde unter Verwendung des pK18*mobsacB*-Vektorsystems etabliert. Die Deletion wurde mittels Test-PCR bestätigt. Das PCR-Produkt im Referenzstamm war 2.500 bp groß, das in dem Deletionsstamm $\Delta rshA$ war mit 2.280 bp rund 220 bp (Größe der Deletion) kleiner (Abb. IV-4 B).

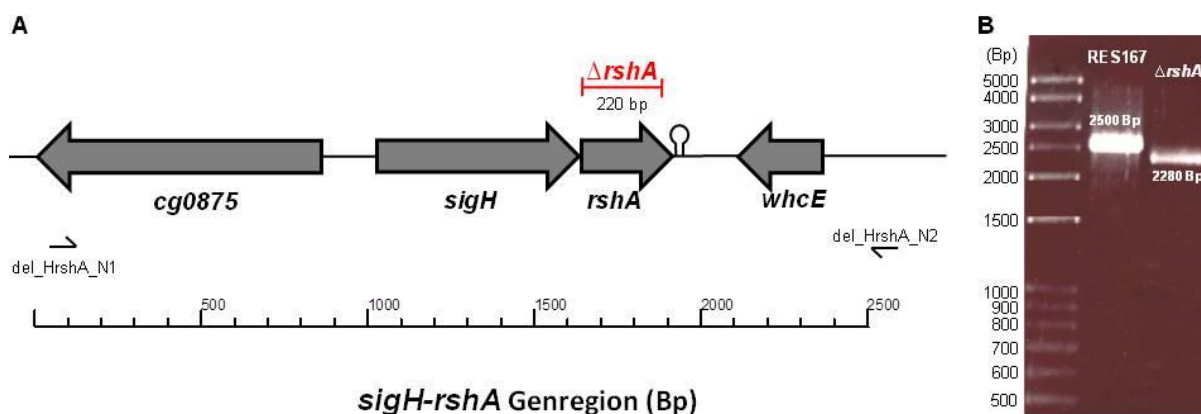


Abb. IV-4 Genregion *sigH-rshA*, inklusive Kartierung der $\Delta rshA$ -Deletion und entsprechendem Deletionsnachweis per Test-PCR. (a) Genregion von *sigH-rshA* mit deletiertem Bereich im $\Delta rshA$ -Stamm. Die zum Nachweis der Deletion verwendeten Primer sind eingezeichnet. Die Position des Transkriptionsterminators ist durch ein Schlaufen-Symbol dargestellt. (b) Deletionsnachweis mittels Test-PCR. Das PCR-Produkt der $\Delta rshA$ -Deletionsmutante ist um 220 bp kleiner als das Produkt des Wildtyps (Marker 100 bp-Leiter, erweitert bis 5.000 bp, MBBL, Bielefeld, Deutschland).

Es wurde erwartet, dass SigH in diesem Stamm konstitutiv aktiv ist und die SigH-abhängigen Gene ohne Stressapplizierung transkribiert werden, da SigH nicht mehr durch den Anti-Sigmafaktor RshA gebunden und dadurch inaktiviert wird. Hierzu wurde eine vergleichende *Microarray*-Analyse durchgeführt. Die Hybridisierung der Gesamtgenom-*Microarrays* von *C. glutamicum* wurde mit cDNA durchgeführt. Als *Template* für die Herstellung der cDNA diente Gesamt-RNA, die aus *C. glutamicum* RES167 und dem entsprechenden $\Delta rshA$ -Deletionsstamm isoliert wurden. Die Kultivierung erfolgte unter Standardbedingungen im Luftschtler bei 300 U/min und 30 °C in CGXII-Minimalmedium. Der m/a-Plot der Hybridisierungen ist in Abb. IV-5 gezeigt. Die dazugehörigen differentiell transkribierten Gene mit einer mindestens zweifach erhöhten Transkriptmenge oder Zugehörigkeit einer der

aufgeführten Funktionsklassen sind in **Tab. IV.1** aufgelistet und nach ihrer Funktion sortiert. Insgesamt weisen 83 Gene in 61 putativen transkriptionellen Einheiten im Vergleich zum Referenzstamm einen vergrößerten mRNA-Pool in der $\Delta rshA$ -Mutante auf (Anhang Tab. 1). Die größten Änderungen der relativen mRNA-Menge wurde für Gene beobachtet, die bereits als SigH-abhängig beschrieben worden waren (Ehira *et al.*, 2009). Diese Daten bestätigen somit die Annahme, dass der Sigmafaktor SigH in der $\Delta rshA$ -Deletionsmutante aktiv vorliegt und durch Bindung an das RNA-Polymerase-Core-Enzym (RNAP-Core Enzym) die Transkription SigH-abhängiger Gene ermöglicht. Grund für die Aktivität ist das Fehlen des funktionalen *rshA*-Genprodukts in der Mutante. Hierdurch wird gezeigt, dass RshA in *C. glutamicum* der verantwortliche Anti-Sigmafaktor ist, der die Aktivität von SigH *in vivo* kontrolliert.

Obwohl die meisten der differentiell transkribierten Gene bereits in früheren Arbeiten als SigH-abhängig beschrieben worden sind, konnten in dieser Arbeit zusätzliche Gene, wie *mshC* (*cg1709*, Mycothiolbiosynthese) und *mca* (*cg1127*, Mycothiol-Konjugat Amidase) als Gene mit einem stark erhöhten mRNA-Pool und *qor2* (*cg1553*, Quinon-Oxidoreduktase) als Gen mit einem gering erhöhten mRNA-Pool in der $\Delta rshA$ -Mutante identifiziert werden. Die genannten Gene sind alle offensichtlich an der Aufrechterhaltung der Redox-Homöostase beteiligt. Dazu passend wurden diese Gene als die am stärksten transkribierten Gene während Diamid-induziertem Disulfidstress identifiziert (Daten nicht gezeigt).

Interessanterweise wiesen einige Hitzestress-zugehörige Gene, welche bereits als SigH-abhängig beschrieben worden sind (*dnaJ2*, *clpB*, *clpP1* und *clpP2*; Ehira *et al.*, 2009; Engels *et al.*, 2004), einen nur leicht erhöhten mRNA-Pool im Deletionsstamm $\Delta rshA$ auf. Einige der Gene, die als Mitglieder des SigH-Regulons beschrieben worden sind, überschritten den gewählten Grenzwert (m-Wert von 0,6 bzw. 1,5-fache Änderung) nicht. Zu diesen nicht-signifikant differentiell transkribierten Genen zählen das *dnaK-grpE*-Operon, *clpC*, *sigB* (nicht-essentieller Sigmafaktor) und die meisten Gene des *suf*-Clusters (Ehira *et al.*, 2009). Weiter konnte keine differentielle mRNA-Menge von *clgR*, einem Regulator der Hitzestressantwort, beobachtet werden, obwohl dieser von einem SigH-abhängigen Promotor erkannt und transkribiert wird (Engels *et al.*, 2004). Auch Ehira *et al.* konnten keine veränderte Transkription des Gens *clgR* in ihren Deletionsstudien beobachten (Ehira *et al.*, 2009). Dieser Widerspruch lässt sich vermutlich mit der negativen Kontrolle durch zusätzliche Regulationssysteme in Abwesenheit eines spezifischen Hitzestresses erklären.

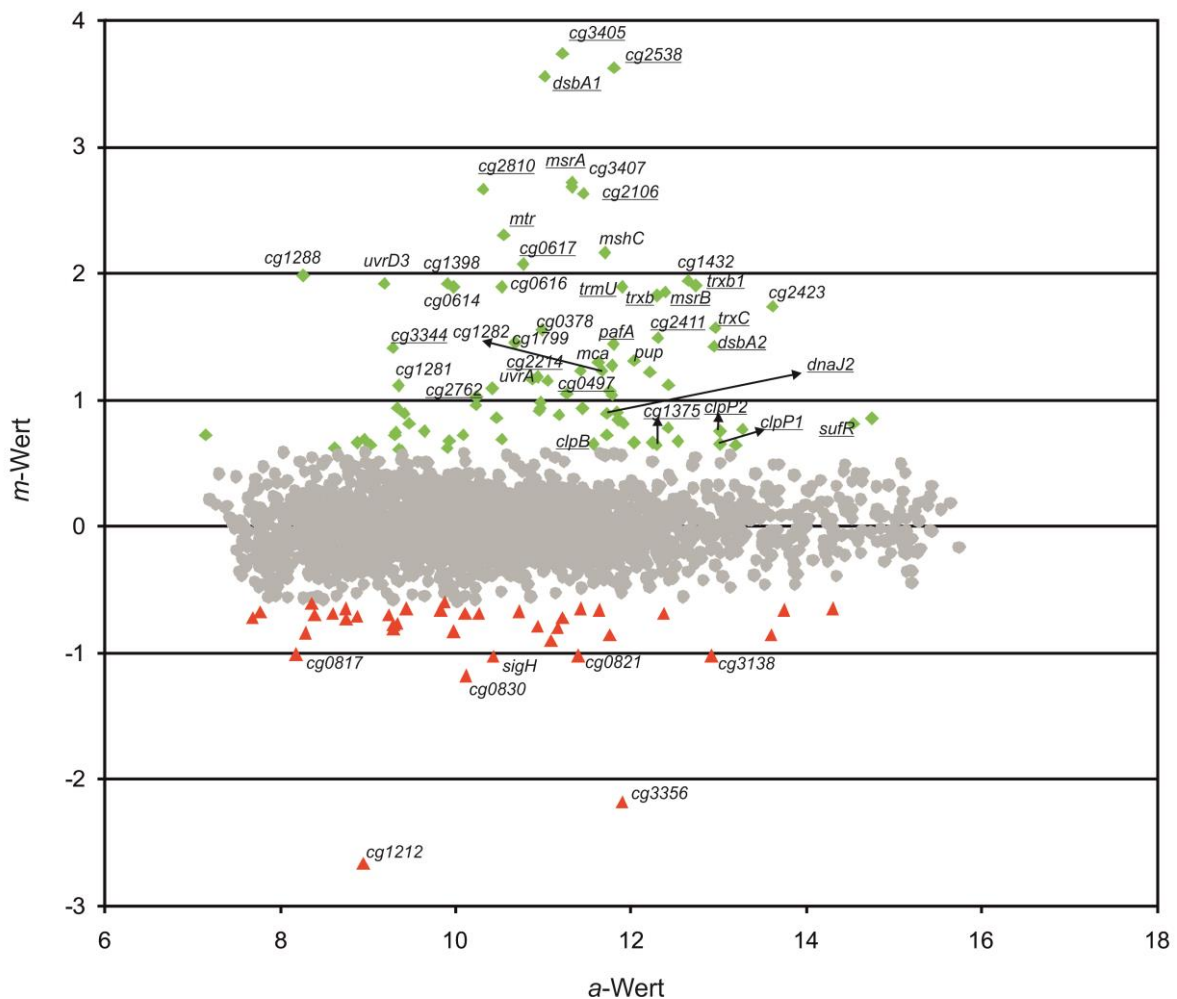


Abb. IV-5 Microarray-Analyse von *C. glutamicum* RES167 im Vergleich mit der $\Delta rshA$ -Deletionsmutante. Darstellung der *Microarray*-Hybridisierung $\Delta rshA$ gegen RES167 als m/a-Plot (m-Wert = $\log_2(\text{Signalintensitäts-Verhältnis})$; a-Wert = mittlere Signalintensität). Für die Hybridisierung wurde Gesamt-RNA von je zwei biologischen Replikaten isoliert, die bis zur exponentiellen Phase in CGXII angezogen wurden. Gene mit erhöhter mRNA-Menge im $\Delta rshA$ -Deletionsstamm haben einen positiven m-Wert, während Gene, die eine höhere mRNA-Menge in der Referenz RES167 aufweisen, einen negativen m-Wert besitzen; grüne Rauten = m-RNA-Menge erhöht in $\Delta rshA$; rote Dreiecke = erniedrigte m-RNA-Menge in $\Delta rshA$; graue Punkte = Gene mit keinem differentiellen mRNA-Pool. M-Werte höher als +0,6 oder kleiner als -0,6 (entspricht einem Mengenunterschied von 1,52-fach bzw. 0,66-fach) gelten als signifikant. Die Namen bzw. *cg*-Nummern relevanter Gene sind im m/a-Plot eingetragen, unterstrichene Gene wurden bereits als SigH-abhängig beschrieben.

Zu den Genen, die hier erstmalig als Mitglieder des SigH-RshA-Regulationsnetzwerkes beschrieben wurden, zählen auch *uvrA* (*cg1560*) und *uvrC* (*cg1790*), beides Untereinheiten der Exonuklease ABC (DNA-Reperatursystem). Weiterhin wurde *uvrD3* (*cg1555*) als dazugehörig identifiziert, eines von drei Genen, die DNA-Helikasen kodieren und Ähnlichkeit zu den UvrD-Proteinen aus *C. glutamicum* besitzen. Ein weiterer neu identifizierter Bestandteil

des Regulationsnetzwerkes ist ein Gencluster (*cg0184-cg0186*), das laut automatischer Annotaton eventuell an der Reparatur alkylierter DNA beteiligt ist. Zusammen mit der Identifizierung eines putativ LexA-regulierten Promotors stromaufwärts des *sigH-rshA*-Operons, lässt sich damit zum ersten Mal eine Verbindung zwischen DNA-Schäden bzw. deren Reparatur und dem SigH-Regulationsnetzwerk herstellen.

Andere neu identifizierte Gene kodieren für Komponenten der Proteasom-Maschinerie *pup* (*cg1689*; kodiert für ein prokariotisches Ubiquitin-ähnliches Protein) und *cg0998* (eine Trypsin-ähnliche Serinprotease). Alle genannten Gene wiesen in dem Δ *rshA*-Deletionsstamm einen vergrößerten mRNA-Pool im Vergleich zum Referenzstamm RES167 auf (Tab. IV.1).

Tab. IV.1 Gene mit erhöhtem mRNA-Pool in *C. glutamicum* Δ *rshA* verglichen mit *C. glutamicum* RES167 (Referenz), nach Funktion sortiert.*

CDS ^a	Gen	Vorhergesagte Funktion	Änderung ^b
Disulfid-Stress Gene			
<u><i>cg3405</i></u> ^{◇*}		NADPH:quinone Reduktase	13,3
<u><i>cg2538</i></u>		Alkanal Monooxygenase (FMN-zugehörig)	12,3
<u><i>cg2838</i></u>	<i>dsbA1</i>	Putative Dithiol-Disulfide Isomerase	11,7
<u><i>cg3236</i></u>	<i>msrA</i>	Protein-Methionin-S-oxid Reduktase	6,6
<u><i>cg2194</i></u>	<i>mtr</i>	Putative NADPH-abhängige Mycothiol Reduktase	4,9
<u><i>cg1709</i></u> ^{-*}	<i>mshC</i>	Putative 1-D-Myo-Inositol-2-Amino-2-Deoxy-Alpha-D-Glucopyranosid-L-cystein Ligase	4,5
<u><i>cg3299</i></u>	<i>trxB1</i>	Thioredoxin (TRX)	3,7
<u><i>cg2078</i></u>	<i>msrB</i>	Peptide Methionin Sulfoxide Reduktase	3,6
<u><i>cg3422</i></u> ^{○*}	<i>trxB</i>	Thioredoxin Reduktase	3,5
<u><i>cg3423</i></u> [○]	<i>trxC</i>	Thioredoxin	3,0
<u><i>cg2661</i></u>	<i>dsbA2</i>	Putative Dithiol-Disulfid Isomerase	2,7
<u><i>cg3344</i></u>		Putative Nitroreduktase	2,7
<u><i>cg1127</i></u> ^{-*}	<i>mca</i>	Putative Mycothiol S-Konjugate Amidase	2,5
<u><i>cg2214</i></u>		Putative Fe-S-Cluster Redox Enzym	2,3
<u><i>cg0497</i></u> ^{□*}	<i>hemA</i>	Glutamyl-tRNA Reduktase	2,1
<u><i>cg1765</i></u>	<i>sufR</i>	Transkriptioneller Repressor des Suf-Operon	1,8
<u><i>cg1553</i></u>	<i>qor2</i>	Quinon Oxidorreduktase involviert in Disulfid-Stress Antwort	1,6
<u><i>cg1375</i></u>		Putatives Thioredoxin	1,6
Hitze-Stress Gene			
<u><i>cg2515</i></u>	<i>dnaJ2</i>	Chaperon	1,9
<u><i>cg2644</i></u> [○]	<i>clpP2</i>	Endopeptidase Clp, proteolytische Untereinheit	1,7
<u><i>cg3079</i></u> ^{●*}	<i>clpB</i>	Putative ATP-abhängige Protease (Hitze-Schock Protein)	1,6
<u><i>cg2645</i></u> ^{○*}	<i>clpP1</i>	Endopeptidase Clp, proteolytische Untereinheit	1,6
SOS und DNA Reparatur Gene			
<u><i>cg1555</i></u>	<i>uvrD3</i>	DNA/RNA Helikase	3,8
<u><i>cg1560</i></u>	<i>uvrA</i>	Excinuclease ABC, ATPase subunit A	2,2
<u><i>cg0184</i></u> ^{^*}		Putatives RNA-bindendes Protein	2,0
<u><i>cg0185</i></u> [^]		Putative Glyoxalase	1,9
<u><i>cg0186</i></u> [^]		Putative methylierte-DNA--Protein-Cystein Methyltransferase	1,9
<u><i>cg1795</i></u>	<i>uvrC</i>	Excinuklease Untereinheit C	1,5
Proteasom Komponenten			
<u><i>cg1688</i></u> [◊]	<i>pafA</i>	Putative Proteasom-Komponente	2,7
<u><i>cg1689</i></u> ^{◊*}	<i>pup</i>	Prokaryotisches Ubiquitin-ähnliches Protein	2,5
<u><i>cg0998</i></u>		Trypsin-ähnliche Serine Protease	1,5
Gene mit anderen Funktionen			

Fortsetzung Tab. IV.1

CDS ^a	Gen	Vorhergesagte Funktion	Änderung ^b
<i>cg3407</i> ◇		Putatives Membranprotein	6,4
<i>cg2106</i>		Konserviertes hypothetisches Protein	6,2
<i>cg0617†*</i>		Hypothetisches Protein	4,2
<i>cg1288</i>		Putative <i>multidrug</i> -Efflux Permease, MFS-type	3,9
<i>cg1432</i>	<i>ilvD</i>	Dihydroxy-Säure Dehydratase	3,8
<i>cg1398‡</i>		Konserviertes hypothetisches Protein	3,7
<i>cg0614†</i>		Hypothetisches Protein	3,7
<i>cg0616†</i>	<i>fdhD</i>	Putative Formate Dehydrogenase, FdhD-Familie	3,7
<i>cg1397†*</i>	<i>trmU</i>	tRNA (5-Methylaminomethyl-2-Thiouridylat)-Methyltransferase	3,7
<i>cg2423</i>	<i>lipA</i>	Lipoyl Synthetase	3,3
<i>cg0378</i>		Putatives Phagen-assoziiertes Protein	2,9
<i>cg2411</i>		Konserviertes hypothetisches Protein	2,8
<i>cg1799*</i>	<i>ribC</i>	Riboflavin Synthase, Alpha-Untereinheit	2,7
<i>cg2247</i>		Hypothetisches Protein	2,4
<i>cg1282</i>		Konserviertes hypothetisches Protein	2,4
<i>cg2127</i>		Hypothetisches Protein	2,4
<i>cg3424</i> ○	<i>cwlM</i>	N-Acetylmuramoyl-L-Alanin Amidase	2,3
<i>cg1798*</i>	<i>ribA</i>	Putative GTP Cyclohydrolase II/3,4-Dihydroxy-2-Butanon-4-Phosphat Synthase	2,2
<i>cg2835</i>		Putative Acetyltransferase	2,2
<i>cg1281</i>		ABC-type putativer <i>multidrug</i> -Transporter, ATPase und Permease Untereinheit	2,2
<i>cg1687</i> ◊		Putatives transkriptionell regulatorisches Protein	2,1
<i>cg1797*</i>	<i>ribH</i>	Riboflavin Synthase, Beta-Untereinheit	2,1
<i>cg1779</i>	<i>opcA</i>	Glukose-6-phosphate 1-Dehydrogenase Untereinheit	2,1
<i>cg2762</i>	<i>murI</i>	Glutamat Racemase	2,0

^aGene, die demselben putativen Operon angehören, sind mit dem gleichen Symbol gekennzeichnet. Das erste Gen eines Operons ist mit einem Sternchen gekennzeichnet.

Unterstrichene Gene wurden bereits als SigH-abhängig beschrieben ((Ehira *et al.*, 2009; Engels *et al.*, 2004)

^bÄnderung, Signalintensitätsverhältniss der verglichenen Proben; definiert als $2^{(m-Wert)}$

*Es sind nur Gene enthalten, deren Transkriptmenge mehr als zweifach erhöht ist, oder die zu einer der aufgeführten Funktionsklassen zählen

Von den Genen mit einem erniedrigten mRNA-Pool überschritten nur 7 den normalerweise üblichen Grenzwert von $m < -1$, was einer Halbierung des Pools entspricht. Diese Gene kodieren für Membranproteine, Transporter und für hypothetische Proteine (Anhang Tab. 1). Interessanterweise weist auch das *sigH*-Transkript selbst einen geringeren Pool (0,5-fach) in der $\Delta rshA$ -Deletionsmutante als im Referenzstamm RES167 auf. Da dieses Ergebnis unerwartet war, wurde der Promotor P_{sigH} durch PCR-Amplifikation und anschließender Sequenzierung der 5'-stromaufwärts liegenden Region auf Mutationen in diesem Stamm untersucht. Es konnten keine Mutationen in dem untersuchten Bereich (315 bp stromaufwärts des *sigH*-Translationsstarts) identifiziert werden (Daten nicht gezeigt). Es kann vermutet werden, dass das *sigH*-Transkript in der $\Delta rshA$ -Mutante aufgrund der Strukturänderung durch die Deletion weniger stabil ist.

Um die neu identifizierten, potentiell SigH-abhängigen Gene zu validieren, wurden für Folgeuntersuchungen diejenigen ausgewählt, von denen neue Einblicke in das SigH-Regulon erwartet wurden. Deshalb wurden Gene, die potentiell an der Antwort auf Disulfidstress, am Proteinabbau oder an der SOS-Antwort beteiligt sind, mittels der sensitiveren RT-qPCR untersucht.

1.4. Die differentielle Transkription ausgewählter SigH-abhängiger Gene konnte durch quantitative RT-PCR bestätigt werden

Die vorstehend beschriebene *Microarray*-Analyse lieferte eine Reihe neuer Kandidatengene für das SigH-Regulon. Um diese Ergebnisse durch eine sensitivere Methode zu bestätigen, wurden RT-qPCR-Analysen für *mshC*, *mca* und *mtr* (beteiligt an der Mycothiol-Synthese und am -Recycling (Newton und Fahey, 2008; Newton *et al.*, 2008)), *pup* (kodiert ein Ortholog zu dem kürzlich identifizierten prokaryotischen Ubiquitin-ähnlichem Protein aus *M. tuberculosis* (Burns *et al.*, 2009), sowie *uvrA* und *uvrD3* (SOS-response) durchgeführt. Zusätzlich wurden zwei der in der Deletionsmutante $\Delta rshA$ am stärksten in ihrer Expression verstärkten Gene, namentlich *dsbA1* (Thiol-Disulfid Isomerase) und *cg3405* (NADPH:Quinon Reduktase), die an der Verteidigung gegenüber Disulfidstress beteiligt sein könnten, ausgewählt. Weiterhin wurde die kürzlich beschriebene kleine Antisense-RNA *arnA*, für die gezeigt wurde, das sie von einem SigH-abhängigen, Hitzeschock-induzierbaren Promotor aus transkribiert wird, mit in die RT-qPCR-Analysen eingebunden. Das *arnA* konnte durch *Microarray*-Analysen nicht gemessen werden, da auf den verwendeten *Microarrays* nur Sonden für Protein-kodierende Gene vorhanden sind (Hüser *et al.*, 2003). Die sehr starke transkriptionelle Induktion, die in der *Microarray*-Analyse beobachtet wurde, konnte für die beiden Gene *dsbA1* und *cg3405* mit einer 60-fach bzw. 20-fach höheren Transkriptmenge in der Mutante $\Delta rshA$ im Vergleich zur Referenz RES167 bestätigt werden (Abb. IV-1). Die Gene *mshC*, *mca*, *mtr*, *uvrD3* und *arnA* waren 3- bis 4-fach induziert. Die schwächsten Induktionswerte wurden für die Gene *pup* und *uvrA* beobachtet, die eine nur 2-fach erhöhte Transkriptmenge in der Deletionsmutante $\Delta rshA$ im Vergleich zum Referenzstamm aufwiesen. Die nicht erwartete Reduktion der *sigH*-Transkriptmenge konnte durch die sensitivere RT-qPCR bestätigt werden. Die Transkriptmenge von *sigH* in $\Delta rshA$ beträgt die Hälfte der Transkriptmenge im Vergleichsstamm, was genau den Werten der *Microarray*-Analyse entspricht.

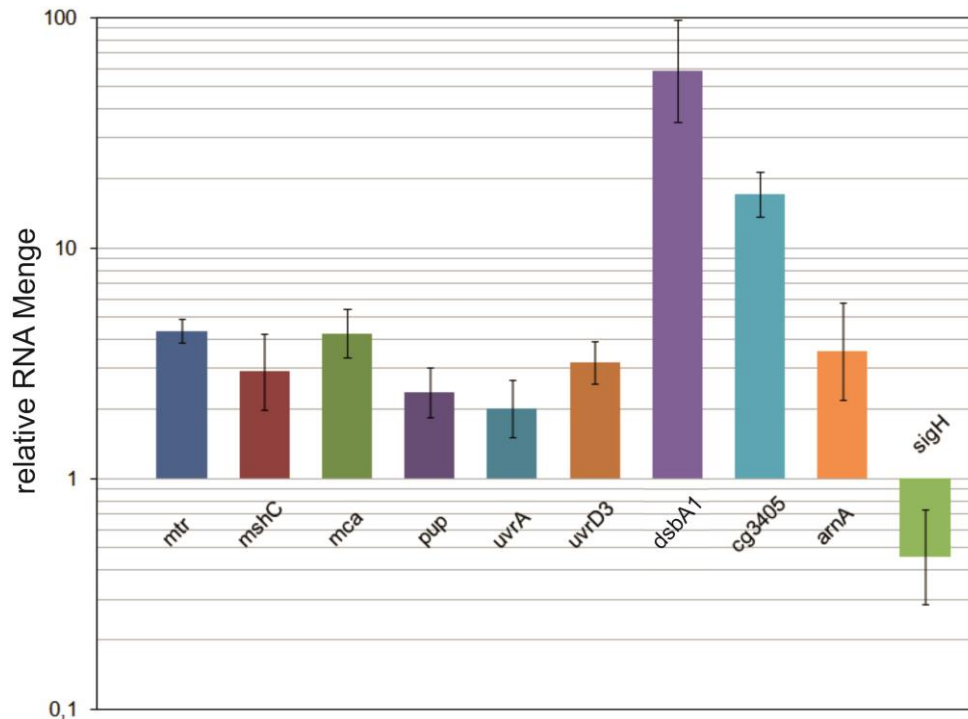


Abb. IV-6 Relative Transkriptmengen ausgewählter potentiell SigH-abhängiger Gene in *C. glutamicum* $\Delta rshA$. Die relativen Transkriptmengen wurden mit RT-qPCR bestimmt. Die Daten des Stammes RES167 wurden als Referenz verwendet und die entsprechenden Werte auf der logarithmischen Skala auf 1 gesetzt. Es wurden drei biologische Replikate von $\Delta rshA$ und vier Replikate von dem Referenzstamm RES167 jeweils doppelt gemessen. SD-Werte sind als Fehlerbalken dargestellt.

1.5. Experimentelle Bestimmung SigH-abhängiger Promotoren und Ableitung einer Promotorkonsensussequenz

Mehrere der Gene, die höhere Transkriptmengen im $\Delta rshA$ -Stamm im Vergleich zum Referenzstamm in der *Microarray*-Analyse und/oder in der RT-qPCR aufwiesen, wurden für die Bestimmung der TSP durch PEX-Analysen ausgewählt. Diese Analysen wurden von Miroslav Patek und Kollegen durchgeführt.

Es wurden die Regionen 300 - 400 bp stromaufwärts der Initiationscodons der zu analysierenden Gene genutzt, um transkriptionelle Fusionen mit dem Gen *cat* im Vektor pET2 herzustellen. Die TSP wurden innerhalb der stromaufwärts liegenden Fragmente der Gene *mshC*, *mca*, *dnaJ2*, *uvrA* und *uvrD3* (alle tragen potentiell SigH-abhängige Promotoren) jeweils 141 bp, 207 bp, 100 bp, 46 bp und 56 bp stromaufwärts des Initiationscodons identifiziert. Beispielhaft sind hier die Ergebnisse der PEX-Analyse für *dnaJ2* und *uvrA* in Abbildung 6 (Abb. IV-7) gezeigt.

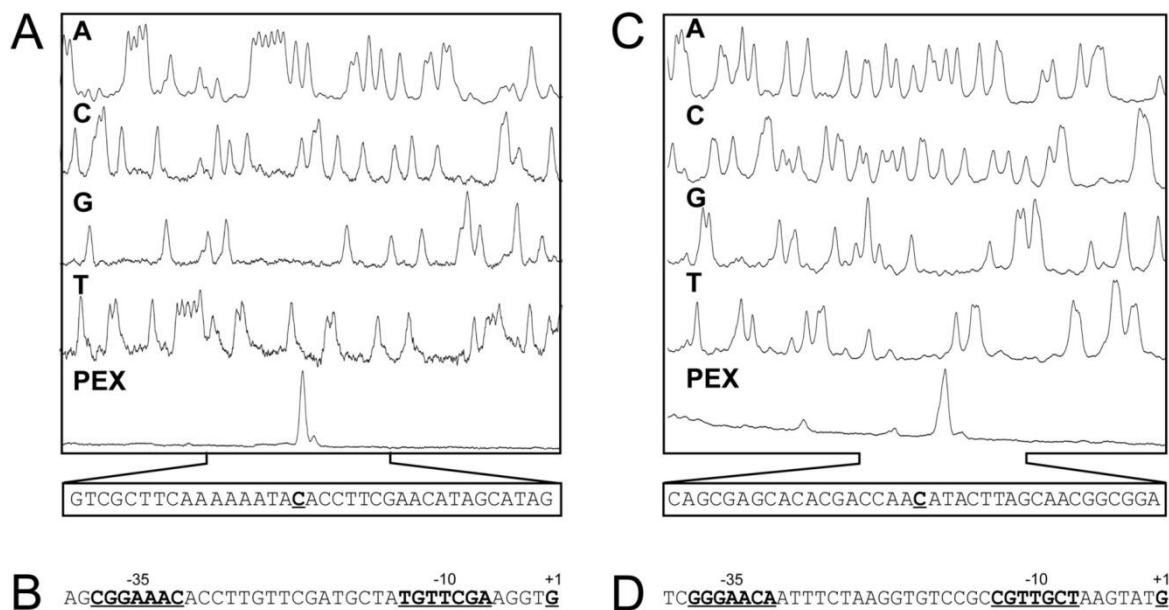


Abb. IV-7 Identifikation der Transkriptionsstartpunkte der Gene *uvrA* und *dnaJ2* durch PEX. (a) Identifikation des *uvrA*-TSP. Ein Auszug der Nukleotidsequenz abgeleitet aus den Sequenzier-Peaks ist unterhalb der Peaks gezeigt, TSP ist fett geschrieben und unterstrichen. (b) *uvrA*-Promotorsequenz. TSP (+1) und die -35- und -10-Promotorelemente sind fett geschrieben und unterstrichen. (c) Identifikation des *dnaJ2* TSP. (d) *dnaJ2*-Promotorsequenz. Wichtig: Die Sequenzen in (b) und (d) sind revers-komplementär zu denen in (a) und (c).

Die den identifizierten TSP entsprechenden -10- und -35-Regionen, die kompatibel zu der Konsensussequenz SigH-abhängiger Promotoren sind (Ehira *et al.*, 2009; Patek und Nesvera, 2011), wurden in dem zu erwartenden Abstand zu den TSP gefunden (Abb. IV-8). Zusätzlich zu diesen TSP wurden weitere Transkriptionsstarts innerhalb der Fragmente von *mca* und *pup* identifiziert, deren dazugehörige Promotoren als SigA-abhängig zu bewerten sind (Daten nicht gezeigt).

Die weiteren Analysen fanden ohne die Arbeitsgruppe um Miroslav Patek statt. Um weitere SigH-abhängige Promotoren zu identifizieren wurden die stromaufwärts vom Translationsstart liegenden Sequenzen der Gene, die in den vorangegangenen Analysen eine höhere Transkriptmenge in der $\Delta rshA$ -Mutante relativ zum Referenzstamm RES167 aufwiesen, für eine Motivsuche mittels des Programms Bioprospector (Liu *et al.*, 2001) herangezogen. Für das notwendige Trainingsset wurden alle bisher beschriebenen SigH-abhängigen Promotoren (Ehira *et al.*, 2009), der Promotor von *arnA* (Zemanova *et al.*, 2008) und die in dieser Arbeit bestimmten Promotoren *rshA*, *mshC*, *mca*, *dnaJ2*, *uvrA* und *uvrD3* verwendet. Es wurden zwei 10-bp Motive mit einem Abstand von 15 – 23 bp gesucht. Unter Verwendung des Programms Bioprospector konnten 10 weitere transkriptionelle Einheiten, die ein konserviertes

SigH-abhängiges Promotormotiv stromaufwärts des Startcodons enthalten, identifiziert werden (Abb. IV-8). In den stromaufwärts des Translationsstarts liegenden Bereichen der anderen 26 analysierten transkriptionellen Einheiten konnten keine weiteren Motive identifiziert werden. Deren Transkription scheint offensichtlich kein direkter Effekt der SigH-Aktivität zu sein. Hierbei handelt es sich wahrscheinlich um einen sekundären Effekt, der unter den gewählten Bedingungen auftritt.

Gen-Name/ID	-35	-10	+1	Referenz
<i>clpP1</i>	P1	GTTTCAT TGAAAT ACGCGGGTAGTCTGGTGAC CATTGAA CCAAA		(Engels et al., 2004)
<i>clpC</i>	P2	AAAGTCT TGGAAGT TTTGCC--CAATAAGGG CGTTAAAGT GGGT		(Engels et al., 2004)
<i>clpR</i>	P1	TAAACT GGGAACA AATTTT--AGGGAAGGG AGTTGAA CTTAAC		(Engels et al., 2004)
<i>dnaK</i>	P2	TCTAGT GGGAACA ACTTTG--TAGCATTCGG CGTTGTC ATATA		(Barreiro et al., 2004; Ehira et al., 2009)
<i>sigB</i>		GGCCTT GGGAAC TTTTGT--GGAAGCAGT CGTTGAA CCCTTTG		(Halgasova et al., 2001)
<i>arnA</i>	P2	TGTGT GAGTA AAAGCTGCG--GACATAGTAT TGTTCTT TCAGGCTG		(Zemanová et al., 2008)
<i>trxB</i>		AACTGAT TGGAAGT TTTTCA--AAGTGTCTG CGTTGAA AACGGTG		(Ehira et al., 2009)
<i>trxB1</i>		TGGCC GGGAATA ACTACA--GTCCCGCTGAA AGTTGGT CTATATAAG		(Nakunst et al., 2007)
<i>trxC</i>		AATGT CGGAAT TCCCAGG--AGTCCCGTCAT TGTTAAT TAGGAG		(Ehira et al., 2009)
<i>suFR</i>		GGAC CGGAAT GAAATTA--GGGAACACT TGTTGTC TAAAGGTG		(Ehira et al., 2009)
<i>clpB</i>	P1	CTTGAGT GGAACA TACTCA--ACTCTTTGT CGTTA TAGTATTA		(Ehira et al., 2009)
<i>msrB</i>		GCTGGAT TGGAAT TTTTCAG--CGCGACCA TGGTTGGG STCTATTG		(Ehira et al., 2009)
<i>cgR_1554</i>		GTGTGAT TGGAAT AACGTTA--ACAATAAGTT TGTTACAT GGTGTG		(Ehira et al., 2009)
<i>cgR_2964</i>		CGGGGG GGGAAT GGAAAA--GTACGCTT TGTTCA TATAGCG		(Ehira et al., 2009)
<i>cgR_1317</i>		GATTT CGGAAC ATGCGGA--TACGCTACGT TGTTGAG ATCAATTA		(Ehira et al., 2009)
<i>cgR_2078</i>		CAAA TGGAAT AGGGGTG--CACACTCAT CGTTGAA AGGAATCA		(Ehira et al., 2009)
<i>cgR_1297</i>		TAGCT GGGAAT TAGCTTTG--TACTTAAACT TGTTGTT TTAAGTG		(Ehira et al., 2009)
<i>cgR_0627</i>		ACGCC GGGAAT TTCCCG--GCCCGCTT TGTTGAA TAAACG		(Ehira et al., 2009)
<i>cgR_2320</i>		CGGTG AGGAAC TAAAAAA--CTCATCACCGT TGTTGAG ATCAAGTG		(Ehira et al., 2009)
<i>cgR_2183</i>		ATTGAT TGGAACA AGAAAG--GTACCCAGT TGTTGCA AAGGAGG		(Ehira et al., 2009)
<i>cgR_2451</i>		CACTA TGGAAT AAATTC--TGTCTACAG CGTTAC AGTAAATG		(Ehira et al., 2009)
<i>cgR_2903</i>		TGATAT GGGAAC TAAATTTGGG--AGCATTCAC CGTTA ACGAAGT		(Ehira et al., 2009; vorhergesagt)
<i>mtr</i>		GTACT TGGAAGC TTTTTAT--AGTCATAT CGTTGAG ATACGTG		(Ehira et al., 2009; vorhergesagt)
<i>msrA</i>		CAAAC CGGAAT AATCGGC--AGCTGAGTAG GGTTG AGGTCATAA		(Ehira et al., 2009; vorhergesagt)
<i>hemA</i>		GGCTC CGGAAT TAAACTC--ACTGAGCGAA TGTTGAT CAAGATGC		(Ehira et al., 2009; vorhergesagt)
<i>cgR_2470</i>		ATCGCG GGGAAT AGTTAGC--GGCGCATGT GGTTGGA AATATGA		(Ehira et al., 2009; vorhergesagt)
<i>cgR_2417</i>		GGCTC GGGAAT TATTTGG--TGCAACGAGG CGTTGAA CAACAAGA		(Ehira et al., 2009; vorhergesagt)
<i>cgR_1848</i>		CTCGT CGGAAT TTTTTT--CCTCAAGCG GGTTAA CTCACTGT		(Ehira et al., 2009; vorhergesagt)
<i>cgR_1753</i>		ATTT TAGGAAT TTTTTCG--AGGGTCCCC CGTTGTA CTGAGCAG		(Ehira et al., 2009; vorhergesagt)
<i>rshA</i>		CCATCG TGGAAGA AAAAACAG--CTCCGAGGA TGTTAA AGGAAGTAG		(diese Arbeit; experimentell bestimmt)
<i>dnaJ2</i>		CTACT CGGAACA AATTTCT--AAGGTGTCCCG CGTTGCT TAAGTATG		(diese Arbeit; experimentell bestimmt)
<i>uvrA</i>		GACTAG CGGAACA CCCTTG--TTCGATGCTA TGTTGAA AGGTTG		(diese Arbeit; experimentell bestimmt)
<i>mca</i>		GGTGT TGGAAT TTCGAA--GCGCCTGAT CGTTGGA TGGAGAG		(diese Arbeit; experimentell bestimmt)
<i>uvrD3</i>		GCTAT TGGAAT GATTGAT--AGCTCCCA AGTTGTA TCTATTC		(diese Arbeit; experimentell bestimmt)
<i>mshC</i>		GGTT TGGAAT AGGTGCA--CTGGCGG CGTTGAA AGTTTCAG		(diese Arbeit; experimentell bestimmt)
<i>pup</i>		GTGTGAT TGGAAT AACGTTA--ACAATAAGTT TGTTACAT AGTGTGG		(diese Arbeit; vorhergesagt)
<i>cg3344</i>		TGATCG GGGAACA AATTTG--AAGGTTTT TCAGTTGCT ATAGGTAT		(diese Arbeit; vorhergesagt)
<i>cg0378</i>		TTACGAT TGGAACA TTTTTTG--AAGAATACCT CGTTGAA TCTAGTGC		(diese Arbeit; vorhergesagt)
<i>cg2835</i>		TTCT TAGGAAT AAAAACA--GGGTGCTTT TGTTCAA AAGGTATA		(diese Arbeit; vorhergesagt)
<i>ispG</i>		ACGAT CGGAAT TGAACTG--CCGGCACCGA TGTTAA AGAGGTGA		(diese Arbeit; vorhergesagt)
<i>cg2434</i>		CCCC GGGAAC ATCACGA--ATAACAACCG AGTTAC AGCCCGCAA		(diese Arbeit; vorhergesagt)
<i>lipA</i>		AGTAC AGGAAT TTTGTGCG--GTGGGGCGCT CGTTGAA CGGAAGTA		(diese Arbeit; vorhergesagt)
<i>guaB2</i>		CGTT TGGAAT AGCGTGT--GAGTGCAG CGTTAT CCGATTAA		(diese Arbeit; vorhergesagt)
<i>cg0184</i>		CGGGCT GGGAAT TTTCATG--ATCTTCGAGGG GGTTGTA GGAATCGG		(diese Arbeit; vorhergesagt)
<i>zwf</i>		CATGAA TGAAAT TAGTTGATCTTATGTGG CGTTAC ACATCTTTC		(diese Arbeit; vorhergesagt)
Konsensus				
<i>C. glutamicum</i>		KGGAAYW		YGTRRW
<i>M. tuberculosis</i>		GGGAAYA	n (16-19)	CGTT (Raman et al., 2001)
<i>S. coelicolor</i>		GGGAAT		GTTG (Paget et al., 2001)

Abb. IV-8 Sequenzen der mutmaßlichen SigH-abhängigen Promotoren von *C. glutamicum*. Putative -10- und -35-Regionen (Abstand 16 - 19 Nukleotide) und TSP (+1) sind in fetter Schrift hervorgehoben. Striche symbolisieren Lücken die eingeführt wurden, um das *Alignment* der -35er-Elemente zu ermöglichen. Die Positionen in der *C. glutamicum* Konsensussequenz mit einem Auftreten eines einzelnen Nukleotids in über 80 % der verwendeten Sequenzen sind fett geschrieben, K = G oder T; Y = C oder T; R = A oder G; W = A oder T. Die von Halgasova ermittelte Sequenz des *sigB*-Promotors ist von *C. glutamicum* CCM251 (Halgasova et al., 2001), die von Ehira ermittelten Sequenzen von *C. glutamicum* R (Ehira et al., 2009). Die übrigen Sequenzen stammen aus *C. glutamicum* ATCC 13032 (J Kalinowski et al., 2003).

Insgesamt konnten neben zehn, durch Motivsuche identifizierten, SigH-abhängigen Promotoren, sechs SigH-abhängige Promotoren durch die präzise Identifizierung von TSP mittels PEX lokalisiert werden. Basierend auf diesen neuen Sequenzen konnte mit den bereits publizierten Sequenzen eine verfeinerte Promotorkonsensussequenz für SigH-abhängige Promotoren erstellt werden, die auf nun 45 Sequenzen beruht (Abb. IV-9).



Abb. IV-9 Nukleotidverteilung innerhalb der -35- und -10-Kernregionen SigH-abhängiger Promotoren in *C. glutamicum*. Das prozentuale Auftreten eines Nukleotids an einer Position ist durch die Größe des Nukleotidsymbols (A, C, G, T) dargestellt. Hierfür wurde das Programm WebLogo genutzt (Crooks *et al.*, 2004). Diese Analyse basiert auf 45 mutmaßlich SigH-abhängigen Promotoren.

2. Tiefenanalyse des *Corynebacterium glutamicum* SigH-Regulationsnetzwerkes mittels RNA-Sequenzierung der Deletionsmutanten RES167 $\Delta sigH$ und $\Delta rshA$

Neue methodische Entwicklungen in dem Bereich der Transkriptomik (RNA-Seq) ermöglichen eine höhere Sensitivität und Auflösung bei der Analyse und ergeben dadurch ein präziseres Bild bei der Betrachtung von Transkriptomen durch die Basen-genaue Identifizierung von TSP und den sich daraus ergebenden Promotoren (Pátek *et al.*, 2013). Die Vorteile der RNA-Seq nutzend, wurde hier das Regulationsnetzwerk des Sigmafaktors SigH im Detail erneut untersucht.

2.1. Erstellung der cDNA-Bibliotheken von mRNA des Referenzstammes *C. glutamicum* RES167 und der Deletionsstämme *C. glutamicum* RES167 $\Delta sigH$ und $\Delta rshA$

Um die TSP und damit die Promotoren zu identifizieren, die in Abhängigkeit von SigH transkribiert werden, wurde hier die Methode der nativen 5'-Enden-RNA-Seq genutzt. Bei dieser Methode werden Transkripte mit nativen (nicht durch Prozessierung entstandenen) 5'-Enden angereichert. Die Sequenzierung der entsprechenden cDNAs erlaubt die eindeutige Identifizierung der TSP und damit der Promotoren.

Die RNA zur Erstellung der erforderlichen cDNA-Bibliotheken wurden von je 3 biologischen Replikaten der Stämme *C. glutamicum* RES167, *C. glutamicum* RES167 $\Delta sigH$ und *C. glutamicum* RES167 $\Delta rshA$ isoliert, die in Minimalmedium CGXII bis zur exponentiellen Phase angezogen wurden. Die biologischen Replikate wurden nach der RNA-Isolierung vereinigt. Anschließend wurden je 10 μg der vereinigten RNA für die Erstellung der nativen 5'-Enden cDNA-Bibliotheken eingesetzt. Die Konzentration der cDNA-Bibliotheken betragen für *C. glutamicum* RES167 1.955,4 pmol/mL für *C. glutamicum* RES167 $\Delta rshA$, 2.689,6 pmol/mL für *C. glutamicum* RES167 $\Delta sigH$ 4.315,5 pmol/mL. Die Größe der cDNAs lag zwischen 150 und 800 bp, wobei der Hauptanteil der cDNAs eine Größe von 150 bp bis 500 bp aufwies.

Die Anreicherung der nativen Transkripte erfolgte mit Hilfe der Terminator-Exonuklease und die Sequenzierung der erstellten cDNA-Bibliotheken mit dem Illumina Genome Analyzer IIx. Das Mapping der cDNA Reads auf das *C. glutamicum* Referenzgenom von *C. glutamicum* ATCC 13032 (Kalinowski *et al.*, 2003) erfolgte mit Hilfe der Software SARUMAN (Jochen Blom *et al.*, 2011) und die Visualisierung der gemappten Reads wurde mit VAMP (Hilker *et al.*, 2013, in Vorbereitung) durchgeführt.

2.2. Mapping-Statistik der RNA-Seq Läufe der Stämme *C. glutamicum* RES167, *C. glutamicum* RES167 $\Delta sigH$ und *C. glutamicum* RES167 $\Delta rshA$

Die Sequenzierung der cDNA-Bibliotheken der Stämme *C. glutamicum* RES167 und der abgeleiteten Deletionsmutanten $\Delta sigH$ und $\Delta rshA$ resultierte in Millionen an sequenzierten cDNA Reads der Länge 35 Basen. Hiervon konnten durch SARUMAN (Jochen Blom et al., 2011) für *C. glutamicum* RES167 1.010.007 cDNA-Reads auf das Referenzgenom *C. glutamicum* ATCC 13032 (J Kalinowski et al., 2003) gemappt werden. Hiervon mappen 743.011 einfach und 275.996 an mehreren Positionen im Genom, was zu einer Gesamtzahl von 2.186.439 mappings für *C. glutamicum* RES167 führt (Tab. IV.2). Die Gesamtzahl an Mappings betrug für RES167 $\Delta sigH$ 4.924.033 und für $\Delta rshA$ 6.683.905. Die genaue Aufschlüsselung in einfach mappende und mehrfach mappende Reads ist in Tabelle Tab. IV.2 aufgeführt. Die erstellten RNA-Seq-Datensätze wurden für die Identifizierung von TSP genutzt.

Tab. IV.2 Mapping-Statistik der RNA-Seq Läufe der Stämme *C. glutamicum* RES167, $\Delta sigH$ und $\Delta rshA$

RNA-Seq-Lauf	Gesamtzahl Mappings	Einfach mappende RReads	Mehrfach mappende RReads	Gesamtzahl mappende cDNA-Reads
RES167 $\Delta sigH$	4.924.033	1.844.466	585.011	2.429.477
RES167 $\Delta rshA$	6.683.905	3.591.884	613.307	4.205.191
RES167	2.186.439	734.011	275.996	1.010.007

2.3. Identifizierung der Transkriptionsstartpunkte durch Analyse der RNA-Seq-Datensätze von *C. glutamicum* RES167, *C. glutamicum* RES167 $\Delta sigH$ und *C. glutamicum* RES167 $\Delta rshA$

Durch die Analyse des RNA-Seq-Datensatzes (Mapping-Datensatz) von *C. glutamicum* RES167 konnten 1.847 TSP Skript-gestützt identifiziert werden. Hiervon entfielen 183 TSP auf rRNA-Regionen. Diese Regionen werden normalerweise nicht beachtet, da sie etliche Terminator Nuklease-resistente Abschnitte enthalten und so reich an falsch-positiven TSP-Signalen sind. Die Analyse des $\Delta rshA$ -Datensatzes ergab nach Abzug der TSP, die auf rRNA-Regionen entfallen (417), 4.835 TSP. Für $\Delta sigH$ konnten 3.260 TSP identifiziert werden, exklusive der rRNA TSP (370). Die automatisch identifizierten TSP umfassen neben den für die weitere Analyse ausgeschlossen TSP, die zu rRNA- und tRNA-Genen gehören noch

alternative TSP. Hierbei handelt es sich um TSP, die weniger als 5 bp und mehr als 10 bp von der -10 er-Region eines Promoters entfernt sind (Patek, persönliche Mitteilung).

2.4. Bestätigung der transkriptionellen Organisation des *sigH-rshA*-Operons

Um die transkriptionelle Organisation des *sigH-rshA*-Operons mittels RNA-Seq zu analysieren, wurden die nativen 5'-Enden RNA-Seq-Datensätze von *C. glutamicum* RES167, RES167 $\Delta rshA$ und RES167 $\Delta sigH$ durch das Programm VAMP (Hilker *et al.*, 2013, in Vorbereitung) visualisiert (Abb. IV-10). Die Positionen der automatisch identifizierten TSP für diese Region sind deckungsgleich mit einem plötzlichen Anstieg der *mappenden Reads* an der entsprechenden Position. Die durch PEX-Analyse identifizierten SigA-abhängigen TSP (1, 2 und 3) stromaufwärts von *sigH* konnten in allen drei Datensätzen identifiziert werden, während der SigH-abhängige TSP stromaufwärts von *rshA* nur in dem RNA-Seq-Datensatz der Deletionsmutante RES167 $\Delta rshA$ identifiziert werden konnte (siehe Track 2 Abb. IV-10).

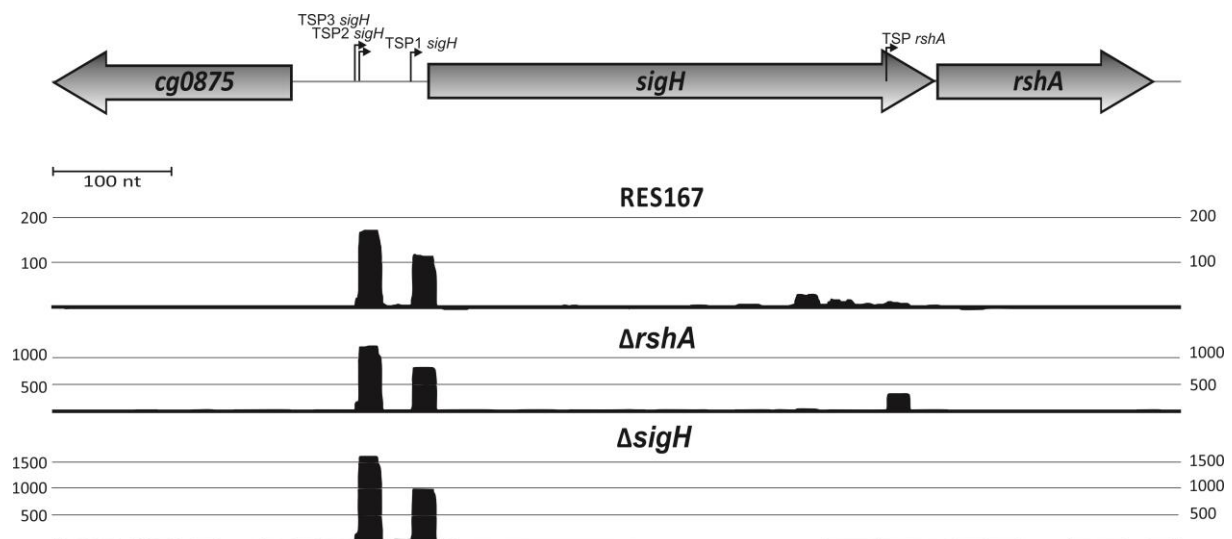


Abb. IV-10 Visualisierte native 5'-Enden RNA-Seq-Daten des genomischen Bereichs *sigH-rshA* von *C. glutamicum* RES167, RES167 $\Delta sigH$ und RES167 $\Delta rshA$. TSP wurden durch einen Anstieg der Anzahl *mappenden Reads* an einer Position identifiziert und sind durch Pfeile gekennzeichnet. In den drei Datensätzen konnten 3 SigA-abhängige TSP (TSP1 *sigH*, TSP2 *sigH* und TSP3 *sigH*) stromaufwärts von *sigH* identifiziert werden. Ein zusätzlicher TSP konnte in dem Deletionsstam RES167 $\Delta rshA$ stromaufwärts von *rshA* identifiziert werden. Das *Mapping* der sequenzierten *Reads* wurde durch SARUMAN durchgeführt (Jochen Blom *et al.*, 2011) und die Visualisierung erfolgte durch VAMP (Hilker *et al.*, 2013, in Vorbereitung).

In **Abb. IV-11** ist die Verteilung der *mappenden Reads* stromaufwärts des *sigH*-Translationsstartes im RNA-Seq-Datensatz von RES167 $\Delta rshA$ auf Sequenzebene dargestellt.

An den Positionen des TSP1 (Pos. 805859) und des TSP2 (Pos. 805792) ist ein deutlicher Anstieg der *mappenden Reads* von 52 auf 806 bzw. 169 auf 910 *Reads* zu erkennen, wodurch die Positionen automatisch als TSP identifiziert werden konnten. Deutlich weniger *mappende Reads* (154) sind an der Position des TSP3 (4 bp stromaufwärts des TSP2) vorhanden.

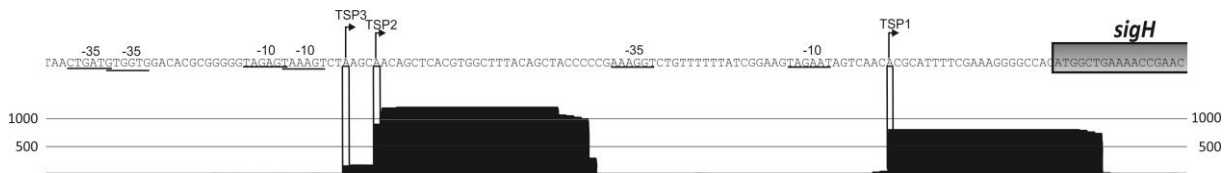


Abb. IV-11 Visualisierte native 5'-Enden RNA-Seq-Daten des genomischen Bereichs stromaufwärts von *sigH* von *C. glutamicum* RES167 $\Delta rshA$. TSP wurden durch einen Anstieg der Anzahl *mappenden Reads* an einer Position identifiziert. Das entsprechende Start-Nukleotid ist über Striche mit den *mappenden Reads*, die zur Identifikation führten verbunden. Die TSP sind zusätzlich durch Pfeile gekennzeichnet. In dem RNA-Seq-Datensatz von RES167 $\Delta rshA$ konnten 3 SigA-abhängige TSP (TSP1, TSP2 und TSP3) stromaufwärts von *sigH* identifiziert werden. Die entsprechenden –10- und –35-Regionen der Promotoren sind unterstrichen und benannt. Das Mapping der sequenzierten *Reads* wurde durch SARUMAN durchgeführt (Jochen Blom et al., 2011) und die Visualisierung erfolgte durch VAMP (Hilker et al., 2013, in Vorbereitung).

Wie bereits in Kapitel IV.1.2 beschrieben, handelt es sich bei den entsprechenden Promotoren der TSP jeweils um SigA-abhängige Promotoren. Wobei der zu TSP3 zugehörige Promotor, wie hier durch die geringe Anzahl an *mappenden Reads* im Vergleich zu der Anzahl *mappenden Reads* der TSP1 und 2 gezeigt wurde, eine geringere Aktivität aufweist. Dieses Ergebnis deckt sich mit den Ergebnissen der PEX-Analyse, in der das PEX-Signal für TSP3 ebenfalls deutlich geringer war (Abb. IV-3).

Der TSP4, der durch PEX-Analyse unter der Verwendung von Gesamt-RNA aus *C. glutamicum* pET2::*sigH4* und den Primer CM4 und CM5 131 bp stromaufwärts des *sigH*-Translationstarts identifiziert und durch Plasmid-basierter 5'-RACE bestätigt werden konnte (IV.1.2), konnte in keinem der 5'-angereicherten RNA-Seq-Datensätze identifiziert werden.

Für das *rshA*-Gen konnte in dem RNA-Seq-Datensatz RES167 $\Delta rshA$ ein TSP 63 bp stromaufwärts des *rshA*-Translationstarts identifiziert werden. Es ist ein deutlicher Anstieg an *mappenden Reads* an der Position des *rshA*-TSP zu erkennen. In **Abb. IV-12** sind die *mappenden Reads* des RNA-Seq-Datensatzes RES167 $\Delta rshA$ mittels VAMP visualisiert. Sieben bp stromaufwärts dieser Region befindet sich die –10-Region des SigH-abhängigen Promotors, was in Übereinstimmung mit den Abständen der TSP zur –10-Region aller bisher beschriebenen SigH-abhängigen Gene ist. Dieser Abstand schwankt für die identifizierten TSP und den dazugehörigen Promotoren zwischen 5 und 10 bp (Abb. IV-8).

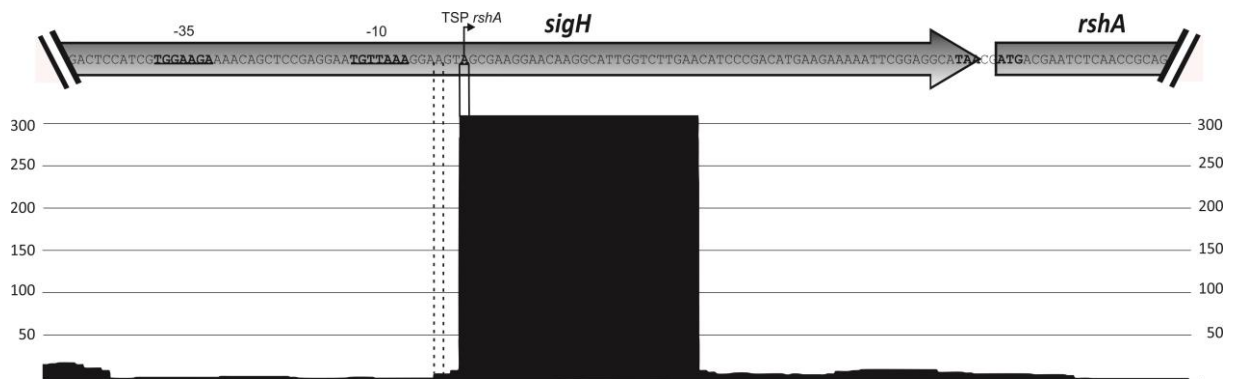


Abb. IV-12 Visualisierte native 5'-Enden RNA-Seq-Daten des genomischen Bereichs stromaufwärts von *rshA* von *C. glutamicum*. TSP wurden durch einen Anstieg der Anzahl *mappender Reads* an einer Position identifiziert. Das entsprechende Startnukleotid ist über Striche mit den *mappenden Reads*, die zur Identifikation führten, verbunden. Der TSP ist zusätzlich durch einen Pfeil gekennzeichnet. In dem RNA-Seq-Datensatz von RES167 $\Delta rshA$ konnte ein TSP stromaufwärts von *rshA* identifiziert werden (TSP *rshA*). Der durch PEX identifizierte TSP ist durch eine gestrichelte Linie gekennzeichnet. Die Anzahl *mappender Reads* an dieser Position reichte nicht für eine Identifikation dieses TSP aus. Die entsprechenden -10- und -35-Regionen des bereits als SigH-abhängig definierten Promotors sind unterstrichen und benannt. Das *Mapping* der sequenzierten *Reads* wurde durch SARUMAN durchgeführt (Jochen Blom et al., 2011) und die Visualisierung erfolgte durch VAMP (Hilker et al., 2013, in Vorbereitung).

Der durch die Analyse der RNA-Seq-Daten identifizierte TSP für *rshA* unterscheidet sich von den TSP, die durch PEX bestimmt werden konnten. In der PEX-Analyse dieser Promotoregion wurden zwei TSP identifiziert; TSP1 62 bp (G) und TSP2 66 bp (A) stromaufwärts des *rshA*-Translationsstarts (Abb. IV-3). Durch die Analyse der RNA-Seq-Daten konnte ein einzelner definierter TSP (TSP *rshA*) 63 bp (A) stromaufwärts des *rshA* Translationsstart im perfekten Abstand zur -10-Promotorregion eindeutig identifiziert werden.

2.5. Experimentelle Bestätigung der vorhergesagten SigH-abhängigen Promotoren und Korrektur einiger experimentell bestimmter TSP durch die TSP-Identifizierung per RNA-Seq des Deletionsstamms RES167 $\Delta rshA$

Um die bereits beschriebenen SigH-abhängigen TSP (Abb. IV-8) per RNA-Seq-Analyse zu validieren bzw. die vorhergesagten TSP experimentell zu bestätigen, wurden die über die Auswertung der RNA-Seq-Daten identifizierten TSP des *C. glutamicum* RES167 $\Delta rshA$ -Stammes mit den bereits beschriebenen TSP abgeglichen. Da durch die *Microarray*-Analyse des Deletionsstammes RES167 $\Delta rshA$ gezeigt werden konnte, dass die Mehrheit der beschriebenen SigH-abhängigen Gene durch die Deletion des Anti-Sigmafaktors RshA einen erhöhten mRNA-Pool aufwies, wurde davon ausgegangen, dass die entsprechenden TSP im RNA-Seq-Datensatz von RES167 $\Delta rshA$ identifiziert werden können. Weiter wurde davon ausgegangen, dass auch die TSP der Gene *dnaK*, *clpC*, *sigB*, *sufR* und *clgR*, für die keine differentielle Transkription per *Microarray*-Analyse des Deletionsstammes RES167 $\Delta rshA$ nachgewiesen werden konnte, mit der sensitiveren Methode der RNA-Seq nachweisen zu können. Von den 45 beschriebenen SigH-abhängigen Promotoren konnten, bis auf sieben, alle über die Identifizierung der TSP im RNA-Seq-Datensatz bestätigt werden, darunter auch TSP der Gene *dnaK*, *clpC*, *sigB*, *sufR* und *clgR*. Von 18 vorhergesagten SigH-abhängigen Promotoren (Abb. IV-8) konnten 15 durch die Bestimmung des TSP experimentell bestätigt werden (Tab. IV.3). Für das Gen *cg3344* wurden zwei unterschiedliche SigH-abhängige Promotoren *in silico* vorhergesagt. Hiervon konnte nur der in dieser Arbeit vorhergesagte Promotor und nicht der von Ehira und Mitarbeitern gemutmaßte Promotor durch die Identifizierung des TSP im RNA-Seq-Datensatz experimentell bestätigt werden.

Neben dem experimentellen Nachweis der vorhergesagten TSP und damit deren Promotoren wurden die TSP, die bereits in anderen Arbeiten und dieser Arbeit experimentell durch andere Methoden bestimmten worden sind, mit denen der RNA-Seq-Datensätze identifizierten TSP verglichen. In diesem globalen Ansatz konnten die TSP für alle Gene, bis auf *dnaK*, *trxB1*, *clpC*, *clgR* und *rshA*, bestätigt werden (Tab. IV.3). Für *dnaK* wurde durch Verwendung der RNA-Seq-Daten ein TSP, der zwei bp stromaufwärts des durch Ehira und Mitarbeitern durch 5'-RACE bestimmten TSP liegt, identifiziert. Dieser TSP liegt im perfekten Abstand von 7 bp zur -10-Promotorregion. Somit wurde der TSP für *dnaK* auf die Position 120 bp stromaufwärts des Translationsstarts korrigiert. Der TSP des Gens *trxB1* wurde durch RNA-Seq 2 bp stromaufwärts des durch Nakunst *et al.*, (2007) publizierten TSP identifiziert und besitzt den optimalen Abstand von 7 bp zur entsprechenden -10-Promotorregion. Für die TSP der Gene *clpC*, *clgR* und *rshA*, die durch PEX bestimmt wurden (Engels *et al.*, 2004 ; bzw. diese Arbeit),

konnten alternative TSP im Abstand eines Nukleotides identifiziert werden (Tab. IV.3). Für das Gen *clpC* wurde ein TSP ein Basenpaar stromabwärts des durch Engels *et al.* bestimmten TSP (T), identifiziert. Das Nukleotid T ist sehr selten der Ausgangspunkt der Transkription (Patek, persönliche Kommunikation), daher wurde der durch RNA-Seq-Daten identifizierte TSP (ein Guanin) als korrekt angesehen. Auch der TSP von *clgR* wurde ein bp stromabwärts des publizierten TSP (Engels *et al.*, 2004) mit dem Nukleotid G, anstelle von C, identifiziert. Da die große Mehrheit der SigH-abhängigen Gene von einem G oder A aus transkribiert wird, wurde der TSP entsprechend auf TSP G, ein bp stromaufwärts des *clgR*-Translationsstarts, korrigiert. Die Korrektur des TSP von *rshA* wurde bereits in Kapitel IV.2.4 eingehend erläutert. Bei der Korrektur der 5 TSP von *Dank*, *trxB1*, *clpC*, *clgR* und *rshA* fällt auf, dass zwei der korrigierten TSP durch 5'-RACE bestimmt worden ist und die übrigen durch PEX.

Tab. IV.3 Promotoren und TSP SigH-abhängiger Gene per RNA-Seq von dem Deletionsstamm RES167 $\Delta rshA$ bestimmt – Experimentelle Bestätigung der vorhergesagten SigH-abhängigen Promotoren und Korrektur einiger experimentell bestimmter TSP

Kodier-region (CDS)	Alternative CDS	Gen-name	Position Genom	relative Anzahl mappende Reads	Abstand Start-codon	Sequenz (-39 -- +1)	Referenz	TSP-Identifizierungsmethode	Neue Referenz ^a (TSP-Identifizierung)
cg2078		<i>msrB</i>	1969473	15,32	121	GAT TGGAAT TTTTTCAGCG-CGACCGTT GGTTGGGG ACTATT G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg2538	cgR_2183	-	2418654	26,57	42	GAT TGGAACA AAGAAAGGT-GCCCAGT CTGTTGCA AAGAAG G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg3100		<i>dnaK</i>	2959823	309,76	120	AGT TGGAACA ACTTTGTGG-CATTTAC CGTTGCT TATATA TG	(Ehira <i>et al.</i> , 2009)	5'-RACE	diese Arbeit (RNA-Seq)**
cg3422		<i>trxB</i>	3273611	305,93	68	TGA TGGAAGT TTTTTCAAAG-TGTC TGACCGTTGAA AACGGT G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg3299		<i>trxB1</i>	3151821	978,84	31	CGGGAATA ACTACAGTC-CGCTGAA AGTTGGT CTATATA A	(Nakunst <i>et al.</i> , 2007)	5'-RACE	diese Arbeit (RNA-Seq)**
cg1765		<i>sufR</i>	1655074	38,3	30	CGGGAATG GGAATTAGGGAACACT TGTTGTC TAAAGGT G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg3079		<i>clpB</i>	2939558	84,98	91	GAG TGGAACA TACTCAACT-CTTTGT CGTTATAG ATATTA A	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg3405	cgR_2964	-	3258874	91,68	36	ACGG GGGAATG GAAAAAGTA-CGCTTGG TGTTCA TATAGC G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg1397	cgR_1317	<i>trmU</i>	1305259	47,16	38	TC GGGAACA TGCGGATAC-GCTACGT TGTTGAG ATTAATTA A	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg2411	cgR_2078	-	2302853	27,53	254	ATC GGGAATAG GGGTGCA-CACTTCAT CGTTGAA AGGAAT C	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg1375	cgR_1297	-	1283606	69,66	56	TAG GGATTAG CTTTGTA-CTTAACT TGTTGTT TTTAAGT G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg0617	cgR_0627	-	544705	90,01	36	CAG GGAAAT TTCCGCGCCGCTTCCT TGTTTAA AATAAC G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg2661	cgR_2320	<i>dsbA2</i>	2538097	427,54	190	GAGGAACT AAAAA ACTCAT CACCGT TGTTGAG ATCAAGT G	(Ehira <i>et al.</i> , 2009)	5'-RACE	
cg0877		<i>rshA</i>	806441	72,05	63	CG TGGAAGA AAAACAGCT-CCGAGGAA TGTTAAAG GAAGT AG	(Busche <i>et al.</i> , 2012)	PEX	diese Arbeit (RNA-Seq)**
cg2515		<i>dnaJ2</i>	2395613	519,46	93	TC GGGAACA AATTTCTAAGGTGTCCG CGTTGCT AAGTAT G	(Busche <i>et al.</i> , 2012)	PEX	
cg2963		<i>clpC</i>	2820286	136,69	38	AGT CTGGAAGT TTTGCCCA--ATAAGGG CGTTAAAG TGGGT G	(Engels <i>et al.</i> , 2004)	PEX	diese Arbeit (RNA-Seq)**

Fortsetzung Tab. IV.3

Kodier-region (CDS)	Alternative CDS	Gen-name	Position Genom	relative Anzahl mappende Reads	Abstand Start-codon	Sequenz (-39 -- +1)	Referenz	TSP-Identifizierungsmethode	Neue Referenz ³ (TSP-Identifizierung)
cg2152		<i>clgR</i>	2039986	154,16	1	ACTGGGAACAAATTTTAGG-GAAAGGGAGTTGAACCTAAC <u>G</u>	(Engels <i>et al.</i> , 2004)	PEX	diese Arbeit (RNA-Seq)**
cg1560		<i>uvrA</i>	1443908	30,64	46	ACTAGCGGAAACACCTTGTTTCGATGCTATGTTTCGAAGGT <u>G</u>	(Busche <i>et al.</i> , 2012)	PEX	
cg1127		<i>mca</i>	1046412	338,73	207	GTTTGGAAATGTTTCGAAGC-GCCTGATGCGTTGGATGGAGAG <u>G</u>	(Busche <i>et al.</i> , 2012)	PEX	
cg1555		<i>uvrD3</i>	1441999	32,08	56	ATCTGGAATGATTGATAG-CTCCCAAGTGTGTATCTATT <u>C</u>	(Busche <i>et al.</i> , 2012)	PEX	
cg1709		<i>mshC</i>	1599116	39,02	141	TTGTGGAATAGGTGCACT-GGCGGCTTGGTTGAAGTTTCAG <u>G</u>	(Busche <i>et al.</i> , 2012)	PEX	
cg2102		<i>sigB</i>	1990877	417,25	25	CTTGGGAACTTTTTTGGA-AGCAGTCCGTTGAACCTCTT <u>G</u>	(Halgasova <i>et al.</i> , 2001)	S1-Nuklease-mapping	
cg0497		<i>hemA</i>	438316	394,51	199	CCGGAATTAAACTCACTGAGCGAAA TGGTTGATCAAGAT <u>G</u> C	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg0184		-	158815	39,98	55	CTGGGAATGTTTCATGATCTTCGAGGGGTTGTAGGAATC <u>G</u>	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg3344	cgR_2903	-	3191301	500,55	0	TGCGGAACAAATTTGAAGGTTTTTCAGTTGCTATAGGT <u>A</u>	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg2194		<i>mtr</i>	2086894	30,16	72	CTTGGGAAGCTTTTTATAGTCATATGCGTTGAGATACGT <u>G</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg3236		<i>msrA</i>	3099488	521,62	36	CCC GGAATA AATTGGCAG-CTAAGTAGGGTTGAAGGGCATA <u>AA</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg2838	cgR_2470	<i>dsbA1</i>	2699850	68,7	46	CGCGGAATAGTTAGCGG-GCGCATGTGGTTGGGATATATGA <u>A</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg2762	cgR_2417	<i>murl</i>	2631352	34,47	233	GAGGAATTATTTGGTGCAACGAGGCTGTTGAACAACAAG <u>GA</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg2214	cgR_1848	-	2105774	95,75	173	TCGGGAATATTTTTTCTCCAAGCGGGTTAACTCACTGT <u>GT</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg2106	cgR_1753	-	1995404	23,7	67	TAGGGATTATTTTCGAG-GGTCCCCAGTTGTACTGAGCA <u>AG</u>	(Ehira <i>et al.</i> , 2009)*		diese Arbeit (RNA-Seq)***
cg1689		<i>pup</i>	1576737	35,67	134	GATGGATTAAACGTTAAC-AATAAGTTTGTACATAGTGT <u>GG</u>	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg0378		-	329842	825,64	92	GATGGGAACATTTTTGAA-GAATACCTCGTTGAATCTAGT <u>GC</u>	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***

Fortsetzung Tab. IV.3

Kodier-region (CDS)	Alternative CDS	Gen-name	Position Genom	relative Anzahl mappende Reads	Abstand Start-codon	Sequenz (-39 -- +1)	Referenz	TSP-Identifizierungsmethode	Neue Referenz ^a (TSP-Identifizierung)
cg2835	-		2695784	322,21	56	TAGGAAATAAAACAGGGTGTCTTTG TGTTCAA AAGGTAT A	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg2206	<i>ispG</i>		2098676	9,81	68	TAC CGGAATT GAACTGCC-GGCACCGAT TGTTAAA AGAG G TGA	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg2423	<i>lipA</i>		2313882	277,21	94	G AGGAATT TTGTCCGGTGGGGCGCCT CGTTGAA GCGAAGT A	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***
cg0699	<i>guaB2</i>		616737	92,4	60	TG TGGAAT AGCGTGTGAGCT-GCAGC GGTTAT CCGATTT AA	(Busche <i>et al.</i> , 2012)*		diese Arbeit (RNA-Seq)***

* vorhergesagter TSP.

^a Neue Referenz: bei Korrektur eines bereits experimentell identifizierten TSP oder experimenteller Bestätigung vorhergesagter TSP.

** Bei Korrektur eines bereits experimentell identifizierten TSP bleibt das Originalzitat bestehen und zusätzlich diese Arbeit zitiert.

*** Bei experimenteller Bestätigung eines vorhergesagten TSP durch RNA-Seq wird das Originalzitat durch die neue Referenz ersetzt.

Transkriptionsstartpunkte (TSP): **fett geschrieben und unterstrichen.**

Durch RNA-Seq korrigierte nicht mehr gültige TSP: **grau und fett geschrieben.**

Durch RNA-Seq korrigierte oder neu bestimmte TSP: **rot und fett geschrieben und unterstrichen.**

-10- und -35-Promotorregion **fett geschrieben.**

Zu den nicht identifizierten TSP zählen die der Promotoren der Gene *arnA* (Zemanova *et al.*, 2008), *trxC*, *cgR_2903* (*cg3344*), *cg2810* (Ehira *et al.*, 2009), *zwf*, *cg2434* (Busche *et al.*, 2012) und *clpP1* (Engels *et al.*, 2004). Bei *arnA* handelt es sich um eine kleine Antisense-RNA (Zemanova *et al.*, 2008), deren Transkript bei der hier verwendeten Methode der RNA-Isolierung vermutlich nicht im finalen RNA-Pool zur Erstellung der cDNA-Bibliotheken vorhanden war. Die weiteren Gene *zwf*, *cg2434*, *cg2810*, *trxC* und *clpP1*, für die kein SigH-abhängiger TSP identifiziert werden konnte, wiesen bereits im *Microarray* RES167 (Referenz) gegen RES167 $\Delta rshA$ nur eine geringere oder keine Änderung in ihrer mRNA Menge auf (Tab. IV.1). Die geringe Änderung der Transkriptmenge dieser Gene lässt sich wohl am besten mit der negativen Kontrolle zusätzlicher Regulationsysteme in Abwesenheit des spezifischen Stresses erklären und hat zur Folge, dass die entsprechenden Promotoren im RES167 $\Delta rshA$ -Stamm nicht voll aktiv sind. Hierdurch ist die Anzahl der SigH-abhängigen Transkripte dieser Gene zu gering, um per RNA-Seq-Analyse einen TSP identifizieren zu können. Für das Gen *cg3344* konnte der in dieser Arbeit vorhergesagte Promotor über die Identifizierung des stromabwärts liegenden TSP identifiziert werden, nicht aber der durch Ehira und Mitarbeiter vorhergesagte.

2.6. Einige SigH-abhängige Promotoren sind in dem Deletionsstamm RES167 $\Delta sigH$ weiterhin aktiv

Um sicher zu stellen, dass die TSP, die in Kapitel IV.2.5 durch die Auswertung der RNA-Seq-Daten des Deletionsstammes RES167 $\Delta rshA$ identifiziert wurden, exklusiv auf die SigH-vermittelte Transkription zurückzuführen sind, wurde überprüft, ob diese TSP im RNA-Seq-Datensatz der Deletionsmutante RES167 $\Delta sigH$ vorhanden sind (Tab. IV.2). Es wurde erwartet, dass die TSP durch die Deletion des Sigmafaktors SigH in diesem Datensatz nicht identifiziert werden können. Die Mehrheit der TSP konnte im RNA-Seq-Datensatz von RES167 $\Delta sigH$, wie erwartet, nicht gefunden werden. Dennoch konnten zehn mutmaßlich SigH-abhängige TSP in beiden RNA-Seq-Datensätzen identifiziert werden und weisen auf eine bestehende Aktivität der SigH-abhängigen Promotoren in Abwesenheit von SigH hin (Tab. IV.4). Hierbei handelt es sich um die SigH-abhängigen TSP der Gene *cg3344*, *cg0378*, *trxB*, *uvrA* und *sigB*, sowie der an der Hitzestressantwort beteiligten Gene *dnaJ2*, *clpB*, *dnaK*, *clpC* und *clgR*. Die relative Anzahl der *mappingen Reads* an den identifizierten TSP unterschied sich zwischen den beiden Datensätzen teils stark. Bis auf den TSP von *sigB* wurden für alle TSP mehr *Reads* in der RES167 $\Delta rshA$ Mutante *gemappt* als in der RES167 $\Delta sigH$ Mutante. Dies ist gleichbedeutend mit einer erhöhten Aktivität der Promotoren. Das Verhältnis der relativen Anzahl *mappender Reads* der beiden Datensätze schwankte hierbei von ca. 10 für *cg3344* bis zu 1,2 für *uvrA* (Tab. IV.4). Um zu überprüfen, ob sich die Transkriptionsinitiation auch im Referenzstamm RES167 von der im Deletionsderivat $\Delta sigH$ an den SigH-abhängigen TSP unterscheidet, wurden auch hier die entsprechenden RNA-Seq-Datensätze verglichen.

Tab. IV.4 Promotoren SigH-abhängiger Gene die im Deletionsstamm RES167 $\Delta sigH$ aktiv sind.

Kodier-region	Gen-name	Position Genom	relative Anzahl mappende Reads $\Delta rshA$	relative Anzahl mappende Reads $\Delta sigH$	Verhältniss $\Delta rshA/\Delta sigH$	Abstand Startcodon	Sequenz (-39 -- +1)	Referenz ^a
cg3344	-	3191301	500,55	52,81	9,5	0	TGCC GGGAACA AAATTTGAAG-GTTTTTC AGTTGCT TATAGGT A	diese Arbeit
cg0378	-	329842	825,64	110,61	7,5	92	GAT TGGAAACA TTTTTGAA-GAATACCT CGTTGAA TCTAGT G	diese Arbeit
cg2515	dnaJ2	2395613	519,46	102,71	5,1	93	TC GGGAACA AATTTCTAAGGTGTCCGCC CGTTGCT AAGTAT G	(Busche <i>et al.</i> , 2012)
cg3079	clpB	2939558	84,98	23,29	3,6	91	GAG TGGAAACA TACTCAACT-C ^{TTT} GTG CGTTATA GATAT T	(Ehira <i>et al.</i> , 2009)
cg3100	dnaK	2959823	309,76	88,57	3,5	120	AGT TGGAAACA ACTTTGTGG-CATTTAC CGTTGCT TATATAT G	(Ehira <i>et al.</i> , 2009)*
cg2963	clpC	2820286	136,69	51,98	2,6	38	AGT TGGAAAGT TTTGCCCA--ATAAGGG CGTTAAA GTGGGT G	(Engels <i>et al.</i> , 2004)*
cg2152	clgR	2039986	154,16	67,36	2,3	1	ACT GGGAACA AAATTTTAGG-GAAAGGG AGTTGAA CCTAAC G	(Engels <i>et al.</i> , 2004)*
cg3422	trxB	3273611	305,93	171,73	1,8	68	TGAT TGGAAAGT TTTTCAAAG-TGTCTGAC CGTTGAA AACGGT G	(Ehira <i>et al.</i> , 2009)
cg1560	uvrA	1443908	30,64	25,78	1,2	46	ACTAG CGGAACA CC ^{TT} TGTTTC--GATGCTAT TGTTCGA AGGT G	(Busche <i>et al.</i> , 2012)
cg2102	sigB	1990877	417,25	451,16	0,9	25	CTT GGGAAC TTTTGTGGA-AGCAGTCC CGTTGAA CCTCTT G	(Halgasova <i>et al.</i> , 2001)

^aIdentifikation des Promotors

*Position des TSP in dieser Arbeit korrigiert

Transkriptionsstartpunkte (TSP): **fett geschrieben und unterstrichen**

-10 und -35 Promotorregion **fett geschrieben**

Erweiterte -35 (T-exklusiv) *kursiv geschrieben*

Die sechste Position der -35 Region **rot und fett geschrieben**

In **Abb. IV-13** ist die Verteilung der *mappenden Reads* für die RNA-Seq-Datensätze der Stämme RES167, RES167 $\Delta rshA$ und RES167 $\Delta sigH$ schematisch für die stromaufwärts liegende Region des Gens *dnaK* dargestellt. Die Transkription des Gens *dnaK* wird von zwei sich überlappenden Promotoren unterschiedlicher Klassen gesteuert. Der P1 Promotor wird von dem *Housekeeping*-Sigmafaktor SigA erkannt und der P2 Promotor von dem ECF-Sigmafaktor SigH (Pátek *et al.*, 2013). Während die Deletion des *sigH*-Gens zu einer Verringerung der *gemappten Reads* am 5'-Ende des SigH-abhängigen Transkripts (TSP 2) führte, zeigte die Deletion des Anti-Sigmafaktors RshA eine drastische Erhöhung der *mappenden Reads* an dieser Position. Dies weist zum einen auf die Abhängigkeit der Transkription von SigH an diesem Promotor hin, zum anderen aber auch auf die Erkennung durch einen weiteren Sigmafaktor, da trotz Abwesenheit von SigH die Transkription nicht völlig ausgeschaltet war. Die relative Anzahl *mappender Reads* an der Position des SigA-abhängigen TSP (TSP1) in den RNA-Seq-Datensätzen RES167 $\Delta sigH$ und RES167 blieb unverändert. Die erhöhte Menge an *mappenden Reads* für den Datensatz RES167 $\Delta rshA$ an dieser Position ist ein Resultat der Überlagerung und damit der Addition der *Reads*, die zugehörig zum 5'-Ende der SigA-abhängigen und der SigH-abhängigen Transkripte sind.

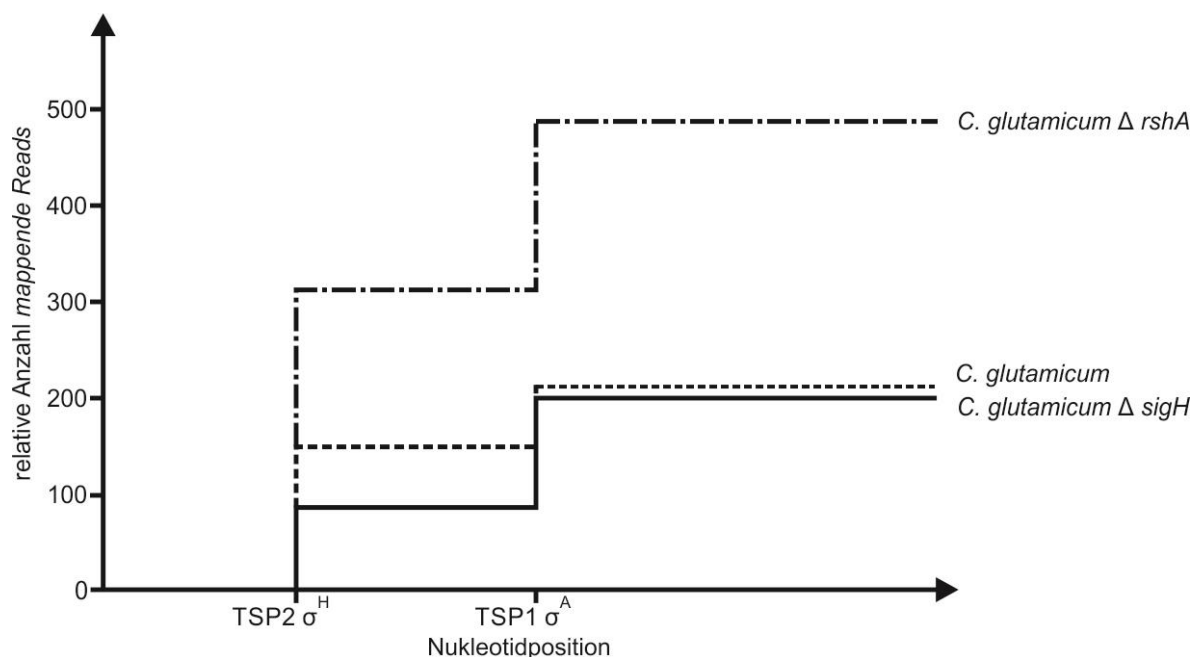


Abb. IV-13 Profile der *mappenden Reads* im Bereich der zwei TSP des *C. glutamicum* Gens *dnaK*. Die relative Anzahl der *mappenden Reads* der RNA-Seq-Datensätze der Stämme RES167, $\Delta rshA$ und $\Delta sigH$, die an den 5'-Enden der SigA- bzw. SigH-abhängigen Transkripte *mappen* ist schematisch dargestellt (TSP1 σ^A ; TSP2 σ^H). Die y-Achse gibt die relative Transkriptionseffizienz als die relative Anzahl mappender *Reads* (Anzahl der *Reads* normalisiert auf die Gesamtanzahl der *mappenden Reads* eines RNA-Seq-Datensatzes) an. Die Deletionsmutante $\Delta rshA$ (mit fehlendem Anti-Sigmafaktor RshA) weist eine erhöhte Aktivität des SigH-spezifischen Promotors P2 auf, während die Aktivität des SigA-getriebenen Promotors P1 in der Deletionsmutante $\Delta sigH$ und $\Delta rshA$ nicht beeinflusst ist (die erhöhte relative Anzahl an *mappenden Reads* an TSP2 im RNA-Seq-Datensatz von $\Delta rshA$ ist nur das Resultat der Addition von *Reads* an den Positionen TSP1 und TSP2). Abb. nach (Pátek, Holatko, [Busche et al.](#), 2013)

Es bleibt zu klären, welcher Sigmafaktor für die Erhaltung der Transkription an den genannten SigH-abhängigen Promotoren bei Deletion des ECF-Sigmafaktors SigH ist.

2.7. Die aktiven SigH-abhängigen Promotoren in der Deletionsmutante $\Delta sigH$ unterscheiden sich am 3'-Ende –35-Region von den übrigen SigH-abhängigen Promotoren

Bei dem Vergleich der –10- und –35-Region der Promotoren, von denen keine Transkription in der $\Delta sigH$ -Mutante ausgeht, und denen, die weiterhin aktiv sind (Tab. IV.4), fiel auf, dass sich die –10 Region nicht signifikant unterscheidet.

Die –35-Region unterschied sich jedoch an der sechsten Position des –35-Motifs. Während die meisten Promotoren, die keine Aktivität in RES167 $\Delta sigH$ aufweisen, an Position 6 des –35-Motifs das Nukleotid G oder T besitzen (Tab. IV.3), befindet sich bei den Promotoren, die in der RES167 $\Delta sigH$ Mutante weiterhin aktiv sind, das Nukleotid C dieser an dieser Stelle der Promotorregion (Tab. IV.4).

Eine Ausnahme stellen die Promotoren P_{trxB} und P_{clpC} dar. Sie besitzen, wie auch die Promotoren, die in der $\Delta sigH$ -Mutante nicht aktiv sind, ein G an Position 6 der –35-Region, aber unterscheiden sich von diesen in der erweiterten –35-Region, die exklusiv aus Thyminen besteht (Tab. IV.4).

3. Entwicklung eines genomweiten *in vitro* Transkriptions/Sequenzierungsansatzes (ROSE) und Anwendung auf die Regulationsnetzwerke von SigH und SigE in *Corynebacterium glutamicum*

Ein intrinsisches Problem bei der Enschlüsselung von hierarchischen und überlappenden Regulationsnetzwerken von Sigmafaktoren liegt in der *in vivo* Analyse von Mutanten oder Stressbedingungen, um regulierte Gene zu identifizieren (*Top-Down*). Es ist durch die klassisch angewandten Methoden nicht möglich, zwischen primären und sekundären Effekten der Stressbedingungen oder Mutationen zu unterscheiden. Um dieses Problem zu umgehen, wurde seit Jahrzehnten zusätzlich die *in vitro* Transkriptionsanalyse von einzelnen Kandidatengen-Promotoren eingesetzt.

Im Folgenden wurde dieser Ansatz auf ein genomweites Level ausgeweitet. Hierzu wurden die bereits *in vivo* erzielten Ergebnisse (*Top-Down*) zur Untersuchung des SigH-Regulationsnetzwerkes mit dem komplementären *in vitro* Ansatz einer genomweiten *in vitro* Transkriptionsanalyse mit verschiedenen aufgereinigten ECF-Sigmafaktoren und dem RNAP-Core Enzym (*Bottom-Up*) kombiniert.

Im Detail wurde die bereits existierende genomweite *in vitro* Transkriptionsanalyse, ROMA (*Run-Off Transcription/Microarray Analysis*) (Maclellan *et al.*, 2009), auf ein höher auflösendes Level ausgeweitet. Bei ROMA wird aufgereinigte RNA Polymerase mit und ohne aufgereinigten Sigmafaktoren eingesetzt, um Transkriptionsprodukte auf Basis genomischer DNA zu identifizieren. Die *in vitro* transkribierte RNA wird mit Hilfe von *Microarrays* untersucht und die differentiell transkribierten Gene in Folgeanalysen einzeln auf TSP überprüft.

Die hier durchgeführte Weiterentwicklung der Methode bestand darin, die Analyse der RNAs anstelle von *Microarrays* mit RNA-Seq von nativen 5'-Enden mRNAs durchzuführen. Diese weiterentwickelte Methode wurde in Anlehnung an den Vorgänger als ROSE (*Run-Off Transcription-Sequencing*) bezeichnet. Die Nutzung der RNA-Seq als Analysemethode ermöglicht die Einzelbasen-genaue Auflösung der für den eingesetzten Sigmafaktor spezifischen TSP auf einem genomweiten Level.

3.1. Die ECF-Sigmafaktoren SigC, SigD, SigE, SigH und der *Housekeeping*-Sigmafaktor SigA konnten als native Proteine für die Nutzung in ROSE aufgereinigt werden

Die ECF-Sigmafaktoren SigC, SigD, SigE und SigH wurden neben dem *Housekeeping*-Sigmafaktor SigA mittels *IMPACT*-System (*Intein Mediated Purification with an Affinity Chitin-binding Tag*) (Kapitel III.13.1) *Tag*-frei aufgereinigt. Hierbei werden die Proteine mit einem selbstspaltenden *Intein-Tag* und einer Chitinbindedomäne versehen und über eine Chitinsäule aufgereinigt. Die Abspaltung des *Tags* erfolgt über eine induzierte Selbstspaltung des *Inteins* durch Dithiothreitol. Die Sigmafaktoren wurden sowohl über den N-Terminus als auch über den C-Terminus aufgereinigt um zu bestimmen, ob es Unterschiede bei der Verwendung der *Tags* in Bezug auf Menge und Reinheit des isolierten Sigmafaktors gibt.

Es konnten keine Unterschiede zwischen der Verwendung eines C- oder N-terminalen *Tags* gefunden werden. Die Konzentrationen schwankten bei N-terminal-aufgereinigten Sigmafaktoren zwischen 13,83 mg/mL für SigD und 0,61 mg/mL für SigH (Tab. IV.5). Daten zu der C-terminalen Aufreinigung sind nicht gezeigt.

Tab. IV.5 Konzentrationen der aufgereinigten Sigmafaktoren (N-terminal) per *IMPACT*-System

Sigmafaktor	Konzentration [mg/mL]
SigA	2,06
SigC	0,75
SigD	13,83
SigE	2,57
SigH	0,61

Zur Identifikation und der Überprüfung auf Reinheit wurden die aufgereinigten Sigmafaktoren zunächst per SDS-PAGE aufgetrennt (Abb. IV-14). Die Sigmafaktoren konnten ohne Verunreinigungen hochrein isoliert werden.

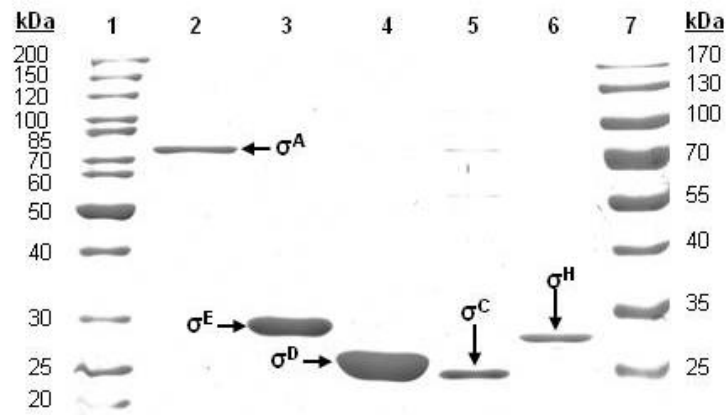


Abb. IV-14 Isolierte *C. glutamicum* Sigmafaktoren (Coomassie-Blau gefärbte SDS-PAGE). Die Aufreinigung erfolgte mittels *IMPACT*-System (New England Biolabs, USA) aus *E. coli* ER2566 Zellen. Spuren: 1 Proteingrößenmarker; 2 SigA (5 µg); 3 SigE (10 µg); 4 SigD (25 µg); 5 SigC (5 µg); 6 SigH (5 µg); 7 Proteingrößenmarker. Die molekulare Größe ist in Kilodalton (kD) links und rechts neben der SDS-PAGE angegeben.

Da die Größenbestimmung der Sigmafaktoren zum Teil zu einer höheren molekularen Masse führte als erwartet (SigA 55,5 kD; SigE 24,2 kD; SigD 20,7 kD; SigC 21,4 kD; SigH 23,2 kD), wurde die Identität der Sigmafaktorproteine durch eine massenspektrometrische Analyse (*peptide mass fingerprinting*) mittels MALDI-TOF-MS überprüft. Die Identität der einzelnen Sigmafaktoren konnte eindeutig bestätigt werden, da ausreichend Peptidmassen identifiziert werden konnten, um die jeweiligen mit einer hohen Wahrscheinlichkeit von mehr als 99,95 % zu identifizieren. Die ECF-Sigmafaktoren konnten nun für die *in vitro* Transkription eingesetzt werden.

3.2. Direkte Identifizierung von exklusiv und partiell SigH-abhängigen Promotoren und Nachweis der direkten Kontrolle durch SigH

In den vorhergegangenen *Top-Down*-Untersuchungen konnte nicht direkt zwischen primären und sekundären Effekten der Deletion von *rshA* unterschieden werden. Die SigH-abhängigen Gene in dieser und vorangegangenen Arbeiten wurden über die stromaufwärts liegenden definierten SigH-abhängigen Promotoren und entsprechender Sigmafaktor- und Anti-Sigmafaktor-Mutanten identifiziert (Busche *et al.*, 2012; Ehira *et al.*, 2009). Einige der bekannten SigH-abhängigen Gene konnten in der Analyse der Anti-Sigmafaktormutante RES167 $\Delta rshA$ durch die zusätzliche Regulation weiterer Transkriptionsfaktoren in dem (bewusst eingesetzten) ungestressten Zustand nicht als differentiell transkribiert identifiziert werden (Busche *et al.*, 2012).

Um zu bestätigen, dass die Transkription der identifizierten SigH-abhängigen Promotoren der tatsächlich primäre Effekt des ECF-Sigmafaktors SigH ist und die im ungestressten Zustand in der Anti-Sigmafaktormutante RES167 $\Delta rshA$ nicht differentiell transkribierten SigH-abhängigen Gene auch SigH-abhängig transkribiert werden, wurden ROSE-Experimente unter Verwendung des ECF-Sigmafaktors SigH durchgeführt. Durch die Rekonstruktion eines Sigmafaktornetzwerks *in vitro* aus aufgereinigten Einzelkomponenten gelingt die Ausblendung sekundärer Effekte. Als Negativkontrolle wurden Ansätze mit RNAP-Core Enzym ohne hinzugefügten Sigmafaktor mitgeführt.

Die *Run-Off*-Transkriptionen wurden mit *EcoRI* geschnittener *C. glutamicum* ATCC 13032 gDNA als *Template* durchgeführt. Es wurden jeweils zehn unabhängige *Run-Off*-Transkriptionen mit RNAP-Core Enzym (*E. coli*), ohne Zugabe eines Sigmafaktors, und mit SigHI bzw. SigHII (über N-terminalen bzw. C-terminalen *Tags* aufgereinigt) gesättigter RNAP, wie in Material und Methoden beschrieben, durchgeführt. Nach der Isolierung der RNA wurden die zehn unabhängigen Proben vereinigt und jeweils 1 μ g RNA für die Erstellung von cDNA-Bibliotheken eingesetzt.

Die Sequenzierung der cDNA-Bibliotheken resultierte in 100.000 – 200.000 *Reads* für die *Run-Off*-Transkription mit der RNAP-SigHI und RNAP-SigHII. Die Anzahl an sequenzierten *Reads* lag unter der Verwendung des RNAP-Core Enzyms (Negativkontrolle) bei 700.000. Die Durchführung des *Mapping* auf das Referenzgenom *C. glutamicum* ATCC 13032 (Kalinowski *et al.*, 2003) mittels SARUMAN (Jochen Blom *et al.*, 2011) führte für RNAP-SigHI zu 152.000, für RNAP-SigHII zu 70.622 und für das RNAP-Core Enzym zu 632.778 *gemappten Reads*. Insgesamt umfasst der ROSE-Datensatz von SigH 252.341 *mappende Reads* (Datensätze von SigHI und SigHII zusammengefasst) und der Datensatz von RNAP-Core Enzym 721.942

mappende Reads. Die genaue Aufschlüsselung in einfach *mappende* und mehrfach *mappende Reads* ist in Tabelle Tab. IV.6 aufgeführt.

Tab. IV.6 Mapping-Statistik der ROSE-Experimente

RNA-Seq-Lauf	Gesamtzahl Mappings	Einfach <i>mappende Reads</i>	Mehrfach <i>mappende Reads</i>	Gesamtzahl <i>mappender cDNA-Reads</i>
ROSE SigH_I	172949	144952	7048	152000
ROSE SigH_II	79392	67489	3133	70622
ROSE SigE	110175	82456	6010	88466
ROSE core-RNAP	721942	605123	27655	632778

Die erstellten ROSE-Datensätze wurden für die Identifizierung von TSP genutzt. Aufgrund der hohen Spezifität dieser *in vitro* Datensätze in Bezug auf die auf SigH-Abhängigkeit beschränkten TSP bzw. Promotoren war es hier möglich die Identifikation der TSP mit einer sensitiven Einstellung des verwendeten Skripts durchzuführen. Als TSP wurden alle Positionen angenommen, an denen mindestens vier Startpunkte kumuliert sind und an der Position direkt stromaufwärts zweifach weniger Startpunkte auftreten

Die durch die Analyse der ROSE-Daten identifizierten TSP wurden mit den TSP der Transkriptomanalyse der Deletionsstämme RES167 $\Delta sigH$ und RES167 $\Delta rshA$ abgeglichen. Beim Abgleich einer Reihe in dieser Arbeit identifizierten SigH-abhängigen TSP mit den durch die ROSE-Analyse mit dem ECF-Sigmafaktor SigH (ROSE-SigH) identifizierten TSP konnte für die meisten TSP eine exakte Übereinstimmung zwischen den *in vitro* und *in vivo* Daten festgestellt werden. In dem ROSE-SigH-Datensatz konnten, wie erwartet, spezifisch nur die TSP zugehörig zu SigH-abhängigen Promotoren identifiziert werden. Als Beispiel sollen hier die zwei TSP bzw. Promotoren des Gens *dnaK* dienen, welche bereits in IV.2.6 und einem Übersichtsartikel (Pátek *et al.*, 2013) auf Basis der RNA-Seq-Datensätze RES167 $\Delta rshA$ und RES167 $\Delta sigH$ als SigH- bzw. SigA-abhängig beschrieben worden sind. In den *in vivo* RNA-Seq-Datensätzen (RES167 $\Delta rshA$ und RES167 $\Delta sigH$) konnten zwei TSP stromaufwärts von *dnaK* identifiziert werden, wobei die Transkription an dem SigH-abhängige TSP in der Deletionsmutante RES167 $\Delta rshA$ verstärkt ist (Abb. IV-13 und Abb. IV-15). Ausschließlich dieser TSP konnte in dem *in vitro* Datensatz ROSE-SigH identifiziert werden (Abb. IV-15). Dies bestätigt die erwartete Spezifität des ROSE-Ansatzes und zeigt die basengenaue Übereinstimmung von *in vitro* und *in vivo* Daten.

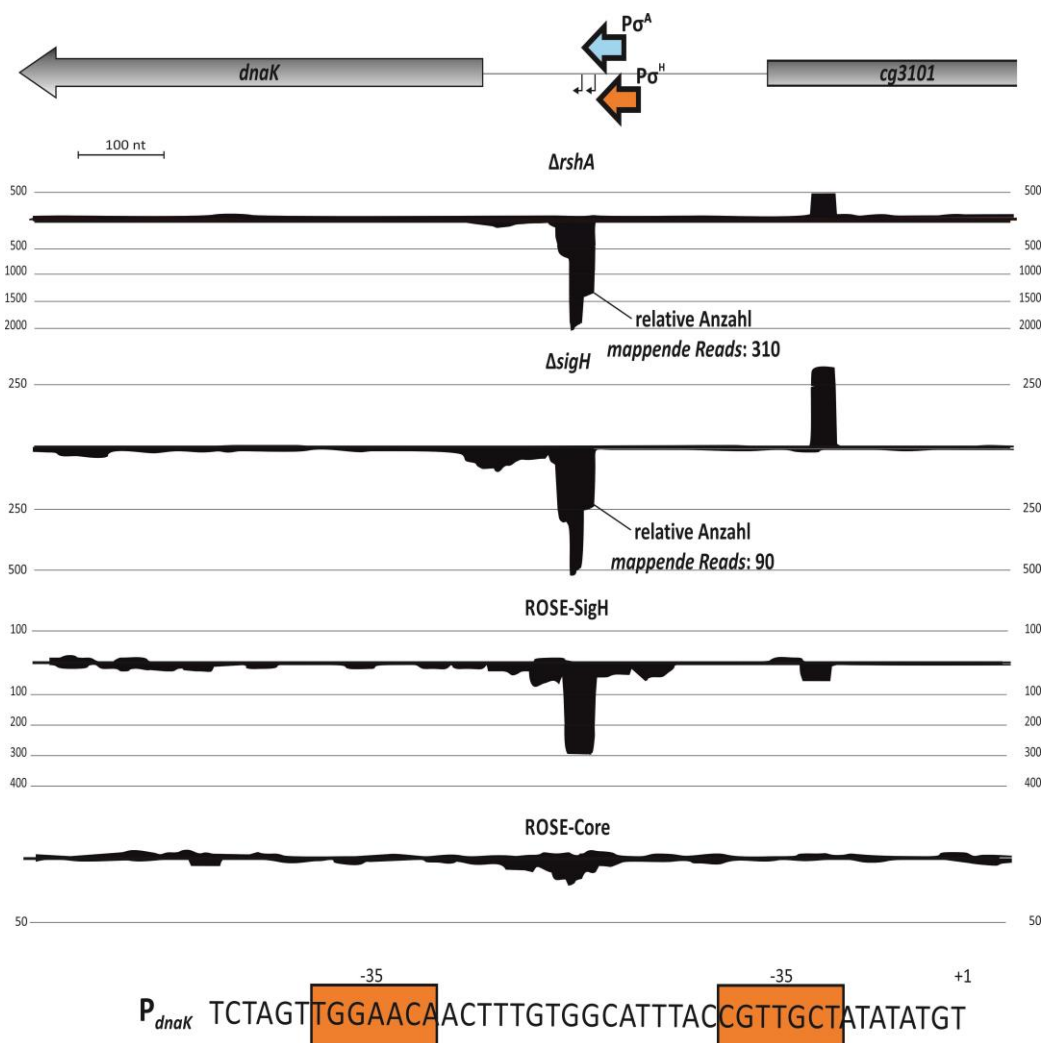


Abb. IV-15 Visualisierung der nativen 5'-Enden RNA-Seq Datensätze der *dnaK*-Genregion der *C. glutamicum* RES167 Deletionsderivate $\Delta rshA$, $\Delta sigH$ und der *in vitro* Datensätze von ROSE-SigH und ROSE-Core (Kontrolle). Die genomische Region des Gens *dnaK* ist oberhalb der Tracks schematisch abgebildet. Die *mappenden Reads* in den unterschiedliche Datensätzen sind schwarz dargestellt und die Tracks nach Datensatz benannt. TSP wurden durch einen Anstieg der Anzahl *mappender Reads* an einer Position identifiziert. TSP sind in der schematischen Darstellung der genomischen Region durch einen geknickten Pfeil gekennzeichnet. In der Negativkontrolle ROSE-Core konnte kein TSP identifiziert werden. Zwei identische TSP konnten in Orientierung des *dnaK*-Gens in den Datensätzen $\Delta rshA$ und $\Delta sigH$, aber nicht in dem *in vitro* Datensatz ROSE-SigH, der spezifisch für SigH ist, identifiziert werden. Der entsprechende SigA-abhängige Promotor (P_{σ^A}) wurde in der genomischen Region durch einen blauen Pfeil dargestellt und benannt. Ein weiterer TSP konnte ebenfalls in den *in vivo* Transkriptom Datensatz von $\Delta rshA$ und $\Delta sigH$ und zusätzlich in den *in vitro* Datensätzen ROSE-SigH identifiziert werden. Dies zeigt, dass SigH für die Erkennung dieses Promotors verantwortlich ist und partiell für die Initiierung der Transkription notwendig ist. Die Reduzierung der relativen Anzahl *mappender Reads* an diesem TSP im Vergleich von $\Delta rshA$ zu $\Delta sigH$ ist auf die Deletion des ECF-Sigmafaktors SigH zurück zuführen. Der entsprechende partiell SigH-abhängige Promotor ist durch einen orangenen Pfeil dargestellt und benannt (P_{σ^H}) und dessen Sequenz am unteren Ende der Abbildung dargestellt. Die -35- und -10 -Region des Promotors sind mit orangenen Boxen hervorgehoben und benannt. Der TSP ist in der Sequenz durch +1 gekennzeichnet. Die Visualisierung der RNA-Seq-Datensätze wurde durch VAMP (Hilker *et al.*, 2013, in Vorbereitung) verwirklicht.

Neben der exakten Übereinstimmung der Position der *in vivo* und *in vitro* identifizierten TSP konnten in Einzelfällen eine Abweichung der TSP-Positionen von 1 - 2 bp festgestellt werden, was keinen Einfluss auf die Identifikation des Promotors hat, aber bei dem kompletten Abgleich der identifizierten TSP der *in vivo* und *in vitro* Datensätze berücksichtigt wurde. Die Positionen der ROSE-SigH-TSP wurden ± 2 Positionen mit den Positionen der TSP der Transkriptomanalysen der Deletionsmutanten RES167 $\Delta sigH$ und RES167 $\Delta rshA$ verglichen. Für TSP der Mutante RES167 $\Delta rshA$ wurden nur solche für den Abgleich verwendet, die durch mehr als zehn *mappende Reads* (relativ) identifiziert worden sind. Es konnte davon ausgegangen werden, dass diese nativen 5'-Enden durch das Fehlen des Anti-Sigmafaktors RshA vermehrt transkribiert werden. Für die bereits bekannten Gene lag dieser Wert in 98 % der Fälle über zehn (vgl. Tab. IV.3).

Es wurden zwei unterschiedliche Vergleiche durchgeführt:

- I) Die Position eines identifizierten TSP im ROSE-Datensatz SigH musste identisch mit der Position eines identifizierten TSP im RNA-Seq-Datensatz RES167 $\Delta rshA$ sein, durfte zusätzlich aber nicht in dem RNA-Seq-Datensatz RES167 $\Delta sigH$ vorhanden sein.
- II) Ein TSP musste in den Datensätzen ROSE-SigH, RNA-Seq-Datensatz RES167 $\Delta rshA$ und RNA-Seq-Datensatz RES167 $\Delta sigH$ identifiziert werden.

Mit dem Vergleich I) werden TSP identifiziert, die exklusiv SigH-abhängig sind, da sie zum einen nur in dem *in vivo* Datensatz RES167 $\Delta rshA$ (*Top-Down*) vorhanden sind und zum anderen im SigH-spezifischen *in vitro* RNA-Seq-Datensatz ROSE-SigH (*Bottom-Up*). Mit dem Vergleich II) werden TSP und damit Promotoren identifiziert die partiell SigH-abhängig sind, da die Transkription in RES167 $\Delta sigH$ erhalten bleibt. In jedem Fall wird durch die Verwendung von ROSE die tatsächliche primäre Aktion von SigH aufgedeckt und ausgeschlossen, dass es sich um sekundäre Effekte handeln könnte. Nach dem Abgleich I) und II) wurden die Listen der TSP in Bezug auf rRNA-Regionen, alternative Starts (aufeinander folgende TSP in einem Bereich von 6 bp) überprüft und bereinigt. Neben einer visuellen Kontrolle der TSP bzw. der Promotoren per VAMP und dem Abgleich mit dem ROSE-Core-RNAP-Datensatz wurden die Promotorensequenzen mittels der Software Improbizer (Ao *et al.*, 2004; www.soe.ucsc.edu/~kent/improbizer) auf konservierte Motive hin überprüft. Es wurden nur Promotoren akzeptiert, die dem bekannten SigH-abhängigen Promotorkonsensusmotiv entsprachen. Hierdurch sollte ausgeschlossen werden, dass es sich um falsch positive TSP handelt, die nicht nativen 5'-Enden entsprechen. Dies ist denkbar, wenn die isolierte mRNA sehr starke sekundär Strukturen aufweist, die den Abbau nicht nativer 5'-Enden durch die 5'-abhängige Terminator-Exonuklease im Verlauf der Erstellung der cDNA-Bibliothek für RNA-Seq (Kapitel III.12.8.5) verhindert.

Insgesamt konnten so 31 TSP und damit Promotoren bestimmt werden, die eindeutig und exklusiv von SigH abhängig sind (Tab. IV.7). Hiervon konnten 14 Promotoren erstmalig identifiziert werden. Insgesamt 17 dieser Promotoren waren bereits durch *Microarray*- oder RNA-Seq-Analysen von Deletionsmutanten des Sigmafaktors SigH bzw. des Anti-Sigmafaktors RshA als SigH-abhängig bekannt (Busche *et al.*, 2012; Ehira *et al.*, 2009; Nakunst *et al.*, 2007). Die Promotoren besitzen mit wenigen Ausnahmen (P_{mtr} , P_{2ctaC} und P_{cg2247}) als –35-Region die Sequenz g/tGGAAT und das konservierte –10-Motiv GTT. Auf diese Weise konnten hier erstmalig experimentell TSP und damit Promotoren durch die Nutzung des *Bottom-Up*-Ansatzes ROSE primär und ausschließlich SigH zugewiesen werden.

Tab. IV.7 Exklusiv SigH-abhängige Promotoren (Identifikation per ROSE-SigH)

Kodier-region	Gen-name	-35-Region	-10-Region	Vorhergesagte Funktion	Referenz
<i>cg1555</i>	<i>uvrD3</i>	TGGAATG	TGTTGTA	DNA/RNA Helikase	(Busche <i>et al.</i> , 2012)
<i>cg1709</i>	<i>mshC</i>	TGGAATA	GGTTGAA	Putative 1-D-Myo-Inositol-2-Amino-2-Deoxy-alpha-D-Glukopyranosid--L-Cystein Ligase	(Busche <i>et al.</i> , 2012)
<i>cg1127</i>	-	TGGAATG	CGTTGGA	Putative Mycothiol S-Konjugat Amidase	(Busche <i>et al.</i> , 2012)
<i>cg1560</i>	<i>uvrA</i>	GGGAATA	AATTGAA	Exconuklease ABC, ATPase Untereinheit A	(Busche <i>et al.</i> , 2012)
<i>cg2411</i>	-	GGGAATA	CGTTGAA	Konserviertes hypothetisches Protein	(Ehira <i>et al.</i> , 2009)
<i>cg0617</i>	-	GGGAATT	TGTTTAA	Hypothetisches Protein	(Ehira <i>et al.</i> , 2009)
<i>cg2078</i>	<i>msrB</i>	TGGAATT	GGTTGGG	Methionin Sulfoxid Reduktase	(Ehira <i>et al.</i> , 2009)
<i>cg3405</i>	-	GGGAATG	TGTTTCA	NADPH:Quinon Reduktase, Zn-abhängige Oxidoreduktase	(Ehira <i>et al.</i> , 2009)
<i>cg1765</i>	-	GGGAATG	TGTTGTC	Putatives regulatorisches Protein	(Ehira <i>et al.</i> , 2009)
<i>cg1375</i>	-	GGGATTA	TGTTGTT	Putatives Thioredoxin	(Ehira <i>et al.</i> , 2009)
<i>cg3299</i>	<i>trxB1</i>	GGGAATA	AGTTGGT	Thioredoxin (TRX)	(Ehira <i>et al.</i> , 2009)
<i>cg2214</i>	-	GGGAATA	GGTTTAA	Putative Fe-S-cluster Redox Enzym	diese Arbeit (RNAseq)
<i>cg2206</i>	<i>ispG</i>	CGGAATT	TGTTAAA	4-Hydroxy-3-Methylbut-2-en-1-yl Diphosphat Synthase	diese Arbeit (RNAseq)
<i>cg2194</i>	<i>mtr</i>	GGGAAGC	CGTTGAG	Putative NADPH-abhängige Mycothiol Reduktase	diese Arbeit (RNAseq)
<i>cg2835</i>	-	AGGAAAT	TGTTCAA	Putative Acetyltransferase	diese Arbeit (RNAseq)
<i>cg1689</i>	<i>pup</i>	TGGATTA	TGTTACA	Prokaryotisches Ubiquitin-ähnliches Protein	diese Arbeit (RNAseq)
<i>cg2423</i>	<i>lipA</i>	AGGAATT	CGTTGAA	Lipoylsynthetase	diese Arbeit (RNAseq)
<i>cg0699</i>	<i>guaB2</i>	TGGAATA	GGTTATC	IMP Dehydrogenase	diese Arbeit (ROSE-SigH)
<i>cg3236</i>	<i>msrA</i>	CGGAATA	GGTTGAA	Protein-Methionine-S-oxid Reduktase	diese Arbeit (ROSE-SigH)
<i>cg1432</i>	<i>livD</i>	GGGAATA	CGTTGAA	Dihydroxy-Säure Dehydratase	diese Arbeit (ROSE-SigH)
<i>cg0454</i>	-	GGGAATT	GGTTGTG	Putativer transkriptioneller Regulator, TetR-Familie	diese Arbeit (ROSE-SigH)
<i>as-cg3357</i>	-	TGGAATA	AGTTGTT	Antisense RNA zu <i>cg3357</i> ; <i>trpP</i> ; Permease	diese Arbeit (ROSE-SigH)
<i>as-cg2053</i>	-	TGGAATA	GGTTAAC	Antisense RNA zu <i>cg2052</i> ; Putatives sekretiertes Protein; Signalpeptid	diese Arbeit (ROSE-SigH)
<i>cg1282</i>	-	GGGAATG	CGTTGTA	Konserviertes hypothetisches Protein	diese Arbeit (ROSE-SigH)
<i>cg2409_P1</i>	<i>ctaC</i>	GGGAATT	GGTTGAT	Cytochrome C Oxidase Untereinheit II	diese Arbeit (ROSE-SigH)
<i>cg2263</i>	-	GGGAATT	TGTTATC	Hypothetisches Protein	diese Arbeit (ROSE-SigH)
<i>cg2409_P2</i>	<i>ctaC</i>	TGGAAGA	TGTTAAA	Cytochrome C Oxidase Untereinheit II	diese Arbeit (ROSE-SigH)

Fortsetzung Tab. IV.7

Kodier-region	Gen-name	-35-Region	-10-Region	Vorhergesagte Funktion	Referenz
<i>cg1095</i>	-	TGGAAAT	CGTTCTT	Hypothetisches Protein; verkürztes Protein durch SigH-abhängigen Start	diese Arbeit (ROSE-SigH)
<i>as-cg2375</i>	-	ATGAAAT	TGTTTGA	Antisense RNA zu <i>cg2375</i> ; <i>ftsI</i> ; Penicillin-bindendes Protein	diese Arbeit (ROSE-SigH)
<i>cg2247</i>	-	CGGAAGC	GGTTGAA	Hypothetisches Protein	diese Arbeit (ROSE-SigH)
<i>cg1386</i>	<i>fixA</i>	TTGAAAT	TGTTGAT	Putatives Elektron transfer-Flavoprotein, Beta Untereinheit	diese Arbeit (ROSE-SigH)

Hierunter befinden sich die Promotoren von Genen der SOS-Stressantwort, wie *uvrA* (Exonuklease) und *uvrD3* (DNA/RNA Helikase) (Abb. IV-16).

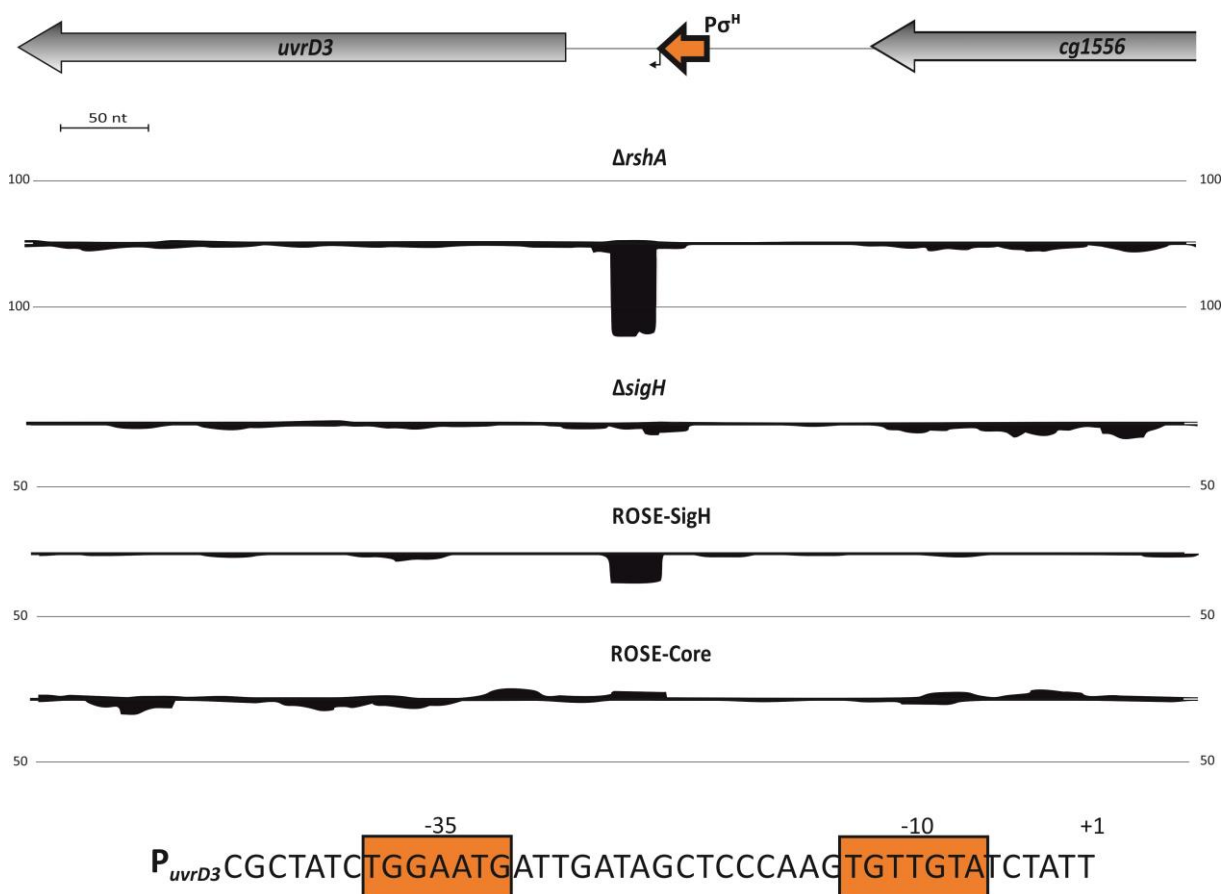


Abb. IV-16 Visualisierung der nativen 5'-Enden RNA-Seq-Datensätze im Bereich des Gens *uvrD3* der Stämme *C. glutamicum* RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der *in vitro* Datensätze von ROSE-SigH, und ROSE RNAP-Core (Kontrolle) Die genomische Region des Gens *uvrD3* ist oberhalb der Tracks abgebildet. Die *mappenden Reads* in den unterschiedlichen Datensätzen sind schwarz dargestellt und nach dem Datensatz benannt. TSP wurden durch einen Anstieg der Anzahl *mappender Reads* an einer Position identifiziert. Es konnte nur in dem Datensatz RES167 $\Delta rshA$ und dem *in vitro* Datensatz ROSE-SigH ein TSP identifiziert werden. So konnte bewiesen werden, dass SigH für die Transkriptioninitiation dieses TSP verantwortlich ist. Der TSP ist in der schematischen Darstellung der genomischen Region durch einen geknickten Pfeil gekennzeichnet. Der entsprechende SigH-abhängige Promotor ist durch einen orangenen Pfeil dargestellt und dessen Sequenz am unteren Ende der Abbildung dargestellt. Die -35- und -10-Region des Promotors sind mit orangenen Boxen hervorgehoben und benannt. Der TSP ist in der Sequenz durch +1 gekennzeichnet. Die Visualisierung der RNA-Seq-Datensätze wurde durch VAMP (Hilker *et al.*, 2013, in Vorbereitung) verwirklicht.

Weiter handelte es sich um die bereits bekannten SigH-abhängigen Promotoren der Gene, die für Proteine kodieren, die zum einen an der Mycothiolsynthese bzw. des Mycothiolrecyclings beteiligt sind (*mshC*, *mtr*, *cg1127*) und zum anderen Gene, die neben Mycothiol als Thioredoxine oder Oxidoreduktasen vermutlich an der Aufrechterhaltung des Redoxgleichgewichts in der Zelle unter Thiol-oxidierenden Bedingungen (z.B. Diamidstress) beteiligt sind. Hierzu zählen *trxB1* (Thioredoxin) (Abb. IV-17), welches zusätzlich einen SigA-abhängigen Promotor besitzt, *cg1375* (putatives Thioredoxin), *cg3405* (Oxidoreduktase), *msrB* (Methioninsulfoxid Reduktase) sowie *cg2214* (putatives Fe-S Cluster Redox-Enzym).

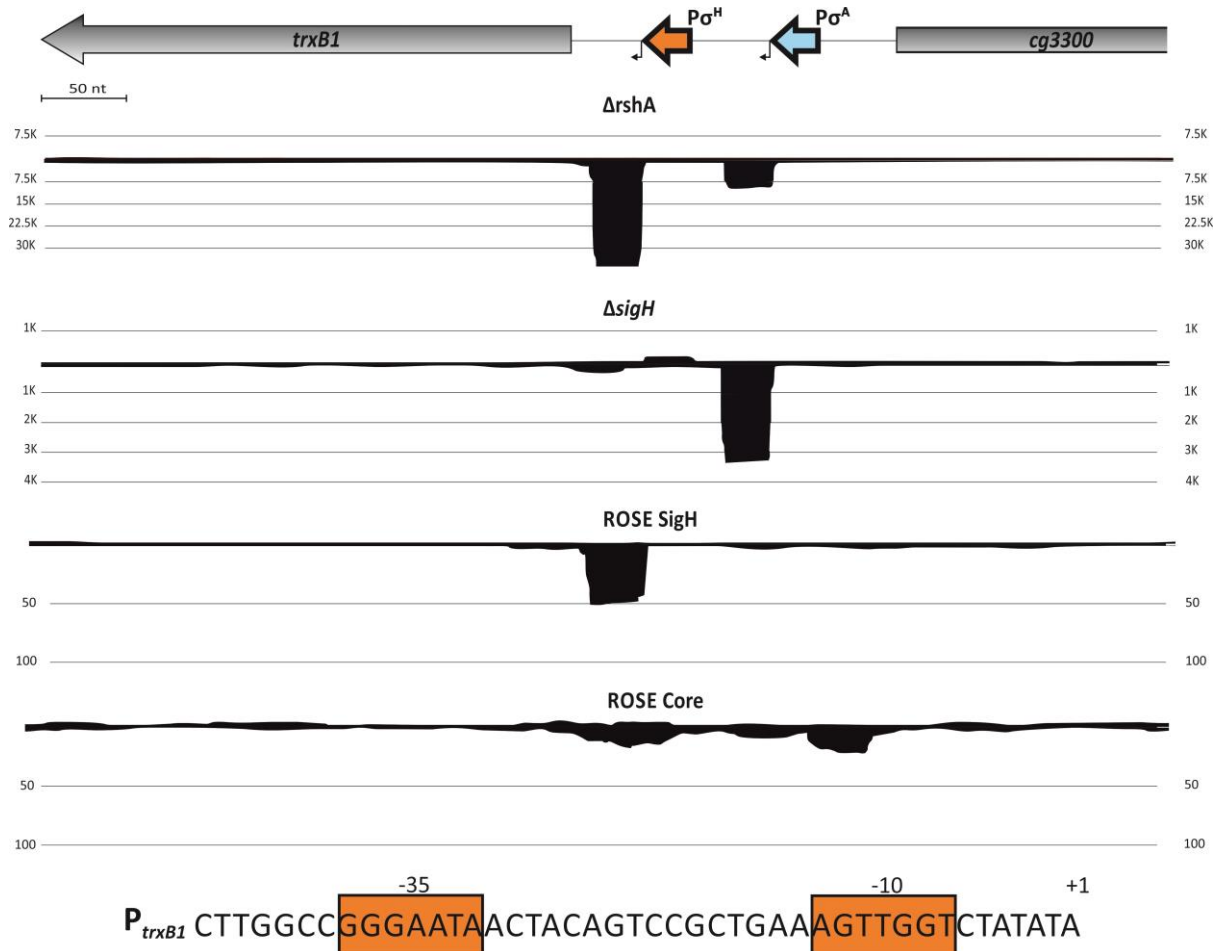


Abb. IV-17 Visualisierung der nativen 5'-Enden RNA-Seq-Datensätzen im Bereich des Gens *trxB1* von *in vivo*-Transkriptomen der Stämme *C. glutamicum* RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der *in vitro* Transkriptome von ROSE-SigH, and ROSE-Core (Kontrolle). Die genomische Region des Gens *trxB1* ist oberhalb der Tracks abgebildet. Die *mappenden Reads* in den unterschiedlichen Datensätzen sind schwarz dargestellt und nach dem Datensatz benannt. TSP wurden durch einen Anstieg der Anzahl *mappender Reads* an einer Position identifiziert. TSP sind in der schematischen Darstellung der genomischen Region durch einen geknickten Pfeil gekennzeichnet. In der Negativkontrolle ROSE-Core-RNAP konnte kein TSP identifiziert werden. Zwei identische TSP konnten in den Datensätzen RES167 $\Delta rshA$ und RES167 $\Delta sigH$, aber nicht im SigH-spezifischen ROSE-SigH-Datensatz identifiziert werden. Der entsprechende SigA-abhängige Promotor ($P\sigma^A$) wird in der genomischen Region durch einen blauen Pfeil dargestellt und benannt. Ein weiterer TSP konnte im *in vivo* Transkriptomdatensatz von RES167 $\Delta rshA$ und im *in vitro* Datensatz ROSE-SigH identifiziert werden. Dies zeigt, dass SigH exklusiv für die Erkennung dieses Promotors verantwortlich ist. Der SigH-abhängige Promotor ist durch einen orangenen Pfeil dargestellt und benannt ($P\sigma^H$) und dessen Sequenz am unteren Ende der Abbildung dargestellt. Die -35- und -10-Region des Promotors sind mit orangenen Boxen hervorgehoben und benannt. Der TSP ist in der Sequenz durch +1 gekennzeichnet. Die Visualisierung der RNA-Seq-Datensätze wurde durch VAMP (Hilker *et al.*, 2013, in Vorbereitung) verwirklicht.

Unter den Genen, die einen erstmalig identifizierten, exklusiv SigH-abhängigen Promotor besitzen, befindet sich *msrA* (paralog zu *msrB*), welches für eine Protein-Methionin-S-Oxid Reduktase kodiert, die im Zusammenspiel mit Thioredoxin oxidierte Methioninreste reduzieren kann und somit Bestandteil der Thioredoxin-bedingten Bekämpfung oxidativen Stresses ist. Weiter konnten exklusiv SigH-abhängige Promotoren stromaufwärts der Gene *ctaC* und *fixA*, deren Produkte an Elektronentransportprozessen beteiligt sind, identifiziert werden.

In der durchgeführten Analyse konnten erstmalig SigH-abhängige Promotoren stromaufwärts potentieller *Antisense*-RNAs identifiziert werden (*as-cg2357*, *as-cg2053*, *as-cg3357*). Die Promotoren sind in *Antisense*-Orientierung zu den Genen *ftsI* (Transpeptidase involviert in die Zellwandbiosynthese), *trpP* (*cg3357*; Tryptophan-Permease) und *cg2053* (putatives sekretiertes Protein) identifiziert worden.

Für Gene ohne klare Annotation (hypothetische Proteine oder Proteine unbekannter Funktion), die einen exklusiv SigH-abhängigen Promotor besitzen (*cg1282*, *cg0454*, *cg2263*, *cg1095*, *cg2267*) bleibt die genaue Funktion zu klären.

In dem zweiten Vergleich wurden TSP und damit Promotoren identifiziert, die partiell SigH-abhängig sind, da die Transkription in RES167 $\Delta sigH$ teilweise erhalten bleibt. Hier konnten 21 TSP und deren partiell SigH-abhängige Promotoren identifiziert werden (Tab. IV.8). Die identifizierten Promotoren besitzen, bis auf eine Ausnahme (P_{clpC}), die Nukleotidfolge g/tGGAAC als -35-Region und unterscheiden sich somit von den exklusiv SigH-abhängigen Promotoren, die die Nukleotidfolge g/tGGAAT besitzen. Der Promotor des Gens *clpC* besitzt als -35-Motiv TGGAAAG, welches auch in Einzelfällen für exklusiv SigH-abhängige Promotoren beobachtet werden konnte (P_{mtr} , P_{ctaC} und P_{cg2247}). Dem -35-Kernmotiv folgt in dem *clpC*-Promotor allerdings eine T-Nukleotidfolge (Anhang Tab. 3), die diesen Promotor von exklusiv SigH-abhängigen Promotoren unterscheidet, die weder an Position 7 des -35-Motivs ein T noch nachfolgend einen T-*Stretch* besitzen (Anhang Tab. 5). Von den insgesamt 21 identifizierten partiell SigH-abhängigen Genen wurden bereits acht als partiell SigH-abhängig durch RNA-Seq-Daten identifiziert (Abschnitt IV.2.6). Der direkte Nachweis der Abhängigkeit erfolgte allerdings erst hier, durch die Nutzung des ROSE-SigH-Datensatzes. Es handelt sich um Gene, die für Proteine kodieren, die an der Hitzestressantwort beteiligt sind bzw. Bestandteil des Proteasoms sind (*dnaK*, *clpB*, *clpC*, *clgR*, *dnaJ2*). Weiter konnte *in vitro* bestätigt werden, dass der Promotor des für den nicht-essentiellen Sigmafaktor kodierenden *sigB*-Gens von SigH erkannt wird.

Die neu identifizierten Promotoren liegen vor einem Gen, das für die ATP-abhängige Clp-Protease kodiert (*clpX*), beziehungsweise stromaufwärts von Membranprotein-kodierenden

Genen (*cg1277*, *cg1640*, *cg2380*, *cg2657*). Es konnten weiterhin drei Promotoren in *Anti-sense*-Orientierung ($P_{as-cg0849}$, $P_{as-cg3313}$, $P_{as-cg3395}$) zu proteinkodierenden Genen (*rmIA2*, Mannose-1-Phosphat Guanylyltransferase (GDP); *mrcB*, Membranständige Carboxypeptidase; *proP*, Prolin/Betain Permease, MFS-Typ) identifiziert werden. Zusätzlich wurden zwei partiell SigH-abhängige Promotoren stromaufwärts von *sigA* (*Housekeeping*-Sigmafaktor SigA) und *cseE* (Anti-Sigmafaktor des ECF-Sigmafaktors SigE) erstmalig identifiziert. Somit steht neben SigB (alternativer Sigmafaktor) (Ehira et al., 2009) auch SigA zum Teil unter der Kontrolle des ECF-Sigmafaktors SigH. Zusätzlich zeigt die teilweise SigH-abhängige Transkription des Gens *cseE*, dass neben SigM (Nakunst et al., 2007) eine weitere Verknüpfung des Regulationsnetzwerkes von SigH zu einem ECF-Sigmafaktor besteht.

Tab. IV.8 Partiiell SigH-abhängige Promotoren (Identifikation per ROSE-SigH)

Kodierregion	Gen-name	-35-Region	-10-Region	relative Anzahl ^a mappender Reads RES167 $\Delta rshA$	relative Anzahl ^a mappender Reads RES167 $\Delta sigH$	Verhältniss $\Delta rshA/\Delta sigH$ mappender Reads	Vorhergesagte Funktion	TSP ROSE- SigH	Referenz
cg3344	-	GGGAACAA	AGTTGCT	501	53	9,5	Putative Nitroreduktase	+	diese Arbeit (RNAseq)
cg2092	sigA	GGGAACAA	TGTTGTA	3688	441	8,4	RNA-Polymerase Sigmafaktor Sigma-A (<i>housekeeping</i> Sigmafaktor – Gruppe 1)	+	diese Arbeit (ROSE-SigH)
cg0378	-	TGGAACAT	CGTTGAA	826	111	7,5	Putatives Phagen-assoziiertes Protein	+	diese Arbeit (RNAseq)
cg2515	dnaJ2	GGGAACAA	CGTTGCT	519	103	5,1	Chaperon	+	(Busche et al., 2012)
cg1397	trmU	GGGAACAT	TGTTGAG	47	12	3,8	tRNA (5-Methylaminomethyl-2-Thiouridylat)-Methyltransferase	+	(Ehira et al., 2009)
cg3079	clpB	TGGAACAT	CGTTATA	85	23	3,6	Putative ATP-abhängige Protease, Hitze-schock Protein	+	(Ehira et al., 2009)
cg2963	clpC	TGGAAGTT	CGTTAAA	137	52	2,6	Putative ATP-abhängige Protease, Hitze-schock Protein	+	(Engels et al., 2004)
cg2152	clgR	GGGAACAA	AGTTGAA	154	67	2,3	Clp Protease Gen Aktivator	+	(Engels et al., 2004)
as-cg0849	-	TGGAATTT	GGTTCAA	23	11	2,1	Antisense RNA zu <i>cg0849</i> ; <i>rmlA2</i> ; Mannose-1-Phosphate Guanylyltransferase (GDP)	+	diese Arbeit (ROSE-SigH)
cg2984	ftsH	GGGAACGT	CGTTTGA	40	20	2,0	Zell-Teilungs Protein (ATP-abhängige Zn Metallopeptidase)	+	diese Arbeit (ROSE-SigH)
cg1272	cseE	GGGAACCTT	CGTTTTG	253	136	1,9	Konserviertes hypothetisches Protein	+	diese Arbeit (ROSE-SigH)
cg2657	-	GGGAACCTT	CGTTACT	495	294	1,7	Putatives Membranprotein	+	diese Arbeit (ROSE-SigH)
cg1640	-	GGGAACAA	CGTTCTG	15	9	1,6	Putatives Membranprotein	+	diese Arbeit (ROSE-SigH)
cg0948	serC	AGGAACCTG	AGTTGTG	20	14	1,4	Phosphoserin Transaminase, AT KlasseIV	+	diese Arbeit (ROSE-SigH)
cg1277	-	GGGAACCA	CGTTGTA	13	10	1,2	konserviertes putatives Membranprotein	+	diese Arbeit (ROSE-SigH)
as-cg3313	-	AGGAACCA	CGTTTCC	11	16	0,7	Antisense RNA zu <i>cg3313</i> ; <i>mrcB</i> ; Putative Membran Carboxypeptidase	+	diese Arbeit (ROSE-SigH)
cg2102	sigB	GGGAACCTT	CGTTGAA	417	451	0,9	RNA-Polymerase Sigmafaktor Sigma-B (alternativer Sigmafaktor – Gruppe 2)	+	(Halgasova et al., 2001)
as-cg3395	-	GGGAACAA	TGTTAGT	8	12	0,7	Antisense RNA zu <i>cg3395</i> ; <i>proP</i> ; Putative Proline/Betain Permease, MFS-Typ	+	diese Arbeit (ROSE-SigH)
cg2380	-	GGGAACAT	CGTTAAT	277	319	0,9	Putatives Membranprotein	+	diese Arbeit (ROSE-SigH)
cg2661	dsbA2	AGGAACCTA	TGTTGAG	428	-*	-	Thiol-Disulfid Isomerase	+	(Ehira et al., 2009)
cg2620	clpX	TGGAACAA	TGTTGGT	70	-**	-	ATP-abhängige Clp Protease	+	diese Arbeit (ROSE-SigH)

^a relative Anzahl: *mappender Reads*/Gesamtzahl *mappender Reads* in Millionen; * unter Grenzwert; **Überlagerung mit SigA-abhängigen Transkri

3.3. Nachweis der zusätzlichen und direkten Kontrolle partiell SigH-abhängiger Promotoren durch den ECF-Sigmafaktor SigE und die Identifizierung putativ exklusiv SigE-abhängiger Promotoren

Der Grund für die verbliebene Transkription SigH-abhängiger Gene bei einer Deletion des ECF-Sigmafaktors SigH konnte durch die bisher durchgeführten Analysen nicht geklärt werden. Die entsprechenden Promotoren wurden daher als partiell SigH-abhängig bezeichnet. Da für das nah verwandte Aktinobakterium *M. tuberculosis* eine duale Kontrolle SigH-abhängiger Gene durch SigH und SigE (ortholog zu SigE aus *C. glutamicum*) beschrieben ist (Song *et al.*, 2008), wurde ein weiteres ROSE-Experiment mit SigE-gesättigter *E. coli*-RNAP (ROSE-SigE) durchgeführt. Dabei sollte überprüft werden, ob die Transkription partiell SigH-abhängiger Gene in der Deletionsmutante RES167 $\Delta sigH$ durch SigE ermöglicht wird. Die Mapping-Statistik des ROSE-SigE-Datensatzes ist in Tabelle Tab. IV.6 erfasst.

Auch für den ROSE-SigE-Datensatz konnte, durch den ersten manuellen Abgleich einer Reihe in dieser Arbeit identifizierten partiell SigH-abhängigen TSP mit den durch ROSE-SigE identifizierten TSP eine exakte Übereinstimmung der *in vivo* und *in vitro* identifizierten TSP festgestellt werden. Wie in Abb. IV-18 zu sehen gibt es nur einen spezifischen TSP stromaufwärts des *Dank*-Gens, der mit den *in vivo* bestimmten partiell SigH-abhängigen TSP (RNA-Seq-Datensätze RES167 $\Delta rshA$ und RES167 $\Delta sigH$) und dem ebenfalls *in vitro* identifizierten TSP des ROSE-Ansatzes ROSE-SigH übereinstimmt.

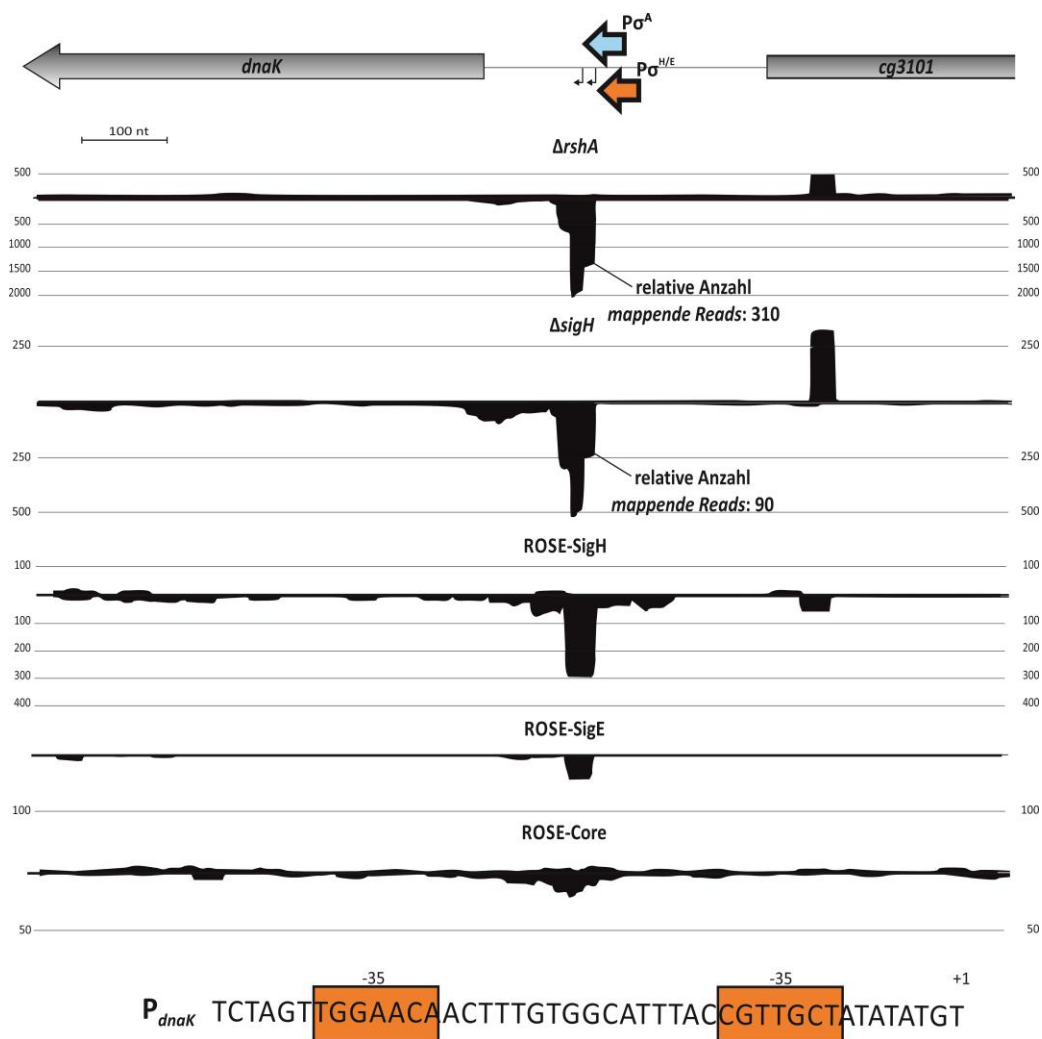


Abb. IV-18 Visualisierung der nativen 5'-Enden RNA-Seq-Datensätze im Bereich des Gens *dnaK* der *C. glutamicum* RES167 Deletionsderivate $\Delta rshA$, $\Delta sigH$ und der *in vitro* Datensätze von ROSE-SigH, ROSE-SigE und ROSE-Core (Kontrolle). Die genomische Region des Gens *dnaK* ist oberhalb der Tracks schematisch abgebildet. Die mappenden Reads in den unterschiedlichen Datensätzen sind schwarz dargestellt und die Tracks nach Datensatz benannt. TSP wurden durch einen Anstieg der Anzahl mappende Reads an einer Position identifiziert. TSP sind in der schematischen Darstellung der genomischen Region durch einen geknickten Pfeil gekennzeichnet. In der Orientierung des *dnaK* Gens in den Datensätzen $\Delta rshA$ und $\Delta sigH$, aber in keinem der *in vitro* Datensätze, die spezifisch für SigH bzw. SigE sind, identifiziert werden. Der entsprechende SigA-abhängige Promotor (P_{σ^A}) wurde in der genomischen Region durch einen blauen Pfeil dargestellt und benannt. Ein weiterer TSP konnte ebenfalls in den *in vivo* Transkriptom Datensätzen von RES167 $\Delta rshA$ und $\Delta sigH$ und zusätzlich in den *in vitro* Datensätzen ROSE-SigH und ROSE-SigE identifiziert werden. Dies zeigt, dass SigH und SigE für die Erkennung dieses Promotors verantwortlich sind und zur Initiierung der Transkription notwendig sind. Die Reduzierung der relativen Anzahl mappende Reads im Vergleich an diesem TSP von $\Delta rshA$ zu $\Delta sigH$ ist auf die Deletion des ECF-Sigmafaktors SigH zurückzuführen. Die bestehende Transkription ist somit abhängig von SigE. Der entsprechende SigH/SigE-abhängige Promotor ist durch einen orangenen Pfeil dargestellt und benannt ($P_{\sigma^{H/E}}$) und dessen Sequenz am unteren Ende der Abbildung dargestellt. Die -35 und -10 Region des Promotors sind mit orangenen Boxen hervorgehoben und benannt. Der TSP ist in der Sequenz durch +1 gekennzeichnet. Die Visualisierung der RNA-Seq-Datensätze wurde durch VAMP (Hilker et al., 2013, in Vorbereitung) verwirklicht.

In Einzelfällen konnte eine Abweichung von 1 - 2 bp der *in vitro* bestimmten TSP zu den auf *in vivo* Daten basierenden Identifikation der TSP festgestellt werden. Die Positionen der ROSE-SigE-TSP wurden daher mit einer Abweichung von ± 2 Positionen mit den Positionen der TSP der Transkriptomanalyse der Deletionsmutanten RES167 $\Delta sigH$ und RES167 $\Delta rshA$ abgeglichen. Es wurde folgender Vergleich, entsprechend zu der Verarbeitung der ROSE-SigH-Transkriptomdaten, durchgeführt: Die TSP mussten in dem *in vitro* Datensatz ROSE-SigE, wie auch den *in vivo* Datensätzen RNA-Seq RES167 $\Delta rshA$ und RNA-Seq RES167 $\Delta sigH$ identifiziert worden sein. Die gleichzeitige Übereinstimmung mit einem TSP des *in vitro* Datensatz ROSE-SigH war optional.

So wurden sowohl solche TSP und damit deren Promotoren identifiziert, die partiell SigH-abhängig sind, als auch solche, die der exklusiven Kontrolle durch den ECF-Sigmafaktor SigE unterliegen. Die Aufarbeitung der Rohdaten erfolgte analog zu der von ROSE-SigH.

Tab. IV.9 SigH- und zusätzlich SigE-abhängige Promotoren (Identifikation per ROSE-SigH und ROSE-SigE)

Kodierregion	Genname	-35-Region	-10-Region	relative Anzahl ^a mappende Reads RES167 $\Delta rshA$	relative Anzahl ^a mappende Reads RES167 $\Delta sigH$	Verhältniss $\Delta rshA/\Delta sigH$ mappende Reads	Vorhergesagte Funktion	TSP ROSE- SigH	TSP ROSE- SigE	Referenz
cg3344	-	GGGAACAA	AGTTGCT	501	53	9,5	Putative Nitroreduktase	+	+	diese Arbeit (RNAseq)
cg2092	<i>sigA</i>	GGGAACAA	TGTTGTA	3688	441	8,4	RNA-Polymerase Sigmafaktor Sigma-A (<i>housekeeping</i> Sigmafaktor – Gruppe 1)	+	+	diese Arbeit (ROSE-SigH)
cg0378	-	TGGAACAT	CGTTGAA	826	111	7,5	Putatives Phagen- assoziiertes Protein	+	+	diese Arbeit (RNAseq)
cg1397	<i>trmU</i>	GGGAACAT	TGTTGAG	47	12	3,8	tRNA (5- Methylaminomethyl-2- Thiouridylat)- Methyltransferase	+	+	(Ehira et al., 2009)
cg3079	<i>clpB</i>	TGGAACAT	CGTTATA	85	23	3,6	Putative ATP-abhängige Protease, Hitze-schock Protein	+	+	(Ehira et al., 2009)
cg3100	<i>dnaK</i>	TGGAACAA	CGTTGCT	310	89	3,5	Chaperon DnaK,	+	+	diese Arbeit (RNAseq)
cg2984	<i>ftsH</i>	GGGAACGT	CGTTTGA	40	20	2	Zell-Teilungs Protein (ATP- abhängige Zn Metallopeptidase); Protease	+	+	diese Arbeit (ROSE-SigH)
cg1272	<i>cseE</i>	GGGAACCTT	CGTTTTG	253	136	1,9	Konserviertes hypothetisches Protein	+	+	diese Arbeit (ROSE-SigH)
cg2657	-	GGGAACCTT	CGTTACT	495	294	1,7	Putatives Membranprotein	+	+	diese Arbeit (ROSE-SigH)
cg1640	-	GGGAACAA	CGTTCTG	15	9	1,6	Putatives Membranprotein	+	+	diese Arbeit (ROSE-SigH)
cg1277	-	GGGAACCA	CGTTGTA	13	10	1,2	konserviertes putatives Membranprotein	+	+	diese Arbeit (ROSE-SigH)
cg2102	<i>sigB</i>	GGGAACCTT	CGTTGAA	417	451	0,9	RNA-Polymerase Sigmafaktor Sigma-B (alternativer Sigmafaktor – Gruppe 2)	+	+	(Halgasova et al., 2001)
cg2380	-	GGGAACAT	CGTTAAT	277	319	0,9	Putatives Membranprotein	+	+	diese Arbeit (ROSE-SigH)
cg2661	<i>dsbA2</i>	AGGAACTA	TGTTGAG	428	-	-	Thiol-Disulfid Isomerase	+	+	(Ehira et al., 2009)

Für 14 der 21 partiell SigH-abhängigen Promotoren (Anhang Tab. 3) konnte durch den *Bottom-Up*-Ansatz ROSE-SigE eine eindeutige, zusätzliche Transkription abhängig von SigE nachgewiesen werden (Tab. IV.9). Die zusätzliche Transkription durch diese Promotoren in Abhängigkeit von SigE ist somit der Grund für die verbleibende Transkription der partiell SigH-abhängigen Gene in der Deletionsmutante RES167 $\Delta sigH$.

Damit wurden zum ersten Mal direkt SigE-abhängige Promotoren in *C. glutamicum* nachgewiesen. Zusätzlich konnte die Überlappung der SigH- und der SigE-Regulationsnetzwerke, wie für *M. tuberculosis* bereits beschrieben, für *C. glutamicum* gezeigt werden (Song *et al.*, 2008). Auffällig ist, dass die hier durch ROSE-SigE identifizierten SigE-abhängigen Promotoren auch mit den übrigen partiell SigH-abhängigen Promotoren (P_{dnaJ2} , P_{clpC} , P_{clpX} , P_{clgR} , P_{serC} , $P_{as-cg3313}$ und $P_{as-cg3395}$) mit Ausnahme von P_{clpC} identische –35-(g/TGGAAC) und –10-Promotormotive (GTT) aufweisen (Anhang Tab. 3). In diesem Zusammenhang konnte für das Gen *clgR* (**Clp Gen Regulator**), einem Regulator der Hitzestressantwort, bei einer manuellen Kontrolle der ROSE-SigE-Daten derselbe TSP wie bei ROSE-SigH identifiziert werden. Dieser TSP wurde nicht automatisch identifiziert und stellt somit nur einen potentiellen SigE-abhängigen TSP dar. Für die übrigen partiell SigH-abhängigen Promotoren wurden keinerlei *Reads* an entsprechender Position im ROSE-SigE-Datensatz *gemappt*.

Basierend auf den eindeutigen Ergebnissen beinhaltet das SigE-Regulon in *C. glutamicum* neben den Sigmafaktorgen *sigA*, *sigB* und dem SigE-Anti-Sigmafaktorgen *cseE*-Gene des SigH-Regulationsnetzwerkes, die für Proteine kodieren, die an der Hitzestressantwort beteiligt sind oder Bestandteil des Proteasoms darstellen (*dnaK-grpE*-Operon, *clpB*, *ftsH* und potentiell *clgR*) sowie solche, die Membranproteine kodieren (*cg1277*, *cg1640*, *cg2380* und *cg2657*; Tab. IV.9). Dies ist in Übereinstimmung mit der Beobachtung, dass eine $\Delta sigE$ -Mutante sensitiv gegenüber Membranstress ist (Park *et al.*, 2008).

Das Gen für den Anti-Sigmafaktor CseE, der die Aktivität von SigE reguliert (Park *et al.*, 2008), wird zum einen konstitutiv und zum anderen abhängig von einem SigH/SigE-abhängigen Promotor transkribiert (Abb. IV-19). Durch Analyse des RES167 RNA-Seq-Datensatzes konnten zwei TSP stromaufwärts von *sigE* identifiziert werden. TSP1 stimmt mit dem Startnukleotid A der Kodiersequenz des Gens überein und als TSP2 wurde das Nukleotid G 116 bp stromaufwärts des Startcodons identifiziert (Daten nicht gezeigt). Die putativen –10-Regionen der entsprechenden Promotoren, CAAAAT (P1) und TATAGT (P2) besitzen eine hohe Übereinstimmung mit der Konsensussequenz der

–10-Region (TANANT) SigA-abhängiger Promotoren (Patek und Nesvera, 2011). Zusätzlich zu der eindeutigen –10-Region besitzt der Promotor P1 eine –35-Region (TAGAT), die Ähnlichkeit zu der Konsensussequenz SigA-abhängiger Promotoren hat (Patek und Nesvera, 2011). Die Kontrolle des Sigmafaktorgens *sigE* steht somit, wie das *sigH*-Gen, unter der Kontrolle multipler SigA-abhängiger Promotoren. Laut Park und Mitarbeitern werden die Gene *sigE* und *cseE* in Form eines Operons transkribiert (Park *et al.*, 2008). Das ECF-Sigmafaktor-Anti-Sigmafaktor-Operon *sigE-cseE* besitzt damit eine ähnliche transkriptionelle Organisation wie *sigH-rshA*, wobei hier kein zusätzlich SigA-abhängiger Promotor stromaufwärts des Gens *rshA* zu finden ist.

Neben der Identifizierung SigH/SigE-abhängiger Promotoren, konnten drei TSP identifiziert werden, die nicht zu den SigH-abhängigen Promotoren gehören. Die entsprechenden TSP konnten nur in dem *in vitro* ROSE-SigE-Datensatz identifiziert werden. Bei den potentiell exklusiv SigE-abhängigen Genen handelt es sich um *cg1121* (Permease vom MFS-Typ), *cg3309* (putatives sekretiertes Protein) und *amtR* (Regulator des Stickstoffmetabolismus) (Tab. IV.10). Die Promotoren besitzen alle ein ähnliches –35-Promotormotiv (GGGAA-CT), welches auch bei nachgewiesenen SigH/SigE-abhängigen Promotoren der Gene *cseE*, *sigB*, *dsbA2* und *cg2657* (Membranprotein) vorhanden ist (Tab. IV.9). Weiter ist die Transkription an den TSP der Gene *cg3309* und *amtR* in der Deletionsmutante RES167 $\Delta rshA$ im Vergleich zu der Deletionsmutante RES167 $\Delta sigH$ wie bei partiell SigH-abhängigen Genen verstärkt (Tab. IV.10). Eine Abhängigkeit der Promotoren von SigH konnte hier aber nicht gezeigt werden.

Tab. IV.10 Putative exklusiv SigE-abhängige Promotoren (Identifikation per ROSE-SigE)

Kodierregion	Genname	-35-Region	-10-Region	relative Anzahl mappende Reads $\Delta rshA$	relative Anzahl mappende Reads $\Delta sigH$	Verhältnis $\Delta rshA/\Delta sigH$ mappende Reads	Vorhergesagte Funktion	TSP ROSE-SigE	Referenz
cg3309	-	GGGAACT	CGTTAAG	39,74	17,88	2,2	Putatives sekretiertes Protein	+	diese Arbeit (ROSE-SigE)
cg0986	<i>amtR</i>	-GAAACT	CGTTGGA	20,35	14,97	1,4	Regulator des Stickstoffmetabolismus	+	diese Arbeit (ROSE-SigE)
cg1121	-	GGGAACT	CGTTAAC	106,53	101,46	1,0	Permease, MFS-Typ	+	diese Arbeit (ROSE-SigE)

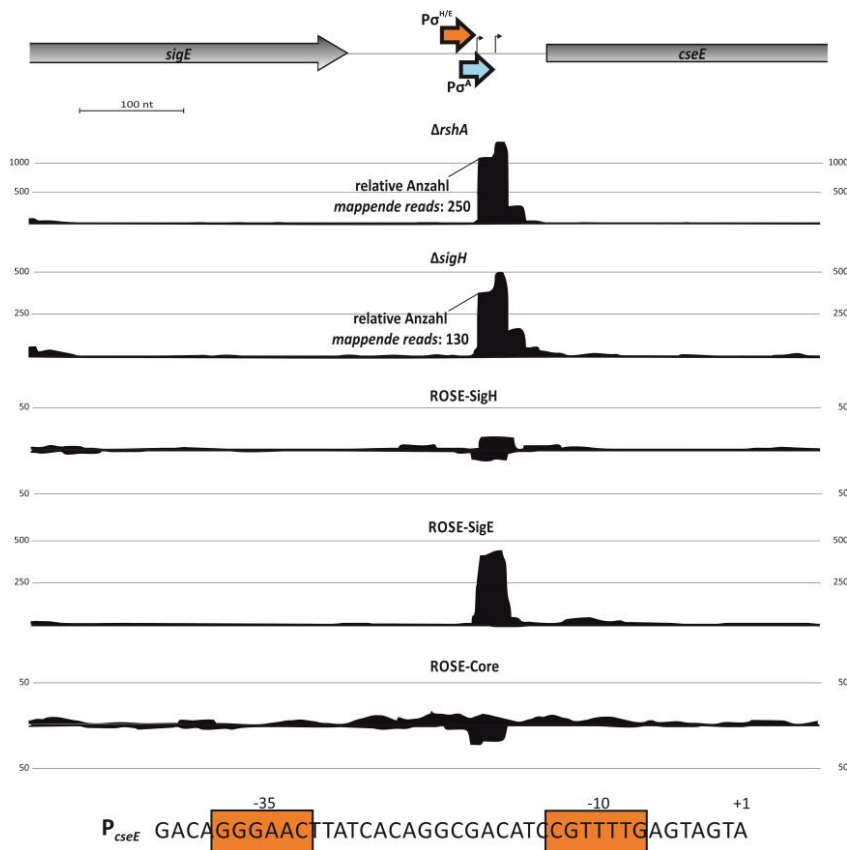


Abb. IV-19 Visualisierung der RNA-Seq-Datensätze der nativen 5'-Enden im Bereich der genomischen Region *sigE-cseE* der Stämme *C. glutamicum* RES167 $\Delta rshA$, RES167 $\Delta sigH$ und der *in vitro* Datensätze von ROSE-SigH, ROSE-SigE und ROSE-Core-RNAP (Kontrolle). Die genomische Region *sigE-cseE* ist oberhalb der Tracks abgebildet. Die mappenden Reads in den unterschiedlichen Datensätzen sind schwarz dargestellt und nach dem Datensatz benannt. TSP wurden durch einen Anstieg der Anzahl mappende Reads an einer Position identifiziert. TSP sind in der schematischen Darstellung der genomischen Region durch einen geknickten Pfeil gekennzeichnet. In der Negativkontrolle ROSE-Core-RNAP konnte kein TSP identifiziert werden. Zwei identische TSP konnten in den Datensätzen RES167 $\Delta rshA$ und RES167 $\Delta sigH$, aber in keinem der *in vitro* Datensätze, die spezifisch für SigH bzw. SigE sind, identifiziert werden. Der entsprechende SigA-abhängige Promotor (P_{σ^A}) wurde in der genomischen Region durch einen blauen Pfeil dargestellt und benannt. Ein weiterer TSP konnte ebenfalls in den *in vivo* Transkriptomdatensatz von RES167 $\Delta rshA$ und RES167 $\Delta sigH$ und zusätzlich in den *in vitro* Datensätzen ROSE-SigH und ROSE-SigE identifiziert werden. Dies zeigt, dass SigH und SigE für die Erkennung dieses Promotors verantwortlich und zur Initiierung der Transkription notwendig sind. Die Reduzierung der relativen Anzahl mappende Reads an diesem TSP, im Vergleich von RES167 $\Delta rshA$ zu RES167 $\Delta sigH$ ist auf die Deletion des ECF-Sigmafaktors SigH zurückzuführen. Die bestehende Transkription ist somit abhängig von SigE. Der entsprechend SigH/SigE-abhängige Promotor ist durch einen orangenen Pfeil dargestellt, benannt ($P_{\sigma^{H/E}}$) und dessen Sequenz am unteren Ende der Abbildung dargestellt. Die -35- und -10-Region des Promotors sind mit orangenen Boxen hervorgehoben und benannt. Der TSP ist in der Sequenz durch +1 gekennzeichnet. Die Visualisierung der RNA-Seq-Datensätze wurde durch VAMP (Hilker *et al* 2013, in Vorbereitung) verwirklicht.

3.4. Das globale Transkriptionsprofil der Deletionsmutante $\Delta cseE$ liefert eine Bestätigung für die Kontrolle *in vitro* identifizierter SigE-abhängiger Gene durch SigE *in vivo* und validiert CseE als Anti-Sigmafaktor von SigE

Um die Gene, die durch die hier durchgeführten ROSE-Analysen als partiell oder exklusiv SigE-abhängig identifiziert worden sind, *in vivo* auf SigE-abhängige Transkription hin zu überprüfen, wurde wie bei der Untersuchung des SigH-Regulationsnetzwerkes vorgegangen. Park und Mitarbeiter erzeugten eine *cseE*-Deletionsmutante und zogen durch physiologische Untersuchungen dieser Mutante und den Nachweis der direkten Interaktion zwischen Sigmafaktor SigE und Anti-Sigmafaktor CseE *in vitro* den Schluss, dass CseE der Anti-Sigmafaktor von SigE ist (Park *et al.*, 2008). Die Deletion $\Delta cseE$ wurde in dieser Arbeit unter Verwendung des *pK18mobsacB*-Systems im Genom von *C. glutamicum* RES167 etabliert. Der Nachweis der erfolgreichen Deletion wurde mittels Test-PCR erbracht. Das PCR-Produkt im Referenzstamm war 1.965 bp groß, das in dem Deletionsstamm $\Delta cseE$ war mit 1.470 bp rund 495 bp (Größe der Deletion) kleiner (Abb. IV-20 B).

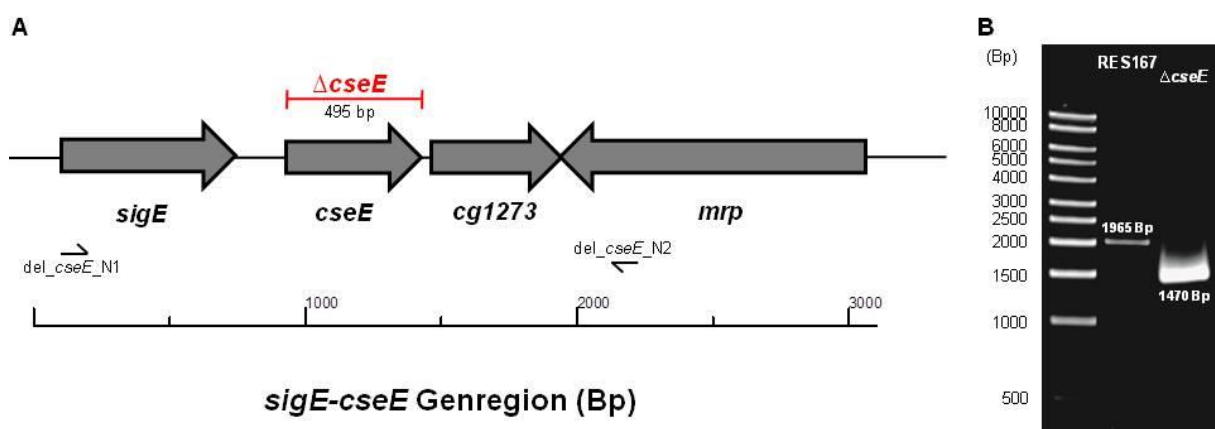


Abb. IV-20 Genregion *sigE-cseE* inklusive Kartierung der Deletion $\Delta cseE$ und entsprechendem Deletionsnachweis per Test-PCR. (a) Genregion von *sigE-cseE* mit deletiertem Bereich im $\Delta cseE$ -Stamm. Die zum Nachweis der Deletion verwendeten Primer sind eingezeichnet. (b) Deletionsnachweis mittels Test-PCR. Das PCR-Produkt der Deletionsmutante $\Delta cseE$ ist um 495 Basenpaare (Bp) kleiner als das Produkt des Wildtyps.

Es wurde davon ausgegangen, dass eine Deletion des Anti-Sigmafaktors zu einer Dereprimierung des SigE-Regulationsnetzwerkes und somit verstärkter Transkription der Zielgene führt. Es wurde eine vergleichende Microarray-Analyse durchgeführt. Die Hybridisierung der Gesamtgenom-Microarrays von *C. glutamicum* wurde mit cDNA durchgeführt. Als Template für die Herstellung der cDNA mittels reverser Transkriptase diente Gesamt-RNA, die aus dem Referenzstamm *C. glutamicum* RES167 bzw. dem abgeleiteten *cseE*-

Deletionsstamm isoliert wurden, durchgeführt. Die Kultivierung erfolgte dabei unter Standardbedingungen im Luftschüttler bei 300 U/min und 30 °C in CASO-Vollmedium, da eine Anzucht der Deletionsmutante in CGXII Minimalmedium nicht möglich war. Die Auswertung der Microarray-Daten erfolgte wie in Material und Methoden beschrieben mit dem Unterschied, dass der Grenzwert für eine signifikante differentielle Transkription auf $m \geq 1$ gesetzt wurde. Der Signalintensitätsverhältnis (m) / Signalintensität (a) Plot der Hybridisierungen ist in Abb. IV 21 gezeigt. Die differentiell transkribierten Gene Anhang Tab. 2 aufgeführt. Insgesamt wurden 85 Gene im Deletionsstamm $\Delta cseE$ im Vergleich zum Referenzstamm stärker transkribiert und 91 Gene schwächer. Zu den im Deletionsderivat stärker transkribierten Genen zählten 11 von 14 in vitro eindeutig als SigH/SigE-abhängig beschriebenen Gene (Tab. IV.9). Für die drei SigH/SigE-abhängigen Gene *sigA*, *trmU* und *dsbA2* konnte keine differentielle Transkription beobachtet werden. Bei den im Vergleich zur Referenz verstärkt transkribierten SigH/SigE-abhängigen Genen handelt es sich um Gene, die für Membranproteine (*cg1640*, *cg2380*, *cg2657*, *cg1277*), Chaperone (*dnaK*), Proteasen (*clpB*, *ftsH*), eine putative Nitroreduktase (*cg3344*) und ein Phagen-assoziiertes Protein (*cg0378*) kodieren. In Übereinstimmung damit befindet sich auch der transkriptionelle Aktivator der *clpP*-Proteasegene, *clgR*, unter den verstärkt transkribierten Genen, der einen potentiell SigH/SigE-abhängigen Promotor besitzt. Weiterhin konnte eine verstärkte Transkription des Gens *sigB*, kodierend für den alternativen Sigmafaktor SigB, in der $\Delta cseE$ -Mutante festgestellt werden. Der Nachweis der SigE-abhängigen Transkription des Anti-Sigmafaktorgens *cseE* (*cg1272*) konnte indirekt durch die erhöhte Transkriptmenge des Folgegens *cg1273* in der Deletionsmutante beobachtet werden (Abb. IV-20 A). Dies ist der erste in vivo Beleg für die SigE-abhängige Transkription des Anti-Sigmafaktorgens *cseE*. Die Transkription des Anti-Sigmafaktorgens durch den eigenen ECF-Sigmafaktor konnte in dieser Arbeit auch für SigH-*rshA* gezeigt werden. Zum Teil handelt es sich bei den identifizierten SigH/SigE-abhängigen Genen um die ersten Gene eines Operons (z.B. das *dnaK-grpE*-Operon). Die Folgegene (14) waren im Vergleich zum Referenzstamms ebenfalls verstärkt transkribiert.

Neben den eindeutig SigH/SigE-abhängigen Genen konnte auch für die partiell SigH-abhängigen Gene *dnaJ2* (Chaperon) und *clpC* (Protease), für die keine zusätzliche Transkription durch SigE-gesättigte RNAP gezeigt werden konnte, eine relative erhöhte mRNA-Menge in der Deletionsmutante $\Delta cseE$ im Vergleich zum Referenzstamm beobachtet werden. Es befanden sich keine weiteren SigH-abhängigen Gene unter den differentiell transkribierten Genen.

Die durch ROSE-SigE identifizierten, putativ exklusiv SigE-abhängigen Gene *amtR* und *cg3309*, inklusive der im Operon enthaltenen Gene, konnten mit m-Werten von 1 bzw. 1,7 für das erste Gen im Operon als differentiell transkribiert identifiziert werden. Das Gen *cg1121* (Permease, MFS-Typ), wurde nur geringfügig stärker als im Referenzstamm transkribiert ($m = 0,86$).

Somit war die SigE-abhängige Transkription an den durch *in vitro* Analysen als SigE-abhängig beschriebenen Promotoren *in vivo* bestätigt. Weiter ist die Annahme, dass der Sigmafaktor SigE in der Deletionsmutante $\Delta cseE$ aktiv ist und durch Bindung an das RNAP-Core Enzym die Transkription SigE-abhängiger Gene ermöglicht damit bestätigt.

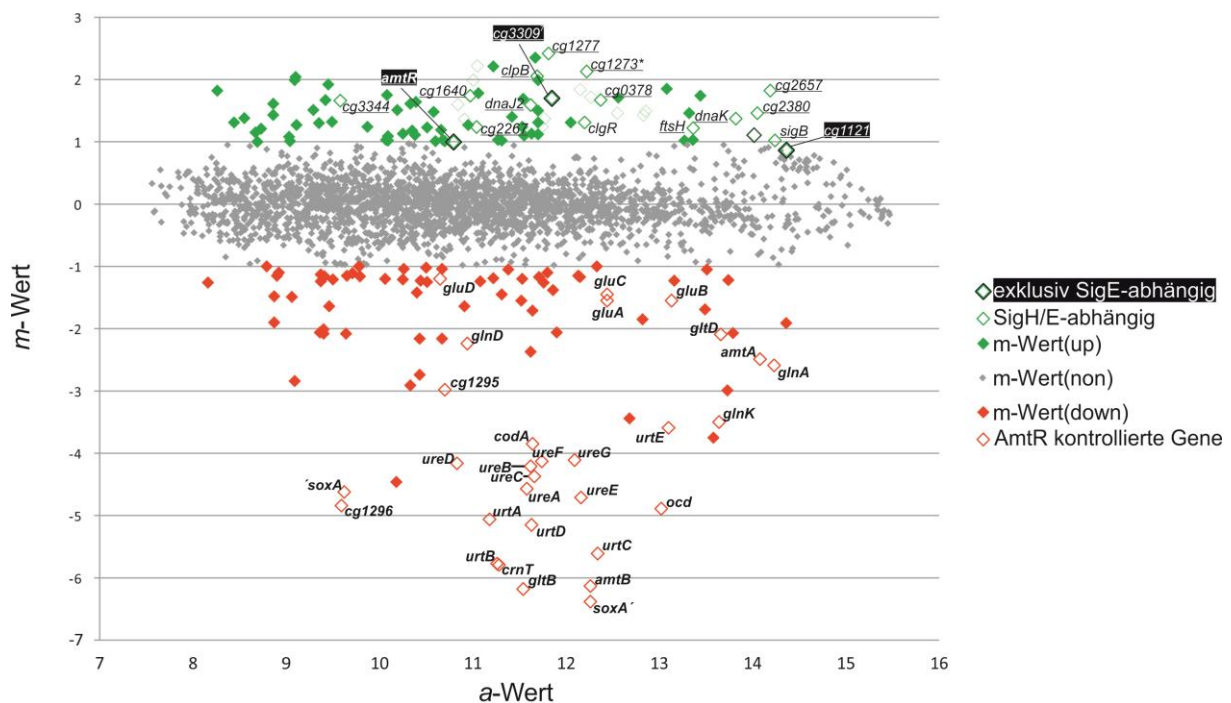


Abb. IV-21 *Microarray-Analyse von C. glutamicum RES167 im Vergleich mit der Deletionsmutante $\Delta cseE$.* Darstellung der *Microarray*-Hybridisierung $\Delta cseE$ gegen RES167 als m/a-Plot (m -Wert = $\log_2(\text{Signalintensitäts-Verhältnis})$; a -Wert = mittlere Signalintensität). Für die Hybridisierung wurde Gesamt-RNA von je zwei biologischen Replikaten, die bis zur exponentiellen Phase in CASO Vollmedium angezogen wurden, isoliert. Gene mit erhöhter mRNA Menge im Deletionsstamm $\Delta cseE$ im Vergleich zum Referenzstamm haben ein positives Signalintensitätsverhältnis, während Gene die eine höhere mRNA Menge in der Referenz RES167 haben ein negatives Signalintensitätsverhältnis; grüne Rauten (stärker in $\Delta cseE$ transkribiert); SigH/SigE-abhängige Gene (leere grüne Rauten) und exklusiv SigE-abhängige Gene (schwarz umrandete leere grüne Rauten) sind durch Varianten grüner Rauten entsprechend der Legende hervorgehoben und nur das erste Gen eines Operons ist benannt und entsprechend der Legende hervorgehoben – Ausnahme *cg1273** (transkriptionelle Einheit mit dem deletierten Gen *cseE*); rote Dreiecke (schwächer in $\Delta cseE$ transkribiert); AmtR regulierte Gene sind durch leere rote Rauten markiert und benannt; Gene mit keiner differentiellen Transkription sind als graue Punkte dargestellt. m -Werte $\geq +1$ oder ≤ -1 (entspricht einer Änderung von 2 bzw. 0,5) gelten als signifikant.

Die in der relativen Transkriptmenge im Vergleich zum Referenzstamm drastisch verringerten Gene im Deletionsstamm $\Delta cseE$ ($m < -3$) sind, bis auf drei Gene, alle durch den AmtR-Repressor reguliert (21) (Anhang Tab. 2). AmtR ist der zentrale negative Regulator des Stickstoffmetabolismus in *C. glutamicum* (Jakoby et al., 2000). Insgesamt sind 31 von 39 Mitgliedern des AmtR-Regulons (Hasselt et al., 2011; Pauling et al., 2012) schwächer ($m \leq 1$) als im Referenzstamm transkribiert (Abb. IV-21 und Anhang Tab. 2).

Die verringerte Transkriptmenge in $\Delta cseE$ ist vermutlich ein direkter Effekt der induzierten SigE-abhängigen Expression von amtR.

3.5. Die differentielle SigE-abhängige Transkription der Gene *amtR* und *cseE* konnte durch quantitative *real-time* RT-PCR in den Deletionsderivaten $\Delta cseE$ und $\Delta sigE$ im Vergleich zum Referenzstamm bestätigt werden

Um die Ergebnisse der vergleichenden Transkriptomanalyse der $\Delta cseE$ -Mutante gegen die Referenz RES167 für die Gene *amtR* und *cseE* mit einer sensitiveren Methode zu validieren, wurden quantitative *real-time* RT-PCR-Analysen durchgeführt. In diese Analysen wurde zusätzlich der Deletionsstamm RES167 $\Delta sigE$ eingeschlossen um zu überprüfen, ob die Deletion des ECF-Sigmafaktorgens einen Einfluss auf die Transkription der beiden Gene hat. Für die Stämme RES167 $\Delta cseE$ und RES167 wurde die gleiche RNA wie für die *Microarray*-Analysen mit einem zusätzlichen biologischen Replikat eingesetzt. Die Kultivierung und Isolierung der RNA des Stammes RES167 $\Delta sigE$ erfolgte mit drei biologischen Replikaten analog zu der Anzucht der Stämme RES167 $\Delta cseE$ und RES167. Für die Messung der Transkriptmenge des Gens *cseE* wurden Primer genutzt die direkt stromabwärts des Gens bzw. des deletierten Bereichs in $\Delta cseE$ binden. Die relative Transkriptmenge des *cseE*-Gens wurde somit in den folgenden Messungen indirekt gemessen.

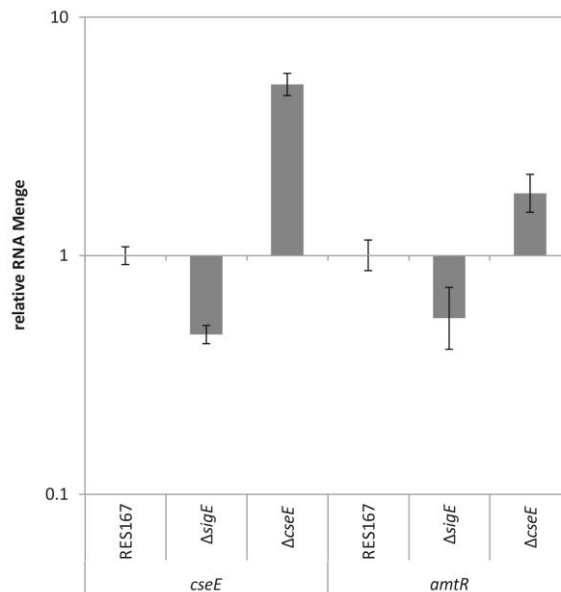


Abb. IV-22 Relative Transkriptmengen der Gene *amtR* und *cseE* in *C. glutamicum* Δ*cseE*/*C. glutamicum* RES167 und *C. glutamicum* Δ*sigE*/*C. glutamicum* RES167. Die relativen Transkriptlevel wurden mit q-RT-PCR bestimmt. Die Daten des RES167 Stammes sind als Referenz verwendet und die entsprechenden Werte auf der logarithmischen Skala auf 1 gesetzt worden. Es wurden je drei biologische Replikate pro Stamm doppelt analysiert. Die Standardabweichungen sind als Fehlerbalken dargestellt.

Durch die durchgeführte q-RT-PCR-Analyse konnte die erhöhte RNA-Menge der Gene *amtR* und *cseE* in RES167 Δ*cseE* mit einer 2-fachen bzw. 4,4-fachen Erhöhung gegenüber dem Referenzstammes nachgewiesen werden (Abb. IV-22). Dies entspricht in etwa den Werten des *Microarrays*. Die Transkriptmenge von *cseE* und *amtR* in der Δ*sigE*-Mutante betrug etwa die Hälfte der Transkriptmenge im Vergleichsstamm, wobei die Werte der 3 biologischen Replikate für die Transkriptmenge des Gens *amtR* stark schwanken. Dennoch konnte so der Einfluss der Deletion des ECF-Sigmafaktorgens im ungestressten Zustand auf *amtR* und *cseE* nachgewiesen werden.

3.6. Ableitung der Promotorkonsensussequenzen exklusiv SigH-abhängiger und SigH/SigE-abhängiger Promotoren

Für die Erstellung der Sequenzlogos der exklusiv SigH- bzw. SigH/SigE-abhängigen Promotoren wurden nur durch ROSE-Experimente eindeutig experimentell validierte Promotoren verwendet. Es wurde die Software WebLogo 3.2 (Crooks *et al.*, 2004; <http://weblogo.threeplusone.com/>) für die Erstellung der Sequenzlogos genutzt.

Der Datensatz zur Erstellung des Sequenzlogos exklusiv SigH-abhängiger Promotoren bestand aus insgesamt 31 identifizierten Promotoren (Tab. IV.7 bzw. Anhang Tab. 5). Basierend auf dieser umfassenden Anzahl exklusiv SigH-abhängiger Promotoren konnte eine klar definierte Konsensussequenz (g/tGGAAT –16-19bp– GTTg) ermittelt werden (Abb. IV-23), wobei das T an Position 6 des –35-Element in drei Promotoren durch ein G ersetzt ist (P_{cg2247} , P_{2ctaC} und P_{mtr} – Tab. IV.7). Auffällig ist neben den definierten –35- und –10-Kernmotiven die stromabwärts folgende AT-reiche Region des –35-Promotorelements, die als erweiterte –35-Region bezeichnet werden kann.

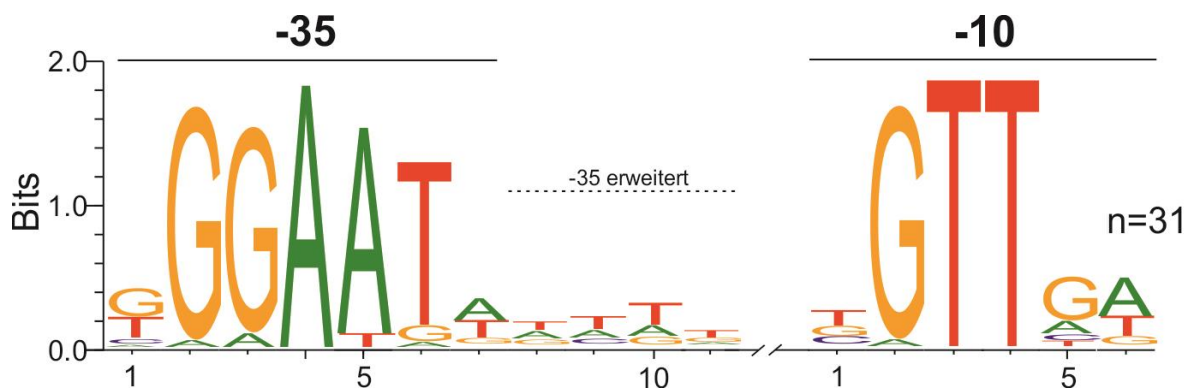


Abb. IV-23 DNA-Sequenzlogo exklusiv SigH-abhängiger Promotoren. Das Sequenzlogo wurde auf Basis von insgesamt 40 in dieser Arbeit identifizierten exklusiv SigH-abhängigen Promotoren durch WebLogo 3.2 (Crooks *et al.*, 2004; <http://weblogo.threeplusone.com/>) erstellt. Der Informationsgehalt eines Nucleotides an einer bestimmten Position ist in Bits durch die Höhe des Nucleotidsymbols dargestellt (A, C, G, T).

Der Datensatz zur Erstellung des Sequenzlogos SigH/SigE-abhängiger Promotoren bestand aus insgesamt 14 identifizierten Promotoren (Tab. IV.9 bzw. Anhang Tab. 4). Basierend auf der umfassenden Anzahl erstmalig identifizierter SigH- und zusätzlich SigE-abhängiger Promotoren konnte die klar definierte Konsensussequenz (g/tGGAAC –16-19bp– GTTg) ermittelt werden (Abb. IV-24). Exklusiv SigH-abhängige Promotoren und SigH/SigE-abhängige Promotoren haben somit eine identische –35-Kernkonsensus-Sequenz mit der Nukleotidfolge kGGAA.

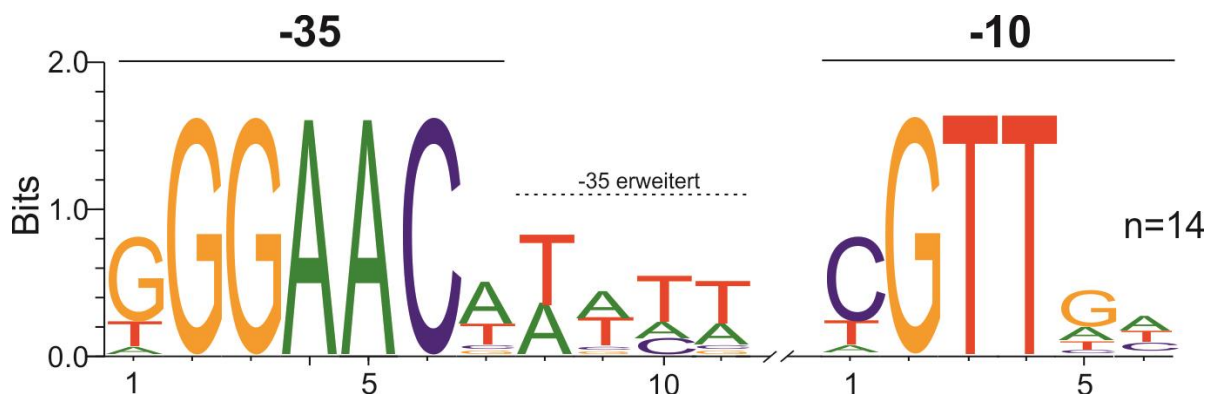


Abb. IV-24 DNA-Sequenzlogo SigH/SigE-abhängiger Promotoren. Das Sequenzlogo wurde auf Basis von insgesamt 31 in dieser Arbeit identifizierten SigH/SigE-abhängigen Promotoren durch WebLogo 3.2 (Crooks et al., 2004; <http://weblogo.threeplusone.com/>) erstellt. Der Informationsgehalt eines Nukleotids an einer bestimmten Position ist in Bits durch die Höhe des Nukleotidsymbols dargestellt (A, C, G, T).

Auch die partiell SigH-abhängigen Promotoren (P_{dnaJ2} , P_{clpC} , P_{clpX} , P_{clgR} , P_{serC} , $P_{as-cg3313}$ und $P_{as-cg3395}$) mit Ausnahme von P_{clpC} (–35-Motiv: GGAAG), für den keine zusätzliche Transkription durch SigE-gesättigte RNAP gezeigt werden konnte, weisen die –35- (g/tGGAAC) und –10-Promotormotive (GTT) auf (Anhang Tab. 3). Die SigH/SigE-abhängigen Promotoren zeigen, wie die SigH-exklusiven Promotoren, eine konservierte erweiterte –35-Region auf. Es handelt sich hier um eine T-reiche Region.

Die Konsensussequenz exklusiv SigH-abhängiger Promotoren unterscheidet sich in der Position 6 des –35-Promotorelements von SigH/SigE-abhängigen Promotoren. Die –35-Kernkonsensus-Sequenz ist mit der Nukleotidfolge kGGAA und die –10-Region mit der Sequenz GTT identisch. Während exklusiv SigH-abhängige Promotoren fast ausschließlich ein T an dieser Stelle vorweisen, in nur drei Promotoren ist ein G lokalisiert, ist an dieser Position in SigH/SigE-abhängigen Promotoren exklusiv ein C lokalisiert.

Für die putativ exklusiv SigE-abhängigen Promotoren (P_{cg1121} , P_{cg3309} und P_{amtR}) konnte auf Grund der geringen Anzahl an identifizierten Promotorsequenzen kein aussagekräftiges Sequenzlogo erstellt werden (Tab. IV.10). Diese Promotoren, die *in vitro* nur durch ROSE-SigE und nicht durch ROSE-SigH identifiziert werden konnten, besitzen nach dem -35-Kernmotiv g/tGGAAC SigH/SigE-abhängiger Promotoren immer ein T (P_{cg1121} , GGGAACT; P_{cg3309} , GGGAACT und P_{amtR} , GGGAACT). Es konnte aber für die SigH/SigE-abhängigen Promotoren der Gene *cseE*, *cg2657*, *sigB* und *dsbA2*, die dasselbe -35-Promotormotiv g/tGGAAC besitzen, *in vitro* durch ROSE-SigH eine Erkennung des Motivs durch den ECF-Sigmafaktor SigH beobachtet werden (Tab. IV.9 und Abb. IV-24).

V. Diskussion

Da in dieser Arbeit sowohl eine neue Methode entwickelt und etabliert, als auch umfassend neue Erkenntnisse über die Regulationsnetzwerke der ECF-Sigmafaktoren aus *Corynebacterium glutamicum* erzielt werden konnten, ist die Diskussion dieser Arbeit in zwei Abschnitte aufgeteilt. Der erste Abschnitt befasst sich mit den methodischen Entwicklungen dieser Arbeit, die zur Identifikation der ECF-Sigmafaktornetzwerke SigH und SigE geführt haben. Die eingesetzten und entwickelten Methoden werden darin hinsichtlich ihrer Vor- und Nachteile diskutiert. Im zweiten Abschnitt werden die Ergebnisse der Arbeit in Bezug auf die neuen Erkenntnisse der Regulationsnetzwerke der ECF-Sigmafaktoren aus *C. glutamicum* zusammengefasst und diskutiert.

1. Die Kombination von *Bottom-Up*- und *Top-Down*-Methoden zur Analyse von ECF-Sigmafaktornetzwerken ermöglicht zusammen mit RNA-Sequenzierung eine neue Betrachtungsweise von transkriptionellen Regulationsnetzwerken

In dieser Arbeit wurde neben der klassischen Transkriptomanalyse von gezielt hergestellten Mutanten unter Verwendung von *Microarrays* und der neueren Methode der RNA-Sequenzierung (RNA-Seq), die zusammenfassend als *Top-Down*-Methode beschrieben werden können, die neu entwickelte Methode ROSE (*Run-Off Transcription/RNA-Seq*) verwendet, die als *Bottom-Up*-Ansatz zu bezeichnen ist. Beide Methoden bieten sowohl Vorteile als auch Nachteile (Abb. V-1).

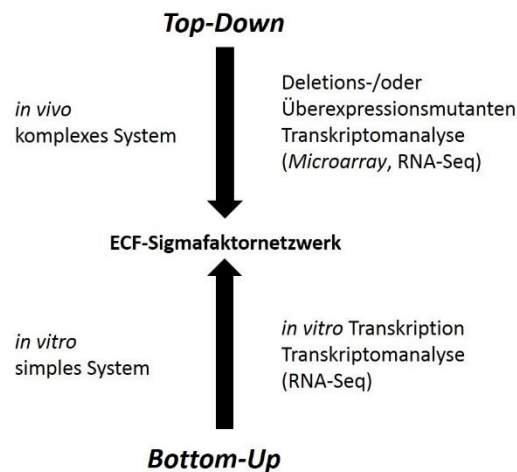


Abb. V-1 Konzept der Kombination aus *Top-Down*- und *Bottom-Up*-Ansatz zur Analyse von ECF-Sigmafaktornetzwerken.

Der Vorteil der klassischen *Top-Down*-Methode, die Auswirkungen einer Gendeletion oder einer Wachstumsbedingung *in vivo* auf Transkriptomebene zu untersuchen, ist gleichzeitig der Nachteil dieser Methode. Es steht außer Frage, dass die komplexe Antwort der Zelle auf eine Deletion bzw. veränderten Wachstumsbedingungen *in vivo* untersucht werden muss, um die teils komplexe Antwort der Zelle als Ganzes betrachten zu können. Die Komplexität des Systems *in vivo* stellt aber auch ein Hauptproblem für konventionelle oder klassische Methoden zur Analyse transkriptioneller Regulation durch einen einzigen Regulator dar. Die

Schwierigkeit ist es, direkte Effekte einer genetischen Veränderung oder von chemischen bzw. physikalischen Behandlungen von Zellen von nachgeschalteten, sekundären Effekten *in vivo* (*Top-Down*) zu unterscheiden. Hierzu zählt die zusätzliche Regulation durch weitere Transkriptionsfaktoren (TF) neben dem zu untersuchenden Regulator, die dessen Wirkung bei der „falschen“ Wahl an Versuchsbedingungen maskieren können. Zum anderen können Regulatoren, die unter normalen Wachstumsbedingungen inaktiv gehalten werden, auf diese Weise nicht untersucht werden, da die Deletion unter diesen Bedingungen keinen Effekt hat. Dies ist auch der Fall für ECF-Sigmafaktoren, die in den meisten Fällen von Anti-Sigmafaktoren unter Standardbedingungen inaktiv gehalten werden (Staron *et al.*, 2009). In dieser Arbeit wurden die genannten Probleme auf zweierlei Arten umgangen.

Zum einen wurde zu der Untersuchung der Regulationsnetzwerke der ECF-Sigmafaktoren SigH und SigE nicht nur Deletionsmutanten der entsprechenden Gene erstellt und untersucht, sondern auch Mutanten in den entsprechenden Anti-Sigmafaktorgenen *rshA* bzw. *cseE*.

So wurde erreicht, dass SigH bzw. SigE in den entsprechenden Anti-Sigmafaktormutanten konstitutiv aktiv sind und so die SigH- bzw. SigE-abhängigen Gene ohne Stressapplizierung transkribiert werden, da diese nicht mehr durch deren Anti-Sigmafaktoren RshA bzw. CseE gebunden und dadurch inaktiviert werden. Es wurden zwei vergleichende *Microarray*-Analysen durchgeführt und so die Induktion der Transkription der Mehrheit der SigH- bzw. SigE-abhängigen Gene durch die Deletion der Anti-Sigmafaktoren in Abwesenheit von Stress demonstriert. Die Untersuchung der ECF-Sigmafaktor-Regulons durch die Deletion der entsprechenden Anti-Sigmafaktoren ermöglichte die Analyse der Netzwerke unter Ausblendung der Effekte, die durch die Verwendung eines Stresszustandes zusätzlich auftreten. Ein solches Vorgehen wurde in der Literatur bis jetzt nicht erwähnt. Neben der Analyse mittels *Microarrays* wurden die Deletionsmutanten RES167 $\Delta sigH$ und $\Delta rshA$ durch die sensitivere Methode RNA-Seq untersucht und so die Transkriptionsstartpunkte (TSP) SigH-abhängiger Gene bestimmt und unterschiedliche Aktivität der SigH-abhängigen Promotoren in den Deletionsderivaten beobachtet. Ergebnisse, die durch klassische *Microarray*-Analysen nicht erzielt worden wären.

Da durch die *in vivo* Analyse (*Top-Down*) nicht sichergestellt werden konnte, dass es sich ausschließlich um die primären Effekte der Erkennung von Promotoren der beiden ECF-Sigmafaktoren handelt, wurde zusätzlich die in dieser Arbeit entwickelte und erstmalig eingesetzte *in vitro* (*Bottom-Up*) Methode ROSE eingesetzt (Abb. V-1). Diese Methode erlaubte die Analyse der Aktivität der ECF-Sigmafaktoren in Abwesenheit aller *in vivo* auftretenden sekundären Effekte bzw. ohne den Einfluss zusätzlicher Repressoren, wie es für

den Vorläufer der Methode **Run-Off-Transcription/Microarray Analyses** (ROMA) bereits beschrieben ist (Maclellan *et al.*, 2009). Der Einsatz von RNA-Seq anstelle von *Microarrays* erlaubte die direkte Identifizierung von TSP und damit der entsprechenden Promotoren. Hierdurch war es in dieser Arbeit erstmals möglich die SigH-Abhängigkeit SigH-abhängiger Promotoren, die zusätzliche Erkennung einiger SigH-abhängigen Promotoren durch SigE und die potentielle exklusive Abhängigkeit einiger Promotoren von SigE *in vitro* nachzuweisen.

Letztendlich konnte durch die Kombination der *Bottom-Up- (in vitro)* und *Top-Down-Methodik (in vivo)* sichergestellt werden, dass die *in vitro* identifizierten Promotoren auch *in vivo* durch einen der ECF-Sigmafaktoren erkannt werden und so zu einer Transkription der entsprechenden Gene führt.

1.1. RNA-Sequenzierung ermöglicht die genomweite qualitative und quantitative Betrachtungsweise von Promotoren

Die Methode der RNA-Seq ermöglicht neben der quantitativen Analyse die Identifizierung der TSP eines bakteriellen Transkriptom und damit auch die Lokalisierung der entsprechenden Promotoren auf genomweiter Ebene und bietet damit einen entscheidenden Vorteil gegenüber klassischen Methoden, wie der *Microarray*-Analyse (Hoen *et al.*, 2008). In dieser Arbeit wurde eine Modifikation, die differentielle RNA-Seq (dRNA-Seq) (Sharma *et al.*, 2010) eingesetzt (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung). Bei dRNA-Seq werden nur native Transkripte sequenziert, wodurch sich die Methode besonders zur eindeutigen Identifizierung von TSP und damit Promotoren eignet.

In dieser Arbeit konnten durch den Einsatz von RNA-Seq 37 der 45 bereits beschriebenen SigH-abhängigen Promotoren (Busche *et al.*, 2012) identifiziert werden. Für 15 Promotoren hiervon wurden die bis dahin lediglich vorhergesagten TSP experimentell bestätigt. Durch einen Vergleich der relativen Anzahl *gemappter Reads* der RNA-Seq-Datensätze an den Positionen der identifizierten TSP konnte gezeigt werden, dass die Deletion von *sigH* oder *rshA* Einfluß auf die Promotoraktivität der entsprechenden Promotoren haben und die Vermutung aufgestellt werden, dass diese Promotoren zusätzlich von noch einem anderen Sigmafaktor erkannt werden. Hier konnte insbesondere auch der Einfluss der Deletionen auf die Promotoraktivität von Genen, die in der *Micorarray*-Analyse des Deletionsstammes RES167 $\Delta rshA$ nicht auffällig waren (*dnaK*, *clpC*, *sigB*, *sufR* und *clgR*), nachgewiesen werden. Dies zeigt deutlich, dass RNA-Seq wesentlich sensitiver und hochauflösender als eine *Microarray*-

Analyse ist und neben der Bestimmung von TSP auch quantitative Aussagen zur Promotoraktivität zulässt. Ein weiterer Vorteil ist, dass RNA-Seq als Analysemethode des Transkriptom nicht wie ein *Microarray* auf die Vorauswahl potentieller *Targets* angewiesen ist (*Microarray-Design*). So war es in dieser Arbeit erstmals möglich die SigH-abhängige Transkription von *anti-sense* RNA festzustellen.

Bei dem Vergleich der TSP, die durch RNA-Seq bestimmt worden sind, mit denen, die bereits in anderen Arbeiten bestimmt worden waren (Halgasova *et al.*, 2001; Nakunst *et al.*, 2007; Park *et al.*, 2008; Ehira *et al.*, 2009; Busche *et al.*, 2012;), fiel auf, dass es vereinzelt Abweichungen von ein bis zwei Nukleotiden gab. Insbesondere die Ergebnisse der *Primer-Extension* (PEX)-Methode zeigten Abweichungen. Ausschlaggebend ist hier, dass PEX nicht die Präzision anderer Methoden aufweist und teilweise ein *Shift* von einer Base bei der Identifizierung eines TSP vorkommen kann (Patek, persönliche Mitteilung). Gegenüber PEX besitzt RNA-Seq einen weiteren Vorteil. Die Transkripte, die durch RNA-Seq analysiert werden, werden *in vivo* auf Basis von gDNA transkribiert. Das heißt die einem Promotor angrenzenden Bereiche haben Einfluss auf die Promotoraktivität. Die Analyse von TSP per PEX basiert hingegen auf der Nutzung von Vektoren, in der die zu untersuchenden DNA-Abschnitte kloniert sind. Hier wird eine Auswahl eines DNA-Abschnittes getroffen auf dessen Basis die Transkription *in vivo* abläuft. Dies kann einen Einfluss auf die Transkription haben und sich somit von der tatsächlichen Situation *in vivo* unterscheiden. Weiter ist der zu untersuchende DNA-Bereich durch die Nutzung eines Vektors in mehrfacher Kopienzahl vorhanden. Hierbei kann nicht ausgeschlossen werden, dass Titrierungseffekte in Bezug auf DNA-bindende Proteine auftreten können. Ein Beispiel hierfür ist der Promotor P4 des Gens *sigH*. Dieser konnte nicht durch RNA-Seq, aber durch PEX nachgewiesen werden. In beiden Analysen wurde RNA von unter Standardbedingungen wachsenden *C. glutamicum* Kulturen isoliert und verwendet. Da P4 des Gens *sigH* mit einer SOS-Box überlappt, sollte dieser Promotor unter diesen Bedingungen geblockt sein, was durch RNA-Seq, aber nicht durch PEX, gezeigt werden konnte. Dies bedeutet, dass LexA entweder aufgrund der Konformation des DNA-Abschnittes innerhalb des Vektors nicht binden konnte oder aber das unter Standardbedingungen nicht genug LexA in der Zelle vorhanden war, um diesen Promotor vollständig zu blockieren und so ein TSP per PEX identifiziert werden konnte. RNA-Seq eignet sich somit deutlich besser, um die tatsächlich in der Zelle vorherrschende Transkription zu analysieren und dies genomweit.

1.2. **Run-off-Transcription/RNA-Sequencing (ROSE) ermöglicht die Genomweite Identifizierung von Promotoren unter Ausblendung aller sekundären Effekte**

In dieser Arbeit wurde die Methode ROMA zur Untersuchung von ECF-Sigmafaktor-Regulationnetzwerken weiterentwickelt. ROMA erlaubt die Analyse der Aktivität eines Sigmafaktors oder TF in Abwesenheit aller *in vivo* auftretenden sekundären Effekte zu untersuchen (MacLellan *et al.*, 2009). Die Weiterentwicklung durch den Einsatz von RNA-Seq erlaubte die direkte Identifizierung von TSP und damit der entsprechenden Promotoren in Abhängigkeit von ECF-Sigmafaktoren. Durch die Nutzung dieser erstmalig eingesetzten Methode ROSE konnten einige der Beschränkungen der Methode ROMA beseitigt werden. Durch auftretendes *Read-Through* in der *in vitro* Transkription können Transkripte entstehen, die *in vivo* nicht auftreten und so zu falsch positiven Ergebnissen in der *Microarray*-Analyse der ROMA-Methode führen. Insbesondere der *Read-Through* in konvergente Gene ist *in vitro* häufiger zu beobachten als *in vivo* (MacLellan *et al.*, 2009). Da in dieser Arbeit nicht die kompletten Transkripte der *in vitro* Transkription untersucht wurden, sondern ausschließlich die nativen 5'-Enden der Transkripte und hierüber die direkte Identifizierung der TSP stattfand, hat die Beschränkung durch *Read-Through* der *in vitro* Transkription keinen Einfluß auf die Ergebnisse. Der offensichtlichste Vorteil der Methode ROSE ist hierbei die direkte Identifizierung von TSP. ROMA ermöglicht lediglich die Identifizierung möglicher Target-Gene die in Abhängigkeit eines Sigmafaktors transkribiert werden. Die Bestimmung der TSP muss einzeln für jedes dieser Gene erfolgen.

Ein Problem dieser Arbeit bei der Nutzung von ROSE war, dass für die *in vitro* Transkription nicht das RNA-Polymerase (RNAP)-Core Enzym von *C. glutamicum* eingesetzt werden konnte, sondern das RNAP-Core Enzym aus *Escherichia coli*. Diese Herangehensweise wurde zwar bereits für *Mycobacterium tuberculosis* erfolgreich beschrieben (Rodrigue *et al.*, 2007). Auf Grund des heterolog Zusammengesetzten RNAP-Holo Enzyms bestanden Zweifel ob eine erfolgreiche und spezifische *in vitro* Transkription durch die eingesetzten ECF-Sigmafaktoren aus *C. glutamicum* möglich ist. Diese Zweifel konnten aber durch die hier erzielten Ergebnisse ausgeräumt werden, wie auch für *M. tuberculosis* gezeigt. das RNAP-Core Enzym aus *E. coli* scheint in Kombination mit aktinobakteriellen Sigmafaktoren zu funktionieren.

In dieser Arbeit konnte gezeigt werden, dass ROSE eine essentielle Methode bei der Untersuchung von ECF-Sigmafaktornetzwerken ist, da diese zumindest im Fall von *C.*

glutamicum zum Großteil die Promotoren nur sequenzspezifisch erkennen. Speziell bei den sich überlappenden Regulons und Regulons die eine Vielzahl von zusätzliche Regulatoren umfassen ist der Einsatz dieser Methode entscheidend. Nur durch den Einsatz von ROSE konnte eindeutig geklärt werden, welche Promotoren von den ECF-Sigmafaktoren SigH und SigE in *C. glutamicum* erkannt werden.

1.3. Vergleich der Methoden zur Identifizierung und zum Vergleich von ECF-Sigmafaktornetzwerken

In dieser Arbeit wurden drei Methoden zur Identifizierung des SigH- bzw. SigE-Regulationsnetzwerkes eingesetzt. Wie oben beschrieben wurden in dieser Arbeit *Microarrays* eingesetzt, die es ermöglichten die Erhöhung der relativen Transkriptmenge von Genen durch die Deletion der Anti-Sigmafaktorgene *rshA* und *cseE* zu detektieren. Daneben wurden die ebenfalls auf *Microarray*-Daten beruhenden Ergebnisse anderer Arbeitsgruppen in dieser Arbeit ebenso berücksichtigt (Ehira *et al.*, 2009; Park *et al.*, 2008; Nakunst *et al.*, 2007). Als sensitivere Ergänzung zu den hier durchgeführten *Microarray*-Analysen und zur Bestimmung der TSP auf einem genomweiten Level wurden die isolierten RNA-Spezies der Deletionsstämme RES167 $\Delta sigH$ und $\Delta rshA$ mittels der RNA-Seq auf TSP und Änderungen der Promotoraktivität der potentiell SigH-abhängigen Promotoren überprüft. Hier konnten die TSP für 37 der bereits bekannten SigH-abhängigen Promotoren bestätigt bzw. um wenige Nukleotide korrigiert werden. Zuletzt konnte durch die Verwendung von ROSE die Mehrheit der SigH-abhängige Promotoren *in vitro* identifiziert werden. Neben den bereits bekannten SigH-abhängigen Promotoren konnte durch die Kombination der sensitiven *Top-Down*-Methode RNA-Seq der Deletionsstämme RES167 $\Delta rshA$ und $\Delta sigH$ und des *Bottom-Up*-Ansatzes ROSE-SigH 23 SigH-abhängige Promotoren identifiziert werden, die in den *Microarray*-Analysen nicht auffällig waren. Der Grund hierfür ist zum einen die höhere Sensitivität der Methode RNA-Seq und zum anderen, dass ROSE es ermöglicht, die Aktivität Sigmafaktor-gesättigter RNAP unter Ausschluss negativer Regulatoren, die *in vivo* Einfluss haben, zu analysieren.

Insgesamt konnten so 60 Promotoren identifiziert werden, die SigH-abhängig sind und auf Basis der Operonstruktur für die Transkription von insgesamt 135 Gene verantwortlich sind. Dabei führte die Kombination der *Bottom-Up*- (*in vitro*) und *Top-Down*-Methodik (*in vivo*) zu einem größeren Umfang der identifizierten Promotoren bei gleichzeitiger Kreuzvalidierung. Ferner konnten unter Verwendung von ROSE-SigE 14 Promotoren identifiziert werden, die zusätzlich zu SigH von SigE erkannt werden, die insgesamt die Transkription von 34 Gene

regulieren, sowie drei Promotoren, die putativ exklusiv SigE-abhängig sind (insgesamt sieben Gene). ROSE konnte somit die Vermutung, dass einige SigH-abhängige Promotoren durch einen weiteren Sigmafaktor erkannt werden, bestätigen und den ECF-Sigmafaktor SigE als verantwortlich ermitteln.

Beim Vergleich der Ergebnisse der einzelnen Methoden zeigt sich, dass mit nur einer der drei Methoden nicht alle identifizierten Promotoren bestimmt werden konnten (Abb. V-2). Wie auch in einer Studie zu SigW in *B. subtilis*, in der ROMA in Kombination mit Microarray-Analyse von Deletionsmutanten und *in silico* Promotorbestimmung verwendet wurde, konnten nicht alle Gene mit nur einer dieser Methoden identifiziert werden (Cao *et al.*, 2002). Dies lässt sich zum einen durch den vorteilhaften Einsatz der *Bottom-Up*-Methode ROSE bzw. ROMA erklären, aber auch durch den Optimierungsbedarf der erstmalig in dieser Arbeit eingesetzten Methode erklären. Die unterschiedlichen Ergebnisse zeigen zudem deutlich, wie wichtig sich ergänzende Ansätze zur Analyse von ECF-Sigmafaktor-Netzwerken bzw. Regulationsnetzwerken im Allgemeinen sind.

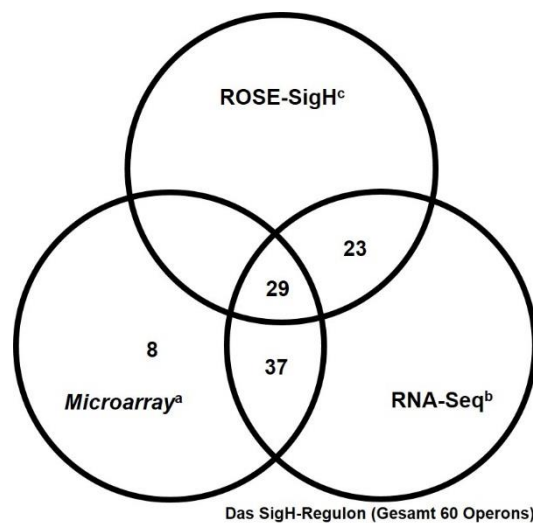


Abb. V-2 Grafische Darstellung der Überschneidungen zwischen den drei angewandten Methoden zur Bestimmung des SigH-Regulon in *C. glutamicum*. Für die 60 SigH-abhängigen Promotoren ist die Anzahl der Promotoren, die durch eine der Methoden oder durch den Kreuzabgleich mehrerer Methoden identifiziert werden konnten angegeben. A) *in vivo* Microarray-Analyse (Busche *et al.*, 2012; Ehira *et al.*, 2009; Engels *et al.*, 2004); B) *in vivo* RNA-Seq der Deletionsderivate RES167 $\Delta sigH$ und RES167 $\Delta rshA$; C) *in vitro* ROSE-SigH.

Die Gründe für die unterschiedlichen Ergebnisse der Methoden werden im Folgenden diskutiert. Der Unterschied zwischen den Ergebnissen der *in vivo Microarray*-Untersuchungen und der ROSE-Methodik war erwartet worden, da ROSE eingesetzt wurde, um die primären Effekte eines ECF-Sigmafaktors auf die Transkription unter Ausschluss sekundärer Effekte untersuchen zu können. Die Gene, die unter den gewählten Bedingungen in den *in vivo Microarrays* nicht auffällig waren, sind offensichtlich unter diesen Bedingungen durch die Aktivität anderer Regulationsmechanismen beeinflusst und daher nicht deutlich differentiell transkribiert.

Die acht Gene, die in *Microarray*-Analysen auffällig waren, für die aber per RNA-Seq keine SigH-abhängige Transkription nachgewiesen werden konnte (Abb. V-2), wiesen bereits im *Microarray* RES167 (Referenz) gegen RES167 $\Delta rshA$ nur eine geringere oder keine Änderung in ihrer relativen mRNA-Menge im Vergleich zur Referenz auf. Offenbar ist die Anzahl der SigH-abhängigen Transkripte dieser Gene zu gering, um per RNA-Seq-Analyse einen TSP identifizieren zu können. Eine Ausnahme stellt der Promotor des Gens *cg3344* dar, hier konnte zwar der durch Ehira und Mitarbeiter identifizierte Promotor nicht detektiert werden. In der hier durchgeführten Analyse wurde ein anderer TSP bestimmt, der mit dem Startnukleotid des Gens zusammenfällt.

Der Grund für die gleichzeitig ausbleibende Transkription dieser Gene in den ROSE-Experimenten ist eventuell die minimale Abweichung von der -35 -Kernkonsensussequenz kGGAA SigH-abhängiger Promotoren (0). Die Promotomotive dieser Gene weichen in einem Nukleotid von dem Konsensusmotiv ab (Abb. IV-9). Eine Ausnahme stellt P_{clpP_1} dar. Hier weicht die -10 -Region (ATTG) in einem Nukleotid von der -10 -Konsensussequenz (GTT) ab.

An dieser Stelle kann spekuliert werden, dass es bei Promotoren, die eine Abweichung in den Kernmotiven aufweisen, zur Erkennung durch die ECF-Sigmafaktoren SigH und SigE zusätzlich einer bestimmten Konformation der DNA bedarf. Dies würde erklären, warum eine Identifizierung mittels ROSE nicht möglich war, da in den ROSE-Analysen lineare DNA als Template genutzt wird und somit nur die rein sequenzspezifische Interaktion zwischen Sigmafaktor-gesättigter RNAP und Promotor nachgewiesen werden kann. Dies müsste aber noch experimentell gezeigt werden.

Die Promotoren, die per *Microarray*-Analyse und RNA-Seq identifiziert werden konnten (37), die aber nicht durch ROSE-SigH bestätigt werden konnten, sind entweder nicht direkt von SigH

abhängig oder die Gründe sind in der Versuchsdurchführung des ROSE-Ansatzes zu finden (Abb. V-2). Es handelt sich um die Promotoren der Gene *rshA*, *dsbA1*, *murl*, *trxB*, *cg0184*, *cg0497*, *cg2106* und *cg2538* (Anhang Tab. 3).

Von diesen acht Genen besitzen drei (*rshA*, *cg0497* und *dsbA1*) eine *EcoRI*-Schnittstelle in dem Bereich +300 bis -200 des Translationstarts. Da für die *in vitro* Run-Off-Transkription *EcoRI*-gespaltene gDNA eingesetzt wurde, sind die entstehenden Transkripte dieser Gene vermutlich zu klein, um mit dem hier genutzten Protokoll der RNA-Seq identifiziert zu werden. In Folgeuntersuchungen sollte auf durch mechanisches Scheren fragmentierte gDNA zurückgegriffen werden.

Die übrigen per *Microarray* und RNA-Seq identifizierten SigH-abhängigen, aber nicht *in vitro* transkribierten Gene weisen keine *EcoRI*-Schnittstellen im Bereich des TSP auf. Bei diesen Transkripten könnte, wie bereits erwähnt, der Fehler in der Wahl linearer DNA als *Template* für die ROSE-Analysen liegen. Um dies aufzuklären, sollten die ROSE-Experimente mit Gesamt-DNA unterschiedlicher Konformationen durchgeführt werden oder eine *in vitro* Transkriptionsanalyse auf Basis von *supercoiled* Plasmiden, die den fraglichen Promotor tragen, durchgeführt werden (Maciag *et al.*, 2011). Für die große Mehrheit der in dieser Arbeit identifizierten Promotoren ist die Konformation der DNA aber nicht ausschlaggebend. Hier scheint allein die Promotorsequenz für die Identifizierung durch den ECF-Sigmafaktor SigH bzw. SigE verantwortlich zu sein.

So stellt die etablierte ROSE-Methodik zwar sicherlich eine entscheidende Erweiterung im methodischen Spektrum zur genomweiten Analyse von Sigmafaktor-Netzwerken dar, sie bedarf aber noch der methodischen Optimierung.

2. Das SigH-Regulationsnetzwerk in *Corynebacterium glutamicum*

SigH zählt zu den SigR-ähnlichen Proteinen der ECF-Gruppe ECF12, die vorwiegend in Actinobakterien zu finden sind. Diesen ECF-Sigmafaktoren ist gemein, dass sie mit wenigen Ausnahmen mit einem RsrA-ähnlichen Anti-Sigmafaktor assoziiert sind (Staron et al., 2009). Es ist bekannt, dass SigH in *C. glutamicum* und in anderen Actinomyceten die Antwort auf Hitze- und oxidativen Stress reguliert (Engels et al., 2004; Kim et al., 2005).

Für die nahe verwandten Actinobakterien *Mycobacterium tuberculosis* und *Streptomyces coelicolor* konnte gezeigt werden, dass die Aktivität des SigH-Orthologs durch den Anti-Sigmafaktor RshA bzw. RsrA kontrolliert wird (Song et al., 2003; Kang et al., 1999). Der Anti-Sigmafaktor und gleichzeitige Stresssensor bindet in *M. tuberculosis* unter ungestressten Bedingungen den ECF-Sigmafaktor SigH und unterbindet so die SigH-abhängige Transkription von Genen. In *C. glutamicum* ist das zu *M. tuberculosis* homologe Anti-Sigmafaktorgen *rshA* nur zwei bp stromabwärts des *sigH*-Gens angeordnet. In *S. coelicolor* wird das *sigR-rsrA*-Operon von multiplen Promotoren aus transkribiert.

Ob RshA in *C. glutamicum* die gleiche Funktion, wie in *M. tuberculosis* erfüllt und ob eine ähnliche transkriptionelle Organisation des Operons vorliegt, war bisher nicht gezeigt worden.

In dieser Arbeit wurde die Untersuchung des Regulationsnetzwerks von SigH in *C. glutamicum* aus einem anderen Blickwinkel als bisher durchgeführt. Hier wurde das Regulationsnetzwerk von SigH zum einen durch die Deletion des Anti-Sigmafaktors RshA in Abwesenheit von Stress induziert und durch *Micorarrays* und RNA-Seq untersucht, und zum anderen wurde die Erkennung SigH-abhängiger Promotoren durch SigH mittels ROSE-SigH *in vitro* nachgewiesen. Die Kombination aus RNA-Seq und ROSE ermöglichte weiterhin die zusätzliche Identifizierung von SigH-Promotoren, deren Gene in den *Micorarray*-Analysen nicht differentiell transkribiert waren.

2.1. Das *sigH-rshA*-Operon in *Corynebacterium glutamicum* weist eine komplexe transkriptionelle Organisation und eine Autoregulation auf

In dieser Arbeit wurde die im Vergleich zur Referenz relativ erhöhte Transkriptmenge der Mehrheit der bekannten SigH-abhängigen Genen bzw. eine verstärkte Aktivität der entsprechenden Promotoren durch die Deletion des putativen Anti-Sigmafaktorgens *rshA* in Abwesenheit von Stress demonstriert. Damit übereinstimmend führte die Deletion des ECF-Sigmafaktorgens *sigH* zu einem Verlust oder der Verringerung der Promotoraktivität. Durch ROSE-SigH konnte bewiesen werden, dass SigH für diese Effekte direkt verantwortlich ist. Die Proteinsequenz von RshA ist ähnlich zu der anderer Anti-Sigmafaktoren aus *M. tuberculosis* (RshA; (T. Song et al., 2003)) oder aus *S. coelicolor* (RsrA; (Kang et al., 1999)). RshA aus *C. glutamicum* besitzt alle konservierten Cysteinreste innerhalb der ZAS-Domäne. Diese Aminosäurereste vermitteln die Interaktion mit dem SigH-Protein, eine Tatsache, die für RshA/SigH in *M. tuberculosis* (T. Song et al., 2003) und RsrA/SigR in *S. coelicolor* (Kang et al., 1999), aber auch für andere Mitglieder der ZAS-Domäne enthaltenden Proteinfamilie innerhalb der Actinobakterien (Jung et al., 2011) experimentell bewiesen wurde. Die eindeutige transkriptionelle Aktivierung der bekannten SigH-abhängigen Gene in der erstellten *rshA*-Deletionsmutante in Zusammenhang mit dem Beweis der primären SigH-Abhängigkeit dieser Gene durch ROSE-SigH ist ein weiterer Beleg dafür, dass RshA in *C. glutamicum* die Funktion eines Anti-Sigmafaktors, wie RshA in *M. tuberculosis* und RsrA in *S. coelicolor*, ausübt.

Auch die *sigH-rshA*-Genorganisation ist in allen sequenzierten *Corynebacterium* Spezies, die in der NCBI-Datenbank hinterlegt sind, und dem weiter entfernt verwandten Actinobakterium *S. coelicolor* konserviert. Für *M. tuberculosis* liegt eine ähnliche Anordnung der beiden Gene vor, mit der Einschränkung dass hier ein Gen, welches für ein Protein mit unbekannter Funktion kodiert, zwischen dem *sigH*- und *rshA*-Gen liegt.

Die transkriptionelle Organisation des *sigH-rshA*-Operons von *C. glutamicum* weist Ähnlichkeiten zur transkriptionellen Organisation des Operons in *M. tuberculosis* und *S. coelicolor* auf, ist aber nicht identisch. In *C. glutamicum* wird das gesamte Operon von vier *Housekeeping*-Promotoren, erkannt durch SigA, transkribiert. Der Grund für die Verwendung von multiplen Promotoren mag in der Sicherstellung der Möglichkeit zum *Fine-Tuning* liegen. Dieses *Fine-Tuning* kann entweder durch die Aktivität von Transkriptionsfaktoren, die die entsprechenden Promotoren beeinflussen, oder durch die unterschiedliche Affinität der Sigmafaktoren SigA und SigB zu den Promotoren geregelt werden. SigB ist der nicht-essentielle Sigmafaktor von *C. glutamicum*, der auch *Housekeeping*-Promotoren erkennt

(Larisch et al., 2007). Die experimentellen Beobachtungen decken sich mit dieser Annahme: Es wurde mit Reporter-Fusions-Studien (*P-sigH::cat*) gezeigt, dass die Aktivität der *sigH*-Promotoren in der stationären Phase und nach oxidativem Stress (Kim et al., 2005) erhöht ist, aber keine signifikante Änderung in der relativen Anzahl an *sigH*-Transkripten nach Hitzeschock (Barreiro et al., 2009) oder in der Transitionsphase des Wachstums (Nakunst et al., 2007) stattfindet. In *S. coelicolor* wird das *sigR-rsrA*-Operon ebenfalls von multiplen Promotoren aus transkribiert. Hier gibt es neben einem TSP für *sigR*, der von dem Housekeeping-Sigmafaktor SigA abhängig ist, einen weiteren, der von SigR^{Sc} selbst abhängig ist (Kim et al., 2009). Für *M. tuberculosis* ist für die Transkription von *sigH* nur eine Autoregulation durch SigH bekannt (T. Song et al., 2003).

Die zusätzliche Regulation von *sigH* in *C. glutamicum* wird vermutlich unter anderem durch den Promotor, der am weitesten stromaufwärts vom *sigH*-Translationstart liegt, vermittelt. Dieser Promotor überlappt mit einer SOS-Box (Jochmann et al., 2009) und wird in Abwesenheit von DNA-schädigenden Agentien wahrscheinlich durch den LexA-Repressor geblockt.

Den Hauptunterschied in der transkriptionellen Regulation des Operons *sigH-rshA* von *C. glutamicum* zur Regulation in den verwandten Actinobakterien *S. coelicolor* und *M. tuberculosis* lässt sich in der Beobachtung, dass das *rshA* Gen in *C. glutamicum* zusätzlich als monocistronisches Transkript transkribiert wird, finden. Es konnte gezeigt werden, dass diese Transkription SigH-abhängig ist. Interessant ist hierbei, dass der SigH-abhängige Promotor in allen bis zum aktuellen Stand dieser Arbeit sequenzierten Corynebakterien mit Ausnahme von *C. jeikeium* und *C. urealyticum*, zweier pathogener Vertreter, zu finden ist. Es scheint somit eine Besonderheit der nicht-pathogenen Corynebakterien zu sein.

Diese transkriptionelle Organisation hat sich vermutlich entwickelt, um SigH auch ohne Stressapplikation konstitutiv zu exprimieren. In Abwesenheit von Stress ist SigH aber durch die Bindung von RshA inaktiviert. Dies gewährleistet eine besonders schnelle Aktivierung von SigH unter Stressbedingungen, die die RshA-Bindung aufheben. Die zusätzliche Transkription des *rshA*-Gens durch einen SigH-abhängigen Promotor kann dann sicherstellen, dass unter den spezifischen Stressbedingungen, unter denen SigH aktiv ist, immer ein Überschuss an RshA über SigH vorliegt. Dadurch kann dann die schnellstmögliche Abschaltung der SigH-abhängigen Transkription durch die Reaktivierung des im Überschuss vorliegenden Anti-Sigmafaktors RshA nach Beendigung des Stresszustandes gewährleistet werden.

2.2. Das bekannte SigH-Regulationsnetzwerk konnte durch die globale Expressionsanalyse der *rshA*-Deletionsmutante und dem Einsatz von *Run-Off-Transcription/RNA-Sequencing* validiert und erweitert werden

SigH ist einer der Hauptregulatoren in *C. glutamicum*. Dies wird speziell bei Hitzestress klar, bei dem aber auch noch eine große Anzahl an verschiedenen anderen transkriptionellen Regulatoren beteiligt sind (Patek und Nesvera, 2011). Um das Regulon von SigH zu untersuchen, wurde in dieser Arbeit nicht der klassische Weg gewählt, die Folgen der Sigmafaktoraktivität durch die Nutzung eines induzierenden Stresses zu untersuchen, sondern folgender: SigH wurde von der negativen Regulation des Anti-Sigma-Faktors RshA durch dessen Deletion entkoppelt. So konnte das Regulon in Abwesenheit von Stress und den damit einhergehenden möglichen Sekundäreffekten auf die Transkription von Genen untersucht werden. In Kombination wurde zusätzlich die Methode ROSE unter Verwendung des ECF-Sigmafaktors SigH eingesetzt, um sekundäre Effekte ausschließen, zu können.

In der Deletionsmutant $\Delta rshA$ im Vergleich zum Referenzstamm (RES167) konnte durch *Microarray*-Analysen eine erhöhte Transkription aller bereits beschriebenen, SigH-abhängigen Genen (Ehira *et al.*, 2009), mit der Ausnahme des *dnaK-grpE* Operons, *clpC*, *sigB* und *sufR* beobachtet werden. Ebenso wie bei Ehira und Mitarbeitern, die mit Überexpression und Deletion des Gens *sigH* arbeiteten, konnte auch in dieser Arbeit kein Nachweis für eine differentielle Transkription des SigH-abhängigen Gens *clgR* erbracht werden. Die schwache transkriptionelle Induktion einiger SigH-abhängiger, Hitzeschock-relevanter Gene und die fehlende Induktion oben erwähnter Gene kann durch dominierende Sekundär-Effekte hervorgerufen durch andere bekannte Transkriptionsregulatoren erklärt werden, die unter den genutzten Standardbedingungen die Transkription beeinflussen. Zu diesen transkriptionellen Regulatoren zählen ClgR, HrcA, HspR und SufR (Barreiro *et al.*, 2009; Ehira *et al.*, 2009; Engels *et al.*, 2004) (Abb. V-3). Die zusätzliche Beteiligung dieser Regulatoren ermöglicht erst unter Hitze- oder oxidativem Stress die SigH-abhängige Transkription der unter ungestressten Bedingungen geblockten Promotoren (s.o.). Diese Erklärung könnte sich auch für das Gen *sigB*, das für den nicht essentiellen Sigmafaktor SigB kodiert, als richtig herausstellen. SigB ist in die Genexpression während der Transitionsphase des Wachstums involviert und die Probenahme für die *Microarray*-Analysen dieser Arbeit fand während der exponentiellen Wachstumsphase statt. Das heisst, wie für die anderen in den hier durchgeführten *Microarray*-Analysen nicht induzierten Gene bereits diskutiert, dass auch für die transkriptionelle

Aktivierung von *sigB* durch SigH eventuell weitere, in diesem Fall Wachstumsphasenspezifische Faktoren nötig sein können. Durch den Einsatz der sensitiveren Methode der RNA-Seq konnte für die oben genannte Gene allerdings eine Erhöhung der Promotoraktivität gezeigt werden und durch ROSE-SigH der Nachweis erbracht werden, dass SigH für die Erkennung der Promotoren verantwortlich ist.

Neben dem alternativen Sigmafaktor SigB konnte auch das Gen für den *Housekeeping*-Sigmafaktor SigA als Bestandteil des SigH-Regulationsnetzwerkes in *C. glutamicum* identifiziert werden. In *S. coelicolor* konnte ebenfalls eine SigR-abhängige Transkription des *Housekeeping*-Sigmafaktorgens festgestellt werden. Es konnte klar gezeigt werden, dass SigR für die Aufrechterhaltung der Aktivität des *Housekeeping*-Sigmafaktors während Thioloxidativem Stress nötig ist (Kim *et al.*, 2012).

Als weiterer Bestandteil des Sigmafaktor-Netzwerkes in *C. glutamicum* konnte auch das SigE-Regulon mit SigH in Verbindung gebracht werden. Hier wird im Unterschied zu *M. tuberculosis* nicht das gesamte Operon aus ECF-Sigmafaktor-/Anti-Sigmafaktorgen SigH-abhängig transkribiert sondern lediglich das Anti-Sigmafaktorgen *cseE*.

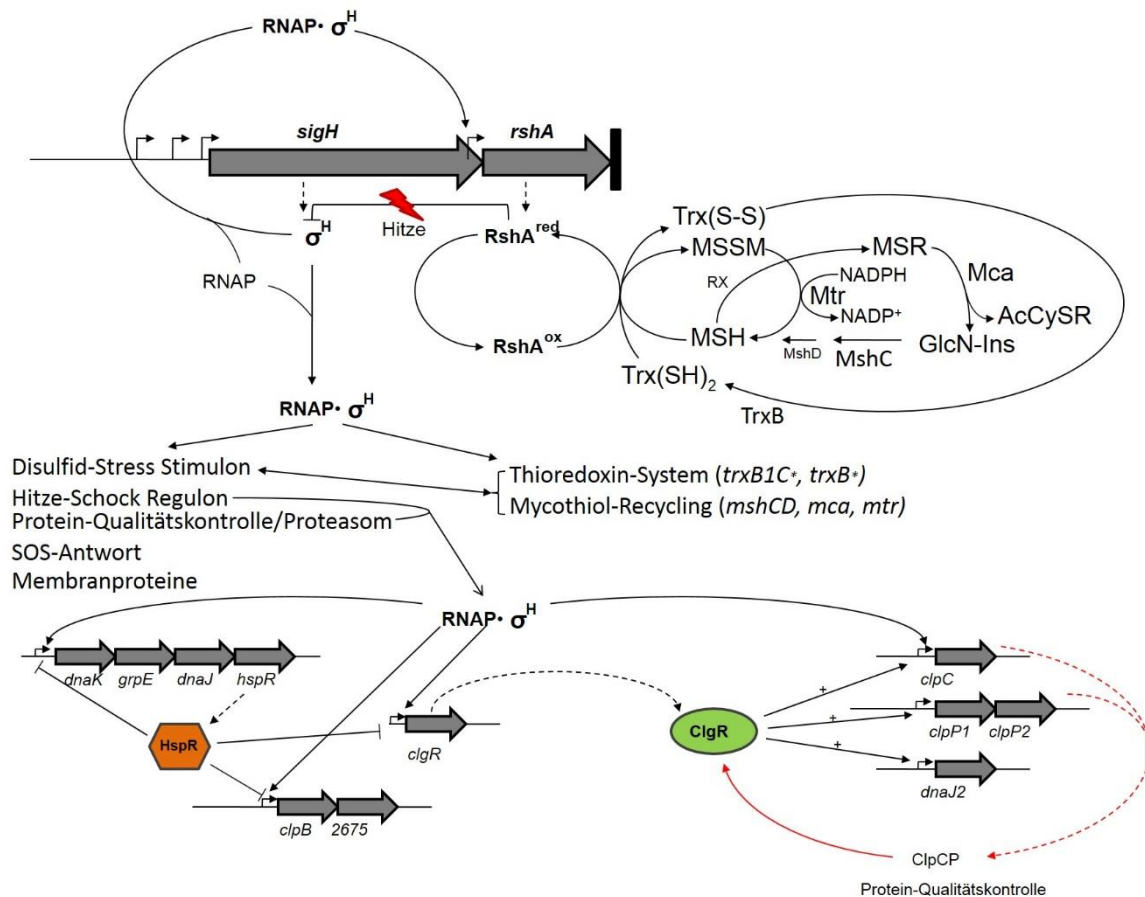


Abb. V-3 Erweitertes Modell des SigH Regulationsnetzwerkes in *C. glutamicum*. Bedingungen, die den Zell-Pool an Thiole durch Oxidation oder Alkylierung verbrauchen führen zu einer Oxidation der Cystein-Reste innerhalb von RshA. RshA^{ox} bindet nicht mehr an SigH, welches nun an das RNA Polymerase (RNAP) core-Enzym bindet und die Transkription seines Regulons induziert. Die direkte Induktion der Gene *trxB*/*trxB1*, *mca*, *mshC* und *mtr*, die an der Disulfid-Stressantwort beteiligt sind, generieren und recyceln/reduzieren die Thiole Thioredoxin (Trx) und Mycothiol (MSH) um die Oxidation von RshA umzukehren sowie die Thiol Redox-Balance und die Bindung von RshA an SigH wiederherzustellen. Die direkte Induktion von *rshA* als separates Transkript verstärkt die Abschaltung der SigH-abhängigen Stressantwort nachdem die Zellen den Stress überwunden haben. Das SigH-Regulon umfasst einen Teil der SOS-Antwort, des Hitzeschock-Regulons, hierunter auch die Regulationsnetzwerke HspR und ClgR, die für die Protein-Qualitäts-Kontrolle verantwortlich sind. Weiter umfasst das SigH-Regulon eine Reihe von Membranproteinen. (*) die SigH-Abhängigkeit dieser Gene konnte allerdings nicht durch die ROSE SigH Analyse nachgewiesen werden.

Weitere in dieser Arbeit identifizierte Gene, die bisher nicht als Mitglieder des SigH-Regulons beschrieben worden sind umfassen solche, die an der Mycothiol- (MSH) Synthese und an dem MSH-Recycling beteiligt sind. Neben Thioredoxin (Trx) ist Mycothiol das wichtigste Thiol mit kleiner Molekülmasse in Corynebakterien, Mycobakterien und Streptomycceten (Newton *et al.*, 1996). Die Biosynthese von MSH in *C. glutamicum* und die zwei essentiellen Gene *mshC* und *mshD*, die Bestandteil dieses Stoffwechselweges sind, wurden bereits beschrieben (Feng *et al.*, 2006). Durch das Vorgehen in dieser Arbeit konnte eine SigH-abhängige Hochregulation von *mshC*, kodierend für das zweite Protein in der MSH-Synthese und von *mca*, wie auch *mtr*,

die beide am Mycothiol-Recycling beteiligt, gezeigt werden (Abb. V-3). *Mca* ist hierbei das erste Gen bei dem MSH-Recycling und wurde bereits als SigH-abhängig beschrieben (Ehira *et al.*, 2009). Es kodiert eine Mycothiol-S-Konjugat-Amidase (*Mca*), die die Addukte der Reaktion von MSH mit Elektrophilen (MSR) spaltet um *N*-Acetyl-Cystein-S-Konjugat (AcCySR) und das Pseudodisaccharid 1-O-(2-Amino-2-Deoxy- α -D-Glucopyranosyl)-D-Myo-Inositol zu erhalten (GlcN-Ins) (Newton und Fahey, 2008; Newton *et al.*, 2008). AcCySR wird sekretiert während GlcN-Ins in der Zelle verbleibt. GlcN-Ins ist das Substrat für *MshC*, wodurch es wieder in den MSH-Stoffwechselweges integriert wird. Durch die nachfolgende enzymatische Aktivität von *MshD* wird MSH gebildet (Feng *et al.*, 2006). Eine weitere wichtige Funktion von *Mca* ist die Degradation von MSH um Cysteine (von Acetylcystein [AcCys]) für andere metabolische Stoffwechselwege zur Verfügung zu stellen (Newton *et al.*, 2008). Wie oben bereits erwähnt ist *mshD* in den hier durchgeführten Untersuchungen nicht als SigH-abhängig identifiziert worden.

In *M. tuberculosis* zeigt sich ein anderes Bild. Hier sind alle Gene des MSH-Stoffwechselweges konstitutiv transkribiert (Newton und Fahey, 2008). Ein ähnlicher Mechanismus ist auch für *S. coelicolor* beschrieben mit dem Unterschied, dass neben *mca*, auch *mshA* ein direktes Ziel von SigR ist und die Gene *mshB*, *mshC* und *mshD* zwar SigR-abhängig, aber anscheinend nicht direkt induziert werden (Park und Roe, 2008).

Neben den Genen die direkt mit dem Disulfid-Stress Stimulon in Verbindung gebracht werden können, wurden einer Reihe von Genen ohne klare Annotation (*cg1282*, *cg0454*, *cg2263*, *cg1095*, *cg2267*) als SigH-abhängig identifiziert, die in einer komparativen Transkriptionsanalyse in der mit 2 mM Diamidgestresste Zellen mit ungestressten Zellen verglichen wurden, eine relativ erhöhte Transkriptmenge festgestellt werden (Daten nicht gezeigt). Diese Gene spielen somit offenbar eine Rolle bei der oxidativen Stressantwort an der SigH beteiligt ist (Busche *et al.*, 2012). Die genaue Funktion muss allerdings geklärt werden.

Das SOS-Regulon vieler verschiedener Bakterien, inklusive *E. coli*, ist in mannigfaltige zelluläre Prozesse, wie zum Beispiel Nukleotid-Exzision und Rekombinations-Reparatur involviert (Walker, 1996). Durch die Deletion des Gens, welches den SOS-Antwort Regulator LexA in *C. glutamicum* kodiert, konnte die SOS-Stressantwort in *C. glutamicum* beschrieben werden (Jochmann *et al.*, 2009). Hierbei zeigte nur eines der *uvr* Gene, nämlich *uvrC*, eine differentielle Transkription in den *Microarray*-Analysen.

Durch die in dieser Arbeit verwendete Herangehensweise konnte eine SigH-abhängige Induktion von drei *uvr*-Genen (*uvrA*, *uvrC*, *uvrD3*) beobachtet werden. Die Induktion der Transkription von *uvrC* war hierbei am geringsten, am wahrscheinlichsten aufgrund der verbleibenden Repression durch LexA unter den gewählten Bedingungen. Wie erwähnt (Jochmann *et al.*, 2009) hängt der Grad der Induktion der SOS Genexpression letztendlich von vier Parametern ab: (i) der Affinität von LexA zur SOS-Box, (ii) der Lokalisation der SOS-Box relativ zum Promotor, (iii) der Promotorstärke und (iv) dem Vorhandensein zusätzlicher konstitutiver Promotoren (Friedberg *et al.*, 2005; Schnarr *et al.*, 1991; Walker, 1996). Es erscheint offensichtlich, dass die SOS-Antwort mit dem Hitze-Stress- und Thiol-oxidativem-Stress-Verteidigungssystemen in Verbindung steht und dass SigH eine wichtige Rolle spielt. Diese Aussage steht im Einklang mit einem Vorschlag von Barreiro *et al.* (Barreiro *et al.*, 2009). Starker Hitzestress (50 °C) könnte unter anderem zu DNA-Schäden führen, was mit einer Ablösung des LexA von der SOS-Box einherginge und so einen weiteren SigH-abhängigen *sigH* Promotor freigeben würde.

In dieser Arbeit wurde festgestellt, dass das SigH Regulationsnetzwerk noch weitere wichtige Funktionen der Stress-Antwort kontrolliert. Unter anderem die Kontrolle einer Reihe von Membranproteinen, die in der Antwort auf Hitzestress eine Rolle einnehmen könnten.

Im Zusammenhang mit der offensichtlichen Beteiligung des SigH-Regulon an Vorgängen der Membran, die eventuell im Zusammenhang mit Hitzestress auftreten können, ist es sehr interessant, dass einige Gene, die für Proteine kodieren, die Bestandteil der Membran sind von SigH-abhängig transkribierten *anti-sense* RNAs reguliert werden könnten. Es konnten Promotoren in *anti-sense* Orientierung zu den Membranprotein-Genen *mrcB*, *ftsI* und *cg2052* identifiziert werden. Die Beobachtung, dass einige *anti-sense* RNAs Bestandteil des SigH-Regulationsnetzwerkes sind, ist neu. Ihre möglichen regulatorischen Funktionen müssen allerdings durch Folgeuntersuchungen bestätigt werden.

Eine weitere besonders interessante neue Entdeckung ist die verstärkte Transkription von Komponenten des Proteasoms in der $\Delta rshA$ Deletionsmutante. Das actinobakterielle Proteasom umfasst Funktionen für die Pupylierung (ein Prozess der ähnlich zu der eukaryotischen Ubiquitinierung, die Proteine für den kontrollierten Abbau markiert). Die hier durchgeführte Studie integriert nun die Gene für die kürzlich identifizierten Pupylierungskomponenten Pup (prokaryotisches Ubiquitin-ähnliches Protein) und PafA (*proteasome accessory factor*, verantwortlich für die Pup-Konjugation) (Burns *et al.*, 2009) in das SigH-Regulon und weist damit auf eine wichtige Rolle von SigH in der Protein-Qualitätskontrolle hin (Abb. V-3).

Basierend auf den Daten, die in diesen vorangegangenen Arbeiten erhoben wurden, konnte ein erweitertes Modell des SigH-Regulons in *C. glutamicum* erstellt werden (Abb. V-3). Dieses umfasst nun die direkte Kontrolle der Stress-Antwort auf Disulfid- und Hitze-Stress durch RshA. Hierbei sind das Thioredoxin-System und das Mycothiol-Recycling-System involviert, welches für die Bewältigung der herrschenden Thiol-oxidierenden Bedingungen verantwortlich ist. In einem ungestressten Zustand ist SigH durch die reduzierte Form von RshA reversibel gebunden und somit inaktiv. Das Aufbrechen des SigH-RshA Komplexes geschieht in *C. glutamicum* unter oxidativem oder schwerem Hitze-Stress durch eine Konformationsänderung von RshA. Das freigegebene SigH kann nun mit der RNAP-Core Enzym ein funktionales RNAP-Holo Enzym formen und so die Stress-Antwort durch die Transkription SigH-abhängiger Gene initiieren. Darunter befinden sich die Gene, die an der Antwort zu Disulfid- und Hitzestress, darunter auch Membranproteine, beteiligt sind. Diese *feed-forward* Induktion der Expression des Anti-Sigma Faktors RshA ermöglicht der Zelle die schnellstmögliche Abschaltung der SigH-basierenden Stress-Antwort nach Beendigung des Stresszustandes. RshA, als der Stress-sensierende „Redox-Schaltknopf“, ist ein Target der biochemischen Stoffwechselwege, die durch Gene des SigH-Netzwerkes kodiert sind. Es handelt sich hierbei um die Reduktionskomponenten Thioredoxin (Trx) und Mycothiol (MSH). Die direkte Induktion der Expression von *trxB1C* generiert das Thiol Trx und die Genprodukte von *trxB*, *mtr*, *mca* und *mshC* reduzieren und/oder recyceln Trx bzw. MSH. Diese Thiole sind nun zusammen mit Reduktasen und anderen Reduktionskomponenten in der Lage, die Thiol-Redox-Balance wieder herzustellen und im speziellen die Oxidation von Cystein-Resten von RshA rückgängig zu machen. Nachdem die Redox-Homöostase erreicht ist wird RshA in diesem geschlossenen Kreislauf reduziert um seine Funktionalität wieder zu erlangen und SigH zu binden. Ein ganz ähnliches Modell wurde auch für die Thiol-Oxidations Stress-Antwort in *S. coelicolor* von Newton und Mitarbeitern im Jahre 2008 entwickelt (Newton und Fahey, 2008; Newton *et al.*, 2008). Das durch SigH kontrollierte transkriptionelle Regulationsnetzwerk ist hochgradig mit anderen Regulatoren verbunden, die die Genexpression in Antwort auf unterschiedliche physikalische oder chemische Trigger modulieren. Ein Beispiel hierfür ist das Hitze-Schock Regulationsnetzwerk, welches die Regulatoren HspR und ClgR umfasst. Ein weiteres Beispiel ist die Aufgabe von SigH für die Aufrechterhaltung der SigA-Transkription während Stressbedingung zu sorgen und die Vernetzung des SigH-Regulationnetzwerkes mit dem alternativen Sigmafaktor SigB, dem ECF-Sigmafaktor SigM und über den Anti-Sigmafaktor CseE mit dem Regulationnetzwerk von SigE sicherzustellen.

3. Das SigE-Regulationsnetzwerk in *Corynebacterium glutamicum*

Der ECF-Sigmafaktor SigE aus *C. glutamicum* wird zu der ECF-Gruppe ECF14 gezählt, die auf Actinobakterien beschränkt ist. Die ECF-Sigmafaktoren dieser Gruppe sind wie meisten ECF-Sigmafaktoren in einem Operon mit ihrem Anti-Sigmafaktor organisiert (Staron *et al.*, 2009).

Für *M. tuberculosis* SigE (ortholog zu *C. glutamicum* SigE), ist bekannt, dass das Sigmafaktor-Anti-Sigmafaktor-Operon (*sigE-rseA*) von drei verschiedenen Promotoren aus transkribiert wird (Sachdeva), die von MprA/MprB, SigA und SigH reguliert werden. Weiter ist bekannt, dass SigE posttranskriptionell von dem Anti-Sigmafaktor RseA reguliert wird (Donà *et al.*, 2008). Die Inaktivierung von RseA und damit die Reaktivierung von SigE findet durch Phosphorylierung des Anti-Sigmafaktors und anschließender Spaltung durch die Proteasen ClpC1P2 statt (Barik *et al.*, 2010). Auch in *C. glutamicum* ist die Funktion von CseE als Anti-Sigmafaktor von SigE und die die Transkription von *sigE-cseE* in Form eines Operons bekannt (Park *et al.*, 2008). Es waren aber weder die TSP des Gens *sigE* bekannt, noch die Regulation der *sigE*-Transkription oder die Inaktivierung des Anti-Sigmafaktors CseE verstanden.

SigE wird in *M. tuberculosis* unter anderem durch Hitze und SDS-Stress induziert (Sachdeva *et al.*, 2010). Die entsprechend regulierten Gene kodieren für Proteine unterschiedlicher Klassen, wie Transkriptionregulatoren (SigB, MprA und MprB), Enzyme der Lipidsynthese und klassische Hitzeschock-Proteine (Sachdeva *et al.*, 2010).

Ein ähnliches Regulon wird auch für *C. glutamicum* SigE angenommen, da eine $\Delta sigE$ -Mutante sensitiv gegenüber Hitze, SDS, EDTA, Lysozym und verschiedenen Antibiotika ist (Park *et al.*, 2008). Bis auf das Gen einer putativen Methyltransferase (*cg1266*) war aber noch kein Mitglied des SigE-Regulons bekannt.

Durch die Analysen SigH-abhängiger Promotoren durch RNA-Seq der Deletionsderivate RES167 $\Delta sigH$ und $\Delta rshA$ wurden eine Reihe von Promotoren identifiziert, deren Aktivität in RES167 $\Delta rshA$ erhöht und in RES167 $\Delta sigH$ verringert, aber detektierbar war. Hier wurde die Vermutung aufgestellt, dass ein weiterer Sigmafaktor für die verbleibende Transkription verantwortlich ist. Da aus *M. tuberculosis* bekannt ist, dass SigH und SigE ähnliche Promotoren erkennen (Song *et al.*, 2008) wurde durch ROSE untersucht, ob dies auch für *C. glutamicum* gilt und ob so die verbleibende Transkription in der RES167 $\Delta sigH$ aufgeklärt werden kann. Als komplementären *in vivo* Ansatz wurde analog zu der Untersuchung des

SigH-Regulationsnetzwerkes die Deletionsmutante RES167 $\Delta cseE$ erstellt und mittels *Microarray*-Analyse eine differentielle Transkriptom-Analyse im Vergleich zum Referenzstamm durchgeführt. Es wurde erwartet, dass das Regulationsnetzwerk von SigE durch die Deletion des Anti-Sigmafaktors in Abwesenheit von Stress induziert ist.

3.1. Das *sigE-cseE* Operon in *Corynebacterium glutamicum* weist eine ähnliche transkriptionale Organisation wie das *sigH-rshA* Operon und eine Autoregulation auf

Bislang war bekannt, dass *sigE* und *cseE* in Form eines Operons transkribiert werden. Durch diese Arbeit konnte gezeigt werden, dass die Transkription von zwei SigA-abhängigen Promotoren ausgeht, wobei der TSP des P1 mit dem Translationsstart von *sigE* übereinstimmt (Abb. V-4). Die Transkription des orthologen Operons in *M. tuberculosis* zeigt zum Teil einen ähnlichen Aufbau. Hier findet die Transkription ebenfalls in Abhängigkeit von zwei SigA-abhängigen Promotoren statt aber im Gegensatz zu *C. glutamicum* befindet sich ein zusätzlicher SigH-abhängiger Promotor stromaufwärts von *sigE* (Sachdeva *et al.*, 2010).

Für *M. tuberculosis* ist bekannt, dass der SigA-abhängige Promotor P1 für eine schwache Transkription des Operons im ungestressten Zustand sorgt. Der Promotor P2 wird von MprAB TCS reguliert und ist für eine verstärkte Transkription des Operons unter Oberflächen-Stress verantwortlich, während der Promotor P3 SigH-abhängig ist und für eine verstärkte Transkription des Operons unter oxidativem Stress und Hitze-Schock verantwortlich ist (Sachdeva *et al.*, 2010). Da die Transkription des Gens *sigE* in *C. glutamicum* während des exponentiellen Wachstums schwach ist (Park *et al.*, 2008), kann darauf geschlossen werden, dass zumindest einer der SigA-abhängigen Promotoren aktiv ist. Es kann weiter angenommen werden dass der zusätzliche SigA-abhängige Promotor ähnlich wie bei *M. tuberculosis* zusätzlich reguliert wird, da auch in *C. glutamicum* die Transkriptmenge des Gens *sigE* nach Zelloberflächen-Stress oder Hitze-Schock erhöht ist (Park *et al.*, 2008). Wie diese Regulation erfolgt ist allerdings nicht klar.

Analog zu der transkriptionellen Regulation des Operons *sigH-rshA* in *C. glutamicum* konnte für *sigE-cseE* festgestellt werden, dass das *cseE* Gen zusätzlich als monocistronisches Transkript transkribiert wird. Es konnte gezeigt werden, dass diese Transkription sowohl SigE-, als auch SigH-abhängig ist (Abb. V-4). Diese zusätzliche Regulation des Anti-Sigmafaktorgens durch den eigenen ECF-Sigmafaktor bzw. SigH ist, wie bei *rshA* bei den orthologen Genen in anderen Actinobakterien außerhalb der Corynebakterien nicht zu finden. Die Besonderheit der transkriptionellen Organisation des *sigE-cseE* Operons zu dem *sigH-*

rshA Operon in *C. glutamicum* ist, dass in den anderen bis zum Zeitpunkt dieser Arbeit sequenzierten Corynebakterien kein ECF-Sigmafaktor-abhängiger Promotor stromaufwärts von *cseE* vorhanden ist. Die zusätzliche Regulation des Anti-Sigmafaktors von SigE scheint nur in *C. glutamicum* vorhanden zu sein.

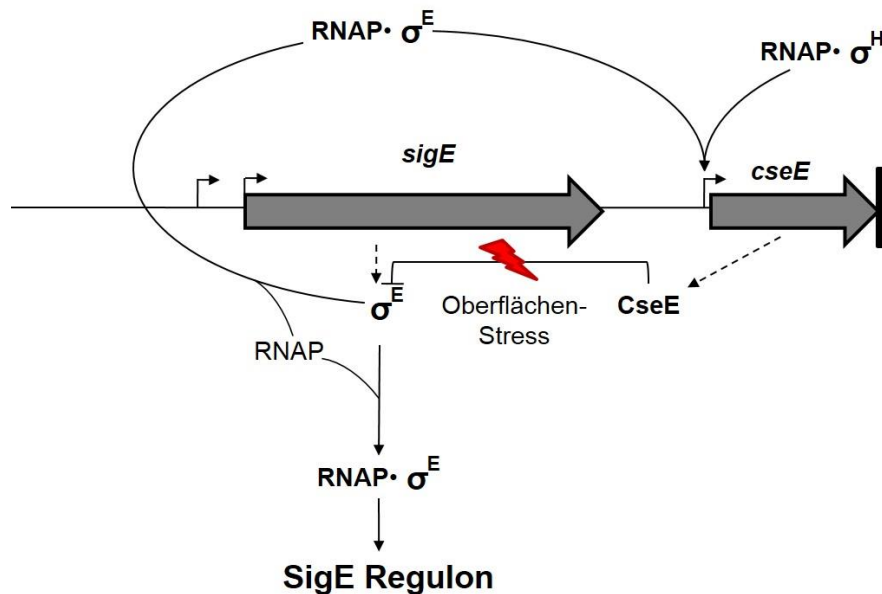


Abb. V-4 Transkriptionelle Organisation des *sigE-cseE* Operons. Die Transkription des gesamten Operons findet durch SigA-abhängige Promotoren statt. Im ungestressten Zustand wird der ECF-Sigmafaktor SigE durch den Anti-Sigmafaktor CseE gebunden und verhindert somit eine Anlagerung von SigE an das RNAP-Core Enzym. Es findet keine Transkription SigE-abhängiger Gene statt. Durch Oberflächenstress wird der Anti-Sigmafaktor durch einen unbekanntem Mechanismus inaktiviert und bindet nicht mehr an SigE. Der ECF-Sigmafaktor bindet an das RNAP-Core Enzym und das SigE-Regulon kann transkribiert werden. Ein Bestandteil des SigE-Regulons ist das Anti-Sigmafaktorgen, welches zusätzlich Abhängigkeit von SigH transkribiert wird. Hierdurch wird bei Aktivierung des SigE-Regulons und des SigH-Regulons bei Stress ein Überschuss des Anti-Sigmafaktors CseE gegenüber SigE gewährleistet, der eine sofortige Inaktivierung des Sigmafaktors SigE ermöglicht, sobald CseE aktiv ist. Promotoren sind mit geneigten Pfeilen und der Terminator als vertikaler schwarzer Balken gekennzeichnet.

Die transkriptionelle Organisation des *sigE-cseE* Operons hat sich vermutlich ähnlich wie bei SigH entwickelt, um SigE auch ohne Stressapplikation konstitutiv zu exprimieren und die Aktivität des Sigmafaktors durch den ebenfalls im Operon konstitutiv exprimierten Anti-Sigmafaktor in Abwesenheit von Stress zu inaktivieren. Diese rein posttranslationale Regulation ist die schnellste Möglichkeit, um auf einen plötzlich auftretenden Stress durch Inaktivierung des Anti-Sigmafaktors reagieren zu können. Daneben scheint das Operon unter Hitze-Schock und Oberflächen-Stress zusätzlich transkriptionell reguliert zu sein, wie es auch für das *sigH-rshA* Operon durch die SOS-Antwort angenommen wird. Die zusätzliche

Transkription des *cseE*-Gens durch einen SigE-abhängigen Promotor kann wie bei SigH-*rshA* sicherstellen, dass unter spezifischen Stressbedingungen, unter denen SigE aktiv ist, immer ein Überschuss an CseE über SigE vorliegt. Dadurch kann dann die schnellstmögliche Abschaltung der SigE-abhängigen Transkription durch die Reaktivierung des im Überschuss vorliegenden Anti-Sigmafaktors CseE nach Beendigung des Stresszustandes gewährleistet werden. Die zusätzliche Transkription von *cseE* durch SigH kann damit begründet werden, dass sich die Regulationsnetzwerke von SigH und SigE zum Teil überschneiden und unter bestimmten Voraussetzungen, in denen nur SigH aktiv ist, sichergestellt werden muss, dass SigE durch einen Überschuss von CseE komplett geblockt wird und es zu keiner Konkurrenz von SigH und SigE am RNAP-Core Enzym kommt.

3.2. Das SigE-Regulationsnetzwerk in *Corynebacterium glutamicum* umfasst Gene für Membranproteine, Chaperone, Proteasen und den Regulator des Stickstoffmetabolismus AmtR

Basierend auf den durch ROSE-SigE gewonnenen Ergebnissen dieser Arbeit können nun erstmalig Mitglieder des SigE-Regulons in *C. glutamicum* benannt werden. Die *in vitro* Ergebnisse konnten durch die *in vivo* Microarray-Analyse der Deletionsmutante RES167 $\Delta cseE$ bestätigt werden. Da in dieser Analyse im Vergleich zum Referenzstamm keine exklusiv SigH-abhängigen Gene differentiell transkribiert waren, konnte ausgeschlossen werden, dass CseE gleichzeitig als Anti-Sigmafaktor von SigH fungiert.

Das SigE-Regulon in *C. glutamicum* umfasst mehrere Komponenten der Sigmafaktor-ausstattung, darunter die Sigmafaktorgene *sigA* und *sigB* und das Anti-Sigmafaktorgen *cseE*, dessen SigE-abhängige Transkription bereits diskutiert wurde.

Da die Transkriptmenge von *sigA* durch die Deletion des Anti-Sigmafaktors CseE im Vergleich zum Wildtyp nicht signifikant erhöht war, kann, wie auch schon bei der Kontrolle durch SigH diskutiert, angenommen werden, dass es sich hierbei um ein *Fine-Tuning* der globalen Transkription handelt, bei der das SigA-Level auch unter Stressbedingungen aufrecht erhalten wird. Die SigE-abhängige Transkription von *sigB* könnte als Vorbereitung auf die nicht optimalen Wachstumsbedingungen, die während der stressbedingten Stationärphase auftreten, interpretiert werden.

Neben den Sigmafaktoren SigA und SigB und dem Anti-Sigmafaktor CseE beinhaltet das SigE-Regulationsnetzwerk Proteine, die an der Hitze-Stressantwort beteiligt sind, sowie weitere Proteasen (DnaK, GrpE, ClpB, FtsH und potentiell *ClgR*; Abb. V-5). Auch das SigE-Regulationnetzwerk in *M. tuberculosis* beinhaltet klassische Hitzeschockproteine (Sachdeva *et al.*, 2010). Die Proteasen ClpC1P2, die zu diesen klassischen Hitzeschockproteinen zählen, sind in *M. tuberculosis* für die proteolytische Inaktivierung des Anti-Sigmafaktors RseA verantwortlich. Dazu muss RseA allerdings phosphoryliert vorliegen (Barik *et al.*, 2010). In *M. tuberculosis* ist der transkriptionelle Regulator Rv2745c für die Aktivierung des *clp*-Regulons in Antwort auf Oberflächenstress verantwortlich und steht unter der Kontrolle von SigE (Barik *et al.*, 2010). In *C. glutamicum* steht das *clp*-Regulon unter der positiven Kontrolle von ClgR (Engels *et al.*, 2004), der potentiell SigE abhängig exprimiert wird. Ein ähnlicher positiver *Feedback Loop*, der zur Aktivierung des SigE-Regulons führt, wäre somit auch für *C. glutamicum* denkbar. Für *C. glutamicum* ist auch denkbar, dass die weiteren, nicht durch ClgR regulierten SigE-abhängigen Proteasen zur proteolytischen Inaktivierung von CseE führen

(Abb. V-5). Für *E. coli* wurde gezeigt, dass die proteolytische Inaktivierung des phosphorylierten Anti-Sigmafaktors RseA durch eine Vielzahl der vorhandenen cytosolischen Proteasen durchgeführt werden kann (Ho und Ellermeier, 2012).

In Übereinstimmung mit der Beobachtung, dass eine $\Delta sigE$ -Mutante sensitiv gegenüber Membranstress ist (Park *et al.*, 2008) steht die Beobachtung, dass Gene für Membranproteine Bestandteil des SigE-Regulons sind. Park und Mitarbeiter begründeten die potentiell instabilere Membran der Deletionsmutante mit dem Fehlen einer putativen Methyltransferase, deren Gen (*cg1266*) SigE-abhängig transkribiert zu werden scheint (Park *et al.*, 2008). *Cg1266* stellte das einzige bis jetzt als SigE-abhängig beschriebenes Gen in *C. glutamicum* dar. Die SigE-abhängige Expression des Gens wurde allerdings nur indirekt durch einen komparativen Proteomikansatz postuliert. Für das Gen *cg1266* konnte in dieser Arbeit keine direkte transkriptionelle Abhängigkeit von SigE-gesättigtem RNAP-Enzym festgestellt werden. Auch in der *Microarray*-Analyse in der die Transkriptome des *C. glutamicum* Deletionsderivat RES167 $\Delta cseE$ mit dem Referenzstamm verglichen wurden konnte keine differentielle Transkription dieses Gens festgestellt werden. Offensichtlich handelt es sich um einen sekundären Effekt der Deletion von *sigE*, die in der Studie von Park und Mitarbeitern zu dem Verlust der entsprechenden Methyltransferase führten (Park *et al.*, 2008). Die Annahme von Park und Mitarbeitern bleibt bestehen, dass der Verlust der Methyltransferase zu einer instabileren Membran führt. Der so entstehende Oberflächenstress wird offenbar mit der direkt SigE-abhängigen Transkription von Membranprotein-Genen beantwortet.

Neben den bereits genannten Mitgliedern des SigE-Regulons, die gleichzeitig auch Bestandteile der SigH-Regulationsnetzwerkes sind, konnten drei Gene identifiziert werden, die *in vitro* exklusiv SigE-abhängig transkribiert werden (Abb. V-5). Bei den potentiell exklusiv SigE-abhängigen Genen handelt es sich um *cg1121* (Permease vom MFS-Typ), *cg3309* (putatives sekretiertes Protein) und um *amtR* (Regulator des Stickstoffmetabolismus). Die Gene *cg1121* und *cg3309* können mit der Hauptaufgabe des SigE-Regulons, der Antwort auf Membranstress in Verbindung gebracht werden. Das Gen *amtR* stellt eine Ausnahme dar, da es nicht mit Membranstress oder Hitzeschock in Verbindung gebracht werden kann. Es ist sehr interessant, dass die geringfügig erhöhte Transkriptmenge von *amtR* durch die Deletion von *cseE* und damit Reaktivierung von SigE zu einer drastischen Repression der Gene des AmtR-Regulons führt. Es wurde bereits vermutet, dass SigE neben der Bekämpfung von Zelloberflächenstress, zusätzlich noch in die Regulation weiterer zellulärer Prozesse involviert ist, da SigE unter Stickstoffmangel (Silberbach *et al.*, 2005) oder in einem *C. glutamicum rel*-Deletionsstamm, der keine Stringente Kontrolle mehr ausüben kann (Brockmann-Gretza und

Kalinowski, 2006) verstärkt transkribiert wird. Die SigE-abhängige Transkription des *amtR*-Gens ist der erste Hinweis, dass SigE bei der Antwort auf Stickstoffmangel beteiligt sein könnte.

4. **Überschneidungen der SigH- und SigE-Regulationnetzwerke in *Corynebacterium glutamicum***

Wie bereits in anderen Arbeiten vermutet (Barreiro *et al.*, 2009) besteht eine Verbindung zwischen den Regulationsnetzwerken von SigH und SigE. Auf Grund der Vielzahl an SigH-abhängig transkribierten Genen inklusive Sigmafaktorgenen wurde postuliert, dass SigH eine zentrale Position in der Regulation von *C. glutamicum* Genen einnimmt, die der Zelle die Antwort auf Hitzestress und oxidativem Stress ermöglicht (Brinkrolf *et al.*, 2010).

In dieser Arbeit konnte nun gezeigt werden, dass diese Überlappung der Regulons existiert. Es konnten Promotoren identifiziert werden, die der dualen Kontrolle der beiden ECF-Sigmafaktoren SigH und SigE unterliegen. Hierbei handelt es sich vor allem um Gene die Mitglieder des Hitzeschockregulons sind und um Gene, die für Proteasen oder Membranproteine kodieren (Abb. V-5). Diese Tatsache stimmt mit dem Wissen überein, dass SigH in *C. glutamicum* und in anderen Actinomyceten die Antwort auf Hitze- und oxidativen Stress reguliert (Engels *et al.*, 2004; Kim *et al.*, 2005) und eine $\Delta sigE$ -Mutante sensitiv gegenüber Hitze, SDS, EDTA, Lysozym und verschiedenen Antibiotika ist (Park *et al.*, 2008). Die Alleinstellungsmerkmale des SigH-Regulons ist die Antwort auf oxidativen Stress unter anderem durch Kontrolle der Mycothiol- bzw. Thioredoxin-Redox-Puffersysteme und der Beteiligung an der SOS-Antwort (Abb. V-5). Das SigE-Regulon umfasst drei exklusiv SigE abhängig transkribierte Gene, wobei zwei der Genprodukte der Bekämpfung von Oberflächenstress dienen könnten und AmtR, den globalen Transkriptionsregulator des Stickstoff-Stoffwechsels (Abb. V-5).

In *M. tuberculosis* liegt eine ähnliche Situation vor. SigH reguliert hier auch das Thioredoxin-Puffer-System und verschiedene Hitzeschockproteine in Antwort auf Hitzeschock und oxidativen Stress (Raman *et al.*, 2001). Phänotypische Analysen konnten für SigE zeigen, dass dieser ECF-Sigmafaktor in *M. tuberculosis* die gleiche Rolle wie *C. glutamicum* SigE in der Antwort auf Zelloberflächenstress spielt (Manganelli *et al.*, 2001). In der Arbeit von Song *et al.* konnte ebenfalls ein gleichzeitige Kontrolle von Hitzeschockgenen und Genen, die Membranproteine kodieren, gefunden werden. Auch hier wurden Gene identifiziert, die nur von SigE abhängig sind und mit der Zellmembran in Zusammenhang stehen (Song *et al.*, 2008).

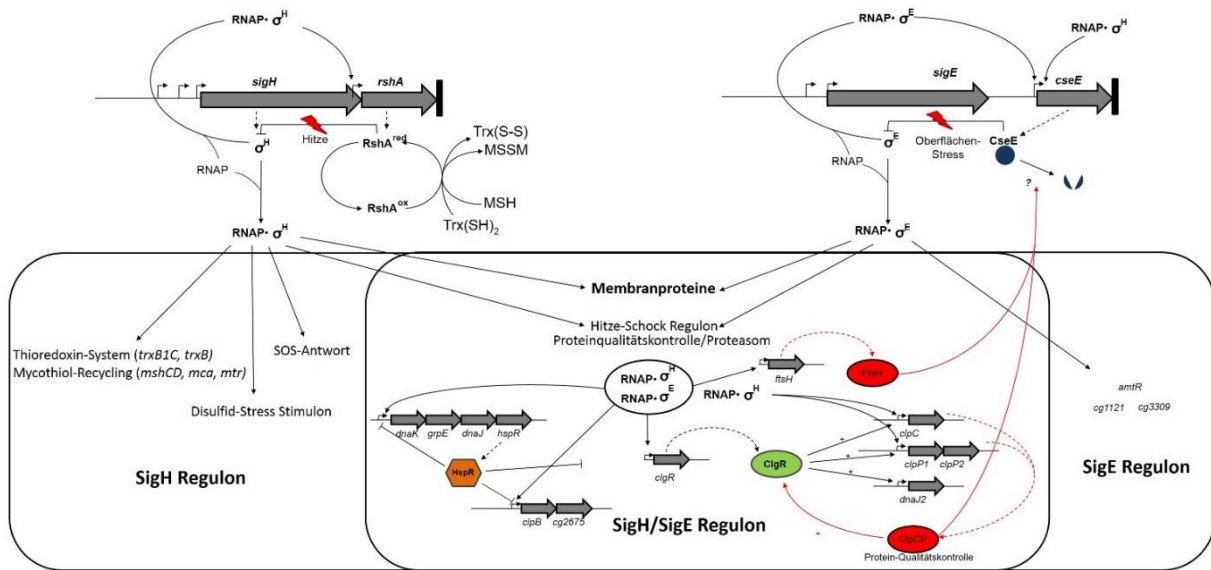


Abb. V-5 Übersicht über die SigH und SigE Regulationsnetzwerke und deren Überschneidung in *C. glutamicum*. Es ist die transkriptionelle Organisation und Regulation der Operons *sigH-rshA* und *sigE-cseE* inklusive ihrer zum Teil überlappenden Regulationsnetzwerke dargestellt. (?) kennzeichnen Spekulationen.

Neben der Verflechtung der beiden Regulationnetzwerke durch die duale Kontrolle von Promotoren hat SigH auch direkten Einfluß auf die Transkription des Anti-Sigmafaktorgens *cseE* und kann eventuell so eine zusätzliche Kontrolle der dual regulierten Promotoren durch SigE unter bestimmten Bedingungen verhindern. In *M. tuberculosis* findet sich eine ähnliche Regulation, nur umfasst diese die SigH-abhängige Transkription des Operons *sigE-rseA* (Sachdeva *et al.*, 2010).

Die Regulationsnetzwerke von SigH und SigE sind also auf verschiedenen Ebenen hierarchisch verflochten. Während das SigH-Netzwerk vermutlich der Bekämpfung der unter vielen Stressbedingungen auftretenden Auslenkung der Redox-Homöostase dient, integriert das partiell untergeordnete SigE-Netzwerk vermutlich die Reaktion auf Oberflächenstress und ein Herunterfahren des Stickstoffmetabolismus unter den spezifischen Induktionsbedingungen für SigE. Da Oberflächenstress sicher auch ohne massive Störung der Redox-Homöostase vorkommen kann, erlaubt die Verknüpfung der Netzwerke so eine flexiblere Antwort.

5. Das 3'-Ende des -35 Promotormotivs bewirkt die Spezifität der Erkennung durch die *Corynebacterium glutamicum* ECF-Sigmafaktoren SigH und SigE

Die Promotoren SigH- und SigE-abhängig transkribierter Gene besitzen die -35 und -10 Promotorkernmotive kGGAA und GTT die auch bei *M. tuberculosis* zu finden sind (Song *et al.*, 2008).

Promotoren, die nur von SigH erkannt werden (Abb. V-6-I) unterscheiden sich von denen die zusätzlich durch SigE erkannt werden (Abb. V-6-III) in der sechsten Position des -35 Promotormotivs. Ein T an dieser Position führt vermutlich zu einer exklusiven Erkennung durch SigH, während ein C an dieser Position zu einer zusätzlichen Erkennung durch SigE führt. Dies entspricht den Beobachtungen für die entsprechenden *M. tuberculosis* Promotoren. Hier wurde für SigH-abhängige Promotoren das -35 Konsensusmotiv gGGAAYA (Raman *et al.*, 2001a) und für SigE-abhängige Promotoren gGGAACYa identifiziert (Song *et al.*, 2008).

Einige wenige SigH-abhängige Promotoren haben anstelle eines Pyrimidins an Position sechs des -35 Motivs das Purin G (Abb. V-6-II). Folgt diesem Motiv ein C oder A wird das Motiv nur durch SigH erkannt (Abb. V-6-IIa), folgt ein T-Abschnitt wird der Promotor offenbar zusätzlich durch SigE erkannt, da diese Promotoren trotz der Deletion von *sigH in vivo* weiterhin aktiv sind (Abb. V-6-IIb). Ein Beweis steht allerdings noch aus. Weiter ist unklar, ob für die evnetuell zusätzliche Erkennung durch SigE nur das T an der siebten Position verantwortlich ist oder tatsächlich eine Folge an Ts. Für *M. tuberculosis* ist das -35 Promotormotiv kGGAAG nicht bekannt.

Neben dieser grundsätzlichen Unterscheidung durch die sechste Position des -35 Promotormotivs scheint die siebte Position zusätzlich einen Einfluß auf die Erkennung durch einen der beiden ECF-Sigmafaktoren zu haben, sofern es sich um einen SigH und zusätzlich SigE-abhängigen Promotor mit dem Motiv gGGAAC handelt. Ist das Nukleotid an Position sieben ein A wird das -35 Motiv klar durch SigH und zusätzlich durch SigE erkannt (Abb. V-6-III). Ist die siebte Position allerdings durch ein T besetzt (Abb. V-6-III und IV) kann es sich in *C. glutamicum* um einen Promotor handeln, der durch SigH und SigE erkannt wird oder der nur von SigE (*in vitro*) erkannt wird. Die Unterscheidung zwischen diesen beiden Promortypen konnte im Rahmen dieser Arbeit nicht aufgeklärt werden. Für SigH- und SigE-abhängige Promotoren in *M. tuberculosis* konnte festgestellt werden, dass das Nukleotid an der siebten Position Auswirkungen auf die optimale Erkennung durch SigH oder SigE hat. In *M. tuberculosis* werden Promotoren, die ein Pyrimidin an dieser Position haben, optimal von SigE und weniger gut von SigH erkannt (Song *et al.*, 2008). Sowohl in *C. glutamicum*, als auch

in *M. tuberculosis* scheint ein T an Position sieben der -35 Promotormotive SigH- und SigE-abhängiger Promotoren besser durch SigE erkannt zu werden, was auch die fehlende *in vitro* Transkription der exklusiv SigE-abhängigen Promotoren durch SigH in *C. glutamicum* teilweise erklären würde. Allerdings müssen wohl noch weitere Elemente für den Ausschluß der Erkennung durch SigH verantwortlich sein. Über die Auswirkungen des Pyrimidins C an dieser Position kann für *C. glutamicum* keine Aussage getroffen werden, da eine solche Zusammensetzung des -35 Promotormotivs nicht bekannt ist.

Ist die siebte Position in den *M. tuberculosis* Promotoren mit dem Motiv gGGAAC durch ein A anstelle von T ersetzt, findet weder *in vitro* noch *in vivo* eine Erkennung durch SigE statt, was auch die -35 Konsensussequenz gGGAAYA SigH-abhängiger Promotoren erklärt. Dies ist ein deutlicher Unterschied zu *C. glutamicum* da hier weiterhin, sowohl *in vivo*, als auch *in vitro* eine Erkennung durch SigE erfolgt (Abb. V-6-III).

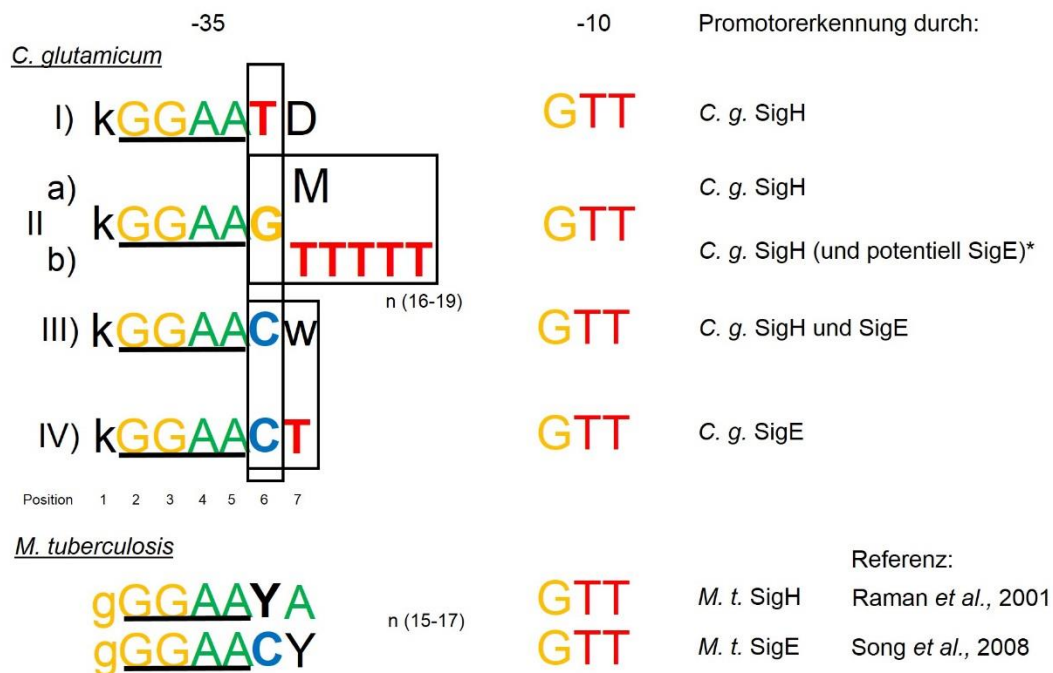


Abb. V-6 Konsensusmotive SigH und SigE-abhängiger Promotoren in *C. glutamicum* und *M. tuberculosis*. Es sind die -35 und -10 Promotormotive SigH- und SigE-abhängiger Promotoren in *C. glutamicum* unter Nutzung des IUPAC-Kodes dargestellt. Als Vergleich sind die Promotor-Konsensussequenzen SigH und SigE-abhängiger Promotoren aus *M. tuberculosis* angegeben. Das -35 Kernmotiv ist unterstrichen, entscheidende Positionen für die Erkennung der -35 Promotormotive sind mit Boxen umgeben; IUPAC-Kode: D= A oder G oder T; M= C oder A; K= G oder T; W= A oder T; Y=C oder T.

Zusammengefasst lässt sich für exklusiv SigH-abhängige Promotoren in *C. glutamicum* folgende Konsensussequenz des -35 Promotormotivs rekonstruieren: kGGAATD bzw. in seltenen Fällen kGGAAGM (Abb. V-6-I und Abb. V-6-IIa). Für Promotoren, die sowohl von SigH als auch SigE erkannt werden lautet die Konsensussequenz des -35 Promotormotivs kGGAACw bzw. kGGAAGT(TTTT), wobei es sich hier nur um zwei Fälle handelt und der experimentelle Nachweis der Kontrolle durch SigE fehlt. Es ist unklar wie sich potentiell exklusiv SigE-abhängige Promotoren mit dem -35 Promotormotiv kGGAAGT von den Promotoren mit identischem -35 Promotorelement unterscheiden, die zusätzlich durch SigH erkannt werden. Es gilt natürlich für diese Interpretationen die Einschränkung, dass der Überspiralisierungsgrad der DANN für die Unterscheidung mitverantwortlich sein könnte. Daher müssen diese Ergebnisse noch durch Mutationen an einzelnen kritischen Positionen der Promotoren sowohl *in vitro* als auch *in vivo* validiert werden.

6. Das Sigmafaktornetzwerk in *Corynebacterium glutamicum*

Auf Basis der bereits veröffentlichten oder durch diese Arbeit erzielten Erkenntnisse in Bezug auf die Promotoren der Sigmafaktorgene in *C. glutamicum* kann nun ein Interaktionsnetzwerk der Sigmafaktoren in *C. glutamicum* erstellt werden. Der *housekeeping* Sigmafaktor SigA ist autoreguliert und, bis auf *sigM*, für die Transkription aller ECF-Sigmafaktorgene in *C. glutamicum* verantwortlich (Abb. V-7). Dies unterscheidet *C. glutamicum* von anderen Actinobakterien, wie *M. tuberculosis* und *S. coelicolor*. In *M. tuberculosis* besitzen die orthologen Gene der ECF-Sigmafaktoren SigE und SigH einen SigH-abhängigen Promotor (Sachdeva *et al.*, 2010). Wie *sigH* in *M. tuberculosis* ist auch das orthologe Gen *sigR* in *S. coelicolor* autoreguliert. Zusätzlich besitzen *sigR* und *M. tuberculosis sigE* SigA-abhängige Promotoren.

Stromaufwärts von *sigH* und *sigE* konnten in *C. glutamicum* multiple SigA-abhängige Promotoren identifiziert werden. Zumindest einer dieser Promotoren ist für eine konstitutive Transkription der Gene verantwortlich, da eine basale Transkription unter Standardbedingungen festgestellt werden konnte (Patek *et al.*, 2012). Wie für andere Promotoren in *C. glutamicum* bereits gezeigt, eröffnen multiple Promotoren die Möglichkeit der zusätzlichen Regulation der Transkription des Gens (Patek *et al.*, 2012). Im Fall von *sigH* wird der Promotor P4 durch die Bindung von LexA an die dortige SOS-Box (Jochmann *et al.*, 2009) blockiert. Diese Art der Regulation konnte für die Promotoren des *sigE* Gens nicht gezeigt werden. Es ist unklar, wie die Transkription durch Zelloberflächen-Stress oder Hitze-Schock verstärkt wird. Die komplexe Regulation der *sigE*-Transkription in Bezug auf die SigA-abhängigen Promotoren in *M. tuberculosis* (ortholog zu *sigE* aus *C. glutamicum*) durch das Zwei-Komponenten-System MprAB an Promotor P2 und der zusätzlichen konstitutiven Expression (P1) (Sachdeva *et al.*, 2010) lässt aber eine ähnliche Kontrolle durch zusätzliche Transkriptionsfaktoren in *C. glutamicum* vermuten, die noch untersucht werden muss.

Weiter besteht die Möglichkeit, dass die multiplen SigA-abhängigen Promotoren von *sigH* und *sigE* unter bestimmten Bedingungen auch von SigB erkannt werden, da SigA und SigB ähnliche Promotoren erkennen (Larisch *et al.*, 2007). Es wird vermutet, dass SigB eine Art *Backup* zu SigA ist, der die Initiierung der Transkription an (bestimmten) SigA-ähnlichen Promotoren während der Stationärphase oder unter anderen nicht optimalen Bedingungen gewährleistet (Patek und Nesvera, 2012). SigB aus *C. glutamicum* besitzt somit eine ähnliche Funktion wie Sig³⁸ (SigS), dem Hauptregulator der allgemeinen Stressantwort in *E. coli* (Becker und Hengge-Aronis, 2001).

Die Erkennung eines Promoters durch den Sigmafaktor SigS aus *E. coli* ist zum Teil von der Konformation der DNA abhängig (Kusano *et al.*, 1996). Dies könnte auch für die effektive Erkennung von SigA-ähnlichen Promotoren durch SigB nötig sein. Es ist somit nicht ausgeschlossen, dass einige der multiplen SigA-abhängigen Promotoren unter bestimmten Voraussetzungen (s.o.) effektiver von SigB erkannt werden und somit die Transkription der Gene *sigH* und *sigE* unter diesen nicht optimalen Wachstumsbedingungen auf SigB-gesättigter RNAP beruht.

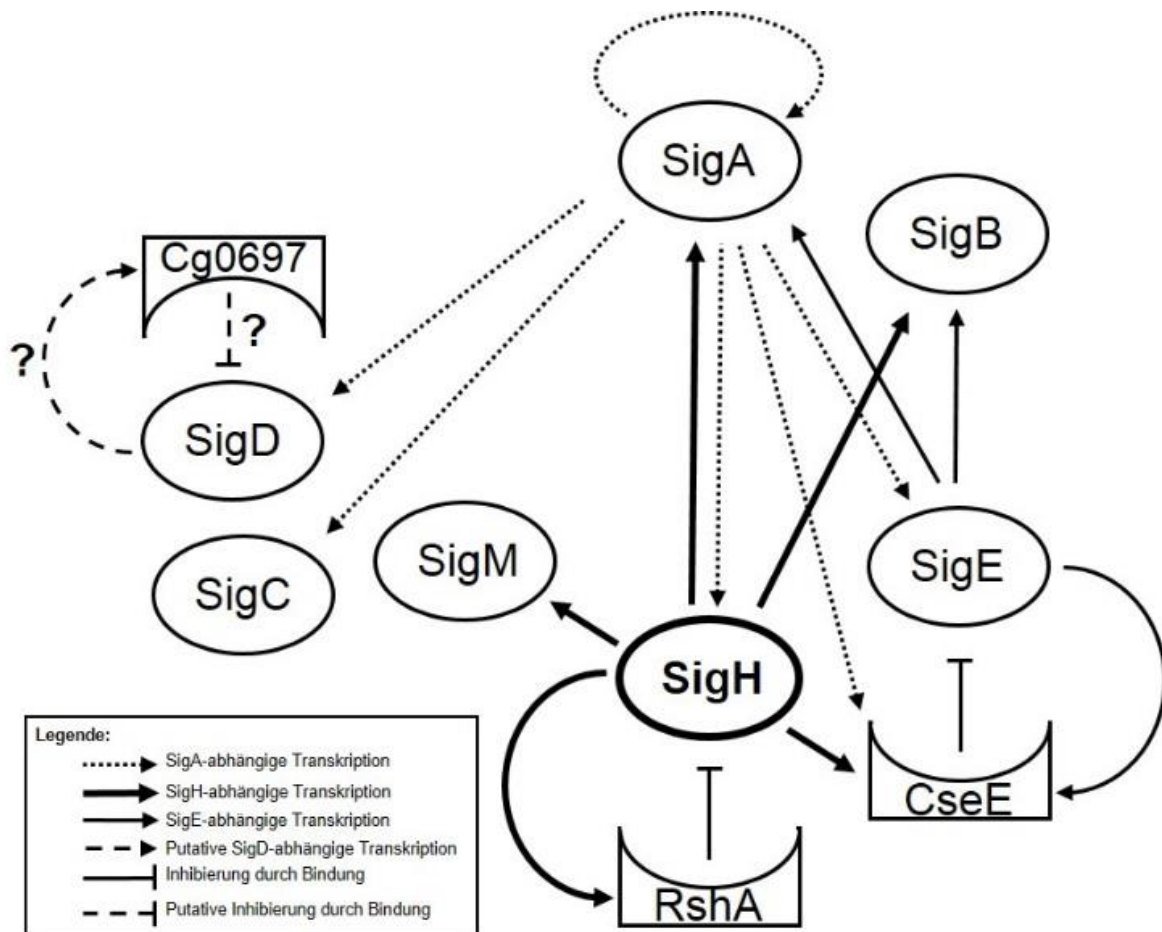


Abb. V-7 Interaktionsnetzwerk der Sigmafaktoren in *C. glutamicum*. Die Sigmafaktoren und Anti-Sigmafaktoren sind durch Ovale und komplementäre Boxen dargestellt und benannt; Pfeile zeigen positive Interaktionen und Linien mit horizontalem Ende eine Bindung und damit Inaktivierung des Sigmafaktors. (?) kennzeichnet spekulative Interaktionen.

Die Gene *sigC* und *sigD* werden von nur einem SigA-abhängigen Promotor aus transkribiert (Pfeifer-Sancar *et al.*, 2013, in Vorbereitung; Abb. V-7). Für *sigD* ist bekannt, dass die relative Transkriptmenge des Gens in einer *rel*-Deletionsmutante im Vergleich zum WT verringert ist (Brockmann-Gretza und Kalinowski, 2006). Somit scheint hier eine Regulation des SigA-abhängigen Promotors in Zusammenhang mit der Stringenten Kontrolle vorzuliegen. Für diese Vermutung spricht auch, dass das orthologe Gen *sigD* in *M. tuberculosis* unter

Nährstoffmangel, einer Bedingung, die zumindest teilweise Rel-abhängig ist, verstärkt transkribiert wird (Betts *et al.*, 2002). Daneben wird *sigD* in *M. tuberculosis* aber während des exponentiellen Wachstums und der Stationärphase konstitutiv exprimiert (Sachdeva *et al.*, 2010). Diese zusätzliche basale Expression scheint für *C. glutamicum sigD* nicht vorzuliegen, da hier nur ein Promotor identifiziert wurde.

Für das Gen *sigC* ist nicht bekannt unter welchen Bedingungen es verstärkt transkribiert wird. Daher bleibt es ungeklärt, ob der SigA-abhängige Promotor unter Standardbedingungen geblockt wird oder ob *sigC* konstitutiv transkribiert wird (Abb. V-7). In *M. tuberculosis* ist die konstitutive Transkription des Gens *sigC* bekannt, da das Transkript von *sigC* in *M. tuberculosis* unter Standardbedingungen von zwei SigA-abhängige. Promotoren aus transkribiert wird (Chang *et al.*, 2012) und es das häufigste Transkript der ECF-Sigmafaktoren in *M. tuberculosis* ist (Sachdeva *et al.*, 2010).

SigH ist einer der Hauptregulatoren in *C. glutamicum*, da dieser ECF-Sigmafaktor ein Regulon kontrolliert, dass neben einer Vielzahl von Genen zur Stressabwehr von oxidativem Stress und Hitzestress mehrere Transkriptionsfaktoren reguliert (Schröder *et al.*, 2010). Es konnte gezeigt werden, dass SigH auch im Interaktionsnetzwerk der Sigmafaktoren in *C. glutamicum* eine zentrale Rolle einnimmt (Abb. V-7). Zum einen ist SigH alleinig für die Transkription des ECF-Sigmafaktorgens *sigM* (Nakunst *et al.*, 2007) verantwortlich und zum anderen für die zusätzliche Transkription der Sigmafaktor Gene *sigA* (diese Arbeit) und *sigB* (holgasova). Neben der SigH-abhängigen Transkription von Sigmafaktoren konnte in dieser Arbeit erstmalig gezeigt werden, dass die Anti-Sigmafaktorgene *rshA* und *cseE* SigH-abhängig transkribiert werden, was die zentrale Rolle des ECF-Sigmafaktors SigH in der Regulation der ECF-Sigmafaktoraktivität zeigt (Abb. V-7).

Zusätzlich zu SigH erkennt auch der ECF-Sigmafaktor SigE die entsprechenden Promotoren stromaufwärts der Gene *sigA* und *sigB* und des ihm zugehörigen Anti-Sigmafaktorgens *cseE*. Die Regulation der Aktivität von ECF-Sigmafaktoren findet im Allgemeinen auf Proteinebene durch Anti-Sigmafaktoren statt (Staron *et al.*, 2009). Dies ist insbesondere wichtig, da die ECF-Sigmafaktorgene in den meisten Fällen, neben zusätzlich regulierter Transkription unter Stressbedingungen, während der exponentiellen Wachstumsphase konstitutiv transkribiert werden. Sowohl SigE als auch SigH werden durch die Anti-Sigmafaktoren CseE (Park *et al.*, 2008) bzw. RshA (Busche *et al.*, 2012) im ungestressten Zustand durch Bindung inaktiv gehalten werden (Abb. V-7). Auf Proteinebene wird reguliert, welcher der beiden ECF-Sigmafaktoren aktiv ist und für die Transkription der Ziel-Gene verantwortlich ist.

Für SigD ist eine ähnliche Regulation denkbar, da ein Gen für einen entsprechenden Anti-Sigmafaktor direkt stromabwärts des *sigD* Gens identifiziert wurde (Staron *et al.*, 2009). Die Funktion als Anti-Sigmafaktors von SigD ist allerdings nicht bewiesen. Die ECF-Sigmafaktoren SigM und SigC besitzen in *C. glutamicum* vermutlich keinen Anti-Sigmafaktor. Da *sigM* im Gegensatz zu den anderen Sigmafaktoren aber keinen SigA-abhängigen Promotor besitzt, scheint die Regulation auf Transkriptionsebene durch die SigH-abhängige Transkription ausreichend zu sein. Im Fall von *sigC*, das eventuell wie in *M. tuberculosis* konstitutiv transkribiert wird, ist die notwendige Regulation unter Standardbedingungen nicht bekannt. In *M. tuberculosis* werden zwei Möglichkeiten der Regulation diskutiert. Zum einen wird über eine post-transkriptionelle Regulation nachgedacht, bei der das *sigC*-Transkript nur mit einer geringen Effizienz translatiert wird (Manganelli *et al.*, 1999). Zum anderen wird die Möglichkeit einer post-translationalen Regulation durch Autoinhibition des ECF-Sigmafaktors SigC diskutiert. Für diese Annahme spricht, dass eine Interaktion der Sigmafaktordomänen 2 und 4 *in vitro* gezeigt werden konnte (Thakur und Gopal, 2005). Es wird nun angenommen, dass die Interaktion der Domänen zu einer Autoinhibition von SigC führen und ein zusätzlicher Faktor notwendig ist, um SigC zu reaktivieren (Thakur *et al.*, 2007).

Damit konnte in dieser Arbeit unter Zuhilfenahme eines optimierten Methodenspektrums umfangreiches Wissen zu einzelnen ECF-Sigmafaktor-Netzwerken und ihrer Verknüpfung gewonnen werden. Es zeigen sich bei der Verknüpfung viele Parallelen zu orthologen Sigmafaktoren in den verwandten Actinobakterien *M. tuberculosis* und *S. coelicolor*.

7. *Corynebacterium glutamicum* als Modell-Organismus

C. glutamicum ist aufgrund seines einfachen Sets von nur fünf ECF-Sigmafaktoren, die alle unterschiedlichen Familien bzw. Gruppen zugeordnet werden können, besonders für die Untersuchung der ECF-Regulationsnetzwerke geeignet und kann auch in dieser Hinsicht als Modellorganismus der Aktinomyzeten betrachtet werden. Es hat sich gezeigt, dass es zwischen orthologen ECF-Sigmafaktoren in Aktinomyzeten viele Parallelen in der Genorganisation, dem Mechanismus der Inaktivierung und in den entsprechenden Regulationnetzwerken gibt. Die Netzwerke enthalten Gene mit ähnlichen Funktionen in der Stressbekämpfung und werden über ähnliche Promotormotive gesteuert. Kleinere Unterschiede in den Regulationsnetzwerken sind sicherlich den besonderen Anpassungen an den Lebensraum und die Lebensweise geschuldet. Auch die Genomgröße kann dabei eine Rolle spielen. So ist es auffällig, dass es in den größeren Aktinomyzeten Genomen zu einer Expansion bestimmter Sigmafaktorfamilien gekommen ist. Auf diese Weise wird sicherlich noch ein noch größeres Maß an Feinabstimmung erreicht. Die Vielzahl von paralogen ECF-Sigmafaktoren stellt aber auch ein großes Hindernis für die experimentelle Bearbeitung einzelner Paraloge *in vivo* dar.

So ist es besonders von Interesse, die auf Basis von *C. glutamicum* als Modellorganismus entwickelten Methoden auf diese Actinobakterien zu übertragen und die Vorteile der RNA-Seq als *Top-Down* Methode in Kombination mit der *Bottom-up* Methode ROSE zur Entschlüsselung der Regulationsnetzwerke in diesen wesentlich komplexer organisierten Actinobakterien zu untersuchen. Speziell für Streptomyceten mit ihren großen Genomen und der damit verbundenen enormen Ausstattung an Sigmafaktoren wird sich die hier entwickelte Herangehensweise als nützlich erweisen.

8. Ausblick

Im Rahmen dieser Arbeit konnten durch ein optimiertes Methodenspektrum die ECF-Sigmafaktor-Netzwerke von SigH und SigE, ihre Überschneidung und die transkriptionelle Organisation der beiden Sigmafaktor/Anti-Sigmafaktor Operons vertieft untersucht werden. Die Ergebnisse dieser Arbeit führen nun zusammen mit Erkenntnissen aus der Literatur zu einem besseren Verständnis der Funktionen und Verknüpfung aller Sigmafaktoren in *C. glutamicum*. In dieser Arbeit stellte hierbei die etablierte ROSE-Methodik eine entscheidende Erweiterung im methodischen Spektrum zur genomweiten Analyse und Differenzierung der ECF-Sigmafaktor-Netzwerke von SigH und SigE dar.

Bei dem Vergleich der Ergebnisse der einzelnen Methoden zeigte sich aber, dass nicht alle Gene der Sigmafaktor-Netzwerke von SigH und SigE mit nur einer der verwendeten Methoden bestimmt werden konnten. Zwar stellt die *Bottom-Up*-Methode ROSE eine wichtige Ergänzung zu den *Top-Down*-Analysen dar, sie wurde aber erstmalig in dieser Arbeit eingesetzt und hat noch Optimierungsbedarf.

Es konnte festgestellt werden, dass einige SigH-abhängige Promotoren *in vivo* aktiv waren, aber nicht in den *in vitro* ROSE-Analysen. Wie diskutiert wurde, könnte die in den ROSE-Analysen als Template genutzte lineare DNA der Grund für die ausbleibende Aktivität sein. Um dies zu klären, sollte die gesamte ROSE-Analyse mit DNA unterschiedlicher Konformationen durchgeführt werden. Alternativ dazu könnte man auch die Promotoren die in der ROSE-Analyse nicht aktiv waren, einzeln durch *in vitro* Transkription von in Plasmiden klonierten Promotoren untersuchen. Durch die Nutzung von *EcoRI*-gespaltener chromosomale DNA in den ROSE-Analysen konnten einige Transkripte aufgrund der Lage der *EcoRI*-Schnittstelle nicht analysiert werden. Hier sollte in Folgeanalysen zufällig mechanisch fragmentierte chromosomale DNA eingesetzt werden.

Die in dieser Arbeit identifizierten Positionen des -35 Promotormotivs, die für die Unterscheidung von SigH- und SigE-abhängigen Promotoren sind, sollten durch Promotormutations-Analysen sowohl *in vitro*, als auch *in vivo* bestätigt werden.

Nach erfolgreicher methodischer Optimierung der ROSE-Methodik ergeben sich vielfältige Möglichkeiten zum Einsatz dieser Methode:

Um das Sigmafaktor-Netzwerk in *C. glutamicum* vollständig aufzuklären, sollte die in dieser Arbeit verwendete Kombination aus *Top-Down*- und *Bottom-Up*-Analyse auf die bisher nicht untersuchten ECF-Sigmafaktor-Netzwerke SigC und SigD angewendet werden. Daneben sollte ROSE eingesetzt werden, um aufzuklären, wie sich die Promotoren, die von dem *Housekeeping* Sigmafaktor SigA und dem alternativen Sigmafaktor SigB, unterscheiden. Da angenommen wird, dass SigB Promotoren eventuell abhängig vom Konformationsstatus der DNA erkennt, sollte man hier zum einen lineare DNA einsetzen, um die rein auf der Sequenz basierenden Erkennung zu klären und zum anderen supercoiled DNA, um die Frage zu klären, in wie weit diese Einfluß auf die Erkennung durch einen der beiden Sigmafaktoren hat.

Neben der Analyse einzelner Sigmafaktor-Netzwerke sollte ROSE eingesetzt werden, um Fragen zur Sigmafaktor-Kompetition und damit verbundenen Affinität der einzelnen Sigmafaktoren zu dem RNAP-Core Enzym aufzuklären. Hier können verschiedene Mischungen von Sigmafaktor-Proteinen zum Einsatz kommen. Allerdings erscheint es dafür notwendig, anstelle des heterologen RNA-Polymerase *core*-Enzyms aus *E. coli* ein *C. glutamicum core*-Enzym einzusetzen. In anderen Labors sind aber bereits Anstrengungen zur Isolierung des *C. glutamicum* RNAP-Core Enzyms unternommen worden (Holátko *et al.*, 2012).

ROSE wurde beim Modell-Organismus *C. glutamicum* entwickelt und soll bei uns in Zukunft auch auf andere Actinobakterien angewendet werden um deren komplexere Sigmafaktor-Netzwerke zu analysieren. Einem Einsatz bei vielen anderen Bakterien steht aber prinzipiell auch nichts im Weg.

Neben dem Einsatz der ROSE-Methodik zur Analyse von Sigmafaktor-Netzwerken kann und sollte die Methode auf die Analyse von weiteren Transkriptionsregulatoren ausgeweitet werden. Hier bieten sich einerseits Regulatoren an, die mit der RNA-Polymerase interagieren und ebenfalls *in vivo* schwer zu analysieren sind, weil die Gene essentiell sind und nicht mutiert werden können. Ein Beispiel hierfür ist in Actinobakterien der essentielle transkriptionelle Regulator der Stringenten Kontrolle CarD (Connolly und Cox, 2009), für den in *M. tuberculosis* eine direkte Interaktion mit dem N-Terminus der RNAP β -Untereinheit gezeigt werden konnte (Stallings und Glickman, 2011).

Grundsätzlich steht aber der Erweiterung von ROSE auf komplexere Regulationssysteme in allen Bakterien nichts entgegen. Es ist daher zu erwarten, dass die ROSE-Methodik einen festen Platz im wissenschaftlichen Methodenspektrum zur Untersuchung von Transkriptionsregulationsnetzwerken erhalten wird.

VI. Literaturverzeichnis

- Ajinomoto**, 2012. Fact sheet: Food Products Business. <http://www1.aminovital.com/en/ir/pdf/Food-Oct2012.pdf>.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.**, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* 25, 3389–402.
- Ao, W., Gaudet, J., Kent, W.J., Muttumu, S., Mango, S.E.**, 2004. Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. *Science (New York, N.Y.)* 305, 1743–6.
- Bansal-Mutalik, R., Nikaido, H.**, 2011. Quantitative lipid composition of cell envelopes of *Corynebacterium glutamicum* elucidated through reverse micelle extraction. *Proceedings of the National Academy of Sciences of the United States of America* 108, 15360–5.
- Bardonnet, N., Blanco, C.**, 1991. Improved vectors for transcriptional signal screening in corynebacteria. *FEMS microbiology letters* 68, 97–102.
- Barik, S., Sureka, K., Mukherjee, P., Basu, J., Kundu, M.**, 2010. RseA, the SigE specific anti-sigma factor of *Mycobacterium tuberculosis*, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Molecular microbiology* 75, 592–606.
- Barreiro, C., Nakunst, D., Huser, A.T., De Paz, H.D., Kalinowski, J., Martin, J.F.**, 2009. Microarray studies reveal a “differential response” to moderate or severe heat shock of the HrcA- and HspR-dependent systems in *Corynebacterium glutamicum*. *Microbiology* 155, 359–372.
- Becker, G., Hengge-Aronis, R.**, 2001. What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). *Molecular microbiology* 39, 1153–65.
- Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A., Duncan, K.**, 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Molecular microbiology* 43, 717–31.
- Birnboim, H.C., Doly, J.**, 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* 7, 1513–23.
- Blom, Jochen, Jakobi, T., Doppmeier, D., Stoye, J., Goesmann, Alexander**, 2011. Exact and complete short read alignment to microbial genomes using GPU programming. *Genome* 1–8.
- Boos, W., Böhm, A.**, 2000. Learning new tricks from an old dog: MalT of the *Escherichia coli* maltose system is part of a complex regulatory network. *Trends in genetics : TIG* 16, 404–9.
- Borukhov, S., Nudler, E.**, 2003. RNA polymerase holoenzyme: structure, function and biological implications. *Current opinion in microbiology* 6, 93–100.
- Borukhov, S., Nudler, E.**, 2008. RNA polymerase: the vehicle of transcription. *Trends in microbiology* 16, 126–34.
- Bradford, M.M.**, 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* 72, 248–54.
- Brinkrolf, K., Schröder, J., Pühler, A., Tauch, A.**, 2010. The transcriptional regulatory repertoire of *Corynebacterium glutamicum*: reconstruction of the network controlling pathways involved in lysine and glutamate production. *Journal of biotechnology* 149, 173–82.
- Brockmann-Gretza, O., Kalinowski, J.**, 2006. Global gene expression during stringent response in *Corynebacterium glutamicum* in presence and absence of the rel gene encoding (p)ppGpp synthase. *BMC genomics* 7, 230.
- Browning, D.F., Busby, S.J.**, 2004. The regulation of bacterial transcription initiation. *Nature reviews. Microbiology* 2, 57–65.
- Burgess, R.R., Travers, A.A., Dunn, J.J., Bautz, E.K.**, 1969. Factor stimulating transcription by RNA polymerase. *Nature* 221, 43–6.

- Burns, K.E., Liu, W.T., Boshoff, H.I., Dorrestein, P.C., Barry 3rd, C.E.**, 2009. Proteasomal protein degradation in Mycobacteria is dependent upon a prokaryotic ubiquitin-like protein. *J Biol Chem* 284, 3069–3075.
- Busche, T., Silar, Radoslav, Pičmanová, M., Pátek, M., Kalinowski, J.**, 2012. Transcriptional regulation of the operon encoding stress-responsive ECF sigma factor SigH and its anti-sigma factor RshA, and control of its regulatory network in *Corynebacterium glutamicum*. *BMC genomics* 13, 445.
- Bustamante, C., Guthold, M., Zhu, X., Yang, G.**, 1999. Facilitated target location on DNA by individual *Escherichia coli* RNA polymerase molecules observed with the scanning force microscope operating in liquid. *The Journal of biological chemistry* 274, 16665–8.
- Cadenas, R.F., Martín, J.F., Gil, J.A.**, 1991. Construction and characterization of promoter-probe vectors for *Corynebacteria* using the kanamycin-resistance reporter gene. *Gene* 98, 117–21.
- Calabrese, F., Molteni, R., Maj, P.F., Cattaneo, A., Gennarelli, M., Racagni, G., Riva, M.A.**, 2007. Chronic duloxetine treatment induces specific changes in the expression of BDNF transcripts and in the subcellular localization of the neurotrophin protein. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 32, 2351–9.
- Calamita, H., Ko, C., Tyagi, Sandeep, Yoshimatsu, T., Morrison, N.E., Bishai, W.R.**, 2005. The *Mycobacterium tuberculosis* SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence 7, 233–244.
- Campbell, E. a, Greenwell, R., Anthony, J.R., Wang, S., Lim, L., Das, K., Sofia, H.J., Donohue, T.J., Darst, S. a**, 2007. A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Molecular cell* 27, 793–805.
- Campbell, E.A., Westblade, L.F., Darst, S.A.**, 2008. Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr Opin Microbiol* 11, 121–127.
- Cao, M., Kobel, P. a, Morshedi, M.M., Wu, M.F.W., Paddon, C., Helmann, John D**, 2002. Defining the *Bacillus subtilis* sigma(W) regulon: a comparative analysis of promoter consensus search, run-off transcription/microarray analysis (ROMA), and transcriptional profiling approaches. *Journal of molecular biology* 316, 443–57.
- Chang, A., Smollett, K.L., Gopaul, K.K., Chan, B.H.Y., Davis, E.O.**, 2012. *Mycobacterium tuberculosis* H37Rv sigC is expressed from two promoters but is not auto-regulatory. *Tuberculosis (Edinburgh, Scotland)* 92, 48–55.
- Ciampi, M.S.**, 2006. Rho-dependent terminators and transcription termination. *Microbiology (Reading, England)* 152, 2515–28.
- Condon, C.**, 2007. Maturation and degradation of RNA in bacteria. *Current opinion in microbiology* 10, 271–8.
- Connolly, L.E., Cox, J.S.**, 2009. CarD tricks and magic spots: mechanisms of stringent control in mycobacteria. *Cell host & microbe* 6, 1–2.
- Constantinides, A.**, 1980. Steroid transformation at high substrate concentrations using immobilized *Corynebacterium simplex* cells. *Biotechnology and bioengineering* 22, 119–36.
- Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E.**, 2004. WebLogo: a sequence logo generator. *Genome Res* 14, 1188–1190.
- de Hase, P.L., Zupancic, M.L., Record, M.T.**, 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *Journal of bacteriology* 180, 3019–25.
- Donà, V., Rodrigue, S., Dainese, E., Palù, G., Gaudreau, L., Manganelli, Riccardo, Provvedi, R.**, 2008. Evidence of complex transcriptional, translational, and posttranslational regulation of the extracytoplasmic function sigma factor sigmaE in *Mycobacterium tuberculosis*. *Journal of bacteriology* 190, 5963–71.
- Dondrup, M., Albaum, S.P., Griebel, T., Henckel, K., Junemann, S., Kahlke, T., Kleindt, C.K., Kuster, H., Linke, B., Mertens, D., Mittard-Runte, V., Neuweger, H., Runte, K.J., Tauch, A, Tille, F., Pühler, A., Goesmann, A**, 2009. EMMA 2--a MAGE-compliant system for the collaborative analysis and integration of microarray data. *BMC Bioinformatics* 10, 50.

- Ehira, S., Shirai, T., Teramoto, H., Inui, M., Yukawa, H., 2008.** Group 2 sigma factor SigB of *Corynebacterium glutamicum* positively regulates glucose metabolism under conditions of oxygen deprivation. *Appl Environ Microbiol* 74, 5146–5152.
- Ehira, S., Teramoto, H., Inui, M., Yukawa, H., 2009.** Regulation of *Corynebacterium glutamicum* heat shock response by the extracytoplasmic-function sigma factor SigH and transcriptional regulators HspR and HrcA. *J Bacteriol* 191, 2964–2972.
- Eikmanns, B.J., Kleinertz, E., Liebl, W., Sahm, H., 1991.** A family of *Corynebacterium glutamicum*/*Escherichia coli* shuttle vectors for cloning, controlled gene expression, and promoter probing. *Gene* 102, 93–8.
- Engels, S., Schweitzer, J.E., Ludwig, C., Bott, M., Schaffer, S., 2004.** *clpC* and *clpP1P2* gene expression in *Corynebacterium glutamicum* is controlled by a regulatory network involving the transcriptional regulators ClgR and HspR as well as the ECF sigma factor SigmaH. *Mol Microbiol* 52, 285–302.
- Esken, J., 2011.** Erstellung eines *in vitro* Transkriptionssystems in *Corynebacterium Glutamicum* zur Analyse von Regulationsnetzwerken extracytoplasmic function (ECF)-Sigmafaktoren. Diplomarbeit an der Universität Bielefeld, Fakultät für Biologie, Lehrstuhl für Genetik.
- Feng, J., Che, Y., Milse, J., Yin, Y.J., Liu, L., Ruckert, C., Shen, X.H., Qi, S.W., Kalinowski, J, Liu, S.J., 2006.** The gene *ncgl2918* encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in *Corynebacterium glutamicum*. *J Biol Chem* 281, 10778–10785.
- Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., Delong, E.F., 2008.** Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences of the United States of America* 105, 3805–10.
- Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., Ellenberger, T., 2005.** DNA Repair and Mutagenesis. American Society for Microbiology, Washington, DC.
- Funke, G., Lawson, P.A., Collins, M.D., 1995.** Heterogeneity within human-derived centers for disease control and prevention (CDC) coryneform group ANF-1-like bacteria and description of *Corynebacterium auris* sp. nov. *International journal of systematic bacteriology* 45, 735–9.
- Gibson, D.G., Smith, H.O., Hutchison, C.A., Venter, J.C., Merryman, C., 2010.** Chemical synthesis of the mouse mitochondrial genome. *Nature methods* 7, 901–3.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009.** Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* 6, 343–5.
- Goodfellow, M., Whitman, W.B., Kämpfer, P., Busse, H.-J., Ludwig, W., Trujillo, M.E., Suzuki, K., 2012.** *Bergey's Manual of Systematic Bacteriology: Volume 5: The Actinobacteria*, 2nd ed. Springer Verlag, New York.
- Grant, S.G., Jessee, J., Bloom, F.R., Hanahan, D., 1990.** Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America* 87, 4645–9.
- Gruber, T M, Gross, C A, 2003.** Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* 57, 441–466.
- Halgasova, N., Bukovska, G., Timko, J., Kormanec, J., 2001.** Cloning and transcriptional characterization of two sigma factor genes, *sigA* and *sigB*, from *Brevibacterium flavum*. *Curr Microbiol* 43, 249–254.
- Hanahan, D., 1983.** Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* 166, 557–80.
- Hansmeier, Nicole, Chao, T.-C., Pühler, A., Tauch, A., Kalinowski, J., 2006.** The cytosolic, cell surface and extracellular proteomes of the biotechnologically important soil bacterium *Corynebacterium efficiens* YS-314 in comparison to those of *Corynebacterium glutamicum* ATCC 13032. *Proteomics* 6, 233–50.

- Hasselt, K., Rankl, S., Worsch, S., Burkovski, A.**, 2011. Adaptation of AmtR-controlled gene expression by modulation of AmtR binding activity in *Corynebacterium glutamicum*. *Journal of biotechnology* 154, 156–62.
- Haugen, S.P., Ross, W., Gourse, R.L.**, 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature reviews. Microbiology* 6, 507–19.
- He, H., Hovey, R., Kane, J., Singh, V., Zahrt, T.C.**, 2006. MprAB is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in *Mycobacterium tuberculosis*. *Journal of bacteriology* 188, 2134–43.
- Helmann, J D**, 1999. Anti-sigma factors. *Current opinion in microbiology* 2, 135–41.
- Helmann, J D, deHaseth, P.L.**, 1999. Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry* 38, 5959–67.
- Helmann, John D**, 2002. The Extracytoplasmic Function (ECF) Sigma Factors.
- Helmann, John D**, 2009. RNA polymerase: a nexus of gene regulation. *Methods (San Diego, Calif.)* 47, 1–5.
- Hilker, R., Blom, J., Goesmann, A.**, 2013. VAMP - Visualization and Analysis of Mapped Sequences. In preparation.
- Ho, T.D., Ellermeier, C.D.**, 2012. Extra cytoplasmic function σ factor activation. *Current opinion in microbiology*.
- Hoer, P.A.C., Ariyurek, Y., Thygesen, H.H., Vreugdenhil, E., Vossen, R.H.A.M., De Menezes, R.X., Boer, J.M., Van Ommen, G.-J.B., Den Dunnen, J.T.**, 2008. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic acids research* 36, e141.
- Holátko, J., Šilar, R., Rabatinová, A.**, 2012. Construction of in vitro transcription system for *Corynebacterium glutamicum* and its use in the recognition of promoters of different classes. *Applied Microbiology*.
- Homuth, G., Heinemann, M., Zuber, U., Schumann, W.**, 1996. The genes of lepA and hemN form a bicistronic operon in *Bacillus subtilis*. *Microbiology* 142 (Pt 7, 1641–1649.
- Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., Pease, L.R.**, 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61–68.
- Hsu, L.M.**, 2009. Monitoring abortive initiation. *Methods (San Diego, Calif.)* 47, 25–36.
- Hüser, A.T., Becker, A., Brune, I., Dondrup, M., Kalinowski, J., Plassmeier, J., Pühler, A, Wieggrabe, I., Tauch, A**, 2003. Development of a *Corynebacterium glutamicum* DNA microarray and validation by genome-wide expression profiling during growth with propionate as carbon source. *J Biotechnol* 106, 269–286.
- Hüttges, A.**, 2005. Charakterisierung des ECF-Sigmafaktors SigE und seines Regulons aus *Corynebacterium glutamicum* mittels DNA-Microarrays. Diplomarbeit an der Universität Bielefeld, Fakultät für Biologie, Lehrstuhl für Genetik.
- Ikeda, M., Nakagawa, S.**, 2003. The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes. *Appl Microbiol Biotechnol* 62, 99–109.
- Ikeda, Masato, Baba, M., Tsukamoto, N., Komatsu, T., Mitsunashi, S., Takeno, S.**, 2009. Elucidation of genes relevant to the microaerobic growth of *Corynebacterium glutamicum*. *Bioscience, biotechnology, and biochemistry* 73, 2806–8.
- Jakob, K., Satorhelyi, P., Lange, C., Wendisch, Volker F, Silakowski, B., Scherer, S., Neuhaus, K.**, 2007. Gene expression analysis of *Corynebacterium glutamicum* subjected to long-term lactic acid adaptation. *Journal of bacteriology* 189, 5582–90.
- Jakoby, M., Nolden, L., Meier-Wagner, J., Krämer, R, Burkovski, A**, 2000. AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. *Molecular microbiology* 37, 964–77.
- Jochmann, N., Kurze, A.K., Czaja, L.F., Brinkrolf, K, Brune, I., Hüser, A.T., Hansmeier, N, Pühler, A., Borovok, I., Tauch, A**, 2009. Genetic makeup of the *Corynebacterium glutamicum* LexA

- regulon deduced from comparative transcriptomics and in vitro DNA band shift assays. *Microbiology* 155, 1459–1477.
- Jung, Y.G., Cho, Y.B., Kim, M.S., Yoo, J.S., Hong, S.H., Roe, J.H.**, 2011. Determinants of redox sensitivity in RsrA, a zinc-containing anti-sigma factor for regulating thiol oxidative stress response. *Nucleic Acids Res* 39, 7586–7597.
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Krämer, R., Linke, B., McHardy, A.C., Meyer, F., Mockel, B., Pfefferle, W., Pühler, A., Rey, D.A., Rückert, C., Rupp, O., Sahm, H., Wendisch, V.F., Wiegrabe, I., Tauch, A.**, 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol* 104, 5–25.
- Kang, J.G., Paget, M.S., Seok, Y.J., Hahn, M.Y., Bae, J.B., Hahn, J.S., Kleanthous, C., Buttner, M.J., Roe, J.H.**, 1999. RsrA, an anti-sigma factor regulated by redox change. *Embo J* 18, 4292–4298.
- Keilhauer, C., Eggeling, L., Sahm, H.**, 1993. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J Bacteriol* 175, 5595–5603.
- Kim, M.-S., Dufour, Y.S., Yoo, J. Sun, Cho, Y.-B., Park, J.-H., Nam, G.-B., Kim, H.M., Lee, K.-L., Donohue, T.J., Roe, J.-H.**, 2012. Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes. *Molecular microbiology* 85, 326–44.
- Kim, M.S., Hahn, M.Y., Cho, Y., Cho, S.N., Roe, J.H.**, 2009. Positive and negative feedback regulatory loops of thiol-oxidative stress response mediated by an unstable isoform of sigmaR in actinomycetes. *Mol Microbiol* 73, 815–825.
- Kim, T.H., Kim, H.J., Park, J.S., Kim, Y., Kim, P., Lee, H.S.**, 2005. Functional analysis of *sigh* expression in *Corynebacterium glutamicum*. *Biochem Biophys Res Commun* 331, 1542–1547.
- Kinoshita, S., Udaka, S., Shimono, M.**, 2004. Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. *The Journal of general and applied microbiology* 50, 331–43.
- Kirchner, Oliver, Tauch, A.**, 2003. Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *Journal of biotechnology* 104, 287–99.
- Kusano, S., Ding, Q., Fujita, N., Ishihama, A.**, 1996. Promoter selectivity of *Escherichia coli* RNA polymerase E sigma 70 and E sigma 38 holoenzymes. Effect of DNA supercoiling. *The Journal of biological chemistry* 271, 1998–2004.
- Laemmli, U.K.**, 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–5.
- Larisch, C., Nakunst, D., Hüser, A.T., Tauch, A., Kalinowski, J.**, 2007. The alternative sigma factor SigB of *Corynebacterium glutamicum* modulates global gene expression during transition from exponential growth to stationary phase. *BMC Genomics* 8, 4.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G.**, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- Lawrence, C.E., Altschul, S.F., Boguski, M.S., Liu, J.S., Neuwald, A.F., Wootton, J.C.**, 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science (New York, N.Y.)* 262, 208–14.
- Lawrence, C.E., Reilly, A.A.**, 1990. An expectation maximization (EM) algorithm for the identification and characterization of common sites in unaligned biopolymer sequences. *Proteins* 7, 41–51.
- Lewis, M.**, 2005. The lac repressor. *Comptes rendus biologiques* 328, 521–48.
- Liebl, W.**, 2005. *Corynebacterium* taxonomy, in: *Handbook of Corynebacterium glutamicum* (Eggeling, L. and Bott, M., Ed.), CRC Press Taylor & Francis Group, pp. 149–186.
- Liebl, W., Bayerl, A., Schein, B., Stillner, U., Schleifer, K.H.**, 1989. High efficiency electroporation of intact *Corynebacterium glutamicum* cells. *FEMS microbiology letters* 53, 299–303.
- Liu, X., Brutlag, D.L., Liu, J.S.**, 2001. BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac Symp Biocomput* 127–138.

- Lonetto, M. a, Brown, K.L., Rudd, K.E., Buttner, M.J.,** 1994. Analysis of the *Streptomyces coelicolor* *sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proceedings of the National Academy of Sciences of the United States of America 91, 7573–7.
- Maciag, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., Landini, P.,** 2011. *In vitro* transcription profiling of the SigmaS subunit of bacterial RNA polymerase: re-definition of the SigmaS regulon and identification of SigmaS-specific promoter sequence elements. Nucleic acids research 70, 1–18.
- MacLellan, S.R., Eiamphungporn, W., Helmann, J. D.,** 2009. ROMA: an in vitro approach to defining target genes for transcription regulators. Methods (San Diego, Calif.) 47, 73–7.
- MacLellan, S.R., Wecke, T., Helmann, John D,** 2008. A previously unidentified sigma factor and two accessory proteins regulate oxalate decarboxylase expression in *Bacillus subtilis*. Molecular microbiology 69, 954–67.
- Madan Babu, M., Teichmann, S.A.,** 2003. Functional determinants of transcription factors in *Escherichia coli*: protein families and binding sites. Trends in genetics : TIG 19, 75–9.
- Manganelli, R, Dubnau, E., Tyagi, S, Kramer, F.R., Smith, I.,** 1999. Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. Molecular microbiology 31, 715–24.
- Manganelli, R, Voskuil, M.I., Schoolnik, G.K., Dubnau, E., Gomez, M., Smith, I.,** 2002. Role of the extracytoplasmic-function sigma factor sigma(H) in *Mycobacterium tuberculosis* global gene expression. Mol Microbiol 45, 365–374.
- Manganelli, R, Voskuil, M.I., Schoolnik, G.K., Smith, I.,** 2001. The *Mycobacterium tuberculosis* ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. Molecular microbiology 41, 423–37.
- Meyer, Folker, Goesmann, Alexander, McHardy, Alice C, Bartels, Daniela, Bekel, Thomas, Clausen, J., Kalinowski, J., Linke, Burkhard, Rupp, Oliver, Giegerich, R., Pühler, A.,** 2003. GenDB--an open source genome annotation system for prokaryote genomes. Nucleic acids research 31, 2187–95.
- Morrison, D.A.,** 1977. Transformation in *Escherichia coli*: cryogenic preservation of competent cells. Journal of bacteriology 132, 349–51.
- Nakamura, Y., Nishio, Y., Ikeo, K., Gojobori, T.,** 2003. The genome stability in *Corynebacterium* species due to lack of the recombinational repair system. Gene 317, 149–55.
- Nakunst, D., Larisch, C., Huser, A.T., Tauch, A, Pühler, A., Kalinowski, J,** 2007. The extracytoplasmic function-type sigma factor SigM of *Corynebacterium glutamicum* ATCC 13032 is involved in transcription of disulfide stress-related genes. J Bacteriol 189, 4696–4707.
- Newton, G.L., Arnold, K., Price, M.S., Sherrill, C., Delcardayre, S.B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R.C., Davis, C.,** 1996. Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. J Bacteriol 178, 1990–1995.
- Newton, G.L., Buchmeier, N., Fahey, R.C.,** 2008. Biosynthesis and functions of mycothiol, the unique protective thiol of *Actinobacteria*. Microbiol Mol Biol Rev 72, 471–494.
- Newton, G.L., Fahey, R.C.,** 2008. Regulation of mycothiol metabolism by sigma(R) and the thiol redox sensor anti-sigma factor RsrA. Mol Microbiol 68, 805–809.
- Nolden, L., Beckers, G., Möckel, B., Pfefferle, W., Nampoothiri, K.M., Krämera, R., Burkovskia, A.,** 2000. Urease of *Corynebacterium glutamicum*: organization of corresponding genes and investigation of activity. FEMS microbiology letters 189, 305–10.
- Notredame, C., Higgins, D.G., Heringa, J.,** 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology 302, 205–17.
- Oguiza, J.A., Marcos, A.T., Martín, J.F.,** 1997. Transcriptional analysis of the *sigA* and *sigB* genes of *Brevibacterium lactofermentum*. FEMS microbiology letters 153, 111–7.
- Olvera, L., Mendoza-Vargas, A., Flores, N., Olvera, M., Sigala, J.C., Gosset, G., Morett, E., Bolívar, F.,** 2009. Transcription analysis of central metabolism genes in *Escherichia coli*. Possible roles of sigma38 in their expression, as a response to carbon limitation. PloS one 4, e7466.

- Pabo, C.O., Sauer, R.T.**, 1992. Transcription factors: structural families and principles of DNA recognition. *Annual review of biochemistry* 61, 1053–95.
- Paget, M.S., Molle, V., Cohen, G., Aharonowitz, Y., Buttner, M.J.**, 2001. Defining the disulphide stress response in *Streptomyces coelicolor* A3(2): identification of the sigmaR regulon. *Mol Microbiol* 42, 1007–1020.
- Park, J.H., Roe, J.H.**, 2008. Mycothiol regulates and is regulated by a thiol-specific antisigma factor RsrA and sigma(R) in *Streptomyces coelicolor*. *Mol Microbiol* 68, 861–870.
- Park, S.D., Youn, J.W., Kim, Y.J., Lee, S.M., Kim, Y., Lee, H.S.**, 2008. *Corynebacterium glutamicum* SigmaE is involved in responses to cell surface stresses and its activity is controlled by the anti-sigma factor CseE. *Microbiology* 154, 915–923.
- Passalacqua, K.D., Varadarajan, A., Ondov, B.D., Okou, D.T., Zwick, M.E., Bergman, N.H.**, 2009. Structure and complexity of a bacterial transcriptome. *Journal of bacteriology* 191, 3203–11.
- Pátek, M., Holátko, J., Busche, T., Kalinowski, J., Nešvera, J.**, 2013. *Corynebacterium glutamicum* promoters: a practical approach. *Microbial biotechnology*.
- Pátek, M., Nesvera, J.**, 2011. Sigma factors and promoters in *Corynebacterium glutamicum*. *J Biotechnol* 154, 101–113.
- Pátek, M., Nesvera, Jan, Guyonvarch, A., Reyes, O., Leblon, G.**, 2003. Promoters of *Corynebacterium glutamicum*. *Journal of biotechnology* 104, 311–23.
- Pauling, J., Röttger, R., Tauch, A., Azevedo, V., Baumbach, J.**, 2012. CoryneRegNet 6.0--Updated database content, new analysis methods and novel features focusing on community demands. *Nucleic acids research* 40, D610–4.
- Pérez-Rueda, E., Collado-Vides, J.**, 2001. Common history at the origin of the position-function correlation in transcriptional regulators in archaea and bacteria. *Journal of molecular evolution* 53, 172–9.
- Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S.**, 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–67.
- Perkins, T.T., Kingsley, R. a, Fookes, M.C., Gardner, P.P., James, K.D., Yu, L., Assefa, S. a, He, M., Croucher, N.J., Pickard, D.J., Maskell, D.J., Parkhill, J., Choudhary, J., Thomson, N.R., Dougan, G.**, 2009. A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. *PLoS genetics* 5, e1000569.
- Pfeifer-Sancar, K., Rückert, C., Kalinowski, J.**, 2013. Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. In preparation.
- Qiu, J., Helmann, J D.**, 2001. The -10 region is a key promoter specificity determinant for the *Bacillus subtilis* extracytoplasmic-function sigma factors sigma(X) and sigma(W). *Journal of bacteriology* 183, 1921–7.
- Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs Jr., W.R., Husson, R.N.**, 2001a. The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J Bacteriol* 183, 6119–6125.
- Rodrigue, S., Brodeur, J., Jacques, P.-E., Gervais, A.L., Brzezinski, R., Gaudreau, L.**, 2007. Identification of mycobacterial sigma factor binding sites by chromatin immunoprecipitation assays. *Journal of bacteriology* 189, 1505–13.
- Rodrigue, S., Provvedi, R., Jacques, P.-E., Gaudreau, L., Manganelli, Riccardo.**, 2006. The sigma factors of *Mycobacterium tuberculosis*. *FEMS microbiology reviews* 30, 926–41.
- Sachdeva, P., Misra, R., Tyagi, A.K., Singh, Y.**, 2010. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *The FEBS journal* 277, 605–26.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A.**, 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (New York, N.Y.)* 239, 487–91.
- Sambrook, J., Fritsch, E.F., Maniatis, T.**, 1989. *Molecular Cloning: a Laboratory Manual*, 2nd Edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

- Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., Pühler, A., 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145, 69–73.
- Schaffer, S., Weil, B., Nguyen, V.D., Dongmann, G., Günther, K., Nickolaus, M., Hermann, T., Bott, M., 2001. A high-resolution reference map for cytoplasmic and membrane-associated proteins of *Corynebacterium glutamicum*. *Electrophoresis* 22, 4404–22.
- Schnarr, M., Oertel-Buchheit, P., Kazmaier, M., Granger-Schnarr, M., 1991. DNA binding properties of the LexA repressor. *Biochimie* 73, 423–431.
- Schröder, J., Maus, I., Trost, E., Tauch, A., 2011. Complete genome sequence of *Corynebacterium variabile* DSM 44702 isolated from the surface of smear-ripened cheeses and insights into cheese ripening and flavor generation. *BMC genomics* 12, 545.
- Schröder, J., Tauch, A., 2010. Transcriptional regulation of gene expression in *Corynebacterium glutamicum*: the role of global, master and local regulators in the modular and hierarchical gene regulatory network. *FEMS microbiology reviews* 34, 685–737.
- Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., Stadler, P.F., Vogel, J., 2010. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464, 250–5.
- Sharma, U.K., Chatterji, D., 2010. Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of sigma activity. *FEMS microbiology reviews* 34, 646–57.
- Silberbach, M., Hüser, A., Kalinowski, J., Pühler, A., Walter, B., Krämer, Reinhard, Burkovski, A., 2005. DNA microarray analysis of the nitrogen starvation response of *Corynebacterium glutamicum*. *Journal of biotechnology* 119, 357–67.
- Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C.D., Vogel, J., 2008. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS genetics* 4, e1000163.
- Song, T., Dove, S.L., Lee, K.H., Husson, R.N., 2003. RshA, an anti-sigma factor that regulates the activity of the mycobacterial stress response sigma factor SigH. *Mol Microbiol* 50, 949–959.
- Song, T., Song, S.E., Raman, S., Anaya, M., Husson, R.N., 2008. Critical role of a single position in the -35 element for promoter recognition by *Mycobacterium tuberculosis* SigE and SigH. *J Bacteriol* 190, 2227–2230.
- Sorek, Rotem, Cossart, P., 2010. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nature reviews. Genetics* 11, 9–16.
- Soshi, A., 2008. Charakterisierung des Sigmafaktors SigH und seines Regulationsnetzwerks aus *Corynebacterium glutamicum* ATCC 13032. Diplomarbeit an der Universität Bielefeld, Fakultät für Biologie, Lehrstuhl für Genetik.
- Stallings, C.L., Glickman, M.S., 2011. CarD: a new RNA polymerase modulator in mycobacteria. *Transcription* 2, 15–8.
- Staroń, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H., Mascher, T., 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Molecular microbiology* 74, 557–81.
- Stock, J., Da Re, S., 2000. Signal transduction: response regulators on and off. *Current biology : CB* 10, R420–4.
- Tauch, A., Kaiser, O., Hain, T., Goesmann, A., Weisshaar, B., Albersmeier, A., Bekel, T., Bischoff, N., Brune, I., Chakraborty, T., Kalinowski, J., Meyer, F., Rupp, O., Schneiker, S., Viehoveer, P., Pühler, A., 2005. Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora. *J Bacteriol* 187, 4671–4682.
- Tauch, A., Kassing, F., Kalinowski, J., Pühler, A., 1995a. The *Corynebacterium xerosis* composite transposon Tn5432 consists of two identical insertion sequences, designated IS1249, flanking the erythromycin resistance gene *ermCX*. *Plasmid* 34, 119–31.

- Tauch, A, Kassing, F., Kalinowski, J, Pühler, A.,** 1995b. The *Corynebacterium xerosis* composite transposon Tn5432 consists of two identical insertion sequences, designated IS1249, flanking the erythromycin resistance gene *ermCX*. Plasmid 34, 119–131.
- Tauch, A, Kirchner, O, Löffler, B., Gotker, S., Pühler, A., Kalinowski, J,** 2002. Efficient electrotransformation of *Corynebacterium diphtheriae* with a mini-replicon derived from the *Corynebacterium glutamicum* plasmid pGA1. Curr Microbiol 45, 362–367.
- Tauch, A, Kirchner, O, Wehmeier, L., Kalinowski, J, Pühler, A.,** 1994. *Corynebacterium glutamicum* DNA is subjected to methylation-restriction in *Escherichia coli*. FEMS Microbiol Lett 123, 343–347.
- Thakur, K.G., Gopal, B.,** 2005. Crystallization and preliminary X-ray diffraction studies of two domains of a bilobed extra-cytoplasmic function sigma factor SigC from Mycobacterium tuberculosis. Acta crystallographica. Section F, Structural biology and crystallization communications 61, 779–81.
- Thakur, K.G., Joshi, A.M., Gopal, B.,** 2007. Structural and biophysical studies on two promoter recognition domains of the extra-cytoplasmic function sigma factor sigma(C) from Mycobacterium tuberculosis. The Journal of biological chemistry 282, 4711–8.
- Ulrich, L.E., Koonin, E. V, Zhulin, I.B.,** 2005. One-component systems dominate signal transduction in prokaryotes. Trends in microbiology 13, 52–6.
- Ulrich, L.E., Zhulin, I.B.,** 2010. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. Nucleic acids research 38, D401–7.
- Van Vliet, A.H.M.,** 2010. Next generation sequencing of microbial transcriptomes: challenges and opportunities. FEMS microbiology letters 302, 1–7.
- Vasicova, P., Abrhamova, Z., Nesvera, J, Patek, M., Sahm, H., Eikmanns, B.,** 1998. Integrative and autonomously replicating vectors for analysis of promoters in *Corynebacterium glutamicum*. BIOTECHNOLOGY TECHNIQUES 12, 743–746.
- Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., Van Sinderen, D.,** 2007. Genomics of *Actinobacteria* : tracing the evolutionary history of an ancient phylum. Microbiology and molecular biology reviews : MMBR 71, 495–548.
- Walker, G.C.,** 1996. The SOS response of *Escherichia coli*., in: Neidhardt, F., Curtis II, I.R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaecheter, M., Umberger, H. (Eds.), *Escherichia Coli and Salmonella: Cellular and Molecular Biology*. American Society for Microbiology, Washington, D.C.
- Weber, H., Polen, T., Heuveling, J., Wendisch, Volker F, Hengge, R.,** 2005. Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. Journal of bacteriology 187, 1591–603.
- Wurtzel, Omri, Sapra, R., Chen, F., Zhu, Yiwen, Simmons, B. a, Sorek, Rotem,** 2010. A single-base resolution map of an archaeal transcriptome. Genome research 20, 133–41.
- Yanisch-Perron, C., Vieira, J., Messing, J.,** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33, 103–19.
- Yoder-Himes, D.R., Chain, P.S.G., Zhu, Y, Wurtzel, O, Rubin, E.M., Tiedje, J.M., Sorek, R,** 2009. Mapping the Burkholderia cenocepacia niche response via high-throughput sequencing. Proceedings of the National Academy of Sciences of the United States of America 106, 3976–81.
- Yokota, T., Otori, T., Kodama, T.,** 1987. Purification and properties of haloalkane dehalogenase from *Corynebacterium sp.* strain m15-3. Journal of bacteriology 169, 4049–54.
- Young, B.A., Gruber, Tanja M, Gross, Carol A,** 2004. Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. Science (New York, N.Y.) 303, 1382–4.
- Yukawa, H., Omumasaba, C.A., Nonaka, H., Kos, P., Okai, N., Suzuki, N., Suda, M., Tsuge, Y., Watanabe, J., Ikeda, Y., Vertes, A.A., Inui, M.,** 2007. Comparative analysis of the *Corynebacterium glutamicum* group and complete genome sequence of strain R. Microbiology 153, 1042–1058.

- Zemanova, M., Kaderabkova, P., Patek, M., Knoppova, M., Silar, R, Nesvera, J,** 2008.
Chromosomally encoded small antisense RNA in *Corynebacterium glutamicum*. FEMS Microbiol Lett 279, 195–201.
- Zheng, D., Constantinidou, C., Hobman, J.L., Minchin, S.D.,** 2004. Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling 32, 5874–5893.

VII. Anhang

1. Publikationsliste

Teile dieser Arbeit sind in folgenden Publikationen veröffentlicht:

Busche, T., Silar, R., Pičmanová, M., Pátek, M., Kalinowski, J, 2012. Transcriptional regulation of the operon encoding stress-responsive ECF sigma factor SigH and its anti-sigma factor RshA, and control of its regulatory network in *Corynebacterium glutamicum*. BMC genomics **13**, 445.

Pátek, M., Holátko, J., Busche, T., Kalinowski, J., Nešvera, J., 2013. *Corynebacterium glutamicum* promoters: a practical approach. Microbial biotechnology (2):103-17

2. Ausführliche Darstellung differentiell transkribierter Gene aus den *Microarray*-Experimenten

Anhang Tab. 1 Differentiell transkribierte Gene in *C. glutamicum* RES167 Δ rshA im Vergleich zu *C. glutamicum* RES167 (Referenz) sortiert nach Änderung.

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
cg3405		NADPH:quinone reductase Zn-dependent oxidoreductase	2.9E-03	3.73	11.22	13.27
cg2538		Alkanal monooxygenase (FMN-linked)	3.0E-02	3.62	11.81	12.3
cg2838	<i>dsbA1</i>	Putative dithiol-disulfide isomerase	1.2E-03	3.55	11.02	11.71
cg3236	<i>msrA</i>	Protein-methionine-S-oxide reductase	9.4E-04	2.72	11.33	6.59
cg3407		Putative membrane protein	2.4E-03	2.68	11.33	6.41
cg2810		Putative secondary H ⁺ /Na ⁺ :glutamate/dicarboxylate symporter, dicarboxylate/amino acid:cation symporter (DAACS) family	8.3E-06	2.66	10.31	6.32
cg2106		Conserved hypothetical protein	1.2E-06	2.63	11.46	6.19
cg2194	<i>mtr</i>	Putative NADPH-dependent mycothiol reductase	2.1E-03	2.3	10.55	4.92
cg1709	<i>mshC</i>	Putative 1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside--L-cysteine ligase	8.3E-04	2.16	11.7	4.47
cg0617		Hypothetical protein	2.0E-03	2.07	10.77	4.2
cg1288		Putative multidrug efflux permease, MFS-type	4.0E-06	1.98	8.26	3.94
cg1432	<i>ivdD</i>	Dihydroxy-acid dehydratase	6.3E-04	1.94	12.65	3.84
cg1398		Conserved hypothetical protein	1.0E-03	1.92	9.91	3.78
cg1555	<i>uvrD3</i>	DNA/RNA helicase, superfamily I	5.7E-07	1.92	9.18	3.78
cg3299	<i>trxB1</i>	Thioredoxin (TRX)	2.6E-05	1.9	12.74	3.73
cg0614		Hypothetical protein	4.2E-04	1.89	9.98	3.71
cg0616	<i>fdhD</i>	Putative formate dehydrogenase, FdhD-family	2.9E-02	1.89	10.53	3.71
cg1397	<i>trmU</i>	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	3.3E-04	1.89	11.9	3.71
cg2078	<i>msrB</i>	Peptide methionine sulfoxide reductase	1.7E-04	1.85	12.39	3.61
cg3422	<i>trxB</i>	Thioredoxin reductase	4.3E-03	1.82	12.3	3.53
cg2423	<i>lipA</i>	Lipoyl synthetase	4.2E-03	1.74	13.62	3.34
cg3423	<i>trxC</i>	Thioredoxin	3.2E-03	1.57	12.96	2.97
cg0378		Putative phage-associated protein	3.2E-03	1.55	10.98	2.93
cg2411		Conserved hypothetical protein, HesB/YadR/YfhF family	8.2E-03	1.49	12.31	2.81
cg1799	<i>ribC</i>	Riboflavin synthase, alpha chain	3.5E-02	1.45	10.68	2.73
cg1688	<i>pafA</i>	Putative proteasome component	5.1E-03	1.44	11.8	2.71
cg2661	<i>dsbA2</i>	Thiol-disulfide isomerase	6.1E-03	1.42	12.95	2.68
cg3344		Putative nitroreductase	1.0E-02	1.41	9.28	2.66
cg1689	<i>pup</i>	prokaryotic ubiquitin-like protein	1.9E-04	1.31	12.04	2.48
cg1127	<i>mca</i>	Putative mycothiol S-conjugate amidase	4.9E-03	1.3	11.63	2.46
cg2247		Hypothetical protein	1.6E-03	1.27	11.79	2.41
cg1282		Conserved hypothetical protein	5.7E-03	1.23	11.67	2.35
cg2127		Hypothetical protein	1.9E-02	1.23	11.42	2.35

Fortsetzung Anhang Tab. 1

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<i>cg3424</i>	<i>cwlM</i>	N-Acetylmuramoyl-L-alanine amidase	2.5E-02	1.22	12.22	2.33
<i>cg2214</i>		Putative Fe-S-cluster redox enzyme	6.1E-03	1.18	10.94	2.27
<i>cg1560</i>	<i>uvrA</i>	Excinuclease ABC, ATPase subunit A	2.2E-03	1.16	10.88	2.23
<i>cg1798</i>	<i>ribA</i>	Putative GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphatesynthase	1.3E-04	1.15	11.05	2.22
<i>cg2835</i>		Putative acetyltransferase	1.6E-03	1.12	12.43	2.17
<i>cg1281</i>		ABC-type putative multidrug transporter, ATPase and permease subunit	2.2E-03	1.11	9.35	2.16
<i>cg1687</i>		Putative transcriptional regulatory protein	8.3E-03	1.09	10.42	2.13
<i>cg1797</i>	<i>ribH</i>	Riboflavin synthase, beta chain	1.1E-03	1.07	11.76	2.1
<i>cg0497</i>	<i>hemA</i>	Glutamyl-tRNA reductase	3.3E-02	1.05	11.27	2.07
<i>cg1779</i>	<i>opcA</i>	Glucose-6-phosphate 1-dehydrogenase subunit	1.9E-06	1.04	11.78	2.06
<i>cg2762</i>	<i>murI</i>	Glutamate racemase	3.2E-03	1.02	10.24	2.03
<i>cg3078</i>		Hypothetical protein	9.1E-04	0.98	10.97	1.97
<i>cg0184</i>		Conserved hypothetical protein	3.2E-07	0.96	10.23	1.95
<i>cg1411</i>	<i>rhsA</i>	ABC-type ribose transporter, ATPase subunit (TC 3.A.1.2.1)	5.1E-03	0.94	9.33	1.92
<i>cg0185</i>		Putative glyoxalase	8.2E-06	0.93	11.45	1.91
<i>cg0186</i>		Putative methylated-DNA--protein-cysteine methyltransferase	1.5E-06	0.93	10.97	1.91
<i>cg2636</i>	<i>catA</i>	Catechol 1,2-dioxygenase	4.4E-03	0.91	10.95	1.88
<i>cg1780</i>	<i>pgl</i>	6-Phosphogluconolactonase	6.4E-07	0.9	11.84	1.87
<i>cg1413</i>	<i>rhsB</i>	ABC-type ribose transporter, substrate-binding lipoprotein (TC 3.A.1.2.1)	6.5E-05	0.89	9.41	1.85
<i>cg2515</i>	<i>dnaJ2</i>	Chaperone, contain C-terminal Zn-finger domain	1.1E-03	0.89	11.73	1.85
<i>cg0498</i>	<i>hemC</i>	Porphobilinogen deaminase	3.5E-05	0.88	11.18	1.84
<i>cg2665</i>		Hypothetical protein	1.1E-03	0.86	10.47	1.82
<i>cg2181</i>		ABC-type putative dipeptide/oligopeptide transporter, substrate-binding lipoprotein	4.0E-03	0.85	14.75	1.8
<i>cg1128</i>		Hypothetical protein, similar to ribosomal protein S2	1.6E-04	0.84	11.86	1.79
<i>cg1139</i>		Allophanate hydrolase subunit 2	1.5E-03	0.81	9.47	1.75
<i>cg1708</i>		Conserved hypothetical protein	7.1E-03	0.81	11.92	1.75
<i>cg1765</i>	<i>sufR</i>	Transcriptional repressor of suf operon	8.6E-03	0.81	14.54	1.75
<i>cg2560</i>	<i>aceA</i>	Isocitrate lyase	3.0E-07	0.78	12.42	1.72
<i>cg2183</i>		ABC-type putative dipeptide/oligopeptide transporter, permease subunit	7.3E-03	0.77	13.27	1.71
<i>cg2434</i>		Putative monooxygenase, luciferase	4.5E-02	0.75	9.64	1.68
<i>cg2644</i>	<i>clpP2</i>	Endopeptidase Clp, proteolytic subunit	2.6E-04	0.75	13.02	1.68
<i>cg0380</i>		Hypothetical protein	6.7E-04	0.74	9.32	1.67
<i>cg0228</i>		Two-component system, sensory histidine kinase, putative pseudogene	4.0E-07	0.72	10.08	1.65
<i>cg1412</i>	<i>rhsC</i>	ABC-type ribose transporter, permease subunit (TC 3.A.1.2.1)	3.8E-07	0.72	9.31	1.65
<i>cg1778</i>	<i>zwf</i>	Glucose-6-phosphate 1-dehydrogenase	7.8E-04	0.72	11.73	1.65
<i>cg0043</i>		ABC-type putative manganese/zinc transporter, ATPase subunit	9.4E-08	0.72	7.15	1.65
<i>cg1482</i>		Putative Zn-dependent hydrolase	6.7E-04	0.69	8.96	1.61

Fortsetzung Anhang Tab. 1

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<i>cg1686</i>		Putative transcriptional regulatory protein	2.7E-04	0.69	10.53	1.61
<i>cg1553</i>	<i>qor2</i>	quinone oxidoreductase involve in disulfide stress response	5.6E-05	0.68	9.93	1.6
<i>cg2514</i>		Conserved hypothetical protein	1.8E-04	0.67	12.54	1.59
<i>cg1375</i>		Putative thioredoxin	3.1E-03	0.66	12.25	1.58
<i>cg2206</i>	<i>ispG</i>	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	5.3E-05	0.66	12.04	1.58
<i>cg2546</i>		Putative secondary C4-dicarboxylate transporter, tripartite ATP-independent transporter (TRAP-T) family	1.6E-05	0.66	8.87	1.58
<i>cg2645</i>	<i>clpP1</i>	Endopeptidase Clp, proteolytic subunit	3.9E-04	0.65	13.02	1.57
<i>cg3079</i>	<i>clpB</i>	Putative ATP-dependent protease (heat shock protein)	3.0E-02	0.65	11.57	1.57
<i>cg0699</i>	<i>guaB2</i>	IMP dehydrogenase	2.6E-06	0.64	12.29	1.56
<i>cg2184</i>		ABC-type putative dipeptide/oligopeptide transporter, ATPase subunit	2.2E-05	0.64	13.2	1.56
<i>cg3077</i>		Putative membrane protein	1.1E-03	0.64	9.03	1.56
<i>cg0998</i>		Trypsin-like serine protease	1.3E-03	0.62	9.9	1.54
<i>cg1410</i>	<i>rhsR</i>	Transcriptional repressor of the ribose importer RhsACBD, LacI-family	1.0E-04	0.62	8.62	1.54
<i>cg1795</i>	<i>uvrC</i>	Excinuclease subunit C	9.1E-05	0.61	9.35	1.53
<i>cg1464</i>		Putative transcriptional regulator, HTH_3-family	4.3E-06	0.6	9.38	1.52
<i>cg0768</i>		ABC-type putative iron-siderophore transporter, ATPase subunit	9.2E-03	-0.6	9.87	0.66
<i>cg0411</i>		Putative membrane protein	3.8E-03	-0.61	8.36	0.66
<i>cg0811</i>	<i>dtsR2</i>	Acetyl/propionyl-CoA carboxylase, beta chain	1.1E-03	-0.65	14.3	0.64
<i>cg0814</i>	<i>birA</i>	Bifunctional biotin--[acetyl-CoA-carboxylase] synthetase/biotin operon repressor	1.7E-05	-0.65	11.43	0.64
<i>cg0839</i>		Hypothetical protein	4.4E-07	-0.65	9.44	0.64
<i>cg2069</i>	<i>psp1</i>	Putative secreted protein	5.1E-03	-0.65	8.74	0.64
<i>cg0527</i>	<i>glyR</i>	Transcriptional activator of glyA, ArsR-family	5.2E-04	-0.66	9.83	0.63
<i>cg0812</i>	<i>dtsR1</i>	Acetyl/propionyl-CoA carboxylase, beta chain	1.5E-03	-0.66	13.75	0.63
<i>cg0816</i>	<i>purK</i>	Phosphoribosylaminoimidazole carboxylase, ATPase subunit	2.5E-06	-0.66	11.64	0.63
<i>cg0840</i>		Conserved hypothetical protein	3.7E-02	-0.67	10.72	0.63
<i>cg0076</i>		Hypothetical protein	1.7E-03	-0.68	7.77	0.62
<i>cg0833</i>		Conserved hypothetical protein	4.9E-02	-0.69	10.26	0.62
<i>cg0921</i>		Siderophore-interacting protein	3.4E-03	-0.69	10.11	0.62
<i>cg2657</i>		Putative membrane protein, putative pseudogen	9.0E-04	-0.69	12.37	0.62
<i>cg3127</i>	<i>tctC</i>	citrate uptake transporter, substrate binding protein	1.2E-02	-0.69	8.6	0.62
<i>cg0825</i>		Putative short chain dehydrogenase related to 3-oxoacyl-(acyl-carrier protein) reductase	5.8E-05	-0.7	8.39	0.62
<i>cg1668</i>		Putative membrane protein	2.4E-03	-0.7	9.23	0.62
<i>cg0843</i>		Putative helicase	3.6E-04	-0.71	8.88	0.61
<i>cg0826</i>		Putative membrane protein	1.3E-05	-0.72	11.22	0.61

Fortsetzung Anhang Tab. 1

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<i>cg1002</i>		Conserved putative secreted protein	3.4E-03	-0.72	7.68	0.61
<i>cg3129</i>		ABC-type transporter, ATPase subunit	4.1E-08	-0.73	8.75	0.6
<i>cg0838</i>		Putative helicase	3.8E-03	-0.77	9.33	0.59
<i>cg0819</i>		Hypothetical protein	6.9E-05	-0.78	9.29	0.58
<i>cg0820</i>	<i>purE</i>	Phosphoribosylaminoimidazole carboxylase, catalytic subunit	9.7E-03	-0.79	10.93	0.58
<i>cg0832</i>		ABC-type putative sugar transporter, permease subunit	3.1E-02	-0.8	11.16	0.57
<i>cg0828</i>		Putative dihydrofolate reductase	1.8E-02	-0.81	9.29	0.57
<i>cg0842</i>		Putative DNA helicase	2.1E-04	-0.83	9.98	0.56
<i>cg0470</i>		Conserved secreted protein	2.3E-05	-0.84	8.29	0.56
<i>cg0810</i>		Conserved hypothetical protein	6.5E-06	-0.86	13.6	0.55
<i>cg3139</i>		Conserved hypothetical protein	2.2E-06	-0.86	11.76	0.55
<i>cg3140</i>	<i>tagA1</i>	DNA-3-methyladenine glycosylase I	1.5E-04	-0.9	11.09	0.54
<i>cg0817</i>	<i>kup</i>	Putative secondary K ⁺ uptake permease, K ⁺ uptake permease (KUP) family	2.5E-04	-1.01	8.18	0.5
<i>cg0821</i>		Conserved hypothetical protein	8.5E-06	-1.02	11.4	0.49
<i>cg3138</i>		Putative membrane protease subunit, stomatin/prohibitin homolog	1.9E-06	-1.02	12.92	0.49
<i>cg0876</i>	<i>sigH</i>	RNA polymerase sigma factor, ECF-family	1.3E-05	-1.03	10.43	0.49
<i>cg0830</i>		Putative membrane protein	1.4E-05	-1.18	10.12	0.44
<i>cg3356</i>		Putative secondary H ⁺ /Na ⁺ :glutamate/dicarboxylate symporter, dicarboxylate/amino acid:cation symporter (DAACS) family	1.4E-04	-2.18	11.9	0.22
<i>cg1212</i>		Putative antibiotic efflux permease, MFS-type	2.9E-05	-2.66	8.95	0.16

Komparative *Microarray*-Analyses mit Gesamt-RNA von *C. glutamicum* RES167 und dem Deletionsderivat $\Delta rshA$, unter Standardbedingungen (30 °C) im Luftschüttler bis zur exponentiellen Phase angezogen. Gene mit einem m-Wert > 1 bzw. m < -1 sind signifikant in ihrem mRNA Pool im Vergleich zur Referenz verändert. Die differentiell transkribierten Gene sind in dieser Tabelle aufgelistet. a) Änderung, *signal intensity ratio* definiert durch $2^{(m\text{-Wert})}$.

Anhang Tab. 2 Differentiell transkribierte Gene in *C. glutamicum* RES167 Δ *cseE* im Vergleich zu *C. glutamicum* RES167 (Referenz) sortiert nach Änderung.

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<u>cg1277²</u>		Conserved putative membrane protein	1.7E-07	2.42	11.81	5.35
<i>cg1412</i>	<i>rbsC</i>	ABC-type ribose transporter, permease subunit (TC 3.A.1.2.1)	8.6E-08	2.35	11.67	5.10
<u>cg3078⁵</u>		Hypothetical protein	1.2E-07	2.22	11.05	4.66
<i>cg1413</i>	<i>rbsB</i>	ABC-type ribose transporter, substrate-binding lipoprotein (TC 3.A.1.2.1)	8.7E-07	2.21	11.22	4.63
<u>cg1273[*]</u>		Putative twin arginine targeting (Tat) Preprotein translocase subunit	1.4E-06	2.13	12.22	4.38
<u>cg3079⁵</u>	<i>clpB</i>	Putative ATP-dependent protease (heat shock protein)	1.1E-06	2.05	11.69	4.14
<i>cg3077</i>		Putative membrane protein	5.8E-06	2.04	9.1	4.11
<u>cg3098⁶</u>	<i>dnaJ</i>	Chaperone DnaJ, heat shock protein	4.1E-07	1.99	11.01	3.97
<i>cg0223</i>		Putative sugar/metabolite permease, MFS-type	6.8E-03	1.99	9.09	3.97
<i>cg1411</i>	<i>rbsA</i>	ABC-type ribose transporter, ATPase subunit (TC 3.A.1.2.1)	6.0E-08	1.99	11.7	3.97
<i>cg2845</i>	<i>pstC</i>	ABC-type putative phosphate transporter, permease subunit (TC 3.A.1.7.1)	9.2E-05	1.92	9.45	3.78
<i>cg3429</i>		Putative 60 KD inner membrane protein, SpoIIJ homolog, cytochrome oxidase biogenesis (Oxa1) family	4.7E-07	1.85	13.08	3.61
<u>cg3099⁶</u>	<i>grpE</i>	Chaperone GrpE, heat shock protein	2.2E-07	1.84	12.15	3.58
<u>cg2657</u>		Putative membrane protein, putative pseudogen	9.0E-09	1.82	14.19	3.53
<i>cg3391</i>	<i>idhA1</i>	Inositol 2-dehydrogenase	3.5E-02	1.82	8.26	3.53
<u>cg1640³</u>		Putative membrane protein, containing a CBS domain	2.2E-07	1.78	11.06	3.43
<i>cg0785</i>		Hypothetical protein	4.2E-05	1.75	10.08	3.36
<i>cg2514</i>		Conserved hypothetical protein	6.1E-08	1.74	13.44	3.34
<u>cg0380¹</u>		Hypothetical protein	2.5E-05	1.73	10.97	3.32
<u>cg1278²</u>		Conserved putative secreted protein	2.1E-06	1.71	12.27	3.27
<i>cg3195</i>		Putative flavin-containing monooxygenase	4.9E-04	1.71	12.56	3.27
<u>cg3309⁷</u>		Putative secreted protein	4.2E-05	1.7	11.85	3.25
<i>cg2610</i>		ABC-type putative dipeptide/oligopeptide transporter, substrate-binding lipoprotein	1.3E-03	1.69	11.54	3.23
<u>cg0378¹</u>		Putative phage-associated protein	2.2E-05	1.67	12.37	3.18
<i>cg0825</i>		Putative short chain dehydrogenase related to 3-oxoacyl-(acyl-carrier protein) reductase	2.5E-05	1.67	9.42	3.18
<u>cg3344</u>		Putative nitroreductase	8.2E-05	1.66	9.58	3.16
<i>cg2846</i>	<i>pstS</i>	ABC-type putative phosphate transporter, substrate-binding lipoprotein (TC 3.A.1.7.1)	3.2E-06	1.64	10.39	3.12
<i>cg1227</i>		Putative membrane protein	6.7E-03	1.61	8.86	3.05
<i>cg3117</i>	<i>cysX</i>	Ferredoxin-like protein, involved in electron-transfer	1.7E-06	1.61	10.33	3.05
<u>cg1639³</u>		Putative membrane protein, containing a CBS domain	4.2E-07	1.6	10.84	3.03
<u>cg2515</u>	<i>dnaJ2</i>	Chaperone, contain C-terminal Zn-finger domain	3.3E-06	1.59	11.62	3.01
<i>cg3118</i>	<i>cysI</i>	Ferredoxin-sulfite reductase	6.2E-05	1.51	10.19	2.85
<i>cg3395</i>	<i>proP</i>	Putative proline/betaine permease, MFS-type	1.4E-02	1.51	9.29	2.85
<i>cg2513</i>	<i>phoH2</i>	Phosphate starvation-inducible protein, PhoH-like	1.2E-05	1.5	11.7	2.83
<u>cg2983⁴</u>	<i>folE</i>	GTP cyclohydrolase I	5.1E-08	1.49	12.85	2.81
<i>cg0201</i>	<i>iolB</i>	Uncharacterized enzyme involved in inositol metabolism	1.2E-04	1.48	10.58	2.79

Fortsetzung Anhang Tab. 2

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
cg2380		Putative membrane protein	5.7E-06	1.46	14.05	2.75
cg2982⁴	folP1	Dihydropteroate synthase	2.0E-07	1.46	12.55	2.75
cg3226		Putative MFS-type L-lactate permease	2.5E-02	1.46	13.32	2.75
cg2981⁴	folX	Dihydroneopterin aldolase	1.6E-08	1.43	12.83	2.69
cg0678		Conserved hypothetical protein	9.3E-03	1.43	8.86	2.69
cg0204	<i>iolG</i>	Myo-inositol 2-dehydrogenase	7.3E-04	1.4	11.42	2.64
cg0767		Siderophore-interacting protein	5.2E-03	1.38	8.55	2.60
cg2979⁴	folK	2-Amino-4-hydroxy-6-hydroxymethylidihydropteridine diphosphokinase	1.2E-05	1.37	11.77	2.58
cg3100⁶	dnaK	Chaperone DnaK, heat shock protein	8.3E-09	1.37	13.82	2.58
cg3097⁶	hspR	Putative transcriptional regulator, MerR-family	2.9E-04	1.36	10.91	2.57
cg0199	<i>iolA</i>	Methylmalonate-semialdehyde dehydrogenase	5.3E-03	1.32	9.49	2.50
cg2152	clgR	Transcriptional activator of Clp protease genes	2.0E-04	1.31	12.2	2.48
cg0205	<i>iolH</i>	Myo-inositol catabolism protein	4.0E-03	1.31	12.05	2.48
cg1361	<i>atpI</i>	Hypothetical protein	4.2E-02	1.31	8.44	2.48
cg1545	<i>uriT</i>	Putative multidrug efflux permease, MFS-type	2.6E-06	1.31	11.7	2.48
cg0197	<i>iolC</i>	Putative 5-dehydro-2-deoxygluconokinase	2.1E-02	1.3	9.35	2.46
cg0202	<i>iolD</i>	Putative acetolactate synthase, large subunit	2.7E-03	1.27	10.95	2.41
cg0784		Putative cell wall-associated hydrolase	4.5E-02	1.27	9.11	2.41
cg2978⁴		Putative membrane protein	6.3E-07	1.25	11.75	2.38
cg1476	<i>thiC</i>	Thiamine biosynthesis protein ThiC	3.3E-04	1.24	9.87	2.36
cg2267		Putative membrane protein	7.6E-06	1.24	11.04	2.36
cg2512		Putative metal-dependent hydrolase	7.3E-06	1.24	11.53	2.36
cg3253	<i>mcbR</i>	Global transcriptional repressor of sulfur metabolism, TetR-family	1.1E-05	1.23	10.51	2.35
cg2984⁴	ftsH	Cell-division protein (ATP-dependent ZN metalloproteinase)	4.9E-04	1.22	13.35	2.33
cg0345		Putative metal-dependent amidohydrolase of the TIM-barrel fold	1.6E-02	1.21	8.73	2.31
cg2843	<i>pstB</i>	ABC-type putative phosphate transporter, ATPase subunit (TC 3.A.1.7.1)	2.1E-04	1.19	10.66	2.28
cg1546	<i>rbsK1</i>	Putative ribokinase	1.1E-03	1.18	10.35	2.27
cg1478		Hypothetical protein	3.3E-02	1.15	8.67	2.22
cg0203	<i>iolE</i>	Putative myo-inosose-2 dehydratase	1.6E-03	1.13	11.63	2.19
cg2938		ABC-type putative dipeptide/oligopeptide transporter, permease subunit	8.7E-04	1.13	10.25	2.19
cg2441		Permease of the major facilitator superfamily, N-terminal region - putative pseudogene	5.2E-06	1.12	11.7	2.17
cg3306⁷	rplI	50S ribosomal protein L9	4.7E-05	1.11	14.02	2.16
cg1410	<i>rbsR</i>	Transcriptional repressor of the ribose importer RbsACBD, LacI-family	6.0E-05	1.11	10.38	2.16
cg2511		Putative membrane protein containing CBS domain	2.7E-07	1.1	11.55	2.14
cg1506		Putative membrane protein	5.5E-05	1.09	10.09	2.13
cg2836	<i>sucD</i>	Succinate--CoA ligase (ADP-forming), alpha subunit	2.2E-02	1.08	9.03	2.11
cg0198		Conserved hypothetical protein, probably involved in myo-inositol metabolism	8.9E-03	1.03	10.07	2.04
cg0652	<i>rpsM</i>	30S ribosomal protein S13	8.7E-06	1.03	13.36	2.04
cg0755	<i>metY</i>	O-Acetylhomoserine sulfhydrylase	7.2E-04	1.03	11.27	2.04
cg3119	<i>fpr2</i>	Ferredoxin--NADP(+) reductase	4.6E-04	1.03	10.68	2.04

Fortsetzung Anhang Tab. 2

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
cg2102	sigB	RNA polymerase sigma factor rpoD (Sigma-A).	1.4E-04	1.02	14.24	2.03
cg1543	uriH	Uridine preferring nucleoside hydrolase	8.2E-06	1.02	13.27	2.03
cg2939		ABC-type putative dipeptide/oligopeptide transporter, ATPase subunit	1.9E-03	1.02	11.31	2.03
cg3024	mrpA	Putative secondary Na ⁺ /H ⁺ antiporter, monovalent cation:proton antiporter-3 (CPA3) family	1.3E-03	1.02	10.09	2.03
cg2636	catA	Catechol 1,2-dioxygenase	1.0E-04	1.01	10.6	2.01
cg2837	sucC	Succinate--CoA ligase (ADP-forming), beta subunit	1.1E-02	1.01	10.7	2.01
cg3390		Putative sugar phosphate isomerase/epimerase	1.4E-02	1.01	9.04	2.01
cg0986	amtR	Transcriptional repressor of nitrogen metabolism, TetR-family	3.8E-06	1	10.8	1.99
cg1542		Putative membrane protein	1.1E-02	1	8.69	2.00
cg2118		Transcriptional regulator protein, DeoR-family	3.7E-02	-1	9.78	0.50
cg3212		Putative carboxymuconolactone decarboxylase subunit	1.2E-02	-1	12.33	0.50
cg3375		Predicted nucleoside-diphosphate-sugar epimerase	6.1E-03	-1	8.79	0.50
cg1626		Conserved hypothetical protein	3.9E-04	-1.02	10.5	0.49
cg0088	citP	Putative secondary Mg ²⁺ /H ⁺ :citrate transporter, CitMHS-family	2.9E-04	-1.04	10.26	0.49
cg1300	cydB	Cytochrome d ubiquinol oxidase subunit II	5.4E-05	-1.04	10.67	0.49
cg1906		Hypothetical protein	3.6E-08	-1.05	11.38	0.48
cg1911		Putative secreted protein	1.7E-03	-1.05	13.51	0.48
cg1907		Putative phosphopantothenoylcysteine synthetase/decarboxylase	1.3E-05	-1.1	11.8	0.47
cg2797		Conserved hypothetical protein	1.9E-02	-1.1	8.92	0.47
cg2165		Putative secreted protein	4.9E-04	-1.11	9.71	0.46
cg3215	glpQ1	Glycerophosphodiester phosphodiesterase	7.6E-03	-1.13	9.37	0.46
cg0105		Hypothetical protein	1.7E-03	-1.15	8.9	0.45
cg0922		ABC-type putative iron-siderophore transporter, substrate-binding lipoprotein	9.1E-03	-1.15	9.65	0.45
cg3000		Putative thiosulfate sulfurtransferase	3.3E-02	-1.15	12.13	0.45
cg0798	prpC1	Putative (methyl)citrate synthase	1.9E-03	-1.16	9.79	0.45
cg0897	pdxR	Pyridoxine biosynthesis transcriptional regulator, GntR-family/valine--pyruvate transaminase, AT class I	5.9E-03	-1.16	9.41	0.45
cg1905		Hypothetical protein	6.4E-07	-1.17	12.15	0.44
cg2525		Putative membrane protein	1.6E-05	-1.17	11.71	0.44
cg1373		Putative glyoxalase	9.3E-05	-1.19	11.22	0.44
cg1090	ggtB	Putative gamma-glutamyltranspeptidase	1.5E-02	-1.2	10.06	0.44
cg1606	pyrG	CTP synthetase	9.9E-08	-1.2	11.53	0.44
cg2139	gluD	ABC-type glutamate transporter, permease subunit (TC 3.A.1.3.9)	5.5E-04	-1.2	10.65	0.44
cg0998		Trypsin-like serine protease	7.8E-04	-1.21	10.25	0.43
cg3354		3-hydroxybenzoate 6-monooxygenase	1.6E-03	-1.21	9.5	0.43
cg2782	ftn	Ferritin-like protein	2.1E-03	-1.22	13.74	0.43
cg1835	aroE3	Shikimate 5-dehydrogenase	1.4E-06	-1.23	13.16	0.43
cg2434		Putative monooxygenase, luciferase	9.9E-06	-1.23	10.44	0.43
cg0044		ABC-type putative sugar transporter, substrate-binding lipoprotein	8.1E-03	-1.24	11.08	0.42
cg0390		Putative multidrug efflux permease, MFS-type	9.3E-03	-1.24	9.37	0.42
cg1167		Putative acetyltransferase, GNAT-family	2.9E-05	-1.25	10.51	0.42

Fortsetzung Anhang Tab. 2

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<i>cg1944</i>		Hypothetical protein	3.8E-07	-1.26	11.76	0.42
	<i>tet</i>	Tetracycline resistance protein	6.9E-04	-1.26	8.16	0.42
<i>cg2986</i>	<i>mesJ</i>	Conserved hypothetical protein, contains ATPase of the PP-loop superfamily	9.6E-05	-1.38	11.86	0.38
<i>cg0958</i>		Putative secreted protein	9.4E-05	-1.42	10.4	0.37
<i>cg2138</i>	<i>gluC</i>	ABC-type glutamate transporter, permease subunit (TC 3.A.1.3.9)	1.9E-04	-1.45	12.44	0.37
<i>cg3366</i>	<i>rmpA</i>	Phosphotransferase system (PTS), putative ribitol/mannitol-specific enzyme IIB component	3.3E-05	-1.45	11.31	0.37
<i>cg1287</i>		Conserved hypothetical protein	8.3E-05	-1.48	8.87	0.36
<i>cg2917</i>		Conserved hypothetical protein	5.2E-06	-1.49	9.06	0.36
<i>cg2136</i>	<i>gluA</i>	ABC-type glutamate transporter, ATPase subunit (TC 3.A.1.3.9)	8.3E-07	-1.55	12.44	0.34
<i>cg2137</i>	<i>gluB</i>	ABC-type glutamate transporter, substrate-binding lipoprotein (TC 3.A.1.3.9)	8.6E-05	-1.55	13.13	0.34
<i>cg3323</i>	<i>ino1</i>	Inositol-3-phosphate synthase	1.7E-02	-1.55	11.52	0.34
<i>cg0043</i>		ABC-type putative manganese/zinc transporter, ATPase subunit	3.4E-04	-1.64	9.46	0.32
<i>cg0566</i>	<i>gabT</i>	Putative 4-aminobutyrate aminotransferase, AT class II	6.8E-06	-1.64	10.91	0.32
<i>cg0953</i>	<i>mctC</i>	secondary acetate/propionate/pyruvate transporter	1.6E-04	-1.69	13.49	0.31
<i>cg0806</i>		Conserved hypothetical protein	4.4E-03	-1.71	11.64	0.31
<i>cg1479</i>	<i>glgP1</i>	Putative glycogen phosphorylase	8.5E-08	-1.85	12.82	0.28
<i>cg1526</i>		Putative multidrug efflux permease, MFS-type	3.1E-04	-1.9	8.87	0.27
<i>cg0899</i>	<i>pdxT</i>	pyridoxal 5'-phosphate (PLP) synthase subunit, glutamine amidotransferase	9.6E-07	-1.91	14.36	0.27
<i>cg3303</i>		Putative transcriptional regulator, PadR-family	3.1E-02	-2.01	9.4	0.25
<i>cg1492</i>	<i>gpsA</i>	Glycerol-3-phosphate dehydrogenase (NAD(P)(+))	1.5E-06	-2.06	11.9	0.24
<i>cg3125</i>	<i>tctA</i>	citrate uptake transporter, membrane subunit	2.3E-04	-2.06	9.36	0.24
<i>cg0952</i>		Putative integral membrane protein	4.5E-05	-2.07	13.79	0.24
<i>cg1077</i>		Putative permease, MFS-type	3.4E-04	-2.08	9.64	0.24
<i>cg3127</i>	<i>tctC</i>	citrate uptake transporter, substrate binding protein	8.6E-06	-2.08	9.4	0.24
<i>cg0230</i>	<i>gltD</i>	Glutamate synthase (NADPH), small chain	3.8E-07	-2.09	13.66	0.23
<i>cg2054</i>		Putative membrane protein	9.5E-05	-2.16	10.43	0.22
<i>cg3365</i>	<i>rmpC</i>	Phosphotransferase system (PTS), putative ribitol/mannitol-specific enzyme IIC component	2.1E-07	-2.16	10.67	0.22
<i>cg2258</i>	<i>glnD</i>	Putative protein PII uridylyltransferase	1.4E-08	-2.24	10.94	0.21
<i>cg1256</i>	<i>dapD</i>	2,3,4,5-Tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	8.1E-06	-2.37	11.62	0.19
<i>cg1785</i>	<i>amt</i>	Putative secondary ammonium transporter, Amt-family	6.0E-09	-2.49	14.08	0.18
<i>cg2429</i>	<i>glnA</i>	Glutamate--ammonia ligase	1.0E-07	-2.59	14.23	0.17
<i>cg0567</i>	<i>gabD2</i>	Putative succinate-semialdehyde dehydrogenase (NAD(P)(+))	9.6E-06	-2.74	10.43	0.15
<i>cg3126</i>	<i>tctB</i>	citrate uptake transporter, membrane subunit	2.4E-05	-2.84	9.09	0.14
<i>cg0568</i>		Putative secondary aromatic amino acid permease, hydroxy/aromatic amino acid permease (HAAAP) family	4.9E-06	-2.91	10.33	0.13
<i>cg1295</i>		Putative hydrolase/acyltransferase, alpha/beta hydrolase superfamily	2.6E-07	-2.98	10.7	0.13
<i>cg2564</i>		conserved hypothetical protein	5.4E-08	-2.99	13.73	0.13
<i>cg1091</i>		Hypothetical protein	1.5E-09	-3.44	12.68	0.09
<i>cg2260</i>	<i>glnK</i>	Nitrogen regulatory protein PII	1.3E-09	-3.5	13.64	0.09

Fortsetzung Anhang Tab. 2

CDS	Gen-name	Vorhergesagte Funktion	p-Wert	m-Wert	a-Wert	Änderung ^a
<i>cg1066</i>	<i>urtE</i>	ABC-type putative branched-chain amino acid transporter, ATPase subunit	1.9E-09	-3.59	13.1	0.08
<i>cg0898</i>	<i>pdxS</i>	pyridoxal 5'-phosphate (PLP) synthase subunit	1.1E-06	-3.75	13.58	0.07
<i>cg0104</i>	<i>codA</i>	cytosine deaminase	6.7E-08	-3.85	11.64	0.07
<i>cg0118</i>	<i>ureG</i>	Urease accessory protein UreG	1.2E-07	-4.11	12.09	0.06
<i>cg0117</i>	<i>ureF</i>	Urease accessory protein UreF	3.8E-08	-4.13	11.74	0.06
<i>cg0119</i>	<i>ureD</i>	Urease accessory protein UreD	1.3E-07	-4.16	10.83	0.06
<i>cg0114</i>	<i>ureB</i>	Urease beta subunit	4.6E-09	-4.21	11.62	0.05
<i>cg0115</i>	<i>ureC</i>	Urease alpha subunit	5.0E-11	-4.37	11.66	0.05
<i>cg1212</i>		Putative antibiotic efflux permease, MFS-type	3.4E-07	-4.46	10.18	0.05
<i>cg0113</i>	<i>ureA</i>	Urease gamma subunit	4.9E-08	-4.57	11.58	0.04
<i>cg1781</i>	<i>'soxA</i>	Putative oxidase, pseudogene (C-terminal fragment)	4.3E-09	-4.62	9.62	0.04
<i>cg0116</i>	<i>ureE</i>	Urease accessory protein UreE	2.9E-08	-4.71	12.16	0.04
<i>cg1296</i>		Conserved hypothetical protein, putative non-ribosomal peptide synthetase module	1.4E-08	-4.84	9.59	0.03
<i>cg1784</i>	<i>ocd</i>	Putative ornithine cyclodeaminase	7.5E-10	-4.89	13.02	0.03
<i>cg1061</i>	<i>urtA</i>	ABC-type putative branched-chain amino acid transporter, substrate-binding lipoprotein	9.6E-10	-5.06	11.18	0.03
<i>cg1065</i>	<i>urtD</i>	ABC-type putative branched-chain amino acid transporter, ATPase subunit	6.2E-07	-5.15	11.63	0.03
<i>cg1064</i>	<i>urtC</i>	ABC-type putative branched-chain amino acid transporter, permease subunit	2.1E-09	-5.61	12.34	0.02
<i>cg1062</i>	<i>urtB</i>	ABC-type putative branched-chain amino acid transporter, permease subunit	1.2E-08	-5.77	11.26	0.02
<i>cg0103</i>	<i>cmT</i>	Putative drug efflux permease, MFS-type	1.4E-09	-5.79	11.28	0.02
<i>cg2261</i>	<i>amtB</i>	Putative secondary ammonium transporter, Amt-family	2.3E-09	-6.13	12.26	0.01
<i>cg0229</i>	<i>gltB</i>	Glutamate synthase (NADPH), large chain	1.6E-08	-6.18	11.54	0.01
<i>cg1783</i>	<i>soxA'</i>	Putative oxidase, pseudogene (N-terminal fragment)	2.7E-09	-6.38	12.26	0.01

**cg1273* ist mit dem deletierten *cseE* (*cg1272*) in einer transkriptionellen Einheit organisiert.

Gene die einen exklusiv SigE- bzw. SigH/SigH-abhängigen Promotor stromaufwärts besitzen sind fett geschrieben.

Gene (nur SigH/SigE-Regulon) eines Operons sind mit der gleichen hochgestellten Zahl gekennzeichnet.

Das erste Gen in einem Operon bzw. einzeln transkribierte Gene sind unterstrichen.

Gene des AmtR-Regulons sind rot hervorgehoben.

Komparative *Microarray*-Analysen mit Gesamt-RNA von *C. glutamicum* RES167 und dem Deletionsderivat $\Delta cseE$, unter Standardbedingungen (30 °C) im Luftschüttler bis zur exponentiellen Phase angezogen. Gene mit einem m-Wert > 1 bzw. m < -1 sind signifikant in ihrem mRNA Pool im Vergleich zur Referenz verändert. Die differentiell transkribierten Gene sind in dieser Tabelle aufgelistet. a) Änderung, *signal intensity ratio* definiert durch $2^{(m\text{-Wert})}$.

3. Zusammenfassung der RNA-Seq- (*Top-Down*) und ROSE-Ergebnisse (*Bottom-Up*)

Anhang Tab. 3 Partiiell SigH-abhängige Promotoren

Identifikation/Nachweis der Promotoren											
RNA-Seq <i>in vivo</i> Datensätze ^a	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$	Referenz ^d	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
+	+	+		+	(Halgasova et al., 2001)	<i>cg2102</i>	<i>sigB</i>	CTTGGGAACCTTTTTGTGGAAGCAGTCCGTTGAACCTCTTG	1990877	25	RNA polymerase sigma factor rpoD (Sigma-A).
+	+	+	+	+	(Ehira et al., 2009)	<i>cg3079</i>	<i>clpB</i>	GAGTGGAACATACTCAACTCTTTGTGCGTTATAGATATTA	2939558	91	Putative ATP-dependent protease (heat shock protein)
+	+	+		+	(Ehira et al., 2009)	<i>cg3100</i>	<i>dnaK</i>	AGTTGGAACAACCTTTGTGGCATTACCCTTGCTATATATG	2959823	120	Chaperone DnaK, heat shock protein
+	+	+	+	+	diese Arbeit (RNAseq)	<i>cg3344</i>	-	TGCGGGAACAATTTGAAGTTTTTTCAGTTGCTATAGGTA	3191301	0	Putative nitroreductase
+	+	+		+	diese Arbeit (ROSE)	<i>cg1272</i>	<i>cseE</i>	ACAGGGAACCTTATCACAGGCGACATCCGTTTTGAGTAGTA	1167979	65	Conserved hypothetical protein
+	+	+		+	diese Arbeit (ROSE)	<i>cg1277</i>	-	GTCTACGGGAACCACTTTCACCAAGCGTGTACGTTGTAATG	1172127	34	Conserved putative membrane protein
+	+	+		+	diese Arbeit (ROSE)	<i>cg1640</i>	-	AAGCGGGAACAAAAAATACCATAATCGTTCTGATTCCA	1529558	178	Putative membrane protein, containing a CBS domain
+	+	+			diese Arbeit (ROSE)	<i>cg2092</i>	<i>sigA</i>	CTGGGGAACAATAACGTTTCCCTATTGTTGATATAGGTA	1981787	136	RNA polymerase sigma factor rpoD (Sigma-A).
+	+	+		+	diese Arbeit (ROSE)	<i>cg2657</i>	-	AACAGGGAACCTTTAACGAAAGACTTGCCTTACTTTAGTA	2534617	22	Putative membrane protein, putative pseudogen
+	+	+		+	diese Arbeit (ROSE)	<i>cg2984</i>	<i>ftsH</i>	AGAGGGAACGTTTTGGCCCTCAACTTCGTTGATTAGTTG	2843112	77	Cell-division protein (ATP-dependent ZN metalloproteinase)
+	+	+	+		(Ehira et al., 2009)	<i>cg2661</i>	<i>dsbA2</i>	GAGGAACATAAAAACTCATCACCGTTGTTGAGATCAAGTG	2538097	190	thiol-disulfide isomerase
+	+	+		+	diese Arbeit (ROSE)	<i>cg2380</i>	-	CGGGGAACATTCGGATATTCTGGCGTTAATAAGGCATA	2268612	41	Putative membrane protein
+	+	+	+		(Ehira et al., 2009)	<i>cg1397</i>	<i>trmU</i>	TCGGGAACATGCGGATACGCTACGTTGTTGAGATTAATTA	1305259	38	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
+	+	+	+	+	diese Arbeit (RNAseq)	<i>cg0378</i>	-	CGATGGAACATTTTTGAAGAATACCTCGTTGAATCTACTG	329842	92	Putative phage-associated protein
+	+	+	+	+	(Busche et al., 2012)	<i>cg2515</i>	<i>dnaJ2</i>	TCGGGAACAATTTCTAAGGTGCCCGTTGCTAAGTATG	2395613	93	Chaperone, contain C-terminal Zn-finger domain
+	+	+		+	diese Arbeit (RNAseq)	<i>cg2152</i>	<i>clgR</i>	ACTGGGAACAATTTTAGGAAAGGAGTTGAACCTAACG	2039986	1	Transcriptional activator of Clp protease genes
+	+	+		+	diese Arbeit (RNAseq)	<i>cg2963</i>	<i>clpC</i>	AGTCTGGAAGTTTTGCCAATAAGGCGTTAAAGTGGGTG	2820286	38	Putative ATP-dependent protease (heat shock protein)

Fortsetzung Anhang Tab. 3

Identifikation/Nachweis der Promotoren											
RNA-Seq <i>in vivo</i> Datensätze ^a	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$	Referenz ^d	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
+	+				diese Arbeit (ROSE)	cg2620	<i>clpX</i>	AGATGGAACAACCCATTGCCAATATTGTTGGTTAGAGTTG	2502483	46	ATP-dependent clp protease ATPase subunit
+	+				diese Arbeit (ROSE)	<i>as-cg3313</i>	-	AGGAACCACAACCGCTGCCACCGACGTTTCCGATTTGAGC	3161284		Antisense RNA zu <i>cg3313</i> ; <i>mrcB</i> ; Putative membrane carboxypeptidase
+	+				diese Arbeit (ROSE)	<i>as-cg3395</i>	-	GGGAACAACGTTGACGTCACTCGACTGTTAGTCTTGGTTCG	3247347		Antisense RNA zu <i>cg3395</i> ; <i>proP</i> ; Putative proline/betaine permease, MFS-type
+	+				diese Arbeit (ROSE)	<i>cg0948</i>	<i>serC</i>	GGTGCTAGGAACTGACAGCTTCAGGGTTATAGTTGTTGGG	878643	60	Phosphoserine transaminase, AT class IV

^a)Identifizierung eines TSP in den *in vivo* RNA-Seq-Datensätzen RES167 $\Delta rshA$ und $\Delta sigH$.

^b)Relativ erhöhtes Transkriptlevel in RES167 $\Delta rshA$ im Vergleich zu RES167.

^c)Relativ erhöhtes Transkriptlevel in RES167 $\Delta cseE$ im Vergleich zu RES167.

^d)Referenz für Promotoridentifizierung

Anhang Tab. 4 SigH- und SigE- abhängigen Promotoren

Identifikation/Nachweis der Promotoren

RNA-Seq <i>in vivo</i> Datensätze	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA$	Microarray Referenz vs $\Delta cseE$	Referenz	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
+	+	+		+	(Halgasova et al., 2001)	<i>cg2102</i>	<i>sigB</i>	CTTGGAACCTTTTTGTGGAAGCAGTCCGTTGAACCTCTTG	1990877	25	RNA polymerase sigma factor rpoD (Sigma-A).
+	+	+	+	+	(Ehira et al., 2009)	<i>cg3079</i>	<i>clpB</i>	GAGTGAACATACTCAACTCTTTGTGCGTTATAGATATTA	2939558	91	Putative ATP-dependent protease (heat shock protein)
+	+	+		+	(Ehira et al., 2009)	<i>cg3100</i>	<i>dnaK</i>	AGTTGGAACAACCTTTGTGGCATTACCCTTGCTATATATG	2959823	120	Chaperone DnaK, heat shock protein
+	+	+	+	+	diese Arbeit (RNAseq)	<i>cg3344</i>	-	TGCGGGAACAAATTTGAAGTTTTTCAGTTGCTATAGGTA	3191301	0	Putative nitroreductase
+	+	+		+	diese Arbeit (ROSE)	<i>cg1272</i>	<i>cseE</i>	ACAGGGAACCTTATCACAGCGACATCCGTTTTGAGTACTA	1167979	65	Conserved hypothetical protein
+	+	+		+	diese Arbeit (ROSE)	<i>cg1277</i>	-	GTCTACGGGAACCATTTACCAGCGTGTACGTTGTAATG	1172127	34	Conserved putative membrane protein
+	+	+		+	diese Arbeit (ROSE)	<i>cg1640</i>	-	AAGCGGGAACAAAAAATACCCATAATCGTTCGATTCCA	1529558	178	Putative membrane protein, containing a CBS domain
+	+	+		+	diese Arbeit (ROSE)	<i>cg2092</i>	<i>sigA</i>	CTGGGGAACAAATACGTTTCCTATTGTTGTATATAGGTA	1981787	136	RNA polymerase sigma factor rpoD (Sigma-A).
+	+	+		+	diese Arbeit (ROSE)	<i>cg2657</i>	-	AACAGGGAACCTTTAACGAAAGACTTGTCTTACTTTAGTA	2534617	22	Putative membrane protein, putative pseudogen
+	+	+		+	diese Arbeit (ROSE)	<i>cg2984</i>	<i>ftsH</i>	AGAGGGAACGTTTTGGCCCTCACTTCGTTGATTAGTTG	2843112	77	Cell-division protein (ATP-dependent ZN metalloproteinase)
+	+	+	+		(Ehira et al., 2009)	<i>cg2661</i>	<i>dsbA2</i>	GAGGAACTAAAAAATCATCACCGTTGTTGAGATCAAGTG	2538097	190	thiol-disulfide isomerase
+	+	+		+	diese Arbeit (ROSE)	<i>cg2380</i>	-	CGGGGAACATCTCGGATATTCTGGGCGTTAATAAGGCATA	2268612	41	Putative membrane protein
+	+	+	+		(Ehira et al., 2009)	<i>cg1397</i>	<i>trmU</i>	TCGGGAACATGCGGATACGCTACGTTGTTGAGATTAATTA	1305259	38	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
+	+	+	+	+	diese Arbeit (RNAseq)	<i>cg0378</i>	-	CGATGGAACATTTTTGAAGAATACCTCGTTGAATCTAGTG	329842	92	Putative phage-associated protein

Anhang Tab. 5 SigH-exklusive Promotoren

Identifikation/Nachweis der Promotoren

RNA-Seq <i>in vivo</i> Datensätze ^a	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$	Referenz ^d	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
+	+		+		(Ehira <i>et al.</i> , 2009)	<i>cg0617</i>	-	CAGGGAATTTCCGCGCCCGCTTCCCTGTTTAAATAAACG	544705	36	Hypothetical protein
+	+		+		diese Arbeit (RNAseq)	<i>cg0699</i>	<i>guaB2</i>	TGTGGAATAGCGTGTGAGCTGCAGCGGTTATCCGATTTAA	616737	60	IMP dehydrogenase
+	+		+		(Busche <i>et al.</i> , 2012)	<i>cg1127</i>	<i>mca</i>	GTTTGGAAATGTTTCGAAGCGCCTGATGCGTTGGATGGAGAG	1046412	207	Putative mycothiol S-conjugate amidase
+	+		+		(Ehira <i>et al.</i> , 2009)	<i>cg1375</i>	-	TAGGGATTAGCTTTGTACTTAAACTTGTGTTTTTAAGTG	1283606	56	Putative thioredoxin DNA/RNA helicase, superfamily I
+	+		+		(Busche <i>et al.</i> , 2012)	<i>cg1555</i>	<i>uvrD3</i>	ATCTGGAATGATTGATAGCTCCCAAGTGTGTATCTATTC	1441999	56	DNA/RNA helicase, superfamily I
+	+		+		diese Arbeit (RNAseq)	<i>cg1689</i>	<i>pup</i>	GATGGATTAACGTTAACAATAAGTTTGTACATAGTGTGG	1576737	134	Conserved hypothetical protein
+	+		+		(Busche <i>et al.</i> , 2012)	<i>cg1709</i>	<i>mshC</i>	TTGTGGAATAGGTGCACTGGCGGTTTGGTTGAAGTTTCAG	1599116	141	Putative 1-D-myo-inosityl-2-amino-2-deoxy-alpha-D-glucopyranoside--L-cysteine ligase
+	+		+		(Ehira <i>et al.</i> , 2009)	<i>cg1765</i>	<i>sufR</i>	CGGGAATGGAATTAGGGAACACTTGTGTTGTCTAAAGGTG	1655074	30	Putative regulator protein
+	+		+		(Ehira <i>et al.</i> , 2009)	<i>cg2078</i>	<i>msrB</i>	GATGGAATTTTCAGCGCAGCGTTGGTTGGGACTATTG	1969473	121	Peptide methionine sulfoxide reductase
+	+		+		diese Arbeit (RNAseq)	<i>cg2194</i>	<i>mtr</i>	CTTGGGAAGCTTTTTATAGTCATATGCGTTGAGATACGTG	2086894	72	Putative NADPH-dependent mycothiol reductase

Fortsetzung Anhang Tab. 5

Identifikation/Nachweis der Promotoren					Referenz ^d	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
RNA-Seq <i>in vivo</i> Datensätze ^a	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$							
+	+		+		diese Arbeit (RNAseq)	cg2206	<i>ispG</i>	TACGGAAATTGAACTGCCGGCACCCGATGTTAAAAGAGGTGA	2098676	68	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
+	+		+		diese Arbeit (RNAseq)	cg2214	-	TCCGGAAATATTTTTCTCCAAGCGGGTTAACTCACTGT	2105774	173	Putative Fe-S-cluster redox enzyme
+	+		+		(Ehira <i>et al.</i> , 2009)	cg2411	-	ATCGGGAATAGGGGTGCACACTTCATCGTTGAAAGGAATC	2302853	254	Conserved hypothetical protein, HesB/YadR/YfhF family
+	+		+		diese Arbeit (RNAseq)	cg2423	<i>lipA</i>	GAGGAATTTTGTCCGGTGGGGCCCTCGTTGAAGCGAAGTA	2313882	94	Lipoyl synthetase
+	+		+		diese Arbeit (RNAseq)	cg2835	-	TAGGAAATAAAACAGGGTCTCTTTGTGTTCAAAGGTATA	2695784	56	Putative acetyltransferase
+	+		+		diese Arbeit (RNAseq)	cg3236	<i>msrA</i>	CCCGGAATAATGGCAGCTAAGTAGGGTTGAAGGGCATAA	3099488	36	Protein-methionine-S-oxide reductase
+	+		+		(Nakunst <i>et al.</i> , 2007)	cg3299	<i>trxB1</i>	CGGGAAATACTACAGTCCGCTGAAAGTTGGTCTATATA	3151823	31	Thioredoxin (TRX)
+	+		+		(Ehira <i>et al.</i> , 2009)	cg3405	-	ACGGGGGAATGGAAAAAGTACGCTTGGTGTTCATATAGCG	3258874	36	NADPH:quinone reductase Zn-dependent oxidoreductase
+	+		+		(Busche <i>et al.</i> , 2012)	cg1560	<i>uvrA</i>	TAGGGAAATACATGACCCCTCAAATTCAAATGAATCACCATG	1443568	386	Excinuclease ABC, ATPase subunit A
+	+				diese Arbeit (ROSE)	cg1095	-	GGTGGAAATAAATACACTGATCAGGTCGTTCTGAAACTG	1017617		hypothetisches Protein
+	+				diese Arbeit (ROSE)	as-cg2375	-	TTATGAAATTCTACAAAAGACACTTCTGTTTGAGTTATTG	2262886		Antisense RNA zu cg2375; <i>ftsI</i> ; Penicillin-binding protein
+	+				diese Arbeit (ROSE)	cg0454	-	GTAGGGAAATTTTTTCGGCCCATGTGCGGTTGTGTCCTAAG	400419	80	Putative transcriptional regulator, TetR-family

Fortsetzung Anhang Tab. 5

Identifikation/Nachweis der Promotoren											
RNA-Seq <i>in vivo</i> Datensätze ^a	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$	Referenz ^d	Kodier-region	Gen-name	Sequence (-39 -- +1)	Position Genom	Abstand Start-codon	Funktion
+	+		+		diese Arbeit (ROSE)	cg1282	-	ATGGGAATGAATGGACCTGACCAAGACGTTGTATCTTTTG	1182549	244	Conserved hypothetical protein
+	+		+		diese Arbeit (ROSE)	cg2247	-	GCACGGAAGCAATTTGCCAAGTAAATGGTTGAAAGGAATA	2131538	9	Hypothetical protein
+	+				diese Arbeit (ROSE)	as-cg2052	-	CCATGGAATATTTGGAAAAGCAACAAGGTTAACATCACCA	1943327		Antisense RNA zu cg2052; Putative secreted protein; Signal peptide
+	+				diese Arbeit (ROSE)	as-cg3357	-	CGCTGGAATATTTGGTGGGATAGAGTTGTTGTTATGACGG	3205910		Antisense RNA zu cg3357; <i>trpP</i> ; <i>Permease, tryptophan-specific</i>
+	+				diese Arbeit (ROSE)	cg1386	fixA	TTGAAATTTTAGTGCGTGGTGTGATGTTGATCCTGATCA	1296568	335	Putative electron transfer flavoprotein, beta subunit
+	+		+		diese Arbeit (ROSE)	cg1432	ilvD	GATGGGAATACAACACAACAGAAAGCGTTGAATCCCAAG	1336937	188	Dihydroxy-acid dehydratase
+	+				diese Arbeit (ROSE)	cg2263	-	GAGGGAATTGCATCTCCATTGGAGTTGTTATCCATCCACG	2149096	0	Hypothetical protein
+	+				diese Arbeit (ROSE)	cg2409_P1	ctaC	GGGGGAATTTCAAGGGCAATCAAAGGTTGATGGTCTGTG	2299889	220	Cytochrome c oxidase subunit II
+	+				diese Arbeit (ROSE)	cg2409_P2	ctaC	TGTTGGAAGAAATAAGACGCTGCATTTGTTAAATCTCGTG	2300097	428	Cytochrome c oxidase subunit II

- a) Identifizierung eines TSP in den *in vivo* RNA-Seq-Datensätzen RES167 $\Delta rshA$ und $\Delta sigH$.
 b) Relativ erhöhtes Transkriptlevel in RES167 $\Delta rshA$ im Vergleich zu RES167.
 c) Relativ erhöhtes Transkriptlevel in RES167 $\Delta cseE$ im Vergleich zu RES167.
 d) Referenz für Promotoridentifizierung

Anhang Tab. 6 Potentiell exklusiv SigE-abhängige Promotoren

RNA-Seq <i>in vivo</i> Datensätze ^a	Identifikation/Nachweis der Promotoren				Referenz ^d	Kodier- region	Gen- name	Sequence (-39 -- +1)	Position Genom	Abstand Start- codon	Funktion
	ROSE-SigH (<i>in vitro</i>)	ROSE-SigE (<i>in vitro</i>)	Microarray Referenz vs $\Delta rshA^b$	Microarray Referenz vs $\Delta cseE^c$							
+		+		+	diese Arbeit (ROSE)	cg1121	-	GTGGGGAAC TATACTGATCCTTGATGCGTTAACTTGATTG	1043461	36	Permease, MFS- type
+		+		+	diese Arbeit (ROSE)	cg3309	-	CGTGGGAAC TTTCTCTAACTAAGTGCCTTAAGGAAGGTA	3158581	17	Putative secreted protein
+		+		+	diese Arbeit (ROSE)	cg0986	amtR	GGTGGAAAC TTATGAGGGTCTTCCTTCGTTGGACATGGTG	924929	397	Transcriptional repressor of nitrogen metabolism, TetR- family

a) Identifizierung eines TSP in den *in vivo* RNA-Seq-Datensätzen RES167 $\Delta rshA$ und $\Delta sigH$.

b) Relativ erhöhtes Transkriptlevel in RES167 $\Delta rshA$ im Vergleich zu RES167.

c) Relativ erhöhtes Transkriptlevel in RES167 $\Delta cseE$ im Vergleich zu RES167.

d) Referenz für Promotoridentifizierung

Danksagung

Herrn Dr. Jörn Kalinowski danke ich für die Bereitstellung des interessanten Themas meiner Dissertation und für die hervorragende wissenschaftliche Betreuung. Desweiteren danke ich ihm für die Zeit, die ich in seiner Arbeitsgruppe verbringen durfte und dafür, dass er mich zu jeder Zeit mit Rat und Tat unterstützt hat.

Ein Dankeschön gebührt allen Mitgliedern der „Coryne-Gruppe“ für ein tolles Arbeitsklima, der ständigen Bereitschaft zu konstruktiven Diskussionen und natürlich auch für die spassigen Zeiten neben der Arbeit. Herausheben möchte ich Marcus, Freddy, Armin, Jens und Nils (um nur bei den Männern zu bleiben), die in all den Jahren zu guten Freunden geworden sind.

Ein besonderes Dankeschön geht in diesem Zusammenhang auch an alle TAs, die mich in meiner Arbeit unterstützt haben. Hier sind Eva, Yvonne und Annika ganz vorne mit dabei.

Ein besonders großer Dank geht natürlich auch an die Post-Docs Karina, Marcus, Jens und Christian, bei denen ich mich speziell für die anregenden Gespräche und die konstruktiven Verbesserungsvorschläge beim Schreiben meiner Arbeit bedanken möchte.

Darüber hinaus danke ich unseren tschechischen Kooperationspartnern Mirek, Rado, Jirka, und Jan für die perfekte Zusammenarbeit.

Zuletzt aber und vor allen anderen möchte ich mich ganz herzlich bei meiner Familie bedanken. Vielen Dank das Ihr immer für mich dagewesen seit. Ohne Euch wäre das Alles nicht möglich gewesen. Jetzt habe ich die Arbeit abgegeben und wieder mehr Zeit für meine Frau Jessie, meine beiden Mäuse Raphael und Joline und natürlich auch für meine Brüder Christopher und Michael.

Ein dickes Danke geht zum Schluß aber natürlich auch an alle Freunde raus.

Erklärung

Mit der vorliegenden Dissertation beantrage ich beim Promotionsausschuss der Fakultät für Biologie der Universität Bielefeld die Eröffnung des Promotionsverfahrens nach § 9 der Promotionsordnung in der Fassung vom 10. Mai 2006.

Die Anfertigung dieser Dissertation mit dem Titel „Analyse von Regulationsnetzwerken der *Extracytoplasmic Function* (ECF)-Sigmafaktoren in *Corynebacterium glutamicum*“ erfolgte unter Betreuung von Dr. Jörn Kalinowski in der AG "Mikrobielle Genomik und Biotechnologie" an der Fakultät für Biologie der Universität Bielefeld.

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Dissertation, die anderen Quellen im Wortlaut oder dem Sinn nach entnommen wurden, sind durch Angaben der Herkunft kenntlich gemacht. Dies gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen sowie für Quellen aus dem Internet.

Weiterhin erkläre ich, dass die vorliegende Dissertation weder vollständig noch in Auszügen einer anderen Fakultät mit dem Ziel vorgelegt worden ist, einen akademischen Titel zu erwerben. Ich bewerbe mich hiermit erstmalig um den Doktorgrad der Naturwissenschaften der Universität Bielefeld.

Bielefeld, den 09.07.2013

Tobias Busche

