

Optimierung der Sekretion von rekombinanten Proteinen
aus *Escherichia coli*
durch Analyse der Genexpression

Von der Naturwissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades
Doktor der Naturwissenschaften
– Dr. rer. nat. –
genehmigte Dissertation
von

Dipl.-Chem. Michael Reck,
geboren am 07.05.1978 in Gehrden.

Hannover, im Februar 2009

Referent: Prof. Dr. T. Scheper
Institut für Technische Chemie
Gottfried Wilhelm Leibniz Universität Hannover

Korreferent: PD. Dr.-Ing. K. Friehs
Technische Fakultät
Universität Bielefeld

Tag der Promotion: 22.05.2009

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel durchgeführt habe.

Hannover, im Februar 2009

Inhaltsverzeichnis

1	Einführung	3
2	Theoretische Grundlagen	4
2.1	Einführung	4
2.2	Typ I Sekretionsmechanismus	4
2.3	Typ II Sekretionsmechanismus	5
2.3.1	SecB-abhängiger Sekretionspfad	6
2.3.2	SRP-Pathway	8
2.3.3	Tat-Pathway	8
2.4	Extrazelluläre Produktion	12
2.5	<i>Bacteriocin release proteins</i>	13
3	Genexpressionsanalytik	16
3.1	Übersicht	16
3.2	Hybridisierung	16
3.3	DNA-Microarray-Technologie	17
3.3.1	Übersicht Microarray-Technologie	17
3.3.2	Oberflächen	18
3.3.3	Computer-gestütztes Design von Oligonukleotiden	19
3.3.4	Hybridisierung und <i>Labeling</i>	20
3.3.5	Waschen von Microarrays	22
3.3.6	Scannen von Microarrays	22
3.4	Auswertemethoden für Microarrays	23
3.4.1	Primäranalyse	23
3.4.2	Sekundäranalyse	25
3.4.3	<i>Within-Array</i> -Normalisierung	26
3.4.4	<i>Between-Array</i> -Normalisierung	29
3.4.5	Identifizierung differentiell exprimierter Gene	29
3.4.6	Genomik und Clusteranalysen	31
3.4.7	<i>K-means Clustering</i>	33
3.5	Quantitative <i>real-time</i> PCR	34
4	Experimentelle Arbeiten	36
4.1	Evaluierung der Variabilitäten von Chipexperimenten	37
4.2	Probenprozessierung und experimentelles Design	43
4.3	Übersicht über die verwendeten Sekretionstämme	44
4.4	Analyse der extrazellulären, Tat-spezifischen Sekretion	45
4.4.1	Induktion mit 100 mg/l <i>L</i> -Arabinose	45
4.4.2	Induktion mit 1000 mg/l <i>L</i> -Arabinose	54
4.4.3	Zusammenfassung	58
4.5	Analyse der extrazellulären, Sec-spezifischen Sekretion	58
4.5.1	Induktion mit 1000 mg/l <i>L</i> -Arabinose	60
4.5.2	Induktion mit 25 mg/l <i>L</i> -Arabinose	62
4.5.3	Induktionsniveau 25 mg/l <i>L</i> -Arabinose; Aufnahme einer feingliedrigeren Kinetik	65

4.5.4	Cluster Analyse	67
4.5.5	Transkriptioneller Vergleich der Zeitpunkte 30 und 60 Minuten nach Induktion mit dem präinduktiven Zustand	72
4.5.6	Induktionsniveau 1000 mg/l; feingliedrige Kinetik	74
4.5.7	Cluster-Analyse	76
4.5.8	Zusammenfassung	79
4.6	Analyse der periplasmatischen, Sec-spezifischen Sekretion	80
4.6.1	Clusteranalyse	83
4.6.2	Zusammenfassung	86
4.7	Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien	86
4.7.1	Einführung	86
4.7.2	Klonierung des Sekretionsreporterplasmids pBADtorAGFPSsrA	88
4.7.3	Klonierung weiterer GFP-Stämme zur Charakterisierung und Verifikation des etablierten Sekretionsreportersystems	95
4.7.4	Fluoreszenzmikroskopische Untersuchung	99
4.7.5	Durchflußzytometrie	99
4.7.6	SDS-PAGE und Zellfraktionierung	103
4.7.7	Bioreaktorkultivierung	105
4.7.8	RNA-Isolation	105
4.7.9	Transkriptionelle Analyse des Stammes pBADtorAGFPSsrA	106
4.7.10	Quantitative <i>real-time</i> PCR	109
4.7.11	<i>Nothern Blot</i> Analyse der GFP-Expression	123
4.7.12	Inkorporation von DIG-dCTP während der PCR	125
4.7.13	Inkorporation von DIG-dUTP mittels Reverser Transkription	126
4.7.14	Generierung von DIG-Sonden mittels <i>in vitro</i> Transkription	126
4.7.15	Zusammenfassung	130
4.8	Etablierung verschiedener arabinoseinduzierbarer RFP <i>E. coli</i> Stämme	131
4.9	Optimierungsansätze für Microarray-Applikationen	136
4.10	<i>Quantum Dots</i>	136
4.11	Etablierung eines Systems zur absoluten Quantifizierung von DNA-Microarrays	142
5	Zusammenfassung und Ausblick	160
6	Literatur	162
A	Anhang	175
A.1	Protokolle	175
A.1.1	<i>Spotting</i> von Oligonukleotiden auf Aldehyd-funktionalisierte <i>Glasslides</i>	175
A.1.2	RNA-Isolation aus <i>E. coli</i> mittels Phenol-Chloroform Extraktion	176
A.1.3	RNA-Isolation aus <i>E. coli</i> mittels des <i>InnuprepTM</i> RNA Mini Kits (Analytik Jena)	176
A.1.4	Direktes <i>Labeling</i> mit der Superscript III Reversen Transkriptase	176
A.1.5	Indirektes <i>Labeling</i> mit dem <i>MicromaxTMTSATM</i> <i>Labeling Kit</i>	177
A.1.6	Indirekte Markierung der cDNA mit Streptavidin funktionalisierten <i>Quantum Dot</i> -Partikeln	179
A.1.7	Plasmidisolierung mit dem <i>Peq-Lab Miniprep Kit II</i>	180
A.1.8	Herstellung von LB-Agarplatten	180

A.1.9	Herstellung chemisch kompetenter Zellen mittels TSS-Methode . . .	181
A.1.10	Herstellung chemisch kompetenter Zellen mittels $CaCl_2$ -Methode . .	181
A.1.11	Transformation chemisch kompetenter Zellen mittels Hitzeschock . .	181
A.1.12	Ligation mittels <i>cut-ligation</i> Methode	182
A.1.13	Ligation mittels kohäsiver Enden	182
A.1.14	Restriktionsverdau	182
A.1.15	Aufreinigung mittels <i>PCR-Purification Kit</i>	182
A.1.16	Gelaufreinigung	183
A.1.17	Generierung von cDNA für PCR-Applikationen	184
A.1.18	<i>Real-time</i> PCR	184
A.1.19	PCR mit <i>Taq</i> -Polymerase	185
A.1.20	PCR mittels <i>Pfu</i> -Polymerase	185
A.1.21	Kolonie-PCR	186
A.1.22	Generierung DIG-markierter DNA-Sonden mittels PCR	186
A.1.23	Generierung DIG-markierter cDNA mittels Reverser Transkription .	186
A.1.24	Generierung von DIG-markierten RNA-Sonden mittels <i>in vitro</i> Tran- skription	187
A.1.25	SDS-PAGE	187
A.1.26	<i>Nothern Blotting</i>	188
A.1.27	Generierung von RNA mittels <i>in vitro</i> Transkription	189
A.1.28	Affinitätschromatographische Aufreinigung über Vivapure Metall-Chelat <i>Spin Columns</i>	189
A.1.29	Phenol-Chloroform Aufreinigung	190
A.1.30	Zellfraktionierung	190
B	Puffer und Lösungen	190
B.1	Puffer und Lösungen	190
C	PCR-Primer	193
C.0.1	Absolute Quantifizierung	193
C.0.2	GFP-Klonierung	194
C.0.3	RFP-Klonierung	195
C.0.4	<i>Real-time</i> PCR	195
D	Mathematische Formeln und Statistik	199
D.1	Berechnung von Inkorporationsfrequenzen	199
D.2	Statistik	199

Abbildungsverzeichnis

1	Typ II Sekretionsmechanismus	6
2	Auswertung von Microarrays	20
3	Indirektes und direktes cDNA- <i>Labeling</i>	21
4	<i>Alignment</i>	24
5	Segmentation und <i>Spot</i> formen	24
6	<i>Scatterplot</i>	27
7	<i>MA-Plot</i>	28
8	Verfahren zur <i>Between-Array</i> -Normalisierung	30
9	Hierarchisches <i>Clustering</i>	32
10	Verbindungsmethoden	33
11	Amplifikationskurven <i>real-time</i> PCR	34
12	Experimentelles Design zur Evaluierung der Variabilitäten von Chipexperimenten	37
13	Gelelektrophoretische Analyse der RNA-Isolationen	39
14	Absorptionsspektren cDNAs	40
15	<i>Logratios</i> von fünf verschiedenen Genen des Parallelexperimentes	42
16	<i>Loop-Design</i> mit <i>Common Reference</i>	43
17	Schema der Sekretion des Stammes MG1655 pBADLppBRPtorAmalE	45
18	Regulierte Gene der Multifun-Hauptklassen	46
19	Relativer Vergleich der Multifun Hauptklassen-regulierte Gene	47
20	Relativer Vergleich der Multifun-Klasse Stessadaption	48
21	Relativer Vergleich der Multifun-Klasse Lokalisierung der Genprodukte	49
22	Relativer Vergleich der Multifun-Klasse Zellprozesse	50
23	Relativer Vergleich der Multifun-Klasse Zellstruktur	51
24	Schema der Sekretion des Stammes Ara1655 pBADLppBRPphoAbla	59
25	Übersicht über die Anzahl der regulierten Gene I	66
26	<i>K-means</i> Clusteranalyse I	67
27	Regulationskaskade Initiierung der SOS-Antwort	69
28	<i>Logratios</i> verschiedener <i>housekeeping</i> -Gene	74
29	<i>Logratios</i> verschiedener Arabinoseimportgene	75
30	Übersicht über die Anzahl regulierter Gene II	76
31	Multifun-Hauptklassen: negativ regulierte Gene	76
32	<i>K-means</i> Clusteranalyse II	78
33	Schema der Sekretion des Stammes Ara1655 pBADphoAbla	81
34	Übersicht über die Anzahl der regulierten Gene III	81
35	Regulierte abwehrinvolvierte Gene-Vergleich	82
36	<i>K-means</i> Clusteranalyse III	84
37	Sekretionskonstrukt torAGFPSrA	88
38	Klonierungsstrategien	89
39	Test verschiedener <i>Enhancer</i> -Zusätze für die PCR	90
40	Selbstligationsexperiment	92
41	Ligation torA mit GFPSrA	93
42	Analytischer Restriktionsverdau Plasmid pBADtorAGFPSrA	94
43	Theoretisches Fragmentationsmuster Plasmid pBADtorAGFPSrA	95
44	GFP-Plasmide	96

45	Subklonierung der GFP-Konstrukte in pGEM5Zf(+)	97
46	Kolonie-PCR GFP-Konstrukte	97
47	Analytischer Restriktionsverdau der generierten GFP-Konstrukte	98
48	Fluoreszenzmikroskopische Untersuchung GFP	99
49	Schematischer Aufbau eines Durchflußzytometers	100
50	Durchflußzytometrische Analyse GFP-Stämme	101
51	Kinetische Evaluierung der Fluoreszenz von pBADtorAGFPSSrA	102
52	SDS-PAGE der Zellfraktionen verschiedener GFP-Stämme	104
53	Schema der Sekretion des Stammes XL1 pBADtorAGFPSSrA	107
54	Optimierung der <i>Annealing</i> -Temperatur	112
55	Schmelzkurvenanalyse von PCR-Amplifikaten	113
56	Bestimmung des Ct-Wertes	114
57	Evaluierung hinsichtlich DNA-Kontaminationen	115
58	Amplifikationskurven für eine 1:10 Verdünnungsreihe des GFP- <i>Templates</i>	115
59	GFP-Primereffizienz	116
60	Parallelexperiment zur Evaluierung der Variabilitäten in der quantitativen <i>real-time</i> PCR	117
61	Normalisierung über <i>gapdH</i>	118
62	Ermittelte Regulationen quantitative <i>real-time</i> PCR	118
63	Amplifikationskurven GFP	122
64	Vergleich <i>real-time</i> PCR vs. Microarray	123
65	Generierung von DIG-Sonden mittels PCR	125
66	DIG-Markierung mittels <i>in vitro</i> Transkription	126
67	<i>Dot-Blot</i> -Analyse der GFP-Expression	127
68	Mock-Hybridisierung	129
69	Denaturierende Gelelektrophorese von RNA	129
70	<i>Nothern Blot</i> -Analyse der GFP-Expression	130
71	Zelllysate dreier induzierter, RFP-codierender Stämme	132
72	RFP-Plasmide	133
73	Schüttelkolbenkultivierungen verschiedener RFP-Stämme	134
74	Fluoreszenzmikroskopische Analyse TOP10 pBADtorARFPSSrA	134
75	Durchflußzytometrische Analyse dreier RFP-codierender Stämme	135
76	Scan eines sekretionsspezifischen <i>Quantum Dot</i> DNA-Microarrays	137
77	Evaluierung verschiedener initialer RNA-Mengen	138
78	Evaluierung verschiedener Dotierungsverhältnisse	139
79	Vergleich der SNRs zwischen einem konventionellen Microarray und einem Q605-Microarray	140
80	Vergleich der resultierenden SNRs von Q655 gegen Q605	141
81	Synthese von <i>in vitro</i> Transkripten	144
82	Zwei Klonierungsstrategien zur Generierung der <i>in vitro</i> Transkriptionsplasmide	146
83	Gelelektrophoretische Analyse PCR- <i>Inserts</i>	149
84	Optimierung der <i>Annealing</i> -Temperatur für das <i>dnaK-Insert</i>	150
85	Gelelektrophoretische Analyse der generierten Transkriptionsplasmide	150
86	Linearisierung der Transkriptionsplasmide-Strategie	151
87	Detektionsprobleme von <i>in vitro</i> RNA-Transkripten	152
88	Inkorporationsfrequenzen von Cy5-dCTP und Reaktionseffizienzen der cDNA-Synthese	154

89	Verdünnungsreihe Signalpeptidase 2-cDNA	155
90	Linearer Zusammenhang zwischen Signalwerten und cDNA-Konzentration ei- ner cDNA-Spezies	156
91	Visualisierung von zehn verschiedenen <i>in vitro</i> RNA-Transkripten	157
92	Absorptionsspektren von zehn verschiedenen, Cy5-markierten cDNA-Spezies	157
93	Regressionsgrade zur Bestimmung der RNA-Kopien aus dem normalisierten Signalwert	159
94	Proteinsekretions-Pathways in <i>E. coli</i>	160

Tabellenverzeichnis

1	Konzentrationen und Reinheiten der RNA-Präparationen des Parallelexperimentes	38
2	Farbstoffkonzentrationen und Inkorporationsfrequenzen des Parallelexperiments	41
3	Übersicht über die verwendeten Sekretionsstämme	44
4	Vergleich der Regulation von 30 ausgewählten Genen	83
5	Optimierte <i>Annealing</i> -Temperaturen	111
6	Gradensteigungen und Korrelationskoeffizienten zur Evaluierung der Primereffizienzen	116
7	Genexpressionsquotienten-Vergleich verschiedener Normalisierungen	119
8	Evaluierte Genexpressionsquotienten	120
9	Ausgewählte Gene bzw. Genabschnitte für die <i>in vitro</i> Transkription	143
10	Vergleich dreier Klonierungsstrategien	148
11	Vergleich verschiedener cDNA-Ansätze aus <i>in vitro</i> Transkripten	153
12	<i>Multiplexing</i> -Versuch	158

Danksagung

In erster Linie danke ich Herrn Prof. Dr. Thomas Scheper für die Möglichkeit, am TCI eine Promotionsstelle antreten und ein hoch aktuelles Thema bearbeiten zu können. Herrn Prof. Dr. Bernd Hitzmann danke ich für die Mithilfe am Projekt und für die Funktion als Drittprüfer. Mein besonderer Dank gilt Herrn Dr. Frank Stahl und Herrn PD Dr. Karl Friehs für die nachhaltige Unterstützung in fachlichen Fragen, Förderung und ein ausgezeichnetes Arbeitsklima. Dr. Benjamin Sommer danke ich für seine kollegiale Mitarbeit an einem hoch interessanten Forschungsthema und für die Bereitstellung der Proben für die Microarray-Analysen. Cornelia möchte ich für die intensive Unterstützung bei der Microarray-Auswertung und für eine sehr lustige Bürozeit danken. Martin Paehler danke ich für eine sehr angenehme Zeit im Labor und Büro sowie für seine vielen wertvollen Tips.

Meiner Freundin Sonja möchte ich für die Geduld, die vielen fachlichen Diskussionen und Tips, sowie für Ihre Unterstützung ganz besonders danken. Schön das es Dich gibt. Mike gilt mein Dank für die erquickende Zeit am TCI, für die Hilfe bei den Bioreaktorkultivierungen und Tex sowie für die moralische Unterstützung bei der irreversiblen Konversion diverser neuronaler Netzwerke. Aus einer äquivalenten intrinsischen Motivation heraus danke ich Herrn Kptl. z. See Thomsen, Dr. H.C. und S. Plissken.

Allen, die mich bei dieser Arbeit unterstützt haben, danke ich, ohne sie hier namentlich aufzuführen.

Abkürzungen

AP	Alkalische Phosphatase
APS	Ammoniumpersulfat
ATP	Adenosintriphosphat
BSA	Rinderserumalbumin <i>bovine serum albumine</i>
bp	Basenpaare
BRP	<i>bacteriocin release protein</i>
CV	normierte Standardabweichung <i>coefficient of variance</i>
DEPC	Diethylpyrocarbonat
DIG	Digoxigenin
DMF	Dimehtylformamid
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleotidtriphosphat
EDTA	Ethylendiamintetraessigsäure
GAPDH	Glycerinaldehyd-3-phosphat-Dehydrogenase
GFP	Grün fluoreszierendes Protein
GST	Glutathion-S-Transferase
IgG	Immunglobulin G
IPTG	Isopropyl- β -thiogalactosid
log	Logarithmus
LPS	Lipopolysaccharid
PBS	Phosphatgepufferte Salzlösung <i>phosphate buffered saline</i>
PCR	Polymerase Kettenreaktion
<i>Pfu</i> -Polymerase	DNA-abhängige DNA-Polymerase aus <i>Pyrococcus furiosus</i>
PMT	Sekundärelektronenvervielfacher <i>photo multiplier tube</i>
RFP	Rot fluoreszierendes Protein
RNA	Ribonukleinsäure <i>ribonucleic acid</i>
rNTP	Ribonukleotidtriphosphat
RT	Reverse Transkriptase
Sec	<i>general secretory</i>
SNR	Signal-Rauschverhältnis <i>signal-to-noise ratio</i>
<i>Taq</i> -Polymerase	DNA-abhängige DNA-Polymerase aus <i>Thermus aquaticus</i>
Tat	<i>twin arginine translocator</i>
Tris	Tris(hydroxymethyl)aminoethan
X-Gal	5-Brom-4-Chlor-3-Indoxyl- β - <i>D</i> -Galactopyranosid

Für Längen-, Massen-, Stoffmengen- und Zeitangaben werden SI-Einheiten verwendet sowie Vorsilben für dezimale Vielfache und Teile der SI-Einheiten nach DIN 1301.

Zusammenfassung

Escherichia coli stellt aufgrund hoher Wachstumsraten, der Toleranz gegenüber vielfältigen Kultivierungsbedingungen und effektiven genetischen Modifizierungsmöglichkeiten nach wie vor einen der wichtigsten Produzenten rekombinanter Proteine für die biotechnologische Produktion und Forschung dar. Zur Optimierung der Produktausbeuten weisen Sekretionsstrategien gegenüber der konventionellen cytoplasmatischen Expression viele Vorteile auf, wie ein signifikant vereinfachtes *downstream processing* bei oftmals deutlich erhöhten Produktmengen an biologisch aktivem und löslichem Protein. Die für derartige Applikationen nutzbaren bakteriellen Hauptsekretionspfade Sec und Tat unterliegen allerdings limitierten Exportkapazitäten. Ein detailliertes mechanistisches Verständnis der komplexen molekularen Interaktionen des Sekretionsvorgangs ist in diesem Kontext höchst erstrebenswert. Zu diesem Zweck wurde der Sekretionsvorgang auf Transkriptomebene mittels kommerzieller *whole-genome* DNA-Microarrays evaluiert. Anhand verschiedener induzierbarer Sekretionssysteme erfolgte sowohl die Analyse der Tat- und Sec-spezifischen Translokation ins Periplasma als auch die extrazelluläre Sekretion mittels Stämmen, die sogenannte *bacteriocin release proteins* (BRPs) coexprimieren. Anhand der Daten aus den *whole-genome* Experimenten wurde ein sekretionsspezifischer *low-density* Microarray mit 109 genspezifischen Oligonukleotiden etabliert und zur routinemäßigen Applikation gebracht. Es konnte gezeigt werden, dass die differentielle Genexpression der sekretionsspezifischen Gene innerhalb weniger Minuten stattfindet. Aufgrund der BRP-Expression wurden substantielle regulatorische Effekte auf Membranstress-assoziierte Gene beobachtet, als Beispiel sei hier das Rfa- und das Phagenschockoperon *pspABCDE* sowie die Gene *spy* und *lolAB* genannt. Zudem fand die differentielle Expression von Komponenten der LPS- und Peptidoglykan-Biosynthese sowie weiterer Gene statt, die in der Aufrechterhaltung der Membranintegritäten fungieren. Eine Initiation der zellulären Stress- und SOS-Antwort wurde anhand der differentiellen Expression der Gene *lexA*, *recA* und *soxS* beobachtet. Für die Sec- und Tat-spezifische Sekretion wurde die Regulation von verschiedenen Chaperonen (FkpA, GroES, GroEL), von Komponenten der bakteriellen Translokase (SecY, SecE, TatD), sowie von Genen, die mechanistisch in den Sekretionsvorgang involviert sein könnten (*mltE*, *imp*), nachgewiesen. Durch Coexpression von Chaperonen und Translokase-Komponenten konnte eine signifikante Steigerung der Sekretionseffizienzen erzielt werden. Deshalb stellen sie interessante Ziele für Optimierungsstrategien dar. Zur besseren Vergleichbarkeit von Expressionsprofilen und funktionellen Interaktionen wurden die regulierten Gene einer *k-means* Clusteranalyse unterzogen. Die Verifikation der Chip-generierten Daten mittels quantitativer *real-time* PCR und *Northern Blotting* wurde exemplarisch anhand des Datensatzes einer *low-density* Microarray-Analyse gezeigt. Zusätzlich wurde die Methode der Chiptechnologie selbst weiterentwickelt, indem neue innovative und multiple Fluorophore, sogenannte *Quantum Dots*, eingesetzt wurden und ein System zur absoluten Quantifizierung etabliert wurde.

Schlüsselbegriffe: Proteinsekretion, Sec-Pathway, Tat-Pathway, *bacteriocin release proteins*, *Quantum Dots*, DNA-Microarray Technologie, Clustering, Bioreaktorkultivierung

Abstract

E. coli still remains one of the most important hosts for the production of recombinant proteins in science and biotechnology industry. Secretion of recombinant proteins comprises many advantages over conventional cytoplasmic expression strategies as e.g. a significantly facilitated downstream processing. In addition more soluble product with an enhanced biological activity can be obtained in most cases. Utilizing the natural secretion pathways Sec and Tat is limited by the secretion-capacity of the translocase. Therefore, a detailed insight into the complex underlying regulatory mechanisms of the secretion process is desirable to resolve this bottleneck for the recombinant protein production. For that purpose the secretion-process is evaluated on the transcriptom-level using commercially available whole-genome DNA-microarrays. Different secretion-systems for the Sec- and Tat-specific translocation into the periplasm and the secretion into the extracellular space, mediated by coexpressed bacteriocin release proteins (BRPs), were evaluated. The results of the whole-genome experiments were used to develop and apply a secretion-specific low-density array with 109 gene-specific oligonucleotides. It was validated that the differential expression of secretion specific genes occurs within few minutes. BRP-expression causes substantial regulatory effects of genes involved in the response to membrane stress as e.g. the Rfa- and the phage-shock operon as well as genes *spy* and *lolAB*. Furthermore genes involved in maintaining the membrane-integrity or coding for components of the Lipopolysaccharide- and peptidoglycan-biosynthesis were differentially expressed. An initiation of stress and SOS-response was observed by the regulation of genes *lexA*, *soxS* and *recA*. For the Sec- and Tat-specific secretion the regulation of different chaperones (e.g. GroES, GroEL and FkpA), components of the bacterial translocase (e.g. SecY, SecE, Tat D) and genes potentially involved in the secretion-mechanism (e.g. *mltE*, *imp*) were observed. Coexpression-strategies of translocase-components and chaperones aiming at an enhanced secretion were proved to be successful in many cases. Therefore, these regulated genes were highly promising candidates for optimisation strategies. For the identification of similar expression-profiles and functional interactions, the regulated genes were clustered using k-means algorithm. The verification of the microarray-data with quantitative real-time PCR and northern blotting was performed exemplarily by a low-density data-set. In addition the DNA-microarray technology was not only used but optimized by establishing a new innovative microarray-detection system based on quantum-dot fluorophores and an array-based system for absolute quantification.

Keywords: protein secretion, sec-pathway, tat-pathway, bacteriocin release proteins, Quantum Dots, DNA-microarray technology, clustering, bioreactor cultivations

1 Einführung

Die Herstellung von rekombinanten Proteinen für Forschung, industrielle Biotechnologie und medizinische Applikationen gewinnt zunehmend an Bedeutung und wirtschaftlicher Relevanz. Allein der Umsatz von gentechnisch erzeugten Arzneimitteln betrug 2005 2,06 Mrd. Euro, dies entspricht einem Anteil an allen Wirkstoffen von 3,6 Prozent. Das Marktvolumen für Enzyme ist in den letzten Jahren um 50 Prozent gewachsen und die Prognosen namhafter Wirtschaftsinstitutionen (Festel Capital 2004, McKinsey 2003) sagen für die weiße oder industrielle Biotechnologie weitere hohe Wachstumsraten voraus. Besonders unter dem Aspekt nachhaltiger industrieller Produktion werden biotechnologische Prozesse weiter an Bedeutung gewinnen und haben in einigen Bereichen bereits die Marktführerschaft erlangt (Positionspapier der DECHEMA: Weiße Biotechnologie (2004)). Seit dem prominenten Beispiel der rekombinanten Insulin-Produktion in *Escherichia coli* im Jahre 1978 wurden über 70 rekombinante Proteine für die humane Medikation in der BRD zugelassen (Stand 2005). Nach wie vor stellt das gram-negative Bakterium *Escherichia coli* einen der wichtigsten Wirtsorganismen zur Produktion rekombinanter Proteine dar. Neben einer detailliert erforschten Genetik liegen die Gründe dafür in einer hohen Produktivität, der Toleranz gegenüber vielfältigen Kultivierungsbedingungen sowie der Existenz etablierter und effizienter Aufreinigungsmethoden (Holland *et al.* (1990), Rinas *et al.* (2004)). Konventionelle cytoplasmatische Expressionsstrategien weisen gegenüber der Sekretion rekombinanter Proteine ins Kulturmedium oder Periplasma gram-negativer Prokaryonten entscheidende Nachteile auf. Neben einem signifikant vereinfachten *downstream processing* ist die Proteaseaktivität im extrazellulären Raum deutlich reduziert. Zudem verhindern die im Cytoplasma vorherrschenden reduzierenden Bedingungen oftmals eine korrekte Ausbildung von Disulfidbrücken, welche für die Proteinfaltung entscheidend sein können. Besonders für Wachstumsfaktoren und Antikörperfragmente besitzt dieses Faktum Relevanz (Fernandez *et al.* (2000)). Die Herstellung mancher rekombinanter Proteine wird durch Sekretionsstrategien erst ermöglicht. Dies gilt für Fälle, in denen das cytoplasmatisch lokalisierte Protein für den Wirt toxisch ist oder einem schnellen proteolytischen Abbau unterliegt (Binet *et al.* (1997)). Zudem tritt bei der prokaryontischen Expression einiger eukaryontischer Proteine, bedingt durch die N-terminale Methionin-Elongation, eine drastische Reduktion ihrer biologischen Aktivität auf (Qui *et al.* (1998)). Aufgrund der post-sekretorischen Prozessierung des Sekretionssubstrates durch die Signalpeptidasen kann diese Problematik ebenfalls umgangen werden. Manche Sekretionspfade diskriminieren aktiv zwischen korrekt gefaltetem und fehlgefaltetem Substrat und sezernieren somit ausschließlich die biologisch aktive Form des rekombinanten Proteins. Dies entspricht einer drastischen Steigerung der spezifischen Aktivität und ist besonders für eine biochemische Charakterisierung und für die Strukturaufklärung von immenser Bedeutung (Pedelacq *et al.* (2002), Goulding *et al.* (2003)). Zusammengefasst erhöhen Sekretionsstrategien häufig die Ausbeute an biologisch aktiven Protein bei drastisch vereinfachten Aufreinigungsbedingungen. In *E. coli* existieren zwei Hauptsekretionspfade Sec und Tat, welche Substrate mit einer spezifischen N-terminalen Signalsequenz ins Periplasma translokieren (Mori *et al.* (2001), Palmer *et al.* (2003), Brüser *et al.* (2007)). Um eine extrazelluläre Lokalisierung der ins Periplasma gelangten Proteine zu realisieren hat sich die Coexpression sogenannter *bacteriocin release proteins* (BRPs) als besonders vorteilhaft erwiesen (Abrahamsen *et al.* (1986), Rahman *et al.* (2005)). Unglücklicherweise unterliegen die natürlichen Sekretionspfade in *E. coli* einer Kapazitätslimitierung, da das Bakterium im natürlichen Habitat nur geringe Mengen und wenige Substrate sekretiert. Aus diesem Grund ist ein de-

tailliertes Verständnis der komplexen molekularen Vorgänge und Interaktionen, denen der Sekretionsprozess unterliegt, wünschenswert. Zur Steigerung der Sekretionseffizienzen haben sich verschiedene Optimierungsstrategien als erfolgreich erwiesen. Diese beinhalteten die Co-expression von Komponenten der bakteriellen Translokase oder Chaperonen (Schaffner *et al.* (2001), Li *et al.* (2003), Perez-Rodriguez *et al.* (2007)). Bis dato existieren derartige Optimierungsstrategien hauptsächlich auf Basis empirischer und unsystematischer Evaluierungen im Rahmen proteinanalytischer Studien. Erstmals sollen in dieser Arbeit die der Proteinsekretion zugrunde liegenden komplexen biochemischen Interaktionen in der Zelle ganzheitlich und systematisch auf globaler Transkriptomenebene erfasst werden. Aufgrund des hochparallelen und flexiblen Analysenformats finden dabei Microarray-basierte Detektionssysteme als Schlüsseltechnik Anwendung. Die transkriptionellen Analysen basieren auf der Verwendung unterschiedlicher Sekretionsstämme, welche für verschiedene Reporterproteine des Sec- und Tat-Pfades codieren. Die extrazelluläre Lokalisierung der Reporterproteine wird mittels Coexpression der angesprochenen BRPs realisiert. Basierend auf den Daten von *whole-genome* Experimenten soll ein applikationsspezifischer *low-density* Microarray designed und etabliert werden. Für die Interpretation der komplexen Microarray-Daten soll neben einer Datenbankanalyse ein Daten-*Clustering* durchgeführt werden. Clusteranalysen stellen ein wichtiges Werkzeug zur Identifizierung ähnlicher Expressionsverläufe oder Funktionalitäten von Genen dar. Anhand der generierten Daten sollte es möglich sein, systematische Optimierungsstrategien zur Generierung von Sekretionsstämmen mit einer erhöhten Sekretionseffizienz zu etablieren. Im zweiten Teil der Arbeit ist der Fokus auf neue Detektions- und Applikationsmöglichkeiten für DNA-Microarrays gerichtet.

2 Theoretische Grundlagen

2.1 Einführung

Das gram-negative Enterobakterium *Escherichia coli* (*E. coli*) ist von zwei Zellmembranen umgeben, der äußeren Membran welche, das Bakterium von seiner Umwelt separiert, und der inneren Cytoplasmamembran. Zwischen diesen liegt die strukturgebende Komponente des Peptidoglykans, eine dünne Schicht, die hauptsächlich aus den zwei Zuckerderivaten N-Acetylglucosamin und N-Acetylmuraminsäure sowie einer kleinen Gruppe von Aminosäuren besteht. Der sich aufgrund der Lokalisation der beiden Membranen ergebende Zwischenraum wird Periplasma genannt. Die Sekretion rekombinanter Proteine in das Kulturmedium hat daher obligat durch diese zwei Membranen zu erfolgen, welche sich hinsichtlich Struktur und Zusammensetzung substantiell unterscheiden. Während des Translokationsprozesses muss dabei die Integrität und Funktionalität der Membran gewahrt bleiben. In *E. coli* sind bis dato fünf verschiedene natürliche Sekretionstypen bekannt, von denen aber lediglich die Typen I und II für die Sekretion rekombinanter Proteine von Bedeutung sind. Beim Typ I erfolgt die extrazelluläre Sekretion in einem Schritt durch beide Membranen, wohingegen bei der Translokation von Typ II ein periplasmatisches Intermediat auftritt (Mergulaho *et al.* (2005)).

2.2 Typ I Sekretionsmechanismus

Der Typ I Sekretionsmechanismus wird von gram-negativen Bakterien zur Sekretion hochmolekularer Toxine und Exoenzyme verwendet. Die Exportmaschinerie besteht aus den

beiden inneren Membranproteinen HlyB und HlyD, welche zur Familie der Transporter der ABC-bindenden Kassette gehören und einem äußeren Membranprotein TolC (Blight *et al.* (1994)). Die Translokation kann jedoch auch von anderen Sekretionspathways beeinflusst werden, im Besonderen soll an dieser Stelle auf den SecB-abhängigen Sekretionspathway verwiesen werden. Der mit Abstand bedeutendste und am häufigsten verwendete Transporter ist das α -Haemolysin-System (HlyA) (Fernandez *et al.* (2000)). Die C-terminale Region von HlyA enthält die notwendige Information für die Translokation des Proteins und wird dem gewünschten rekombinanten Produkt angehängt, welches somit von der Exportmaschinerie als Substrat erkannt wird. Die beiden ABC-Proteine lagern sich unter ATP-Aufnahme zu einem stabilen Komplex zusammen und binden an das cytoplasmatisch lokalisierte Substrat mit der C-terminalen Signalsequenz. Ein TolC-Trimer bildet einen Translokationskanal (Länge 14 nm, Durchmesser 3,5 nm) der beide Membranen verbindet und sich dabei an den Komplex aus HlyB, HlyD und dem rekombinanten Protein bindet. Für die Translokation des Substrates durch den Kanal ist ATP-Hydrolyse notwendig, nach erfolgtem Transport dissoziiert TolC vom HlyB/HlyD-Komplex, wobei die Membranen wieder voneinander separiert werden. Während des Transports durch den Tunnel erfolgt die Ausbildung von Disulfidbrücken, welche sich unabhängig von den membrangebundenen Dsb-Enzymen vollzieht (Gentshev *et al.* (2002)). Rekombinante Proteine im Größenbereich von 50 bis 4000 Aminosäuren wurden erfolgreich mittels dieses Pathways sekretiert, die Abmessungen des Translokationstunnels erlaubt auch die Sekretion teilweise gefalteter Substrate. Bis zu fünf Prozent rekombinantes Protein bezogen auf den Gesamtproteingehalt der Zelle lassen sich so sekretorisch darstellen (Mergulaho *et al.* (2005)). Trotz dieser Vorteile weist die Verwendung dieses Sekretionspfades zwei entscheidende Nachteile auf. So wird die C-terminale Signalsequenz nicht abgespalten, was einen zusätzlichen proteolytischen Schritt zur Produktgewinnung erforderlich macht. Des Weiteren ist es oftmals nötig, durch Coexpression der Komponenten des Sekretionspfades die Transportkapazität zu erhöhen. Trotzdem lässt sich zumeist nur durch die Realisierung niedriger Proteintranslationsraten eine Sättigung der Exportkapazitäten umgehen (Fernandez *et al.* (2000), Gentshev *et al.* (2002)).

2.3 Typ II Sekretionsmechanismus

Bei dem Sekretionsmechanismus vom Typ II handelt es sich um einen zweistufigen Prozess. Für die Translokation durch die Cytoplasmamembran sind drei Pfade bekannt:

1. SecB-abhängiger Pathway
2. Signal recognition particle (SRP)-Pathway
3. Twin arginine translocator (Tat)-Pathway

In dem zweiten Schritt der extrazellulären Sekretion, der Translokation durch die äußere Membran, ist eine spezifische Proteinmaschinerie involviert, welche Sekretion genannt wird. Daneben besteht aber für die rekombinante Proteinproduktion die Möglichkeit sich einer unspezifischen Sekretion des periplasmatischen Inhalts zu bedienen. Hierfür stehen verschiedene chemische und physikalische Methoden zur Verfügung. Besonders der Einsatz von sogenannten *bacteriocin release proteins* soll an dieser Stelle erwähnt werden. Abbildung 1 gibt eine Übersicht über den Typ II-Sekretionsmechanismus.

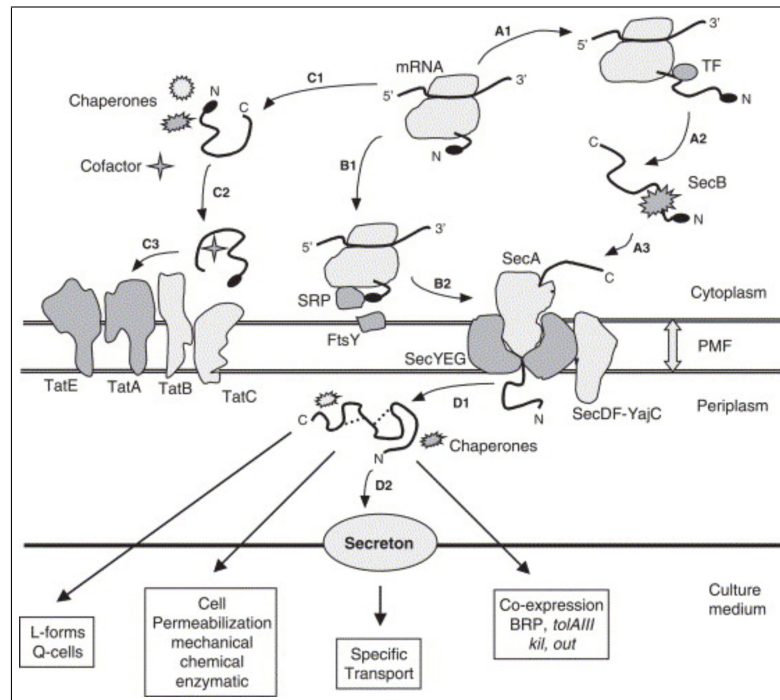


Abb. 1: Translokation rekombinanter Proteine durch die Cytoplasmamembran nach dem Typ II Sekretionsmechanismus. Aus Mergulaho *et al.* (2005).

2.3.1 SecB-abhängiger Sekretionspfad

Die große Mehrheit der von *E. coli* sekretierten Proteine wird mittels dieses Pathways transloziert, daher wird in Sekretionsstrategien zur rekombinanten Proteinproduktion zumeist dieser Pfad verwendet. Die zentralen Membrankomponenten SecYEG fungieren auch beim SRP-Pathway als Translokationstunnel durch die Cytoplasmamembran (siehe auch Abbildung 1). Substrate des Sec-Pathways besitzen eine dreigliedrige, N-terminale Signalsequenz von 18-30 Aminosäuren Länge. Sie besteht aus einem positiv geladenen N-Terminus (N-Region), einem zentralen hydrophoben Teil (H-Region) und einer polaren, helixbrechenden Region (C-Region), welche die Schnittstelle für die prozessierende Signalpeptidase enthält (Christobal *et al.* (1999), Economou *et al.* (1999)). Aufgrund der positiven Ladung des N-Terminus wird vermutet, dass dieser Teil der Signalsequenz von der negativ geladenen Phospholipiddoppelschicht eine elektrostatische Anziehung erfährt und daher dazu fungiert, das Präprotein zur Membran zu dirigieren. Diese Vermutung wurde durch Mutationsstudien bestätigt, welche eine erhöhte Translokationsrate als Folge einer Ladungserhöhung der N-Region nachweisen konnten. Die H-Region umfasst 7-15 Aminosäuren. Es konnte gezeigt werden, dass eine Erhöhung der Hydrophobizität mit einer Erhöhung der Sekretionseffizienz einhergeht und dass eine minimale Hydrophobizität notwendig ist, um eine Proteinsekretion zu vermitteln. Das entstehende, noch ribosomal assoziierte Sekretionsprotein wird von einem sogenannten *Trigger*-Faktor gebunden (Beck *et al.* (2000)). Diese Bindung verhindert, dass das zu sekretierende Protein mit Komponenten des SRP-Pathways interagiert, und sie wird aufrecht erhalten, bis das Präprotein das Ribosom verlässt (Maier *et al.* (2003)). Durch das Chaperon SecB wird das Präprotein in einem Transport-kompetenten Zustand gehalten. Indem sich das Chaperon an die reifende Region des Präproteins anlagert, wird seine

vollständige Faltung verhindert (Ullers *et al.* (2004)). In Gegenwart vom Präprotein bindet SecB an SecA, die periphere Membrankomponente der Translokase. Die Wechselwirkung von SecA mit dem zu sekretierenden Protein wird durch die Signalsequenz erleichtert, welche von SecA spezifisch erkannt wird. Die Bindung von ATP an eine der zwei ATP-bindenden Domänen von SecA führt zur Freisetzung von SecB aus der Membran. Die Formierung der Sec-Komponenten zu einer funktionsfähigen Translokase ist dabei weiterhin unklar, es wurden sowohl mono-, di-, tri- und oligomere Formen postuliert (Breyton *et al.* (2002), Yahr *et al.* (2000)). Die Bindung vom Präprotein an membrangebundenes SecA führt zur Translokation von ca. 20 Aminosäuren. Nachfolgende Bindung von ATP ermöglicht den Zutritt von weiteren 15-20 Aminosäuren in den Translokationskanal. Für die vollständige Translokation des zu sekretierenden Proteins kommen zwei Arten von freier Energie zur Verwendung: ATP-Hydrolyse und protonenmotorische Kraft. Im Translokationstunnel wird das Präprotein unter ATP-Hydrolyse von SecA freigesetzt, welches selber aus der Membran deinsertiert und durch cytosolisches SecA ersetzt werden kann. Sukzessive Runden von ATP-Hydrolyse und Insertion/Deinsertion von SecA translokieren das Präprotein durch die Membran. Sobald das Präprotein zur Hälfte durch den Sekretionstunnel gelangt ist, vermag die protonenmotorische Kraft die Translokation zu komplettieren (auch in Abwesenheit von SecA) (Driessen *et al.* (2001)). Der limitierende Schritt bei der Sekretion mittels des Sec-Pfades und damit auch der *bottleneck* für viele Produktionsprozesse ist die SecA-Freisetzung aus der Membran. Bei einer Blockade des Exportes wird eine Erhöhung der SecA-Synthese beobachtet, weshalb Optimierungsstrategien häufig auf eine Erhöhung der cytosolisch lokalisierten SecA-Konzentration abzielen (Tziatzios *et al.* (2004)). Unglücklicherweise bindet SecA allerdings an seine eigene mRNA, was die Expression reprimiert. Neben dieser Selbstkontrolle der Expression verkompliziert eine weitere Expressionskontrolle, diesmal auf der Translationsebene, den Erfolg einfacher Coexpressions-Strategien erheblich. Die Translation von SecA wird durch das Produkt eines stromaufwärts liegenden sekretorischen Gens (*secM*) kontrolliert. Beide Gene sind im selben Operon organisiert und liegen nach der Transkription auf derselben polycistronischen mRNA vor. Ist die Sekretion von SecM limitiert präsentieren die Ribosomen, welche den SecM-ORF translatieren, die Ribosomenbindedomäne der SecA/SecM-mRNA und stimulieren so die SecA-Translation. Ist hingegen genügend Exportkapazität vorhanden und wird SecM effektiv translokiert, so wird die SecA-Translation reprimiert (Sarker *et al.* (2002)). Um diesem Umstand zu begegnen, wurden Strategien entwickelt, die eine effektive SecM-Sekretion unterbinden. So verhindern zum Beispiel Mutationen im Signalpeptid von SecM die Translokation, wodurch die SecA-Translation konstitutiv stattfindet. Allerdings hat sich gezeigt, dass die Expression von Signalpeptid-mutierten SecM extrem toxisch für die Zelle ist (Nakatogawa *et al.* (2002)). Eine Produktion von einem artifiziellen *secA*-Gen, isoliert von *secM*, ist ebenfalls häufig keine adäquate Lösung, da die Expression weniger effektiv verläuft als vom natürlichen *secAM*-Operon (Nakatogawa *et al.* (2004)). Um die limitierende Translokationskapazität zu erhöhen, hat die Verwendung sogenannter Protein-Lokalisationsmutanten (*prl*-Mutanten) oftmals erstaunliche Erfolge gezeigt. Dabei handelt es sich um Mutationen im *secY* (*prlA*), *secE* (*prlG*), *secG* (*prlH*) oder *secA* (*prlD*)-Gen, also jenen Genen deren Proteinprodukte den Translokationskanal bilden. Es wird angenommen, dass die Mutationen den Tunnel entspannen und so auch für Substrate mit defekten Signalsequenzen passierbar machen. Die bisher effektivste Mutante wurde im *secY*-Gen gefunden (*prlA4*) und führte zu einer zehnfach höheren Sekretion des rekombinanten Produktes verglichen mit dem Wildtyp (van der Wolk *et al.* (1998)). Trotz der vielen Vorteile und Möglichkeiten einer Verwendung des Sec-Pfades für die extrazelluläre Produk-

tion von rekombinanten Proteinen gibt es auch deutliche Limitierungen in der Anwendbarkeit. So geschieht die Sekretion mittels des Sec-Pfades zu großen Teilen posttranslational und sich im Cytoplasma schnell faltende Proteine können nicht exportiert werden (Palmer *et al.* (2003), De Lisa *et al.* (2004)). Die korrekte Faltung und damit 3D-Struktur ist zudem für viele Cofaktor-tragende Enzyme von substantieller Bedeutung. Auch für diese Art von Proteinen kommt eine Sekretion mittels des Sec-Pfades nicht in Betracht. Derartige Proteine lassen sich mittels des *twin-arginine translocator* (Tat)-Pathways exportieren, welcher nachfolgend in Abschnitt 2.3.3 vorgestellt wird.

2.3.2 SRP-Pathway

Der *signal recognition particle* (SRP)-Pathway nutzt ebenfalls den membrangebundenen Teil der Sec-Translokase und soll daher an dieser Stelle nur kurz Erwähnung finden. Eine N-terminale Signalsequenz mit stark hydrophobem Kern dirigiert Substrate zu diesem Sekretionspfad. Der einzige Unterschied zur SecB-abhängigen Signalsequenz besteht in der fehlenden Bindestelle für den *Trigger*-Faktor (siehe Abschnitt 2.3.1). Eine erhöhte Hydrophobizität der Signalsequenz erhöht auch beim SRP-Transport die Effizienz (Peterson *et al.* (2003)). Kotranslational zur der am Ribosom entstehenden Proteinkette wird dieselbe vom Faktor Ffh, assistiert von einer 4.5S rRNA, gebunden. Die Bindestelle für Ffh mit der am Ribosom entstehenden Proteinkette überlappt mit der Bindestelle für den *Trigger*-Faktor, weshalb ein diskriminierender Mechanismus propagiert wurde (Beck *et al.* (2000)). Mechanistisch wird davon ausgegangen, dass SRP- und *Trigger*-Faktor des Sec-Pfades gleichzeitig um die Bindestelle am Ribosom konkurrieren und gebunden bleiben bis die reifende Proteinkette sich vom selben löst. Abhängig von den Eigenschaften des Peptides wird entweder die Bindung von SRP oder des *Trigger*-Faktors stabilisiert. Das Protein wird entweder posttranslational via SecB-Pfad oder als kompletter Ribosom-SRP-Komplex von FtsY zur Membran dirigiert. Während der Interaktion mit den Phospholipiden der Membran wird die GTPase Aktivität von FtsY und Ffh stimuliert. Als Folge davon wird die reifende Proteinkette in den Translokationstunnel freigesetzt. Der Translokationstunnel wird meistens, wie eingangs erwähnt, von den Sec-Komponenten SecYEG gebildet (Tziatzios *et al.* (2004)). Einige Autoren berichten jedoch auch über die Nutzung anderer Translokationsporen (Yahr *et al.* 2000)). Ein SRP-*Targeting* kann durch eine Erhöhung der Hydrophobizität der Signalsequenz erreicht werden. Vorteilhaft gegenüber einer Translokation via des SecB-abhängigen Pfades kann die Tatsache sein, dass der SRP-Pfad in der Lage ist, teilweise gefaltete Proteine zu translokieren (Beck *et al.* (2000)). Sich im Cytoplasma schnell faltende Proteine sind inkompatibel mit einer Sec-Translokation, allerdings ist die Sekretion vollständig gefalteter Proteine nur mittels des Tat-Pfades möglich.

2.3.3 Tat-Pathway

Parallel zum Sec-Translokation und unabhängig von diesem existiert in Bakterien (ein analoges System findet sich auch im Thykaloidlumen der Chloroplasten höherer Pflanzen) ein weiterer Sekretionspfad, welcher als *twin-arginine translocator* (Tat)-Pathway Eingang in die Literatur gefunden hat. Der Name leitet sich von zwei aufeinanderfolgenden, hochkonservierten Argininen in der Signalsequenz der Tat-Substrate her. Dieser Sekretionspfad weist die einzigartige Eigenschaft auf, unabhängig von ATP vollständig gefaltete Proteine durch die Cytoplasmamembran zu translokieren. Natürliche Substrate dieses Pfades sind

z.B. Proteine, die in der respiratorischen Energiegewinnung des Bakteriums über Elektronentransferkaskaden involviert sind (Sanders *et al.* (2001)). Solche Proteine gewinnen ihre katalytischen Aktivitäten durch Bindung bestimmter Cofaktoren im Cytoplasma, welches eine vorherige vollständige Faltung voraussetzt. Auch müssen diese Substrate obligat gefaltet translokiert werden, um vollständig biologisch aktiv zu sein (Berks *et al.* (2000)). Ähnlich dem Sec-Pathway besitzen Substrate, die mittels des Tat-Pfades translokiert werden, eine N-terminale, dreigliedrige Signalsequenz als Sekretionssignal. Allerdings weist diese eine größere Länge (25-50 Aminosäuren), hauptsächlich durch einen ausgedehnteren N-Terminus verursacht, und eine geringere Hydrophobizität in der zentralen H-Region auf. Eine positive Ladung in der C-Region wie sie bei vielen Tat-Signalsequenzen auftritt, allerdings noch nie bei Sec-Signalpeptiden beobachtet wurde, fungiert mit großer Wahrscheinlichkeit als Ausschlusskriterium für den Sec-Transport (Blaudeck *et al.* (2001)). Dieses konnte auch dadurch gezeigt werden, dass Sec-Signalpeptide ihre Fähigkeit verloren, ihr Protein zu exportieren, wenn eine Substitution durch basische Aminosäuren in dieser Region vorgenommen wurde. An der Grenze zwischen der C- und der H-Region befindet sich das schon eingangs erwähnte hochkonservierte Argininpaar, welches Teil einer Consensussequenz S-R-R-X-F-L-K ist. Eine polare Aminosäure (S), eine variable Position X und zwei hydrophobe Aminosäuren flankieren die Arginine, wobei die hydrophoben Positionen essentiell für eine effektive Translokation sind. Die helixbrechende C-Region enthält, analog den Sec-Signalpeptiden, die Erkennungssequenz für die Prozessierung durch die Signalpeptidase nach erfolgter Translokation. Die Positionen -1 und -3 relativ zur Schnittstelle der Peptidase werden dabei von kleinen, ungeladenen Aminosäuren besetzt (Alami *et al.* (2003)). Zwei genetische Loci codieren in *E. coli* für die Komponenten der Tat-Maschinerie; das TatA-Operon, welches die vier Gene *tatABCD* enthält und das monocistronische *TatE*. Die konstitutive Expression der Tat-Komponenten deutet darauf hin, dass jene essentiell unter allen Wachstumsbedingungen sind. *TatA*, *TatB* und *TatE* weisen eine ähnliche Struktur auf. Sie besitzen am N-Terminus eine Transmembran-Helix, gefolgt von einer amphiphilen Helix, welche mit der cytoplasmatischen Seite der Membran assoziiert ist, und schließlich einen in das Cytoplasma hineinragenden C-Terminus variabler Länge (Robinson *et al.* (2004)). Deletionsanalysen haben ergeben, dass der C-Terminus nicht essentiell für die Translokation ist. *TatC* hingegen weist eine unterschiedliche Struktur auf, es handelt sich dabei um ein stark hydrophobes Protein mit sechs Transmembran-Helices sowie jeweils auf der cytoplasmatischen Seite der Membran lokalisierten N- und C-Terminus. Zwischen unterschiedlichen Spezies weist *TatC* die größte Aminosäurenkonservierung auf, die meisten davon in den cytoplasmatischen Loops des Proteins. Dieser Befund ist konsistent mit der Beobachtung, dass *TatC* sehr wahrscheinlich die spezifische Determinante der Substraterkennung ist (Tullman-Ercek *et al.* (2007)). Das Gen *tatD* codiert für ein cytoplasmatisches Protein mit Nukleaseaktivität, welches allerdings keine obligate Rolle im Tat-Export spielt. Für eine funktionelle Translokase sind in *E. coli* mindestens die drei Tat-Komponenten A,B,C notwendig. *TatA* und *TatE* weisen Sequenzhomologien von über 50 Prozent auf und übernehmen überlappende Funktionen. Allerdings wird *TatA* ca. 50-100 fach stärker exprimiert als dies für *TatE* der Fall ist, weshalb auch *tatE*-Mutanten eine sehr viel höhere Sekretion als *tatA*-Mutanten zeigen (Mergulaho *et al.* (2005)). In anderen Spezies liegen häufig auch nur *TatAC*-Systeme vor, *TatA* übernimmt hier die Funktion von *TatB*, da beide zur gleichen Klasse von Proteinen gehören. Auch in *E. coli* besitzt ein *TatAC*-System noch eine gewisse Restaktivität bei artifiziellen Tat-Substraten. Alle Tat-Komponenten zeigen eine hohe Tendenz zur Oligomerisierung und Selbstassoziation. *TatB* bildet mit *TatC* eine multimere Struktur, dieser Komplex bindet

die Signalpeptide von Tat-Substraten. Mittels blauer, nativer PAGE lässt sich dieser Komplex als zwei Banden bei 440 und 580 kDa im Gel sichtbar machen, was darauf hindeutet, dass es sich entweder um zwei verschiedene Komplexe oder um einen Komplex mit zwei verschiedenen Konformationen handelt (Brüser *et al.* (2007)). Die Substratbindung induziert den Einbau von oligomeren TatA in den TatBC/Substrat-Komplex. TatA bildet dabei die Translokationspore, welches die Translokation des Substrates, angetrieben ausschließlich vom pH-Gradienten über der Membran, initiiert. Die Signalsequenz bleibt während der Translokation an TatC gebunden, TatC fungiert dabei, gemäß der Hypothese von Brüser und Sanders, als Motor. Nach erfolgter Translokation wird das Signalpeptid von einer Signalpeptidase abgespalten und das reife, biologisch aktive Protein liegt an seinem Wirkungsort vor. Mittels des Tat-Pfades wurden rekombinante Proteine als Monomere, Homo- und Heterooligomere, auch mit mehr als einer Signalsequenz, im Größenbereich von 9-142 kDa erfolgreich sekretiert (Berks *et al.* (2000)). Dies zeigt die weite Spanne von Dimension und Form der Proteine auf, welche mittels des Tat-Pfades exportiert werden können. Was den Tat-Pathway so außerordentlich interessant für biotechnologische Applikationen macht, ist die Tatsache, dass fast nur gefaltete Substrate translokiert werden (Palmer *et al.* (2003)). Durch diese Diskriminierung vor dem Export ist die Gewinnung von fast ausschließlich biologisch aktivem Protein möglich, was einem Maximum an spezifischer Aktivität des gewonnenen Produktes entspricht. Eine streng diskriminierende Sekretion hinsichtlich des Faltungskriteriums hängt allerdings auch immer im hohen Maße von dem zu sekretierenden Substrat ab (De Lisa *et al.* (2004), Blaudeck *et al.* (2001)). Ein inhärentes Problem für die industrielle Applikation des Tat-Pfades liegt in der geringen Exportkapazität desselben begründet. Die Kombination aus langsamer Translokation (Halbwertszeiten im Bereich einiger Minuten anstelle Sekunden für den Sec-Pfad) und geringer Menge von Tat-Translokase pro Zelle führen zur schnellen Saturierung der Tat-Maschinerie (De Lisa *et al.* (2004)). Coexpressionsstrategien der Tat-Komponenten haben in vielen Fällen zu einer Erhöhung der Sekretionseffizienzen geführt. Jedoch zeigen hohe Translokationsraten einen negativen Effekt auf die Integrität der Membran, unter diesen Bedingungen wird ein hoher transmembraner Protonenfluss beobachtet. Dieses wiederum führt zu einer signifikanten Verringerung des pH-Gradienten und damit zur Absenkung der Triebkraft der Translokation. Coexpression der Tat-Komponenten erhöht daher die Sekretion nur in einem gewissen Maße, die verringerte protonenmotorische Kraft (PMK) wird dann limitierend. Dieses Dilemma lässt sich jedoch durch die Coexpression von PspA umgehen, einem Stress-Faktor, der die Ionendurchlässigkeit der Cytoplasmamembran minimiert. Mit solchen Optimierungsstrategien konnte eine signifikant höhere Exportrate nachgewiesen werden (De Lisa *et al.* (2004)). Weitere Optimierungsmöglichkeiten für die sekretorische Darstellung von rekombinanten Proteinen mittels des Tat-Pathways ergeben sich durch:

1. Wahl des geeigneten Signalpeptides
2. Coexpression von Chaperonen
3. Verwendung heterologer Translokons

Vom Prinzip kann jedes Protein, das mit einer entsprechenden zugänglichen Signalsequenz fusioniert wurde, von der Tat-Translokase exportiert werden. Jedoch ist nicht jedes Protein kompatibel mit dem Translokationsprozess. Dies kann zum Beispiel dann der Fall sein, wenn zur Faltung die Ausbildung von Disulfidbrücken notwendig ist. Im reduzierenden Mi-

lieu des Cytoplasmas ist dieses oftmals unmöglich (Tullman-Ercek *et al.* (2007)). Ein Ausweg kann die Verwendung von *trx/gor E. coli*-Stämmen sein, welche die Ausbildung von Disulfidbrücken im Cytoplasma ermöglichen (Oden *et al.* (1994), gezeigt bei verschiedenen Antikörperfragmenten, alkalischer Phosphatase und weiteren Proteinen). Tat-Signalpeptide alleine sind meistens keine spezifischen Determinanten für eine Tat-Sekretion, da sie oftmals auch eine Translokation mittels des Sec-Pfades erlauben (Blaudeck *et al.* (2001)). Die Spezifität eines Signalpeptides hinsichtlich des Tat-Pfades muss immer in Kombination mit dem zu sekretierenden Protein betrachtet werden. Ein einfacher Test, ob eine Sekretion ausschließlich mittels des Tat-Pfades durchgeführt wird, stellt die Verwendung von Tat-Mutantenstämmen dar, welche keine funktionsfähige Tat-Translokase besitzen. Tritt unter diesen Bedingungen keine Sekretion auf, so lässt sich vermuten, dass der Export ausschließlich nach dem Tat-Mechanismus abläuft. Ein paar generelle Tendenzen können dennoch konstatiert werden. Je schneller sich die Faltung des Proteins im Cytoplasma vollzieht und je basischer die Aminosäuren in der C-Region sind, desto spezifischer ist die Translokation mittels des Tat-Pfades. Auch die Nettoladung der N-terminalen Region des reifen Proteins, welche sich unmittelbar an das Signalpeptid anschließt, spielt eine Rolle in der Tat-Spezifität. Sofern möglich sollte eine Fusionierung des rekombinanten Proteins nicht nur mit dem Signalpeptid allein, sondern zusammen mit den ersten Aminosäuren des natürlichen Substrates, von dem es abstammt, erfolgen. Bei allen Optimierungsbestrebungen hinsichtlich der Signalsequenz des rekombinanten Peptides gilt zu beachten, dass die fusionierte Signalsequenz auch zur Fehlfaltung des Proteins führen kann. In Zweifelsfall sind daher verschiedene Signalpeptide für den anvisierten Sekretionspfad zu testen. Eine oft getroffene Klassifizierung der Signalpeptide in solche, die ausschließlich eine Translokation mittels des Tat-Pfades ermöglichen, und solche, die auch mit dem Sec-Pfad kompatibel sind, erscheint unter diesen Gegebenheiten schwierig bis unmöglich (Tullman-Ercek *et al.* (2007)).

Wahrscheinlich binden viele, wenn nicht alle Tat-Signalpeptide Chaperone, die zur Faltung von Signalpeptid und/oder Substrat notwendig sind. Akkumuliert nun aufgrund limitierender Transportkapazitäten das rekombinante Protein im Cytoplasma, so werden die Chaperone schnell zum limitierenden Faktor und die Signalpeptide sensitiv gegenüber Proteasen, welche Tat-Substrate zu reifer Größe oder gänzlich degradieren (Brüser *et al.* (2007)). Daher erhöhen coexprimierte Chaperone die Menge an sekretiertem rekombinantem Protein oftmals in signifikanter Weise. Als Beispiele seien hier die Coexpression von *torD* bei dem Sekretionskonstrukt *torAGFP* sowie einige Beispiele der Coexpression der allgemeinen Chaperone *DnaK* und *SlyD* genannt (Perez-Rodriguez *et al.* (2007), Li *et al.* (2003)).

Als letzte Optimierungsmöglichkeit sei auf die Verwendung heterologer Translokons hingewiesen. So hat sich z.B. das *TatAC-Translocon* von *Bacillus subtilis* bei Tat-defizienten *E. coli*-Stämmen in vielen Fällen als voll funktionsfähig und dem wirtseigenen Translokon als überlegen erwiesen (Yamabhai *et al.* (2008)). Im Einzelfall muss aber darauf geachtet werden, dass für die korrekte Faltung der rekombinanten Proteine auch Wirtsfaktoren nötig sein können, die in *E. coli* nicht existent sind. Die Translokationseffizienz eines gegebenen Substrates variiert stark bei Verwendung verschiedener Tat-Translokons. Zudem können auch Inkompatibilitäten auftreten, was ein individuelles *Screening* oft unverzichtbar macht.

Zusammenfassend zeigt dieser kurze Überblick über die Proteinsekretion mittels des Tat-Pfades, welche vielversprechenden Möglichkeiten und Eigenschaften dieser Weg zur extracytoplasmatischen Darstellung rekombinanter Proteine aufweist. Trotz aller Limitierungen zeigen die vielfältigen Lösungen und Optimierungsmöglichkeiten, dass die Anwendbarkeit auf biotechnologische Fragestellungen hier erst am Anfang steht. Durch Intensivierung der

Erkenntnisse über den Tat-Pathway lässt sich sicherlich eine signifikante Steigerung der Sekretionseffizienzen und eine Lösung etlicher Sekretionsprobleme erzielen. Dies gilt gerade in Hinsicht auf schwierig zu translokierende Produkte.

2.4 Extrazelluläre Produktion

E. coli sekretiert unter normalen Umweltbedingungen keine Proteine in den extrazellulären Raum. Eine Konsequenz daraus ist, dass im Kulturmedium keine Kontamination mit Wirtskomponenten vorliegt und auch die Proteaseaktivität minimal ist (Kleist *et al.* (2003)). Neben dem Sekretionsmechanismus vom Typ I kann eine extrazelluläre Darstellung rekombinanter Proteine durch unspezifisches Auslaufen des periplasmatischen Inhaltes oder durch den zweiten Schritt vom Typ II-Sekretionsmechanismus realisiert werden. Für letztgenannten Schritt sind 12-16 Proteine notwendig, welche die Bestandteile des Sekretons bilden. Sekretonkomponenten sind hochkonserviert unter verschiedenen gram-negativen Bakterien. Ins Periplasma gelangte Proteine bilden ihre Tertiär- und teilweise sogar Quartärstrukturen aus und werden anschließend von den Sekreton-Komponenten erkannt. Bisher wurden allerdings keine Sekretionssignale bei den gefalteten Proteinen entdeckt, was eine technische Applikation schwierig gestaltet. *E. coli* besitzt einen kompletten Gensatz für das Sekreton, der allerdings unter normalen Laborbedingungen nicht exprimiert wird. Optimierungsstrategien zielen daher auf eine Überexpression der Sekreton-Komponenten und auf der Identifizierung von Umweltbedingungen ab, die deren Expression ermöglichen (Mergulaho *et al.* (2005)). Akkumulierung von rekombinanten Proteinen im Periplasma erzeugt einen beträchtlichen osmotischen Druck und kann somit eine Störung der Membran-Integrität verursachen, was die Permeabilität der äußeren Membran selektiv erhöht. Um die Permeabilität der äußeren Membran für das gewünschte rekombinante Protein zu erhöhen, stehen verschiedene physikalische und chemische Methoden zur Verfügung. Als Beispiele seien hier Ultraschallbehandlung, Variation des pH-Wertes, der Temperatur, der Kulturmedienzusammensetzung und Begasung, Zugabe von Mg^{2+} -Ionen, Ca^{2+} -Ionen, EDTA, Glycin, Triton X-100 oder Lysozym genannt. Ebenso kann die Verwendung von speziellen *E. coli*-Stämmen, sogenannter *leaky-strains*, eine Option für die extrazelluläre Darstellung sein (Gumpert *et al.* (1998)). In diesen Stämmen wird die äußere Membran nur unvollständig ausgebildet oder fehlt im Extremfall zusammen mit dem Periplasma und dem Peptidoglykan vollständig (L-Formen). Allerdings sind derartig mutierte Stämme meist nicht für die biotechnologische Produktion anwendbar, da sie Wachstums-inhibiert und nicht robust genug für Hochdichte-Fermentationen sind. Ein Patent aus neuerer Zeit (Wacker-Chemie 2007) zeigt jedoch einen erfolgreichen Einsatz verschiedener Mutantenstämme in der großtechnischen Produktion. Dieses reflektiert auch die außerordentliche wirtschaftliche Relevanz von Sekretionsstämmen. Daneben stellt der Einsatz von Q-Zellen eine potenzielle Alternative dar. In diesen findet durch Überexpression des plasmid-codierten Zellzyklus-Regulators Rcd praktisch nur die Transkription und Translation der plasmidcodierten Gene statt. Die Ressourcen und Exportkapazitäten der Zelle stehen damit vornehmlich für die rekombinante Proteinproduktion zur Verfügung (Rowe *et al.* (1999)). In vielen Fällen hat sich auch die Coexpression von sogenannten *bacteriocin release proteins (BRP)* als sehr vorteilhaft erwiesen (Miksch *et al.* (2002), Kleist *et al.* (2003)).

2.5 *Bacteriocin release proteins*

Bei den Bacteriocinen wie z.B. der Gruppe der Colicine oder Cloacin DF13, handelt es sich um plasmidcodierte, für *E. coli* und verwandte bakterielle Spezies hochtoxische Proteine. Die exprimierende Zelle schützt sich durch simultane Ausschüttung eines Immunitätsproteins vor den letalen Folgen und erhält so einen Wettbewerbsvorteil gegenüber nichtproduzierenden Zellen. Im natürlichen Habitat weisen ca. 1/3 der *E. coli*-Kulturen colicinogenes Potenzial auf, unter pathogenen *E. coli*-Stämmen, z.B. während einer Infektion, liegt der prozentuale Wert weit höher. Gemäß dem Wirkungsmechanismus der Bacteriocine werden diese in vier verschiedene Gruppen klassifiziert:

1. RNase-Aktivität
2. DNase-Aktivität
3. Porenbildner in der Cytoplasmamembran
4. Inhibitoren der Peptidoglykan- und der LPS-Oberflächenantigensynthese

Die Immunitätsproteine derjenigen Bacteriocine, welche eine Nukleaseaktivität aufweisen, binden an die katalytische Domäne vom Bacteriocin und blockieren so dessen Aktivität. Diese Immunitätsproteine entstehen als äquimolarer Komplex mit dem entsprechenden Bacteriocin. Porenbildende Bacteriocine werden hingegen als Monomer gebildet und ihr Immunitätsprotein ist in der Cytoplasmamembran lokalisiert.

Genetisch liegen Bacteriocine in einem Gencluster mit ihrem Immunitätsprotein und einer dritten Klasse von Proteinen organisiert vor, die für die Freisetzung der Bacteriocine in den extrazellulären Raum verantwortlich sind. Diese werden als *bacteriocin release proteins* (BRPs) bezeichnet (van der Waal *et al.* (1995)). Es handelt sich dabei um kleine, hochexprimierte Lipoproteine mit einer N-terminalen Signalsequenz, die in der äußeren Membran lokalisiert sind. Man unterscheidet 2 Plasmidgruppen, kleine *high-copy* Plasmide, die nicht-transmissierbar sind, und große *low-copy* Plasmide, die hingegen transmissierbar sind. Gene, die für Bacteriocine mit Nukleaseaktivität codieren, stehen zusammen mit ihrem Immunitätsprotein und dem BRP unter der Kontrolle induzierbarer Promotoren, die über die SOS-Antwort der Zelle aktiviert werden können, wie z.B. durch eine UV-Bestrahlung oder Exposition mit Mitomycin C. Das Immunitätsprotein von porenbildenden Bacteriocinen steht hingegen unter der Kontrolle eines konstitutiven Promoters und die Transkription erfolgt in entgegengesetzter Richtung zum Bacteriocin und BRP. Im extrazellulären Raum vorliegende Bacteriocine gelangen über hochaffine Rezeptoren des bakteriellen Aufnahmesystems der äußeren Membran in die Empfängerzelle, wo sie ihre toxische Wirkung entfalten. Diese Rezeptoren fungieren normalerweise für die Aufnahme großer Moleküle wie Vitamin B12 oder Eisensiderophore. Bacteriocine mit unterschiedlichem Wirkmechanismus weisen dennoch die gleiche Rezeptor-Bindedomäne auf, welche in der zentralen Region des Bacteriocins lokalisiert ist. Der N-Terminus des Bacteriocins ist entscheidend für Aufnahme durch die Zellhülle und am C-Terminus befindet sich die katalytische Domäne des Proteins (Oudega *et al.* (1993), van der Waal *et al.* (1995)).

Der Freisetzung von Bacteriocinen in den extrazellulären Raum liegt ein einzigartiger Mechanismus zugrunde, welcher sonst noch nicht bei sekretierten Substraten beobachtet wurde. Durch die kombinierte Wirkung des reifen BRP, seines in der Cytoplasmamembran

akkumulierten stabilen Signalpeptides und der aktivierten, Detergenzien-stabilen Phospholipase A (PldA) der äußeren Membran von *E. coli* wird die Sekretion des Bacteriocines ermöglicht (Luirink *et al.* (1991)). BRPs verschiedener Bacteriocine sind untereinander austauschbar, sie alle vermitteln auch die Freisetzung verwandter Bacteriocine, was auf einen sehr ähnlichen Sekretionsmechanismus schließen lässt. Die BRP vermittelte Sekretion beruht dabei auf einer semispezifischen Erhöhung der Membranpermeabilität, neben dem Bacteriocin werden auch ein kleiner Teil der cytoplasmatischen Proteine und viele periplasmatisch lokalisierte Proteine sekretiert. Dieser Vorgang ist jedoch nicht mit einer vollständigen Lyse vergleichbar, da viele Proteine und größere Moleküle wie die DNA und RNA weiter im Cytoplasma verbleiben. Diesen Umstand macht man sich für die Sekretion rekombinanter Proteine zu nutze (Kleist *et al.* (2002), Rahman *et al.* (2003)).

Die volle Induktion eines Colicin-Genclusters führt zu einer Abnahme der Trübung einer *E. coli*-Kultur und zur Lethalität von auf Agarplatten wachsenden Bakterien. Dieses als Quasi-Lyse bezeichnete Phänomen entsteht nicht durch ungenügenden Schutz durch das Immunitätsprotein oder eine zu große Aktivität des Bacteriocins, sondern einzig durch das reife BRP und/oder dessen stabilen Signalpeptids (Luirink *et al.* (1991), van der Waal *et al.* (1995)). Begleitet wird die Quasi-Lyse von einem sinkenden Membranpotential. Schon bevor die Symptome sichtbar werden wird die Proteinbiosynthese der Zelle und der Transport von Magnesium-Ionen stark inhibiert (Stegehuis *et al.* (1995)). Die beobachteten Effekte, verursacht durch die Induktion der Colicin-Gencluster, gleichen denen einer Bakteriophagen-Infektion oder einer Zellbehandlung mit EDTA. Beide beruhen auf einer Aktivierung der Detergenzien-stabilen Phospholipase A der äußeren Membran von *E. coli* (Dekker *et al.* (1999)). Tatsächlich zeigten PldA-defiziente Mutantenstämme unter Induktion keine Quasi-Lyse und keine Freisetzung von Bacteriocinen. Weiterhin wurde diese These dadurch verifiziert, dass die Aktivität von PldA während der Quasi-Lyse um das 30-40 fache ansteigt (Pugsley *et al.* (1993)). Elektronenmikroskopische Untersuchungen haben gezeigt, dass die äußere Membran durch die Quasi-Lyse größere Öffnungen aufweist. Durch die PldA-Aktivität wird das Hauptphospholipid der äußeren Membran PE abgebaut. Die Degradationsprodukte, freie Fettsäuren und Lyso-PE steigen signifikant in den induzierten Zellen an. PldA-defiziente Systeme können allerdings durch Zugabe von Triton X-100 oder durch die Verwendung von tolQ-Stämmen (Permeabilität für periplasmatisch lokalisierte Proteine) die Fähigkeit zur Bacteriocinfreisetzung zurückgewinnen, auch hier tritt Quasi-Lyse auf. In beiden Fällen ist aber obligat das BRP für eine Freisetzung vom Bacteriocin notwendig, was für eine zusätzliche Modifizierung der Zellmembran durch das BRP spricht. Das BRP wird als Vorläufermolekül von 45-52 Aminosäuren mit einer N-terminalen Signalsequenz exprimiert. Die Signalsequenz vermittelt eine Translokation durch die Cytoplasmamembran mittels des Sec-Pfades; SecA, SecB und SecY sind essentiell für den Export fast aller BRPs. Im Bereich der Schnittstelle für die prozessierende Signalpeptidase befindet sich eine hochkonservierte Region, welche als Lipobox bezeichnet wird: Leu-X-Y-Cys. Bei X und Y handelt es sich um kleine, neutrale Aminosäuren und Cystein stellt die erste Aminosäure des reifen BRP dar. Während der Translokation des BRP durch die Cytoplasmamembran wird die Sulfhydryl-Gruppe des Cysteins kovalent mit einer Diacylglycerly-Gruppe gekoppelt, anschließend wird an der periplasmatischen Seite der Cytoplasmamembran durch die Signalpeptidase II das Signalpeptid vom BRP gespalten. Schließlich wird noch eine Fettsäure kovalent an die freie Aminogruppe des Cysteins der Lipobox gebunden. Nach erfolgter Lipidmodifizierung erhält das BRP seine volle biologische Aktivität (Quasi-Lyse und Bacteriocin-Freisetzung werden sonst nicht beobachtet). Die Lipobox ist essentiell für die Lipidmodifikation und Prozessie-

nung durch die Signalpeptidase II. Eine Substitution des Cysteins der Lipobox gegen andere Aminosäuren verhindert eine Lipidmodifizierung des BRPs (van der Waal *et al.* (1995)).

Ebenso ist die Prozessierung durch die Signalpeptidase bei einigen BRPs, besonders bei Cloacin DF13, ein sehr langsamer Prozess im Bereich von einigen Minuten. Unprozessiertes BRP lässt sich neben prozessiertem BRP und dem Signalpeptid in *pulse-chase*-Experimenten mittels SDS-PAGE visualisieren (Luirink *et al.* (1991)). Für gewöhnlich ist die Abspaltung der Signalpeptide durch die Signalpeptidase II ein sehr schneller Vorgang, beim als Beispiel fungierenden LPP, dem häufigsten Lipoprotein, beträgt die Halbwertszeit wenige Sekunden. Die Lipidmodifikation der BRPs ist auch für ihre Lokalisierung in der äußeren Membran essentiell, bei einem Austausch des Cysteins der Lipobox gegen andere Aminosäuren konnte das BRP ausschließlich in der Cytoplasmamembran nachgewiesen werden. Jedoch sind auch weitere Bereiche des reifen BRPs für die korrekte Lokalisierung verantwortlich. Ungewöhnlich ist die Tatsache, dass das Signalpeptid der meisten BRPs stabil ist und nicht unmittelbar nach der Abspaltung durch die Signalpeptidase proteolytisch degradiert wird. Statt dessen akkumuliert es exklusiv in der Cytoplasmamembran. Diese stabilen Signalpeptide sind teilweise für die Quasi-Lyse und Lethalität verantwortlich, wie deren Expression als separate Einheit zeigen konnte (Van der Waal *et al.* (1992)). Allerdings vermögen die stabilen Signalpeptide alleine nicht die Freisetzung von Bacteriocinen zu vermitteln (van der Waal *et al.* (1995)). Für die ungewöhnlich hohe Stabilität der Signalpeptide scheint der N-Terminus und ein am C-Terminus befindliches Alanin verantwortlich. Es wurden verschiedene Hybrid-Signalpeptide z.B. mit Teilen des instabilen Signalpeptides vom Lpp-Protein dargestellt, um zu ermitteln, welche Bereiche für eine effiziente Bacteriocin-Freisetzung benötigt werden. Einige diese Hybridsignalpeptide zeigen großes Potenzial bei der Freisetzung des periplasmatischen Inhalts, ohne die negativen Effekte der Quasi-Lyse zu vermitteln (van der Waal *et al.* (1998)).

Der genaue Mechanismus der Bacteriocinfreisetzung über zwei bakterielle Membranen hinweg, ist noch nicht vollständig geklärt. Es wird davon ausgegangen, dass die stabilen Signalpeptide, das reife BRP und PldA in der Bildung von Transmembranporen kooperieren. So akkumulieren die Bacteriocine Colicin A und Cloacin DF13 im Cytoplasma PldA-defizienter Stämme. Gleiches wird phänotypisch beobachtet, wenn eine Expression der genannten Bacteriocine ohne BRP oder nur mit dem Signalpeptid des BRPs erfolgt, was auf eine simultane Translokation über zwei Membranen, ohne periplasmatisches Intermediat, hindeutet (van der Waal *et al.* (1995)). Eine Visualisierung der Transmembranporen mittels elektronenmikroskopischer Methoden ist bis jetzt noch nicht gelungen. Indirekte Nachweise für deren Existenz wurden jedoch häufiger beobachtet. So sind zum Beispiel die Cytoplasma- und äußere Membran bei Bacteriocin-freisetzenden Bakterien sehr schwer zu separieren. Eine gängige Methode bedient sich dabei der isopyknen Saccharose-Dichtegradientenzentrifugation. Proteine die charakteristisch für eine diskrete Membranfraktion sind, konnten in einer intermediären Phase mittlerer Dichte nachgewiesen werden, welche sonst nicht auftritt. *In vitro* Experimente mit den Signalpeptiden von PhoE und dem M13-Phagenhüllprotein konnten zudem zeigen, dass akkumulierende Signalpeptide in einem artifiziellen Membransystem die natürliche Doppelschichtstruktur auflösen. In diesem System sind Enzyme, welche die Signalpeptide degradieren, nicht vorhanden, weshalb diese an sich unstablen Modellsignalpeptide akkumulieren (van der Waal *et al.* (1995)).

Die einzigartigen Eigenschaften der BRPs lassen sich für die Freisetzung heterologer Proteine aus dem Periplasma von *E. coli* nutzen. Diese Applikation wurde durch eine extrazelluläre Darstellung verschiedenster rekombinanter Proteine unterschiedlichster Größe

verifiziert (Miksch *et al.* (2002), Kleist *et al.* (2003), Beshay *et al.* (2008), Rahman *et al.* (2003)). Durch Fusionierung des Zielproteins mit einer N-terminalen Signalsequenz für einen der erwähnten Sekretionspfade vom Typ II und simultane Expression eines BRPs lässt sich die extrazelluläre Lokalisierung realisieren. Dabei hat es sich als vorteilhaft erwiesen, ein 1-Plasmid-System zu verwenden, bei dem sich das BRP und das Zielprotein auf dem gleichen Plasmid unter Kontrolle verschiedener, induzierbarer Promotoren befinden (Mergulaho *et al.* (2005)). Die Verwendung verschiedener, induzierbarer Promotoren verfolgt den Zweck, die Expressionsniveaus möglichst optimal und unabhängig voneinander zu halten. Eine zu hohe Expression der BRPs würde eine Quasi-Lyse verursachen, welches negative Folgen für die Produktivität und die Aufreinigung beinhaltet. Hingegen sollte das rekombinante Protein in möglichst großer Konzentration im Periplasma vorliegen, da auch andere periplasmatisch lokalisierte Proteine durch die BRP-Expression in den extrazellulären Raum gelangen. Somit sollte auf eine möglichst effektive Expression und Prozessierung geachtet werden. Weitere Optimierungsmöglichkeiten stellen die Verwendung Protease-defizienter Stämme und optimierte Kultivierungsbedingungen dar, welche einer Quasi-Lyse entgegenwirken (z.B. Zusatz zweiwertiger Ionen zum Medium) (Sommer *et al.* (2008)). Eine weitere Möglichkeit liegt, wie schon eingangs erwähnt, in der Nutzung von Hybridsignalpeptid-modifizierten BRPs, welche funktionell in der Freisetzung des periplasmatischen Inhalts sind, aber erst bei weit höheren Expressionsniveaus eine Quasi-Lyse verursachen.

3 Genexpressionsanalytik

3.1 Übersicht

Die wichtigsten Methoden zur Analyse der Genexpression sind:

1. DNA-Microarrays
2. Quantitative *real-time* PCR
3. Northern Blots
4. Ribonuklease Protektions Assay (RPA)
5. SAGE

Im Folgenden sollen die DNA-Microarrays und die quantitative *real-time* PCR genauer erläutert werden, zunächst wird jedoch die Hybridisierung als Grundlage von Genexpressionsanalysen kurz vorgestellt.

3.2 Hybridisierung

Die Grundlage jeglicher Genexpressionsanalysen ist die sequenzspezifische Hybridisierung einer Sonde mit dem zu detektierenden Nukleinsäuretarget. Bei den Sonden kann es sich um Oligodesoxynukleotide, DNA-Fragmente, PCR-Produkte, *in vitro* RNA-Transkripte oder artifizielle Sonden wie z.B. Peptidnukleinsäuren (PNAs) handeln. Die Hybridisierung kann dabei in homogener Phase zwischen zwei Molekülen in Lösung erfolgen (RPA, *real-time* PCR) oder zwischen einem immobilisierten Fängermolekül (Membran, Mikroarray) und dem in Lösung befindlichen Zielmolekül.

Die Hybridisierung von Nucleinsäuren beruht auf der komplementären Basenpaarung über Wasserstoffbrücken. Adenosin bildet mit Thymin bzw. Uracil zwei Wasserstoffbrücken aus, wohingegen Cytosin und Guanin über drei Brücken paaren, was einen entsprechend stabileren Duplexstrang zur Folge hat. Die Schmelztemperatur ist als diejenige Temperatur definiert, bei der die Hälfte der Nucleinsäureduplexe dissoziiert in den Einzelsträngen vorliegt. Sie ist abhängig von der Länge und Basenzusammensetzung des hybridisierenden Sequenzabschnittes, der Salzkonzentration, der Konzentration helixdestabilisierender Reagenzien (z.B. Formamid) und der Art der hybridisierenden Duplexe (RNA-RNA, RNA-DNA oder DNA-DNA). Generell weisen RNA-RNA-Hybride die größte Schmelztemperatur auf, gefolgt von den DNA-RNA-Heteroduplexen. Die Kinetik der Hybridisierung ist abhängig von der Duplexlänge und der Diffusionsrate der Sonde bzw. des Nucleinsäuretargets. Kleine Moleküle weisen die höchsten Diffusionskoeffizienten auf, so ist bei Oligonukleotidsonden die Hybridisierung deutlich schneller als bei längeren Nucleinsäuresonden. Oligonukleotidsonden weisen allerdings einen Sensitivitäts-Nachteil auf, da der hybridisierende Bereich kleiner ist und die Anzahl der Markierungen aufgrund der Sondenlänge limitiert (Southern *et al.* (1992)). Allerdings lässt sich durch Verwendung von Oligonukleotidkassetten oder terminalem Anheften mehrerer Markierungen (*tailing*) die Sensitivität deutlich steigern. Weitere die Kinetik beeinflussende Faktoren sind die Ionenstärke, die Viskosität der Hybridisierungslösung, Basenfehlpaarungen und die Konzentration helixdestabilisierender Reagenzien. Eine Erhöhung der Hybridisierungsrate lässt sich z.B. durch Additive wie Dextranulphat oder Polyethylenglykol erzielen, diese wirken wasserentziehend und steigern so die effektive Konzentration. Wesentlich für die Detektion eines Nucleinsäuretargets ist eine möglichst spezifische Hybridisierung, diese wird von der Stringenz der Reaktionsbedingungen und der Stabilität des Hybridkomplexes beeinflusst. Stringente Bedingungen sind solche unter denen (fast) nur ideal basengepaarte Nucleinsäureduplexe auftreten. Die Stringenz wird bestimmt von der Ionenstärke, der Temperatur und der Konzentration helixdestabilisierender Reagenzien. Monovalente Kationen wie Natrium-Ionen erniedrigen die Stringenz, da durch die positive Ladung die negative Ladung des Zucker-Phosphat-Rückgrates abgeschirmt wird. Der Schmelzpunkt eines Nucleinsäureduplexes verringert sich je ein Prozent Basenfehlpaarung um bis zu fünf Grad, daher ist für eine hohe Stringenz eine möglichst hohe Temperatur zu wählen. In dieser Hinsicht ist besonders die Verwendung von PNA-Sonden zu erwähnen, hier erfolgt die Diskriminierung aufgrund von Basenfehlpaarung besonders effektiv. Der Grund dafür liegt im peptidähnlichen, ungeladenen Rückgrat dieser artifiziellen Nucleinsäureanaloge, welches keine intermolekulare, elektrostatische Abstoßung zwischen den hybridisierenden Molekülen generiert (Lottspeich F.: Bioanalytik (1998)).

3.3 DNA-Microarray-Technologie

3.3.1 Übersicht Microarray-Technologie

Seit der vollständigen Sequenzierung des menschlichen Genoms und derer vieler anderer Organismen steht eine große Anzahl an Sequenzinformationen zur Verfügung. Die Zuordnung von biologischen Funktionen zu diesen strukturellen Genomdaten hat allerdings gerade erst begonnen (Übergang von *structural* zu *functional genomics*). Zur Realisierung dieser Aufgabe stellen Microarrays aufgrund ihres hohen Parallelisierungsgrades eine außerordentlich wertvolle Analysenmethode dar. Microarrays haben sich durch Miniaturisierung z.B. aus den Macroarrays auf Nylonmembranen entwickelt (*Reversed Northern*). Motiviert wurde dieser Vorgang durch den Wunsch, ein Hochdurchsatz-*Screening* mit minimalem Probenmemgen

für verschiedene Applikationen zu ermöglichen, um so eine Zeit- und Ressourcenersparnis zu realisieren (Skena *et al.* (1995)). Allerdings stellt die Interpretation und Bearbeitung der immensen Datenmengen, welche durch Microarray-Experimente generiert werden, einen nicht zu unterschätzenden Arbeitsaufwand dar. Generell werden auf Microarrays Interaktionen zwischen verschiedenen, auf einer festen Oberfläche immobilisierten Fängermolekülen (*probes*) und komplementären, zur Quantifizierung markierten Zielmolekülen untersucht. Mögliche Interaktionspartner sind DNA-, RNA-, PNA- und Proteinmoleküle, aber auch kleine organische oder anorganische Verbindungen (z.B. *drug-target-screening*). Eine Klassifizierung der Microarrays kann nach der Art des immobilisierten Fängermoleküls oder nach dem wissenschaftlichen Applikationsbereich erfolgen. Beispiele für letzteres sind das *drug-target-screening* oder die Transkriptomanalyse. Die Stärke des Formates liegt in der parallelen Ermittlung vieler relativer Signalstärken, die absolute Quantifizierung stellt nach wie vor eine besondere Herausforderung dar (Eickhoff *et al.* (1999)). DNA-Microarrays ermöglichen den Expressionsstatus von tausenden von Genen simultan zu erfassen und haben sich als wichtiges *Screening-Tool* in der Molekularbiologie und medizinischen Diagnostik etabliert. Eine der häufigsten Anwendungen von DNA-Microarrays stellt die differentielle Genexpressionanalyse dar (Skena *et al.* (1995)). Dabei wird das Transkriptom zweier oder mehrerer Zustände miteinander verglichen und die Unterschiede ermittelt. Eine mögliche Applikation ist zum Beispiel der Vergleich des Transkriptoms von Tumorgewebe mit dem Gewebe von phänotypisch nicht-entarteten Zellen. Gene, die im Tumorgewebe differentiell exprimiert sind, weisen auf einen möglichen Zusammenhang mit der Entartung hin und sind somit potenzielle Kandidatengene.

3.3.2 Oberflächen

Ein DNA-Microarray besteht aus einer festen Oberfläche, zumeist ein Mikroskopträger mit einer modifizierten Glasoberfläche, auf dem spezifische Sonden immobilisiert sind. Bei den Sondenformaten kann es sich um längere PCR-Produkte oder um Oligonukleotide handeln. Unter den Oligonukleotiden erfolgt weiterhin eine Unterscheidung nach kurzen (20-30nt) oder längeren Sonden (50-70nt), die jeweils für unterschiedliche Applikationen Einsatz finden. Die Deponierung der Sonden auf der Glasoberfläche, nach Reihen, Spalten und Blöcken geordnet, erfolgt mittels sogenannter *Arayer* oder durch *in situ*-Synthese z.B. mittels photolithographischer Masken. Liegen keine Sequenzinformationen über den untersuchten Organismus vor, so müssen obligat cDNA-Produkte als Sonde verwendet werden. Längere Oligonukleotide finden meist in der klassischen Applikation der differentiellen Genexpressionsanalyse Verwendung, wohingegen kürzere Oligonukleotide ihre Stärke z.B. in der Detektion von SNPs (*single nucleotide polymorphisms*) aufweisen. Fehlpaarungen von nur einer Base können durch Wahl entsprechend stringenter Hybridisierungsbedingungen so detektiert werden. Zumeist sind diese Sonden über *Spacer* variabler Länge mit einer primären aliphatischen Aminofunktion kovalent an die chemisch modifizierten Glasoberflächen gekoppelt. Die *Spacer* werden entweder über das 5'-Ende der Primer bei der PCR-Reaktion oder bei der Synthese der Oligonukleotide eingefügt. Bei der Oligonukleotidsynthese besteht alternativ auch die Möglichkeit einer 3'-Modifizierung des Oligonukleotids. Eine gängige Modifizierung der Glasoberfläche erfolgt z.B. durch chemische Konversion der Silanolgruppen des Glases in Aldehydfunktionalitäten. Diese Aldehydgruppen reagieren mit der Aminofunktion der Sonden *spacer* in einer spontanen Reaktion zu einem Imin. Dieses wird durch Reduktion z.B. mit Borhydriden zu einem stabilen sekundären Amin umgewandelt und die Sonde so kovalent

an die Chipoberfläche fixiert. DNA-Microarrays sind für viele Standardapplikationen bereits kommerziell erhältlich. Für spezielle wissenschaftliche Fragestellungen müssen sie aber zumeist selbst designed und produziert bzw. in Auftrag gegeben werden. Dies ermöglicht ein flexibles Design z.B. mit mehreren Genreplikaten und eine Fokussierung auf die relevanten Gene. Eine gängige Vorgehensweise ist z.B. die Identifizierung relevanter Gene über kommerzielle *whole-genome* Microarrays, um dann über spezifische *low-density* Microarrays eine tiefere Analyse durchzuführen (Schadt *et al.* (2000)).

3.3.3 Computer-gestütztes Design von Oligonukleotiden

Oligonukleotide für Microarrays repräsentieren einen kurzen Abschnitt des Zielgenes, dessen Expression gemessen werden soll. Normalerweise werden sie innerhalb weniger hundert Basenpaare vom 3'-Ende desselben designed, da die 5'-Polyadenylierung der mRNA für ein Oligo-dT-*Priming* bei der cDNA-Synthese genutzt wird. Zudem handelt es sich zumeist um jenen Bereich, der für das betreffende Gen am spezifischsten ist und die geringsten Homologien mit anderen Genen aufweist. Ideale Oligonukleotide müssen die folgenden drei Eigenschaften aufweisen:

1. Sensitivität
2. Spezifität
3. Isothermales Verhalten

Eine sensitive Sonde ergibt in Gegenwart des komplementären *Targets* ein starkes Signal. Die Sensitivität ergibt sich im Wesentlichen aus zwei Faktoren. Die Sonde darf keine intramolekularen und intermolekularen (mit benachbarten, identischen Sondenmolekülen) Sekundärstrukturen ausbilden und muss am komplementären *Target* binden können. Dieses kann durch Sekundärstrukturen im selbigen erschwert werden. Eine spezifische Sonde zeigt keine Kreuzhybridisierungen und im Idealfall kein Signal, wenn das komplementäre Zielmolekül nicht in der Hybridisierungslösung vorhanden ist. Es sind wiederum zwei Faktoren, welche sich auf die Spezifität des Oligonukleotides auswirken, die Kreuzhybridisierung mit anderen *Targets* und die nicht-spezifische Bindung aufgrund von Sequenzen mit niedriger Komplexität (z.B. G-Quartets). Isothermale Sonden zeigen ein gleiches Verhalten unter den Hybridisierungsbedingungen (bestimmt im Wesentlichen von Temperatur, Salz- und Formamidkonzentration) des Microarray-Experimentes. Der Parameter, welcher dieses hinlänglich beschreibt, ist die Schmelztemperatur der entstehenden Duplexe aus Sonde und *Target*. Diese sollte nach Möglichkeit für alle Oligonukleotide auf dem Microarray gleich sein. Viele dieser Anforderungen einer Sonde lassen sich durch Verwendung von Computeralgorithmen sicher stellen. Es existieren verlässliche Programme zur Vorhersage von internen Sekundärstrukturen (z.B. MFold), Homologien des Oligonukleotides mit anderen *Targets* (z.B. BLAST) und zur Vorhersage der Schmelztemperatur (z.B. *base-stacking* Modell). Sekundärstrukturen im *Target* und nicht-spezifische Interaktionen lassen sich hingegen schwerer erfassen und vorhersagen. Kane *et al.* (2000) haben für das Design von 50mer Oligonukleotiden für Microarray-Applikationen konkrete Kriterien bestimmt anhand derer die in diesem Abschnitt beschriebenen Voraussetzungen für ein spezifisches und sensibles Sondendesign erfüllt werden.

3.3.4 Hybridisierung und *Labeling*

Ein Chipexperiment setzt sich aus folgenden fünf Schritten zusammen:

1. Probenpräparation und *Labeling*
2. Hybridisierung
3. Waschen
4. Scannen
5. Analyse (Primär, Sekundär und Tertiäranalyse)

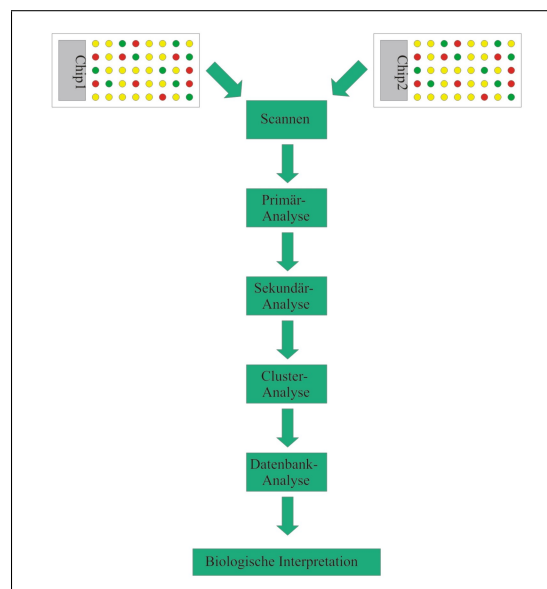


Abb. 2: Auswertung von Microarrays. Schematisch ist die Vorgehensweise nach der experimentellen Durchführung bis zur biologischen Interpretation der Daten gezeigt.

Die Abbildung 2 zeigt schematisch den Gang der Microarray-Analyse nach den eigentlichen experimentellen Schritten. Mittels immobilisierter DNA-Sonden lassen sich sowohl Proteine (Aptamerchips) als auch Nukleinsäuren als Zielmoleküle detektieren. Dazu muss das *Target* mit einer quantifizierbaren Markierung (*Labeling*) versehen werden. Im Bereich der Genexpressionsanalysen dient dazu die mRNA, welche entweder mittels chemischer Modifikation mit fluoreszierenden, radioaktiven oder anderen Haptenen wie z.B. Biotin oder Digoxigenin gekoppelt oder enzymatisch mittels reverser Transkription in cDNA oder cRNA konvertiert wird. Man unterscheidet direkte und indirekte *Labelingsysteme*, je nachdem ob die signalgebende Reportergruppe direkt oder über ein weiteres Molekül an das Nukleinsäuretarget gebunden ist (Schäferling *et al.* (2008)).

Die Verwendung von RNA als *Target* beinhaltet den Vorteil einer stärkeren Basenpaarung und damit die Möglichkeit stringenter Bedingungen einzusetzen. Dennoch wird meistens cDNA verwendet, da diese robuster gegen Hydrolyse und ubiquitär vorkommende RNAsen ist. Die Orientierung der generierten Zielnukleinsäure (*sense* oder *antisense* Strang) besitzt Implikationen für das Design der Fängermoleküle auf dem Chip, diese müssen eine komplementäre Sequenz aufweisen. Aufgrund der Tatsache, dass zumeist cDNA als

Target für Microarray-Applikationen Verwendung findet, wird die Mehrzahl der Chips mit *sense*-Oligonukleotiden *gespottet*. Bei der Synthese der cDNA werden dem Reaktionsansatz fluoreszierende, radioaktive oder Hapten-modifizierte Desoxynukleotidtriphosphate (dNTPs) zugesetzt, welche von dem Enzym Reverse Transkriptase in die entstehende cDNA-Kette inkorporiert werden (Abbildung 3). Durch unterschiedliche Dotierungsverhältnisse von markiertem (meistens als modifiziertes dCTP vorliegend) zu unmarkiertem dNTP lässt sich eine unterschiedliche Markierungsdichte der cDNA realisieren. Handelt es sich, wie z.B. bei den Fluorophoren, um sterisch große Moleküle, so erfolgt der Einbau der markierten dNTPs signifikant langsamer als der jener unmarkierten dNTPs. Für *Multiplexing*-Ansätze auf dem gleichen Chip besitzt dieses ebenfalls Relevanz, wenn mehrere, sterisch unterschiedlich anspruchsvolle Modifizierungsgruppen für verschiedene Proben verwendet werden sollen. Als Beispiel kann hier die Markierung der cDNA mit Cy3-dCTP und Cy5-dCTP fungieren. Durch ein ausgedehnteres Π -Elektronensystem ist das Cy5-Molekül sterisch anspruchsvoller und wird von der Reversen Transkriptase mit geringerer Effizienz eingebaut. Dieser Tatsache muss Rechnung getragen werden, da eine unterschiedliche Signalstärke bei einem Vergleich zweier Proben sonst auch das Resultat unterschiedlicher Einbauraten sein kann. Durch ein entsprechendes Design der Experimente lässt sich diese Problematik aber einfach umgehen, indem nämlich jede Probe sowohl mit Cy3-dCTP als auch mit Cy5-dCTP gelabelt wird (*Dye-Swap*). Es wurden zudem spezielle, modifizierte Reverse Transkriptasen generiert (z.B. Label-Star RT von Qiagen), welche den Einbau sterisch anspruchsvoller, modifizierter dNTPs mit größerer Effizienz vollziehen.

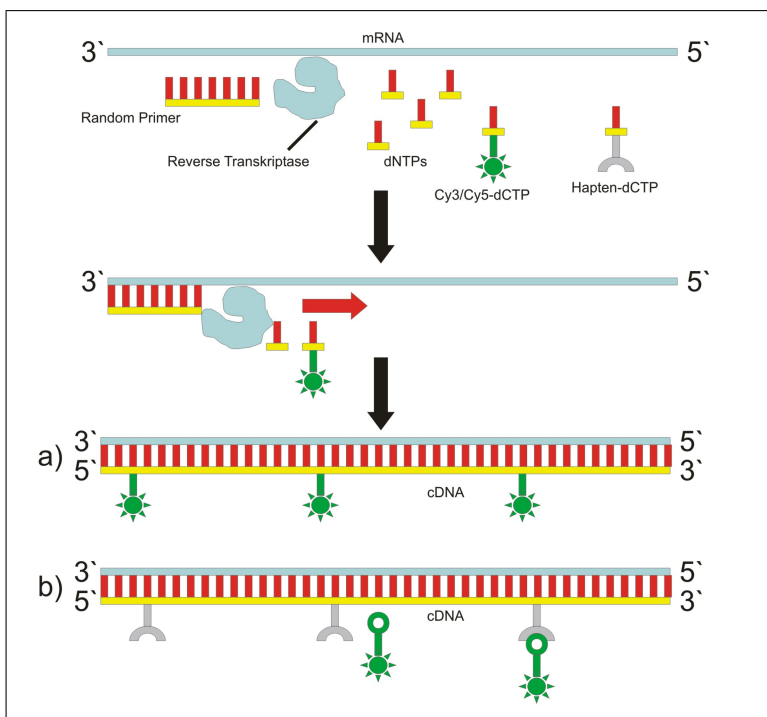


Abb. 3: Direktes a) und indirektes b) *Labeling* der cDNA mittels Reverser Transkription.

Das so markierte *Target* wird in Hybridisierungspuffer aufgenommen und auf die Chipoberfläche transferriert. Bei den Hybridisierungspuffern handelt es sich zumeist um SSC-Puffer basierte, formamidhaltige Lösungen. Je nach Formamid- und Salzkonzentration wird

eine Hybridisierungstemperatur von 42–65°C gewählt und zumeist über Nacht hybridisiert, damit sich das Hybridisierungsgleichgewicht einstellen kann. Man unterscheidet manuelle von automatisierten Hybridisierungen, welche in sogenannten Hybridisierungsstationen durchgeführt werden. Während bei den manuellen Hybridisierungen in einer feuchten Kammer oder einem Schüttler ein Deckglaschen auf dem Microarray als Verdunstungsschutz fungiert und die Hybridisierungsreaktion zumeist diffusionskontrolliert verläuft, sorgt ein konvektiver Flüssigkeitsstrom über Mikropumpen in den Hybridisierungsstationen für eine gleichmäßige Verteilung der markierten Zielmoleküle über der Oberfläche. Dementsprechend sind die Resultate einer automatisierten Hybridisierung meistens reproduzierbarer. Aufgrund der applizierten, sehr kleinen Volumina ist ein effektiver Verdunstungsschutz von essentieller Bedeutung, da eine Eintrocknung der Hybridisierungslösung in extremen Hintergründen resultiert.

3.3.5 Waschen von Microarrays

Zur Eliminierung von unspezifisch an die Oberfläche adsorbierten *Targets* besitzt der Waschvorgang einen großen Einfluss auf das generierte SNR. Die dafür verwendeten Waschpuffer sollten sterilfiltriert werden und der Waschvorgang in einer möglichst staubfreien Atmosphäre geschehen. Meistens finden Natriumchlorid- und Natriumcitrat-haltige Puffer, denen z.T. geringe Mengen an Detergenzien zugesetzt werden, Verwendung. Die Waschtemperatur stellt ebenfalls einen kritischen Parameter für hinreichend stringente Waschbedingungen dar.

3.3.6 Scannen von Microarrays

Der Scanvorgang liefert ein digitales Abbild vom Chip und wird entweder in Fluoreszenz-Laserscannern oder mittels CCD-Kameras durchgeführt. Dabei müssen neben sehr starken auch schwache Signale detektiert werden. Dies ist oftmals nicht durch einen einzigen Scan zu realisieren. Um eine Adaption an variierende Umweltbedingungen zu ermöglichen, stehen viele Gene unter der Kontrolle induzierbarer Promotoren. Die Expressionsrate eines Genes ist somit häufig starken Schwankungen unterworfen (manche Gene werden nach Induktion 1000-fach stärker transkribiert). In einer *E. coli* Zelle liegen seltene Transkripte (1–10 Kopien) neben Transkripten großer Häufigkeit vor (ca. 10^5 Kopien) (Stryer J.: Biochemie (1996)), was einen enormen Anspruch an den dynamischen Bereich der Analysenmethode stellt. Durch Scan-Einstellungen (PMT, Laserpower) mit hohen Empfindlichkeiten laufen die starken Signale in die Sättigung. *Vice versa* können bei niedriger Empfindlichkeit zwar die Unterschiede der starken Signale korrekt erfasst werden, dabei gehen aber die schwachen Signale verloren. Technisch gesehen ist der reale dynamische Bereich einer Analysenmethode die Speicherkapazität dividiert durch den Gesamthintergrund. Verfügt der Scanner zum Beispiel über einen 16-Bit-Konverter so lassen sich Werte von 0–65355 darstellen. Leider schränkt jedoch der inhärent vorliegende Hintergrund, welcher durch unspezifische Adsorption von Zielmolekülen und von der Autofluoreszenz der Microarray-Oberfläche herrührt, den dynamischen Bereich stark ein. Schon ein Hintergrund von 100 bedeutet im gewählten Beispiel, dass der dynamische Bereich statt 65355 nur noch 653 beträgt. Der Hintergrund schränkt den dynamischen Bereich stark ein und sollte daher möglichst minimiert werden.

Das *Image* des Microarrays ist aus einzelnen Pixeln als kleinste Einheit zusammengesetzt. Jedes Pixel repräsentiert einen numerisch quantifizierbaren Wert der Fluoreszenzintensität. Die Pixelgröße hängt dabei von der gewählten Auflösung ab. Um einen *Spot* ge-

trennt von seiner Umgebung darstellen zu können, muss er aus einer Mindestanzahl von Pixeln bestehen. Als Faustformel haben sich 50-100 Pixel bewährt, für einen *Spot* mit einem Durchmesser von 100 μm bedeutet dieses, dass ein Scan mit einer Auflösung von 10 μm ausreichend viele Pixel pro *Spot* generiert. Durch eine erhöhte Auflösung lassen sich die *Spots* besser darstellen, allerdings geht dieses einher mit einer Verkleinerung der Pixelfläche. Diese ist aus zwei Gründen nicht beliebig zu verkleinern. Erstens sinkt die Empfindlichkeit des Systems, da sich auf der verkleinerten Fläche weniger signalgebende Fluorophore befinden und zweitens ist der dynamische Bereich des Detektionssystems direkt abhängig von der Pixelfläche. Die Speicherkapazität eines Pixelfeldes ist definiert als: $1000 \cdot (\text{Pixelfläche}[\mu\text{m}])^2$. Für die meisten Applikationen sind Auflösungen im Bereich von 5-10 μm gut geeignet um robuste Daten zu generieren.

3.4 Auswertemethoden für Microarrays

Die wichtigsten, in dieser Arbeit verwendeten Auswertemethoden werden im Folgenden kurz erläutert.

3.4.1 Primäranalyse

Das vom Scanner generierte *Image* des Microarrays stellt die Rohdaten des Chipexperimentes dar. Computeralgorithmen konvertieren es in numerische Daten, welche eine Quantifizierung der Genexpression ermöglichen. Zunächst muss die Position der *Spots* auf dem Array identifiziert werden. Dazu bedient man sich einer Maske, *Grid* genannt, welche das vollständige *Spotting*-Muster der Oligonukleotide auf dem Chip repräsentiert. Dieses Muster ist unterteilt in Blöcke und den darin enthaltenden *Spots* und dient zur Adressierung der Signale auf dem Chip zu dem entsprechenden *Gen*target. Ein Block wurde dabei jeweils von dem gleichen *Pin* des *Spotters* gedruckt. Das Grid wird passend über das digitale Abbild des Chips gelegt (*Alignment*). Zuerst werden dabei die einzelnen Blöcke abgeglichen und anschließend die einzelnen *Spots* innerhalb eines Blocks. Beispielhaft ist ein *Alignment* in Abbildung 4 gezeigt.

Besonders bei Chips, die mittels *Kontaktprintern* hergestellt wurden, kann das *Alignment* schwierig sein. Blöcke können als Folge einer fehlerhaften Abstimmung der *Pins* in ihren Positionen zueinander verschoben sein. Ungleichmäßige Abstände zwischen den *Spots* oder eine Kurve der *Spots* innerhalb eines Blocks deuten auf Unebenheiten der Glasoberfläche oder auf eine Verschiebung der *Pins* im Druckkopf hin. Eine ungleichmäßige *Spotgröße* entsteht durch Transferrierung unterschiedlicher Volumina Flüssigkeit auf die Chipoberfläche. Diese Fehler von der ideal geometrischen Anordnung stellen eine Herausforderung an die Primäranalysensoftware dar und verlangen häufig eine manuelle Korrektur. Der nachfolgende Schritt besteht aus der Identifizierung der Pixel welche zum *Spot* gehören, dies bezeichnet man als *Segmentation*. Computeralgorithmen verwenden dafür vier verschiedene Methoden als Grundlage:

1. Identifizierung über eine Kreisstruktur fester Größe
2. Identifizierung über eine Kreisstruktur variabler Größe
3. Identifizierung über das Histogramm
4. Identifizierung über eine adaptive Formgebung

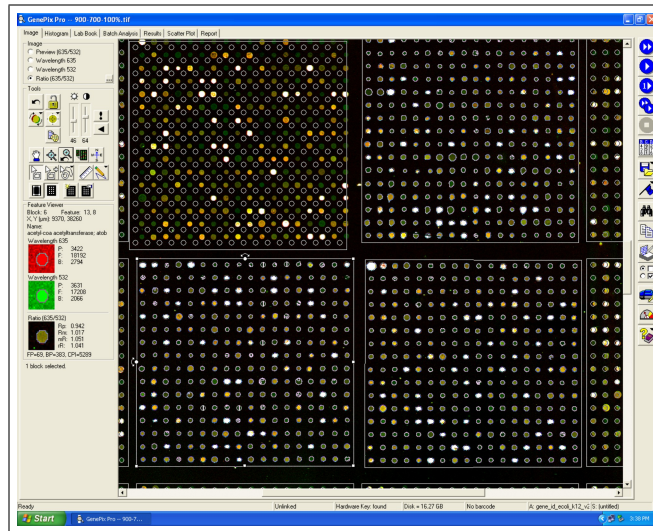


Abb. 4: Alignment des Grids mit dem Image des Microarrays mittels einer kommerziellen Software.

Die Illustration der verschiedenen Verfahren ist in Abbildung 5 gezeigt. Eine Identifizierung über Kreisstrukturen fester Größe sollte vermieden werden, da *Spots* im Realfall häufig eine variierende Größe aufweisen. Daneben entspricht die *Spot*morphologie häufig nicht derer eines idealen Kreises z.B. aufgrund sogenannter *Donut*-Effekte. Deswegen finden besonders die beiden letztgenannten Algorithmen Verwendung. Bei der Histogramm-Segmentation wird ein Kreis über *Spot* und Hintergrund gelegt und ein Histogramm der Intensität der Pixel vom *Spot* erstellt. Die Pixel mit den größten und kleinsten Intensitäten werden nicht bei der Quantifizierung der *Spot*intensität berücksichtigt. Dieses Verfahren generiert verlässliche Daten bei irregulär geformten *Spots*, zeigt aber Schwächen wenn die *Spot*größe zu klein für die zirkuläre Maske ist. Die genauesten Ergebnisse werden zur Zeit durch Algorithmen zur adaptiven Formgebung generiert. Der dazu verwendete Algorithmus benötigt eine kleine Anzahl von Startpixeln im Zentrum des *Spots*. Der *Spot* wird davon ausgehend auf benachbarte Pixel mit ähnlichen Intensitäten ausgeweitet.

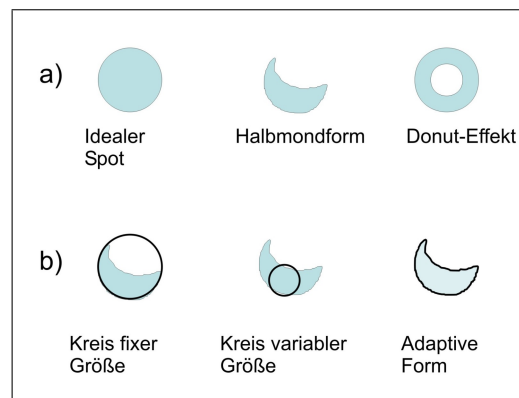


Abb. 5: Verschiedene *Spot*formen (a) und Verfahren zur Identifizierung der *Spot*pixel (b).

Identifizierung der Hintergrundpixel

Der bei Microarray-Experimenten auftretende Hintergrund ist das additive Resultat aus unspezifischer Adsorption und Eigenfluoreszenz der Chipoberfläche. Zumeist werden zu seiner numerischen Kalkulation Pixel verwendet, welche sich in Nachbarschaft des *Spots* befinden. Dabei kommen unterschiedliche Algorithmen zur Anwendung. Es existieren mittlerweile sogar Auswerteverfahren, welche sich zur Normalisierung der Hintergrundfluoreszenz bedienen.

Die Signalintensität des *Spots* und des Hintergrundes wird durch Mittelwertbildung über die zugehörigen Pixel ermittelt. Dabei kann entweder der Median oder der *Mean* als Mittelwert fungieren. Der Median generiert dabei zumeist die robusteren Werte, da er per Definition weniger ausreißeranfällig ist. Wenige sehr intensive Pixel eines *Spots*, z.B. verursacht durch stark fluoreszierende Artefakte, verändern den Mean signifikant, haben aber keine Auswirkungen auf den Median. Neben den Signalintensitäten von *Spot* und Hintergrund gibt das verwendete Softwarepaket zur primären Datenanalyse zumeist noch die Standardabweichungen von *Spot*- und Hintergrundintensität, den *Spot*durchmesser, die Anzahl der zum *Spot* gehörigen Pixel und *Flags* an. *Flags* fungieren als eine Art Qualitätskontrolle bei *Spots*, die bestimmten Qualitätskriterien nicht genügen. Dies kann z.B. der Fall sein, wenn die Standardabweichung relativ zur Intensität des *Spots* sehr hoch ist (*bad feature*), die *Spot*intensität geringer als die des Hintergrundes ist (*negative feature*), oder sehr klein ist (*dark feature*). Es lassen sich aber auch manuelle Qualitätskriterien erstellen, anhand derer die Software die *Spots* überprüft (*manual flag*).

3.4.2 Sekundäranalyse

Bevor die aus der Primäranalyse generierten Daten benutzt werden, um eine wissenschaftliche Fragestellung zu beantworten, muss eine hohe Datenqualität sichergestellt sein. Dies liegt zum Teil auch darin begründet, dass Microarray-Experimente vielen Variabilitäten unterliegen. Eine Validierung der Daten, auch mit anderen Analysemethoden, ist daher wünschenswert. Als erster Schritt zu diesem Zweck erfolgt die Eliminierung der *geflaggten* Daten. Zumeist werden diese Daten einfach aus der Analyse entfernt, was allerdings einen Verlust potenziell wertvoller Daten beinhaltet. Eine Alternative besteht darin, manuell die *geflaggten Spots* zu untersuchen und den Grund ihrer Klassifizierung zu ermitteln. Gegebenfalls erfolgt dann eine neue Primäranalyse mit anderen Parametern für diesen *Spot*. Dieses Verfahren erfordert allerdings einen hohen Zeit- und Ressourcenaufwand. Der nächste Schritt der Sekundäranalyse ist die Subtraktion des Hintergrundes von der Signalintensität. Diese validierte Methodik funktioniert zuverlässig, einzig bei *Spots* deren Intensität geringer als die des Hintergrundes ist erfordert sie eine differenzierte Betrachtung. Entweder können diese *Spots* von der Analyse entfernt werden, die niedrigste verfügbare Signalintensität eins zugewiesen bekommen, oder durch spezielle Algorithmen (z.B. Bayesian-Algorithmus) analysiert werden. Die Rohdaten der *ungeflaggten* und hintergrundbereinigten Daten werden nun in Logarithmen konvertiert. Zum einen soll damit eine Gleichverteilung der *Spots* über die Intensitätsskala erreicht werden und die Variabilitäten konstant und unabhängig vom Intensitätsniveau sein. Zudem sollten die Intensitäten und experimentellen Fehler einer Normalverteilung nach Gauß folgen. Da aus empirischen Gründen häufig ein Genexpressionsquotient von 2 als Grenzwert für regulierte Gene angewandt wird, verwendet man meist den Logarithmus zur Basis 2. Nichtregulierte Gene weisen so ein *logratio* von 0, zweifach regulierte von 1 oder -1 auf.

3.4.3 *Within-Array-Normalisierung*

Die gebräuchteste Chipapplikation beinhaltet den Vergleich zweier unterschiedlich markierter Proben auf einem Microarray. Erfolgt die Markierung z.B. mit zwei unterschiedlichen Fluoreszenzfarbstoffen, so ergeben sich daraus systematische Fehlerquellen, die zu hohen Variabilitäten führen können. Zwei Proben werden mit zwei unterschiedlichen Farbstoffen in zwei separaten chemischen Reaktionen markiert. Ihre Anregung erfolgt zudem über zwei verschiedene Laser und bei zwei verschiedenen Wellenlängen. Zusammengefasst existieren folgende Quellen für Variabilitäten:

- Die Inkorporation der Fluorophore kann unterschiedlich sein, möglicherweise variiert die Inkorporationsrate auch bei unterschiedlich häufigen mRNA-Spezies
- Die Fluoreszenzfarbstoffe können ein unterschiedliches Emissionsverhalten aufweisen, auch verursacht durch Anregung von zwei verschiedenen Lasern bei zwei verschiedenen Wellenlängen
- Die Emissionen der zwei Fluorophore (zwei verschiedene Wellenlängen und damit verschiedene Energien der Strahlung) werden eventuell unterschiedlich im PMT detektiert, auch abhängig von den jeweiligen Intensitäten der Strahlung
- Die Intensitäten der Farbstoffe können auf unterschiedlichen Positionen des Arrays unterschiedlich sein z.B. verursacht durch nicht-planare Bereiche auf dem Chip

Daher muss sichergestellt werden, dass eine gemessene differentielle Expression nicht das Resultat von systematischen Fehlern des Experimentators und des experimentellen Designs durch Verwendung zweier unterschiedlicher Farbstoffe ist. Man bedient sich zu diesem Zwecke dreier verschiedener Normalisierungsmethoden:

1. Lineare Regression
2. Lineare Regression der *logratios* gegen die mittlere Signalintensität
3. Nicht-lineare Regression

Für Positionseffekte auf Microarrays existieren separate Korrekturmöglichkeiten, zum Beispiel über eine 2D- oder *block-by-block* Loess-Regression. Alle hier beschriebenen Normalisierungsmethoden machen die Kernannahme, dass die Mehrheit der Gene auf dem Microarray nicht differentiell exprimiert ist. Daraus folgt, dass solche Normalisierungsmethoden meistens nur für *whole-genome* Microarrays anwendbar sind, da diese Voraussetzung für applikationsspezifische *low-density* Chips häufig nicht gegeben ist. Für derartige Chips bedient man sich anderer Normalisierungsverfahren, zum Beispiel über Referenzgene, artifizielle RNA-Spezies oder über die Hintergrundintensität.

Lineare Regression

Die lineare Regression erfolgt durch Auftragung der Daten in einem sogenannten *Scatterplot*. Die hintergrundbereinigten Fluoreszenzintensitäten eines Kanals werden dabei gegen die des anderen Kanals aufgetragen. Ein Beispiel für einen *Scatterplot* ist in Abbildung 6 gezeigt.

Die lineare Regression durch alle Datenpunkte sollte eine Gerade durch den Ursprung mit der Steigung 1 ergeben. Abweichungen davon repräsentieren ein ungleiches Verhalten

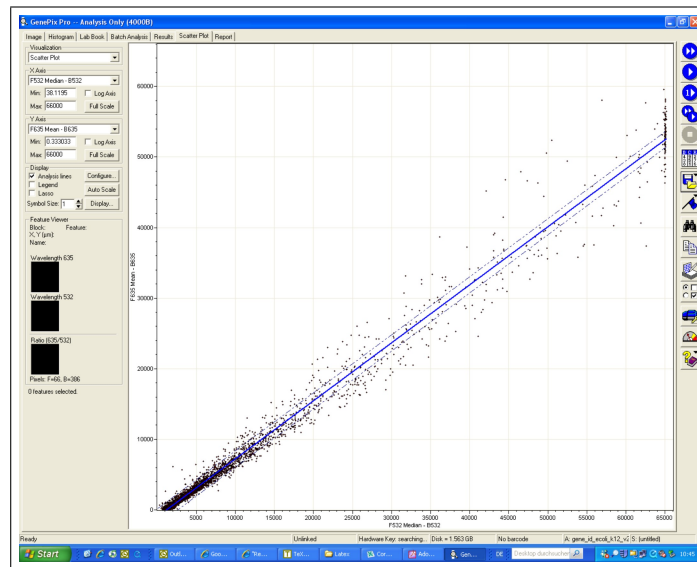


Abb. 6: Scatterplot eines Microarrayexperimentes. Aufgetragen ist die Signalintensität des Cy5-Kanals (Ordinate) gegen die Signalintensität des Cy3-Kanals (Abzisse).

der Fluoreszenzkanäle. Ein Achsenabschnitt ungleich 0 deutet so auf eine stärkere Fluoreszenzantwort eines Kanals bei kleinen Intensitäten hin, für eine Steigung größer 1 gilt die getroffene Aussage für hohe Intensitäten. Abweichungen von der Geraden deuten auf nicht-Linearitäten der Fluoreszenzantwort hin. Die Normalisierung der Daten erfolgt über die ermittelte Regressionsgerade, die Intensitäten eines Fluoreszenzkanals werden anhand ihrer neu berechnet. Die Normalisierung über lineare Regression funktioniert zuverlässig bei Daten, die einen linearen Zusammenhang genügen, weist allerdings zwei Nachteile auf. Die lineare Regression behandelt die beiden Fluoreszenzkanäle unterschiedlich und würde bei einer inversen Auftragung ein anderes Ergebnis generieren. Außerdem vermag das menschliche Auge nicht-lineare Abweichung von diagonalen Linien schlechter zu realisieren als von vertikalen oder horizontalen Linien, wodurch Nicht-Linearitäten schlechter erkannt werden. Deshalb bedient man sich häufiger der linearen Regression der *logratios* gegen die mittlere Signalintensität über den sogenannten *MA-Plot*. Beim *MA-Plot* ist das *logratio* gegen den Mittelwert der logarithmierten Fluoreszenzintensitäten beider Farbstoffkanäle aufgetragen. Ein *MA-Plot* vor und nach durchgeführter Normalisierung ist in Abbildung 7 gezeigt. Bei einem identischen Verhalten der beiden Kanäle sollten die Datenpunkte symmetrisch um eine horizontale Linie durch 0 verteilt sein. Abweichungen repräsentieren unterschiedliches Verhalten der beiden Kanäle. Lineare und nicht-lineare Normalisierungen über den *MA-Plot* behandeln beide Fluoreszenzkanäle gleich und generieren robustere und reproduzierbarere Daten. Die lineare Normalisierung erfolgt dadurch, dass eine Gerade durch die Daten gelegt wird und die Werte der Geradengleichung am entsprechenden Punkt von dem *logratio* jedes Gens subtrahiert werden.

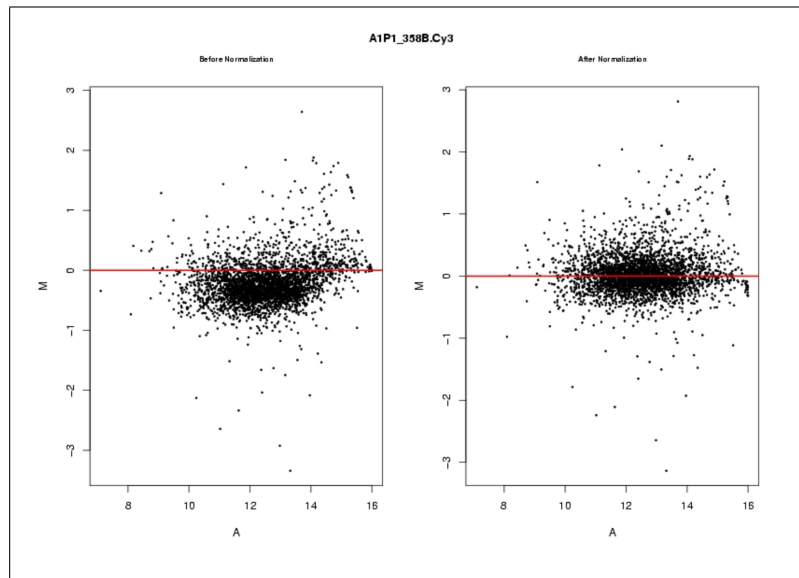


Abb. 7: MA-Plot eines Microarrayexperimentes vor (linkes Bild) und nach (rechtes Bild) durchgeführter Normalisierung.

Nicht lineare Regression

Das Verhältnis zweier Fluoreszenzkanäle bei Microarray-Experimenten folgt oftmals nicht einem linearen Zusammenhang. Meistens wird für die Normalisierung eine Loess-Regression verwendet. Dieser Terminus steht für *locally weighted polynomial regression* und funktioniert über eine große Anzahl lokaler Regressionen in überlappenden Regionen des gesamten Datenbereichs. Die einzelnen Regressionen werden anschließend zu einer glatten Kurve vereint. Loess-Regressionen sind in vielen *Softwaretools* zur Datenanalyse von Microarrays implementiert. Einige Parameter der Loess-Regression müssen vom Anwender selbst gewählt werden, am bedeutendsten ist die Fensterbreite der lokalen Regression. Wird diese zu schmal gewählt reagiert die entstehende Kurve zu sensitiv auf lokale Schwankungen, bei einer zu großen Fensterbreite tritt der gegenteilige Effekt auf.

Korrektur von Positionseffekten

Positionseffekte liegen dann vor, wenn in manchen Regionen des Chip ein Kanal gegenüber den anderen eine konsistent höhere Fluoreszenzintensität aufweist, welche nicht das Resultat von einer differentiellen Expression ist. In diesen Regionen ist das *logratio* dann konsistent negativ bzw. positiv. Derartige Effekte sind das Resultat einer fehlenden Planarität der Chipoberfläche bzw. einer nicht vollständig horizontalen Position im Scanner. Die Tiefe des Fokus der beiden verwendeten Laser hängt proportional von der Wellenlänge ab. In einigen Regionen können beide Kanäle im Fokus sein, während in anderen Regionen sich nur der Kanal mit der größeren Wellenlänge noch im Fokus befindet. Zur Normalisierung derartiger Effekte existieren im Wesentlichen zwei Methoden, die 2D-Loess Regression und die *block-by-block* Loess Regression. Bei der 2D-Loess Regression handelt es sich grundsätzlich um die genauere Methode, da die Grenze von Positionseffekten sich auch innerhalb eines Blocks befinden kann.

3.4.4 *Between-Array*-Normalisierung

Um Proben von verschiedenen Chips statistisch signifikant vergleichen zu können, bedarf es weiterer Normalisierungsverfahren, da sich jede Hybridisierung leicht voneinander unterscheiden kann. Die Visualisierung verschiedener Verteilungen erfolgt in der Microarray-Datenanalyse häufig über sogenannte *Box-Plots*. In ihnen wird ein Vergleich der Verteilungen der logarithmierten Intensitäten oder *logratios* von Genen auf unterschiedlichen Microarrays durchgeführt. Die Box repräsentiert dabei die Standardabweichung, die horizontalen Linien repräsentieren die Extremwerte und die symmetrisch durch die Box verlaufende Linie den Mittelwert der Verteilung. Die nun getroffene Annahme geht davon aus, dass die Variationen in der Verteilung der Arrays das Resultat von experimentellen Fehlern sind und nicht die biologischen Variabilitäten repräsentieren. Es existieren drei Methoden zur Normalisierung:

1. Skalierung
2. Zentrierung
3. Verteilungsnormalisierung

Eine Übersicht über die drei Normalisierungsmethoden wird in Abbildung 8 gegeben. Die Skalierung stellt sicher, dass der Mittelwert aller Verteilungen gleich ist. Dies erfolgt durch Subtraktion des Mittelwertes der Verteilung von allen *logratios* oder logarithmierten Signalintensitäten der einzelnen Gene. Daraus resultiert ein Mittelwert der Verteilung von 0. Eine Alternative stellt die Verwendung des Medians dar, da dieser weniger anfällig für Ausreißer ist. Die Zentrierung führt dazu, dass der Mittelwert und die Standardabweichung der Verteilung für alle Chips gleich ist. Nach Subtraktion des Mittelwertes erfolgt die Division durch die Standardabweichung für jeden Datenpunkt, so dass der Mittelwert der Verteilung 0 und die Standardabweichung 1 ist. Alternativ bietet sich wiederum die Verwendung des Medians anstelle des *Means* an. Bei der Verteilungsnormalisierung wird zusätzlich zu den Normalisierungen der Zentrierung die Verteilung der Daten auf jedem Chip angeglichen. Nach der Zentrierung der Daten wird für jeden Chip eine Sortierung der Daten nach der Größe des Wertes von klein nach groß vorgenommen. Anschließend erfolgt die Berechnung einer neuen Verteilung bei der der kleinste Wert der Mittelwert des am schwächsten exprimierten Gens ist, der zweitniedrigste Wert der Mittelwert des am zweitniedrigsten exprimierten Gens und schließlich der höchste Wert der Mittelwert des am stärksten exprimierten Gens. Final erfolgt eine Substitution jedes Messwertes durch den korrespondierenden Mittelwert der neuen Verteilung.

3.4.5 Identifizierung differentiell exprimierter Gene

Nachdem die Qualität der generierten Daten durch *Within-Array*- und *Between-Array*-Normalisierung optimiert wurde, kann die statistische Identifizierung der regulierten Gene erfolgen. In früheren Experimenten wurde dazu ein empirisch ermittelter Grenzwert des Genexpressionsquotienten von 2 bzw. 1/2 verwendet. Bei größeren bzw. kleineren Quotienten wurde ein Gen als differentiell exprimiert betrachtet. Diese rein empirische Vorgehensweise genügt allerdings nicht statistischen Anforderungen an ein signifikantes Ergebnis. Neben dem reinen Wert des Quotienten sollten Variabilitäten zwischen den Experimenten und die Anzahl der Wiederholungen in Betracht gezogen werden. Die experimentellen Daten werden zu diesem Zweck einem Hypothesen-Test unterzogen, welcher auf der Nullhypothese beruht.

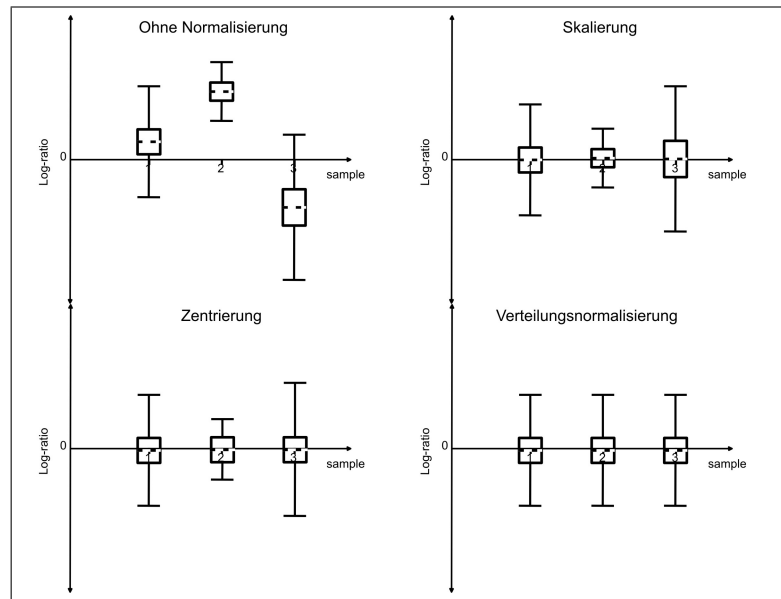


Abb. 8: Verfahren zur *Between-Array-Normalisierung*. Gezeigt ist ein Datensatz ohne Normalisierung (links oben) sowie nach Anwendung der Zentrierung (links unten), der Skalierung (rechts oben) und Verteilungsnormierung (rechts unten).

Diese geht davon aus, dass die ermittelte Regulation eines Gens nicht das Resultat eines biologischen Effekts ist, sondern aufgrund von Unterschieden unter den Individuen oder systematischen Messfehlern auftritt. Die Wahrscheinlichkeit diese differentielle Genexpression durch Zufall aus den Messwerten zu erhalten wird durch den p-Wert beschrieben. Je niedriger dieser Wert ist, desto unwahrscheinlicher ist es, dass das beobachtete Ergebnis das Resultat eines Zufalls ist und desto signifikanter ist das Ergebnis. Ein p-Wert von 0,02 bedeutet zum Beispiel eine zweiprozentige Wahrscheinlichkeit. p-Werte werden häufig mittels einer t-Statistik durch Vergleich mit einer t-Verteilung mit entsprechender Anzahl an Freiheitsgraden kalkuliert. Für die Anwendung eines t-Testes muss zuerst bestimmt werden, ob es sich bei dem vorliegenden Datensatz um gepaarte oder ungepaarte Daten handelt. Gepaarte Daten sind zum Beispiel Werte eines Zeitverlaufes, da sie von der gleichen Probe, nur zu unterschiedlichen Zeitpunkten stammen. Ungepaarte Daten sind völlig unabhängig voneinander und stehen in keinerlei Beziehung zueinander. Es kann sich zum Beispiel um den Vergleich von verschiedenen Individuen der gleichen Art mit einem unterschiedlichen Phänotyp handeln. Der t-Test für ungepaarte Daten unterscheidet sich etwas von demjenigen, welcher für gepaarte Daten angewandt wird. Ein t-Test für gepaarte Daten weist folgende mathematische Beschreibung auf:

$$t = \frac{\bar{x}}{s/\sqrt{n}}$$

\bar{x} ist der Mittelwert des *logratios*, s die Standardabweichung und n die Anzahl der Proben. Der p-Wert wird daraus durch einen Vergleich mit einer t-Verteilung mit einer ausreichenden Anzahl an Freiheitsgraden kalkuliert. Die Freiheitsgrade sind dabei die Anzahl der unabhängigen Variablen. t-Tests sind ebenfalls in vielen Software-Paketen wie R oder Matlab implementiert. Voraussetzung für die Anwendbarkeit eines t-Test ist die Normalverteilung der Daten.

3.4.6 Genomik und Clusteranalysen

Der große Vorteil der Microarray-Analyse liegt nicht in der isolierten Betrachtung einzelner Gene, sondern in der Möglichkeit, die Interaktionen verschiedener Gene, welche einen Phänotyp verursachen, zu erfassen. Biochemische Vorgänge sind zumeist höchst komplex und entstehen durch Wechselwirkung vieler verschiedener Komponenten. Die simultane Betrachtungsmöglichkeit des gesamten Transkriptoms einer Zelle stellt sich hierbei von unschätzbarem Wert heraus. Die Ähnlichkeit von Genen oder Proben lässt sich durch Vergleich ihrer Expressionsprofile ermitteln und wird über den Abstand zueinander quantifiziert. Es existieren drei Abstandmaße zu diesem Zweck. Die Wahl des geeigneten Maßes für die jeweilige Applikation hat entscheidenden Einfluss auf das generierte Ergebnis.

1. Korrelationskoeffizient
2. Spearman Korrelation
3. Euklidische Distanz

Der Korrelationskoeffizient nimmt Werte zwischen -1 und 1 an, wobei +1 einer perfekten positiven Korrelation, 0 keinerlei Korrelation und -1 einer perfekten negativen Korrelation entspricht. Bei zentrierten Daten wird der Korrelationskoeffizient r über folgende mathematische Gleichung definiert:

$$r = \sum x_i \cdot y_i$$

Geometrisch stellt der Korrelationskoeffizient das Skalarprodukt zweier Vektoren im n -dimensionalen Raum dar. Parallele Vektoren weisen, abhängig von ihrer Orientierung, eine Korrelation von +1 oder -1 auf, orthogonale Vektoren von 0. Für das Distanzmaß des Korrelationskoeffizienten gilt:

$$d(x, y) = 1 - (r(x, y))$$

Ein Problem, das sich bei der Verwendung des Korrelationskoeffizienten als Abstandsmaß ergibt, ist seine Anfälligkeit für Ausreißer. Besonders für Microarray-Analysen ist dieses von Bedeutung, da die Daten aufgrund der großen Anzahl parallel studierter Gene und vieler Variabilitäten häufig ein starkes Rauschen aufweisen. Daher ist die Verwendung der ausreißerunanfälligen Spearman Korrelation bei Microarray-Analysen dem Korrelationskoeffizienten häufig vorzuziehen. Die euklidische Distanz beschreibt den geradlinigen Abstand zweier Punkte im 2- oder 3-dimensionalen Raum und kann über den Satz von Pythagoras errechnet werden. Ein Problem der Verwendung der euklidischen Distanz liegt darin begründet, dass sie abhängig von der Größenordnung der Daten ist. Ähnliche Expressionsprofile auf unterschiedlichem Niveau weisen eine hohe euklidische Distanz auf. Gleichwohl besteht eine signifikante Ähnlichkeit zwischen ihnen. Dieses Problem lässt sich durch die Verwendung zentrierter Daten umgehen.

Cluster-Analysen sind ein sehr häufig eingesetztes Mittel zur Identifizierung ähnlicher Expressionsverläufe von Proben oder Genen. Man unterscheidet das hierarchische Clustern von nicht-hierarchischen Clustermethoden. Durch die enorme Vereinfachung der großen und unübersichtlichen Datenmenge stellen sie ein wertvolles Mittel zur biologischen Interpretation der Daten da. Gruppen von ähnlichen Genen werden identifiziert und können detailliert analysiert werden.

Beim hierarchischen Clustern werden die Proben oder Gene in einer Baumstruktur organisiert. Ähnliche Expressionsverläufe erscheinen nahe beieinander im sogenannten Dendrogramm, während nicht-korrelierte Daten weiter entfernt erscheinen. Das Ergebnis einer hierarchischen Clusteranalyse ist in Abbildung 9 wiedergegeben. Die Organisation in Dendrogramm wurde schon erfolgreich beim Vergleich von Sequenzdaten oder anderer Parameter in der Verwandtschaftsanalyse durchgeführt (phylogenetische Bäume). Beim hierarchischen *Clustering* wird die Distanzmatrix zwischen den einzelnen Genen betrachtet und die zwei Gene mit dem geringsten Abstand zueinander in einem neuen Cluster klassifiziert. Nun erfolgt eine erneute Bestimmung der Abstandsmatrix zwischen dem neu gebildeten Cluster und den weiteren Genen. Wiederum werden die nächsten Gene bzw. Cluster zu einem neuen Cluster vereint. Dieser Vorgang wird solange wiederholt bis alle Gene miteinander vernetzt sind.

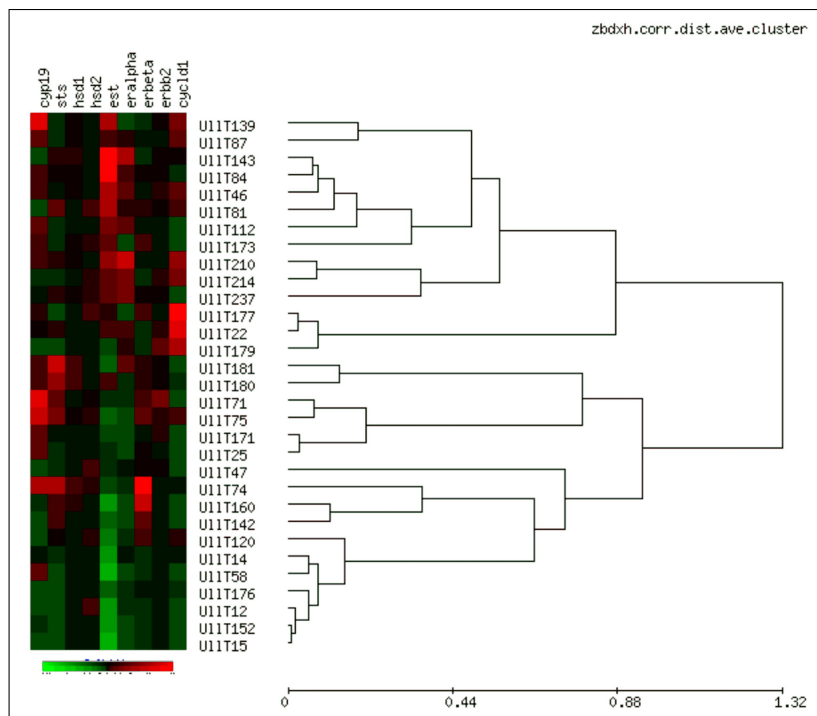


Abb. 9: Hierarchisches *Clustering*. Gene, die eine hohe Similarität zueinander aufweisen, erscheinen nahe beieinander im Dendrogramm.

Die Bestimmung des Abstandes zwischen zwei Clustern kann auf dreierlei Arten erfolgen: *single*, *average* und *complete linkage*. Die drei Arten den Abstand zwischen Clustern zu bestimmen sind in Abbildung 10 illustriert. Bei der Bestimmung über *single linkage* werden die beiden Gene, welche den geringsten Abstand zwischen den Clustern aufweisen, als Abstandsmaß verwendet. Für die Bestimmung über *complete linkage* fungieren die beiden Gene mit dem größten Abstand zueinander als Abstandsmaß und beim *average linkage*-Verfahren werden von allen Genen die Abstände zwischen den Clustern ermittelt und daraus der Mittelwert gebildet.

Das generierte Clusterergebnis sollte auf seine Eignung hinsichtlich einer biologischen Interpretation validiert werden. Dieses kann sowohl visuell erfolgen, als auch anhand bekannter und biologisch sinnvoller Zusammenhänge (z.B. Gene eines Stoffwechselpfades finden sich in

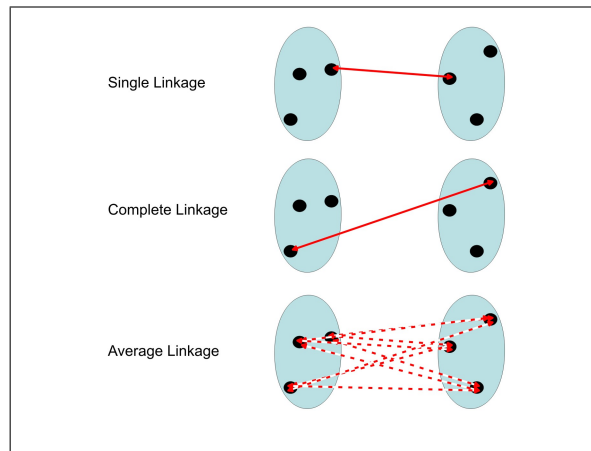


Abb. 10: Verbindungsmethoden zur Abstandsbestimmung zwischen verschiedenen Clustern und Genen.

einem Cluster wieder). Daneben lassen sich allerdings auch statistische Methoden verwenden um die Güte des Clusterergebnisses zu verifizieren.

3.4.7 *K-means Clustering*

Dieser Clusteralgorithmus unterscheidet sich vom hierarchischen Clustern auf dreierlei Arten:

1. Die Anzahl der Cluster muss im Voraus festgelegt werden
2. Zu Beginn erfolgt eine zufällige Verteilung der Gene in die jeweiligen Cluster. Jeder *k-means* Lauf kann so zu etwas variierenden Ergebnissen führen
3. Es existiert keine hierarchische Struktur oder Verbindung zwischen den Clustern, lediglich ähnliche Expressionsprofile fallen in ein Cluster

Die Ergebnisse der ersten Berechnung liefern die Grundlage für die zweite Berechnung. Der Algorithmus verläuft über folgende Teilschritte:

1. Wahl der Anzahl an Clustern (k)
2. Zufällige Verteilung der Genexpressionsprofile in eines der k Cluster
3. Für jedes Profil wird der Abstand zwischen ihm und dem Zentroid der verschiedenen Cluster bestimmt
4. Wenn das Profil näher an einem anderen Cluster als dem bis dato zugehörigen liegt, wird es dem neuen Cluster zugeordnet

Die Validierung der *k-means* Ergebnisse kann wiederum visuell, über die biologische Sinnhaftigkeit und über statistische Methoden erfolgen. Zusätzlich sollte allerdings verifiziert werden, dass ein erneutes Clustern mit gleichem k -Wert zu einem äquivalenten Ergebnis führt.

3.5 Quantitative real-time PCR

Bei der *real-time* PCR handelt es sich um die *state-of-the-art* Technik zur Nukleinsäurequantifizierung, Mutationsdetektion, Genotypisierung und Chimärenanalyse. Diese Methode weist die größte Sensitivität aller Genexpressionsanalysetechniken, eine hohe Reproduzierbarkeit und einen enormen dynamischen Bereich über mehr als 8 Zehnerpotenzen auf. Das Prinzip basiert dabei auf dem einer konventionellen PCR. Durch Zusatz von Fluorophoren, die entweder als sequenzunspezifische, in dsDNA interkalierende Farbstoffe vorliegen oder an genspezifische Sonden gekoppelt sind, lässt sich der gesamte Amplifikationsprozess der PCR in *real-time* beobachten und analysieren. Die daraus abgeleiteten Amplifikationskurven können zur relativen oder absoluten Quantifizierung der analysierten Nukleinsäure verwendet werden. In einer nachfolgenden Schmelzkurvenanalyse kann das amplifizierte Produkt über seine Schmelztemperatur, welche eine Funktion aus Produktlänge und Basenzusammensetzung darstellt, näher charakterisiert werden. Des Weiteren können unspezifische Koamplifikate und Primerdimere detektiert werden.

Quantitative Analyse

Die Amplifikation mittels PCR lässt sich in drei Phasen einteilen, eine initiale *lag*-Phase, die exponentielle Phase, die schließlich in eine Plateau-Phase mündet. Dieses Amplifikationsprofil ist in Abbildung 11 dargestellt. Nur die exponentielle Phase kann zur exakten Nukleinsäurequantifizierung verwendet werden. Prinzipiell lässt sich die exponentielle Phase durch Extrapolation auch ab Zyklus 0 darstellen, der dabei gemachte Fehler ist allerdings zu groß. Das erhaltene Fluoreszenzsignal ist dabei proportional zur Menge des PCR-Produktes. Bei einer idealen PCR findet in jedem Zyklus eine Verdoppelung des Produktes statt. Mathematisch wird dieses daher mittels einer Exponentialfunktion mit der Basis 2 beschrieben. Der Betrag der Basis wird als Effizienz der Amplifikation bezeichnet, eine ideale PCR verläuft daher mit einer Effizienz von zwei. Reale PCR-Amplifikationen weisen allerdings Effizienzen auf, welche kleiner als zwei sind. Meistens nimmt die Effizienz, abhängig von amplifizierter Sequenz und verwendeten Primer, Werte zwischen 1,7 und 1,9 an.

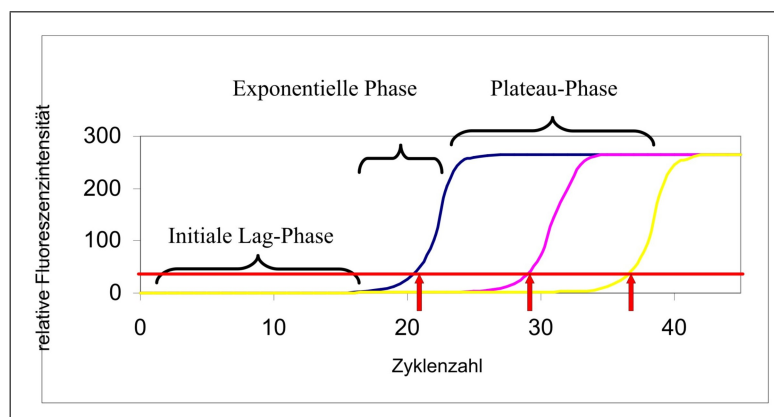


Abb. 11: Amplifikationskurven in der *real-time* PCR. Der Amplifikationsprozess lässt sich in die *lag*-Phase, die exponentielle Phase und die Plateau-Phase unterteilen.

Gründe für die verringerte Effizienz liegen in der Existenz von PCR-Inhibitoren und der suboptimalen Wahrscheinlichkeit für das Zusammentreffen von Primern, *Template* und Polymerase sowie Fehlern beim Amplifikationsprozess. Niedrigere Effizienzen verringern die

Sensitivität, steigern allerdings die Genauigkeit der ermittelten Werte. Für das erhaltene Fluoreszenzsignal S ergibt sich daher folgende mathematische Beschreibung:

$$S = k \cdot N_0 \cdot E^C$$

Der Proportionalitätsfaktor k stellt dabei den funktionellen Zusammenhang zwischen der Menge an PCR-Produkt und der daraus resultierenden Signalintensität S dar. N bezeichnet die Anzahl der initial vorliegenden Templatemolekülen, E die Effizienz und C die Zyklenanzahl der PCR. Daraus leitet sich folgender logarithmischer Zusammenhang zwischen Zyklenzahl und initial vorliegenden *Templatemolekülen* her:

$$C = -(\log E)^{-1} \cdot (\log k + \log N_0 - \log S)$$

mit $m = -(\log E)^{-1}$ und $b = -(\log E)^{-1} \cdot (\log k + \log S)$ vereinfacht sich obige Beziehung zu :

$$C = m \cdot \log N_0 + b$$

Diese Gleichung beschreibt den linearen Zusammenhang zwischen dem Logarithmus der *Templatekonzentration* und der Zyklenzahl. In der *real-time* PCR wird zur Quantifizierung eine Zyklenzahl verwendet, bei der das Fluoreszenzsignal signifikant über dem Hintergrundrauschen liegt und der Gesamtprozess sich noch in der exponentiellen Phase der Amplifikation befindet. Dieser Wert wird als C_T -Wert bezeichnet ($T = \text{threshold}$). Durch Auftragung der ermittelten C_T -Werte einer Verdünnungsreihe gegen den Logarithmus der (relativen) Templatekonzentration lässt sich so aus der Steigung m der Geraden die Effizienz der Amplifikation ermitteln:

$$E = 10^{\left(\frac{-1}{m}\right)}$$

Die differentielle Expression eines Genes in zwei Zuständen, ausgedrückt über den Quotienten der Templatemoleküle, lässt sich wie folgt ermitteln:

$$\frac{N_{01}}{N_{02}} = \frac{E^{C_{T2}}}{E^{C_{T1}}} = E^{C_{T2}-C_{T1}} = E^{\Delta C_T}$$

Für die Verifikation der differentiellen Genexpression beim Vergleich zweier Zuständen (z.B. Kontrolle vs. Probe) verwendet man als Referenz eine artifizielle RNA-Spezies oder mehrere *housekeeping*-Gene. Damit soll sicher gestellt werden, dass die festgestellte Regulation nicht das Resultat von experimentellen Fehlern ist. Für den Quotienten der Genexpression unter Einbezug einer Referenzprobe ergibt sich daher Folgendes:

$$Q = \frac{(E_P)^{\Delta C_T(\text{Kontrolle}-\text{Probe})}}{(E_R)^{\Delta C_T(\text{Kontrolle}-\text{Probe})}}$$

E_P beschreibt die Amplifikationseffizienz der Probe und E_R diejenige der Referenz. Diese Auswertemethode wurde erstmals von Pfaffl *et al.* (2001) beschrieben. Eine weitere, häufig verwendete Methode ist die Delta-Delta-Ct-Methode. Diese geht vereinfachend davon aus, dass die Amplifikationseffizienz von Probe und Referenz gleich bzw. 2 ist. Damit vereinfacht sich die obrige Gleichung zu:

$$Q = 2^{(\Delta C_t(\text{Probe}) - \Delta C_t(\text{Kontrolle}))}$$

Da die Näherung gleicher Effizienzen bzw. sogar einer Effizienz von 2 mit sehr großen Fehlern behaftet sein kann, ist die Pfaffl-Methode in jedem Fall vorzuziehen.

Detektionsformate in der real-time PCR

Prinzipiell unterscheidet man bei Reporterfluorophoren für *real-time* PCR-Applikationen zwischen doppelstrangspezifischen Fluoreszenzfarbstoffen und sequenzspezifischen Fluoreszenzsonden. Im Folgenden werden ausschließlich die doppelstrangspezifischen Fluoreszenzfarbstoffe behandelt, für Erläuterungen zu sequenzspezifischen Sonden sei auf die Literatur verwiesen (z.B. Wilhelm *et al.* (2003)). Bei den Farbstoffen findet am häufigsten der asymmetrische Cyaninfarbstoff *SybrGreen I* Verwendung. Dieses Fluorophor interkaliert weitgehend sequenzunspezifisch in die kleine Furche doppelsträngiger DNA. Die Fluoreszenzintensität des gebundenen Farbstoff übersteigt die des freien Fluorophor um mehr als den Faktor 1000. Das Anregungsmaximum liegt dabei bei 480 nm und das Emissionsmaximum bei 520 nm, das Fluorophor lässt sich daher mit konventionellen Argon-Lasern anregen. *Quenching*-Effekte wie sie bei *SybrGreen I* beschrieben wurden sind in der *real-time* Detektion nicht von Bedeutung, da das PCR-Produkt mit dem Farbstoff gesättigt ist (Farbstoff/Basenpaarverhältnis größer 2) und das *Quenching* proportional zur Menge des PCR-Produktes ist. Durch die Verwendung sequenzunspezifischer Farbstoffe wird *de facto* jede doppelsträngige DNA-Spezies nachgewiesen, daher auch unspezifische Koamplifikate und Primerdimere. Dieses Faktum wird häufig als genereller Nachteil bei der Verwendung von sequenzunspezifischen Farbstoffen angeführt. Für ein *real-time* Experiment muss diese Fehlerquelle allerdings unabhängig vom Detektionsformat ausgeschlossen werden. Koamplifizierte Produkte führen zu fehlerhaften Amplifikationseffizienzen und Quantifizierungsergebnissen. Zudem werden solche Produkte durch die Schmelzkurvenanalyse detektiert. Neben *SybrGreen I* existieren eine Reihe weiterer sequenzunspezifischer Farbstoffe mit verschiedenen Vor- und Nachteilen. Das allgemeine Prinzip auf dem diese beruhen ist jedoch identisch.

4 Experimentelle Arbeiten

Ziel dieser Arbeit war es mittels grundlegender Genexpressionsanalysen ein besseres Bild der Zusammenhänge im Sekretionsgeschehen zu erlangen. Für ein tiefergehendes mechanistisches Verständnis der Hauptsekretionspathways Sec und Tat ist die Kenntnis des Transkriptionstatus der Pathwaygene zu jedem Zeitpunkt der Sekretion entscheidend. Zusätzlich waren die Auswirkungen der BRP-Expression auf die Transkription der Zelle Ziel der Analysen. Aufgrund der Komplexität der involvierten Vorgänge ist eine globale Betrachtungsweise auf Transkriptomebene indiziert. Für die angestrebten Evaluierungen sollten daher DNA-Microarray Analysen mittels *whole-genome* Arrays fungieren. Aufgrund des hohen Parallelisierungsgrades und der Flexibilität des Analysendesigns stellt die DNA-Microarray-Technologie eine ideale Methodik für die gewünschte Applikation dar. Um die Qualität der erhaltenen Daten einschätzen zu können, wurden zunächst die bei Chipexperimenten auftretenden Variabilitäten untersucht.

4.1 Evaluierung der Variabilitäten von Chipexperimenten

Chipexperimente unterliegen in hohem Maße inhärenten Variabilitäten, die das additive Resultat der sukzessiv durchgeführten Prozessschritte in der experimentellen Durchführung sind (Schadt *et al.* (2000)). Als Quellen von Variabilitäten sind dabei vornehmlich zu nennen:

1. RNA-Isolierung
2. Reverse Transkription und/oder Fluoreszenzmarkierung
3. Hybridisierung
4. Variabilitäten aufgrund der Verwendung unterschiedlicher Markierungen
5. Experimentelle Faktoren

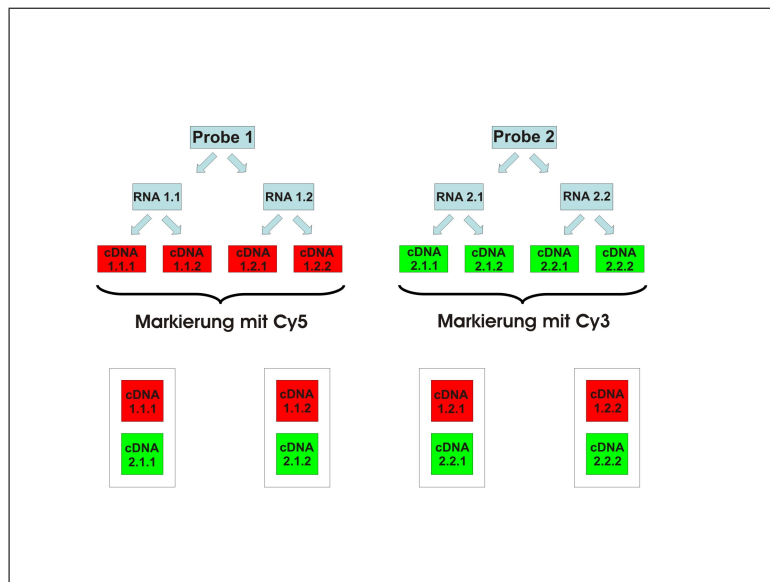


Abb. 12: Experimentelles Design zur Evaluierung der Variabilitäten von Chipexperimenten. Aus zwei biologischen Proben werden jeweils vier uniform markierte cDNAs generiert, welche aus zwei RNA-Isolationen der gleichen biologischen Probe hervorgingen. Die unterschiedlich markierten cDNAs werden paarweise gegeneinander auf Microarrays hybridisiert. Mittels dieses Experimentes lassen sich die Variabilitäten der RNA-Isolierung, der Reversen Transkription und der Hybridisierung evaluieren.

Zur Verifikation der experimentellen Methodik und der Identifikation von experimentellen Schritten, die mit besonders hohen Variabilitäten belastet sind, wurde daher nachfolgend beschriebenes Experiment designed und durchgeführt (siehe auch Abbildung 12). Zwei biologische Proben unterschiedlicher Zeitpunkte einer reproduzierbaren *E. coli*-Kultivierung im Bioreaktor dienten als Ausgangsmaterial. Aufgrund der hohen Individuenzahl, aus denen jede Probe besteht, wird davon ausgegangen, dass die biologischen Schwankungen zwischen den unterschiedlichen Individuen innerhalb der Population hinreichend erfasst sind. Aus jeder Probe wurde in drei getrennten Ansätzen RNA isoliert (in Abbildung 12 sind nur die zwei RNA-Isolationen gezeigt, welche in cDNA konvertiert werden) und photometrisch Konzentration und Reinheit bestimmt. Zwei RNA-Präparationen jeder Probe wurden anschließend

Tabelle 1: Die resultierenden Konzentrationen und Reinheiten, ausgedrückt durch den Quotienten A260/A280, für die sechs verschiedenen RNA-Isolationen der beiden biologischen Proben des Parallelexperimentes.

	Konzentration (ng/ μ l)	A260/A280
RNA 1.1	2459,02	2,16
RNA 1.2	2399,70	2,16
RNA 1.3	2427,69	2,18
RNA 2.1	2997,83	2,10
RNA 2.2	3014,66	2,14
RNA 2.3	3007,81	2,13

in jeweils zwei unabhängigen Ansätzen in cDNA konvertiert und durch Zusatz von Cy3- bzw. Cy5-dCTP zum Reaktionsansatz fluoreszenzmarkiert. Dabei wurde so verfahren, dass die cDNAs der ersten biologischen Probe alle Cy3-markiert vorlagen und jene cDNAs der zweiten Probe entsprechend Cy5-markiert waren. Aus einer biologischen Probe resultieren so vier uniform markierte cDNAs, welche aus zwei unterschiedlichen RNA-Präparationen hervorgegangen sind. Die Vorgehensweise ist schematisch in Abbildung 12 dargestellt.

Um eine maximale Reproduzierbarkeit hinsichtlich der RNA-Isolierung zu erzielen, wurde ein kommerziell erhältliches RNA-Isolierungskit (*InnuPrepTM* RNA Minikit, Analytik Jena) verwendet. Es existieren zwei generelle Möglichkeiten, um den differentiellen Transkriptionsstatus eines Gens zu erfassen. Zum einen kann ein Vergleich von RNA-Präparationen erfolgen, welche aus der gleichen Zellzahl gewonnen wurden oder für jede Probe die gleiche Menge an Gesamt-RNA oder mRNA verglichen werden. Die gebräuchlichere Methode besteht in der Verwendung gleicher RNA-Mengen, da die RNA-Isolierung aus gleichen Zellzahlen stets großen Variabilitäten unterliegt. Im Folgenden wurden Genexpressionanalysen immer auf der Basis gleicher RNA-Mengen durchgeführt.

Das zur RNA-Isolation verwendete Protokoll ist im Anhang (A.1.3) wiedergegeben. Von der so generierten RNA wurden Konzentration und Reinheit photometrisch bestimmt und die molekulare Integrität gelelektrophoretisch verifiziert. Die RNA-Präparationen enthielten nach der Isolierung noch Reste von Plasmid- und chromosomaler DNA, die bei der photometrischen Konzentrationsbestimmung mit erfasst würden. Daher erfolgte eine Eliminierung dieser DNA-Kontaminationen durch DNase I-Behandlung für 30 Minuten bei 37°C. Nachfolgend wurden die DNase I-behandelten RNA-Präparationen durch eine Phenol-Chloroform-Extraktion gereinigt und so die DNase I quantitativ entfernt. Die Quantität dieses Schrittes ist besonders bedeutsam, da Reste von DNase I die neu synthetisierte cDNA im nachfolgenden RT-Schritt hydrolysieren können. Tabelle 1 gibt die erzielten Konzentrationen und Reinheiten der verschiedenen RNA-Isolationen wieder.

Die ermittelten Konzentrationswerte der RNA-Isolationen, welche aus einer biologischen Probe hervorgingen, waren nahezu identisch. Die RNA-Isolationen der zweiten biologischen Probe, welche zu einem späteren Zeitpunkt der Kultivierung entnommen wurde, generierten konsistent höhere Konzentrationen als jene der ersten Probe. Der Quotient A260/A280, als Maß für die Reinheit der RNA-Präparation, nahm für alle Proben Werte im Bereich von

4.1 Evaluierung der Variabilitäten von Chipexperimenten

2,1 an. Für eine reine RNA-Lösung beträgt der Quotient 2, Abweichungen zu niedrigeren Werten deuten auf Proteinkontaminationen hin. Sowohl die geringe Variabilität der Konzentrationen zwischen den RNA-Isolationen der gleichen biologischen Probe als auch die hohen Reinheitsquotienten geben erste Hinweise auf eine hohe Güte und Reproduzierbarkeit der Isolation. Das in Abbildung 13 vorgestellte Gelbild bestätigt diese Annahme.

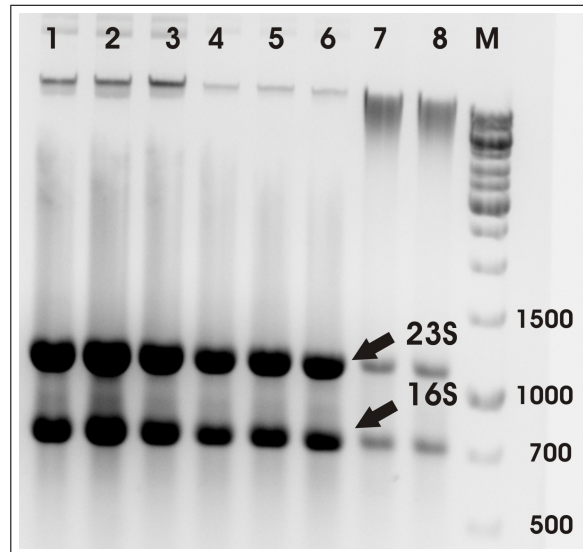


Abb. 13: Gelelektrophoretische Analyse der RNA-Isolationen aus den biologischen Proben 1 (früherer Kultivierungszeitpunkt) und 2 (späterer Kultivierungszeitpunkt). Gezeigt ist ein 1,5%iges Agarosegel, die Visualisierung der RNA erfolgte durch Ethidiumbromid-Färbung. Deutlich zu erkennen sind die beiden 23S und 16S rRNA-Banden mit nahezu identischen Intensitäten für die Spuren 1-3 (RNA aus Probe 2) und 4-6 (RNA aus Probe 1).

Auf den Spuren 1-3 sind drei RNA-Präparationen der zweiten biologischen Probe aufgetragen, auf den Spuren 4-6 jene der ersten biologischen Probe. Es sind für jede Probe zwei diskrete Banden hoher Intensität erkenntlich. Dabei handelt es sich um die 16S- und 23S-Banden der ribosomalen RNA, welche den Hauptbestandteil der Gesamt-RNA einer Zelle ausmachen. Wie aus dem Gelbild ersichtlich, sind kaum Unterschiede in der Bandenintensität zwischen den einzelnen RNA-Isolationen einer Probe visuell zu erfassen. Die Banden der zweiten Probe weisen eine etwas höhere Intensität auf als jene der ersten Probe. Bei allen aufgetragenen RNA-Isolationen sind noch Banden von Plasmid- und chromosomaler DNA zu erkennen. Sie laufen im Gel im höhermolekularen Bereich, hier knapp ober- und unterhalb der Bandenbeschriftung. Erkentlich an den zwei intakten ribosomalen Banden der 23S- und 16S-RNA, die im ungefähren Intensitätsverhältnis von 2:1 erscheinen, wird die molekulare Integrität der RNA belegt. Bei dem ebenfalls aufgetragenen Marker handelt es sich um einen DNA-Größenstandard, daher kann nicht von der Größe der Markerbanden auf die Größe der beiden ribosomalen RNA-Banden geschlossen werden. Dennoch erlaubt ein Vergleich mit dem Marker die Identifizierung degradierter RNA, diese erscheint unterhalb der 500 bp Bande. Gleiche Mengen an RNA werden nachfolgend in cDNA konvertiert und mit dem Standardprotokoll prozessiert (Anhang A.1.4). Die Konzentrationen an generierter cDNA und inkorporiertem Farbstoff wurden nachfolgend spektrophotometrisch am Nanodrop ermittelt. Die Absorptionsspektren sind in der Abbildung 14 gezeigt.

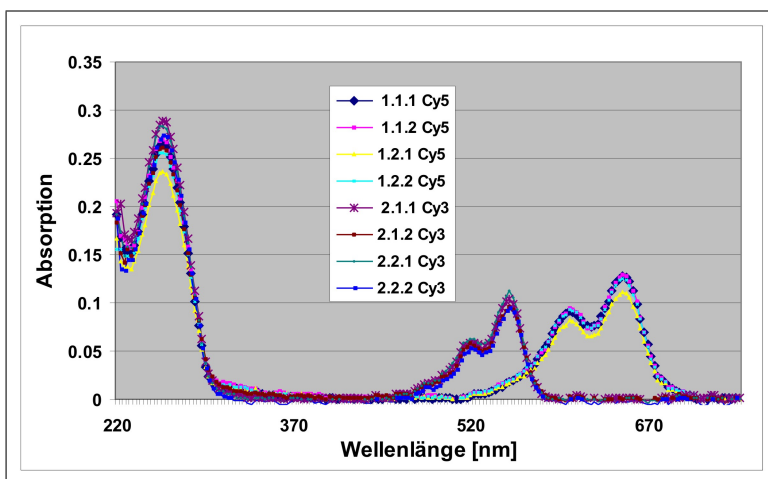


Abb. 14: Gezeigt sind die Absorptionsspektren der vier Cy5- und vier Cy3-markierten cDNAs, welche jeweils aus einer biologischen Probe hervorgingen. Das Absorptionsmaximum bei 260 nm reflektiert die cDNA-Konzentration, wohingegen das Maximum bei 550 nm (für Cy3) bzw. bei 650 nm (für Cy5) auf die inkorporierte Farbstoffmenge schließen lässt. Deutlich ersichtlich sind die fast identischen Kurvenverläufe, welche auf eine hohe Reproduzierbarkeit hinsichtlich dieses Prozessschrittes hindeuten.

Die Spektren für die vier verschiedenen cDNAs jedes Farbstoffes sind fast deckungsgleich und weisen kaum Variabilitäten auf. Die errechneten Werte für die Farbstoff- und cDNA-Konzentrationen sowie die Inkorporationsfrequenzen (FOI) können der Tabelle 2 entnommen werden.

Angegeben sind die spektrophotometrisch ermittelten Werte der cDNA und Cyanin-Farbstoffkonzentration sowie die sich daraus ergebenden Inkorporationsfrequenzen. Die für die Berechnung der Inkorporationsfrequenzen verwendete Formel kann dem Anhang (D.1) entnommen werden. Im Fall des Cyanin5-Farbstoff wurden Konzentrationen zwischen 5,11 pmol/ μ l und 4,38 pmol/ μ l in die cDNA inkorporiert. Drei Farbstoffkonzentrationen liegen in einem sehr engen Bereich zwischen 4,98 pmol/ μ l und 5,11 pmol/ μ l, wohingegen der Wert von 4,38 pmol/ μ l etwas von diesem Bereich differiert. Die cDNAs, welche aus der gleichen biologischen Probe hervorgingen, sollten die gleichen DNA- und Farbstoffkonzentrationen sowie die gleiche Inkorporationsfrequenz aufweisen. Abweichungen davon deuten auf Variabilitäten im RT-Schritt hin bzw. bei signifikanten Abweichungen zwischen den RNA-Präparationen der gleichen biologischen Probe auf Variabilitäten, die diesem Schritt zugrunde liegen. Die Standardabweichung zwischen den cDNAs, welche aus einer RNA-Isolation gewonnen wurden, weist einen größeren Wert auf als jene, die auch die cDNAs der anderen RNA-Isolation mit einbezieht. Die cDNA-Konzentrationen schwanken zwischen 75,51 ng/ μ l und 85,90 ng/ μ l. Die resultierenden Inkorporationsfrequenzen sind jedoch für alle Proben einander recht ähnlich, sie variieren zwischen 16,51 und 17,14.

Für die Cy3-markierten cDNAs ergibt sich ein ähnliches Bild. Die Farbstoffkonzentrationen schwanken mit Werten zwischen 6,12 pmol/ μ l und 7,21 pmol/ μ l diesmal über einen etwas größeren Bereich, als dies für den Cy5-Farbstoff beobachtet wurde. Gleiches gilt für die cDNA-Konzentrationen, welche Werte zwischen 94,54 ng/ μ l und 83,80 ng/ μ l annehmen. Die Inkorporationsfrequenzen variieren zwischen 19,43 und 22,04 und sind damit vom

Tabelle 2: Generierte cDNA- und Farbstoffkonzentrationen sowie die sich daraus ergebenden Inkorporationsfrequenzen (FOI) für die verschiedenen RT-Ansätze des Parallelexperimentes.

	cDNA-Konz. (ng/ μ l)	Konz. Cy5 (pmol/ μ l)	Konz. Cy3 (pmol/ μ l)	FOI
cDNA 1.1.1	84,77	5,11	/	16,95
cDNA 1.1.2	85,90	5,09	/	16,83
cDNA 1.2.1	75,51	4,38	/	16,51
cDNA 1.2.2	82,06	4,98	/	17,14
cDNA 2.1.1	94,54	/	6,78	20,44
cDNA 2.1.2	83,80	/	6,12	20,62
cDNA 2.2.1	93,06	/	7,21	22,04
cDNA 2.2.2	90,80	/	6,24	19,43

Betrag und der Standardabweichung etwas größer als jene für die Cy5-dCTPs. Die Abweichung zwischen den zwei cDNAs, welche aus einer identischen RNA-Isolation hervorgingen, sind wiederum größer als zwischen den cDNAs unterschiedlicher RNA-Präparationen der gleichen biologischen Probe. Dies gibt einen Hinweis darauf, dass die Schwankungen eher in dem RT-Schritt begründet liegen und sich weniger aus den Variabilitäten zwischen den RNA-Präparationen ergeben. Die Inkorporationsfrequenzen sind hingegen fast identisch und zeigen, dass die Markierungsdichte der cDNA nur unwesentlich schwankt. Lediglich die Menge an generierter cDNA variiert in kleinerem Maße. Dieser Befund deckt sich weitgehend mit vielen Studien, welche belegen, dass der Schritt der reversen Transkription eine Quelle für höhere Variabilitäten ist, nicht jedoch die RNA-Isolation (Schadt *et al.* (2000)). Indes lagen die beobachteten Schwankungen weit unterhalb derer, die in anderen Studien ermittelt wurden und bis zu 90 Prozent betragen (Skena *et al.* (1995)). Leider lässt sich durch die photometrische Messung der cDNA- und Farbstoffkonzentration nur das Resultat der Reversen Transkription als Summe für alle *Targets* erfassen. Ob die getroffenen Aussagen auch für ein individuelles *Target* zulässig sind, lässt sich nur extrapolieren. Die RNA-Menge der einzelnen Spezies ist zu gering, um ein separates Umschreiben mit sequenzspezifischen Primern für ein einzelnes *Target* zu realisieren. Die daraus resultierende cDNA-Konzentration wäre zu gering für eine photometrische Analyse. Für die weitergehende Analyse wurden die so generierten cDNAs der beiden Proben paarweise gegeneinander auf sekretionsspezifische *low-density* Microarrays appliziert und über Nacht unter konvektiven Bedingungen im Schüttler hybridisiert. Die Rohdaten der Fluoreszenzintensität der einzelnen *Spots* sollten für alle Chips gleich sein. Ein Vergleich mit den Schwankungen des RT-Schrittes, welcher über die photometrische Messung der cDNA erfasst wurde, erlaubt Rückschlüsse über den Einfluß des Hybridisierungsschrittes. Die letztendlich entscheidende Größe für die Auswertung von DNA-Microarrays stellt das *logratio* dar, da es die Regulation des Genes widerspiegelt. Um den unterschiedlichen Inkorporationsfrequenzen und Fluoreszenzeigenschaften der beiden Cyanin-Farbstoffe Rechnung zu tragen, werden Microarray Experimente häufig im *Dye-Swap* Design durchgeführt d.h. jede Probe wird mit beiden Farbstoffen markiert und die unterschiedlich markierten cDNAs der verschiedenen Proben jeweils gegeneinander hybridisiert. Somit werden farbstoffspezifische Effekte ausgeglichen. Bei dem hier präsentierten

Experiment wurde eine andere Vorgehensweise gewählt, da nicht die explizite Regulation der Gene von Interesse war, sondern nur die Variabilitäten zwischen Chipexperimenten. Die ermittelten *logratios* spiegeln also nicht in jeden Fall die biologische Regulation wieder, da es sich nicht um farbstoffbereinigte Daten handelt. Im Rahmen einer Stichprobe wurden die *logratios* von fünf verschiedenen Genen auf den vier Microarrays miteinander verglichen. Graphisch wird dieser Vergleich in der Abbildung 15 gezeigt. Es wurden dabei Gene gewählt die *logratios* mit einem Betrag größer und kleiner 1 aufwiesen. Zusätzlich wurden Gene in die Auswahl integriert, welche ein *logratio* nahe 0 aufwiesen. Diese Daten wurden keinerlei Normalisierung unterzogen und spiegeln die Rohdaten des Experimentes nach der Primäranalyse wieder.

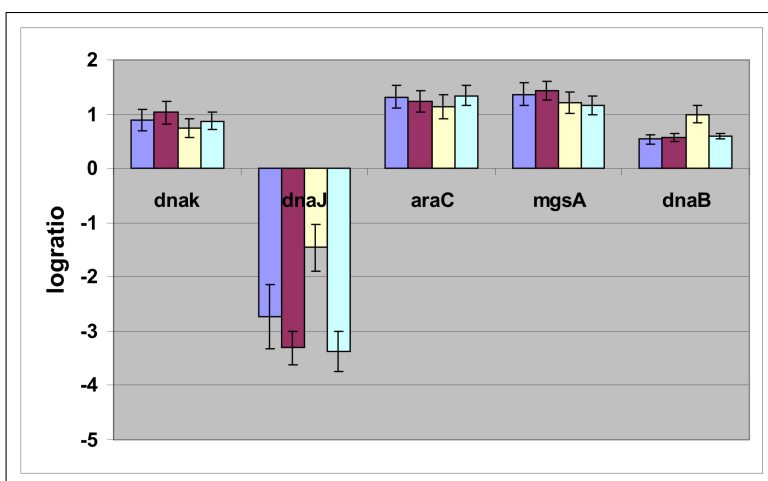


Abb. 15: Dargestellt sind die Ergebnisse aus der Primäranalyse der vier verschiedenen Microarrays des Parallelexperimentes. Das resultierende *logratio* von fünf repräsentativen Genen auf dem sekretionsspezifischen *low-density* Microarray ist für jeden Chip angegeben und nebeneinandergestellt.

Generell ist ersichtlich, dass das Vorzeichen und die Größenordnung der *logratios* für die jeweiligen Gene über alle Chips übereinstimmt. Für die Gene *dnaK*, *araC* und *mgsA* werden unwesentliche Schwankungen gefunden, wie eine Betrachtung der Standardabweichung der *logratios* ersichtlich macht. Für *dnaB* und *dnaJ* ergibt jeweils der gleiche Chip einen statistisch signifikanten Ausreißer. Der Betrag für das negative *logratio* des *dnaJ*-Genes ist deutlich kleiner als für die anderen Chips. Eine umgekehrte Aussage gilt für das positive *logratio* vom *dnaB* Gen, in diesem Fall wird mittels des Chips ein deutlich größeres *logratio* gefunden. Da die *logratios* für beide Gene in die gleiche Richtung gegenüber den anderen Chips abweichen, kann davon ausgegangen werden, dass bei diesem Chip ein Fluoreszenzkanal konsistent intensiver ist als bei den übrigen Chips. Allerdings tritt der beobachtete Effekt nicht bei den drei weiteren Genen der Betrachtung auf, weshalb hier eine regionale Erhöhung der Fluoreszenzintensität eines Kanales vorliegen könnte. Gestützt wird diese Vermutung dadurch, dass beide genspezifischen Oligonukleotide in räumlicher Nähe auf dem Chip lokalisiert sind.

Zusammengefasst zeigen die Ergebnisse, dass die Richtung der *logratios* richtig erfasst wurde. Bei einem der vier Chips existierten allerdings bei zwei der stichprobenartig untersuchten Gene signifikante Ausreißer. Relativierend muss allerdings bedacht werden, dass es sich um die Rohdaten des Chipexperimentes handelt und keinerlei statistische Normalisie-

rung erfolgte, welche zum Ausgleich derartiger Variabilitäten beitragen könnte. Das Experiment zeigt jedoch, dass statistisch signifikante Variabilitäten bei Microarray-Experimenten auftreten können, diese allerdings nicht so groß ausfallen, dass die Richtung einer potenziellen Regulation falsch angegeben wird oder eine differentielle Expression nicht detektiert wird. Für die meisten Gene stimmen generierten Daten der verschiedenen Chips gut überein. Dies zeigt, dass die Microarray-Experimente sehr gut geeignet sind, um die Regulation verschiedener Gene in einer hochparallelen Applikation sicher vorherzusagen. Der große Wert der Methodik liegt im *Screening*, weniger in der absoluten Genauigkeit. Allerdings führen Verbesserungen in der Instrumentarisierung, Standardisierungen in der experimentellen Durchführung und innovative Normalisierungsmethoden zu einer immer besseren Reproduzierbarkeit und Genauigkeit der chipbasierten Daten. Dieser Trend wird sich in Zukunft verstärken (Schäferling *et al.* (2006)). Dennoch sollte der genaue Wert der Regulation mit anderen Methoden verifiziert werden, da Schwankungen in einem gewissen Maße auftreten können. Aufgrund dieses initial durchgeführten Experimentes zur Evaluierung der Reproduzierbarkeit von Chipexperimenten kann die Arbeitshypothese der Eignung der Microarray-Technologie für die gewünschte Applikation der Genexpressionanalyse des Sekretionsvorganges in *E. coli* verifiziert werden.

4.2 Probenprozessierung und experimentelles Design

Für die evaluierten Transkriptomanalysen fungierten RNA-Isolationen, die aus Proben reproduzierbarer Bioreaktorkultivierungen mittels Phenol-Chloroform-Extraktion hervorgingen. 50 μg Gesamt-RNA wurden in cDNA konvertiert und diese durch Zusatz von Cy3- bzw. Cy5-dCTP fluoreszenzmarkiert. Jede Probe wurde dabei in 2 Ansätzen mit beiden Farbstoffen markiert, um farbstoffspezifischen Variabilitäten Rechnung zu tragen. Die farbstoffmarkierten cDNAs wurden im *Loop-Design* auf *E. coli whole-genome* Microarrays (MWG K12 V2 Array) appliziert und nach dem Standardprotokoll (siehe Anhang) behandelt. Das *Loop-Design* stellt das ökonomischste experimentelle Design dar, schematisch ist es in Abbildung 16 gezeigt.

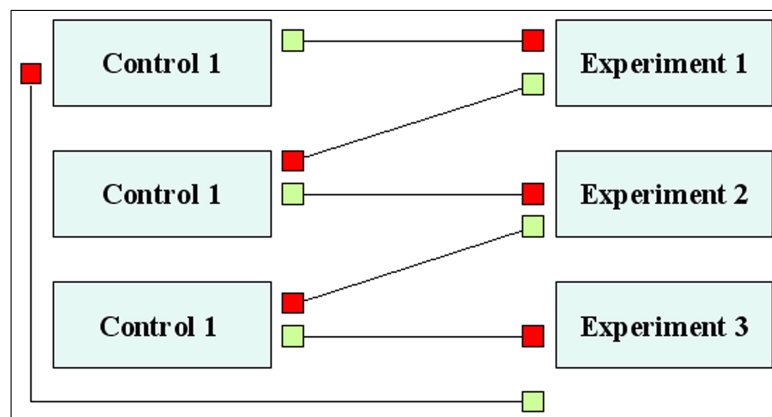


Abb. 16: *Loop-Design* mit *Common Reference*.

Nach dem Scannen der Microarrays erfolgte die Primärauswertung mit Hilfe des Programmes Gene Pix Pro 6.0TM. Geflaggte Daten wurden vor der weiteren Analyse eliminiert und von den hintergrundbereinigten Signalintensitäten der Logarithmus gebildet. Die Sekundäranalyse umfasste zudem eine *Between-Array*-Normalisierung des Datensatzes und

4.3 Übersicht über die verwendeten Sekretionstämme

Tabelle 3: Übersicht über die verwendeten Sekretionsstämme

Stamm	MG1655 pBAD LppBRPtorAmalE	Ara1655 pBAD LppBRPphoAbla	Ara1655 pBAD phoAbla	XL1 pBAD torAGFPSsrA
Reporter -protein	Maltose -bindeprotein	Alkalische Phosphatase β-Lactamase	Alkalische Phosphatase β-Lactamase	GFPSsrA
Ort	extrazellulär	extrazellulär	Periplasma	Periplasma
Induktor- konzentr. (mg/l)	1000 100	1000 25	1000	1000
Analyse	<i>whole-genome</i>	<i>whole-genome</i>	<i>whole-genome</i> <i>low-density</i>	<i>low-density</i>
Ursprung	Universität Bielefeld	Universität Bielefeld	Universität Bielefeld	eigene Klonierung

eine Beseitigung von Sättigungseffekten. Gene, dessen *logratio* einen Betrag größer als 1 aufwies, wurden als reguliert identifiziert. Aufgrund der fehlenden Genreplikate auf den kommerziellen Chips konnte keine t-Statistik angewandt werden. Sämtliche Daten sind beim NCBI Gene Expression Omnibus (GEO) hinterlegt und können dort eingehender studiert werden.

4.3 Übersicht über die verwendeten Sekretionstämme

Die durchgeführten transkriptionellen Sekretionsanalysen basierten auf der Verwendung von vier verschiedenen Sekretionsstämmen, welche in Tabelle 3 präsentiert werden. Um möglichst reproduzierbare Bedingungen zu generieren erfolgte die Kultivierung der verschiedenen Stämme im Bioreaktor (Biostat C, Sartorius Stedim GmbH). Alle Stämme tragen Plasmide, die sich von denen der pBAD-Serie ableiten (Guzman *et al.* (1995)) und beinhalten den arabinoseinduzierbaren Promoter pBAD. Der Arabinosepromoter hat sich als am besten geeignet für die gewünschte Applikation der Genexpressionsanalysen erwiesen, da die Transkription der stromabwärts befindlichen Gene sehr stark induziert werden kann (Faktor größer 1000) und über einen weiten Bereich linear steuerbar ist. Zudem ist die basale Transkription der promoterregulierten Gene sehr gering. In Tabelle 3 sind die verschiedenen analysierten Stämme aufgelistet. Zusätzlich wird die Information ihrer Reporterproteine, die Ziellokalisierung derselben, die untersuchten Induktionsniveaus sowie die Analysenform (*whole-genome* oder *low-density*) angegeben.

Aus den Daten der *whole-genome* Experimenten sowie Literaturangaben über essentielle Sekretionsgene wurde ein sekretionsspezifischer *low-density Array* entworfen. Dessen Applikation wird anhand der transkriptionellen Analyse der Stämme XL1 pBADtorAGFPSsrA sowie Ara 1655 pBADLppBRPphoAbla gezeigt. Um sekretionsinvolvierte Gene zu identifizieren, wurden das präinduktive Transkriptom mit dem verschiedener Zustände nach Induktion der rekombinanten Proteinexpression verglichen. Im Folgenden werden die einzelnen transkriptionellen Vergleiche präsentiert.

4.4 Analyse der extrazellulären, Tat-spezifischen Sekretion

Proben, welche aus zwei induzierten Bioreaktorkultivierungen des Stammes MG1655 pBADLppBRPtorAmalE hervorgingen, fungierten als Basis der im Folgenden beschriebenen Transkriptomanalysen. Der Stamm codiert für das Signalpeptid-modifizierte BRP von dem Plasmid CloDF13 (LppBRP) und dem Sekretionsreporter Maltosebindeprotein (MaE), welches mit der Signalsequenz der N-Oxidtrimethylaminreduktase (TorA) für den Tat-Pathway fusioniert wurde. Beide codierende Gene stehen unter der Kontrolle des Arabinosepromoters. Mittels der BRP-Aktivität wird über den Tat-Pathway ins Periplasma gelangtes MaE in das Kulturmedium sezerniert. Schematisch ist der Sekretionsvorgang für das Reporterprotein in Abbildung 17 gezeigt.

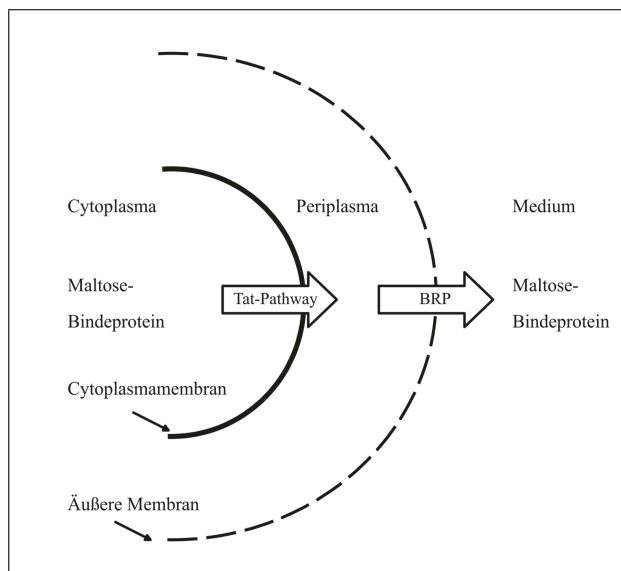


Abb. 17: Schema der Sekretion des Reporterproteins Maltosebindeprotein (MaE) des Stammes MG1655 pBADLppBRPtorAmalE. Das Reporterprotein wird mittels des Tat-Pathways in das Periplasma von *E. coli* translokiert. Durch die Aktivität des LppBRPs gelangt das periplasmatische Intermediat schließlich durch die äußere Membran in den extrazellulären Raum.

Es wurden zwei verschiedene Induktionsniveaus von 100 und 1000 mg/l *L*-Arabinose in Rahmen einer differentiellen Transkriptomanalyse evaluiert. Dem Bioreaktor wurden Proben direkt vor sowie 30 und 60 Minuten nach Induktion entnommen und daraus mittels Phenol-Chloroform-Extraktion RNA isoliert.

4.4.1 Induktion mit 100 mg/l *L*-Arabinose

Die Ergebnisse der transkriptionellen Analyse für ein Induktionsniveau von 100 mg/l werden im Folgenden präsentiert.

Übersicht

30 Minuten nach Induktion wurde im Vergleich zum Zustand vor der Induktion die differentielle Expression von 125 Genen beobachtet, 111 davon waren positiv und 14 negativ reguliert. Die Gesamtanzahl der regulierten Gene repräsentiert 2,8 Prozent der gesamten *E.*

4.4 Analyse der extrazellulären, Tat-spezifischen Sekretion

coli ORFs. In einem analogen Vergleich 60 Minuten nach Induktion wurden 102 Gene als reguliert identifiziert, wovon 86 positiv und 16 negativ reguliert waren. Diese repräsentieren 2,3 Prozent aller *E. coli* Gene. 60 Minuten nach Induktion waren im Vergleich zu dem vorgegangenen Zeitpunkt 30 Minuten nach Induktion nur noch 32 Gene reguliert (19 negativ und 13 positiv), welches einem prozentualen Wert von 0,72 Prozent des gesamten *E. coli* Genoms entspricht.

Analyse mit Hilfe der Multifun-Datenbank

Um den Datensatz einer biologischen Interpretation zugänglich zu machen wurde die Multifun-Datenbank für eine Funktionszuordnung der regulierten Gene verwendet. In dieser Datenbank sind 76 Prozent aller *E. coli* Gene mit Annotationen zu biologische Funktionen, Stoffwechsel- und Interaktionspfaden verfügbar. Bei 66 Prozent aller *E. coli* ORFs wurden diese Funktionen experimentell ermittelt. Damit stellt die Datenbank eine der genauesten und umfangreichsten Möglichkeiten zur biologischen Funktionsanalyse von Microarray-Daten dar (Karp *et al.* (2007)). Aufgrund der Komplexität der generierten Microarray-Datensätze wird so eine enorme Vereinfachung der Interpretation realisiert.

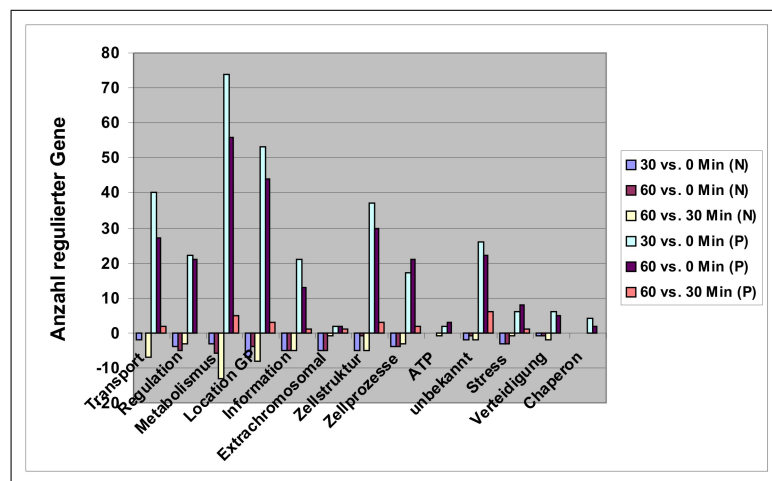


Abb. 18: Anzahl der regulierten Gene über die verschiedenen Multifun-Hauptklassen für die drei evaluierten Vergleiche. Die Auftragung erfolgte separat für die positiv (P) wie auch für die negativ (N) regulierten Gene. Zusätzlich sind die Unterklassen Stress, Chaperone und Verteidigung mit aufgeführt. Die Auftragungen erlauben keinen direkten Vergleich zwischen verschiedenen Klassen, da diesen unterschiedliche Anzahlen an Genen zugeordnet sind. Zusätzlich kann ein Gen mehreren Funktionsklassen zugehörig sein.

Die Multifun Datenbank ordnet die *E. coli* Gene primär in zehn verschiedene Hauptkategorien ein, die sich jeweils in weitere Unterkategorien aufteilen. Ein Gen kann dabei mehreren Kategorien zugehörig sein, so dass die Gesamtzahl der Gene in den verschiedenen Kategorie die Gesamtzahl der *E. coli* Gene übersteigt. Für die Auswertung der Chipdaten gilt damit analog, dass die Summe der Gene in den verschiedenen Multifun-Funktionsklassen die Gesamtzahl der regulierten Gene übersteigt. Zusätzlich kann ein Gen auch mehrmals einer Funktionsklasse zugeordnet sein, sofern es innerhalb dieser verschiedenen Unterkategorien zugehörig ist. Die Anzahl der den jeweiligen Funktionsklassen zugehörigen Genen kann der Abbildung 18 entnommen werden. Die meisten regulierten Gene sind in den Funktionklassen Metabolismus, Lokalisation der Genprodukte, unbekannte Funktion und Zellstruktur

4.4 Analyse der extrazellulären, Tat-spezifischen Sekretion

vorhanden. Dieser Beobachtung gilt jedoch anzumerken, dass diese Funktionsklassen unterschiedliche absolute Anzahlen an Genen enthalten und damit nicht direkt miteinander vergleichbar sind. So sind von den 4460 *E. coli* Genen 2328 Gene der Funktionsklasse Metabolismus zugeordnet, aber nur 518 der Funktionsklasse Zellstruktur. In diesem Kontext sind relative Werte, welche die Anzahl der Gene der jeweiligen Funktionsklasse mit in die Betrachtung einbeziehen, biologisch aussagekräftiger.

Für die positiv und negativ regulierten Gene der drei Vergleiche ist dieses in der Abbildung 19 realisiert. Es wurde für die jeweiligen Funktionsklassen der Quotient aus der Anzahl der regulierten Gene und der Gesamtzahl der Gene, die dieser Funktionsklasse zugeordnet sind gebildet und für die drei transkriptionellen Vergleiche graphisch aufgetragen. Für die hochregulierten Gene ist eindeutig ersichtlich, dass sich die Graphenverläufe der beiden Vergleiche nach Induktion gegen den Referenzzeitpunkt vor Induktion sehr ähnlich sind und sich substantiell von dem Vergleich der beiden postinduktiven Zustände unterscheiden. Die größten regulatorischen Effekte treten direkt nach der Induktion auf und bleiben bis zum Zeitpunkt 60 Minuten nach Induktion erhalten. Zwischen den Zeitpunkten 30 und 60 Minuten nach Induktion findet nur noch eine Regulation weniger Gene statt. Als Effekt der Arabinose-Supplementierung werden fast ausschließlich positiv regulierte Gene beobachtet. Die größten regulatorischen Effekte treten in den Funktionsklassen Zellprozesse, Adaption auf Stress sowie Verteidigung auf. In diesem Befund werden die drastischen Effekte der BRP-Expression reflektiert. Eine positive Regulation der Stress- und Verteidigungsantwort der Zelle scheint in diesem Kontext biologisch sinnvoll. Für den Vergleich 30 Minuten nach Induktion gegen den Zustand vor Induktion wird zusätzlich eine auffallend hohe Anzahl an Chaperonen reguliert vorgefunden. Chaperone assistieren neugebildeten und fehlgefalteten Proteinen, ein funktionelles Faltungsmuster einzunehmen. Ihre Expression wird Stress-induziert stark erhöht und gibt somit einen weiteren Hinweis auf die zelluläre Antwort hinsichtlich der BRP-Expression.

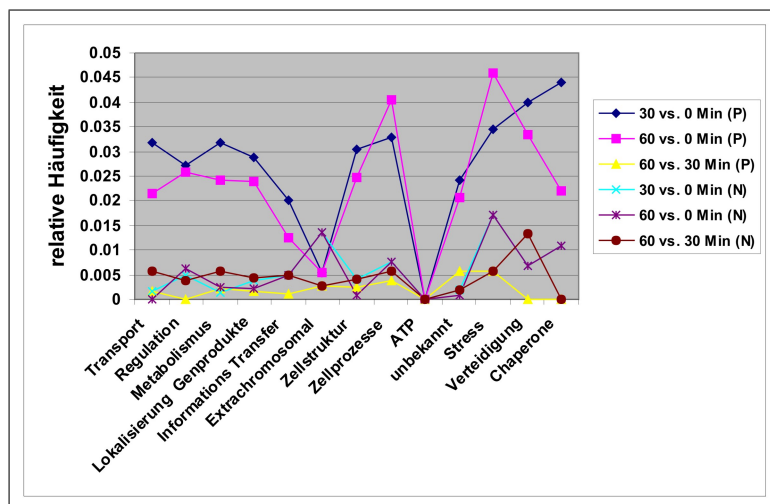


Abb. 19: Zuordnung der regulierten Genen der drei evaluierten Vergleiche zu den jeweiligen Multifun-Funktionsklassen. Aufgetragen sind die Quotienten aus der Anzahl der regulierten Gene und der Gesamtanzahl der einer Funktionsklasse zugehörigen Gene. Neben den Multifun-Hauptklassen wurden die Subklassen Stress, Verteidigung und Chaperone mit in den Vergleich integriert.

Für die negativ regulierten Gene werden derartig eindeutige Tendenzen nicht vorgefunden. Gleichwohl findet die stärkste Regulation ebenfalls in den Funktionsklassen Stress und Verteidigung statt und koinzidiert mit jenen Ergebnissen, die für die hochregulierten Gene gefunden wurden. Zusätzlich findet eine auffallend hohe Regulation bei den extrachromosomalen Genen für die Vergleiche vor/nach Induktion statt. Für eine detaillierte Diskussion von einzelnen Genen wird daher im Folgenden der Fokus auf diese Funktionsklassen gelegt.

Zur Evaluierung aufgrund welcher BRP-induzierter Stressarten eine besonders häufige Genregulation beobachtet wurde, erfolgte eine Aufschlüsselung der zellulären Stressantwort in die einzelnen Stressunterklassen. Ziel war dabei, die BRP-Aktivität einem bereits bekannten zellulären Stress zuzuordnen, um so schon bekannte Optimierungs- und Präventionsmaßnahmen auf die BRP-Reaktorkultivierungen zur Proteinsekretion zu übertragen. Die graphische Auftragung der relativen Werte ist in Abbildung 20 realisiert.

Wie aus der Abbildung 20 ersichtlich, besteht die Stressantwort der Zelle aus verschiedenen Stressarten und kann nicht einem bestimmten Stressstimulus zugeordnet werden. Gene, welche in der Antwort auf thermischen, oxidativen, Nahrungsmangel-induzierten und besonders Austrocknungsstress involviert sind, weisen die größten relativen Häufigkeiten auf. Jene Gene, die dem Austrocknungsstress zugeordnet sind, zeigen ausnahmslos ihre Funktion in der Synthese eines Kapselpolysaccharides, welches ein wichtiges Oberflächenadhäsion bei der bakteriellen Biofilmbildung darstellt. Durch die Biofilmbildung schützt sich das Bakterium vor schädlichen Umwelteinflüssen. Eine Biofilmbildung könnte auch BRP-induziert auftreten. Die Regulation der in der Biofilmbildung involvierten Gene tritt erst 60 Minuten nach Induktion auf. Dieses korreliert sehr gut mit der zunehmenden BRP-Akkumulierung (Sommer *et al.* (2008)). *E. coli* lagert zudem unter Bedingungen hoher Osmolaritäten signifikante Mengen an niedermolekularen Osmolyten, wie z.B. Zuckern, ein (Luirink *et al.* (1991)). Anhand dieses Faktums könnte die Regulation der diskutierten Gene ebenfalls erklärt werden.

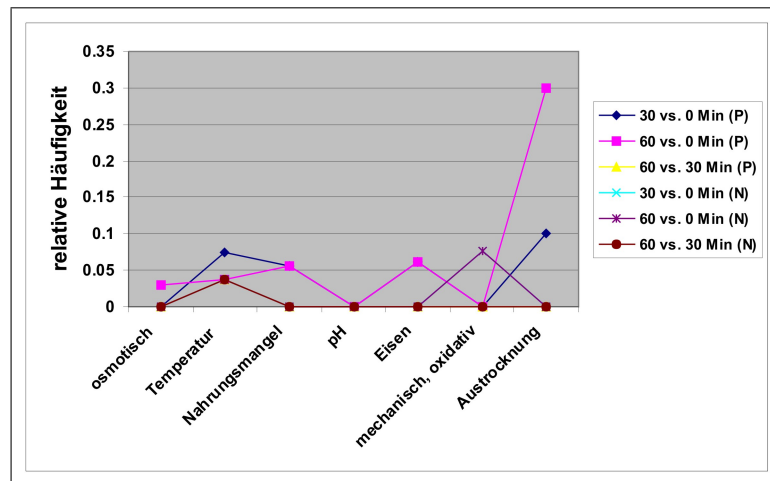


Abb. 20: Betrachtung der Multifun-Klasse Adaption auf Stress für die drei evaluierten Vergleiche, separat für positiv (P) und negativ (N) regulierte Gene durchgeführt. Die differentiell exprimierten Gene, welche dieser Funktionklasse zugehörig waren, wurden hinsichtlich ihrer Verteilung auf die einzelnen Subklassen evaluiert. Für eine bessere Vergleichbarkeit der einzelnen Subklassen untereinander wurden relative Werte, gebildet, welche die unterschiedlichen Genanzahlen in jeder Subklasse mit einbezogen.

Betrachtet man die Lokalisierung der Genprodukte für die regulierten Gene der drei eva-

lierten Analysen in Abbildung 21, so ist auffällig, dass sich für die positiv regulierten Gene der Vergleiche vor/nach Induktion ein hoher prozentualer Anteil der Proteinprodukte in der inneren und äußeren Membran sowie im Periplasma befindet. Dies sind die zellulären Regionen, welche von einer BRP-Expression besonders betroffen sind (van der Waal *et al.* (1995)). Aufgrund des Einflusses der BRP-Expression auf die Membranintegritäten und die semispezifische Freisetzung des periplasmatischen Inhalts wird eine substantielle Zellreaktion in diesen Bereichen erwartet und auch experimentell beobachtet. Relativierend muss allerdings konstatiert werden, dass auch die Arabinosesupplementierung und die damit einhergehende positive Regulation von Transportern, periplasmatischen Bindungsproteinen sowie Porinen zur Substrataufnahme über die äußere Membran einen signifikanten regulatorischen Einfluß auf Gene ausübt, deren Produkte in dieser Region lokalisiert sind. Es ist also in konkreten Fällen differenziert zu betrachten, ob es sich um einen BRP- oder Arabinose-induzierten Regulationseffekt handelt.

Eine genauere Betrachtung der Zellprozesse (Abbildung 22) zeigt, dass die meisten regulierten Gene in die Biofilmbildung, die Virulenz und die Stress- und Verteidigungsantwort involviert sind. Wiederum ereignet sich die deutliche Mehrheit der regulatorischen Effekte bei den positiv regulierten Genen der Vergleiche vor/nach Induktion. Die signifikante Regulation von Genen, die in der Biofilmbildung eine zentrale Rolle spielen, tritt nur für den Vergleich 60 gegen 0 Minuten auf und koinzidiert mit den zuvor gemachten Beobachtungen.

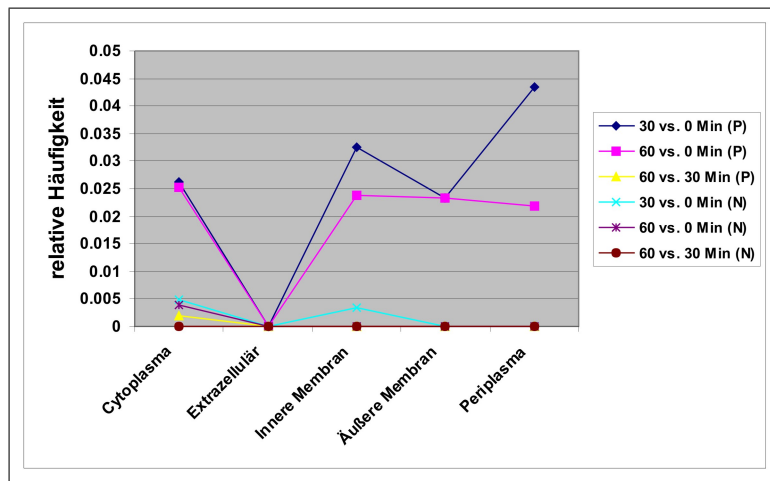


Abb. 21: Betrachtung der Multifun-Klasse Lokalisierung der Genprodukte für die drei evaluierten Vergleiche, separat für positiv (P) und negativ (N) regulierte Gene durchgeführt. Die differentiell exprimierten Gene, welche dieser Funktionsklasse zugehörig waren, wurden hinsichtlich ihrer Verteilung auf die einzelnen Subklassen evaluiert. Für eine bessere Vergleichbarkeit der einzelnen Subklassen untereinander wurden relative Werte gebildet, welche die unterschiedlichen Genanzahlen in jeder Subklasse mit einbezogen.

Schlüsselt man die regulierten Gene in der Multifun-Funktionsklasse Struktur nach ihrer Zugehörigkeit zu den einzelnen Unterklassen auf und bildet zur besseren Vergleichbarkeit erneut relative Werte, so erhält man die in Abbildung 23 gezeigte graphische Darstellung. Die größten regulatorischen Effekte werden für die positiv differentiell exprimierten Gene der drei evaluierten Vergleiche beobachtet. Es wird für den Vergleich 60 vs. 0 Min ein deutliches Maximum bei der Subklasse Kapselbildung beobachtet. Weiterhin signifikant hohe Werte sind für die beiden Vergleiche vor/nach Induktion bei den Unterklassen Oberflächen-

antigene, Flagellum und Peptidoglykan zu verzeichnen. Für die hochregulierten Gene des Vergleichs der beiden postinduktiven Zustände werden signifikante regulatorische Effekte bei den Subklassen Kapselbildung und Pilus identifiziert. Besonders interessant sind die beobachteten Effekte im Bereich Kapselbildung, Oberflächenantigene und Peptidoglykan, da die signifikante Auswirkung der BRP-Expression auf diese Zellkompartimente in der Literatur belegt ist (Van der Waal *et al.* (1995)).

Im Folgenden muss eine detaillierte Betrachtung der einzelnen Gene klären, ob deren Regulation ursächlich aufgrund der BRP-Aktivität auftritt. Aufgrund der breitgefächerten zellulären Stressantwort, der Überlappung Arabinose- und BRP-induzierter Effekte und der Existenz weiterer interessanter regulierter Gene und Operons wird im Folgenden eine detaillierte Diskussion der Genregulation realisiert. Wie eingangs erwähnt, findet eine besondere Fokussierung auf Gene statt, welche die komplexen mechanistischen Vorgänge und Reaktionen auf eine BRP-Expression und die damit einhergehende Proteinsekretion ermöglichen.

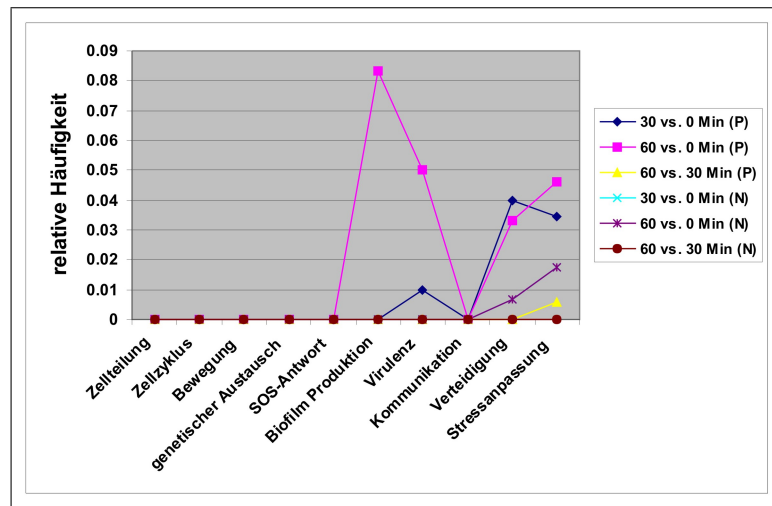


Abb. 22: Betrachtung der Multifun-Klasse Zellprozesse für die drei evaluierten Vergleiche, separat für positiv (P) und negativ (N) regulierte Gene durchgeführt. Die differentiell exprimierte Gene, welche dieser Funktionklasse zugehörig waren, wurden hinsichtlich ihrer Verteilung auf die einzelnen Subklassen evaluiert. Für eine bessere Vergleichbarkeit der einzelnen Subklassen untereinander wurden relative Werte gebildet, welche die unterschiedlichen Genanzahlen jeder Subklasse mit einbezogen.

Detaillierte Analyse

Generell wird eine auffallend hohe Anzahl an Stress- und Abwehr-involvierten Genen vorgefunden, sowie solchen Genen, die in der Synthese des Peptidoglykans bzw. des Lipopolysaccharides der äußeren Membran entscheidende Funktionen inne haben.

Vergleich 30 vs. 0 Min

30 Minuten nach Induktion wird das araBAD-Operon stark positiv reguliert. Die Genprodukte dieses Operons codieren für die Enzyme, welche die initialen Schritte der metabolischen Degradation der Arabinose katalysieren. Somit stellt der Befund eine Positivkontrolle der Chipeergebnisse dar.

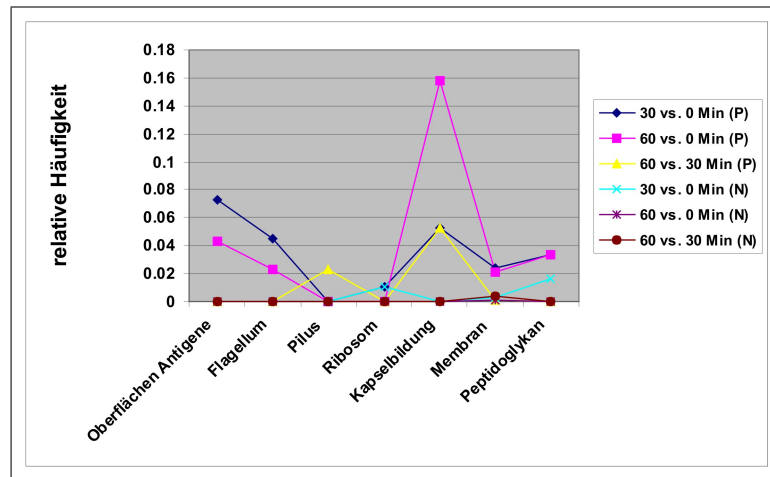


Abb. 23: Betrachtung der Multifun-Klasse Zellstruktur für die drei evaluierten Vergleiche, separat für positiv (P) und negativ (N) regulierte Gene durchgeführt. Die differentiell exprimierten Gene, welche dieser Funktionklasse zugehörig waren, wurden hinsichtlich ihrer Verteilung auf die einzelnen Subklassen evaluiert. Für eine bessere Vergleichbarkeit der einzelnen Subklassen untereinander wurden relative Werte gebildet, welche die unterschiedlichen Genanzahlen in jeder Subklasse mit einbezogen.

Zusätzlich positiv reguliert sind die Gene *araFGH* sowie das Strukturgen *araE*. Die Regulation dieser Gene kann ebenfalls als Positivkontrolle fungieren. Durch die Supplementierung der *L*-Arabinose als Induktor wird dieser mittels zweier verschiedener Aufnahmesysteme in die Zelle transportiert. *E. coli* besitzt zu diesem Zweck ein Aufnahmesystem mit niedriger Substrataffinität, dessen Triebkraft der pH-Gradient über der Membran ist, sowie ein hochaffines, ATP-abhängiges Transportsystem vom Typus der ABC-bindenden Kasette. Der Reporter von ersterem System wird von dem Gen *araE* codiert, wohingegen die Gene *araFGH* für Letzteres codieren. Das Genprodukt von *araF* fungiert dabei als Arabinose-Bindungsprotein, jenes von *araH* als der membrangebundene Teil und schließlich *araG* als ATP-bindende Komponente des ABC-Transporters. Unter physiologischen Bedingungen, bei denen keine *L*-Arabinose im Medium vorliegt, werden diese spezifischen Transporter nicht benötigt und die Genexpression der codierenden Gene entsprechend reprimiert. Bei der Zugabe von *L*-Arabinose findet daher eine starke positive Regulation statt, wie sie auch im Experiment vorgefunden wird. Die Gene *araFGH* wurden nach Induktion besonders stark reguliert vorgefunden, die Arabinoseaufnahme wird daher vornehmlich über die hochaffine ABC-bindende Kasette realisiert. Neben den Genen, welche in die Arabinoseaufnahme und den Arabinosemetabolismus involviert sind, wurden zudem eine auffallend große Anzahl an regulierten Genen identifiziert, welche in der Biosynthese des Lipopolysaccharids der äußeren Membran, der Eisenaufnahme, der Biofilmproduktion und der Adaption an osmotisch und thermisch bedingten Stress fungieren. Eine generelle Auswirkung von zellulären Stress stellt zudem die beobachtete Regulation von Chaperonen dar (Poranen *et al.* (2006); Harcum *et al.* (2006)). Der augenscheinliche Zusammenhang mit der BRP-Expression, welche massive Auswirkungen auf die Membranintegritäten, besonders der äußeren Membran hat, wurde zum Anlaß genommen, diese Gene in besonderer Detailtiefe zu erläutern. Über die verschiedenen evaluierten Microarray-Experimente mit gleichen oder ähnlichen Sekretionsreportersystemen existierten Gene, welche in nahezu allen Vergleichen reguliert vorgefunden

wurden. Diesen Genen galt zusätzlich eine besondere Aufmerksamkeit, zumal ihre Existenz einen Hinweis auf reproduzierbare und biologisch sinnvolle Analysenergebnisse darstellt.

Das Gen *bglX* codiert für eine periplasmatisch lokalisierte β -D-Glukosidase und wurde in nahezu allen Microarrayexperimenten mit dem Sekretionsreportersystem nach Induktion als reguliert vorgefunden. Wie bei dem ebenfalls regulierten *ygcY* handelt es sich dabei um ein osmotisch induzierbares Gen. Eine ganze Reihe von Genen, welche in die Biosynthese des Lipopolysaccharides (LPS) der äußeren Membran und dem Peptidoglykan involviert sind, wird 30 Minuten nach Induktion positiv reguliert. Es handelt sich dabei um die Gene *rfaCD*, *rfbB*, *mltC*, *tolC*, *ybbF*, *galKF* und *B1640*. Das LPS stellt die Hauptkomponente der äußeren Membran gram-negativer Bakterien dar. Seinerseits ist das LPS in die beiden Hauptbestandteile Lipid A und dem phosphorylierten Kernoligosaccharid gegliedert (Brock: Mikrobiologie (2003)). Die Proteinprodukte von *rfaC* und *rfaD*, welche chromosomal in einem Operon organisiert vorliegen, sind direkt an der Biosynthese von Lipid A beteiligt. Gleiches gilt für die Strukturgene *ybbF* und *rffB*. Die BRP-Expression führt zu einer Aktivierung der Phospholipase A durch Dimerisierung, welches eine partielle Degradation der äußeren Membran zur Folge hat (Dekker *et al.* (1999)). In diesem Kontext wird verständlich, dass eine Nachproduktion von Komponenten der äußeren Membran erfolgt, um deren Verlust zu kompensieren und die Membranintegrität zu gewährleisten. Aus sekretorischer Sicht interessant ist die Regulation von *tolC*, welches für ein Porin der äußeren Membran codiert und für die Aufnahme und Freisetzung verschiedener hydrophober und amphiphatischer Moleküle verantwortlich ist. *TolC* bildet in der äußeren Membran ein Trimer und stellt ein Substrat für den Sec-Pathway dar (Mergulaho *et al.* (2005)). Auch eine mögliche Beteiligung an der BRP-Freisetzung kann vermutet werden. In dieser Hinsicht ist auch die Regulation von *mltC* äußerst interessant. Das *mltC*-Gen codiert für eine lytische Endotransglykosylase und stellt ein membrangebundenes Lipoprotein dar. Diese Murein-Hydrolase vermag eine lokalisierte Öffnung der netzartigen Peptidoglykanstruktur zu realisieren und so den Export sterisch anspruchsvoller Komponenten wie DNA, Toxinen oder fimbrischer Proteinen zu ermöglichen. Es ist davon auszugehen, dass es sich hierbei um einen BRP-induzierten Effekt handelt, um die Sekretion periplasmatische Proteine ins Kulturmedium zu realisieren. Die codierenden Gene für *mltC* und dessen Analogon *mltE* finden sich zudem in vielen evaluierten transkriptionellen Vergleichen vor/nach Induktion unter den positiv regulierten Genen wieder. Eine weitere Auffälligkeit ist die positive Regulation vieler stressinduzierbarer Gene, welche in der Eisenaufnahme der Zelle eine Rolle spielen. In dem konkret betrachteten Vergleich sind dieses die Gene *yecI*, *uvry*, *yhea* und *yebI*. Eisenionen spielen eine wichtige Rolle bei der Zellatmung, da sie die Hauptbestandteile der Cytochrome und Schwefel-Eisen-Proteine sind. Die Lokalisierung der Schwefel-Eisenproteine in der inneren Membran und des Eisenaufnahmesystems in der äußeren Membran deutet auf eine mögliche Beeinflussung durch die BRP-Expression hin. Besonders das Gen *yecI* wird nach der Induktion in nahezu allen durchgeführten Analysen als reguliert identifiziert. Diese Genklasse stellt die zelluläre Antwort auf oxidativen Stress dar. Eine zu starke Induktion der BRP-Expression kann zu einer Zellyse führen und beeinträchtigt auch die Integrität der inneren Cytoplasmamembran (Van der Waal *et al.* (1995)). Da im Periplasma gram-negativer Bakterien oxidierende Bedingungen vorherrschen, könnte eine BRP-induzierte Schädigung der Cytoplasmamembran Gründe für die oxidativen Stressantwort liefern. Weitere stressinduzierbare Gene und solche, die in den Abwehrmechanismus der Zelle involviert sind, wurden durch den Induktionsstimulus reguliert. Prominente Beispiele, die auch in weiteren Analysen in Koinzidenz reguliert auftraten, sind *cspa*, *marB*, *evgA*, *ksgA*, *deaD*, *spoT* und *ygdp*. Das Gen *cspA* wurde nach der Induk-

tion besonders stark positiv reguliert. Es codiert für einen Transkriptionsaktivator, der ein RNA-Chaperon darstellt und dessen Expression bei einem Kälteschock induziert wird. Das Protein CsgA verhindert bei tiefen Temperaturen die vermehrte Sekundärstrukturbildung der RNA. Weitere Transkriptionsaktivatoren werden von den positiv regulierten Genen *dead* und *evgA* codiert. Wiederum reflektiert ihre konsistente Regulation über viele verschiedene Vergleiche hinweg eine generelle Zellantwort. Interessant ist in diesem Zusammenhang auch die Regulation des Genes *mdaB*, welches für eine Untereinheit der NADH-Ubiquinon Reduktase codiert. Neben der ATP-Hydrolyse stellt die protonenmotorische Kraft (PMK) die Triebkraft der Sec-abhängigen Translokation dar. Für den Tat-Export ist diese die alleinige Triebfeder. Die PMK wird von dem pH-Gradienten über der inneren Membran verursacht. Zur Aufrechterhaltung dieses Gradienten spielt das von *mdaB* codierte Protein eine wichtige Rolle (Stolpe *et al.* (2004)). Die vermehrte Expression von Chaperonen stellt einen weiteren Indikator für eine Stressantwort der Zelle dar und wurde auch bei der präsentierten Analyse beobachtet. Positiv reguliert wurden die Strukturgene von den Chaperonen *FliT*, *FkpA* und *HybB*. Besonders die Regulation des Hitzeschockgens *fkpA* deutet auf ein biologisch sinnvolles Ergebnis hin, da verschiedene Studien die Analogien zwischen BRP-induzierten Effekten und denen eines Hitzeshocks aufzeigen (Van der Waal *et al.* (1995)). Gemäß dem eingangs Konstatierten finden die größten regulatorischen Effekte bereits relativ kurz nach Induktion statt. Entsprechend sind die wenigen hochregulierten Gene beim Vergleich der beiden postinduktiven Zustände aus sekretorischer Sicht nicht interessant. Einzig und allein die positive Regulation des Genes *osmB* ist erwähnenswert, zumal sich dieses Gen über fast alle transkriptionellen Vergleiche mit unterschiedlichen Sekretionssystemen hinweg als positiv reguliert erweist. Die osmotisch induzierbare Expression und osmoprotektive Wirkung gibt klare Hinweise auf eine Auswirkung hinsichtlich der BRP-Expression. Die positive Auswirkung der *OsmB*-Überexpression auf die BRP-vermittelte Proteinsekretion könnte so mittels einer Coexpressionstrategie evaluiert werden.

Vergleich 60 vs. 0 Min

Analog dem unlängst Gesagten werden die meisten regulierten Gene, wie sie für den Vergleich 30 Minuten gegen den Zustand vor Induktion identifiziert wurden, auch nach 60 Minuten noch positiv reguliert vorgefunden. Die hohe Wiederfindungsrate der regulierten Gene in diesem Vergleich gibt Hinweise auf sinnvolle Chipdaten, da eine zufällige Regulation exakt derselben Gene statistisch unwahrscheinlich ist. Die Positivkontrollen *araBAD* und *araHGF* wurden ebenso wie die Gene *yecI*, *yedA*, *bglX*, *lysP*, *evgA*, *ygfQ*, *tolC*, *dead*, *yheA*, *fkpA*, *cspA*, *mltE* und *spoT* wiedergefunden. Die Wiederfindungsrate zwischen den beiden Vergleichen wurde nur für die aus sekretorischer Sicht interessanter Gene evaluiert. Zusätzlich positiv reguliert waren die Gene des Galaktoseoperons *galEKT*, deren Genprodukte bei der Synthese des Kapselpolysaccharids *Colanic Acid* eine wichtige Rolle spielen. Die Synthese einer Kapsel wird bei der bakteriellen Biofilmbildung als Oberflächenadhäsion beobachtet. Mit der gleichen Begründung kann in diesem Zusammenhang die Regulation des Genes *yedQ* verstanden werden, welches ebenfalls essentiell für die Ausbildung bakterieller Biofilme ist. Die Bildung von Biofilmen, d.h. Kolonien von schleimhüllten Zellen, welche mittels Polysaccharid-Adhäsinen an eine Oberfläche assoziiert sind, stellt einen Abwehrmechanismus des Bakteriums gegen schädliche Umwelteinflüsse dar. Es ergeben sich erhebliche wirtschaftliche und medizinische Implikationen aus der bakteriellen Biofilmbildung, da diese im Biofilm für das humane Immunsystem und Bakteriozide unerreichbar sind (Ren *et al.*

2004, Wang *et al.* (2004)). Durch die drastischen Auswirkungen der BRP-Expression wird dieser Abwehrmechanismus verständlich. Weiterhin positiv differentiell exprimiert ist das Gen *osmB*, was auf Bedingungen mit einer erhöhten Osmolarität hindeutet.

Vergleich 60 vs. 30 Min

Die Zustände 30 und 60 Minuten relativ zum Induktionszeitpunkt ähneln einander fast vollständig, und sinngemäss tritt beim Vergleich zwischen Ihnen keine nennenswerte differentielle Expression sekretorisch interessanter Gene auf.

negativ regulierte Gene

Eine detaillierte Betrachtung der negativ regulierten Gene für die drei evaluierten Vergleiche zeigt im Wesentlichen die Regulation von in den Zellmetabolismus involvierten Genen. Die beiden Vergleiche 30 vs. 0 Min und 60 vs. 0 Min gleichen sich fast völlig. Lediglich für den Vergleich der beiden postinduktiven Zustände ist noch die Regulation des Genes *glpB* erwähnenswert, da dieses in den Glycerinmetabolismus involviert ist (Lin *et al.* (1976)). Durch die Umstellung des Zellmetabolismus auf die energetisch vorteilhaftere Kohlenstoffquelle Arabinose wird verständlich, warum die negative Regulation dieses Gens auftritt.

4.4.2 Induktion mit 1000 mg/l *L*-Arabinose

Für die Fermentation des Stammes MG 1655 pBADtorAmalE LppBRP mit einer Induktionsstärke von 1000 mg/l *L*-Arabinose wurde analog zur vorangegangenen Diskussion ein Vergleich dreier äquivalenter Zeitpunkte evaluiert. Der transkriptionelle Vergleich der Probe, welche 30 Minuten nach Induktion entnommen wurde, gegen jene, die den präinduktiven Zustand repräsentiert, identifizierte 164 regulierte Gene, von denen 78 negativ und 88 positiv reguliert waren. Dies entspricht einem prozentualen Wert von 3,68 Prozent aller *E. coli* Gene. Nach 60 Minuten waren gegenüber der Referenz vor Induktion ebenfalls 164 Gene differentiell exprimiert, wovon allerdings 128 positiv und 38 negativ reguliert wurden. Relativ zur Gesamtzahl aller existenten *E. coli* ORFs entspricht dieses ebenfalls 3,68 Prozent. Schließlich wurden bei dem Vergleich der beiden postinduktiven Zustände 175 regulierte Gene beobachtet, diese setzten sich aus 39 negativ und 136 positiv regulierten Genen zusammen, welche 3,92 Prozent des *E. coli* Genoms repräsentieren.

Mittels der Multifun-Datenbank wurden den regulierten Genen biologische Funktionen zugeordnet. Dabei wurde analog zu der vorherigen transkriptionellen Analyse mit dem gleichen Sekretionsreportersystem verfahren. Es ergaben sich äquivalente Ergebnisse zur vorherigen Transkriptomanalyse, weshalb auf eine gesonderte Darstellung derselben verzichtet wurde. Im Folgenden wird mit der detaillierten Diskussion einzelner regulierter Gene und Operons für die drei evaluierten transkriptionellen Vergleiche fortgefahren.

Vergleich 30 vs. 0 Min

Die positive Regulation der in den Arabinosemetabolismus und der Arabinoseaufnahme involvierten Operons *araBAD* und *araFGH* verifizierte wiederum die biologische Sinnhaftigkeit der generierten Analysenergebnisse. Die Regulation der Gene *lolA* und *lolB*, welche chromosomal in einem Operon organisiert vorliegen, identifizierte den Einfluß der BRP-Expression auf die Konstitution der äußeren Membran. Die beiden Gene codieren für ein periplasmatische Chaperon (*LolA*) und für ein in der äußeren Membran befindliches Lipoprotein (*LolB*),

welche essentiell für die korrekte Lokalisierung und den Transport der Hauptkomponente Lpp der äußeren Membran sind (Matsuyama *et al.* (1995); Yokota *et al.* (1999)). Die experimentell beobachtete Überexpression des lolA-Genes aktiviert das RCS-System, welches die Produktion einer extrazellulären Polysaccharid-Kapsel aus *Colanic Acid* (M-Antigen) bewirkt. Bei der transkriptionellen Analyse einer schwächer induzierten Kultur des gleichen Stammes (Abschnitt 4.4.1) wurde dessen Regulation ebenfalls beobachtet (Wang *et al.* (2004)). Die Auswirkungen der Aktivierung des RCS-Systems lassen sich auch experimentell anhand der Chipdaten beobachten. Der Gencluster galKET codiert für die Enzyme, die eine Schlüsselposition in der Biosynthese der in diesem Zusammenhang erwähnten Polysaccharidkapsel aus *Colanic Acid* einnehmen, und wird in Koinkidenz mit lolA stark reguliert vorgefunden. Die positive Regulation dieser Gene findet sich in vielen Vergleichen wieder und stellt eine mögliche zelluläre Antwort auf die BRP-Expression dar. Eine weitere interessante Beobachtung, welche mit der Regulationskaskade einhergeht, ist die differentielle Expression der Genkomponenten der Tryptophanbiosynthese. LolB besitzt einige hochkonservierte Tryptophan-Domänen, welche für die Substratspezifität entscheidend sind, wie durch *site-directed-mutagenesis*-Analysen belegt wurde (Yokota *et al.* (1999)). Durch die vermehrte Produktion von lolB und den damit einhergehenden erhöhten Tryptophanbedarf könnte die positive Regulation erklärt werden. Weiterhin positiv reguliert sind die osmotisch induzierbaren Gene *otsB*, *osmB* und *osmY*, welches wiederum einen Hinweis auf einen BRP-induzierten osmotischen Stress für die Zelle gibt. Bei dem Osmoprotektanten *OsmB* handelt es sich um ein Lipoprotein, welches zur Kompensation des Verlustes der Membranintegrität der äußeren Membran vermehrt produziert werden könnte. In Koinkidenz mit dieser Beobachtung findet die positive Regulation des Genes *spy* statt. Bei dem Genprodukt handelt es sich um ein Polypeptid, welches vom Cpx-System für Membranstress reguliert wird und bei der Bildung von Sphäroblasten eine zentrale Rolle spielt (Jung *et al.* (1990)). Die semispezifische Freisetzung des periplasmatischen Inhaltes, bedingt durch den BRP-induzierten partiellen Abbau der äußeren Membran, gleicht in vielen Punkten der Sphäroblastenbildung und ruft entsprechend bei der Zelle eine analoge Antwort hervor. Evident zu dem bisher Konstatierten ist auch die beobachtete Regulation des Genes welches für den Sigmafaktor RpoS codiert und dessen Expression unter dem Stimulus von osmotischem Stress und der Biofilmbildung induziert wird (Ren *et al.* (2004)). Die drastischen Folgen der BRP-Expression manifestieren sich ebenso in der Regulation der Gene *nlpD*, *exbB* und des Operons *rfaCF*, welche eine zentrale Rolle in der LPS-Biosynthese spielen. Eine Überexpression von *rfaC* führt zur Biofilmbildung (Bailey *et al.* (1996)), welches experimentell auch durch die Regulation des Gencluster galKET und weiterer involvierter Gene beobachtet wird und gut mit der positiven Regulation von lolA korreliert. Bei *NlpD* handelt es sich um ein äußeres Membranlipoprotein, welches eine hohe Sequenzhomologie zu dem Lpp-Protein aus *Haemophilus somnus* aufweist (Wu *et al.* (2009)). Dieses könnte ebenfalls dazu fungieren, dem BRP-induzierten proteolytischen Abbau der äußeren Membran mittels Phospholipase A entgegenzuwirken. Des weiteren ist *NlpD* für die Aufnahme der Gruppe B-Colicine von essentieller Bedeutung. Zusätzlich wurde die Regulation einer Reihe von Prophagengene beobachtet, deren Proteinprodukte bisher unbekannte Funktionalitäten aufweisen. Die Regulation der Gene B2350 und B2351 trat auch in weiteren transkriptionellen Vergleichen auf und sei hier als Beispiel genannt.

Vergleich 60 vs. 0 Min

Die Ergebnisse für den Vergleich 60 Minuten gegen den Zeitpunkt vor Induktion decken sich in weiten Teilen mit jenen, die vorangehend diskutiert wurden. Die positive Regulation der Operons *araBAD* und *araEFG* als Positivkontrolle wird ebenfalls in diesem Vergleich identifiziert. Gleiches gilt für die substantielle Zellantwort auf die BRP-Expression durch Regulation verschiedener Gene, die in die LPS-Biosynthese oder die Ausbildung von Biofilmen involviert sind. Die Gene bzw. Operons *galKTE*, *lolAB*, *trpCED*, *osmB*, *rfaC* und *spot* werden ebenfalls nach Induktion positiv reguliert. Erwähnenswert ist die zusätzliche differentielle Expression vom Gen *ycgG*, dessen Produkt ebenfalls eine substantielle Bedeutung bei der Ausbildung von Biofilmen spielt und gut mit den zuvor getroffenen Aussagen koinzidiert. Weiterhin sind die Gene *vacJ* und *evgA*, die ihre Bedeutung in der allgemeinen Stressantwort der Zelle haben, positiv reguliert. Aus sekretorischer Sicht ist die Regulation des Genes *secG* bedeutsam, welches für einen Bestandteil des zentralen Translokationstunnel der Sec-Translokase codiert (Economou *et al.* (1999)). Durch die Saturierung der Sekretionskapazitäten des Sec-Pfades aufgrund des Sec-Substrates BRP wird die Regulation dieses Genes verständlich.

Vergleich 60 vs. 30 Min

Im Gegensatz zu den Beobachtungen, welche mit dem gleichen Sekretionsrepotersystem bei einem niedrigeren Induktionsniveau gemacht wurden, findet zwischen den Zuständen 30 und 60 Minuten relativ zum Zeitpunkt der Induktion noch eine signifikante positive Genregulation statt. Besonders interessant ist in diesem Kontext die differentielle Expression des Phagenschockoperons *pspABCDE* und die hohe Anzahl an regulierten Chaperon-Genen wie *grpE*, *stpA*, *hybB*, *hslU* sowie *hslV*. Dies deckt sich gut mit den experimentellen Befunden, die eine maximale BRP-Aktivität nach dieser Zeitspanne nach Induktion detektieren. Durch die Akkumulierung des BRPs über einen bestimmten Schwellenwert hinaus zeigt sich die lytische Wirkung der BRP-Aktivität. Stegehuis *et al.* (1995) beobachteten eine vollständige Inhibition der Proteinbiosynthese 60 Minuten nach Induktion. Konsistent mit diesen Beobachtungen ist auch die Regulation verschiedener osmotisch-induzierbarer Gene wie *betA*, *mdaH*, *osmC* so wie die differentielle Expression von Strukturgenen von Enzymen und Komponenten zur Peptidoglykan- und LPS-Biosynthese wie *ompX*, *htrB*, *mltC*, *yciD*, *tonB*, *slyB* und *dacB*. Das Gen *ompX* codiert für ein Membranporin der äußeren Membran, welches eine wichtige Rolle in der Colicin-Aufnahme spielt (Wu *et al.* (2009)). *TonB* fungiert als Energieumwandler, welcher die PMK über der inneren Membran für den Import von Eisensiderophoren und Vitamin B12 durch die äußere Membran bereitstellt. *ExbB* fungiert in der Interaktion mit *TonB* in der Substrataufnahme über die äußere Membran. Das Protein interagiert mit einer ganzen Reihe von Transportproteinen der äußeren Membran und steht in engem Kontakt mit beiden Membranen (Held *et al.* (2002)). Ein solcher Transporter wird von dem Gen *fecA* codiert, welches ebenfalls positiv reguliert wurde. *FecA* ist in der äußeren Membran lokalisiert und fungiert in der *TonB*-abhängigen Aufnahme von Eisen (Nikaido *et al.* (1998)). Viele weitere Proteine, welche in die bakteriellen Eisenaufnahme und Zellatmung involviert sind, wurden nach dem Induktionsstimulus reguliert vorgefunden, wie unlängst an den Beispielen *exbB*, *fecA* und *tonB* gezeigt wurde. Offenbar verursacht die BRP-Induktion mit ihrer degradativen Auswirkung auf die äußere Membran einen Eisenmangel in der Zelle. Die Eisenaufnahme geschieht über Porine der äußeren Membran, die BRP-induziert verloren

gehen oder in ihrer Funktionalität eingeschränkt werden und infolge dessen vermehrt nachproduziert werden. Die Regulation der beiden Hitzeschockproteine IbpA und IbpB, welche in einem Operon organisiert sind, deutet auf eine weitere zelluläre Stressantwort hin.

Für die negativ regulierten Gene der drei evaluierten Vergleiche erfolgt im folgenden eine detaillierte Diskussion der Stress- und sekretorischen Antwort der Zelle. Zusammenfassend lässt sich für die negativ regulierten Gene eine hohe Anzahl solcher Gene beobachten, welche in der Aminosäuren- und Nukleotidbiosynthese eine Rolle spielen.

negativ regulierte Gene

30 Minuten nach Induktion wird das Transkriptionsniveau der osmotisch induzierbaren Gene *osmC*, *rsbB* und *betA* und Genen wie *ftsI*, *mltE*, *dacD*, welche in die Peptidoglykanbiosynthese involviert sind, reduziert. Weiterhin interessant ist die Regulation der Hitzeschockproteine IbpA und IbpB, die genetisch zusammen in einem Operon vorliegen (Zhao *et al.* (2005)). In die LPS-Biosynthese der äußeren Membran involviert sind die Gene *lpxC*, *B0805* und *ompF*, welche ebenfalls negativ differentiell exprimiert wurden. OmpF ist zudem durch seine Funktion als Membranporin der äußeren Membran hinsichtlich der Proteinsekretion interessant (Wu *et al.* (2009)).

60 Minuten nach Induktion werden die Gene *glpDFKT* negativ reguliert, sie sind in einem Operon organisiert und ihre Proteinprodukte werden in der Zelle für die Metabolisierung von Glycerin als Kohlenstoffquelle benötigt (Lin *et al.* (1976)). Durch die Supplementierung der Arabinose als Induktor findet eine Umstellung des Kohlenstoffmetabolismus von Glycerin auf die energetisch vorteilhaftere Arabinose statt. Als Folge dessen werden jene Gene, die in den Glycerinmetabolismus involviert sind, negativ reguliert. Die negative Regulation der Glycerinmetabolismus-Gene kann somit, analog den Arabinosemetabolismus-Genen, als Positivkontrolle für die biologische Sinnhaftigkeit der Chipdaten fungieren. Weiterhin interessant ist die negative Regulation zweier großer Operons. Die Identifizierung von regulierten Operons gibt ebenfalls einen Hinweis auf biologisch sinnvolle Chipergebnisse, da die zufällige, gleichzeitige Regulation mehrerer Gene eines Operons statistisch höchst unwahrscheinlich ist. Neben dem Operon, welches in den Glycerinmetabolismus involviert ist, sind das Operon *cysACIMW* für die Sulfat-Aufnahme und das Operon *gatABCDY* für die Galactitolmetabolismus negativ reguliert. Kleinere Operons in Form von *pyrBI*, involviert in die Pyrimidin-Biosynthese, und *ylcAB* wurden ebenfalls als differentiell exprimiert identifiziert. Ein BRP-induzierter Einfluß auf die Nukleotidbiosynthese manifestiert sich durch die Regulation von *pyrBI*, wohingegen der Einfluß der BRP-Induktion auf die Konstitution der äußeren Membran durch die Regulation von *ylcAB* verdeutlicht wird. *YlcA* und *YlcB* stellen, genau wie das Proteinprodukt des ebenfalls regulierten Gens *ompF*, Membranporine der äußeren Membran dar. Diese besitzen wichtige Funktionen in der Stoffaufnahme aus dem extrazellulären Raum (Wu *et al.* (2009)). Ihre Funktion in der Colicin-Aufnahme verdeutlicht einen Zusammenhang mit der BRP-Expression. Das zelluläre RCS-System ist maßgeblich in der Entstehung von Biofilmen involviert und das partizipierende Gen *rcsC* wird ebenfalls in dieser Auswahl beobachtet (Brill *et al.* (1988)).

Der Vergleich der transkriptionellen Zustände 30 und 60 Minuten nach Induktion gleicht im Wesentlichen dem unlängst diskutierten Vergleich 60 vs. 0 Minuten. Dieses gibt einen guten Hinweis auf die Reproduzierbarkeit und biologische Sinnhaftigkeit des Datensatzes. Die Operons *cysACDIJNPW* und *gatABC* werden ebenfalls negativ reguliert vorgefunden. Gleiches gilt für das Operon der Pyrimidinbiosynthese *pyrBI*. Das Operon *sdhBC* codiert

für Proteine, die in den Zellmetabolismus und die aeroben Respiration involviert sind. Die generelle negative Regulation ganzer Operons, die eine essentielle Funktion für den Zellmetabolismus aufweisen, korreliert mit der in der Literatur berichteten BRP-induzierten negativen Auswirkung auf den allgemeinen Zellmetabolismus (Luirink *et al.* (1991), Stegehuis *et al.* (1995)). Die größten transkriptionellen Umstellungen in dieser Hinsicht finden, bedingt durch die BRP-Akkumulierung, zwischen den Zuständen 30 und 60 Minuten nach Induktion statt.

4.4.3 Zusammenfassung

Anhand des *E. coli* Stammes MG1655 pBADtorAmalELppBRP erfolgte die Analyse der Genexpression der Tat-spezifischen Sekretion in den extrazellulären Raum. Die regulierten Gene wurden mit Hilfe der Multifun-Datenbank klassifiziert und biologischen Funktionen zugeordnet. Die größte relative Regulation wurde bei den Funktionsklassen Verteidigung, Stress und Zellprozesse beobachtet. Eine Evaluierung der Verteilung der regulierten Gene in die einzelnen Multifun-Subklassen ergab für die Klasse Lokalisierung der Genprodukte hohe prozentuale Anteile im Periplasma und den beiden Zellmembranen. Dabei kann es sich sowohl um einen BRP-induzierten als auch um einen Arabinose-induzierten Effekt handeln. Im Bereich der Zellprozesse wurden die größten regulatorischen Effekte bei den Subklassen Stress, Verteidigung, Virulenz und Biofilmbildung beobachtet. Für die Funktionsklasse Zellstruktur ergab eine äquivalente Evaluierung die größten Effekte bei den Oberflächenantigenen, dem Peptidoglykan und der Kapselbildung. Diese transkriptionellen Auswirkungen wurden vornehmlich der BRP-Expression zugeschrieben. Die beobachtete positive Regulation der primären Arabinosemetabolismusgene araBAD und der Arabinoseimportgene araEFGH wurde als Positivkontrolle für die biologische Verifikation der Analyseergebnisse verwendet. Neben Stress- und Abwehr-involvierten Genen trat die differentielle Expression von verschiedenen Komponenten der LPS- und Peptidoglykan-Biosynthese auf. Zudem wurde die Regulation von verschiedenen Chaperonen (z.B. FkpA, HybB, HslUV und FliT), Membranstressmarkern (z.B. lolAB, spy, pspABCDE) und Komponenten zur Aufrechterhaltung des Protonengradienten über der Cytoplasmamembran (z.B. mdbE) beobachtet. Die Regulation von Genen wie mltC und tolC deutet auf eine mechanistische Beteiligung dieser im Sekretionsvorgang hin. Aus sekretorischer Sicht ist besonders die Regulation der verschiedenen Chaperone und von pspA interessant, da sich diese Gene als erfolgreich in verschiedenen Coexpressionsstrategien erwiesen haben (Li *et al.* (2003), De Lisa *et al.* (2004), Perez-Rodriguez *et al.* (2007)).

4.5 Analyse der extrazellulären, Sec-spezifischen Sekretion

Der Induktor des Arabinosepromoters, *L*-Arabinose, wird von *E. coli* K12 als Kohlenstoffquelle metabolisiert. Als Folge dessen findet ein rasches Abklingen des Induktionsstimulus der promoterregulierten Gene in der Regel schon nach wenigen Stunden statt. Für eine effiziente Sekretionsstrategie im Kontext hoher sekretierter Produktausbeuten sollte die BRP-Expression konstant gehalten werden. Zu diesem Zweck wurden *E. coli* Stämme mit einem defizienten Arabinosemetabolismus entwickelt und sind teilweise kommerziell erhältlich. Eine Deletion der chromosomal lokalisierten Gene araBAD, welche in der primären Arabinosemetabolisierung involviert sind, verhindert die Degradation der *L*-Arabinose. Das

initiale Enzym stellt dabei die vom Gen *araA* codierte Arabinose-Isomerase dar, welche die Konversion von *L*-Arabinose zu *L*-Ribulose katalysiert. Ein erfolgreicher *knock-out* des Arabinosemetabolismus muss daher obligat mindestens dieses Enzym umfassen. Bis dato etablierte Stämme mit defektem Arabinosemetabolismus zeigen jedoch einige Nachteile, wie eine geringe Sekretionseffizienz, uncharakterisierte Mutationen und unvollständige Sequenzinformationen (Sommer *et al.* (2008)). Aus diesem Grunde wurde auf Basis des vollständig sequenzierten und gut charakterisierten *E. coli* Stammes K12 MG1655 ein Derivat konstruiert, welches einen Austausch der Gene *araA* und *araB* gegen das CAT-Gen aufweist, und daher nicht mehr in der Lage ist, Arabinose zu verstoffwechseln. Der Genaustausch erfolgte mittels homologer Rekombination (Sommer *et al.* (2008)). Basierend auf diesem Stamm, welcher als Ara1655 bezeichnet wurde, erfolgte die Analyse der Proteinsekretion anhand des plasmidbasierten Reportersystems pBADLppBRPphoAbla. Das Plasmid codiert für eine Signalpeptid-modifizierte Variante des BRPs von CloDF13 (Luirink *et al.* (1991)), dessen Expression unter der Kontrolle des Arabinose-Promoters steht. Die beiden Reporterproteine alkalische Phosphatase (PhoA) und β -Lactamase (Bla) wurden stromabwärts des Promoters der β -Glukanase kloniert und werden somit konstitutiv exprimiert. Beide Proteine werden mittels des Sec-Pfades in das Periplasma sezerniert und gelangen von dort BRP-induziert in den extrazellulären Raum. Schematisch ist dieser Vorgang in Abbildung 24 gezeigt.

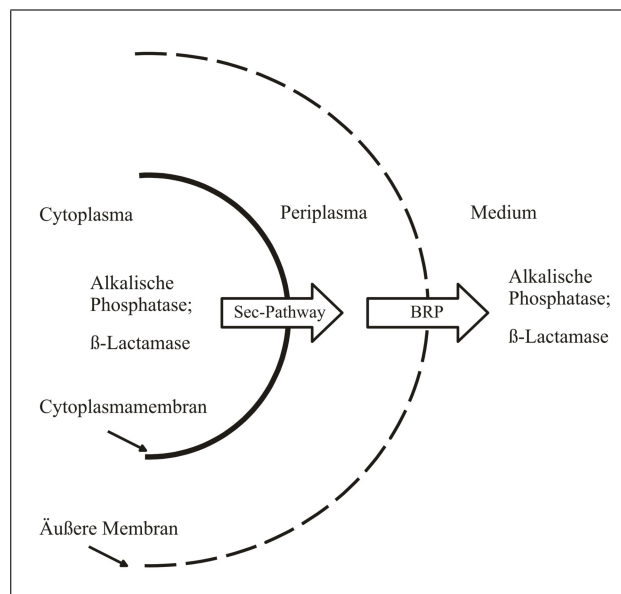


Abb. 24: Schema der Sekretion der Reporterproteine Alkalische Phosphatase und β -Lactamase des Stammes Ara1655 pBADLppBRPphoAbla. Die beiden Reporterproteine werden mittels des Sec-Pathways in das Periplasma von *E. coli* translokiert. Durch die Aktivität des Lpp-BRPs gelangt das periplasmatische Intermediat schließlich durch die äußere Membran in den extrazellulären Raum.

Es wurden reproduzierbare Bioreaktorkultivierungen des beschriebenen Sekretionsstammes durchgeführt und Proben vor, sowie 30 und 60 Minuten nach Induktion dem Reaktor entnommen. Auf Basis dieser Proben erfolgte die transkriptionelle Analyse.

In zwei unabhängigen Reaktorläufen erfolgte die Induktion in der mittleren exponentiellen Phase mit zwei unterschiedlichen Induktionsstärken, es wurden dazu 1000 und 25 mg/l *L*-Arabinose der Kultivierung zugesetzt. Somit lassen sich die Auswirkungen unterschiedlicher Expressionsniveaus des BRPs auf das *E. coli* Transkriptom evaluieren.

4.5.1 Induktion mit 1000 mg/l *L*-Arabinose

Die Ergebnisse der differentiellen Genexpressionsanalyse des Sekretionsstammes bei einem Induktionsniveau von 1000 mg/l *L*-Arabinose werden nachfolgend präsentiert. 30 Minuten nach Induktion wurde die differentielle Expression von 91 Genen identifiziert, dies entspricht einem prozentualen Anteil von 2,04 Prozent aller 4460 ORFs von *E. coli*. Von diesen wurden 58 positiv und 33 negativ reguliert. Beim Vergleich des Transkriptoms 60 Minuten nach Induktion gegen den Zustand vor Induktion wurden 42 runterregulierte und 44 hochregulierte Gene beobachtet. Dies entspricht einer Gesamtzahl von 86 regulierten Genen, welche 1,93 Prozent des *E. coli*-Genomes repräsentieren. Zusätzlich wurde ein Vergleich der beiden Zustände nach Induktion (30 und 60 Minuten) durchgeführt, dabei wurden 78 regulierte Gene identifiziert, die sich in 48 negativ und 30 positiv regulierte Gene gliedern. 1,75 Prozent aller *E. coli* ORFs gehört damit zu den regulierten Genen. Um den Datensatz einer biologischen Interpretation zugänglich zu machen, wurde die Multifun-Datenbank für eine Funktionszuordnung der regulierten Gene verwendet. Es ergaben sich äquivalente Ergebnisse zu den vorangegangenen transkriptionellen Analysen (Abschnitt 4.4.1). Auf eine gesonderte Darstellung derselben wurde daher verzichtet.

Durch die starke Induktion der BRP-Expression mit 1000 mg/l *L*-Arabinose wurde eine substantielle Zellreaktion erwartet. Die BRP-Expression hat signifikante Auswirkung auf die Integrität der beiden Zellmembranen. In experimentellen Studien wurde belegt, dass die Auswirkungen der BRP-Expression ähnliche Zellreaktionen hervorrufen, wie sie beim Hitzeschock oder einer Phageninfektion beobachtet werden (Van der Waal *et al.* (1995)). Nachfolgend erfolgt eine Diskussion von differentiell exprimierten Genen der einzelnen transkriptionellen Evaluierungen, die Hinweise auf diese Zellreaktion geben oder mechanistisch am Sekretionsvorgang beteiligt sein könnten. Zuerst erfolgt die Präsentation der positiv regulierten Gene.

Vergleich 30 vs. 0 Min

Beim Vergleich der transkriptionellen Zustände 30 Minuten nach Induktion gegen dem Zustand vor Induktion werden die Gene *araEFGH*, welche für die Arabinosetransporter codieren, positiv reguliert vorgefunden. Die biologische Sinnhaftigkeit dieser Regulation ist eindeutig gegeben und kann als Positivkontrolle fungieren. Eine Regulation der Gene *araBAD*, wie sie bei der transkriptionellen Analyse der Tat-spezifischen extrazellulären Sekretion beobachtbar wurde (Abschnitt 4.4.1), trat nicht auf. Dieses Kriterium kann daher als Negativkontrolle fungieren, da diese Gene im Ara1655 Stamm nicht mehr exprimiert werden können (*knock-out* des Arabinose-Metabolismus). Als weitere Effekte der Arabinose-Supplementierung wird die Regulation weiterer Gene beobachtet, welche nachfolgend, geordnet nach ihrer biologischen Funktion, stichpunktartig angegeben werden.

Eine detaillierte Diskussion dieser Gene ist bereits vorangehend erfolgt (Abschnitt 4.4.1) oder wird zusammen mit den Ergebnissen einer weiteren transkriptionellen Evaluierung auf Basis desselben Sekretionsstammes in Abschnitt 4.5.2 durchgeführt.

- Biofilm-Gene: galEKM, ycdPQ,
- Hitzeschock-Gene: Ddg, degS, ygeA
- LPS-Biosynthese: lolAB
- Membranstress/SOS-Antwort: spy, vacJ, soxS, uvrCD, galP

Vergleich 60 vs. 0 Min

60 Minuten nach Induktion wurden in guter Übereinstimmung zu dem Vergleich 30 vs. 0 Minuten nahezu alle Gene übereinstimmend reguliert vorgefunden. Zusätzlich sind verschiedene Gene positiv reguliert, welche in die Tryptophanbiosynthese involviert sind. Es handelt sich dabei um die Gene trpCDEI.

Vergleich 60 vs. 30 Min

Ein Vergleich der beiden postinduktiven Zustände identifiziert diese Tryptophan-assoziierten Gene ebenfalls als reguliert. Die ATP-Hydrolyse stellt die entscheidende Triebkraft der Translokation durch die Cytoplasmamembran und die spezifischen Arabinose-Aufnahmesysteme der ABC-bindenden Kasette dar. Hinsichtlich dieser Aspekte wird die Regulation des Genes *atpG* verständlich. Ein sehr interessantes Beispiel für die Applikation der Microarray-Technologie zur Zuordnung von biologischen Funktionen stellt die beobachtete Regulation der Gene *ytfHQR* dar. Basierend auf der Sequenzhomologie wurde den Genprodukten von dem Operon *ytfQRST* die Funktion eines ATP-abhängigen Zuckertransporters zugeschrieben. *YtfR* soll dabei die ATP-bindende Komponente, *YtfT* die membrangebundene Komponente und *YtfQ* das Bindungsprotein sein (Saurin *et al.* (1999)). Bisher existierte kein experimenteller Nachweis, welcher hiermit, durch die Regulation dieser Gene unter dem Stimulus der *L*-Arabinose Supplementierung, erbracht wurde. *OmpA* stellt eines der häufigsten Proteine in der äußeren Membran dar und fungiert als Porin für die Aufnahmen verschiedener Substrate durch die äußere Membran. Am bedeutendsten ist die zentrale Rolle, welche *OmpA* bei der BRP-induzierten, semispezifischen Freisetzung von Bacteriocinen und des periplasmatischen Inhalts spielt (van der Waal *et al.* (1995)). Auffällig ist weiterhin die deutliche Regulation von einigen Prophagen-Genen. Zum Beispiel codieren die Gene B1578 und B1550 für Produkte des Qin-Prophagen. Die Überproduktion des Genproduktes von B1550 verursacht bei SecG-Null-Mutanten den Einbau einer größeren Anzahl ungesättigter Fettsäuren in die Membranen und damit einhergehend eine größere Membranfluidität. Dieses erleichtert die Membranpassage für verschiedene Sekretionssubstrate (Sugai *et al.* (2001)). Bedenkt man, dass die Translokationskapazitäten durch die simultane Sekretion des BRPs, PhoA und Bla saturiert sind, wird die Regulation zur Erhöhung der Sekretionseffizienz verständlich. Weiterhin könnte auch eine Involvierung in die BRP-Aktivität zur Proteinfreisetzung aus dem Periplasma postuliert werden. Möglicherweise stellt die BRP-Expression auch einen Stimulus für Prophagengene dar, aus dem lysogenen Zyklus in den lytischen Zyklus zu wechseln (Lewin B. Gene: (1995)).

Weitere interessante, positiv regulierte Gene werden im Folgenden aufgelistet:

- Biofilm-Gene: cutCF, ycfJ, yceP, hha
- Hitzeschock-Gene: treA, GroEL
- LPS-Biosynthese: lolB
- Membranstress/SOS-Antwort: ompA

Negativ regulierte Gene

Im Folgenden wird der Fokus auf die negativ regulierten Gene gelegt. Generell werden weit weniger stress-induzierbare Gene negativ reguliert. Vornehmlich ist eine Auswirkung der Arabinosezugabe und der damit einhergehenden BRP-Expression auf die Proteinbiosynthese zu beobachten. Dies manifestiert sich in der negativen Regulation vieler Gene, die für Komponenten der Biosynthese verschiedener Aminosäuren wie Histidin, Arginin und Prolin codieren. Der negative Effekt der BRP-Expression auf die Proteinbiosynthese wurde zuerst von Stegehuis *et al.*(1995) beobachtet und koinzidiert mit den hier gefundenen Ergebnissen. Experimentell wurde eine Stunde nach Induktion des Colicin-Genclusters eine signifikante Reduzierung der generellen Proteinbiosynthese beobachtet. Die vollen regulatorischen Auswirkungen auf die Aminosäuren- und Nukleotid-Biosynthese sind im vorliegenden Fall evident dazu ebenfalls nach 60 Minuten auszumachen. Als konkrete Beispiele mögen die Gencluster purBCFRMNTL und carAB fungieren. Der Vergleich der beiden postinduktiven Zustände offeriert keine weitere Regulation von Genen, die ein erweitertes Verständnis von der zellulären Antwort auf die BRP-Expression oder den Sekretionsvorgang bieten. Die meisten negativ regulierten Gene des Vergleiches 60 vs. 0 Minuten werden auch in dieser Analyse wiedergefunden.

4.5.2 Induktion mit 25 mg/l *L*-Arabinose

Der Stamm Ara 1655 pBADLppBRPphoAbla wurde unter analogen Bedingungen zu der vorhergehenden Analyse kultiviert und mit 25 mg/l *L*-Arabinose in der mittleren exponentiellen Wachstumsphase induziert. Wiederum wurden die transkriptionellen Unterschiede von Proben, welche direkt vor sowie 30 und 60 Minuten nach Induktion dem Reaktor entnommen wurden, evaluiert. Es zeigte sich, dass 30 Minuten nach Induktion im Vergleich zu dem präinduktiven Zustand 297 Gene positiv und 81 Gene negativ reguliert waren. Dies entspricht einem prozentualen Anteil von 8,48 Prozent an regulierten Genen, bezogen auf die Gesamtzahl aller in *E. coli* existenten Gene. Für den analogen Vergleich 60 Minuten nach Induktion wurden 210 Gene positiv und 129 Gene als negativ reguliert identifiziert. Diese repräsentieren somit 7,83 Prozent aller 4460 in *E. coli* existierenden ORFs. Beim Vergleich der beiden post-induktiven Proben wurde die positive Regulation von 64 und die negative Regulation von 165 Genen beobachtet, welche zusammen 5,13 Prozent des *E. coli* Genoms repräsentieren. Zusammenfassend lässt sich somit konstatieren, dass der Induktionsstimulus hauptsächlich zu einer positiven Regulation führte. Nach Induktion werden zwischen den beiden Zeitpunkten 30 und 60 Minuten allerdings vornehmlich negativ regulierte Gene beobachtet. Dies kann zum einen ursächlich mit der zunehmenden BRP-Akkumulierung verbunden sein oder auf eine schnelle Adaption der Zelle auf den Induktor hindeuteten.

Um die Komplexität der Datenanalyse zu erleichtern und eine Fokussierung auf besonders stark involvierte Funktionsklassen/Zellregionen zu ermöglichen, wurde der Datensatz mit Hilfe der Multifun-Datenbank biologischen Funktionalitäten zugeordnet. Auf eine separate Darstellung der Ergebnisse wurde verzichtet, da diese äquivalent zu vorangehenden Transkriptomanalysen ausfielen (Abschnitt 4.4.1). Es wurde erwartet, dass das moderatere Induktionsniveau (25 mg/l *L*-Arabinose) und die damit einhergehende geringere zelluläre BRP-Konzentration zu einer schwächeren zellulären Stressantwort und zu geringeren Auswirkungen auf die Membranintegritäten führen würde. Erstaunlicherweise zeigte sich jedoch die Regulation einer ganzen Reihe von Genen, die maßgeblich in die LPS-Biosynthese involviert sind oder ihre Funktion in der äußeren Membran wahrnehmen und essentiell für die Colicin-Aufnahme sind. Der offensichtliche biologische Sinn als Reaktion auf die BRP-Expression offenbart sich bei dem niedrigeren Induktionsniveau also im äquivalenten Maße. Analog zu den vorherigen Beobachtungen zeigt sich die unmittelbare Auswirkung der BRP-Expression erst in vollem Umfang 60 Minuten nach Induktion. Zu diesem Zeitpunkt werden auch die größten regulatorischen Effekte auf die Gene verzeichnet, deren Proteinprodukte für die LPS-Biosynthese eine zentrale Rolle spielen. Zuerst wird der Fokus auf die positiv regulierten Gene gelegt.

Vergleich 30 vs. 0 Min

30 Minuten nach Induktion sind die Gene *tolQ* und *arcAE* gegenüber den Referenzzustand vor Induktion differentiell exprimiert. *TolQ* und *ArcE* sind innere Membranproteine, die Bestandteile des Tol-Pal-Systems zur Aufrechterhaltung der Membranintegrität darstellen (Perouiche *et al.* (1995)). Zusätzlich impliziert die essentielle Bedeutung dieses Systems für die Colicin A Aufnahme einen Zusammenhang mit der BRP-Expression. *TolQ* und die zwei weiteren Membranproteine *TolA* und *TolR* sind strukturell und funktionell homolog zu *TonB*, *ExbB* und *ExbD*, welche ebenfalls unter Ausnutzung der protonenmotorischen Kraft über der inneren Membran Substrate importieren. Die Gene *tonB* und *exbB* wurden ebenfalls in einigen Vergleichen differentiell exprimiert vorgefunden. Dem Tol-Pal-System wird eine zusätzliche Rolle bei Infektionen mit filamentösen Phagen zugeordnet. *TolQ* ist zudem aus sekretorischer Sicht höchst interessant, da es ein Bestandteil des Haemolysin-Sekretionsweges in *E. coli* ist (Blight *et al.* (1994)). In Koinzidenz mit der differentiellen Expression von *tolQ* und *arcE* wird das Gen *btuB* reguliert, welches ebenfalls eine signifikante Bedeutung bei der Aufnahme von Gruppe A Colicinen hat. Das Gen codiert für ein äußeres Membranporin, dass sich für die Aufnahme von Vitamin B12 und von Bakteriophagen verantwortlich zeigt. *BtuB* wird mittels *TonB*-abhängigen Transport in die äußere Membran dirigiert und interagiert dort mit *OmpF* (Jones *et al.* (2003)). Das Gen *pphA* stellt den Aktivator der Hitzeschockantwort der Zelle dar. Interessanterweise wird 30 Minuten nach dessen positiver Regulation die ebenfalls positive differentielle Expression des Gens *rpoH* detektiert, welches für den Hitzeschock-Sigmafaktor codiert. Passend in diesem Kontext wird zudem die Regulation des Gens *ibpB* beobachtet. Ebenso interessant ist die positive Regulation der Gene *glpABT*, die im Glycerinmetabolismus der Zelle eine zentrale Rolle spielen (Lin *et al.* (1976)). Die Induktion dieser Gene wurde bei dem Wildtyp-Stamm MG1655 nicht beobachtet, da nach Induktion eine Umstellung auf den Arabinose-Stoffwechsel erfolgte. Der Stamm Ara1655 vermag Arabinose nicht zu metabolisieren und verwendet daher weiterhin Glycerin als C-Quelle. Insofern scheint es biologisch sinnvoll, dass wachstumsbedingt die Expression dieser Gene erhöht wird. Zusätzlich fällt die differentielle Expression der Arabi-

noseimportgene deutlich schwächer aus als bei dem höheren Induktionsniveau. Auch dieses scheint im Kontext der deutlich geringeren Induktorkonzentration erklärbar und biologisch sinnvoll zu sein. Aufgrund der konstitutiven Expression der Reporterproteine PhoA und Bla, sowie der induktiven Expression des LppBRPs, welches alle Sec-Substrate darstellen, kann davon ausgegangen werden, dass die Sec-Translokase saturiert vorliegt. Insofern wäre eine regulatorische Antwort der Zelle auf diesen Zustand wahrscheinlich. Tatsächlich wird das Gen *secE*, welches für den zentralen membrangebundenen Teil der Translokase codiert, positiv reguliert vorgefunden. In diesem Zusammenhang ist auch die positive Regulation des SRP-Rezeptors FtsY hervorzuheben. Es existieren deutliche experimentelle Hinweise darauf, dass die Translokationen mittels des SRP-Patways den zentralen, membrangebundenen Teil der Sec-Translokase für die Sezernierung der Substrate ins Periplasma verwenden (Beck *et al.* (2000)). Unter dem Stimulus der Saturierung der cytoplasmatischen Sec-Komponenten könnte somit ein Teil des Sec-Substrate in SRP-abhängiger Art und Weise translokiert werden. Mergulaho *et al.* (2005) berichten über äquivalente Beobachtungen, was aufzeigt, wie die Zelle die Saturierung der Translokationskapazitäten durch einen *Bypass* umgeht. Bei RpoE handelt es sich um den Sigmafaktor der RNA-Polymerase, welcher als Reaktion auf Membranstress und Eisenmangel vermehrt ausgeschüttet wird. Koinzidierend dazu tritt die Regulation des Genes *rseC* auf, welches für den Transkriptionsaktivator der Eisenaufnahmegene codiert. Solche Gene wurden z.B. mit *fepG* und *soxR* differentiell exprimiert vorgefunden. Die Initiierung der zellulären Antwort auf Membranstress wurde im Rahmen der evaluierten transkriptionellen Analysen bereits eingehend diskutiert und experimentell bewiesen. Eine weitere äußerst interessante Beobachtung stellt die differentielle Expression des codierenden Genes vom Phagenschock-Repressor *pspA* dar. De Lisa *et al.* (2004) konnten experimentell beweisen, dass eine Coexpression dieses Genes zu einer signifikanten Erhöhung der sekretierten Proteinmenge führt. Die PspA-Aktivität erlaubt aufgrund ihrer membranstabilisierenden Wirkung nicht nur die Realisierung höhere Exportraten (Aufrechterhaltung der PMK als Triebkraft des Exports), sondern schützt die exprimierenden Zellen auch zusätzlich vor den letalen Folgen der BRP-Aktivität. Insofern weist *pspA* ein immenses Potential für eine erfolgreiche Coexpressionstrategie auf. Weitere interessante regulierte Gene für die LPS- und Peptidoglykanbiosynthese sind: B0805, *blc*, *wecB*, *rfaF*, *yaT*, B2250, *msbB*, B1371, *oppD* und *hha*. Auf eine detaillierte Diskussion dieser sei auf das vorangehende Kapitel 4.4.1 verwiesen.

Vergleich 60 vs. 0 Min

Für den zweiten transkriptionellen Vergleich 60 Minuten nach Induktion gegen den Zustand vor Induktion werden nahezu alle unlängst diskutierten Gene ebenfalls als positiv reguliert identifiziert. Diese hohe Wiederfindungsrate gibt einen Hinweis auf einen anhaltenden Induktionsstimulus und zeigt die Reproduzierbarkeit und biologische Sinnhaftigkeit der durchgeführten Analyse auf.

Vergleich 60 vs. 30 Min

Bedingt durch die hohe Anzahl an identischen, regulierten Genen für die beiden Vergleiche 60 und 30 gegen 0 Minuten gibt es im Vergleich der beiden postinduktiven Zustände nur wenig regulatorische Effekte zu beobachten.

Negativ regulierte Gene

Zusammenfassend lässt sich konstatieren, dass bedeutend weniger Stress-induzierbare und mit der Membranfunktion assoziierte Gene negativ reguliert vorgefunden wurden, als dies bei den positiv regulierten Genen der Fall war. Hauptsächlich tritt die Regulation von Genen auf, die in den Zellmetabolismus involviert sind oder wichtige Rollen in der Biosynthese von Aminosäuren und Nukleotiden bekleiden. Die Regulation größerer Operons wie *cysAWPM* und *hisGFQI* verifiziert die biologische Sinnhaftigkeit der evaluierten Ergebnisse.

4.5.3 Induktionsniveau 25 mg/l *L*-Arabinose; Aufnahme einer feingliedrigeren Kinetik

Anhand von Proben, die einer Bioreaktorkultivierung des Stammes Ara1655 pBADLppBR-PphoAbla unmittelbar vor sowie zu verschiedenen Zeitpunkten (5, 10, 20, 30, 45, 60, 90 Minuten) nach der Induktion mit 25 mg/l *L*-Arabinose entnommen wurden, erfolgte eine Evaluierung der differentiellen Genexpression. *E. coli* und andere Prokaryonten vermögen auf externe Stimuli sehr rasch durch Adaption ihrer Genexpression zu reagieren. Das gut charakterisierte *lac*-Operon wird durch Zugabe von Lactose induziert, binnen 2 Minuten entstehen signifikante Mengen mRNA der codierenden Gene *lacAYZ*. Dies stellt einen Wert dar, welcher für die meisten mRNA-Spezies typisch ist (Lewin *et al.* (1995)). Um eine entsprechend rasche Umstellung auf variierende Umweltbedingungen zu vollziehen, sind zudem viele mRNA-Spezies entsprechend instabil und besitzen Halbwertszeiten von nur wenigen Minuten. Nicht benötigte Gene können so binnen sehr kurzer Zeit auf transkriptioneller Ebene stillgelegt werden, während essentielle mRNA-Transkripte schnell synthetisiert werden. Gleiches gilt für die regulatorischen Vorgänge auf translationaler Ebene. Noch während die mRNA von der RNA-Polymerase *de novo* synthetisiert wird, assoziieren die Ribosomen mit dieser und initiieren den Translationsvorgang. Transkription und Translation erfolgen daher in prokaryontischen Organismen mehr oder minder simultan und die Proteinprodukte der codierenden Gene erscheinen ungefähr eine halbe Minute nach ihren mRNA-Spezies in der Zelle (Lewin B., *Gene* (1995)). Aus den angeführten Gründen wurde eine feingliedrigere Kinetik der differentiellen Genexpression mit Probennahmen zu früheren Zeitpunkten nach Induktion evaluiert, um diesen Umstand Rechnung zu tragen. Zusätzlich ergaben proteinanalytische Daten im betrachteten Sekretionssystem, dass ein Hexahistidin-*getagtes* LppBRP schon 2 Minuten nach Induktion mit *L*-Arabinose mittels eines Western-Blot detektierbar war (Sommer *et al.* (2008)). Begünstigt wird die rasche Bildung des LppBRPs durch seine geringe molekulare Größe von nur 3 kDa. Dieses Faktum bestätigt die gewählte Vorgehensweise zusätzlich, da somit eine schnelle, BRP-induzierte Auswirkung auf das Transkriptom wahrscheinlich ist. Ebenso lassen sich so frühe transkriptionelle Effekte, die aufgrund der Arabinose-Supplementierung auftreten, besser verfolgen. Für die im folgenden diskutierte Kinetik wurde die Probe, deren Entnahme unmittelbar vor der Induktion erfolgte, als Referenz verwendet, auf die sich alle Proben die zu späteren Zeitpunkte entnommen wurden bezogen. Die differentielle Genexpression aller evaluierten Vergleiche ist relativ zu dieser angegeben. 5 Minuten nach Induktion wurden 901 Gene differentiell exprimiert, von denen 278 positiv und 623 negativ reguliert waren. Die Gesamtzahl der regulierten Gene repräsentiert 20,2 Prozent aller 4460 ORFs in *E. coli*. Für den analogen Vergleich 10 Minuten nach Induktion wurde die Regulation von 1309 Genen beobachtet (293 positiv, 1016 negativ), was einen prozentualen Anteil von 29,35 Prozent aller in *E. coli* existenten Gene darstellt.

4.5 Analyse der extrazellulären, Sec-spezifischen Sekretion

Für die weiteren evaluierten Zeitpunkte ergaben sich folgende Anzahlen an regulierten Genen:

- 20 vs. 0 Min : 626 (115 positiv; 511 negativ)
- 30 vs. 0 Min : 177 (87 positiv; 90 negativ)
- 45 vs. 0 Min : 302 (265 positiv; 37 negativ)
- 60 vs. 0 Min : 458 (435 positiv; 23 negativ)
- 90 vs. 0 Min : 220 (170 positiv; 50 negativ)

Dies entspricht prozentualen Werten von 14,04 (20 Min); 3,97 (30 Min); 6,77 (45 Min); 10,27 (60 Min) und 4,93 Prozent (90 Min) relativ zu allen existierenden *E. coli* K12 Genen. Graphisch wird die Anzahl der positiv und negativ regulierten Gene für die einzelnen Vergleiche in Abbildung 25 dargestellt.

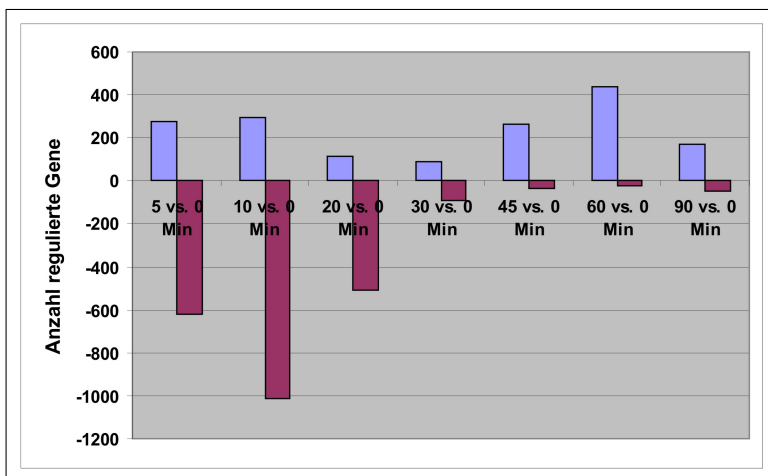


Abb. 25: Anzahl der negativen und positiven regulierten Gene für die verschiedenen evaluierten Vergleiche.

Interessant ist die Beobachtung, dass zu den frühen Zeitpunkten nach Induktion die Mehrzahl der differentiell exprimierten Gene negativ reguliert wurden. 30 Minuten nach Induktion ergab sich dann ein paritätisches Verhältnis von negativ zu positiv regulierten Genen, wohingegen für die späteren Zeitpunkte fast nur positiv regulierte Gene beobachtet wurden. Es fand somit eine substantielle Umstellung der Genexpression von vornehmlich negativer Regulation zu fast ausschließlich positiver Regulation statt. Da die späteren Zeitpunkte 45, 60 und 90 Minuten nach Induktion mit der zunehmenden Akkumulation des BRPs korrelieren, und die substantiellsten Auswirkungen der BRP Expression phänotypisch 60 Minuten nach Induktion beobachtet wurden (Stegehuis *et al.* (1995)), kann vermutet werden, dass es sich bei der Mehrzahl der positiv regulierten Gene um eine zelluläre Antwort auf die BRP-Expression und die rekombinante Proteinsekretion handelt. Analog der Vorgehensweise bei vorherigen Transkriptomanalysen wurden den regulierten Genen mittels der Multifun-Datenbank biologische Funktionen zugeordnet.

4.5.4 Cluster Analyse

Der Datensatz der evaluierten Vergleiche wurde einer Clusteranalyse mittels des *k-means* Algorithmus unterworfen. Als Ähnlichkeitsmaß fungierte der Korrelationskoeffizient, als Rohdaten für das *Clustering* wurden die normierten und hintergrundbereinigten Median-Werte der Signalintensität verwendet. Um den Datensatz überschaubar zu halten und nur signifikant regulierte Gene in die Clusteranalyse zu implementieren, wurden ausschließlich Gene analysiert, die mindestens in drei Zuständen reguliert vorlagen. Beim *k-means Clustering* muss die Anzahl der entstehenden Cluster im Voraus festgelegt werden. Der initiale Schritt des Algorithmus beginnt damit die Daten zufällig einem Zentroid zuzuordnen, was zur Folge hat, dass jeder Clusterlauf zu leicht anderen Ergebnissen führen kann (Stekel: Microarray Bioinformatics (2004)). Im konkreten Fall wurde die Strategie angewendet, das Clusterverfahren mit einer relativ großen Anzahl an Clustern zu beginnen und das Resultat hinsichtlich einander ähnlicher Cluster zu überprüfen. Waren diese vorhanden, so wurde die Clusteranzahl reduziert, bis sich alle Clusterverläufe signifikant voneinander unterschieden. Im vorliegenden Fall wurden fünf verschiedene Cluster gebildet. Neben recht komplexen mathematischen Methoden kann die Sinnhaftigkeit der Clusterergebnisse am biologischen Kontext verifiziert werden (Eisen *et al.* 2000)). Die Präsenz von Operons oder Genen mit ähnlicher oder überschneidender Funktion gibt gute Hinweise auf biologisch sinnvolle Clusterergebnisse. Die graphische Darstellung der Cluster ist in Abbildung 26 gezeigt, zu der Generierung dieser wurde für jeden Zeitpunkt der Mittelwert aus allen normierten *Mean-Signal*werten der Gene gebildet, welche an dem Cluster partizipieren.

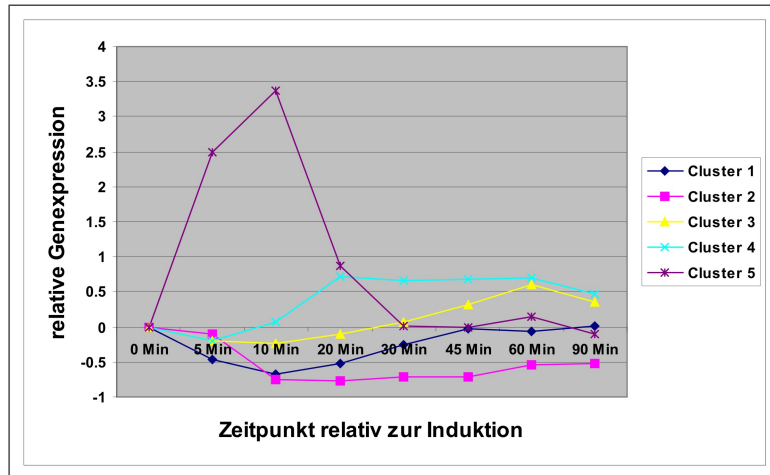


Abb. 26: *K-means* Clusteranalyse mit dem Datensatz der evaluierten transkriptionellen Vergleiche der induzierten Bioreaktorkultivierung. Gezeigt sind die Ergebnisse des Cluster-Algorithmus mit 5 verschiedenen Zentroiden.

Die im Folgenden angegebenen Zeitpunkte sind relativ zum Zeitpunkt der Induktion, welcher als 0 definiert wurde, zu betrachten. Cluster 1 zeigt von Zeitpunkt 0 an einen abfallenden Verlauf der Genexpression, welcher beim Zeitpunkt 10 Minuten in ein Minimum mündet. Anschließend steigt der Kurvenverlauf wieder, bis er 45 Minuten nach Arabino-sezugabe ungefähr den Ausgangswert erreicht hat und konstant bleibt. Für den Cluster 2 wird ein etwas anderer Verlauf beobachtet, der Expressionsverlauf bleibt bis 5 Minuten nach Induktion konstant und fällt dann ebenfalls ab, um ab 10 Minuten relativ zum Induktion-

zeitpunkt mehr oder minder konstant auf dem niedrigeren Niveau zu bleiben. Die Werte für Cluster 3 fallen leicht ab, erreichen 10 Minuten nach Induktion ein schwach ausgeprägtes Minimum und steigen danach an, bis sie 60 Minuten nach Induktion ein Maximum erreichen, um danach wieder leicht abzufallen. Cluster 4 zeigt einen gegenteiligen Verlauf zu Cluster 2, bis 10 Minuten nach Induktion bleibt die Genexpression in etwa konstant, um anschließend ab 20 Minuten nach Induktion auf ein Plateau anzusteigen. Für den Cluster 5 werden die mit Abstand deutlichsten Regulationen beobachtet. Der Graphenverlauf steigt stark an und erreicht ein Maximum 10 Minuten nach Induktion, um nachfolgend bis 30 Minuten nach Induktion wieder auf das Grundniveau abzufallen und konstant zu bleiben. Dieser äußerst interessante Verlauf deutet auf eine substantielle Zellantwort auf den Induktionsstimulus hin. Im Folgenden muss geklärt werden, ob dieser aufgrund der Arabinose-Supplementierung entsteht oder aufgrund der damit einhergehenden BRP-Expression. Zur Verifizierung der biologischen Sinnhaftigkeit und zur biologischen Interpretation der Clusterverläufe wird nachfolgend eine ausführliche Diskussion von Operons und Genen durchgeführt, welche in dem jeweiligen Cluster enthalten sind. Der erste Cluster umfasst 118 Gene. Unter diesen befinden sich eine große Zahl an Prophagengenen wie die Operons *mhpRT*, *ymfER* und die Gene B0247, B0280, B0551 sowie *relE*. *MhpR* und *relE* stammen beide vom Q λ -Phagen und sind in die Kälteschockantwort der Zelle involviert (Gotfredsen *et al.* 1998). Gleiches gilt für die Proteinprodukte der Gene *mspGB*, welche beide in einem Operon organisiert vorliegen. Wie Eingangs erwähnt, indiziert das Vorliegen von Operons ein biologisch sinnvolles Clusterergebnis. Das Genprodukt von B0551 ist insofern interessant, als es einen Transkriptions-Antiterminator darstellt. Für das ebenfalls in dem Cluster befindliche Operon *ybhNR* konnte bisher keine biologische Funktion festgestellt werden. Die Induktion von Prophagengenen kann durch die positive Regulation von *recA*, welches im Cluster 4 organisiert vorliegt, erklärt werden. *RecA* spielt als Aktivator der SOS-Antwort der Zelle die zentrale Rolle (Kuzimov *et al.* (1999)). Das Protein vermittelt die autokatalytische Degradation von drei Repressor-Proteinen, dem Repressor der SOS-Antwort (*LexA*), dem Lambda-Phagenrepressor und dem *UmuD*-Protein. Eine Degradation des Lambda-Phagenrepressors hat die Induktion verschiedener Prophagengene zur Folge, welches auch experimentell beobachtet wird. Ein Teil der experimentell beobachteten Regulationskaskade infolge der *RecA*-Aktivität ist in Abbildung 27 graphisch wiedergegeben.

Die Gene *dsbA*, *malE* und *ynhE* codieren für periplasmatische Chaperone. Eine bedeutende Gruppe in diesem Cluster bilden die Gene, deren Proteinprodukte in die LPS-Biosynthese und Colicin-Aufnahme involviert sind. Es handelt sich dabei um die Gene *wecC*, *cvpA*, *htrB*, *mdoH* und *kdsB*. Besonders *cvpA* ist in diesem Kontext interessant, da es für den Colicin V Zusammenbau und Freisetzung in der äußeren Membran fungiert (Fath *et al.*(1989)). Anmerkwürdig ist auch die Funktion des Genes *mhpR* als Transkriptionsaktivator für verschiedene Gene zur Fettsäurebiosynthese. Im Zusammenhang mit dem BRP-induzierten Einfluß auf die Membranintegritäten scheint dieses biologisch sinnvoll zu sein. Weiterhin wird eine große Zahl an Genen in diesem Cluster identifiziert, die in einer Verteidigungs- und SOS-Antwort der Zelle partizipieren. Bei den wichtigsten Vertretern dieser Klasse handelt es sich um die Gene *xthA*, *yebG* und *muiN*. Die Expression dieser Gene wird durch auftretende DNA-Schäden z.B. nach UV-Exposition induziert und stellt eine SOS-Antwort der Zelle dar (Lin *et al.* (1989)). Besonders bemerkenswert ist das Proteinprodukt von *yebG* aufgrund der Tatsache, dass es den *LexA* Transkriptionsrepressor reguliert. Dieser verhindert die Transkription jener Gene, die für die SOS-Antwort der Zelle verantwortlich sind. Das codierende Gen für *lexA* ist in der vorliegenden transkriptionellen Analyse stets

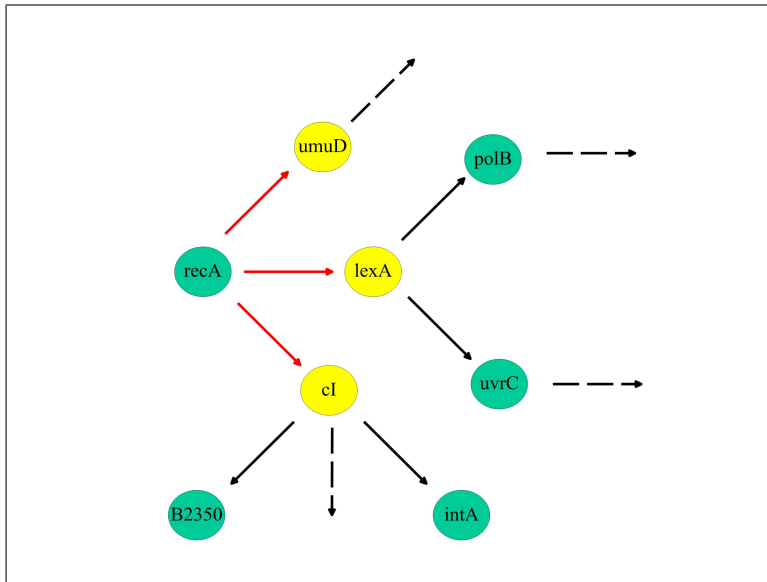


Abb. 27: Regulationsnetzwerk infolge der Erhöhung der *recA*-Transkription. RecA bewirkt die Degradation (rote Pfeile) der Proteine cI, UmuD und LexA, die ihrerseits wiederum Repressoren (gelbe Färbung) für verschiedene Operons und Gene darstellen. Aufgrund der Degradation von cI werden verschiedene Prophagengene induziert, im Schema ist beispielhaft die Regulation von B2350 und *intA* angegeben. Bedingt durch die Proteolyse von *lexA* werden die SOS-Operons *polAB* und *uvrABC* induziert, im Schema beispielhaft durch die experimentell beobachtete Regulation von *uvrC* und *polB* angegeben.

stark negativ reguliert, welches eindeutig ein Anschalten der zellulären SOS-Antwort belegt. Das Gen *lexA* seinerseits findet sich in Cluster 2 wieder. Die Clustergröße beträgt 100 Gene. Es wurde Eingangs erwähnt, dass sich in diesem Cluster die konstant negativ regulierten Gene befinden. Neben der auch experimentell beobachteten Regulation durch *yebG* wird *lexA* hauptsächlich durch das Genprodukt von *recA* reguliert. Dieses reagiert mit *lexA* unter einer autokatalytischen Abspaltung, wodurch die DNA-Bindekapazität von *lexA* signifikant abgeschwächt wird. Die SOS-Operons werden als Folge dessen für die Transkription durch die RNA-Polymerase frei zugänglich, welches auch experimentell anhand der Regulation von Genen wie *uvrC* und *polB* beobachtbar war. RecA ist in Cluster 4 organisiert, was aufgrund der gegenteiligen Expressionsverläufe von Cluster 4 und Cluster 2 biologisch sehr sinnvoll erscheint. Neben einigen Genen, die in den Aminosäure- und Energiemetabolismus der Zelle involviert sind (*glt*, *folA*, *entF* oder *cstC*), finden sich überraschenderweise die Gene *oppF*, *wcaJ*, *rffG* und *slt* in diesem Cluster wieder. *WcaJ* spielt in der Ausbildung der Polysaccharidkapsel eine wichtige Rolle (Grant *et al.* (1969)), *oppF* und *slt* codieren für Proteinprodukte, die bei dem Recycling von Zellwandbestandteilen wichtige Funktionen bekleiden (Höltje *et al.* (1975)).

Im Cluster 3 sind die Gene *araEFG* zusammen lokalisiert. Der biologische Sinn des Clusteregebnisses für das hochaffine Arabinoseaufnahmesystem ist hier augenscheinlich. Das Strukturgen eines bislang nur aufgrund von Sequenzhomologien vorhergesagten Zuckertransporters *ytfR* wird ebenfalls unter diesen Genen gefunden. Auch in vorangegangenen transkriptionellen Analysen wurde eine Regulation desselben festgestellt. Zusammen mit den evidenten Clusterergebnissen kann diese theoretische Vorhersage experimentell untermauert

werden. Weitere in die Zuckeraufnahme beteiligte Gene wie *uhpT*, *rhaD* und *ptsO* wurden innerhalb dieses Clusters identifiziert. Die Gene *amiB* und *mreD* sind ebenfalls in diesem Cluster vorhanden und codieren für Produkte, die für die Peptidoglykan-Beschaffenheit von essentieller Bedeutung sind (Wachi *et al.*(1989)). Das Proteinprodukt von *mreB* ist für die stäbchenförmige Gestalt von *E. coli* verantwortlich, eine Überexpression des codierenden Genes inhibiert die Zellteilung. *AmiB* fungiert als Sec-Substrat und ist bei der Teilung des Murein Septums beteiligt (Wachi *et al.*(1989)). Weiterhin zu Cluster 2 gehörig ist das Gen *soxS*, welches für einen Transkriptionsaktivator der Gene codiert, die als Folge eines oxidativen Stresses induziert werden. Das Gen ist zusätzlich in der bakteriellen Biofilmbildung involviert und wird in vorangegangenen Analysen nach Induktion stets reguliert vorgefunden. Die hohen prozentualen Anteile an regulierten Genen, die in Zusammenhang mit der oxidativen Stressantwort stehen, wurden auch durch die Multifunanalyse der Stressantwort belegt. Ebenfalls aufgrund von oxidativen Stress wird das Gen *yrfI* induziert, das für ein Chaperon codiert, welches desweiteren auch bei einem Hitzeschock exprimiert wird. In der Hitzeschockantwort der Zelle sind ebenfalls das geclusterte Operon *ibpAB* und das Gen *rbfA* involviert. Ein weiteres Hitzeschock-Chaperon wird von dem Gen *yabH* codiert. Neben *DnaJ* fungiert es als Cochaperon mit dem allgemeinen Hitzeschock-Chaperon *DnaK*. Aufgrund der erwiesenermaßen guten Eignung von *DnaK* für Coexpressionsstrategien (Perez-Rodriguez *et al.* (2007)) mit dem Ziel einer erhöhten Proteinsekretion stellt auch dieses Chaperon ein interessantes Ziel dar. Das Vorliegen des größeren Operons *rfaHIPZ* verifiziert wiederum die biologische Sinnhaftigkeit des Clusterresultats. Die Genprodukte dieses Operons besitzen Schlüsselpositionen bei der Synthese des Lipopolysaccharids der äußeren Membran. *RfaH* codiert für einen Antiterminator des Operons, welches für die LPS-Biosynthese, Zusammenbau und Export benötigt wird (Bailey *et al.* (1996)). Hierbei handelt es sich daher um ein potentiell Ziel für eine Coexpressionsstrategie, um der Zelle eine erhöhte Produktion von LPS-Komponenten zu ermöglichen. Somit ließe sich der schädigenden Wirkung der BRP-Expression begegnen. In diesem Kontext sinnvoll scheint auch die Identifizierung der Gene *yhbS*, *sanA*, *smpA*, *yefWQ*, *accA* und *B0834*, deren Proteinprodukte ähnliche Funktionen ausüben. Besonders die Regulation von *accA*, dessen Produkt am ersten Schritt der Fettsäurebiosynthese beteiligt ist (Li *et al.* (1992)), gibt Hinweise auf ein BRP-induziertes Gegensteuern der Zelle zur Wahrung der Membranintegrität. In der Diskussion der Gene von Cluster 2 wurden über das Anschalten der allgemeinen SOS-Antwort der Zelle berichtet. Auch im hier betrachteten Cluster 2 finden sich Vertreter der SOS-Gene wieder, es handelt sich dabei um *uvrB* und *polB*. Beide Genprodukte fungieren in der DNA-Reparatur, *PolB* ist dabei die prozessierende Polymerase. Zusätzlich sind in Cluster 3 die Gene *sbmA*, *hcaT*, *sanA*, und *ygeA* lokalisiert, deren Expression Resistenz gegen verschiedene Cytotoxine verleiht. Die Regulation des Genes *osmB* fand konsistent über fast alle evaluierten Vergleiche vor/nach Induktion, auch bei vorherigen Analysen statt und offenbart die osmoprotektive Zellantwort. Ebenfalls konsistent mit dem bisher formulierten geht die Regulation der Operons *rcsAC*, *galUP* und des Genes *yabH* einher. Das RCS-Operon stellt das Kontrollelement der Gene dar, welche für die Ausbildung einer Kapsel aus *Colanic Acid* und der bakteriellen Biofilmbildung verantwortlich sind.

Im Cluster 4 werden aus sekretorischer Sicht interessante Gene identifiziert, welche für die Funktionalität des Peptidoglykans und des LPS der äußeren Membran wichtig sind. Es sind dies die Gene *lpxA*, *glmS*, *lplA*, *dniR* und *pgpB*. Auf die besondere Bedeutung von *recA* zur Initialisierung der SOS-Antwort in der Wechselwirkung mit *lexA* wurde bereits hingewiesen. Besonders bedeutend für eine biologische Verifikation der Clusterergeb-

nisse ist die Identifizierung des Phospholipase A codierenden Genes *pdlA*. *PdlA* bewirkt die BRP-induzierte partielle Degradation der äußeren Membran und die daraus resultierende semispezifische Freisetzung des periplasmatischen Inhaltes (Dekker *et al.* (1999)). Aus sekretorischer Hinsicht stellt die Identifizierung eine Positivkontrolle dar. Gene die im gleichen Cluster organisiert sind, können potentiell in den Mechanismus der Proteinsekretion eine Rolle spielen. Ein Beispiel ist das periplasmatische Chaperon *LplA*, welches die Translokation der Lipoproteine von der äußeren auf die innere Membran vermittelt und somit auch für die Freisetzung rekombinanter Proteine eine Rolle spielen könnte.

Der graphische Verlauf des Clusters 5 weist das signifikanteste Profil auf. Er beinhaltet Gene, deren Expression kurz nach der Induktion stark zunimmt und anschließend wieder auf das Grundniveau sinkt. Das in nahezu allen Vergleichen zum Nullpunkt am stärksten positive regulierte Gen *eaeH* ist in diesem Cluster organisiert. Die Lokalisierung seines Genprodukts in der äußeren Membran deutet auf einen möglichen Zusammenhang mit der BRP-Expression hin. Interessant ist auch die Existenz des Genes *imp* in dem hier betrachteten Cluster. Hinter dem Terminus *imp* steht die Abkürzung *increased membrane permeability*, was eine Beteiligung an der BRP-induzierten Freisetzung periplasmatischer Proteine impliziert (Wu *et al.* (2006)). Ähnlich bemerkenswert ist das Vorliegen des Genes *lepB*, das für die Signalpeptidase codiert, welche die Signalsequenz von ins Periplasma translokierten Sekretionssubstraten proteolytisch abspaltet (Economou *et al.* (1999)). Koinzidierend mit dem bisher Konstatierten sind in dem Cluster weitere Gene enthalten, die wichtige Funktionen in der Aufrechterhaltung der Membranintegritäten und des Peptidoglykans ausüben. Neben dem schon erwähnten Gen *eaeH* sind dieses die Gene *cutC*, *wzB*, *yabC*, *yaeD*, *rfaB* und *lipA*. Das Operon *mhpEF* codiert für essentielle Produkte der Fettsäurebiosynthese und weist ähnliche Funktionen zu den bisher angeführten Genen auf. Die Gene *wzB* und *yedG* sind zudem in der Ausbildung des extrazellulären K-Antigens und der Biofilmbildung maßgeblich beteiligt. Weiterhin erwähnenswert sind die am Cluster partizipierenden Gene *map*, *ortM*, *ftsK* und *cutC*. Ersteres codiert für ein wichtiges zelluläres Chaperon, wohingegen die Proteinprodukte der drei anderen Gene in die zelluläre SOS-Antwort und Detoxifizierung involviert sind.

Anhand der Clusterergebnisse konnten verschiedene signifikante Expressionsverläufe (z.B. Cluster 5) identifiziert werden und die in dem Cluster organisierten Gene in einen funktionellen Zusammenhang gestellt werden. Aufgrund von äquivalenten Funktionen der Clustergene und der Existenz von Operons innerhalb eines Clusters wurde die Analyse biologisch verifiziert.

In den bisherigen Transkriptom-Analysen (Abschnitte 4.4.1, 4.4.2, 4.5.1 und 4.5.2) wurden drei Proben miteinander verglichen, welche drei verschiedenen Zeitpunkte (0, 30 und 60 Minuten relativ zum Zeitpunkt der Induktion) der Bioreaktorkultivierung repräsentierten. Anhand der Ergebnisse der evaluierten, feingliedrigeren differentiellen Genexpressionsanalyse, welche in diesem Abschnitt präsentiert wurden, konnte gezeigt werden, dass schon 5 Minuten nach Induktion eine substantielle Genregulation stattfindet. Zudem identifizierte die Clusteranalyse einen Expressionsverlauf, welcher ein signifikantes Maximum 10 Minuten nach Induktion aufweist. 30 Minuten nach Induktion wird keine Regulation dieser Gene mehr festgestellt, so dass deren differentielle Genexpression in vorangehenden Analysen nicht erfasst wurde. Dies manifestiert sich auch an der relativ geringen Anzahl an regulierten Genen des Vergleichs 30 gegen 0 Minuten gegenüber den Vergleichen früherer Zeitpunkte mit dem Referenzzustand vor Induktion. Um die generierten Ergebnisse mit denen vorheriger transkriptioneller Analysen (Abschnitte 4.4.1, 4.4.2, 4.5.1 und 4.5.2) vergleichen zu können,

wird im Folgenden ein separater Vergleich des zellulären Transkriptom vor sowie 30 und 60 Minuten nach Induktion präsentiert.

4.5.5 Transkriptioneller Vergleich der Zeitpunkte 30 und 60 Minuten nach Induktion mit dem präinduktiven Zustand

Um die analysierten Daten mit denen vorangehender Analysen zu vergleichen, wurde das Transkriptom der Zustände 30 und 60 Minuten nach Induktion mit dem gewählten Referenzzustand vor Induktion verglichen, um so eine äquivalente Evaluation zu ermöglichen.

Vergleich 30 vs. 0 Min

In einem ersten Vergleich des Transkriptom 30 Minuten nach Induktion gegen den präinduktiven Zustand wurde die Regulation des Operons *yagBFNS* beobachtet. Zusammen mit dem Ergebnis der positiven Regulation der Gene *araFGE* kann die biologische Sinnhaftigkeit der Analyse verifiziert werden. Die Gene *oppF*, *dniR*, *mreD* und *glmS* codieren für Komponenten, die für die Peptidoglykan-Integrität von großer Bedeutung sind und werden im evaluierten Vergleich differentiell exprimiert. Erstaunlicherweise werden demgegenüber einige Gene negativ reguliert, deren Funktion in der Biosynthese des Lipopolysaccharids und von Oberflächenantigenen der äußeren Membran liegt. Konkret sind dieses die Gene *htrB*, *B1998*, *wcaJ*, *yraJ*, *slp* und *rffG*. Anzumerken ist die positive Regulation des Genes der cytoplasmatischen Lon-Protease. Deren differentielle Expression kann durch die Saturierung der Sec-Translokase und der damit verbundenen Akkumulierung von Sec-Substraten im Cytoplasma erklärt werden. Verschiedene Autoren berichten über äquivalente Beobachtungen (Mergulaho et al. 2005, Collins *et al.* (2004)). Interessant anzumerken ist die positive Regulation von den Genen *grpE* und *ppiA*, welche für Chaperone codieren, die in die Faltung der Sec-Substrate zur Generierung eines transportkompetenten Zustandes involviert sein könnten. Die BRP-Expression führt sehr wahrscheinlich zu einer Aktivierung der SOS-Antwort der Zelle. Dies manifestiert sich in der konsistent stark negativen Regulation des Genes *lexA* nach dem Induktionsstimulus. Das Strukturgen *lexA* codiert, wie bereits erwähnt, für den Repressor der Operons der zellulären SOS-Antwort. Die Regulation der SOS-Antwort wird von der Zelle nicht nur auf transkriptioneller Ebene durch Reprimierung der Expression von *lexA* durchgeführt, sondern auch auf Proteinebene durch Aktivierung der *lexA*-Degradation mittels *RecA*. Das Gen *recA* ist nach Induktion konstant positiv reguliert und codiert für ein Bindeprotein des *lexA*-Repressors, welches im Komplex die autokatalytische Proteolyse von freien *LexA*-Molekülen als Coenzym bewirkt. Damit sind eindeutige Hinweise für eine BRP-induzierte SOS-Antwort von *E. coli* gegeben. Deren Auswirkungen lassen sich auch experimentell bestätigen wie z.B. durch die positive Regulation des Genes *uvrC* und *polB*. Das Proteinprodukt von *uvrC* ist an der Beseitigung von DNA-Schäden, wie sie z.B. nach einer UV-Exposition auftreten können, beteiligt, während *PolB* mechanistisch die prozessierende Polymerase für die DNA-Reparatur darstellt (Lin *et al.* (1989)).

Vergleich 60 vs. 0 Min

60 Minuten nach Induktion sind gegenüber dem Referenzzustand vor Induktion viele Operons reguliert, die direkt in der LPS-Biosynthese oder der des Peptidoglykans involviert sind. Als Positivkontrolle kann erneut die beobachtete Regulation der Gene *araHGFEJ* gewertet werden, welche für die Komponenten des ATP-abhängigen (*araHGF*) und pH-Gradienten vermittelten (*araE*) Arabinoseaufnahmesystems codieren. Das Gencluster *murADE* und die Gene *mrcA*, *amiB*, *mltR*, *ftsI* und *anpG* codieren sämtlich für Komponenten bzw. Enzyme,

die in der Peptidoglykan-Beschaffenheit eine zentrale Rolle spielen. Das Gen *mltR* ist insofern besonders interessant, da sein Proteinprodukt eine lokale Öffnung der Netzstruktur des Peptidoglykans initiiert und so an einer BRP-induzierte Freisetzung periplasmatischer Proteine mechanistisch beteiligt sein könnte (Dijkstra *et al.* (1996), Kraft *et al.* (1998)). Eine vergleichbare Diskussion wurde schon anhand des Analogons *mltC* geführt (Abschnitt 4.4.1). Ebenso interessant ist die Regulation von *anpG*, sein Genprodukt fungiert in dem Recycling von Zellwandbestandteilen. Eine Erklärung würde die Zellwandschädigende BRP-Expression liefern. Sehr augenscheinlich ist die Regulation des großen Operons *rfaBIPSZ*, welches direkt in die Lipid A-Biosynthese, der Hauptkomponente der äußeren Membran, involviert ist (Bailey *et al.* (1996)). Evidentes gilt für das Operon *lpxAB*, dessen Proteinprodukte in dem gleichen Pathway ihre Funktion aufweisen. Das Gen *eaeH* gehört in allen Vergleiche zu dem am stärksten positiv regulierten Genen und die Lokalisierung seines Genproduktes in der äußeren Membran indiziert eindeutig einen Zusammenhang mit der BRP-Expression. Die verschiedenen Genprodukte des Operons *ycdJPQS* sind ebenfalls in der äußeren Membran angesiedelt und stellen zudem Schlüsselkomponenten bei der Ausbildung bakterieller Biofilme dar (Ren *et al.* (2004), Wang *et al.* (2004)). In Koinzidenz dazu ist auch das codierende Gen eines Aktivators der Biofilmbildung (*rscC*) positiv reguliert. Sowohl aus sekretorischer Sicht als auch für das mechanistische Verständnis der BRP-Expression interessant ist die Regulation der Gene *tolAQ*. Die Proteine, für welche beide Gene codieren, besitzen zentrale Funktionen bei der Colicin-Freisetzung und sind Bestandteile des Haemolysin-Pfades zur TypI-Proteinsekretion in *E. coli* (Blight *et al.* (1994)). Ein Teil des Phagenschockoperons mit den Genen *pspCDE* wird ebenfalls als reguliert detektiert. Aufgrund der RecA-vermittelten Autoprotolyse des Lambda-Repressors kann die Regulation des Psp-Operons mechanistisch erklärt werden. Die Gene *lipA*, *yogH*, *ytfm*, *yrbI* und *lplA* codieren für Lipoproteine oder Produkte, welche für die Membranintegrität der äußeren Membran entscheidend sind. Zwei für eine Verifizierung der generierten Ergebnisse sehr wichtige regulierte Gene sind *lepA* und *pldA*. Das Gen *lepA* codiert für die Signalpeptidase, welche die Substrate nach der Translokation ins Periplasma proteolytisch von ihrer Signalsequenz trennt (Economou *et al.* (1999)). Im Kontext saturierter Translokationskapazitäten wird die Regulation verständlich. Ein weiteres eindrucksvolles Argument für die biologische Sinnhaftigkeit der Analyseergebnisse wird durch die beobachtete Regulation des Genes *pldA* geliefert. Die Phospholipase A dimerisiert und permeabilisiert die äußere Membran durch eine partielle Hydrolyse der Membranbestandteile. Durch diesen Effekt erfolgt die semispezifische Freisetzung des Colicins und Bestandteilen des Periplasmas (Dekker *et al.* (1999)). Eine weitere Verifizierung der biologischen Sinnhaftigkeit kann durch die differentielle Expression von *rpoH* gegeben werden. Der Sigmafaktor wird bei einem Hitzeschock vermehrt gebildet und sorgt für die vermehrte Transkription von jenen Genen, die in die Hitzeschockantwort der Zelle involviert sind (Zhao *et al.* (2005)). Ein Beispiel für solche Hitzeschock-induzierbaren Gene ist das Operon *ibpAB*, welches auch experimentell im Rahmen dieses Vergleiches als reguliert vorgefunden wird. Die Gene *yeiI*, *feoB* und *febG* spielen im Eisenmetabolismus und der zellulären Stressanpassung eine Rolle und wurden bei weiteren transkriptionellen Analysen des gleichen Stammes als reguliert identifiziert. Gleiches gilt für die Gene *uvrC*, *deaD*, *clpB* und *rseC*. Die Regulation der Operons *fimDFG* und *phoBR* gibt weitere Hinweise auf ein biologisch sinnvolles Resultat.

4.5.6 Induktionsniveau 1000 mg/l; feingliedrige Kinetik

Zur Evaluierung einer weiteren kinetischen Genexpressionsanalyse des Stammes Ara 1655 pBADLppBRPphoAbla wurde dieser im Bioreaktor kultiviert und beim Erreichen der mittleren exponentiellen Wachstumsphase mit 1000 mg/l L-Arabinose induziert. Die Probenahme erfolgte -10, 0, 2, 5, 10, 20, 30, 45 und 60 Minuten relativ zum Zeitpunkt der Induktion. Für die durchgeführten transkriptionellen Analysen geschah die weitere Probenprozessierung analog zu den vorherigen Experimenten.

Die beiden Proben, welche vor der Induktion dem Bioreaktor entnommen wurden, fungierten als Referenzzustand, auf dem sich die differentielle Expression der übrigen Proben bezog. Zur Verifikation der biologischen Sinnhaftigkeit der generierten Ergebnisse wurde die differentielle Genexpression einiger *housekeeping*-Gene näher untersucht. Obwohl die Verwendung von *housekeeping*-Genen zur Standardisierung kontrovers diskutiert wird, da deren Expression unter Einfluß verschiedenen Stimuli variieren kann, erlaubt die Verwendung mehrerer zu diesem Zweck gebräuchlicher Gene dennoch eine verlässliche Normalisierung. Fehlende oder geringe Unterschiede im detektierten Expressionsniveau unterschiedlicher *housekeeping*-Gene verschiedenener biologischer Proben geben zudem deutliche Hinweise auf eine gute Vergleichbarkeit zwischen diesen Proben. Um die Vergleichbarkeit der verschiedenen, auf die Microarray applizierten cDNA-Proben zu demonstrieren, wurden die *logratios* verschiedener Gene, welche in der Literatur für Normalisierungen über *housekeeping*-Gene verwendet werden, für die verschieden evaluierten Vergleiche aufgetragen. Das Ergebnis kann in Abbildung 28 betrachtet werden.

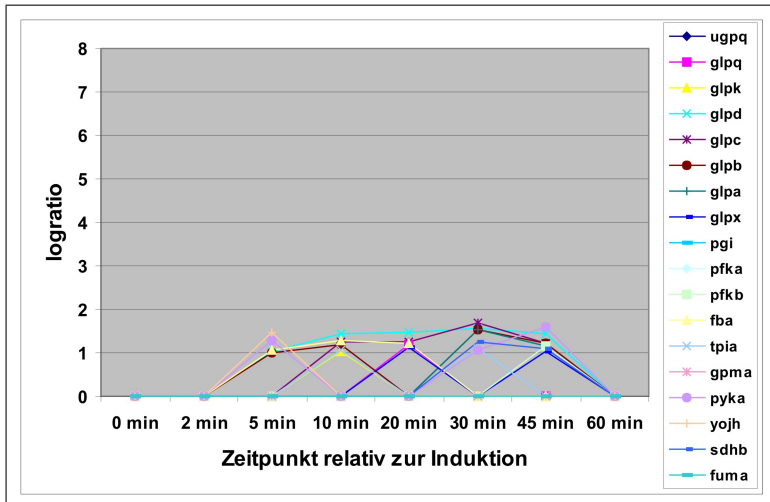


Abb. 28: *Logratios* verschiedener *housekeeping*-Gene für die evaluierten Vergleiche. Evident ist das im Vergleich zu dem Arabinosegenen araFGHE geringe Regulationsniveau (Gezeigt in Abbildung 29).

Wie aus der Abbildung 28 ersichtlich, variiert die Expression der meisten Gene dieser Auswahl nur im geringen Maße, wie die geringen *logratios* von 0 bis 1 indizieren. Manche Gene wie *glpACD*, die in den Glycerin-Metabolismus der Zelle involviert sind, weisen jedoch auch größere *logratios* bis hin zu Werten von 1,7 auf. Dies zeigt eindeutig, dass die Verwendung von nur einem oder wenigen *housekeeping*-Genen als Normalisierungsstandard nicht ausreichend ist, da deren Expression nicht zwangsläufig bei den untersuchten Bedingungen

konstant bleibt. Microarray-Applikationen eignen sich aufgrund ihres Parallelisierungsgrades sehr gut dafür, eine Auswahl an geeigneten Kandidaten für diese Anwendung zu finden. Die Verwendung mehrerer *housekeeping*-Gene generiert zumeist reproduzierbare Daten und findet immer mehr Eingang in andere Genexpressions-Analysentechniken wie z.B. quantitative *real-time* PCR oder *Northern Blotting* (Hugget *et al.* (2005), Wilhelm *et al.* (2003)). Verglichen mit den *logratios* für die als Positivkontrolle fungierenden Arabinosegene araEFGH fallen die Werte der *housekeeping*-Gene gering aus. Die Abbildung 29 zeigt die *logratios* der Gene araBCEFGHJ für die verschiedenen evaluierten Vergleiche.

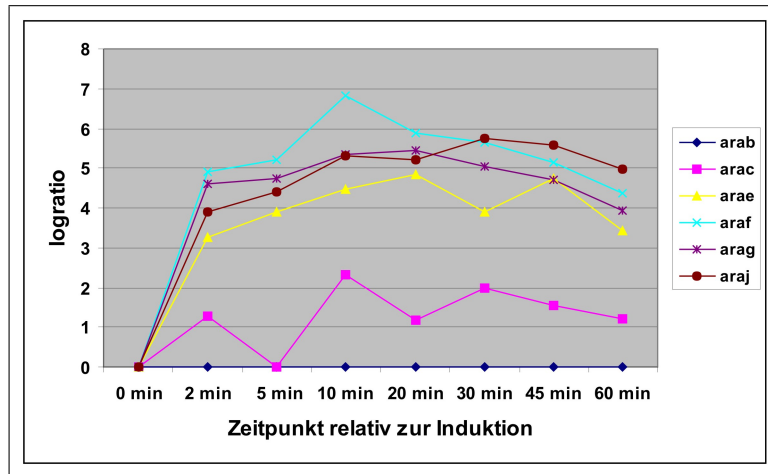


Abb. 29: *Logratios* verschiedener Gene welche in den Arabinose-Import (araFGHE) involviert sind. Zur Veranschaulichung des *knock-outs* des AraBAD-Operons ist stellvertretend das *logratio* von araB aufgetragen. Das Regulatorgen des Arabinoseoperons araC ist ebenfalls gezeigt. Dessen Transkription wird von einem separaten Promoter kontrolliert und korreliert daher nicht mit den übrigen gezeigten Genen.

Deutlich ersichtlich ist das signifikante Ansteigen der Expression von araEFGH 2 Minuten nach Induktion, das Expressionsniveau dieser Gene erhöht sich nachfolgend nur noch wenig, insgesamt wird ein hyperbolischer Verlauf beobachtet. Das Gen araC codiert für den Regulator des Arabinoseoperons und zeigt einen schwankenden Verlauf. Anhand des Genes araB lässt sich der *knock-out* des Arabinoseoperons dokumentieren. Zusammengefasst verlief die Verifikation der generierten Ergebnisse anhand der Negativkontrolle der *housekeeping*-Gene und der Positivkontrolle der Arabinosetransportgene erfolgreich, so dass im Anschluß eine eingehendere Diskussion des Datensatzes erfolgt.

2 Minuten nach Induktion werden 551 Gene differentiell exprimiert, von diesen sind 533 positiv und 18 negativ reguliert. Die regulierten Gene repräsentieren 12,35 Prozent aller in *E. coli* existierenden ORFs. Für die Zeitpunkte 5, 10, 20, 30, 45 und 60 Minuten nach Induktion werden 2257 (2229/28); 1571 (1558/13); 2006 (2000/6); 1805 (1794/11); 1601 (1589/12) und 733 (645/88) regulierte Gene beobachtet. In den Klammern sind jeweils die Anzahlen an positiv regulierten (erste Position) und negativ regulierten Genen (zweite Position) angegeben. Die regulierten Gene entsprechen damit 50,61; 35,22; 44,98; 40,47; 35,90 und 16,43 Prozent aller in *E. coli* K12 existenten Gene. Graphisch ist die Anzahl der regulierten Gene für jeden evaluierten Vergleich in der Abbildung 30 wiedergegeben.

Der Hauptteil der differentiell exprimierten Gene ist positiv reguliert. Auf zellulärer Ebene tritt eine substantielle Regulation bedingt durch den Induktionsstimulus auf, wie die

4.5 Analyse der extrazellulären, Sec-spezifischen Sekretion

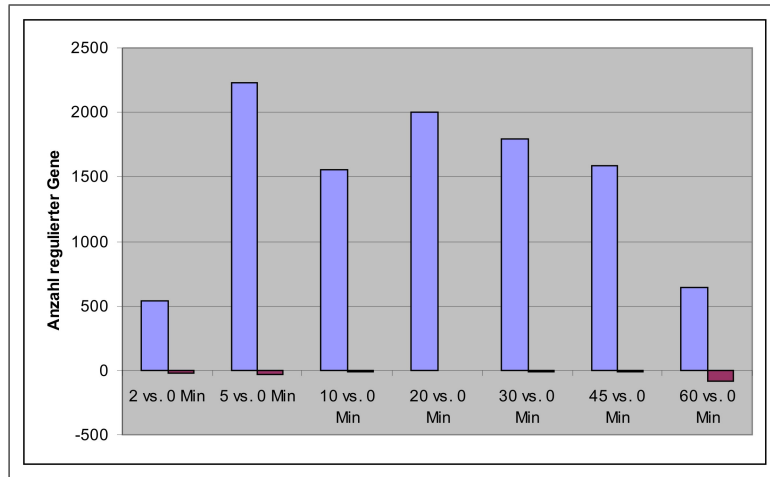


Abb. 30: Anzahl der negativen und positiven regulierten Gene für die verschiedenen evaluierten Vergleiche.

hohen prozentualen Werte an regulierten Genen suggerieren. Die Anzahl an negativ regulierten Genen steigt nach 60 Minuten relativ zum Zeitpunkt der Arabinosezugabe im Vergleich zu den vorhergehenden Zeitpunkten dramatisch an. Hierin könnte sich der von Steghuis *et al.* (1995) beschriebene Effekt der BRP-Expression auf die Proteinbiosynthese widerspiegeln. Anhand der Multifun-Datenbank wurden die regulierten Gene biologischen Funktionen zugeordnet. Abbildung 31 zeigt den sehr deutlichen Anstieg der negativ regulierten Genen für den Zeitpunkt 60 Minuten nach Induktion an. Die meisten negativ regulierten Gene sind der Multifun-Funktionsklasse Metabolismus zugeordnet.

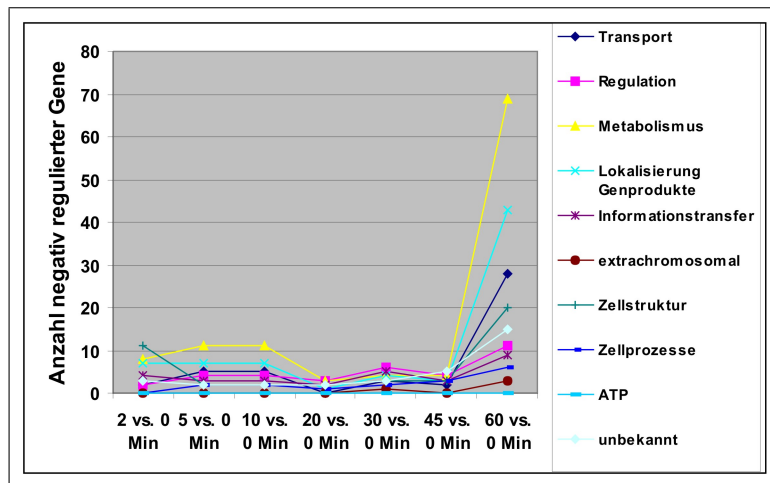


Abb. 31: Anzahl der negativ regulierten Gene der jeweiligen Multifun-Hauptklassen für die verschiedenen evaluierten Vergleiche.

4.5.7 Cluster-Analyse

Der Datensatz wurde einer *k-means* basierten Clusteranalyse unterzogen, bei der der Korrelationskoeffizient als Ähnlichkeitsmaß fungierte. Die Daten von Genen, welche in mindes-

tens drei Vergleichen reguliert waren, wurden dem Algorithmus unterzogen. Die Evaluierung verschiedener Clusteranzahlen ergab bei sieben verschiedenen Clustern keine signifikanten Ähnlichkeiten hinsichtlich der Expressionsverläufe und wurde so als Startparameter für den Algorithmus gewählt. Durch Mittelwertbildung aus den normierten Signalwerten der Gene, aus denen sich der Cluster zusammensetzt, wurde der Expressionsverlauf der verschiedenen Cluster graphisch in Abbildung 32 dargestellt.

Cluster 1 weist einen konstant ansteigenden Expressionsverlauf auf, welcher in ein Maximum 30 Minuten nach Induktion mündet. Anschließend findet ein signifikanter Rückgang der Expression auf das Anfangsniveau statt. Drei Gene, deren Proteinprodukte die Schlüssel-funktion in der Freisetzung von Colicinen einnehmen, sind in diesem Cluster zusammengefasst. Es handelt sich um die Gene *fepA*, *purR* und *yciD*. Zusätzlich werden in diesem Cluster die Gene *secDFG* und *yajC* des Sec-Pathways identifiziert, welches weitere Hinweise auf eine BRP-induzierte Regulation gibt. Der Komplex aus *secDF* und *yajC* assistiert dem zentralen, membrangebundenen SecYEG-Translokator bei der Interaktion mit SecA und dem zu translokierenden Präprotein (Econoumu *et al.* (1999), Nouwen *et al.* (2002)). Die drei codierenden Gene sind genetisch in einem Operon organisiert. Infolge des hohen Induktionsniveaus von *L*-Arabinose und der damit einhergehenden Akkumulierung von dem Sec-Substrat BRP ist eine Saturierung der Exportkapazität wahrscheinlich. Insofern stellt die positive Regulation von Genen, die für Komponenten der Translokase codieren eine biologisch sinnvolle zelluläre Reaktion dar. Ein analoges Expressionsprofil wird für jene Operons beobachtet, welche für die NADH-Ubiquinon Oxidoreduktase I (*nuoHGKI*), Glycerin-3-Phosphat-Dehydrogenase (*glpABC*) und ATP-Synthetasen (*atpDIP*) codieren. Die Ubiquinon-Oxidoreduktase I fungiert als Protonenpumpe, welche den pH-Gradienten über der Cytoplasmamembran aufrecht erhält (Stolpe *et al.* (2004)). Diese protonenmotorische Kraft stellt neben der ATP-Hydrolyse die Triebkraft der Sec-vermittelten Translokation dar. In Koinzidenz zu dieser Beobachtung wird die Regulation von Genen beobachtet, welche für die ATP-Synthetase codieren. Die Glycerin-3-Phosphat-Dehydrogenase ist ein Membran-assoziiertes Enzymkomplex, welcher Glycerin zu Dihydroxyaceton-Phosphat oxidiert und dabei zwei Protonen auf membrangebundenes Ubiquinon überträgt (Lin *et al.* (1976)). Diese Reaktion könnte daher ebenfalls dazu dienen, den Protonengradient über der Membran aufrecht zu erhalten. Zusätzlich sind optimale Konzentrationen an Glycerin-3-Phosphat essentiell für die Biosynthese von Phospholipiden, welche die Hauptbestandteile beider Zellmembranen darstellen. Die Anwesenheit der periplasmatischen Chaperone *dsbACG* innerhalb dieses Clusters ist ebenfalls erwähnenswert. Der Verlust des periplasmatischen Inhaltes, bedingt durch die BRP-Aktivität, könnte ihre Regulation erklären. Zudem könnte aufgrund der periplasmatisch akkumulierenden Sec-Proteine ein erhöhter Bedarf an Chaperonen bestehen, um diesen zur ihrer korrekten Konformation zu verhelfen.

Der Verlauf des Clusters 2 ähnelt dem des Clusters 1, weist allerdings generell ein etwas höheres Expressionsniveau und ein Maximum 45 Minuten nach Induktion auf. Die sukzessiv ansteigende Expression gibt Hinweise auf einen möglichen Zusammenhang mit der Akkumulierung des BRPs. Tatsächlich werden die Gene der beiden Hitzeschockchaperone *DnaK* und *DnaJ* in diesem Cluster identifiziert. Auf die Analogie zwischen der BRP-Expression und einem Hitzeschock durch Änderung der Membranzusammensetzung und der Induktion der Chaperonexpression wurde bereits hingewiesen. Die beiden Chaperone assistieren den Polypeptidketten, einen transportkompetenten Zustand einzunehmen. Sie sind ein zelluläres Detektionssystem für fehlgefaltete Proteine (Lamb *et al.* (2008)). Perez-Rodriguez *et al.* (2007) konnten durch Coexpression von *DnaK* eine signifikante Steigerung der Sekretions-

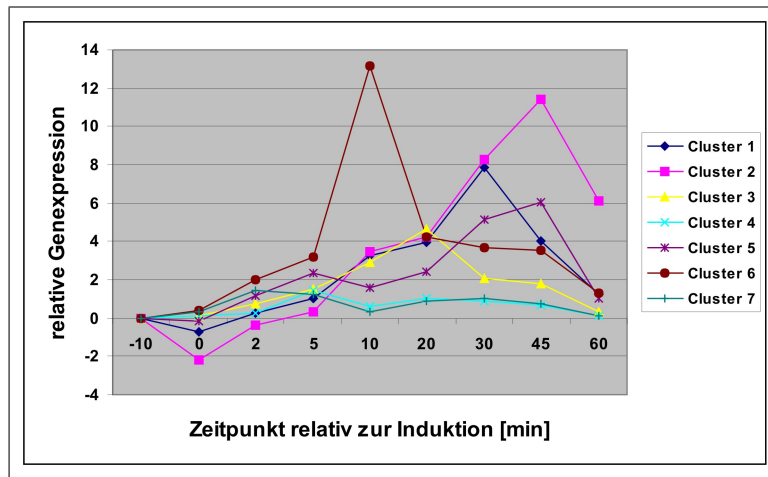


Abb. 32: *K-means* Clusteranalyse mit dem Datensatz der evaluierten transkriptionellen Vergleiche der induzierten Bioreaktorkultivierung. Gezeigt sind die Ergebnisse des Cluster-Algorithmus mit 7 verschiedenen Zentroiden.

effizienz nachweisen. Mit *dsbB* wird ein weiteres Gen innerhalb des Clusters vorgefunden, welches für ein Hitzschockchaperon codiert. Besonders interessant ist die Existenz von *rpoH* in diesem Cluster. Das Gen codiert für den Hitzschock-Sigmafaktor der RNA-Polymerase und verifiziert die beobachtete Regulation der eben genannten Hitzschockgene. Bedingt durch die partielle Degradation der äußeren Membran und dem damit einhergehenden hohen osmotischen Druck auf die Zelle wird die Regulation von *osmB* verständlich. Das Protein *osmB* ist ein Lipoprotein der äußeren Membran, dessen Expression osmotisch induzierbar ist. Seine Überexpression vermittelt der Zelle osmoprotektive Wirkung. Gene die im Cluster 3 organisiert sind, zeigen einen leicht ansteigenden Expressionsverlauf mit einem Maximum 20 Minuten nach dem Induktionsstimulus. Einige Operons wie das Flagellar-Regulon (*fliCPJT*) und das der Acetyltransferasen (*atoABCDE*) (Pauli *et al.* (1972)) werden innerhalb des Cluster vorgefunden. Die Acetyltransferasen besitzen eine zentrale Bedeutung bei der Metabolisierung von Fettsäuren, wie sie durch die BRP-induzierte Phospholipase-Aktivität in der äußeren Membran vermehrt entstehen könnten. Neben diesen werden einige Gene als zugehörig identifiziert, deren Proteinprodukte an der Aufnahme und Metabolisierung von Arabinose beteiligt sind. Der biologische Sinn ihrer positiven Regulation ist augenscheinlich.

Cluster 4 zeigt im Vergleich zu den übrigen Expressionsverläufen ein nicht sehr signifikantes Maximum 5 Minuten nach Induktion. Interessanterweise befinden sich alle Gene des *E. coli* Phagenschock-Operons in diesem Cluster. Die Induktion dieser Gene wird durch verschiedene Arten von Membranstress und durch stark oxidative Bedingungen sowie hohe Osmolaritäten hervorgerufen (Brissette *et al.* (1991)). Die Auswirkungen der BRP-Expression zeigen deutliche Analogien zu denen einer Phageninfektion, weshalb die Induktion dieses Operons ein weiteres Indiz für den biologischen Sinn der generierten Daten darstellt. Im vorherigen Abschnitt 4.5.1 wurde eine mechanistische Erklärung für die Regulation dieses Operons gegeben. Wie das Expressionsprofil von Cluster 2 weist jenes von Cluster 5 ein Maximum 45 Minuten nach der *L*-Arabinosezugabe auf. Innerhalb dieses Clusters wird das *Rfa*-Operon vorgefunden, welches essentiell für die Biosynthese und den Zusammenbau des LPS der äußeren Membran ist. Diese Beobachtung ist im Kontext der dramatischen Auswirkungen der BRP-Expression auf die Membranintegritäten erklärbar. Zudem sind die

Colicin-Gene *tolQ* und *fhuA* innerhalb des Clusters organisiert. TolQ ist ein inneres Membranprotein und verleiht der exprimierenden Zelle durch Bewahrung der Membranintegrität Resistenz gegen Colicine der Gruppe A und filamentöse Phagen. Die gesteigerte Expression unter dem Stimulus der BRP-Akkumulierung ist unmittelbar einsichtig. Die positive Regulation von *tolQ* stellt eine Immunitätsantwort der Zelle dar, um die dramatischen Folgen der BRP-Aktivität zu überleben (Blight *et al.* (1994)).

Ein ausgeprägt signifikantes Maximum weist der Cluster 6 10 Minuten nach Induktion auf. Vor und nach diesem Zeitpunkt fällt die Genexpression dramatisch ab. Einige Arabinoseimportgene befinden sich in diesem Cluster, was die Vermutung nahelegt, dass es sich um eine Arabinose-induzierte Auswirkung auf die Genexpression handelt. Von besonderem Interesse ist die Existenz eines Teil des Operons, welches für die 50S ribosomale Untereinheit der RNA-Polymerase codiert. Die positive Regulation dieser Gene könnte dazu fungieren, der allgemeinen Tendenz zur Repression der Aminosäurebiosynthese durch das BRP entgegenzuwirken. Weiterhin sind viele Gene in dem Cluster enthalten, welche Strukturgene für Chaperone darstellen, als Beispiel seien *maP*, *ecpD*, *rmF*, *mglB*, *htgA* und *ccmC* genannt. SecY, ein Teil des zentralen Translokationstunnels, wird ebenfalls diesem Cluster zugeordnet. Aufgrund des unauffälligen Expressionsprofils von Cluster 7 wurde erwartet, dass in diesem aus sekretorischer Sicht keine interessanten Gene enthalten sind. Die experimentelle Beobachtung stützt diese Annahme.

4.5.8 Zusammenfassung

Anhand der transkriptionellen Analyse des Sekretionssystemes pBADLppBRPBlaPhoA in dem *E. coli*-Stamm Ara 1655 konnte eine substantielle transkriptionelle Regulation aufgrund der Arabinosesupplementierung und der damit einhergehenden BRP-Expression nachgewiesen werden. Diese umfasste die positive Regulation von Genen, welche in die LPS- und Peptidoglykan-Biosynthese, die Biofilmbildung, und die Eisenaufnahme der Zelle involviert waren. Eine Regulationskaskade zur Initiierung der zellulären SOS-Antwort wurde beobachtet. Zusätzlich wurde die differentielle Expression von verschiedenen Chaperonen, Osmoprotektanden und Hitzeschockgenen nachgewiesen. Die Sinnhaftigkeit von deren Regulation wurde anhand der signifikanten Auswirkung der BRP-Expression auf die Membranintegritäten diskutiert und anhand von Positiv- und Negativkontrollen wie dem Arabinoseaufnahmesystem *araFGH* verifiziert. Eine bislang nur auf Sequenzhomologien fußende These der Funktion der Genprodukte von *yftQRST* als Zuckertransporter konnte anhand der Microarray-Daten experimentell verifiziert werden. Zusätzlich wurde die Regulation von einigen Prophagengenen beobachtet, von denen besonders die differentielle Expression von B1550 Hinweise auf einen Zusammenhang mit der Proteinsekretion und eine zelluläre Reaktion zur Erhöhung der Sekretionskapazitäten gibt.

Aus all diesen Beobachtungen lassen sich verschiedene Optimierungstrategien mit dem Ziel einer erhöhten Sekretionseffizienz modifizierter *E. coli* Stämme generieren. Die Coexpression verschiedener Chaperone hat sich im Kontext erhöhter Sekretionskapazitäten häufig als erfolgreich erwiesen. Die beobachtete Regulation von Genen wie *torD*, *fkpA* und *groEL* gibt daher Hinweise auf mögliche Ziele für Coexpressionsstrategien. Eine vermehrte Regulation Biofilm-involvierter Gene tritt besonders unter hohen Expressionsraten der BRPs auf. Die Ausbildung bakterieller Biofilme sollte im Rahmen von Sekretionsstrategien unbedingt vermieden werden, da die Einkapselung der Bakterien in einer Sekretionsprävention resultiert. Ob dieses allerdings im gerührten Bioreaktor in signifikantem Maße auftritt, ist unklar.

Die Verwendung von Mutantenstämmen mit defekten Oberflächeadhäsions-Funktionalitäten wäre in diesem Kontext zu erwähnen. Eine weitere, vielversprechende Zielstrategie wäre die Coexpression von essentiellen Lipoproteinen, welche helfen, die Membranintegrität zu gewährleisten. Beim Lipid A der äußeren Membran handelt es sich zum Beispiel um den hydrophoben Anker des LPS, welcher damit eine essentielle Funktion für die Membranintegrität besitzt. Gleiches gilt auch für die Coexpression von Osmoprotektanden und Hitzeschockgenen, welche ebenfalls eine zielführende Strategie darstellen könnte. Der oxidative Stress, welcher durch den Verlust von Cytochromen und Flavinen der inneren Membran entstehen könnte, stellt ebenfalls einen Ansatzpunkt für Optimierungsstrategien mit dem Ziel, diesen zu minimieren, dar. Die beobachtete Regulation des Phagenschockoperons könnte sowohl BRP- als auch sekretionsinduziert, durch Saturierung der Sec-Translokase erfolgen (Brissette *et al.* (1991)). Es wurde beobachtet, dass die Expression von *pspA* infolge übersättigter Exportkapazitäten signifikant erhöht wird. Durch Coexpressionstrategien mit *pspA* erfolgte der Nachweis, dass man sich diesen Umstand zur Steigerung der sekretierten Proteinmenge zu Nutze machen kann (De Lisa *et al.* (2004)). *PspA* und andere Prophagengene stellen daher ebenfalls vielversprechende Ziele für eine Optimierung der Proteinsekretion dar. Besonders die bereits ausführlich diskutierte positive Regulation von B1550, welche in einer Erhöhung der Membranfluidität mündet, könnte für Sekretionsstrategien ein erfolgsversprechendes Konzept sein.

4.6 Analyse der periplasmatischen, Sec-spezifischen Sekretion

Die Interpretation der Ergebnisse der bisherigen Transkriptom-Analysen wurde durch den Umstand erschwert, dass die beobachtete differentielle Genexpression das additive Resultat aus der Arabinosesupplementierung und der BRP-Expression darstellte. Um die beiden Effekte voneinander abzugrenzen, wurde eine transkriptionelle Analyse von Proben einer induzierten Bioreaktorkultivierung des Stammes Ara 1655 pBADLppBRPphoAbla realisiert. Dieser Stamm ist mit Ausnahme des fehlenden LppBRPs vollständig identisch mit dem Stamm Ara pBADLppBRPphoAbla, welcher in den vorherigen Transkriptomanalysen zum Einsatz kam. Die beobachteten regulatorischen Effekte können so allein der Arabinoszugabe zugeschrieben werden. Durch eine Subtraktion dieser Effekte von denen, welche in induzierten Bioreaktorkultivierungen des BRP-codierenden Stammes beobachtet wurden, sollte es möglich sein, die BRP-induzierten Effekte zu extrahieren.

Schematisch ist die Sec-spezifische Sekretion der beiden konstitutiv exprimierten Reporterproteine Alkalische Phosphatase und β -Lactamase in Abbildung 33 gezeigt.

Für die Gesamtzahlen an regulierten Genen ergaben sich folgende Werte: 2 Minuten nach Induktion waren im Vergleich zum präinduktiven Zustand 56 Gene reguliert, wovon 22 negativ und 34 positiv differentiell exprimiert waren. Dies entspricht einem Anteil von 1,26 Prozent, bezogen auf alle existierenden 4460 ORFs in *E. coli* K12. Für die analogen transkriptionellen Vergleiche (5, 10, 20, 30, 45, 60 Min) mit dem Zustand vor der Induktion wird die Regulation von 122 (57/65); 136 (67/59); 386 (295/91); 392 (205/188); 424 (259/165); 659 (416/243) Genen beobachtet. In den Klammern sind die Anzahlen an negativ (erste Position) und positiv (zweite Position) regulierten Genen angegeben. Relativ auf das *E. coli* Genom bezogen entspricht dies einer differentiellen Expression von 2,73 (5 Min); 3,05 (10 Min); 8,65 (20 Min); 8,79 (30 Min); 9,51 (45 Min) und 14,78 (60 Min) Prozent. Graphisch wird diese Verteilung in der Abbildung 34 wiedergegeben.

Es ist ersichtlich, dass bezogen auf den Referenzzustand vor Induktion, die Anzahl an

4.6 Analyse der periplasmatischen, Sec-spezifischen Sekretion

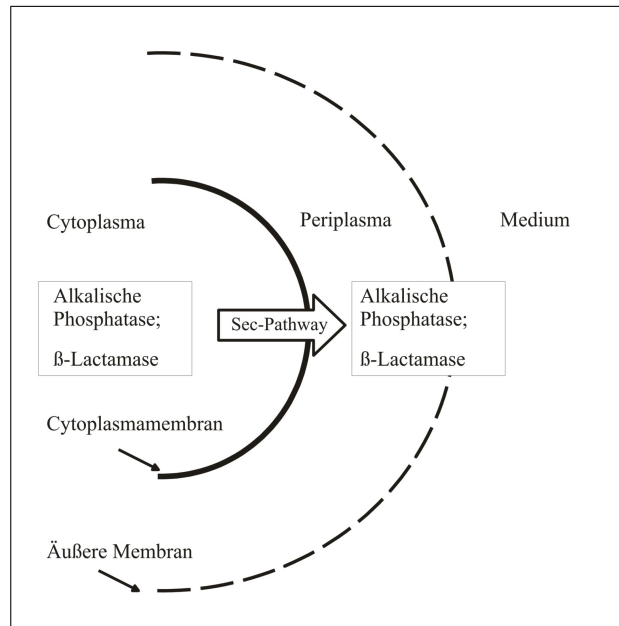


Abb. 33: Schema der Sekretion der Reporterproteine Alkalische Phosphatase und β -Lactamase des Stammes Ara1655 pBADphoAbla. Die beiden Reporterproteine werden mittels des Sec-Pathways in das Periplasma von *E. coli* translokiert.

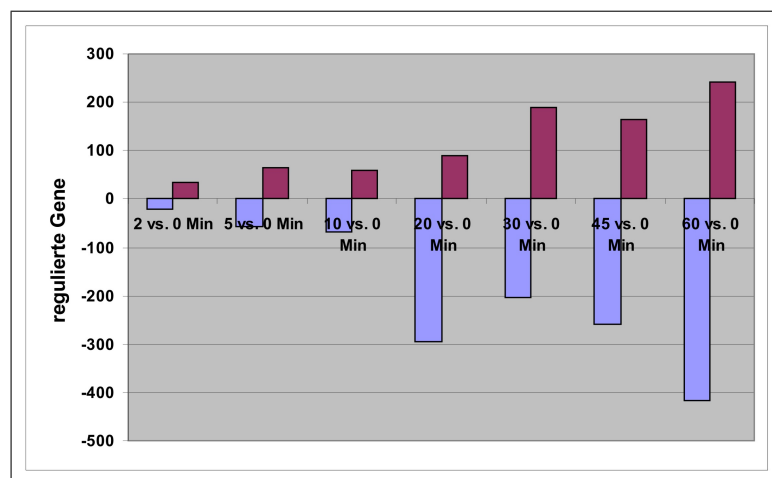


Abb. 34: Anzahl der regulierten Gene (separat für positiv und negativ regulierte Gene durchgeführt) für die verschiedenen evaluierten Vergleiche mit dem präinduktiven Zustand.

regulierten Genen sukzessiv mit wachsender Zeitdauer nach Induktion zunimmt. Dabei sind in den meisten Vergleichen zum präinduktiven Zustand mehr Gene negativ reguliert. Folglich ergibt sich ein differentes Bild verglichen mit den transkriptionellen Analysen anhand BRP-codierender Stämme. Die Gesamtanzahl an regulierten Genen für die meisten Vergleiche ist signifikant geringer. Daneben trat bei den Transkriptomanalysen der BRP-codierenden Stämme nach einer gewissen Zeitdauer nach Induktion eine substantielle, fast ausschließlich positive Regulation auf, welche im vorliegenden Fall nicht beobachtet wird.

Besonders signifikant lassen sich die transkriptionellen Unterschiede zwischen BRP-codierenden und BRP-defizienten Stämmen anhand der Anzahlen an regulierten Genen in den Multifun-Klassen Verteidigung, Äußere Membran, Peptidoglykan und Extrachromosomal demonstrieren. Dies sind die zellulären Kompartimente bzw. die zellulären Prozesse, auf die sich, der Erwartung nach, die BRP-Expression besonders stark auswirken sollte. Zum Vergleich fungierten die Daten aus einer Transkriptomanalyse des vergleichbaren BRP-exprimierenden Stammes.

Exemplarisch ist dieser Sachverhalt für positiv regulierte Gene, die in der Zellverteidigung involviert sind, in Abbildung 35 gezeigt. Es werden für die Kultivierung mit dem BRP-exprimierenden Stamm signifikant mehr regulierte Gene beobachtet als für den äquivalenten Stamm ohne BRP-Expression. Besonders deutlich fällt das Ergebnis für die frühen und späten Vergleiche aus. Interessant zu beobachten ist, dass die späten Vergleiche der Kultivierung des BRP-codierenden Stammes ausschließlich positiv regulierte, Stress-involvierte Gene aufweisen (hier nicht gezeigt). Dies gibt Hinweise auf eine BRP-induzierte Regulation dieser Gene, zumal die späten Zeitpunkte gut mit den Zeitpunkten für das Auftreten phänotypischer Effekte der BRP-Expression, wie sie in der Literatur (Steghuis *et al.*(1995), Van der Waal *et al.* (1991)) belegt sind, übereinstimmen.

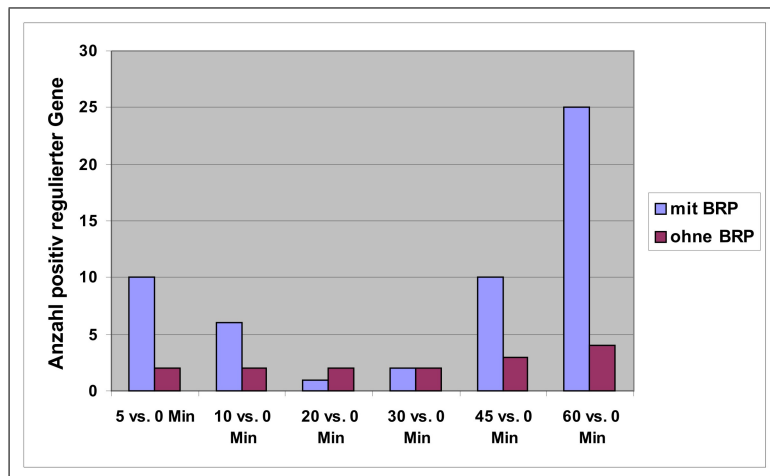


Abb. 35: Vergleich der positiv regulierten Gene, welche in die zelluläre Abwehrantwort involviert sind, zwischen den BRP-codierenden und den BRP-defizienten Stamm.

Die regulierten Gene der vorangehend evaluierten transkriptionellen Analysen welche Hinweise auf einen Zusammenhang mit der BRP-Expression liefern, wurden hinsichtlich ihrer Regulation in der Transkriptomanalyse des BRP-defizienten Stammes überprüft. Übereinstimmende Regulationen geben Hinweise auch einen Arabinose-induzierten Effekt, wohingegen eine fehlende Regulation oder eine Regulation in die entgegengesetzte Richtung

Tabelle 4: Vergleich von 30 in mindestens drei Vergleichen als reguliert identifizierten Genen der Transkriptomanalysen der BRP-codierenden Stämme. Gene, die in der vergleichbaren Analyse des BRP-defizienten Stammes Ara 1655 pBAD_{phoA}bla nicht reguliert sind, werden in der Tabelle mit einem - bezeichnet. Übereinstimmende Regulation ist durch ein P und entgegengesetzte Regulation durch ein N vermerkt.

Gen	30 vs. 0 Min	60 vs. 0 Min	Gen	30 vs. 0 Min	60 vs. 0 Min
imp	-	-	pspB	-	-
spy	-	-	pspC	-	-
lolA	-	-	ygeA	P	P
hemM	N	N	ytfR	P	P
ycdP	-	-	ytfQ	P	P
ycdQ	-	-	tolA	-	-
yecI	-	-	tolQ	-	-
soxS	-	-	mrdB	-	-
bglx	-	-	amiB	-	N
vacJ	-	N	rfaA	-	-
uvrC	-	-	rfaB	-	-
uvrD	-	-	rfaF	-	-
tonB	-	N	mltC	-	-
osmB	-	P	recA	-	-
pspA	-	-	rpoH	-	-

einen BRP-induzierten Effekt impliziert. Für 30 Gene ist dieser Vergleich in der Tabelle 4 durchgeführt. Es erfolgte eine Evaluierung der Vergleiche der postinduktiven Zustände 30 und 60 Minuten mit dem präinduktiven Zustand, da diese Vergleiche für alle durchgeführten Transkriptomanalysen analysiert wurden.

Wie aus der Tabelle 4 ersichtlich ist, sind fast alle Gene, die in der Transkriptomanalyse des BRP-codierenden Stammes als reguliert identifiziert wurden, bei der äquivalenten Evaluierung des BRP-defizienten Stammes nicht oder in eine andere Richtung reguliert. Dieses Ergebnis verifiziert die zuvor getroffenen Aussagen über einen BRP-induzierten Effekt. Lediglich die Gene *osmB* und *ygeA* sowie *ytfR* und *ytfQ* waren übereinstimmend reguliert. Diese Gene sind daher sehr wahrscheinlich als Folge der Arabinose-Supplementierung reguliert. Im Falle von *ytfR* und *ytfQ* wurde dieses auch erwartet, beide codieren für ein hypothetisches zelluläres Zuckerimportsystem und sollten daher nicht als Folge der BRP-Expression reguliert sein. Somit wäre der letzte experimentelle Beweis ihrer Funktionalität als Saccharid-Transportsystem erbracht.

4.6.1 Clusteranalyse

Anhand der normierten und hintergrundbereinigten Signalwerte wurde eine *k-means* Clusteranalyse der verschiedenen evaluierten Vergleiche durchgeführt. Als Ähnlichkeitsmaß fungierte, analog zu der vorherigen Vorgehensweise, der Korrelationskoeffizient. Sieben verschiedene Cluster ergaben keine signifikante Überschneidung in den Clusterverläufen und wurden daher als Startwert für den Algorithmus verwendet. Die graphischen Verläufe der einzelnen

Cluster, generiert aus den Mittelwerten aller an dem jeweiligen Cluster partizipierenden Genen, kann in der Abbildung 36 eingesehen werden.

Die Genexpression jener 101 Gene, die zu Cluster 1 gehörig sind, fällt sukzessiv leicht ab. Unter den zugehörigen Genen codieren einige für Oberflächenantigene. So sind das Operon *rfbDB* und die Gene *dniR*, *ugd*, *rfaD*, *dsbA* und *kdtA* innerhalb des Clusters wiederzufinden. Daneben sind auch Gene in dem Cluster vorhanden, deren Proteinprodukte im Bereich des Peptidoglykan lokalisiert sind und auch bei transkriptionellen Analysen mit BRP-codierenden Stämmen als reguliert identifiziert wurden. Es handelt sich dabei um die Gene *imp*, *ybbF* und *amiB*. Allerdings wiesen die Regulationen dieser Gene nach Induktion in die entgegengesetzte Richtung, so dass ein BRP-induzierter Effekt wahrscheinlich ist. Weiterhin sind im dem Cluster Teile des Tryptophanoperons mit den Genen *trpADE*, des Argininoperons *argCEI* und des Cysteinoperons *cysUPH* zu finden. Das Vorliegen von Genen, die genetisch in einem Operon organisiert sind, deutet auf die Sinnhaftigkeit der generierten Clusterergebnisse hin. Die Gene *car*, *purR*, *vacJ*, *rpsI* und *rplI* wurden zumeist auch bei den Analysen der BRP-codierenden Stämme als differentiell exprimiert identifiziert. Die Richtung der Regulation weist bei den BRP-Stämmen mit Ausnahme von *vacJ* in die gleiche Richtung, so dass von einem Effekt ausgegangen werden kann, welcher aufgrund der Arbinose-Zugabe auftrat. Zusätzlich sind die Gene *livFH* und *accBC* dem Cluster zugehörig, *accB* und *accC* haben Funktionen in der Fettsäuren-Biosynthese.

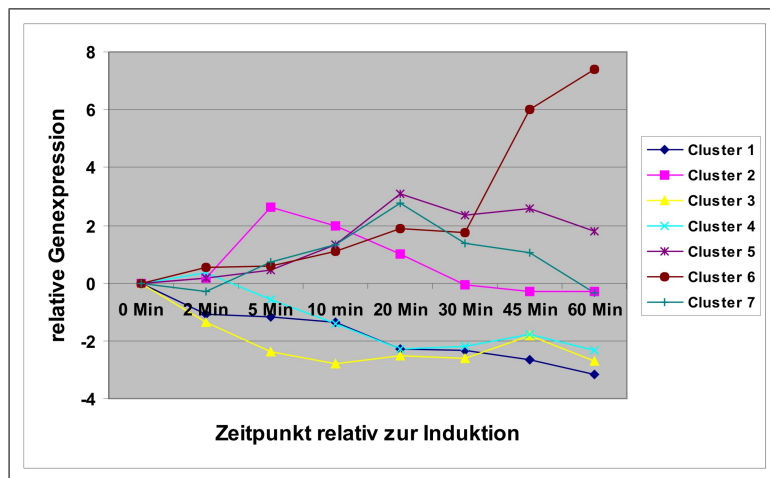


Abb. 36: *K-means* Clusteranalyse mit regulierten Genen der Transkriptomanalyse anhand des Stammes Ara 1655 pBADphoAbla.

Der Cluster 2 beinhaltet nur die geringe Anzahl von 14 Genen, der Verlauf weist ein charakteristisches Maximum 5 Minuten nach Induktion auf. Aus sekretorischer Sicht sind wenig interessante Gene darin enthalten, lediglich die beiden Gene *yiiO* und *hipB* sind in diesem Zusammenhang erwähnenswert. Beide codieren für Proteinprodukte, die an der Reaktion auf extracytoplasmatischen Stress (Cpx-System) beteiligt sind. *YiiO* ist ein negativer Regulator der Cpx-Antwort, die beobachtete positive Regulation unterdrückt also die zelluläre Cpx-Antwort (Otto *et al.* (2002)).

In dem Cluster 3 sind 45 Gene enthalten, die Genexpression der partizipierenden Gene sinkt zuerst bis zum Zeitpunkt 10 Minuten nach Induktion ab und bleibt dann in etwa konstant auf diesem Niveau. 45 Minuten relativ zum Zeitpunkt der Induktion wird ein wenig ausgeprägtes lokales Maximum beobachtet. Bemerkenswert ist die Existenz der Gene

rlpP, tonB und rfaK in diesem Cluster, deren Proteinprodukte alle eine Funktion in der LPS-Biosynthese einnehmen. Besonders die Bedeutung von tonB als energetischer Mediator der Translokation von Bestandteilen der äußeren Membran wurde hingewiesen. Die Regulation dieser Gene wurde in BRP-codierenden Stämmen vornehmlich positiv detektiert und weist auf einen BRP-induzierten Effekt hin.

Für den vierten Cluster wird ein ähnlicher Verlauf zu dem Cluster 3 beobachtet, allerdings sinkt die Genexpression erst etwas später, 5 Minuten nach Induktion, signifikant ab. Die 89 Gene, aus denen der Cluster besteht, bekleiden unterschiedliche zelluläre Funktionen. Analog zu dem bei Cluster 1 Angemerktem werden auch innerhalb dieses Clusters Teile von Operons wiedergefunden welche in der Aminosäurebiosynthese involviert sind. Im konkreten Fall sind dies z.B. ilvIH und leuI für die Leucin-Biosynthese. Interessanterweise wird auch ein Teil der Sec-Translokase, mit dem Genen secA und secF in diesem Cluster organisiert vorgefunden. Zusätzlich sind weitere Gene existent, deren Proteinprodukte in der LPS-Biosynthese eine wichtige Rolle spielen. Konkret zu nennen sind lpxD, ylcB, slt und int. Generell wurde anhand der Clusterergebnisse dieser Transkriptomanalyse festgestellt, dass die Regulation von in die LPS- und Peptidoglykan-Biosynthese involvierten Genen zumeist in eine andere Richtung weist, als dies bei BRP-codierenden Stämmen der Fall war. Dies stellt einen deutlichen Hinweis dafür dar, dass es sich bei der positiven Regulation dieser Gene um einen von der BRP-Expression verursachten Effekt handelt. In Koinzidenz zu dem eben Angemerkten lagen die Gene galFP, sulA und evgA, deren Regulation ebenfalls bei vielen transkriptionellen Vergleichen BRP-codierender Stämme nach Induktion festgestellt wurde, in diesem Cluster organisiert vor.

31 Gene bilden den Cluster 5, ihre Expressionsprofile steigen sukzessiv an, um 20 Minuten nach Induktion ein Maximum zu erreichen und danach wieder leicht abzufallen. Neben dem für ein Lipoprotein codierenden Gen yabC ist ein Teil des Operons zur Metabolisierung von Galactose, galKT, in diesem Cluster befindlich. Ebenso wurde die Regulation des Sekretionsgenes secG und der Chaperon-codierenden Gene clpX und fkpA beobachtet. Die beobachtete vermehrte positive Regulation von verschiedenen Chaperonen tritt also auch als Effekt der Arabinosezugabe auf. Die Gene leuA, lrP und ilvM, welche ebenfalls in dem diskutierten Cluster zu finden sind, haben ihre Bedeutung bei der Biosynthese der Aminosäure Leucin. Eine weitere interessante Beobachtung stellt die Regulation der Gene mglAB und atpAG dar. Die Proteine MglA und MglB fungieren im Kohlenhydratmetabolismus, während atpGA für die ATP-Bereitstellung eine Schlüsselposition einnimmt. Die große Anzahl an vermehrt gebildeten Zuckeraufnahmesystemen von Typus der ABC-bindenden Kasette sowie die konstitutive Expression und Translokation der Sec-Substrate PhoA und Bla bedarf der Bereitstellung größerer Mengen ATP (Economou *et al.* (1999)).

Für den Cluster 6, welcher sich aus 57 Genen zusammensetzt, wird ein besonders signifikanter Verlauf beobachtet. Die Genexpression der enthaltenden Gene steigt zuerst leicht an, um dann 30 Minuten nach Induktion sprunghaft anzuwachsen. Nahezu alle darin enthaltenden Gene codieren für Proteine, die direkt in den Kohlenhydratmetabolismus involviert sind. Als Positivkontrolle für das Clusterresultat kann das simultane Auftreten der Gene/Operons araFGH und araEJ in diesem Cluster fungieren. Das gemeinsame *Clustering* der Gene für das hoch- und niedrigaffine Arabinoseaufnahmesystem stellt offensichtlich ein biologisch sinnvolles Resultat dar. Weiterhin in den Kohlenhydratmetabolismus (speziell im Trehalosemetabolismus) involviert sind die Gene treA, sfa, B2016, mglC, csrA und glgS sowie das Operon ytfRQ. Das letztgenannte Operon wurde schon im vorherigen Abschnitt eingehend diskutiert, bisher ist die Funktion als Zuckertransporter nur aufgrund von

Sequenzhomologien vorhergesagt worden. Auch ein Teil des Galaktoseoperons *galeUS* ist diesem Cluster zugeordnet. Weiterhin auffällig ist die hohe Anzahl an Genen, welche in den Acetatmetabolismus involviert sind. Im konkreten Fall handelt es sich um die Gene *mdhH*, *accB*, *atoB* und *glgS*. Die Regulation des Genes *rpoE* erscheint im Vergleich zu den anderen Genen des Clusters, welche vornehmlich ihre Funktion im Kohlenhydratmetabolismus ausüben, ungewöhnlich. Das Gen codiert für den Sigmafaktor, welcher als Effekt auf einen Hitzeschock oder Membranstress induziert wird. *RpoE* initiiert die Transkription des Chaperons *FkpA*, welches ebenfalls unter den regulierten Genen zu finden ist und in Cluster 5 organisiert vorliegt. Die Zugabe größerer Mengen *L*-Arabinose scheint somit Implikationen auf die periplasmatischen Proteine zu haben, so dass die beobachtete Regulation erklärbar wird.

Im letzten Cluster, welcher aus 85 Genen besteht, sind vornehmlich Gene lokalisiert, welche für Produkte codieren, die ihre Funktion in der Sulfat-Assimilierung haben oder ribosomale Proteine darstellen. Das Operon *cysABDW* sei als Beispiel für erstere Gruppe genannt, wohingegen die Operons *rplADQJ* und *rpsNJ* Beispiele für Letztere sind. Zusätzlich interessant ist die Zuordnung der Hitzeschockgene *ibpAB* und des Chaperon-codierenden Genes *hslV* zu diesem Cluster. Die positive Regulation der Gene *ibpAB* wurde einer BRP-induzierten Wirkung zugeschrieben, allerdings zeigt diese Beobachtung eindeutig, dass auch oder alleinig die Arabinosezugabe deren Induktion bewirkt.

4.6.2 Zusammenfassung

Anhand des Vergleiches BRP-codierender/BRP-defizienter Stamm konnten die Auswirkungen der BRP-Expression von denen der Arabinose-Supplementierung abgegrenzt werden. Es zeigte sich, dass die differentielle Expression der meisten Gene aufgrund der BRP-Expression auftrat. Besonders für die Funktionsklassen Stress, Verteidigung und Extrachromosomal wurden signifikant größere regulatorische Effekte bei den BRP-codierenden Stamm beobachtet. Eine durchgeführte *k-means* Clusteranalyse ergab zudem weniger explizite Expressionsverläufe für die verschiedenen Cluster des BRP-defizienten Stammes. Die Regulation von verschiedenen Zuckerimport-Komponenten und deren gemeinsame Organisation in einem Cluster wurde für die Verifikation der experimentellen Daten verwendet.

4.7 Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien

4.7.1 Einführung

Bedingt durch das immense Interesse an rekombinanten Proteinen im akademischen und industriellen Bereich wurden viele verschiedene Expressionssysteme entwickelt und zur routinemäßigen Applikation gebracht. Verschiedene Studien belegen, dass hohe konstitutive Expressionsraten des rekombinanten Proteins in den meisten Fällen das Wachstum und den Zellmetabolismus in signifikanter Weise beeinträchtigen und dadurch zum Absterben der Zellen führen (Lottspeich F.: Bioanalytik (1995)). Deshalb verwendet man für die Produktion rekombinanter Proteine fast ausschließlich induzierbare Expressionssysteme. Je nach vorliegendem Promoter-Operatorsystem wird die Genexpression durch Metabolitzugabe, Entfernen bestimmter Kohlenstoffquellen oder Temperaturänderungen induziert. Die basale Transkription, d.h. die Transkriptionsrate der stromabwärts von einem Promoter liegenden Gene ohne Induktionstimulus, unterscheidet sich für verschiedene Promotoren deutlich (Singer M.:

Gene und Genome (1992)). Für Genexpressionsstudien ist diese Tatsache von essentieller Bedeutung, da man explizit an der differentiellen Expression der Gene im Vergleich vor und nach der Bildung/Sekretion des rekombinanten Produktes interessiert ist. Findet schon vor der Induktion eine signifikante Expression statt, so vermag sich die Zelle regulatorisch auf diese Situation einzustellen. Aus diesem Grund sollte die basale Transkription möglichst gering und das Induktionsniveau regulierbar sein, um die limitierten Sekretionskapazitäten nicht zu überlasten (Simmons *et al.* (1996)). Das bislang beste Expressionssystem für diese Belange basiert auf dem Arabinosepromoter pBAD. Neben einer sehr geringen basalen Transkription ist die Induktionsstärke des Arabinosepromoters über einen weiten Bereich linear regulierbar, und es lassen sich hohe Transkriptionsraten der promoterregulierten Gene realisieren (Guzman *et al.* (1995)). Aus diesem Grunde wurde das Plasmid pBAD33, welches stromabwärts des Arabinosepromoters eine Multi-Klonierungsstelle aufweist, als Klonierungsvektor für die Etablierung eines Sekretionsreportersystems gewählt. Das 5432 bp große Plasmid codiert u.a. für das Gen der Chloramphenicol-Acetyltransferase (CAT), welches Resistenz gegen Chloramphenicol verleiht und als Selektionsmarker dient. Als Reporterproteine finden zumeist Enzyme Verwendung, da sich diese aufgrund ihrer spezifischen biologischen Aktivität leicht nachweisen und quantifizieren lassen. Daneben werden in jüngerer Vergangenheit im stetig steigendem Maße fluoreszierende Reporterproteine verwendet. Besonders für das grün fluoreszierende Protein (GFP) wurden seine herausragenden Eigenschaften als Reporterprotein in vielen Studien belegt und seine Entdeckung und Applikation mit dem Nobelpreis 2008 gewürdigt. Bei dem aus der Qualle *Aequorea victoria* stammenden GFP handelt es sich um ein monomeres, autofluoreszierendes und globuläres Protein von 26 kDa Größe. Neben einer sehr geringen Zytotoxizität, hohen Quantenausbeuten und damit einhergehenden hohen Fluoreszenzintensitäten behält das Protein auch bei Fusionierung an viele andere Proteine seine biologische Aktivität bei. Zur Ausbildung der Fluoreszenzeigenschaften durchläuft das Protein eine Reihe von posttranslationalen, intramolekularen Reaktionen. Für die Entstehung des Fluorophors ist die Cyclisierung und Autooxidation der Aminosäuren Ser65-Tyr66-Gly67 verantwortlich. Das Emissionsmaximum des reifen GFP liegt bei 509 nm bei einer UV-Anregungswellenlänge von 395 nm. Für seine Fluoreszenz werden keine exogenen Substrate oder Cofaktoren benötigt. Dadurch ist GFP ein einzigartiges Reporterprotein für Genexpressionsstudien, sowie für die Untersuchung von Proteinlokalisationen und Proteinkinetiken in pro- und eukaryontischen Systemen. Durch eine einfache Messung der Fluoreszenzintensität lässt sich die GFP-Menge direkt quantifizieren. Allerdings muss bei einer derartigen Quantifizierung kalkuliert werden, dass ein Teil des GFPs in Form sogenannter *Inclusion Bodies* anfallen kann, welche aufgrund ihrer Fehlfaltung keinen Beitrag zur Gesamtfluoreszenzintensität liefern. Die real exprimierte GFP-Menge liegt in diesen Fällen daher höher, als die experimentell ermittelte Fluoreszenzintensität suggeriert (Bongaerts *et al.* (2002)). Vom Wildtyp-GFP wurden durch Aminosäureaustausch und/oder Optimierung der Codon-Verwendung eine Reihe von spektralen Varianten mit höheren Fluoreszenzintensitäten, verschobenen Anregungs- und Emissionswellenlängen und veränderten Halbwertszeiten im lebenden Organismus generiert. Ein hauptsächliches Problem bei der Verwendung von Wildtyp-GFP ergibt sich aus der Tatsache, dass ein großer Teil des bei 37°C exprimierten Proteins nicht korrekt gefaltet und cyclisiert vorliegt. Durch multiplen Aminosäureaustausch mittels *DNA-Shuffling* wurden thermostabile Varianten geschaffen, welche unter diesen Bedingungen eine korrekte Faltung aufweisen und damit eine signifikant höhere Fluoreszenzintensität generieren (Cramer *et al.* (1995)). Die daraus hervorgegangenen Derivate wurden mit GFPmut1-3 bezeichnet, GFPmut2 wird im Rah-

men dieser Studie als Reporterprotein verwendet. Es beinhaltet die Aminosäureaustausche S65A, V68L und S72A und weist eine 33-fach stärkere Fluoreszenz als Wildtyp-GFP auf. Die mut-Derivate sind besonders wertvoll für *in vivo* Studien, da sie ideal für mikroskopische und durchflußzytometrische Analysen geeignet sind. Neben GFP wurden eine ganze Reihe weiterer fluoreszierender Proteine entdeckt und erfolgreich als Reportersysteme etabliert. Als Sekretionsreportersystem wurde GFP erstmals 2001 erfolgreich eingesetzt. Durch Fusionierung mit dem Signalpeptid der N-Trimethylaminoreduktase (*torA*) wurde GFP mittels des Tat-Pathways in das Periplasma von *E. coli* sezerniert (Thomas *et al.* (2001)). Da die korrekte Faltung essentiell für die Fluoreszenzeigenschaften des Proteins ist, muß die Sekretion obligat mittels des Tat-Pfades erfolgen, da dieser als einziger Sekretionspathway komplett gefaltete Proteine zu translokieren vermag.

4.7.2 Klonierung des Sekretionsreporterplasmids pBADtorAGFPSsrA

Die vielversprechenden Eigenschaften von GFP, die gute Nachweis- und Quantifizierungsmöglichkeit, sowie die Verfügbarkeit effizienter Sekretionsplasmide bedingten die Wahl dieses Reportersystems in der vorliegenden Studie. Die GFPmut2-Variante wurde freundlicherweise von Prof. De Lisa (New York, USA) in Form des Plasmides pETGFPmut2 zur Verfügung gestellt. Die bereits verifizierte Funktionalität der *torA*-Fusion als Sekretionssignal sollte auch im Rahmen dieser Versuche verwendet werden, um GFP ins Periplasma zu translokieren. Um mittels einer Fluoreszenzmessung nur die real sekretierte Menge am Reporterprotein GFP zu erfassen, wurde die GFPmut2-Variante C-terminal mit dem *SsrA-tag* fusioniert. Dieses Degradations-*tag* dirigiert Substrate zum cytoplasmatisch lokalisiertem Clp XP/AP-Proteasesystem (Flynn *et al.* (2001)). Wird durch die Wahl eines geeigneten Induktionsniveaus das Proteasesystem nicht überlastet, so generiert lediglich das durch Sekretion ins Periplasma gelangte GFP eine Fluoreszenzantwort (De Lisa *et al.* (2002)). Das Konstrukt, welches sich aus der C- und N-terminalen Fusion ergibt, ist schematisch in Abbildung 37 wiedergegeben.

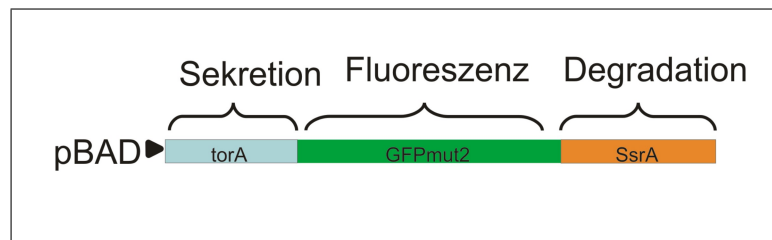


Abb. 37: Das Konstrukt *torAGFPSsrA* welches als Sekretionsreporter im Rahmen der vorliegenden Studie fungiert.

Zur Klonierung des GFP-Konstruktes wurden zwei Strategien entwickelt, von denen sich Strategie 2, welche auf einer initialen *blunt-end* Subklonierung beruht, erfolgreich umsetzen ließ. Beide Strategien werden in der Abbildung 38 präsentiert.

Theoretisch sollten Klonierungen über kohäsive Enden mit deutlich größerer Effizienz erfolgen als dies bei *blunt-end* Klonierungen der Fall ist. (Sambrook: Molecular Cloning (2003)). Insofern überrascht das Misslingen der Klonierungsstrategie 1. Allerdings belegen verschiedene Studien die potenziellen Schwierigkeiten die sich aus der Klonierung von PCR-Produkten über angefügte Restriktionsschnittstellen ergeben (Crowe *et al.* (1991), Barnes *et al.* (1992), Levis *et al.* (1995)).

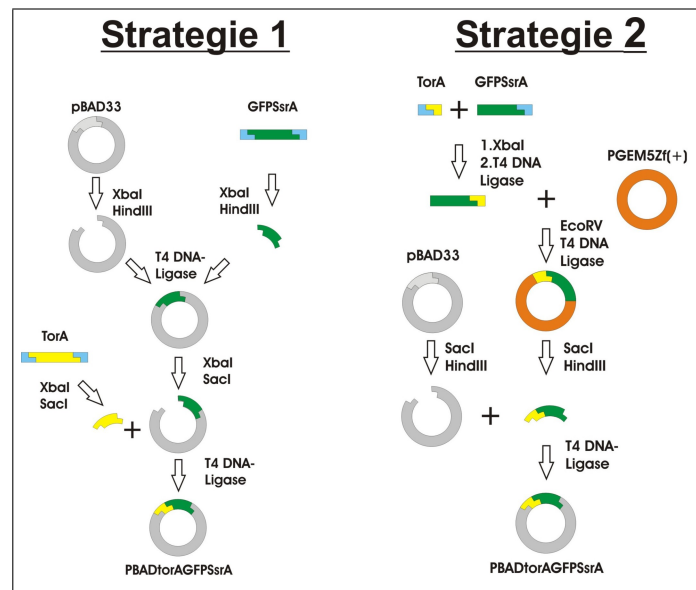


Abb. 38: Zwei Klonierungsstrategien zu Generierung des Sekretionsreporterplasmids pBADtorAGFPsSrA. Die auf der linken Seite gezeigte Strategie 1 bedient sich zweier Klonierungsschritte über kohäsive Enden. Die dazu notwendigen Restriktionsschnittstellen werden über die 5'-Enden der PCR-Primer in die Inserts eingefügt. In Strategie 2, welche auf der rechten Seite präsentiert wird, erfolgt eine initiale Ligation der beiden PCR-generierten Inserts torA und GFPsSrA sowie eine *blunt-end* Subklonierung in den Vektor pGEM5Zf(+). Das aus dem Vektor herausgeschnittene Konstrukt toraGFPsSrA wird final in den Zielvektor pBAD33 kloniert.

Generierung der Klonierungsinserts mittels PCR

Die Generierung der Klonierungsinserts erfolgte mittels PCR. Durch Anfügen der Restriktionsschnittstellen SacI und XbaI bzw. XbaI und HindIII an die 5'-Enden der Primer wurde die Möglichkeit einer direktionalen Klonierung in den Zielvektor geschaffen. Zusätzlich wurden die Primer um weitere 6 Basenpaare stromaufwärts der Restriktionsschnittstelle verlängert. Dieser zusätzliche Überhang ist notwendig, da viele Restriktionsendonukleasen terminal lokalisierte Zielsequenzen nicht zu prozessieren vermögen (Sambrook: Molecular Cloning (2003)).

Das C-terminale SsrA-tag vom GFP wurde an das 5'-Ende des reversen Primers zwischen dem genspezifischen Teil und der Restriktionsschnittstelle implementiert. Die *in silico* designeten Primer wurden mittels des Programmes MFold auf interne Sekundärstrukturen überprüft und mittels des Programmes BLAST auf ihre Spezifität getestet. Das GFP-Plasmid pETGFPmut2 fungierte als Template zur Amplifikation des Zielgenes, während die Signalsequenz von torA aus der cDNA einer *E. coli* Gesamt-RNA Präparation amplifiziert wurde.

Aufgrund des sehr langen 5'-Überhanges, bedingt durch das eingefügte SsrA-tag, ergaben sich anspruchsvolle Amplifikationsbedingungen. Eine Variation der *Annealing*-Temperatur in dem Bereich von 55°C bis 70°C generierte lediglich bei 62°C eine geringe Menge an Amplifikat, welche aber für *downstream*-Applikationen bei weitem nicht ausreichend war.

4.7 Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien

Daher wurden verschiedene Strategien zur Steigerung der Amplifikationseffizienz entwickelt.

1. Verwendung eines neuen *Forward* Primers mit einer längeren genspezifischen Sequenz
2. PCR mit einem zweistufigen Programm (Umkehrung einer *touch-down* PCR). In den ersten 10 Zyklen wurde eine niedrige *Annealing*-Temperatur gewählt, um initial genügend Template zu generieren. Dann wurde die *Annealing*-Temperatur für die weiteren 20 Zyklen erhöht, um der PCR die maximale Effizienz zu verschaffen
3. Zusatz verschiedener Additive als *Enhancer*

Als *Enhancer* wurden folgende Additive bzw. Kombinationen von Ihnen getestet:

- DMSO
- Tween 20
- DMSO + Tween 20
- Betain

Die Verwendung eines längeren *Forward* Primers sowie ein zweistufiges PCR-Programm generierten keine signifikanten Mengen an PCR-Produkt. Letztendlich zielführend war die Verwendung einer Kombination aus DMSO und Tween 20 als *Enhancer*, wie in Abbildung 39 ersichtlich ist. Die Wahl der geeigneten *Annealing*-Temperatur scheint dabei eine untergeordnete Rolle zu spielen.

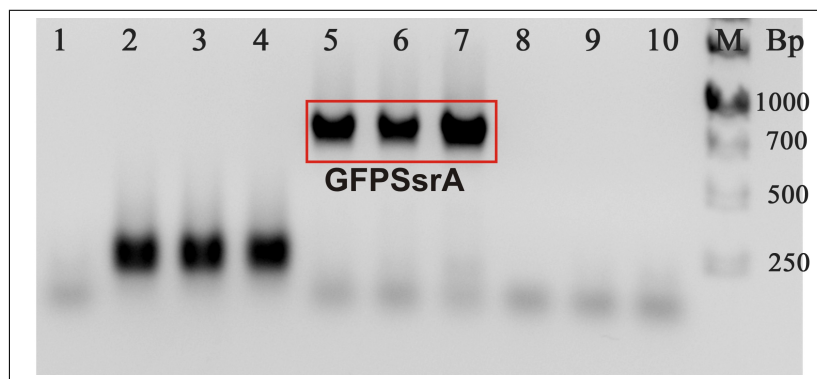


Abb. 39: Test verschiedener *Enhancer*-Zusätze zum PCR-Ansatz zur Generierung des Klonierungsinserts GFPSSrA. Gezeigt ist ein 1,5 prozentiges Agarosegel, die Visualisierung der Banden erfolgte über Ethidiumbromidfärbung. Die Spuren 8-10 zeigen die Ergebnisse für die Amplifikation mit DMSO-Zusatz (Spur 8: 70°C, Spur 9: 65°C, Spur 10: 60°C), die Spuren 5-7 das Amplifikationsresultat mit Zusatz von 1,5 Prozent DMSO und 0,5 Prozent Tween 20 (Spur 5: 70°C, Spur 6: 65°C, Spur 7: 60°C), sowie die Spuren 2-4 die Ergebnisse mit Betain-Zusatz (Spur 2: 70°C, Spur 3: 65°C Spur 4: 60°C). Auf Spur 1 ist eine Negativ-Kontrolle (*Annealing*-Temperatur 60°C) aufgetragen. Die rot umrahmten Banden zeigen das gewünschte Amplifikat.

Wie aus dem Gelbild (Abbildung 39) ersichtlich, wurde mittels DMSO-Zusatz kein Amplifikat erhalten. Der Zusatz von Betain führte zu kürzeren und damit unspezifischen Amplifikaten, unabhängig von der gewählten *Annealing*-Temperatur. Zielführend war die Verwendung einer Kombination aus Tween 20 und DMSO. In allen aufgetragenen Spuren (5-7) sind

intensive Banden des Amplifikates erkennbar. Mit etwas über 700 bp zeigt es die erwartete Länge. Die größte Produktmenge bildete sich bei einer *Annealing*-Temperatur von 60°C. In der auf Spur 1 applizierten Negativkontrolle konnte erwartungsgemäss kein Amplifikat detektiert werden.

Die Verwendung eines längeren *Forward* Primers wurde angestrebt, um den Schmelztemperaturunterschied zwischen *Forward* und *Reverse* Primer zu reduzieren. Die Strategie eines zweistufigen PCR-Programmes wurde angewandt, da eine mögliche Ursache für die fehlgeschlagene Amplifikation in der Tatsache vermutet wurde, dass zu Beginn der PCR nur der genspezifische Teil des Primers binden kann. Denn nur die Plasmid-DNA steht als Matrize zur Verfügung. Es existiert zu diesem Zeitpunkt noch kein elongierter Primer, welcher als vollständig komplementäres *Template* dienen könnte. Für diese Situation ist, aufgrund des langen Überhanges am 5'-Ende des Primers, eine niedrigere *Annealing*-Temperatur optimal. Wenn genügend elongierter Primer als *Template* vorhanden ist, steigt die optimale *Annealing*-Temperatur deutlich an, da jetzt die gesamte Primersequenz komplementär binden kann. Folglich kann die *Annealing*-Temperatur erhöht werden. Der Zusatz von Additiven (Punkt 3) wie DMSO oder Betain kann die Spezifität und Effektivität der PCR-Reaktion drastisch steigern (Sarkar *et al.*(1990); Frackmam *et al.* (1986)). Gründe dafür liegen z.B. in der Prävention von Polymerasen-Agglomerationen durch Detergenzienzugabe. Nach der Amplifikation wurden die beiden Produkte mittels des *PCR-Purification Kits* (Quiagen GmbH) säulenchromatographisch aufgereinigt und spektrophotometrisch am Nanodrop ND1000 Konzentration und Reinheit bestimmt. Eine Sequenzierung des PCR-Produktes bestätigte die korrekte Sequenz des Klonierungsinserts.

Klonierungsstrategie 1

Zur Realisierung der Klonierungsstrategie 1 wurden 3 μg des GFPSsrA-Inserts mit jeweils 3 Units der Restriktionsenzyme XbaI und HindIII versetzt und für 8 Stunden bei 37°C inkubiert. Zur Reinigung wurden die Restriktionsansätze auf ein Agarosegel (1%ig) aufgetragen. Die gewünschten *Insert*bande wurde anschließend mittels des Promega Wizard SV Gelisolutionskits aus dem Gel extrahiert. Der Zielvektor wurde ebenfalls mit den Restriktionsenzymen XbaI und HindIII geschnitten und zur Prävention von Religationen mit alkalischer Phosphatase behandelt. Bei sämtlichen beschriebenen Klonierungen wurden stets Religationskontrollen durchgeführt. Nachfolgend erfolgte eine Isolation und Aufreinigung des linearisierten Vektors aus dem Agarosegel analog dem der *Inserts*.

Die *Inserts* und 200 ng Zielvektor wurden im molaren Verhältnis von 3:1 in einem 20 μL -Ansatz mit T4-DNA-Ligase ligiert. Für Ligationen über kohäsive Enden wurde ein Temperaturgradient (20°C bis 4°C) gewählt. Das ligierte Plasmid wurde nachfolgend in chemisch kompetente XL1-Blue Zellen mittels Hitzeschock transformiert. Das Plasmid pBAD33 codiert für das CAT-Gen, welches Resistenz gegen Chloramphenicol verleiht. Daher erfolgte einen Ausstreichen der Zellen auf chloramphenicolhaltigen Agarplatten. Trotz Transformation mit unterschiedlichen Volumina vom Ligationsansatz wurden keine Kolonien auf den Agarplatten erhalten. Um die Gründe für das Scheitern von Strategie 1 zu evaluieren wurden die Prozessschritte des Restriktionsverdau und der Ligation separat analysiert. Um die Funktionalität der Restriktionsschnittstellen in den PCR-*Inserts* nachzuweisen wurden diese mit den entsprechenden Restriktionsendonukleasen geschnitten und einer Selbstligation unterzogen. Die gelektrophoretische Analyse der Selbstligationsansätze ist in Abbildung 40 gezeigt.

Wie aus der Abbildung 40 ersichtlich, zeigen alle Ligationsansätze die Bildung von Di-

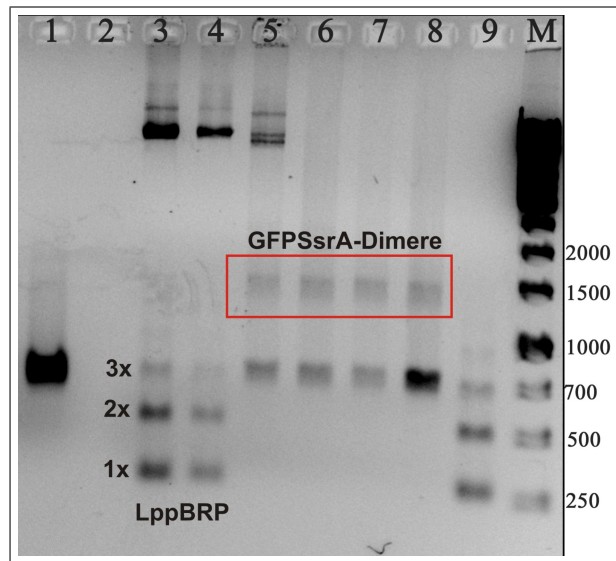


Abb. 40: Gezeigt ist ein einprozentiges Agarosegel, die Visualisierung der DNA-Banden erfolgte durch Färbung mit Ethidiumbromid und UV-Exposition. Aufgetragen sind die Selbstligationsansätze des mit HindIII (Spur 6), XbaI (Spur 7) sowie HindIII und XbaI (Doppelverdau, Spur 8) geschnittenen *Inserts* GFPSSrA. Zum Vergleich ist der Selbstligationsansatz eines weiteren geschnittenen *Inserts* (LppBRP mit HindIII, Spur 9) aufgetragen. Weiterhin sind auf den Spuren 3-5 Ligationsansätze der beiden *Inserts* (LppBRP auf Spur 3 und 4 sowie GFPSSrA auf Spur 5) mit dem Vektor pBAD33 gezeigt. Spur 1 zeigt das PCR-Produkt GFPSSrA.

bzw. Multimeren, was die Funktionalität der Restriktionsschnittstellen belegt. Das mit zwei Restriktionsenzymen prozessierte GFPSSrA-*Insert* (Spur 8) sollte auch höhere Multimere bilden können. Deren Konzentration ist allerdings möglicherweise so gering, dass sie unter die Detektionsgrenze fällt. In Spur 5 ist der Ligationsansatz des GFPSSrA-Konstruktes mit dem Zielvektor pBAD33 aufgetragen. Es ist erkennbar, dass das GFPSSrA-Konstrukt Konkatanere bildet und auch im Größenbereich über 5 kb sind mehrere Banden sichtbar. Die korrekte Funktion der T4-DNA-Ligase unter den gewählten Reaktionsbedingungen ist damit bewiesen. In den Spuren 3 und 4 sind Ligationsansätze einer parallel durchgeführten Klonierung des Zielvektors mit einem anderen *Insert* (LppBRP) aufgetragen. Wie aus dem Bild ersichtlich, bildet das *Insert* auch höhere Multimere aus. Allerdings ist zu beachten, dass das *Insert* deutlich kleiner (216 bp) als das GFP-Konstrukt ist und daher auch in höherem Maße zur Selbstligation tendiert. Ein Selbstligationsansatz dieses *Inserts* ist auf Spur 9 aufgetragen.

Basierend auf diesen Ergebnissen konnte gezeigt werden, dass beide Restriktionsschnittstellen im *Insert* von den Restriktionsendonukleasen XbaI und HindIII prozessiert werden. Zusätzlich belegt die Entstehung von *Insert*dimeren, die auch im Ligationsansatz mit dem Klonierungsvektor auftraten, die Funktionalität der T4-DNA-Ligase. Der Prozessschritt des Restriktionsverdau kann daher für ein Scheitern der Klonierung ausgeschlossen werden. Für eine Analyse des Ligationsschrittes wurden verschiedene molare *Insert*-Vektor Verhältnisse evaluiert. Zusätzlich wurden verschiedene Ligationsbedingungen (16°C, 4°C, Gradient RT bis 4°C) getestet. All diese Maßnahmen waren jedoch nicht zielführend, so dass im Folgenden die Ergebnisse der erfolgreichen Klonierungsstrategie 2 gezeigt werden.

Klonierungsstrategie 2

Die neuerliche Generierung der *Inserts* erfolgte mittels *Pfu*-Polymerase. Im Gegensatz zur *Taq*-Polymerase zeigt diese DNA-Polymerase keine terminale Transferase-Aktivität und damit keine unspezifische Adenosin-Elongation an dem 3'-Ende der amplifizierten Sequenz. Daher lassen sich derartige Amplifikate auch für *blunt-end*-Klonierungen verwenden. Um einen zweiten Klonierungsschritt zu umgehen, wurden die beiden *Inserts* *torA* und *GFPSsrA* zuerst im molaren Verhältnis von 2:1 miteinander ligiert und die entsprechende Bande aus dem Agarosegel ausgeschnitten und aufgereinigt. Das präparative Agarosegel zur Gewinnung des *torAGFPSsrA*-Fragmentes ist in der Abbildung 41 gezeigt.

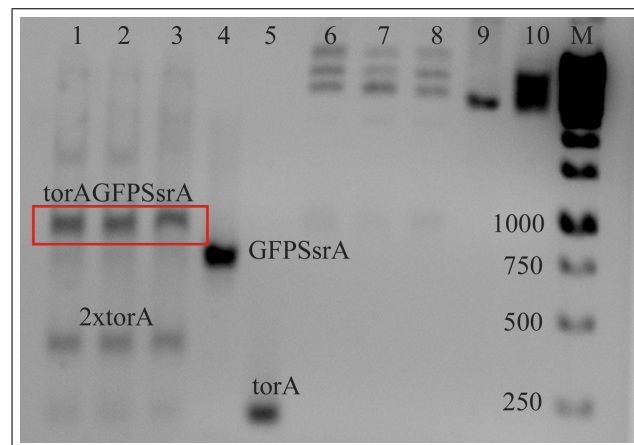


Abb. 41: Gezeigt ist ein einprozentiges Agarosegel, die Visualisierung erfolgte mit Ethidiumbromidfärbung. Die Spuren 1-3 zeigen die Ligationsansätze des Inserts *GFPSsrA* mit dem Insert *torA*, beide mit *XbaI* geschnitten. Zum Vergleich sind die PCR-generierten Inserts auf den Spuren 4 und 5 aufgetragen. Deutlich ist der Größenunterschied zwischen dem Ligationprodukt *torAGFPSsrA* und dem Insert *GFPSsrA* zu sehen. Weiterhin sind Ligationsansätze auf den übrigen Spuren 6-10 aufgetragen.

Auf den Spuren 1-3 sind die Ligationsansätze der mit *XbaI* geschnittenen *Inserts* *GFPSsrA* und *torA* aufgetragen. Zum Vergleich sind die PCR-generierten *Inserts* in daneben befindlichen Spuren 4 und 5 gezeigt. Deutlich ist die Ligationssbande des Konstruktes *torAGFPSsrA* etwas oberhalb derer des *GFPSsrA*-PCR-Produktes zu erkennen. Zusätzlich sind noch Banden von *torA*-Dimeren auszumachen, eine Bande unligiertes *GFPSsrA*-Insert ist kaum visuell zu erfassen. Zusammengefasst ist die Ligation von *GFPSsrA* mit *torA* damit nahezu quantitativ erfolgt. Die gewünschte Bande wurde aus dem Gel extrahiert. Das so generierte *torAGFPSsrA*-Konstrukt wurde in den Vektor *pGEM5Zf(+)* über die *EcoRV* Restriktionsschnittstelle mittels der *cut-ligation* Methode subkloniert (siehe Anhang A.12). Die Spuren 6-8 zeigen die Ligationsansätze zur Subklonierung des *torAGFPSsrA*-Konstruktes in den *pGEM5Zf(+)*-Vektor, welcher in linearisierter Form, nach *EcoRV*-Prozessierung, in daneben befindlicher Spur 9 gezeigt ist. Auf Spur 10 ist der Vektor ohne Linearisierung aufgetragen. In den Ligationsansätzen werden Banden oberhalb der geschnittenen Vektorbande beobachtet. Dies deutet auf eine erfolgreiche Ligation des *torAGFPSsrA*-Konstruktes hin. Die Ligationsansätze wurden in chemisch kompetente *XL1-Blue* Zellen transformiert, die resultierenden Transformationsansätze auf ampicillinhaltigen Agarplatten ausplattiert und über Nacht bei 37°C im Brutschrank inkubiert. Positive Kolonien können durch die

Blau-Weiss-Selektion einfach erkannt werden. Drei durch Kolonie-PCR verifizierte, positive Kolonien wurden gepickt, vermehrt und eine Plasmidpräparation durchgeführt. Aus dem so erhaltenen Plasmid wurde mittels der Restriktionsschnittstellen *SacI* und *XbaI* das gesamte Konstrukt herausgeschnitten, wiederum aus dem Gel aufgereinigt und mit dem ebenfalls mit *SacI* und *HindIII* geschnittenen Vektor pBAD33 ligiert. Der Vorteil dieser Vorgehensweise liegt in der Möglichkeit, die korrekte Prozessierung durch die Restriktionsendonuklease direkt im Agarosegel nachzuweisen. Zudem lassen sich *Inserts* aus Vektor-DNA signifikant einfacher ligieren als PCR-Produkte. Die Transformation des Ligationsansatzes erfolgte erneut in chemisch kompetenten XL1-Blue Zellen mittels Hitzeschock bei 42°C. Parallel wurden als Negativkontrolle kompetente Zellen ohne Zugabe des Ligationsansatzes analog prozessiert und auf Agarplatte mit Chloramphenicol-Zusatz ausgestrichen. Bei der Religationskontrolle waren nach Inkubation im Brutschrank bei 37°C über Nacht 7 Kolonien gewachsen. Die Platte mit der Negativkontrolle enthielt keine Kolonien, wohingegen die Platten mit den Ligationsansätzen 150-250 Kolonien enthielten. Von diesen wurden 5 Kolonien gepickt, im Schüttelkolben vermehrt und durch eine Kolonie-PCR und einem analytischen Restriktionsverdau nach der Plasmidisolierung untersucht.

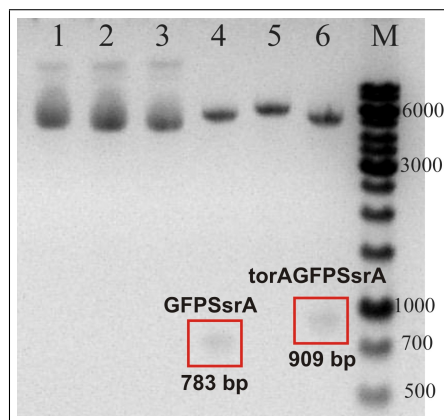


Abb. 42: Analytischer Restriktionsdoppelverdau des Plasmides pBADtorAGFPSsrA mit den Restriktionsendonukleasen *XbaI*, *SacI* und *HindIII*. Die Proben sind auf einem 1,5 prozentigem Agarosegel aufgetragen und mit Ethidiumbromid-Lösung gefärbt.

Die Abbildung 42 zeigt die Ergebnisse des analytischen Restriktionsverdaus, welcher auf einem 1,5 prozentigem Agarosegel aufgetragen wurde. Die Spuren 1-3 zeigen die Plasmidpräparationen dreier positiver Klone. Der mit *XbaI* und *HindIII* geschnittene Vektor ist auf der Spur 4 aufgetragen, für die Enzymkombinationen *SacI* und *XbaI* bzw. *XbaI* und *HindIII* ist selbiges auf den Spuren 5 und 6 gezeigt. Das entstehende Fragmentmuster deckt sich mit dem theoretisch erwarteten Bandenmuster, welches in Abbildung 43 illustriert wird.

Durch die Prozessierung des Vektors pBADtorAGFPSsrA mit *XbaI* und *HindIII* wird das GFPSsrA-Fragment aus der Vektorsequenz herausgeschnitten, welches eine Größe von ca. 780 bp aufweist. Auf Spur 4 wird ein Fragment in der erwarteten Größe gefunden, gleiches gilt für die Behandlung des Vektors mit *SacI* und *HindIII*, welche in der Entstehung des Fragmentes torAGFPSsrA mit einer Größe von ca. 910 bp resultiert und auf Spur 6 wiedergefunden wird. Das torA-Fragment, welches aus einer Behandlung des Vektors mit *SacI* und *XbaI* entsteht, konnte aufgrund der geringen Größe nicht visualisiert werden. Allerdings weist die Vektorbande auf Spur 5 im Vergleich zu den Spuren 4 und 6 die größte

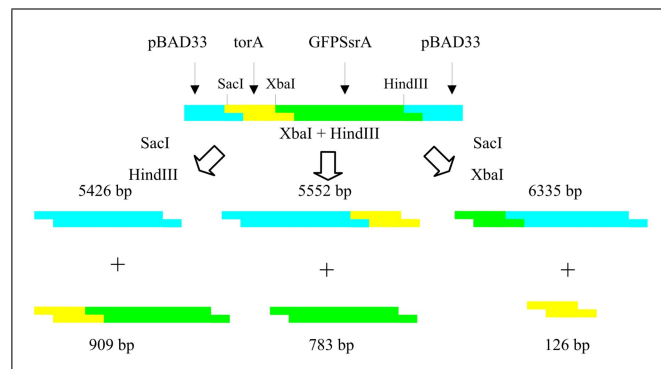


Abb. 43: Das theoretische Fragmentmuster bei der Prozessierung des Plasmides pBADtorAGFPsSrA mit allen möglichen Doppelkombinationen der Restriktionsenzyme XbaI, HindIII und SacI.

Länge auf, was den Verlust des kleinsten Abschnittes aus der Vektorsequenz belegt. Analog weist die Vektorbande auf Spur 4 eine größere Länge auf als jene der Spur 6.

Da sowohl der Restriktionsverdau als auch die Kolonie-PCR ein positives Ergebnis generierten, wurden drei Plasmidminipräparationen zur Sequenzierung geschickt. In Koinkidenz mit dem Ergebnis aus dem Restriktionsverdau und der Kolonie-PCR wurde die korrekte Sequenz des Konstruktes bestätigt.

4.7.3 Klonierung weiterer GFP-Stämme zur Charakterisierung und Verifikation des etablierten Sekretionsreportersystems

Nachdem der Sekretionsvektor pBADtorAGFPsSrA erfolgreich generiert wurde, erfolgte für die nähere Charakterisierung des Sekretionssystems die Konstruktion weiterer, auf der GFPmut2-Variante basierender Plasmide. Die zu klonierenden Konstrukte sind in der Abbildung 44 gezeigt.

Um die Sezernierung des GFPs in das Periplasma auf einfache Art und Weise nachzuweisen, wurde das Konstrukt pBADtorAGFPHis und zur Verifikation pBADGFPHis konstruiert (Fusionierung des GFP-Derivats mit einem Hexahistidin(His-tag)). In einer durchgeführten Zellfraktionierung sollte nur die Periplasmafraktion des Stammes mit dem Plasmid pBADtorAGFPHis GFP enthalten, nicht jedoch jene von pBADGFPHis, da diesem das entsprechende Sekretionssignal fehlt. Der Nachweis kann dabei einfach über die Fluoreszenz und auf Proteinebene über die Aufreinigung und Aufkonzentrierung mittels Immobilisierender Metallionenchelchromatographie (IMAC) erfolgen. Da die Sekretionskapazitäten mittels des Tat-Pfades limitiert sind und zudem damit gerechnet werden muss, dass im Größenbereich von GFP andere Wirtsproteine vorliegen, stellt die Fusionierung mit einem His-tag und die Aufreinigung und Aufkonzentrierung mittels Affinitätschromatographie einen deutliche Vereinfachung des Nachweises dar (Lottspeich F: Bioanalytik (1998)). Eine Alternative stellt der direkte Nachweis über einen Western Blot mittels GFP-Antikörpers dar. Durch den His-tag kann jedoch auch ein immunologischer Nachweis über einen His-tag-spezifischen Antikörper erfolgen. Zu Vergleichszwecken wurden die Konstrukte pBADGFP und pBADtorAGFP generiert. Das Stamm mit dem Plasmid pBADtorAGFP fungiert als Negativkontrolle für die Aufreinigung über IMAC. Die Periplasmafraktion sollte eine grüne Fluoreszenz aufweisen, darf aber nicht an die Säule gebunden werden. Zusätzlich lässt sich

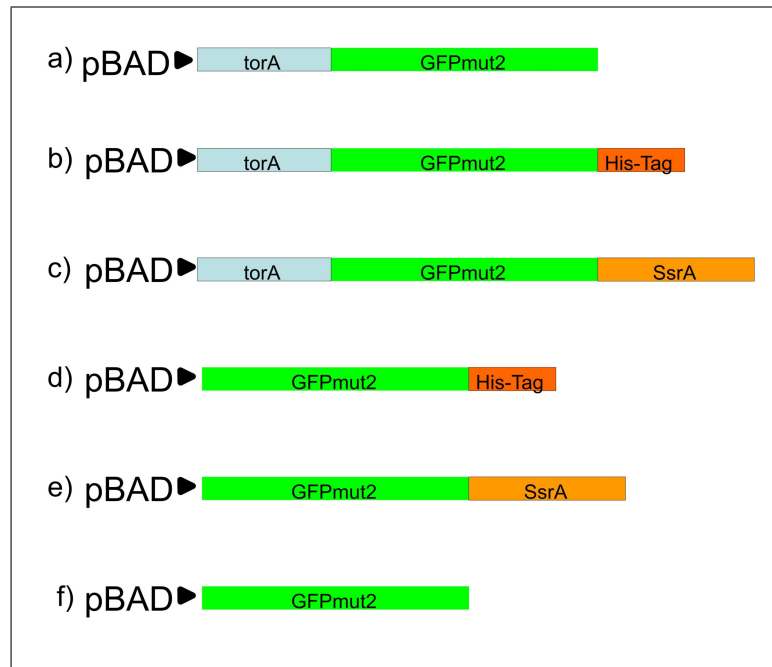


Abb. 44: Verschiedene GFP-Konstrukte zur näheren Charakterisierung und Verifizierung des etablierten Sekretionsreportersystems.

evaluieren, welchen Einfluß der C-terminale Hexahistidin-*tag* auf die Fluoreszenzeigenschaften des GFPs hat. Um die Wirksamkeit des C-terminal lokalisierten SsrA-*tags* zu demonstrieren, wurde der Stamm pBADGFPSSrA konstruiert. Ohne die Signalsequenz zur Sekretion sollte das cytoplasmatisch lokalisierte GFP vollständig von dem Clp-Proteasesystem degradiert werden. Somit lässt sich die Höhe des Induktionsniveaus bestimmen, bei dem das Proteasesystem nicht überlastet wird.

Für die verschiedenen Konstrukte wurde eine analoge Klonierungsstrategie angewandt, wie sie bei dem Konstrukt pBADtorAGFPSSrA erfolgreich zum Einsatz kam. Zuerst wurden die mittels *Pfu*-Polymerase generierten PCR-Amplifikate *blunt-end* in den Vektor pGEM5Zf(+) subkloniert. Abbildung 45 zeigt die so gewonnenen Plasmide und deren Analyse durch einen analytischen Restriktionsverdau.

Spur 1 und 20 enthalten einen Längenstandard, die einzelnen Banden sind in der Bildunterschrift spezifiziert. Der Zielvektor für die Subklonierung pGEM5Zf(+) ist in Spur 18 aufgetragen. Die Spuren 12-17 zeigen die unterschiedlichen GFP-Plasmide. Deutlich sichtbar ist der Größenzuwachs im Vergleich zu pGEM5Zf(+). Die Spuren 1-10 zeigen die analytischen Restriktionsverdau für die jeweiligen Plasmide, aufgetragen neben den aufgereinigten PCR-*Inserts*. Die Konstrukte GFP und GFP_{HIS} wurden mit XbaI und HindIII geschnitten, während GFP_{RBS}, GFP_{RBSHIS} und GFP_{SSrA} mit den Restriktionsenzymen SacI und HindIII behandelt wurden. Alle Plasmide verlieren durch den Restriktionsverdau ein Fragment, welches in der Größe dem amplifizierten PCR-Produkt (ca. 700 bp) entspricht. Da die Restriktionsschnittstellen im Zielvektor nicht existent sind, kann zudem davon ausgegangen werden, dass die Subklonierung erfolgreich verlaufen ist. Eine Kolonie-PCR mit GFP-spezifischen Primern bestätigt das positive Ergebnis (Abbildung 46).

Analog der Klonierung von pBADtorAGFPSSrA wurden die GFP-Konstrukte über die in

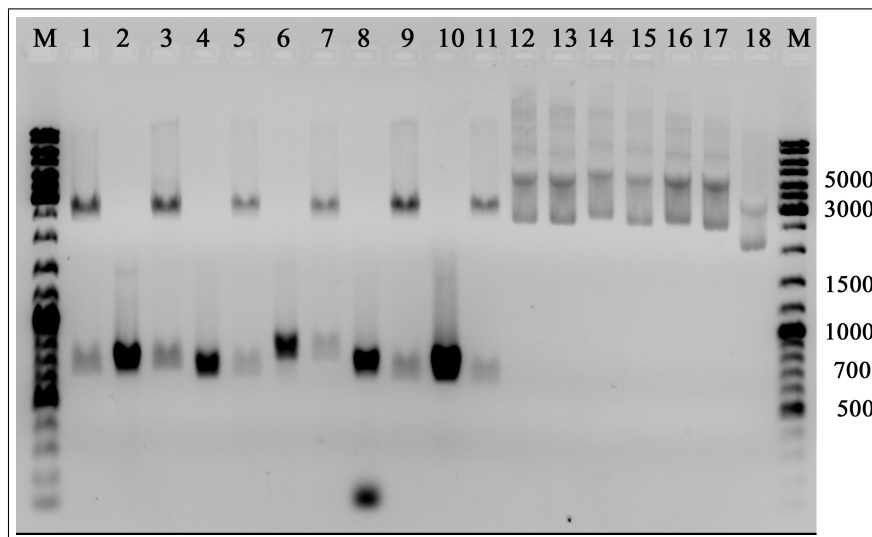


Abb. 45: Die mittels *cut-ligation* Methode generierten Plasmide der verschiedenen GFP-Konstrukte in pGEM5Zf(+)(Spuren 12-17) und deren Analyse mittels Restriktionsdoppelverdau (Spuren 1, 3, 5, 7, 9), aufgetragen auf einem 1,5 prozentigem Agarosegel. Zur Kontrolle sind die PCR-generierten *Inserts* (Spuren 2, 4, 6, 8) mit appliziert. Die Restriktionsverdau auf den den Spuren 1, 3, 5, 7, 9 sind der Reihe nach: GFPSrA1, GFPSrA2, GFPRBSHis, GFPRBS, GFPHis, GFP. Jeweils links neben den zuletzt spezifizierten Spuren sind die PCR-generierten *Inserts* aufgetragen. Die Plasmide auf den Spuren 12-17 sind: GFP, GFPHis, GFPRBSHis, GFPRBS und GFPSrA. Auf Spur 18 ist der Klonierungsvektor pGEM5Zf(+) aufgetragen.

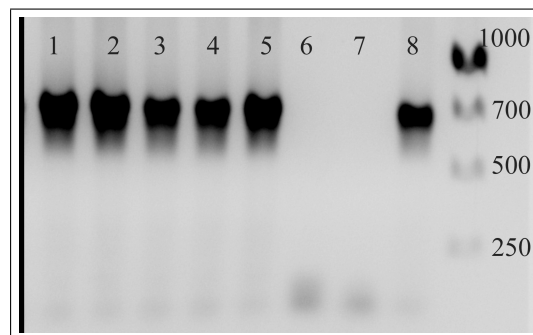


Abb. 46: Kolonie-PCR verschiedene GFP-Konstrukte, aufgetragen auf ein 1,5 prozentiges Agarosegel und gefärbt mit Ethidiumbromid. Ein 1 kb Größenstandard ist zur Zuordnung der Amplifikatlängen aufgetragen, Spur 8 zeigt das PCR-Amplifikat eines positiven torAGFPSrA-Klones, wohingegen Spur 7 das Amplifikationsergebnis eines Klones zeigt, welcher kein GFP-*Insert* trägt. Spur 6 wird von einer Negativkontrolle belegt und die Spuren 1-5 zeigen die Amplifikate positiver Klone mit den restlichen generierten Plasmiden. Von links nach rechts sind dieses: pBAD-torAGFP, pBAD-torAGFPHis, pBAD-RBS-GFPSrA, pBAD-RBS-GFPHis und pBAD-RBS-GFP.

den Vektor eingebrachten Restriktionsschnittstellen XbaI und HindIII bzw. SacI und HindIII aus der Vektorsequenz isoliert und aus dem Agarosegel extrahiert und gereinigt. So erhaltene *Inserts* wurden in den ebenfalls mit SacI und HindIII bzw. XbaI und HindIII geschnittenen Vektor pBADtorAGFPSrA über kohäsive Enden kloniert. Eine Verwendung dieses Vektors wurde gegenüber dem Klonierungsvektor pBAD33 favorisiert, da dieser schon die torA-Sequenz enthielt und eine korrekte Prozessierung durch die Restriktionsendonukleasen durch Verlust des GFPsSrA-Konstruktes leicht auf dem Agarosegel nachzuweisen ist.

Die Ligation mittels eines Temperaturgradienten wurde über Nacht im Kühlraum nach der bereits beschriebenen Methode (Anhang A.13) durchgeführt, die Ligationsansätze anschließend in chemisch kompetente XL1-Blue Zellen mittels Hitzeschock transformiert und diese auf chloramphenicolhaltigen LB-Agarplatten ausgestrichen. Über Nacht gewachsene Klone wurden gepickt, vermehrt und standardmäßig mittels Kolonie-PCR und nach der Plasmidisolierung mittels analytischem Restriktionsverdau untersucht. Die Plasmidpräparationen und das Ergebnis des analytischen Restriktionsverdaus ist, stellvertretend für die Gesamtheit der Konstrukte, für sechs Klone aus zwei verschiedenen Ligationsansätzen in Abbildung 47 gezeigt. Es handelt sich dabei um Klone, die aus Transformationsansätzen zur Generierung der Konstrukte pBADtorAGFP und pBADtorAGFPHis hervorgegangen sind.

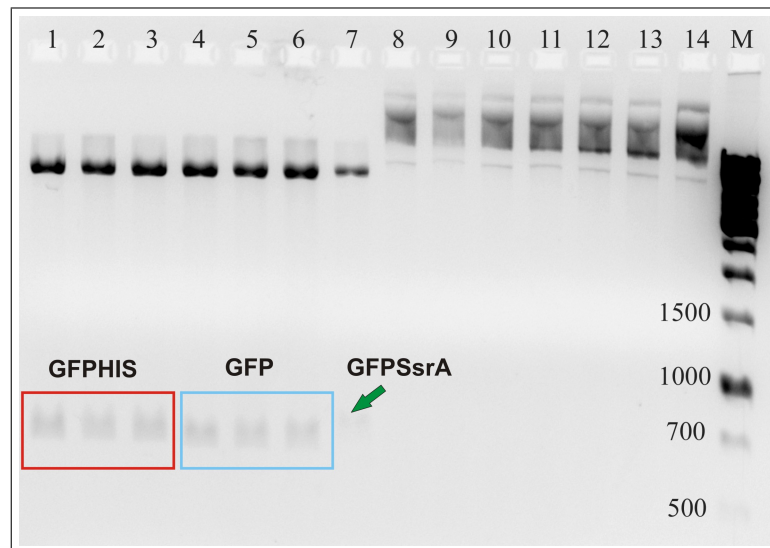


Abb. 47: Analytischer Restriktionsverdau der Plasmidpräparationen von sechs Klonen zur Generierung der Plasmide pBADtorAGFP und pBADtorAGFPHis mit den Restriktionsenzymen XbaI und HindIII (Spuren 1-6). Zu Vergleichszwecken wurde das bereits verifizierte Plasmid pBADtorAGFPSrA analog prozessiert und mit auf das Gel aufgetragen (Spur7). Nebenstehend sind die Plasmidpräparationen der Klone gezeigt (Spuren 8-14).

In den Spuren 11-13 sind die Plasmidpräparationen dreier Klone gezeigt, die das Plasmid pBADtorAGFP tragen, die dazugehörigen Restriktionsverdaus mit XbaI und HindIII sind in den Spuren 4-6 gezeigt. Analoges gilt für das Plasmids pBADtorAGFPHis. Die Plasmidpräparationen sind in den Spuren 8-10 gezeigt und die zugehörigen Restriktionsverdaue auf den Spuren 1-3. Gut zu erkennen ist die Existenz eines durch den Restriktionsverdau entstandenen Fragmentes in Größenbereich von 700 bp, welches in der Größe dem Fragment aus dem zu Vergleichszwecken ebenfalls mit aufgetragenden Restriktionsverdaus des verifizier-

ten Plasmids pBADtorAGFPSsrA auf Spur 7 entspricht. Bei genauer Betrachtungsweise läßt sich der geringe Größenunterschied zwischen den verschiedenen GFP-Konstrukten (GFPSsrA > GFPHIS > GFP) visuell erfassen. Bei allen Restriktionsansätzen wird eine gleich große Bande im Bereich um 6 kb beobachtet. Dabei handelt es sich um den linearisierten pBAD33-Klonierungsvektor. Die Existenz von Restriktionsfragmenten im erwarteten Größenbereich und die wiedergefundenen relativen Größenverhältnisse der *Inserts* sind deutliche Indizien einer erfolgreichen Klonierung.

4.7.4 Fluoreszenzmikroskopische Untersuchung

Nachdem alle Vektorsequenzen durch Sequenzierung verifiziert wurden, erfolgte die Transformation der verschiedenen Plasmide in den Stamm XL1-Blue, um die Funktionalität des Sekretionssystems auf Proteinebene nachzuweisen. Zu diesem Zweck wurde der plasmidtragende Stamm im Schüttelkolben kultiviert und in der exponentiellen Phase beim Erreichen einer OD600 von 0,8 durch Zugabe von *L*-Arabinose in einer Endkonzentration von 0,2 Prozent induziert. Parallel dazu erfolgte die Kultivierung einer uninduzierten Kontrolle des gleichen Stammes. Die Zellen wurden zentrifugiert, in PBS-Puffer resuspendiert und anschließend unter dem Fluoreszenzmikroskop betrachtet. Dadurch, dass die Zellen nicht fixiert wurden, konnte keine genaue Fokussierung auf einzelne Zellen erfolgen. Dennoch belegte die Hintergrundfluoreszenz, gezeigt in Abbildung 48, eindeutig die Anwesenheit von GFP, wie auch ein Vergleich mit dem uninduzierten Stamm oder einer Probe der gleichen Kultur vor Induktion zeigte.

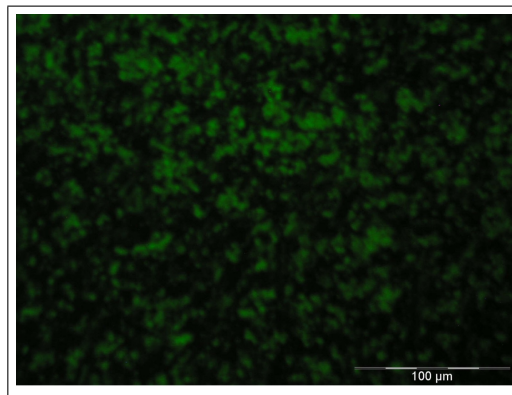


Abb. 48: Fluoreszenzmikroskopische Untersuchung der Hintergrundfluoreszenz des mit 0,2 Prozent *L*-Arabinose induzierten Stammes XL1-Blue pBADtorAGFPSsrA. Deutlich ist eine intensive grüne Fluoreszenz zu erkennen, die beim uninduzierten Stamm nicht auftrat.

4.7.5 Durchflußzytometrie

Um quantifizierbare Werte der GFP-Fluoreszenz zu erlangen, wurden die verschiedenen GFP-exprimierenden Stämme im Durchflußzytometer analysiert. Bei der Durchflußzytometrie handelt es sich um eine Technik zur simultanen Analyse der Größe und Granularität sowie mehrerer Fluoreszenzeigenschaften einer Zelle. Die Analysenrate beträgt dabei bis zu 50000 Zellen pro Minute. Ein Durchflußzytometer gliedert sich in drei funktionelle Abschnitte, ein Flüssigkeitssystem zum Transport der Zellen zur Detektionskammer, ein optisches System zur Anregung und Detektion der Zellen und ein elektronisches System zur

Datenerfassung und Steuerung des Gerätes. Der apparative Aufbau ist in Abbildung 49 wiedergegeben.

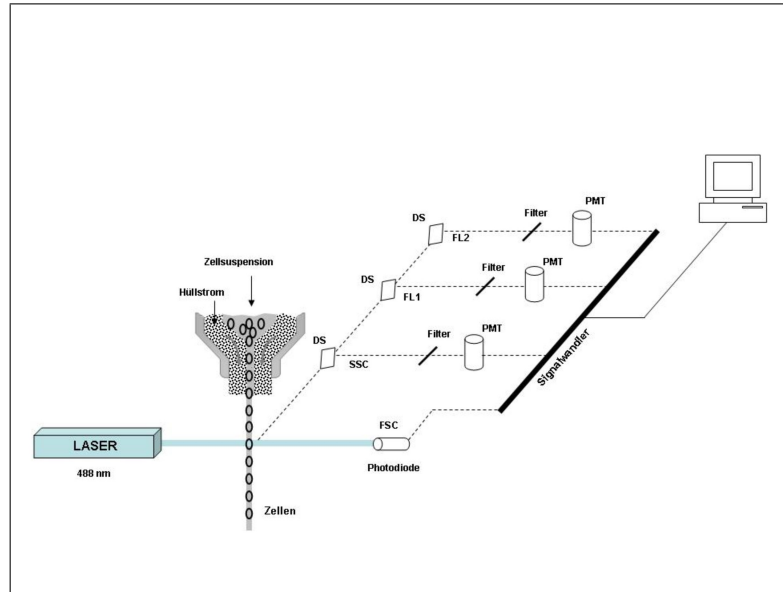


Abb. 49: Aufbau eines Durchflußzytometers. Die mittels eines Hüllstromes separierten Einzelzellen werden von einem orthogonal zum Probenstrom lokalisierten Laser angeregt. Das generierte Streulicht wird über verschiedene optische Anordnungen analysiert und in ein digitales Signal konvertiert.

Die zu analysierende Zellsuspension wird durch eine Düse geleitet und dabei von einem Flüssigkeitsstrom umgeben (Hüllstrom). Der Hüllstrom bewirkt eine hydrodynamische Fokussierung der Zellsuspension zu einer Sequenz von Einzelzellen. Rechtwinklig zum Probenfluß werden die Zellen durch einen Laser (meistens ein Argon-Ionenlaser mit einer Anregungswellenlänge von 488 nm) angeregt und die Lichtstreuung durch die Zellen als Vorwärtsstreulicht (*forward scatter*) sowie im 90° Winkel zu Laserstrahl und Probenstrom als Seitwärtsstreulicht (*sideward scatter*) erfasst. Das Vorwärtsstreulicht ist dabei ein Maß für die Größe und Form der Zellen, wohingegen das Seitwärtsstreulicht aus der Streuung des Laserstrahls an intrazellulären Zellkompartimenten entsteht und daher ein Maß für die Granularität der Zellen ist. Die Anregung von Fluoreszenzmarkern wird ebenfalls über das Seitwärtsstreulicht detektiert. In durchflußzytometrischen Analysen können simultan mehrere Fluoreszenzmarker nachgewiesen werden, die Emissionsspektren werden durch dichroitische Spiegel und Bandpassfilter voneinander separiert, solange der Wellenlängenunterschied der Emissionsmaxima der verschiedenen Fluorophore ausreichend groß ist. Das generierte Streulicht wird in *Photomultiern* detektiert, in elektrische Signale umgewandelt und verstärkt. Anschließend werden die elektrischen Signale prozessiert, als digitalisierter Datensatz ausgegeben und gespeichert. Eine Auftragung der Meßdaten kann dabei ein- oder zweidimensional erfolgen. Die Auftragung von Streulichtparametern oder Fluoreszenzintensitäten gegen die Partikelzahl in Histogrammen wird als eindimensionale Analyse bezeichnet. Bei der zweidimensionalen Auftragung, dem sogenannten *dotplot*, werden hingegen auf der Ordinate und Abzisse die Intensitätsverteilungen zweier Meßparameter aufgetragen. Aus *dotplot* und Histogramm lassen sich Informationen über die Verteilung von Intensitäten für

einen Parameter bzw. die Assoziation von zwei verschiedenen Messparametern generieren. Diese können verwendet werden, um Subpopulationen von Zellen mit bestimmten Eigenschaften zu definieren. Durch ein Verfahren, welches als elektronisches *gating* bezeichnet wird, können anschließend andere Eigenschaften dieser Zellsubpopulation näher untersucht werden.

In einem Parallel-Versuch wurden die Fluoreszenzeigenschaften der GFP-codierenden Plasmide durch eine durchflußzytometrische Analyse evaluiert. Die verschiedenen Stämme wurden in 100 ml LB-Medium in 500 ml Schüttelkolben kultiviert und beim Erreichen einer OD von 0,8 mit 0,2 Prozent *L*-Arabionose (Endkonzentration) induziert. Eine uninduzierte Vergleichsprobe fungierte dabei als Referenz für alle untersuchten Zeitpunkte. Direkt vor der Induktion erfolgte die Entnahme einer Probe und die nachfolgende Analyse im Durchflußzytometer von jeder unterschiedlichen Kultur. Zu diesem Zweck erfolgte die Entnahmen und Zentrifugation von 0,5 ml der Kultur. Das Zellpellet wurde in 0,5 ml PBS resuspendiert und 1 μ l davon, je nach vorliegender Zelldichte, mit 0,5-2 ml sterilfiltriertem PBS verdünnt. 30 und 60 Minuten nach Induktion erfolgte eine erneute Probennahme und durchflußzytometrische Analyse. Zur Veranschaulichung der Ergebnisse sind nicht alle Histogramme wiedergegeben, Abbildung 50 zeigt die Ergebnisse der durchflußzytometrischen Analyse der Proben welche 60 Minuten nach der Induktion entnommem wurden. Zusätzlich wurde die Probe des Stammes mit dem Plasmid pBADtorAGFPsSrA 30 Minuten nach Induktion mit aufgetragen, um das Anwachsen der Fluoreszenz zu dokumentieren.

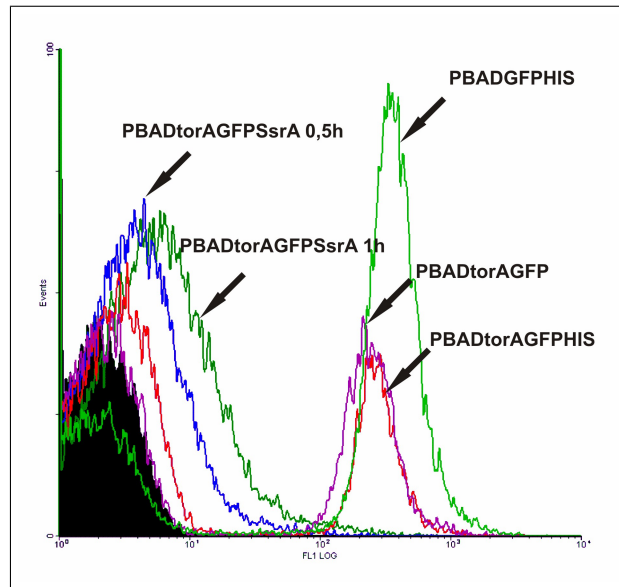


Abb. 50: Durchflußzytometrische Analyse von *E. coli* Stämmen, welche verschiedene GFP-codierende Plasmide tragen.

Der Stamm, welcher das Konstrukt pBADGFPHis trägt, generiert die höchste Fluoreszenzintensität, wohingegen die uninduzierte Referenz und der Stamm pBADGFPsSrA (schwarz ausgefüllte Referenz) keine Fluoreszenz zeigen. Dies belegt, dass die gewählten Induktionsbedingungen das cytoplasmatische Proteasesystem nicht überlasten; alles cytoplasmatisch lokalisierte GFP wird degradiert. Die Fluoreszenzintensität der Stämme pBADGFPHis und pBADtorAGFPHis ist unwesentlich geringer als die der analogen Stäm-

me ohne Hexahistidin-Modifikation (in der Graphik nur für pBADGFPHis gezeigt). Dies deutet auf keinen signifikanten Einfluß des *tags* auf die korrekte Faltung des Proteins hin. Zusätzlich weisen die Stämme, welche die GFP-Plasmide ohne Signalpeptid tragen, eine höhere Fluoreszenzintensität auf als jene mit Signalpeptid (ebenfalls nur für pBADGFPHis gezeigt). Die N-terminale Fusion mit *torA* scheint geringe Auswirkungen auf die Fluoreszenzeigenschaften zu haben. Es kann dabei ein Einfluß auf die Faltung vermutet werden. Der Stamm, welcher das Plasmid pBADtorAGFPSrA trägt, weist nach Induktion die niedrigste, aber eine dennoch signifikante Fluoreszenzintensität auf. In der Differenz der Fluoreszenzintensität zwischen den Stämmen mit den Plasmiden pBADtorAGFP und pBADtorAGFPHis spiegelt sich der Einfluß des SsrA-Degradations-*tags* wieder. Ein zusätzlicher Vergleich mit dem nicht-fluoreszierenden, pBADGFPSrA beinhaltenden Stamm zeigt, dass der Fluoreszenzunterschied zwischen den beiden Stämmen durch das Sekretionssignal entsteht, welches das cytoplasmatisch lokalisierte GFP durch Sezernierung ins Periplasma vor der Degradation bewahrt. Dies stellt einen Beweis für die Funktionalität des Sekretionsreporterplasmides pBADtorAGFPSrA im Sinne der angestrebten Applikation dar. Ausschließlich die sekretierte, ins Periplasma gelangte GFP-Menge wird durch die Messung der Gesamtfluoreszenzintensität erfasst.

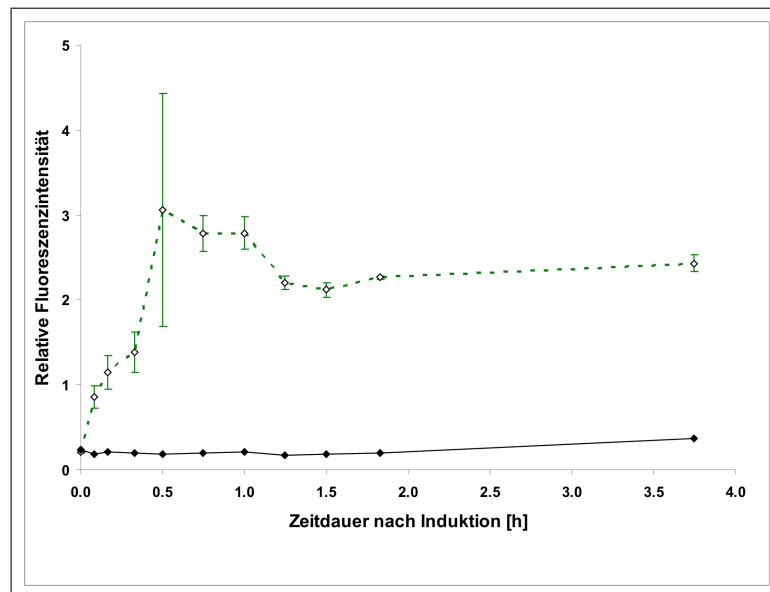


Abb. 51: Kinetik der GFP-Fluoreszenz des Stammes XL1-Blue pBADtorAGFPSrA nach Induktion mit 0,2 Prozent *L*-Arabinose. Gezeigt ist der Mittelwert der relativen Fluoreszenz, ermittelt aus drei identischen Schüttelkolbenkulturen, welche gegen die Zeitdauer nach Induktion aufgetragen ist. Zum Vergleich ist eine uninduzierte Kultur desselben Stammes mit aufgetragen.

Von dem das Konstrukt pBADtorAGFPSrA tragenden Stamm wurde eine genauere Kinetik mit einer Mehrfachbestimmung evaluiert. Die relative grüne Fluoreszenz als Mittelwert dreier parallel durchgeführten Schüttelkolbenkultivierungen ist in Abbildung 51 gegen die Kultivierungsdauer relativ zum Zeitpunkt der Induktion aufgetragen. Als Referenz wurde ein uninduzierter Vergleichsstamm parallel untersucht. Bis ca. 0,3 Stunden nach Induktion folgt der Kurvenverlauf einem logarithmischen Zusammenhang. Ab diesem Zeitpunkt steigt die Fluoreszenz rapide an, um nach dem Erreichen eines Maximums 0,6 Stunden nach In-

duktion auf einen fast konstanten Wert zu fallen. Zwei Stunden nach Induktion bleibt die Fluoreszenzintensität konstant. Es kann davon ausgegangen werden, dass bis zu diesem Zeitpunkt GFP exprimiert wird. Durch die Metabolisierung des Induktors *L*-Arabinose bricht die Expression dann nachfolgend zusammen. Aufgrund der großen Halbwertszeit des GFPs von mehr als 24 Stunden bleibt die Fluoreszenzintensität aber konstant. Das Maximum 0,6 Stunden nach Induktion entsteht sehr wahrscheinlich durch eine kurzfristige Überlastung des Proteasesystems. Für eine Quantifizierung der sekretierten GFP-Stoffmenge über die Gesamtzellfluoreszenz sollte daher ein späterer Zeitpunkt gewählt werden. Ab 2 Stunden nach Induktion bleibt die Fluoreszenzintensität konstant und kann als Maß für das sezernierte GFP verwendet werden. Generierte Mutantenstämme, welche eine erhöhte Sekretionskapazität aufweisen, können somit durch eine einfache Messung der Gesamtzellfluoreszenz und dem Vergleich mit dem unmodifizierten Stamm identifiziert werden.

4.7.6 SDS-PAGE und Zellfraktionierung

Nachdem die durchflußcytometrische Analyse bereits erste indirekte Beweise für die Sekretion des GFP-Konstrukts ins Periplasma ergeben hatte, wurde der Nachweis des GFP mit Hilfe einer Zellfraktionierung durchgeführt. Durch die Zellfraktionierung wird vom gramnegativen Bakterium eine Periplasma- und Cytoplasmafraktion erhalten, welche anschließend mittels SDS-PAGE analysiert wurde. Zusätzlich konnte bei den Stämmen, die das Plasmid pBADGFPHis und pBADtorAGFPHis tragen das GFP aufgrund des Hexahistidin-*tags* mittels Affinitätschromatographie aufkonzentriert und aufgereinigt werden. Bei dem Stamm mit dem Plasmid pBADGFPHis sollte dabei nur in der Cytoplasmafraktion GFP detektiert werden. Dies ist zugleich eine Qualitätskontrolle über die durchgeführte Fraktionierung. Bei Zellfraktionierungen besteht die Möglichkeit, dass einige Zellen lysieren und cytoplasmatischer Inhalt in die Periplasmafraktion gelangt. Zur Verifikation der Fraktionierung wird die Anwesenheit der für die jeweiligen Fraktion charakteristischen Marker und entsprechend deren Abwesenheit in der anderen Fraktion überprüft. Mittels der in Rahmen dieser Studie etablierten Stämme lässt sich visuell durch Auftreten oder Ausbleiben einer Grünfärbung der Fraktionierungserfolg verifizieren. Das gesagte gilt auch für die Verifikation der Aufreinigung mittels immobilisierender Metallionen-Chelatchromatographie (IMAC). Der Durchlauf der Säule nach dem Bindungsschritt war vollständig klar und die Säule intensiv grün gefärbt. Bei der Elution weist das Eluat entsprechend wieder eine intensive Grünfärbung auf und die Säulenpackung wird entfärbt. Durch eine Quantifizierung der GFP-Menge ließe sich somit auch eine Vergleichsstudie über die Effizienz und die Ausbeuten von verschiedenen Aufreinigungen und Fraktionierungen mit einfachen Mitteln durchführen. Die Durchführung der Zellfraktionierung wird im Anhang (A.1.30) beschrieben. Gleiche Zellzahlen der etablierten Stämme mit den Plasmiden pBADGFPHis und pBADtorAGFPHis sowie pBADtorAGFPSrA wurden vor und nach Induktion fraktioniert und ein Teil der Peri- und Cytoplasmafraktion, wie im Anhang beschrieben (Abschnitt A.1.28), über Viva-pure Metal Chelate Spin Columns affinitätschromatographisch aufgereinigt. Das restliche Fraktionsvolumen wurde mittels Trichloressigsäure (10 Prozent Endkonzentration) gefällt, das Präzipitat in 50 μ L Wasser aufgenommen und mit dem gleichen Volumen Lämmli-puffer versetzt. Zur Neutralisation der anhaftenden TCA wurden 1-5 μ L 2M Tris-Base Lösung (pH 9,0) zugesetzt. Die Periplasma und Cytoplasmafraktionen der verschiedenen Proben, sowie die Eluate der affinitätschromatographischen Aufreinigung wurden auf ein Polyacrylamidgel aufgetragen.

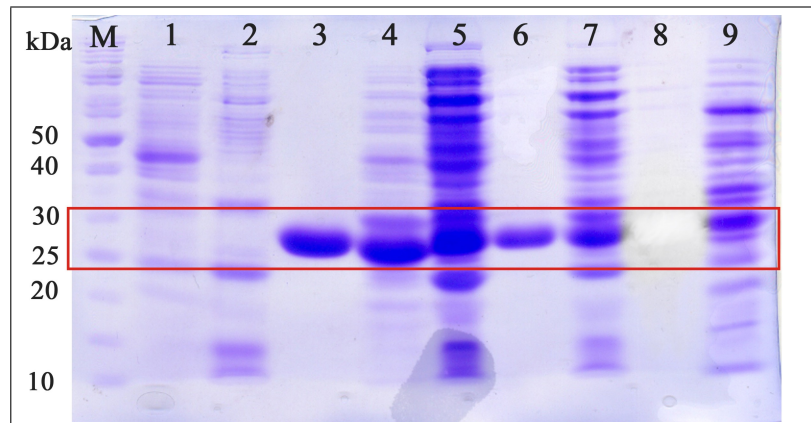


Abb. 52: Untersuchung der Peri- und Cytoplasmafraktionen verschiedener GFP-codierender Stämme mittels SDS-PAGE. Gezeigt ist ein 12,5 prozentiges Polyacrylamid-Gel welches mit Coomassie Blue gefärbt wurde.

Das Ergebnis der SDS-PAGE ist als Scan des Gelbildes in Abbildung 52 wiedergegeben. Die Färbung des Gels erfolgte mit *Coomassie Blue*, die Versuchsvorschriften zur Durchführung der PAGE, Zellfraktionierung und Affinitätschromatographie ist dem Anhang (A.1.25, A.1.28, A.1.30) zu entnehmen. Spur M zeigt einen Proteinmarker im Größenbereich von 10-100 kDa, die Spezifizierung der einzelnen Marker-Banden wurde in der Abbildung 52 vorgenommen. Auf den Spuren 1 und 2 wurden die Peri- und Cytoplasmafraktionen der uninduzierten Vergleichskontrolle aufgetragen. Es ist zu erkennen, dass sich die beiden Fraktionen vom Bandenmuster deutlich unterscheiden. Die Cytoplasmafraktion sollte einen deutlich höheren Proteingehalt und damit intensivere Banden aufweisen. Dass dies sich nicht im Gelbild widerspiegelt, liegt zum einen an der durchgeführten Proteinfällung mittels TCA und zum anderen an einem suboptimalen Zellaufschluss der Cytoplasmafraktion. Der Zellrückstand nach dem Ultraschallaufschluss wies noch eine grüne Fluoreszenz auf, was beweist, dass viele Zellen noch nicht vollständig aufgeschlossen wurden. Im Größenbereich um 30 kDa, in dem die GFP-Bande vermutete wurde, zeigen beide Fraktionen kaum Proteinbanden. Die Spuren 3-6 enthalten die Peri- und Cytoplasmafraktionen des Stammes mit dem Plasmid pBADtorAGFPHis sowie die über IMAC aufgereinigten Fraktionen beider Zellkompartimente. Die über IMAC aufgereinigte Periplasmafraktion ist auf Spur 3 aufgetragen, jene der Cytoplasmafraktion auf Spur 6. Deutlich ist bei beiden eine sehr intensive Bande zu erkennen, welche in der uninduzierten Vergleichsprobe nicht zu erkennen ist. Mit einer Größe von ca. 30 kDa liegt das nachgewiesene Produkt im erwarteten Bereich des GFPs. Da die Eluate der IMAC-Aufreinigung intensiv grün gefärbt waren und sonst keine Proteinbanden in diesen beiden Fraktionen erkennbar waren, ist mit hoher Wahrscheinlichkeit davon auszugehen, dass es sich dabei um die gesuchte GFP-Bande handelt. In den Peri- und Cytoplasmafraktionen ist diese intensive Bande ebenfalls existent. Zur Verifikation der Fraktionierung wurden die über IMAC aufgereinigte Periplasmafraktion des Stammes, welcher das Plasmid pBADGFPHis trägt, ebenfalls in der Spur 8 aufgetragen. Diese Fraktion zeigte keine grüne Fluoreszenz, und es ließ sich entsprechend kein Proteinprodukt über IMAC aufreinigen, wie das Gelbild bestätigt. Erwartungsgemäß wird in der Cytoplasmafraktion dieses Stammes (Spur 7) allerdings dieselbe Bande beobachtet wie bei den IMAC-aufgereinigten Fraktionen des Stammes mit dem Plasmid pBADtorAGFPHis. Zum finalen Beweis der Sekretion in das Periplasma für

das etablierte Sekretionsreportersystem mit dem Plasmid pBADtorAGFPSSrA, ist die Periplasmafraktion des entsprechenden Stammes auf Spur 9 aufgetragen. Wiederum ist deutlich eine Bande in der gesuchten Größenregion zu beobachten, womit eine erfolgreiche Sekretion in das Periplasma bewiesen wäre. Zusammenfassend ist damit eindeutig verifiziert, dass eine Sezernierung des GFP-Konstruktes ins Periplasma erfolgt ist, wie die Gelergebnisse belegen. Ausschließlich die Periplasmafraktionen der Stämme mit den Plasmiden pBADtorAGFPHis und pBADtorAGFPSSrA zeigen auf dem SDS-Gel eine intensive Bande im Größenbereich um 30 kDa, nicht aber die Periplasmafraktion von pBADGFPHis. Mittels IMAC lässt sich diese Bande im Falle von pBADtorAGFPHis deutlich konzentrieren und aufreinen.

4.7.7 Bioreaktorkultivierung

Nachdem der Beweis für die Funktionalität des etablierten Reportersystems für die Proteinsekretion rekombinanter Proteine erbracht war, wurde die differentielle Genexpression unter Bildung und Sekretion des rekombinanten Produktes untersucht. Zu diesem Zweck wurde der Stamm BL21 pBADtorAGFPSSrA im 10L-Maßstab im Biostat C (Sartorius Stedim GmbH) Bioreaktor kultiviert, um reproduzierbare Bedingungen für die Genexpressionsanalyse zu generieren. Als Kohlenstoffquelle fungierte anstelle der üblichen Glucose Glycerin, da die Arabinose-Induktion einer Katabolitrepression durch Glucose unterliegt (Guzman *et al.* (1995)). Die übrigen Medienbestandteile sind im Anhang (Abschnitt B) angegeben, es handelt sich um ein chemisch definiertes Medium mit der Bezeichnung CGSXIIIFS. Die Induktion der Kultur erfolgte in der Mitte der exponentiellen Phase bei einer OD600 von ca. 3 mit 0,1 Prozent *L*-Arabinose pro OD600. Es wurde direkt vor der Induktion eine Referenzprobe entnommen. Bezogen auf den Zeitpunkt der Induktion wurden dem Reaktor 15, 30, 45, 60, 90, 120 und 180 Minuten nach dem Induktionsvorgang Proben entnommen. Um eine Degradation der mRNA zu verhindern und um den jeweiligen Genexpressionszustand der Zellen möglichst schnell zu konservieren, wurde jede Probe für die RNA-Isolierung sofort nach Entnahme aus dem Reaktor in eine eiskalte fünfprozentige phenolische Lösung (in Ethanol) zugesetzt und die Probe im unmittelbaren Anschluss für 15 Minuten bei 3300 g zentrifugiert. Der Medienüberstand wurde entfernt und das Zellpellet sofort in flüssigen Stickstoff überführt. Nachdem alle Proben entnommen und so prozessiert wurden, erfolgte die Transferrierung und Lagerung bei -80°C. Parallel zu den Proben für die RNA-Extraktion wurden Proben zur Bestimmung der optischen Dichte und des Glucosegehalts des Mediums entnommen. Die Online-Analytik erfolgte mittels 2D-Fluoreszenzspektrometrie (Bioview).

4.7.8 RNA-Isolation

Um eine maximale Reproduzierbarkeit der RNA-Extraktion zu generieren, erfolgte die RNA-Isolierung mittels eines kommerziell erhältlichen RNA-Isolierungskits der Firma Analytik Jena. Von der so gewonnenen RNA wurden spektrophotometrisch Konzentration und Reinheit bestimmt sowie deren Integrität mittels Agarosegelelektrophorese überprüft. Da DNA-Kontaminationen bei der spektrophotometrischen Konzentrationsbestimmung der RNA zu einem signifikanten Fehler führt, wurden die RNA-Isolationen einer DNase I-Behandlung unterzogen. Zu diesem Zweck wurden der RNA-Extraktion pro 50 µg RNA 100 Units DNase I (Roche) zugesetzt und für 60 Minuten bei 37°C inkubiert. Anschließend erfolgte eine Phenol-Chloroformextraktion zur quantitativen Entfernung der DNase I aus dem Reaktionsansatz, da diese mit weiteren Applikationen interferiert. Die so präparierte RNA konnte für verschiedene Applikationen im Bereich der Genexpressionsanalysen verwendet werden.

Eine Verifikation der Integrität erfolgte wiederum mittels Gelelektrophorese und die Konzentration und Reinheit wurden am Spektralphotometer Nanodrop ND 1000 bestimmt. Im Folgenden werden die Ergebnisse der *low-density* Microarray-Analysen präsentiert welche auf der Verwendung dieser RNA basierten.

4.7.9 Transkriptionelle Analyse des Stammes pBADtorAGFPSsrA

Eine Analyse mittels applikationsspezifischer Microarrays bietet gegenüber der Verwendung von kommerziellen *whole-genome* Chips viele Vorteile. Neben ökonomischen Faktoren bieten *low-density* Chips die Möglichkeit, eine größere Anzahl von Genreplikaten auf den Chip zu platzieren. Somit können Positionseffekte besser detektiert werden und das Microarray-Experiment unterliegt einer signifikanteren Statistik (es besteht z.B. die Möglichkeit einen Ausreißertest anzuwenden). Die Reduktion auf eine spezifische Fragestellung und die damit einhergehende Verkleinerung des Umfangs an analysierten Genen geht oft auch einher mit einer gesteigerten Sensitivität. Das Phänomen, dass mittels einer *low-density* Analyse mehr regulierte Gene identifiziert werden, findet auch Niederschlag in der Literatur (Scheina *et al.* (1995)). Bedingt durch die ökonomischen Einsparungen ist zudem die Durchführung mehrerer technischer Replikate eines Experimentes möglich. Dieses eröffnet wiederum die Möglichkeit statistisch signifikantere Methoden zur Evaluierung der differentiell exprimierten Gene anzuwenden. Der im Rahmen dieser Studie etablierte, sekretionsspezifische *low-density* Chip umfasste 109 genspezifische Oligonukleotide, welche jeweils in 5 Replikaten auf die Chipoberfläche appliziert wurden. Die Auswahl der repräsentierten Gene erfolgte basierend auf den Ergebnissen der *whole-genome* Experimente und gängigen Literaturangaben zu sekretionsinvolvierten Genen. Für das Design der 50 mer-Oligonukleotide fanden die von Kane *et al.* (2000) eingeführten Kriterien Berücksichtigung. Diese Gensonden wurden von der Firma MWG synthetisiert und kommerziell erworben. Das *Spotting* der Oligonukleotide auf die Chipoberfläche aldehydfunktionalisierter *Glasslides* (Telechem) erfolgte mittels der Affymetrix 417 Arrayers. Stellvertretend wird die Applikation des *low-density* Microarrays anhand der transkriptionellen Analyse der Tat-spezifischen Sekretion ins Periplasma mittels des Stammes XL1 pBADtorAGFPSsrA gezeigt. Ein Schema der Translokation des GFP-Reporterproteins ist in Abbildung 53 gezeigt. Zur Evaluierung einer Kinetik der differentiellen Genexpression nach dem Induktionsvorgang mittels Microarray-Analyse wurde die RNA gemäss dem *TSA-Labeling and Detection Kit* (Perkin Elmer) prozessiert. Es wurde ein *Loop-Design* gewählt, um so mit minimalem Ressourcenaufwand die Genexpression der Zeitpunkte 0, 15, 30, 45, 60, 90, 120, 180 und 240 Minuten relativ zum Zeitpunkt der Induktion erfassen zu können. Das Prozedere der weiteren experimentellen Durchführung kann dem Anhang entnommen werden (A.1.5). Nach dem Scannen der resultierenden Arrays erfolgte die Primär- und Sekundäranalyse nach der im Anhang für die *low-density* Microarrays beschriebenen Methode.

Die generierten Ergebnisse der *low-density* Microarrays werden im Folgenden dargestellt und interpretiert. Die einzelnen, differentiell exprimierten Gene der aufgenommenen Genexpressionskinetik sind bei GEO (Gene Expression Omnibus) hinterlegt.

Generell lässt sich ein signifikanter Einfluß der Induktion durch *L*-Arabinose und die damit einhergehende Expression des torAGFPSsrA-Konstrukts auf die Transkriptionsrate der sekretionsspezifischen Gene beobachten. Ein besonders hohes differentielles Expressionsniveau weisen vor allem die Vertreter der positiv regulierten Gene auf. Es handelt sich dabei neben GFP um die Gene *lacA*, *nuoA*, signalpeptidase 2, *aphA*, *tdcA* und zu späteren Zeit-

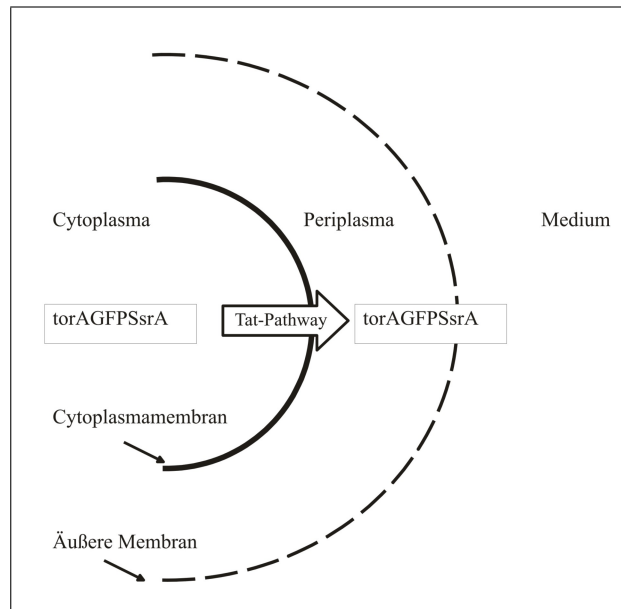


Abb. 53: Schema der Sekretion des Reporterproteins GFPsSrA des Stammes XL1 pBADtorAGFPSsrA. Das Reporterprotein wird mittels des Tat-Pathways in das Periplasma von *E. coli* transloziert.

punkten nach Induktion die Gene, welche für DnaJ und DnaK codieren. Die starke positive Regulation von GFP kann in diesem Fall als Positivkontrolle fungieren, durch die Induktion mit *L*-Arabinose wird die Transkription des plasmidcodierten GFP-Konstrukts initiiert. Aufgrund der Promoterstärke und der niedrigen basalen Transkriptionrate jener Gene, die stromabwärts des Arabinosepromoters lokalisiert sind, wird eine starke positive Regulation beobachtet. Das Gen *lacA* codiert für die β -Galactosid-Transacetylase, ein Enzym das eine Acetylgruppe von Acetyl-CoA auf β -Galactoside überträgt. Die *lacA*-mRNA ist sehr kurzlebig mit einer Halbwertszeit von ca. 3 Minuten (Lewin *et al.* (1991)), so dass eine sehr schnelle Induktion und Repression der Transkription stattfinden kann. Die positive Regulation des *lacA*-Genes geht sinnvoll mit der negativen Regulation des Repressors des *lac*-Operons *lacI* und der positiven Regulation der β -Galactosidase, codiert von *lacZ*, einher. NuoA codiert für eine Untereinheit der NADH-Ubiquinon-Oxidoreduktase I. Dabei handelt es sich um eine NADH-Dehydrogenase, welche den Transfer von Elektronen von NADH zu den Quinon-Pool in der Cytoplasmamembran vermittelt und ein Teil der anaeroben und aeroben Atmungskette der Zelle darstellt. Aus sekretorischer Sicht besonders interessant ist die Funktion als Protonenpumpe, welche durch einen transmembranen Protonenfluss in das Periplasma den Protonengradienten aufrecht erhält, der über der Membran anliegt (Stolpe *et al.* (2004)). Dieser pH-Gradient ist die Triebkraft der Translokation mittels des Tat-Pathways (Palmer *et al.* (2003)). Bedingt durch die intensive Nutzung des Tat-Pfades durch das rekombinant exprimierte GFP-Konstrukt scheint die positive Regulation der Komponente, welche für die Triebkraft der Translokation verantwortlich ist, biologisch höchst sinnvoll. Die in das Periplasma sezernierten GFP-Moleküle werden nachfolgend von der Signalpeptidase 2 proteolytisch prozessiert, so dass sie ihr Signalpeptid verlieren und in der reifen Form vorliegen (Blaudeck *et al.* (2001)). Die positive Regulation des für die Signalpeptidase 2 codierenden Genes deckt sich daher mit den zuvor getroffenen Aussagen und deutet auf einen Stimu-

lus durch die GFP-Sekretion und dessen Akkumulierung im Periplasma hin. Drei bzw. vier Stunden nach Induktion wird keine Regulation dieses Genes mehr beobachtet. Dies lässt sich durch den metabolischen Abbau des Induktors und der damit einhergehenden sinkenden GFP-Expression erklären. 45 Minuten nach Induktion beginnt die positive Regulation der beiden Chaperone DnaK und DnaJ, bis sie 180 und 240 Minuten nach Induktion negativ reguliert werden. Durch die zunehmende Akkumulierung des GFP-Konstrukts vor dem Export im Cytoplasma wird die positive Regulation dieser beiden Gene verständlich. Substrate des Tat-Pfades müssen vor der Translokation obligat gefaltet vorliegen. Bei den Chaperonen handelt es sich um eine Proteinklasse, welche ribosomal neu gebildeten Proteinen beim Faltungsprozess assistiert und auch inkorrekt gefalteten Proteinen zum richtigen Faltungsmuster verhelfen kann. DnaK und DnaJ werden auch unter dem Einfluß eines Hitzeschocks vermehrt exprimiert, um den erschwerten Faltungsbedingungen von Proteinen zu begegnen (Harcum *et al.* (2006)). Die zunehmende Akkumulierung des GFPs, und damit auch die zunehmende Existenz von fehlgefalteten Proteinen erklärt den etwas zeitverzögerten Effekt der positiven Regulation dieser beiden Chaperone. Durch eine Saturierung der Tat-Maschinerie könnte die vermehrte Bildung der Chaperone auch einem Schutzmechanismus vor dem proteolytischen Abbau darstellen, wie er für Chaperone in Verbindung mit Tat-Substraten häufiger propagiert wurde (Brüser *et al.* (2007)). Mittels Coexpressionstrategien von DnaK, DnaJ und TorD konnte experimentell eine signifikant höhere Proteinsekretion mittels des Tat-Pfades nachgewiesen werden (Perez-Rodriguez *et al.* (2007) und Li *et al.* (2003)). Dies manifestiert auch den außerordentlich Wert der Microarray-Technologie als *Screening-Tool*, da diese Gene, wie gezeigt wurde, zu den am stärksten regulierten in der durchgeführten Genexpressionskinetik des Stammes XL1 pBADtorAGFPSrA gehören. Das Gen *apHA* codiert für eine periplasmatisch lokalisierte, acidische Phosphatase. Diese katalysiert die Dephosphorylierung verschiedener organischer Phosphomonoester wie z.B. 3'- und 5'-Nukleotide. Daneben besitzt das Enzym eine Phosphotransferaseaktivität und fungiert zum Transfer von Phosphatgruppen niederenergetischer Phosphomonoester zu Akzeptoren mit freien Hydroxylgruppen. Die periplasmatische Lokalisation deutet auf einen möglichen Zusammenhang mit der Akkumulierung des GFP-Konstruktes im Periplasma hin. Ab 60 Minuten nach Induktion findet die positive Regulation eines Genes statt, welches direkt in den Arabinosemetabolismus involviert ist (*araC*). Es fungiert sowohl als Aktivator als auch als Repressor des Arabinosemetabolismus. Konstant positiv reguliert wird das Gen *ftsL*, dessen Produkt essentiell für das Wachstum und die Zellteilung ist. Da der Zeitpunkt der Induktion in der Mitte der exponentiellen Wachstumsphase lokalisiert ist, erscheint die Regulation dieses Genes biologisch sinnvoll. Da die Translokation des GFP-Konstruktes mittels des Tat-Pfades erfolgt, werden die verschiedenen Tat-Komponenten unter dem Induktionsstimulus im Folgenden detailliert betrachtet. Bedingt durch die limitierten Exportkapazitäten der Tat-Translokase wird die positive Regulation einiger Tat-Komponenten erwartet, um diesen *bottleneck* entgegenzuwirken. Zudem wurde eine positive Regulation einiger Tat-Komponenten experimentell in einigen Studien zur Untersuchung der Proteinsekretion mittels des Tat-Pathways beobachtet (De Lisa *et al.* (2002), Mergulaho *et al.* (2005)). Das Gen *tatA* ist leicht positiv reguliert. Bei den Vergleichen 45, 60, 90 und 120 Minuten nach Induktion gegenüber dem Zustand vor Induktion verschwindet dessen positive Regulation. Das *tatB*-Gen liegt in den meisten Vergleichen unreguliert vor, *tatD* ist leicht negativ reguliert. Eine Ausnahme in der Richtung der Regulation stellt *tatC* dar. In den meisten Vergleichen ist dieses Gen positiv reguliert, die stärkste positive Regulation liegt in den Vergleichen 90 und 120 Minuten nach Induktion gegenüber den Nullpunkt vor. *TatD* ist in den meisten

Vergleichen negativ reguliert und das Transkriptionsniveau von *tatE* ist schließlich gegenüber den präinduktiven Zustand leicht angehoben. Die teilweise negative Regulation der Tat-Komponenten überrascht. Die zentralen Komponenten des Translokationstunnels, bestehend aus TatABD, scheinen bei dem gewählten Induktionsniveau in ausreichender Zahl vorhanden zu sein. Lediglich die Transkriptionsrate der Komponenten TatC, wahrscheinlich verantwortlich für die spezifische Substraterkennung, und TatE werden leicht angehoben. Insgesamt überrascht diese Beobachtung und bedarf daher einer eingehenden Validierung durch weitere Methoden z.B. mittels quantitativer *real-time* PCR. Eine Erklärung könnte die Existenz des C-terminal lokalisierten Degradations-*tags* des Tat-Substrates GFP liefern. Durch die cytoplasmatische Degradation wird die effektive Konzentration des Substrats im Cytoplasma stets auf einem relativ niedrigen Niveau gehalten, so dass es zu keiner signifikanten Sättigung der Tat-Translokase kommt.

Die Diskussion der negativ regulierten Gene gestaltet sich als weitaus schwieriger, da zum einen das Niveau der Regulation geringer ist und zum anderen der biologische Sinn sich nicht immer offensichtlich erschließt. Konsistent negativ reguliert, bis auf die Zeitpunkte 180 und 240 Minuten nach Induktion, ist das Gen *glyA*. Die Serin-Hydroxymethyltransferase für welche das *glyA*-Gen codiert, konvertiert die Aminosäure Serin in Glycin durch Transferierung einer Methylgruppe zu Tetrahydrofolat, wodurch 5,10-Methylen-Tetrahydrofolat gebildet wird (5,10-mTHF). Bei dem 5,10-mTHF handelt es sich um die Hauptquelle für C1-Einheiten in der Zelle. Daher ist GlyA ein Schlüsselenzym für die Biosynthese von Purinen, Thymidinen, Methioninen, Cholin und Fetten. Experimentell wurde belegt, dass GlyA essentiell für das Wachstum auf Glycerin-basiertem Minimalmedium ist (Schirch *et al.* (1985)). Die Supplementierung von Arabinose durch die Induktion führt zu einem Wechsel des Kohlenstoffmetabolismus der Zelle von Glycerin zu Arabinose als C-Quelle. Durch diesen Effekt könnte die negative Regulation des nun nicht mehr essentiellen GlyA verstanden werden. Nachdem die Arabinose vollständig metabolisiert wurde, findet wieder ein Wechsel zum Glycerinmetabolismus statt und somit wird die negative Regulation aufgehoben. Konsistent mit dieser Beobachtung geht die Regulation des Genes *tdcA* einher, welches für den Aktivator des *tdcABCDEF*G-Operons codiert. Dieses Operon ist ebenfalls in der Degradation von Aminosäuren involviert, besonders dem der Aminosäure Serin (z.B. durch TdcB). Die Methylglyoxal-Synthetase katalysiert die Bildung von Methylglyoxal aus Dihydroxyaceton-Phosphat und ist ebenfalls ein Enzym welches in der Katabolisierung von Glycerin involviert ist (Lin *et al.* (1976)). Zusammenfassend lässt sich konstatieren, dass die negative Regulation der meisten Gene, bedingt durch den Induktionsstimulus, das Resultat der Umstellung des Kohlenstoffmetabolismus von Glycerin auf Arabinose ist.

Anhand der Microarray-Analyse mittels sekretionsspezifischer *low-density* Chips wurden einige für Optimierungsstrategien interessante, regulierte Gene identifiziert. Diese Ergebnisse stehen in Koinzidenz zu weiteren Studien (De Lisa *et al.* (2004), Lin *et al.* (2003), Perez-Rodriguez *et al.* (2007)), welche deren Potenzial für Coexpressionsstrategien zur Generierung von Sekretionsstämmen mit erhöhter Translokationskapazität aufzeigen konnten. Zudem beweisen die Ergebnisse den großen Wert der Microarraytechnologie als *Screening*-Methode.

4.7.10 Quantitative *real-time* PCR

Zur Verifizierung der generierten Daten aus den *low-density* Chipexperimenten wurde eine quantitative *real-time* PCR durchgeführt. Zur Etablierung der Methode wurde ein separater Arbeitsplatz geschaffen, welcher ausschließlich für diese Applikation genutzt wurde.

Aufgrund der enormen Sensitivität der Methode besteht ein hohes Kontaminationsrisiko, weshalb die Mastermixe und sämtliche benötigten Reagenzien nur an diesem separaten Arbeitsplatz hergestellt und verwendet wurden. *Real-time* Experimente zur Analyse von *E. coli* mRNAs unterliegen zudem noch der besonderen Schwierigkeit der ubiquitären Präsenz von *E. coli* DNA in biotechnologischen Laborräumen, aufgrund der weitverbreiteten Verwendung als Wirtsorganismus. Zusätzlich besteht im Gegensatz zu eukaryontischen Systemen nicht die Möglichkeit, Primer intronübergreifend zu designen, weshalb *real-time* Applikationen in Prokaryonten weit höhere Anforderungen stellen als dies für eukaryontische Systeme der Fall ist (Pfaffl *et al.* (2001)). Aus den angeführten Gründen wurden spezielle mit einem Filtereinsatz versehene nukleasefreie Pipettenspitzen verwendet. In vorangegangenen Experimenten hat sich gezeigt, dass die autoklavierten Pipettenspitzen eine Kontaminationsquelle darstellen und deshalb nicht verwendet werden sollten. Zur Herstellung der Mastermixe wurden ebenfalls nichtautoklavierte Reaktionsgefäße aus ungeöffneten Originalverpackungen verwendet. Diese sind, bedingt durch den Herstellungsprozess bei sehr hohen Temperaturen (über 200°C), frei von DNA-Kontaminationen.

Die Gene welche nach der Induktion mit *L*-Arabinose mittels der Chipdaten als differenziell experimentiert identifiziert wurden, bildeten den Hauptteil der mittels *real-time* PCR näher untersuchten Kandidaten. Zumeist waren dies positiv regulierte Gene. Die bekanntermaßen starke Regulation des unter der Kontrolle des Arabinosepromoters stehenden GFP-Genes kann dabei als Positivkontrolle fungieren. Weiterhin als stark reguliert identifiziert wurden die Gene, welche für Signalpeptidase 2, LacA und DnaK codieren. Die Gene *tatD*, *rpoA*, *secD* und *dnaA* waren schwach bis mittelstark positiv reguliert. Zu Kontrollzwecken wurden ebenfalls nicht-regulierte Gene mit in die Auswahl integriert. Für diesen Zweck fungierten die Gene *tatA* und *tatE*, zumal deren Transkriptionsstatus als Komponenten der Translokase besonderes Interesse galt. Das Gen, welches für die Serin Hydroxymethyltransferase (*glyA*) codiert, wurde als Vertreter der nach Induktion negativ regulierten Gene mit in die Auswahl integriert. Im Rahmen der quantitativen *real-time* Analyse wurden zwei verschiedene Normalisierungsverfahren evaluiert. Zum einen sollte die Normalisierung über zwei *housekeeping*-Gene erfolgen, zum anderen mittels eines artifiziellen RNA-Standards. Als Vertreter der *housekeeping*-Gene wurden *gapdH* und *gst* verwendet. Bei einem Vergleich verschiedener *housekeeping*-Gene für *real-time* Applikationen haben sich diese als besonders geeignet erwiesen, da ihr Transkriptionsniveau unter verschiedensten Bedingungen relativ konstant blieb (Radonic *et al.* (2004)). Zusätzlich werden sie in natürlichen Zellen stark exprimiert und ergeben somit nach wenigen PCR-Zyklen ein Fluoreszenzsignal, welches für eine sensitive Analyse von Bedeutung ist. Generell wird die Existenz von *housekeeping*-Genen kontrovers diskutiert, da deren Expressionsniveau unter bestimmten Umweltbedingungen doch variiert. Die Verwendung eines *housekeeping*-Genes für einen bestimmten Zweck sollte immer verifiziert werden, besser geeignet ist jedoch die Analyse mit mehreren *housekeeping*-Genen als Normalisierungsstandard. Moderne statistische Applikationen gewichten den Einfluß von verschiedenen *housekeeping*-Genen unterschiedlich. Je nachdem wie stark deren Expressionsniveau innerhalb eines Vergleichs variiert, wird der Normalisierungseinfluß auf die Analyse gewichtet (Van der Sompele *et al.* (2006)). Die Verwendung einer artifiziellen RNA-Spezies stellt eine weitere Möglichkeit zur Normalisierung dar. Gleiche Mengen dieses Standards werden der RNA-Isolation oder den Zellpellet vor der RNA-Isolation zugesetzt und anschließend zusammen mit der zelleigenen RNA-Spezies in cDNA konvertiert. Dieser als schwankungsintensiv identifizierte Prozessschritt wird somit vom Standard mit durchlaufen, und Variabilitäten können so aufgedeckt werden. Die Normalisierung mittels

Tabelle 5: Die experimentell ermittelten, optimalen *Annealing*-Temperaturen für die verschiedenen Primerpaare für quantitative *real-time* PCR-Analysen.

Primerpaar	<i>Annealing</i> -Temperatur
gfp	62°C
rfp	65°C
gst	63°C
gapdh	63°C
tatA	62°C
tatD	64°C
tatE	62°C
secD	63°C
dnaA	64°C
dnaK	62°C
B0421	63°C
glyA	63°C
lacA	62°C

artifizierlicher RNA umgeht einige Problematiken die sich bei der Verwendung von *housekeeping*-Genen als Normalisierungsstandard ergeben. In diesem konkreten Fall fungiert eine mittels *in vitro* Transkription generierte rfp-RNA als Standard. Für alle untersuchten Gene wurden sequenzspezifische Primer *in silico* designed und bei der Firma MWG geordert. Die Amplifikate dieser Primer sollten im Idealfall zwischen 100-200 bp lang sein, da somit eine hohe Amplifikationseffizienz, einhergehend mit einer hohen Detektionssensivität erzielt werden kann (Wilhelm *et al.* (2003)). Zudem kommt es bei kurzen Amplifikaten zu signifikant weniger Abbruchsreaktionen bei der Primer-Elongation durch die *Taq*-Polymerase. Kurze Primer erhöhen somit die Präzision und Sensitivität der Analyse, besonders beim Vergleich der Ergebnisse zwischen Amplifikaten, welche mit verschiedenen Primern generiert wurden. Für das Primerdesign wurden die gängigen Kriterien gewählt, wie sie in Standardwerken der Molekularbiologie angegeben werden (Sambrook: Molecular Cloning (2003)). Alle verwendeten Primer sind im Anhang (Abschnitt C.3) angegeben.

Als Template für die Amplifikationen wurde cDNA aus 2 μg total RNA mittels der Improm M-MLV Reversen Transkriptase (Promega) generiert. Die RNA wurde zuvor durch 60 minütige Inkubation mit DNase I bei 37°C von chromosomalen und Plasmid-DNA-Kontaminationen befreit und das Enzym durch Phenol-Chloroform Extraktion quantitativ entfernt. Die so gewonnene cDNA wurde 1:20 in nukleasefreiem Wasser verdünnt und als Arbeitslösung verwendet bzw. daraus die Verdünnungsreihen angesetzt. Die Zugabe des cDNA-*Templates* zu den Mastermischen erfolgte separat und getrennt von den Räumlichkeiten zur Präparation der Mastermische. Dies stellte den letzten Schritt zur Präparation des Amplifikationsansatzes dar. Als Mastermix wurde der IQ-SybrGreen Mastermix der Firma Biorad verwendet, welcher mit den jeweiligen Primern und nukleasefreiem Wasser versetzt und in eine spezielle 96-Loch-Platte pipettiert wurde. Die Zugabe des *Templates*, oder des Wassers bei den Negativkontrollen, erfolgte als letzter und wiederum räumlich getrennter

Schritt.

Zuerst erfolgte die Evaluation der jeweiligen optimalen *Annealing*-Temperaturen für die verschiedenen Primerpaare. Das verwendete PCR-Protokoll ist dem Anhang zu entnehmen (Abschnitt A.1.18). Für die verschiedenen verwendeten Primerpaare ist die optimale *Annealing*-Temperatur in Tabelle 5 zusammengefasst.

Für nahezu alle Primerpaare liegt diese zwischen 62°C und 64°C, es konnten daher uniforme Bedingungen eingehalten werden. Beispielhaft ist das Ergebnis einer Optimierung in der Abbildung 54 gezeigt. Die nachfolgend durchgeführte Schmelzkurvenanalyse des Amplifikats dient zur Identifizierung von Primerdimeren und unspezifischen Coamplifikaten, welche unter allen Umständen vermieden werden müssen.

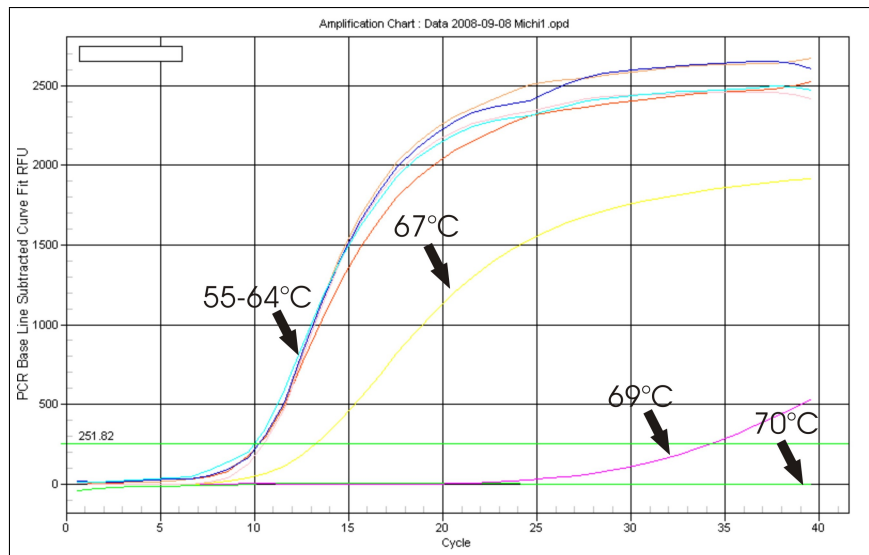


Abb. 54: Optimierung der *Annealing*-Temperatur für das GFP-Primerpaar. Gezeigt sind die resultierenden Amplifikationskurven für verschiedene *Annealing*-Temperaturen. Die Temperatur, welche den niedrigsten Ct-Wert und keine unspezifischen Coamplifikate generiert, wird als optimal betrachtet.

Diejenige *Annealing*-Temperatur, welche den kleinsten Ct-Wert generiert und dabei keine Bildung von unspezifischen Coamplifikaten zeigt, wurde als optimal für das jeweilige Primerpaar angesehen. Die Identifikation des Ct-Wertes erfolgte aus der logarithmischen Auftragung der Amplifikationskurven. Innerhalb des linearen Bereichs dieser Auftragung kann der Schwellenwert der Signalintensität gelegt werden. Eine direkte Identifizierung des Ct-Wertes anhand der Amplifikationskurven gestaltet sich häufig als schwierig, da der Übergang zwischen der exponentiellen und der stationären Amplifikationsphase meist nicht richtig erkannt wird (Abbildung 56). Ein niedriger Ct-Wert ist wünschenswert, um eine möglichst große Sensitivität der Analyse zu erzielen. Dabei handelt es sich zumeist um gegenläufige Tendenzen; eine niedrige *Annealing*-Temperatur generiert niedrige Ct-Werte. Dies ist gleichbedeutend mit einer hohen Sensitivität, die Tendenz zur Bildung unspezifischer Coamplifikate steigt jedoch, da auch nicht perfekt basengepaarte Hybridisierungen des Primers Amplifikate generieren können. Beispielhaft ist eine Schmelzkurvenanalyse in Abbildung 55 gezeigt.

Nachdem für alle Primerpaare die optimale *Annealing*-Temperatur evaluiert war, wurde die Abwesenheit von DNA-Kontaminationen in den RNA-Präparationen verifiziert. Die-

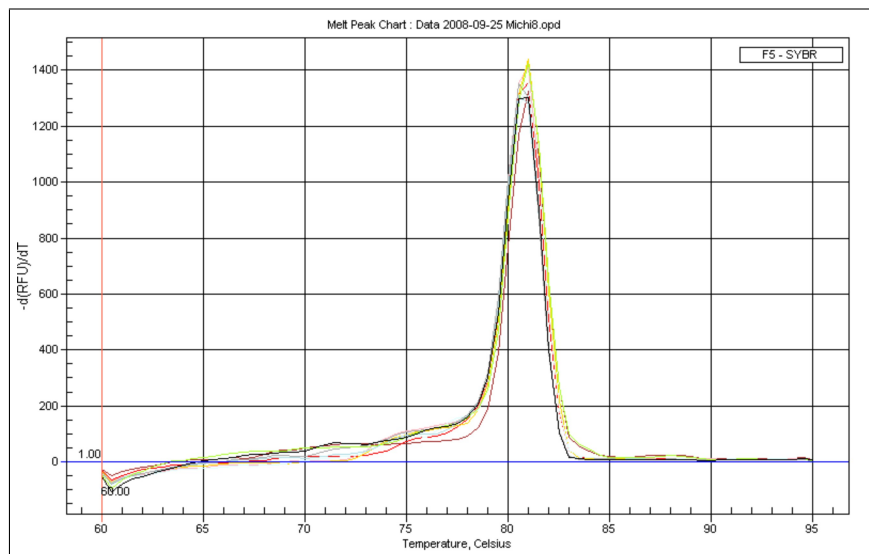


Abb. 55: Schmelzkurvenanalyse von Amplifikaten, welche aus der Optimierung der *Annealing*-Temperatur für GFP-spezifische Primer hervorgegangen sind. Gezeigt ist die erste Ableitung der Schmelzkurve. Klar ersichtlich ist, dass nur ein einziges Amplifikat generiert wurde, was auf eine spezifische Amplifikation schließen lässt.

ses wurden stichprobenartig für chromosomale DNA (repräsentiert durch GAPDH) und Plasmid-DNA (repräsentiert durch GFP) bei fünf verschiedenen RNA-Proben durchgeführt. Dazu wurde die RNA in analogen Konzentrationen eingesetzt wie sie zur Generierung der cDNA verwendet wurde, und als Template dem Amplifikationsansatz zugesetzt. Simultan wurde eine cDNA-Amplifikation als Positivkontrolle sowie Negativkontrollen für die verschiedenen Primerpaare durchgeführt. Alle RNA-Präparationen zeigten bei sehr hohen Ct-Werten die Bildung von Amplifikaten. Die Negativkontrollen generierten auch nach 35 Zyklen kein Amplifikat. Dies belegt, dass geringe Mengen an DNA-Kontaminationen trotz des durchgeführten DNase I-Verdau in der Lösung vorhanden sind. Beispielhaft ist dieses in Abbildung 57 gezeigt.

Ein Vergleich mit den Positivkontrollen zeigte jedoch, dass diese bereits zu weit früheren Zyklen Amplifikate generierten. Der kleinste Unterschied in den Ct-Werten betrug mindestens 16 Zyklen für das GAPDH-Amplifikat. Wenn man eine Effizienz von nur 1,6 für die PCR-Reaktion mit diesem Primerpaar zugrunde legt, bedeutet dies, dass die DNA-Konzentration in der RNA-Präparation ca. 1840 mal geringer ist, als selbige in der generierten cDNA. Bei einer Effizienz von 1,9 beträgt diese sogar nur 1/28844 des Wertes in der cDNA. Der dadurch begangene Fehler liegt weit unterhalb der Variabilitäten, welche sich durch das Pipettieren oder für die Analysemethode an sich ergeben.

Aus diesem Grunde wurde auf eine zusätzliche DNase I-Behandlung der RNA verzichtet, da weitere Prozessschritte immer die Möglichkeit beinhalten, die relativen Verhältnisse der Analyten zueinander zu verändern bzw. produktspezifische Verluste zu generieren. Aufgrund des außerordentlich geringen Fehlers, welcher durch die minimale DNA-Kontamination der RNA verursacht wird, sollte diese Vorgehensweise robuste, reproduzierbare Daten generieren.

Um die verschiedenen Effizienzen der verwendeten Primerpaare zu ermitteln, wurde aus einem cDNA-*Template*, welches initial 1:20 verdünnt wurde, zunächst eine 1:10 Verdün-

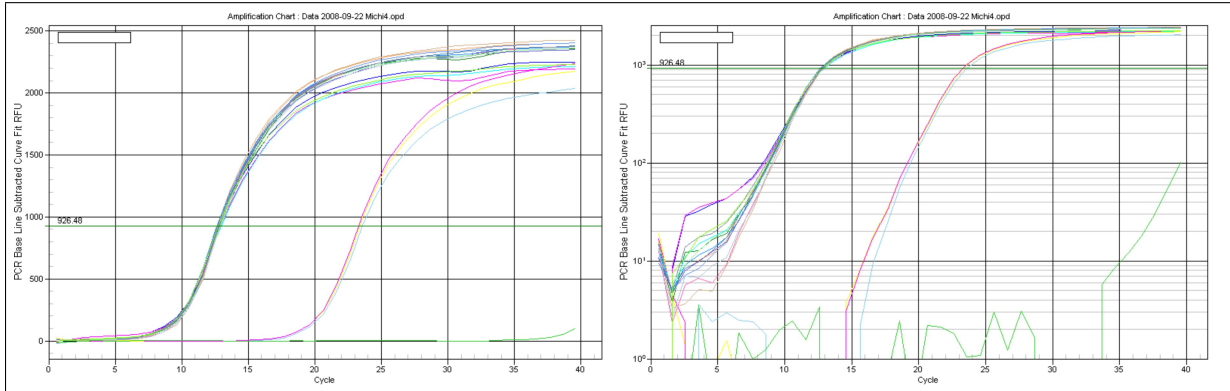


Abb. 56: Mögliche Fehlerquellen bei der Festlegung des Schwellenwertes der Fluoreszenz zur Bestimmung des Ct-Wertes (grüne Linie). Der scheinbar in der exponentiellen Phase befindliche Schwellenwert (linke Seite) ist schon am Übergang zur stationären Phase der Amplifikation situiert, wie die logarithmische Auftragung auf der rechten Seite zeigt.

nungsreihe erstellt. $5 \mu\text{l}$ der jeweiligen Verdünnung fungierten als *Template* für die PCR-Reaktion mit den unterschiedlichen Primern. Jede unterschiedliche Amplifikation wurde als Triplikat durchgeführt, um reproduzierbare Ergebnisse zu generieren. Exemplarisch sind die Amplifikationskurven für das GFP-Primerpaar in Abbildung 58 gezeigt.

Der äquidistante Abstand zwischen den einzelnen Kurven ist leicht zu erkennen. Die Replikate der jeweiligen Amplifikationen zeigen geringe Variabilitäten und ihre Amplifikationskurven liegen fast perfekt übereinander. Aus der Auftragung des Ct-Wertes gegen den Logarithmus der (relativen) Konzentration kann aus der Steigung die Effizienz der Amplifikation ermittelt werden (siehe Kapitel 3.6). Es sollte sich eine lineare Gerade ergeben, deren Korrelationskoeffizient ein Maß für die Güte der generierten Ergebnisse darstellt. Wiederum ist dies stellvertretend durch die GFP-Amplifikation in Abbildung 59 gezeigt.

Der Korrelationskoeffizient ist mit 0,9951 sehr hoch und bestätigt die Güte der Analyse. Aus der Geradensteigung von -3,521 ergibt sich eine Amplifikationseffizienz von 1,923. Bei einer maximalen Effizienz von 2, d.h. einer Produktverdoppelung je Zyklus, stellt dies einen sehr hohen Wert dar. Dies gibt auch Hinweise auf ein gutes Primerdesign und optimale Amplifikationsbedingungen. Die ermittelten Geradensteigungen und Korrelationskoeffizienten, sowie die sich daraus ergebenden Amplifikationseffizienzen sind für jedes Primerpar in der Tabelle 6 angeführt.

Nachdem die Effizienzen der verschiedenen Primerpaare evaluiert wurden, erfolgte die Analyse der verschiedenen generierten cDNA-Ansätze, welche Proben zu unterschiedlichen Fermentationszeitpunkten vor und nach Induktion repräsentierten. RNA-Präparation welche die Zeitpunkte 0, 15, 30, 45 und 60 Minuten, bezogen auf den Zeitpunkt der Induktion, repräsentierten, bildeten das Ausgangsmaterial für die quantitativen *real-time* Analysen.

Um Variabilitäten in dem Schritt der Reversen Transkription zu identifizieren, wurde jede RNA-Probe in drei unabhängigen Reaktionen in cDNA konvertiert. Zusätzlich wurden von der Probe, welche 15 Minuten nach Induktion dem Reaktor entnommen wurde, drei identische RNA-Präparationen erstellt und jeweils in drei unabhängigen Reaktionen in cDNA konvertiert. Hieraus resultierten neun cDNA-Präparationen aus drei verschiedenen RNA-Präparationen einer biologischen Probe. Somit lassen sich Variabilitäten ermitteln, welche

4.7 Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien

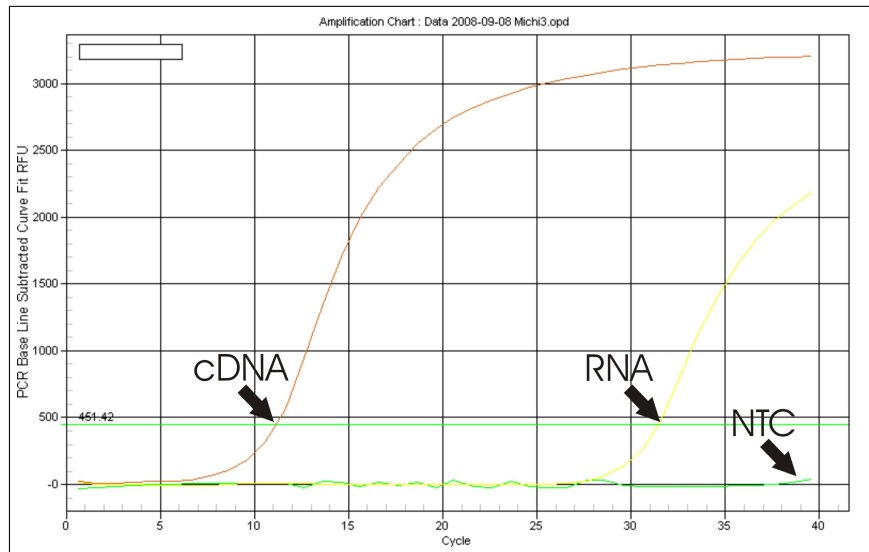


Abb. 57: Evaluierung hinsichtlich DNA-Kontaminationen in den RNA-Präparationen. Gezeigt sind exemplarisch die resultierenden Amplifikationskurven für eine PCR mit GAPDH-spezifischen Primern. Als *Template* fungieren die RNA-Präparation und im Vergleich dazu eine cDNA. Zusätzlich ist eine Negativkontrolle mit Wasser als *Template* gezeigt. Wie aus den Amplifikationskurven ersichtlich ist, sind noch vernachlässigbare Spuren von DNA in der RNA-Präparation enthalten. Aufgrund des großen Ct-Unterschiedes von mindestens 16 Zyklen interferieren diese jedoch in keiner Weise mit der nachfolgenden Analyse.

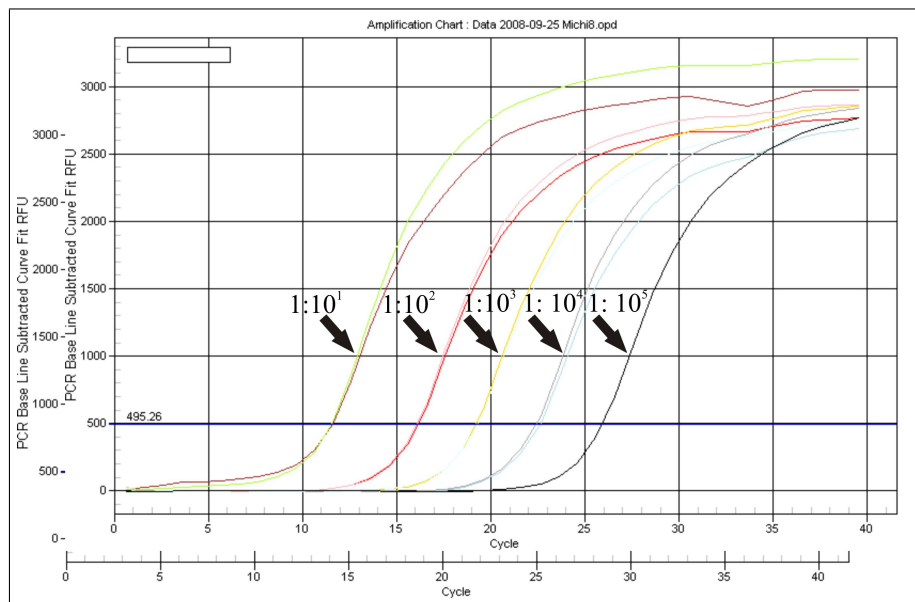


Abb. 58: Resultierende Amplifikationskurven für eine 1:10 Verdünnungsreihe des GFP-*Templates*. Deutlich ersichtlich ist der zumeist äquidistante Abstand der Amplifikationskurven (in Duplikaten) zwischen den einzelnen Verdünnungen.

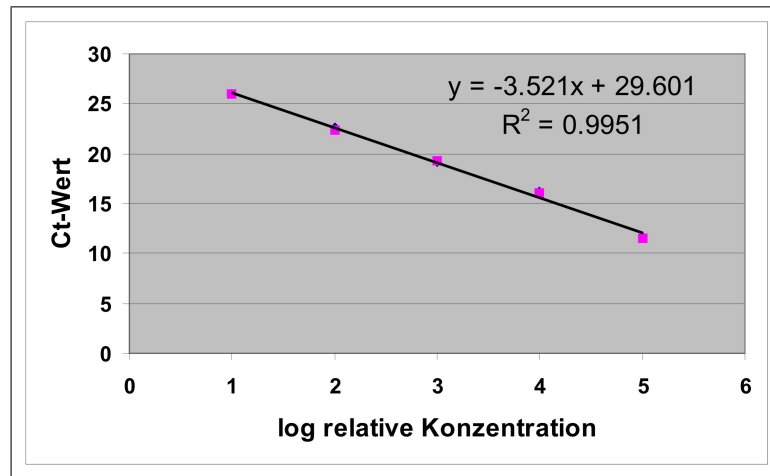


Abb. 59: Auftragung der ermittelten Ct-Werte gegen den Logarithmus der relativen Templatekonzentration für eine GFP-Verdünnungsreihe. Der Versuch wurde als Triplikat durchgeführt, dargestellt ist der Mittelwert aus diesen Versuchen.

Tabelle 6: Die ermittelten Gradensteigungen und Korrelationskoeffizienten der Regressionsgraden für die verschiedenen Verdünnungsreihen zur Evaluierung der Primereffizienzen.

Amplifikat	Steigung	Effizienz	Korrelationskoeffizient
lacA	-4,800	1,626	0,988
gst	-3,689	1,867	0,989
dnaK	-3,961	1,788	0,994
rfp	-4,548	1,659	0,985
gfp	-3,521	1,923	0,995
gapdH	-3,395	1,970	0,997
tatA	-3,667	1,874	0,990
B0421	-3,397	1,970	0,993
tatE	-3,770	1,842	0,986
secD	-3,428	1,958	0,992
glyA	-3,559	1,910	0,986
dnaA	-3,469	1,942	0,990
tatD	-3,617	1,890	0,994

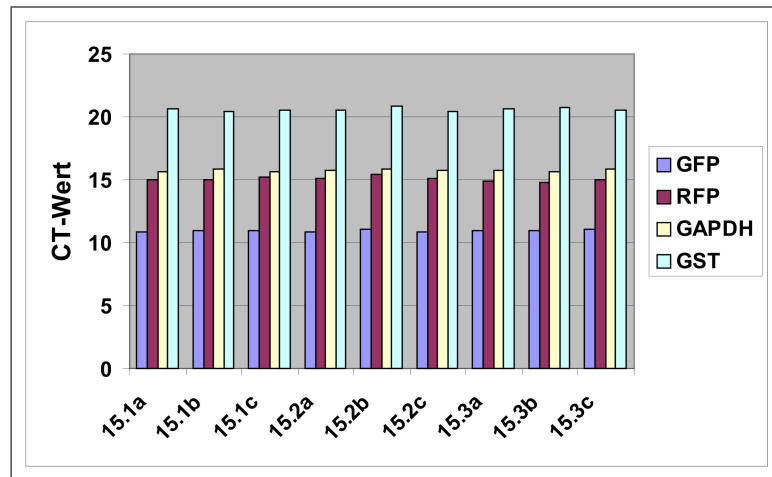


Abb. 60: Neun verschiedene cDNAs welche aus drei verschiedenen RNA-Präparationen einer biologischen Probe hervorgingen sind hinsichtlich des Ct-Wertes für vier verschiedene *Templates* und Primerpaare untersucht. Die hohe Reproduzierbarkeit der Analyse ist an den marginalen Unterschieden im Ct-Wert zu erkennen.

ihre Ursache in der RNA-Präparation haben, und ein Vergleich zwischen den Variabilitäten der beiden Prozessschritte durchführen. Die neun verschiedenen cDNA-Ansätze dienen als *Template* für Amplifikationen mit den Primern RFP, GFP, GAPDH und GST. Abbildung 60 zeigt das Ergebnis der ermittelten Ct-Werte für die unterschiedlichen cDNA-Ansätze.

Aus der Abbildung 60 ist zu entnehmen, dass die Ct-Werte für ein Primerpaar sowohl zwischen cDNAs, welche aus einer RNA-Präparation hervorgegangen sind als auch zwischen cDNAs, die aus verschiedenen RNA-Präparationen hervorgegangen sind, nahezu identisch sind und keine signifikanten Variabilitäten vorliegen. Diese Aussage gilt für alle vier Primerpaare. Da alle cDNA-Ansätze eine biologische Probe repräsentieren, ist dieses Ergebnis zu erwarten. Es belegt die außerordentlich hohe Reproduzierbarkeit und Güte der durchgeführten Analyse. Die Standardabweichung zwischen den drei verschiedenen cDNAs schwankt zwischen 0,043588 für GFP und 0,2165 Zyklen für GST; dies entspricht einer relativen Standardabweichung von 0,4 Prozent für GFP und 1,1 Prozent für GST. Interessant ist die Tatsache, dass für GST die höchsten Ct-Werte detektiert wurden, gleichbedeutend mit der niedrigsten Anzahl an mRNA-Kopien pro Zelle. Dies resultiert in der höchsten relativen Standardabweichung für die GST-Analyse. Umgekehrte Aussage gilt für die GFP-Analyse. Die Standardabweichungen zwischen den cDNAs unterschiedlicher RNA-Präparationen schwanken in der gleichen Größenordnung und sind nicht unterscheidbar von denen, welche zwischen cDNAs vorliegen, die aus einer RNA-Präparation hervorgegangen sind.

Nachdem bewiesen wurde, dass die Variabilitäten, welche aus der Verwendung unterschiedlicher RNA-Präparationen einer biologischen Probe herrühren, vernachlässigbar gering sind, wurden die nachfolgenden *real-time* Analysen für die Probe 15 Minuten nach Induktion lediglich mit drei cDNAs durchgeführt, welche aus einer RNA-Präparation hervorgegangen waren. Es wurden die Ct-Werte für alle neun Primerpaare, jeweils für die fünf verschiedenen Zeitpunkte, in Triplikaten bestimmt. Zur Normalisierung wurden die Analysen für die Primer rfp, gapdH und gst verwendet. Stellvertretend für gapdH werden die Amplifikationskurven aller Proben in Abbildung 61 gezeigt.

4.7 Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien

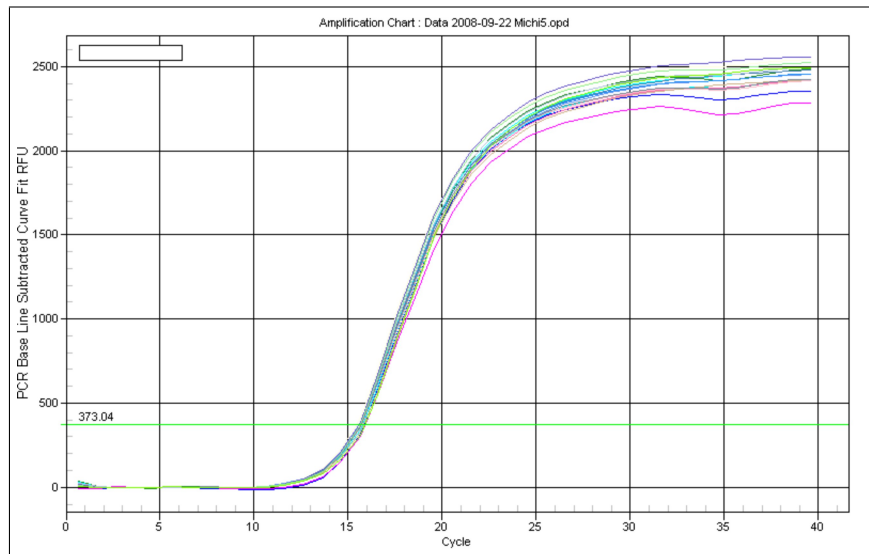


Abb. 61: Die resultierenden Amplifikationskurven bei Verwendung gapdH-spezifischer Primer für alle evaluierten biologischen Proben. Die Amplifikationskurven liegen nahezu perfekt übereinander.

Diese liegen nahezu perfekt übereinander und beweisen, dass gapdH ein gutes *housekeeping*-Gen für diese Applikation darstellt.

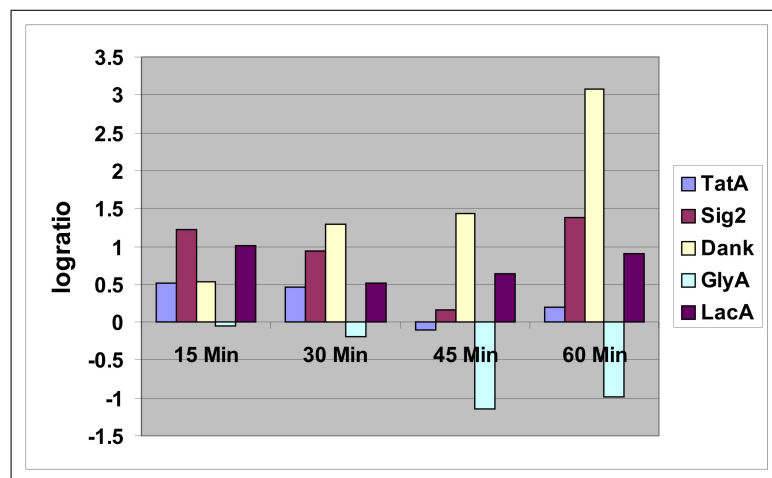


Abb. 62: Visualisierung der Regulation bezogen auf den Zeitpunkt vor Induktion für repräsentative Vertreter der untersuchten Gene für die verschiedenen evaluierten Zeitpunkte nach der Induktion.

Die Auswertung erfolgte gemäß der von Pfaffl *et al.*(2001) eingeführten Methodik. Die Regulation der Gene gfp, tatA, tatD, tatE, dnaK, dnaA, glyA und lacA relativ zum Nullpunkt der Induktion wurde mittels vier verschiedener Normalisierungen bestimmt. Es wurde über die Differenzen des Ct-Wertes von RFP, GAPDH, GST und einer Kombination von GAPDH und GST normalisiert. Die vier verschiedenen Ergebnisse sind, stellvertretend für drei Gene in der Tabelle 7 wiedergegeben. Wie ersichtlich, unterscheiden sich die ermit-

Tabelle 7: Vergleich der ermittelten Genexpressionsquotienten für drei Gene mittels vier verschiedener Normalisierungen relativ zum Zeitpunkt vor Induktion. Die Normalisierungen basieren auf der Verwendung der *housekeeping* Gene *gapdH* und *gst* sowie einer *gespikten* artifiziellen RNA-Spezies (*rfp*).

	RFP-Norm.	GST-Norm.	GAPDH-Norm.	GAPDH+GST-Norm.
tatA				
15	1,814	1,428	1,610	1,512
30	2,055	1,454	1,464	1,459
45	1,355	0,979	0,8354	0,902
60	1,570	1,260	1,130	1,192
B0421				
15	2,659	2,093	2,355	2,216
30	2,734	1,934	1,948	1,941
45	1,748	1,263	1,077	1,163
60	3,14	2,521	2,261	2,383
dnaK				
15	1,843	1,451	1,632	1,536
30	3,226	2,282	2,298	2,290
45	3,659	2,643	2,256	2,434
60	5,372	4,312	3,867	4,077

telten Werte, abhängig von der verwendeten Normalisierungsmethode. Die Richtung und der Größenbereich der Regulation stimmen jedoch nahezu immer überein. Auffällig ist, dass die Normalisierungsfaktoren, welche über die artifiziellen RFP-RNA ermittelt wurden, wesentlich größer sind und stärkeren Schwankungen zwischen den verschiedenen Proben unterliegen, als dies für die *housekeeping* Gene *gapdH* und *gst* der Fall ist. Ein Vergleich der Amplifikationskurven aller Proben für *gapdH* und *gst* belegt, dass die Unterschiede in den Ct-Werten zwischen verschiedenen Proben marginal sind. Im Wesentlichen trifft dies auch für *rfp* zu, jedoch sind die Schwankungen hier im Vergleich deutlich größer. Dass die nahezu perfekte Übereinstimmung der Ct-Werte der *housekeeping* Gene für die verschiedenen Proben aus einer zufälligen Verteilung resultiert ist statistisch höchst unwahrscheinlich. Dies deutet darauf hin, dass die *housekeeping*-Gene unter den gegebenen Bedingungen als Normalisierungstandard gut geeignet sind. Die Schwankungen der Ct-Werte verschiedener RFP-Amplifikationen und die im Vergleich zu *gst* und *gapdH* größeren Standardabweichungen deuten auf Variabilitäten hin. Somit werden im Folgenden die Ergebnisse (Tabelle 8) diskutiert, welche anhand der Normalisierung über die beiden *housekeeping*-Gene generiert wurden. Die Visualisierung der Regulation von fünf verschiedenen Genen, bezogen auf den Referenzzustand vor Induktion ist in Abbildung 62 gezeigt.

Das Gen *tatA* weist 15 und 30 Minuten nach Induktion eine ca. 1,5 fache höhere Expression auf als vor der Induktion. 45 und 60 Minuten nach Induktion bewegt sich das Transkriptionsniveau wieder auf dem Wert, welcher vor der Induktion vorherrschte. Die schwache positive Regulation des *tatA* Genes deutet auf einen möglichen Zusammenhang mit der Translokation des GFP-Substrates hin. Ein sehr ähnlicher Expressionsverlauf wurde

Tabelle 8: Regulation der zehn untersuchten Gene für die Zeitpunkte 0, 15, 30, 45 und 60 Minuten nach Induktion. Die Regulation ist bezogen auf den Zeitpunkt vor Induktion. Eine Normalisierung der Daten erfolgte mittels der von Pfaffl *et al.*(2001) entwickelten Methodik über die beiden *housekeeping*-Gene *gapDH* und *gst*.

Zeitpunkt relativ zur Induktion	tatA	B0421	dnaK	gfp	lacA
0 Min	1	1	1	1	1
15 Min	1,512	2,216	1,536	981,455	2,016
30 Min	1,459	1,941	2,290	1128,266	1,509
45 Min	0,902	1,163	2,434	927,006	1,633
60 Min	1,192	2,384	4,077	1032,203	1,900
Zeitpunkt relativ zur Induktion	tatE	dnaE	glyA	dnaA	tatD
0 Min	1	1	1	1	1
15 Min	1,040	1,406	0,952	1,642	2,029
30 Min	1,279	1,520	0,843	1,604	1,937
45 Min	0,857	0,783	0,465	0,930	1,118
60 Min	1,103	1,058	0,502	1,151	1,304

auch mittels der Microarray-Analyse beobachtet. TatA codiert für den zentralen, membran-gebundenen Teil der Tat-Translokase, welcher den Hauptbestandteil des Translokationstunnels bildet.

Gleiches gilt auch für das Gen *tatE*, dessen mRNA-Menge für die Zeitpunkte 15, 30, 45 und 60 Minuten nach Induktion zwischen dem 1,2 und 0,9 fachen des Wertes vor Induktion schwankt. Somit kann dieses Gen als nicht reguliert betrachtet werden, welches gut mit den Daten aus den Chipexperimenten korreliert. Anhand dieser wird das Gen als sehr schwach positiv reguliert identifiziert. Ein weiterer Teil der Tat-Translokase, *tatD* wies 15 und 30 Minuten nach Induktion eine mRNA-Menge auf, die dem 2,02 bzw. 1,93 fachen des Wertes vor Induktion entsprach. Für die Zeitpunkte 30 und 45 Minuten nach Induktion sank diese auf ein Niveau welches dem 1,12 bzw. 1,34 fachen des Vorinduktionsniveaus entsprach. Diese positive Regulation des *tatD*-Genes, welche 45 Minuten nach Induktion abklingt zeigt einen ähnlichen Verlauf wie jener für das *tatA*-Gen. *TatD* ist allerdings deutlicher reguliert. Insgesamt wird die positive Regulation der Tat-Komponenten unter dem Stimulus der Akkumlierung des zu sekretierenden rekombinanten Proteins erwartet. Viele Untersuchungen belegen, dass die Tat-Maschinerie bei der Sekretion rekombinanter Proteine rasch saturiert ist (De Lisa *et al.* (2004), Robinson *et al.* (2004), Mergulaho *et al.* (2005)). Als Folge dessen wurden verschiedene Tat-Komponenten als positiv reguliert identifiziert. Aufgrund des Degradations-*tags* des GFP-Konstruktes für das cytoplasmatische Clp-System und des moderaten Induktionsniveaus sind die cytoplasmatischen GFP-Konzentrationen gering. Aus diesem Grund sollte die Tat-Translokase nicht in hohem Maße mit Substrat saturiert sein und die positive Regulation der Tat-Komponenten nur schwach ausfallen, was im Wesentlichen auch beobachtet wird. Die stark positive Regulation des GFP-Genes auf das ca. 1000 fache des Transkriptionsniveaus vor Induktion kann als Positivkontrolle fungieren. Dieser

Wert deckt sich gut mit jenen Daten, welche aus Experimenten zur Evaluierung der Stärke des Arabinosepromoters stammen. Für die meisten stromabwärts des Arabinosepromoters lokalisierten Gene wurde eine Regulation auf das 1000-1500 fache Niveau vor Induktion ermittelt (Guzman *et al.* (1995)). Die sehr deutliche Verschiebung des Ct-Wertes nach Induktion wird in der Abbildung 63 gezeigt.

Mittels der Microarray-Experimente wurde die stark positive Regulation ebenfalls detektiert, allerdings ist der ermittelte Genexpressionsquotient vor/nach Induktion mit ca. 100 um den Faktor 10 niedriger. Darin spiegelt sich ein allgemein beobachteter Trend wieder, wonach die mit Microarray-Experimenten ermittelten Genexpressionsquotienten deutlich geringer sind als die vergleichbaren Werte aus quantitativen Genexpressionsanalysen mittels *real-time* PCR. Gründe dafür liegen in den grundsätzlichen Unterschieden einer Hybridisierung komplementärer Nukleinsäuren in homogener Lösung und an einem immobilisierten *Target* an einer heterogenen Phasengrenzfläche. Aufgrund sterischer und elektrostatischer Abstoßung zwischen dem Analyten sinkt die Hybridisierungseffizienz. Außerdem können aufgrund der großen Anzahl von Fluorophoren in räumlicher Nähe *Quenching*-Effekte in Erscheinung treten. Besonders für hochexprimierte *Targets* besitzt dieser Effekt Relevanz.

Im Folgenden wird die Regulation von zwei *E. coli* Genen diskutiert, welche mittels der Microarray-Daten als positiv differentiell exprimiert (*lacA* und signalpeptidase 2) identifiziert wurden. Das Gen *lacA*, welches für die Thiogalactosid Acetyltransferase codiert, wird 15 Minuten nach Induktion ca. zweifach hochreguliert. Diese Regulation gegenüber dem präinduktiven Grundzustand bleibt auch für die drei weiteren Zeitpunkte nach Induktion bestehen. Analoges kann aus den Microarray-Daten abgeleitet werden; diese weisen auf eine konstante positive Regulation des Genes hin. Das Gen *lacA* codiert für ein Enzym welches in der Metabolisierung verschiedener Zuckerderivate durch Übertragung einer Acetylgruppe involviert ist. Durch die Zugabe von *L*-Arabinose wird der Kohlenhydratstoffwechsel der Zelle von Glycerin auf Arabinose als C-Quelle umgestellt. Zur Metabolisierung des Induktors könnte u.a. auch die Thiogalactosid Acetyltransferase involviert sein, wodurch die positive Regulation zu erklären ist. Die Signalpeptidase 2 fungiert als prozessierendes Enzym bei der Proteolyse der Signalsequenzen von ins Periplasma sezernierten Tat-Substraten (Berks *et al.* (2000)). Die gefundene mRNA-Menge steigt nach Induktion auf das 2,22 fache ihres Wertes vor Induktion an und bleibt in etwa auf diesem Niveau. Durch die Translokation des rekombinanten GFP-Konstruktes steigt der Bedarf der Zelle an diesem Enzym über das natürliche Niveau heraus an. Die konsistente positive Regulation dieses Genes wird unter dem Aspekt der Akkumulierung des unprozessierten GFP-Proteines im Periplasma verständlich.

Eine ebenfalls positive Regulation wird für das Gen *dnaK* identifiziert. Dessen Transkriptionsniveau, verglichen mit dem Zustand vor Induktion, steigt sukzessiv von dem 1,54 fachen Wert 15 Minuten nach Induktion auf den ca. vierfachen Wert 60 Minuten nach Induktion an. Ein ähnlicher Verlauf wird auch mittels der Microarray-Analyse beobachtet. *DnaK* codiert für ein Hitzeschock-Chaperon, welches maßgeblich in der Faltung von Proteinen unter harschen Umweltbedingungen involviert ist (Zhao *et al.* (2005)). Durch die zunehmende Akkumulierung von GFP im Cytoplasma wird der Verlauf der Genexpression verständlich. Es wird davon ausgegangen, dass nahezu alle Tat-Substrate vor der Translokation Chaperone binden, welche diesen bei der korrekten Faltung assistieren. Gemäß einem häufig propagierten Mechanismus könnten Chaperone auch dazu dienen, das Tat-Substrat vor proteolytischer Degradation seines Signalpeptides zu schützen, und somit die Effizienz der Translokation zu steigern (Brüser *et al.* (2007)). Coexpressionstrategien von Chaperonen zeigten bei Tat-abhängigen Translokationen oftmals signifikante Steigerungen der Translokationseffizienz.

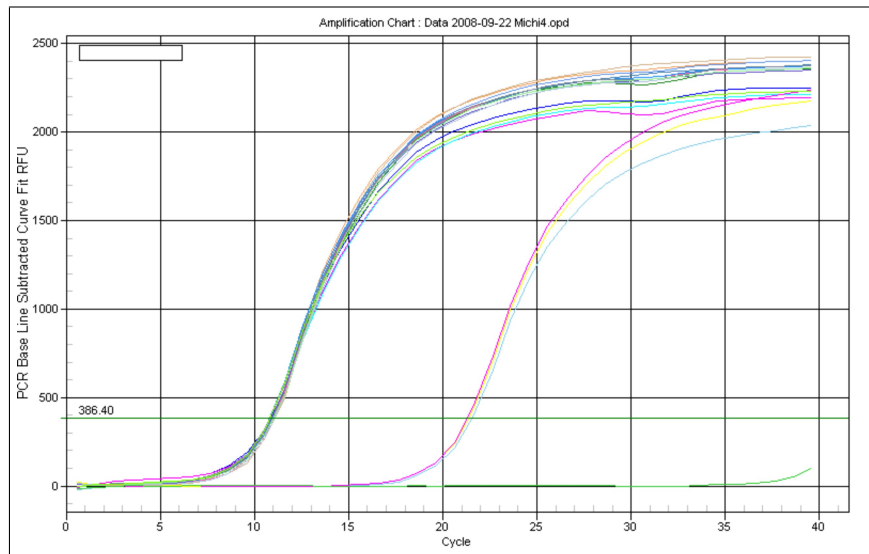


Abb. 63: Die resultierenden Amplifikationskurven für gfp-spezifische Primer. Deutlich zu sehen ist die Verschiebung des Ct-Wertes der Probe vor Induktion (rechte Kurven) zu den Werten der Proben nach Induktion (linke Kurven).

Insbesondere DnaK hat sich unter diesem Aspekt als besonders effizient erwiesen. Perez-Rodriguez *et al.* (2007) konnten durch die Coexpression von DnaK die translokierte GFP-Menge um den Faktor 3 steigern. Als weiteres Beispiel sei auf die Coexpressionstrategie von torD im gleichen System verwiesen (Li *et al.* (2003)). Diese Ergebnisse belegen den außerordentlichen Wert der Microarray-Analyse als *Screening-Tool*, um differentiell exprimierte Gene zu identifizieren. DnaK hat sich nach Induktion als deutlich positiv reguliert erwiesen und diese Ergebnisse wurden durch die Resultate aus quantitativen *real-time* PCR-Analysen gestützt. Dies zeigt wie die generierten Daten für die systematische Entwicklung von Coexpressionstrategien erfolgreich genutzt werden können. Die Gene *secD* und *dnaA* zeigen einen ähnlichen transkriptionellen Verlauf wie er für *tatD* und *tatA* evaluiert wurde. 15 und 30 Minuten nach Induktion sind sie leicht positiv reguliert, um anschließend wieder auf ihr basales Niveau zu fallen. Das Gen, welches für die Serin-Hydroxymethyltransferase codiert wurde mit in die Auswahl aufgenommen, da es sich aus den Chipdaten als negativ reguliert erwiesen hatte. Dieser Trend wird auch mittels der *real-time* PCR-Daten beobachtet. 15 Minuten nach Induktion wird das Gen schwach reprimiert und sein Transkriptionsniveau fällt sukzessiv auf die Hälfte des Niveaus vor Induktion. Wie bereits bei der Diskussion der Microarray-Daten erläutert (Abschnitt 4.7.9), erschließt sich der biologische Sinn aus der Umstellung des Kohlenstoffmetabolismus von Glycerin auf *L*-Arabinose. Ein Vergleich der Microarray-Analyse mit der quantitativen *real-time* PCR-Analyse für vier verschiedene Gene ist in Abbildung 64 gezeigt.

Wie aus der Abbildung ersichtlich, korrelieren die Microarray-generierten Daten sehr gut mit den Ergebnissen, welche mittels quantitativer *real-time* PCR-Analyse erhalten wurden.

Zusammenfassend lässt sich konstatieren, dass die generierten Chipdaten aus den *low-density* Microarray Experimenten im Wesentlichen durch die quantitative *real-time* PCR-Analyse verifiziert wurden. Die Regulation der Gene *gfp*, *lacA*, *tatA*, *dnaK*, *signalpeptidase 2* und *glyA* wurde bestätigt und sowohl der Grad der Regulation als auch die Richtung

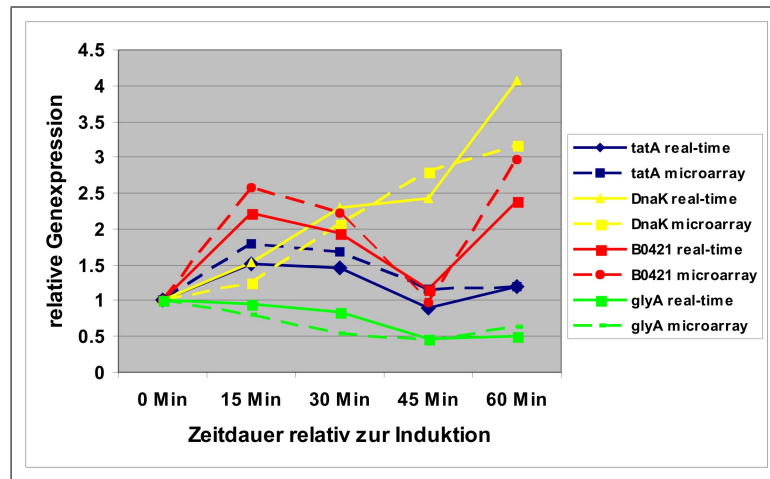


Abb. 64: Vergleich der Expressionsprofile von vier verschiedenen Genen. Gezeigt sind die Ergebnisse welche mittels Microarray-Analyse (gestrichelte Kurven) und quantitativer *real-time* PCR-Analyse (durchgehende Kurven) ermittelt wurden.

in guter Korrelation mit den Microarraydaten identifiziert. Für das sehr stark exprimierte GFP-Gen weicht der chipbasierte Genexpressionsquotient allerdings signifikant von dem mittels *real-time* Analytik ermittelten Wert ab. Diese Abweichung liegt allerdings in der Natur der Analysenmethoden begründet. Einige nur sehr schwach regulierte Gene, wurden mittels Microarray-Analyse nicht identifiziert oder wiesen sogar eine entgegengesetzt gerichtete Regulation auf. Dies zeigt, dass die Microarray-Analyse sehr gut als *Screening*-Methode geeignet ist, um die meisten regulierten Gene reproduzierbar zu erfassen. Allerdings unterliegt die Methodik Limitierungen, wenn der absolute Wert des Genexpression bestimmt werden soll, wie das Beispiel des hoch exprimierten GFP-Genes zeigt. Diese Aussage gilt im Speziellen für nur schwach regulierte Gene. Aufgrund der hohen Variabilitäten von Microarray-Analysen werden diese Gene nicht immer korrekt erfasst und sollten im Zweifelsfall mit anderen Methoden, wie der im Rahmen dieser Arbeit durchgeführten quantitativen *real-time* PCR-Analyse verifiziert werden. Am Beispiel des *dnaK*-Genes manifestiert sich der außerordentliche Nutzen der Microarray-Technologie zur systematischen Evaluierung von potentiellen Zielgenen für Coexpressions- und *knock-out* Strategien. Die quantitative *real-time* PCR generiert sehr genaue und reproduzierbare Daten und eignet sich hervorragend zur Verifikation von Chipdaten. Es konnte gezeigt werden, dass unter der Voraussetzung einer genauen experimentellen Arbeitsweise die Schritte der reversen Transkription und der RNA-Präparation kaum Quellen für Variabilitäten in einem *real-time* Experiment darstellen.

4.7.11 Northern Blot Analyse der GFP-Expression

Als weitere Validierungsmethode für die generierten Microarray-Daten wurde die Methodik des *Northern Blotting* genutzt. Der Transfer und die Immobilisierung von RNA auf einer Membran mit nachfolgender sequenzspezifischer Detektion wird als *Northern Blotting* bezeichnet. Der Name leitet sich vom *Southern Blotting*, dem Transfer und Nachweis von DNA auf Membranen ab, aus dem diese Methode hervorgegangen ist. Die zu untersuchende RNA wird zunächst gelelektrophoretisch unter denaturierenden Bedingungen aufgetrennt und nach der Fixierung auf der Membran durch Hybridisierung mit einer sequenzspezifischen

DNA- oder RNA-Sonde detektiert. Neben der Quantifizierung gewinnt man somit auch über einen Vergleich mit einem Längenstandard die Information der Transkriptlänge. Aufgrund der stärkeren Hybridisierung und der damit einhergehenden größeren Sensitivität ist die Verwendung von RNA-Sonden jener von DNA-Sonden vorzuziehen. Bedingt durch substantielle Unterschiede zwischen DNA und RNA muss die Gelelektrophorese obligat unter denaturierenden Bedingungen erfolgen. Aufgrund der Einzelsträngigkeit des RNA-Moleküls und der thermodynamisch stabileren Basenpaarung tendiert die RNA in stärkerem Maße zur Ausbildung intra- und intermolekularer Sekundärstrukturen als dies bei DNA der Fall ist. Zudem muss aufgrund der deutlich größeren Sensitivität gegenüber Hydrolyse durch Basen oder Säuren sowie gegenüber der enzymatischen Degradation (RNAsen) wesentlich mehr Sorgfalt bei der experimentellen Durchführung getroffen werden. Als Denaturierungsmittel finden Glyoxal oder Formaldehyd Verwendung. Diese werden dem Gel und/oder dem Probenauftragungspuffer zugesetzt. Zur Evaluierung der differentiellen Genexpression werden von jeder Probe gleiche Mengen Gesamt-RNA oder mRNA aufgetragen. Der Transfer vom Agarosegel auf die Membran kann mittels Kapillar-, Vakuum- oder Elektroblothing erfolgen. Als Membranen kommen Nitrocellulose- oder Nylonmembranen zum Einsatz, wobei die Vorteile der Nylonmembranen deutlich überwiegen. Neben der größeren physikalischen Stabilität und einer höheren Bindungskapazität der Membran lässt sich die RNA kovalent immobilisieren. Dies bietet auch die Möglichkeit einer mehrmaligen Verwendung (z.B. durch *Stripping* der Membran zur sukzessiven Detektion mehrere RNA-Spezies). Durch die Verwendung positiv geladener Nylonmembranen erfahren die RNA-Moleküle eine zusätzliche elektrostatische Attraktion beim Transfer und bei der Immobilisierung. Der bei manchen Applikationen auftretende höhere Hintergrund der Nylonmembranen gegenüber der Verwendung von Nitrocellulose lässt sich durch die Wahl geeigneter *Blocking*-Reagenzien vollständig unterdrücken. Nach dem Transfer auf die Membran müssen für die nachfolgende Hybridisierung die Denaturierungsmittel entfernt werden. Dies geschieht durch Schwenken des Gels in stark verdünnter Natronlauge oder durch Backen der Membran bei 120°C. Durch das Backen wird die RNA kovalent an die Membran gebunden, eine Alternative stellt das UV-*Crosslinking* dar. Die nachfolgende Hybridisierung kann mit radioaktiv markierten oder Hapten-modifizierten, sequenzspezifischen Nukleinsäuresonden erfolgen. Eine Quantifizierung erfolgt dann mittels Scanning-Densitometrie, Szintillationszählung der ausgeschnittenen Bande oder durch einen Phosphorimager.

Im Folgenden wird die Etablierung eines Systems zur gleichzeitigen Quantifizierung und Längendetektion von mRNA-Spezies mittels *Northern Blotting* beschrieben. Zur Detektion der mRNA-Spezies werden sequenzspezifische Digoxigenin(DIG)-modifizierte Sonden verwendet. Das Digoxigenin System basiert auf der hochaffinen Wechselwirkung eines DIG-Antikörpers mit Digoxigenin, einem chemisch modifizierten Steroid, welches ausschließlich in der Digitalis-Pflanze zu finden ist. Bedingt durch diese Tatsache resultiert eine sehr hohe Spezifität der Interaktion und ein geringer Hintergrund für das DIG-System. Der DIG-Antikörper ist kovalent mit einer Reportergruppe verbunden, meistens handelt es sich um eine enzymatische Markierung z.B. über die alkalische Phosphataseaktivität. In den hier beschriebenen Versuchen wird die DIG-Funktionalität durch einen mit alkalischer Phosphatase markiertem DIG-Antikörper und nachfolgenden kolorimetrischer Umsetzung der Substrate BCIP/NBT durch die Enzymaktivität nachgewiesen.

Es wurden drei Methoden zur Generierung sequenzspezifischer DIG-markierter Sonden evaluiert:

1. Inkorporation von DIG-dCTP während einer PCR-Reaktion
2. Inkorporation von DIG-CTP durch die Reverse Transkriptase während der cDNA-Synthese mittels sequenzspezifischer Primer
3. Inkorporation von DIG-dUTP während der *in vitro* Transkription

Die ersten beiden Methoden generieren DIG-markierte DNA-Sonden, wobei die zu letzt beschriebene Möglichkeit der DIG-Inkorporation RNA-Sonden erzeugt.

4.7.12 Inkorporation von DIG-dCTP während der PCR

Zur Evaluierung der Möglichkeit, die PCR-Reaktion zur Sondensynthese für das DIG-System zu nutzen, wurden sequenzspezifische Primer für GFP und Signalpeptidase 2 zur Amplifikation der cDNA eines induzierten, GFP-exprimierenden Stammes getestet. Eine parallel durchgeführte und bereits optimierte PCR-Reaktion unter identischen Bedingungen, allerdings ohne DIG-11-dUTP-Zugabe, fungierte als Vergleichsreaktion und Positivkontrolle. Die PCR-Bedingungen sind dem Anhang (A.1.22) zu entnehmen, für die GFP-Primer wurde eine *Annealing*-Temperatur von 58°C und für die Signalpeptidase 2-Primer von 60°C gewählt. Die entstandenen Amplifikate der PCR mit und ohne DIG-dUTP Zusatz im Desoxynukleotidtriphosphat-Mix wurden auf einem 1,5 prozentigen Agarosegel aufgetragen, welches in Abbildung 65 dargestellt ist.

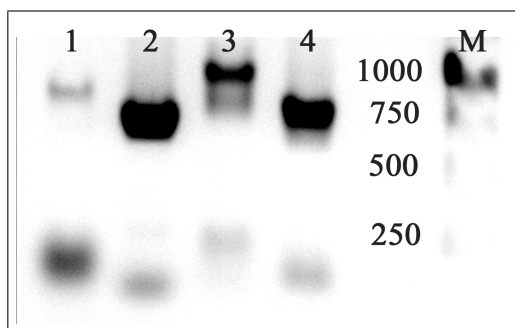


Abb. 65: Vergleich der PCR-Amplifikate zweier verschiedener Primerpaare mit und ohne DIG-dUTP Markierung.

Zum Einordnen der Größenrelation fungiert das Bandenmuster eines 1 kb Marker, welcher auf der Spur mit der Bezeichnung M aufgetragen ist. Auf Spur 4 ist das GFP-Amplifikat, welches ohne DIG-dUTP Zusatz generiert wurde, aufgetragen sowie auf Spur 3 das DIG-markierte Amplifikat. Deutlich ist die Verschiebung der DNA-Bande in den höhermolekularen Bereich durch den Einbau des DIG-dUTP zu erkennen. Aufgrund der signifikant größeren molaren Masse des DIG-dUTP gegenüber dem unmarkierten dTTP tritt der beobachtete *Gelshift* auf. Zu erkennen ist ebenfalls, dass die Bandenintensität des DIG-GFP Amplifikates geringer ist als jene des unmarkierten GFP. Daraus lässt sich schließen, dass die Effizienz der PCR-Reaktion durch den Einbau des sterisch anspruchsvollen DIG-dUTP reduziert wird. Auf Spur 1 und Spur 2 sind die entsprechenden Ergebnisse für die Signalpeptidase 2-Amplifikation aufgetragen, auf Spur 2 befindet sich der Amplifikationsansatz

ohne DIG-dUTP-Zusatz, auf Spur 1 jener mit DIG-dUTP-Zusatz. Die für die Amplifikation mittels GFP-spezifischen Primern getroffenen Aussagen können auch auf dieses Resultat übertragen werden. Die reduzierte Amplifikationseffizienz der *Taq*-Polymerase tritt hier noch stärker in Erscheinung und der Intensitätsunterschied beider Banden fällt noch größer aus als dies bei den GFP-Amplifikaten der Fall ist. Das GFP-Amplifikat wurde zur Abschätzung der DIG-Inkorporation mittels eines *dot blots* (Abbildung 67) näher untersucht.

4.7.13 Inkorporation von DIG-dUTP mittels Reverser Transkription

50 μg einer Gesamt-RNA wurden mittels genspezifischer Primer für GFP in cDNA konvertiert. Durch Zusatz von DIG-dUTP zum Reaktionsansatz wurden die DIG-markierten Nukleotide in die entstehende cDNA-Kette inkorporiert. Die Dotierung mit DIG-dUTP wurde dabei von dem TSA-Protokoll (A.1.5) adaptiert, da die Detektion der hybridisierten Fluorescein- und Biotin-markierten cDNAs ebenfalls über Antikörper erfolgt und eine zu hohe Markierungsdichte für die nachfolgende Antikörperbindung vermieden werden muss. Das Ergebnis der RT-Reaktion ließ sich nicht mittels eines Gelbildes visualisieren, da die generierte spezifische cDNA-Menge unter der Nachweisgrenze lag. Die nähere Untersuchung erfolgte daher mittels eines *dot blots*, welcher in Abbildung 67 gezeigt ist.

4.7.14 Generierung von DIG-Sonden mittels *in vitro* Transkription

Das etablierte System zur Generierung artifizierlicher RNA mittels *in vitro* Transkription (Abschnitt 4.12) wurde zur Herstellung sequenzspezifischer, DIG-markierter RNA-Sonden für GFP- und Signalpeptidase 2-*Targets* modifiziert. Dem Transkriptionsansatz wurden DIG markiertes rUTP im Dotierungsverhältnis von 1,8 : 1 (rUTP : DIG-rUTP) zugesetzt und nachfolgend das etablierte Protokoll angewandt. Die generierte DIG-RNA wurde mittels einer Phenol-Chloroform Extraktion (A.1.29) aufgereinigt, mit 5M Natriumchlorid in Ethanol gefällt und in nukleasefreiem Wasser aufgenommen. Anschließend wurde versucht, den Massenzuwachs der RNA durch Vergleich mit einer unmarkierten *in vitro* RNA, welche aus dem identischen *Template* hervorgegangen ist, zu visualisieren.

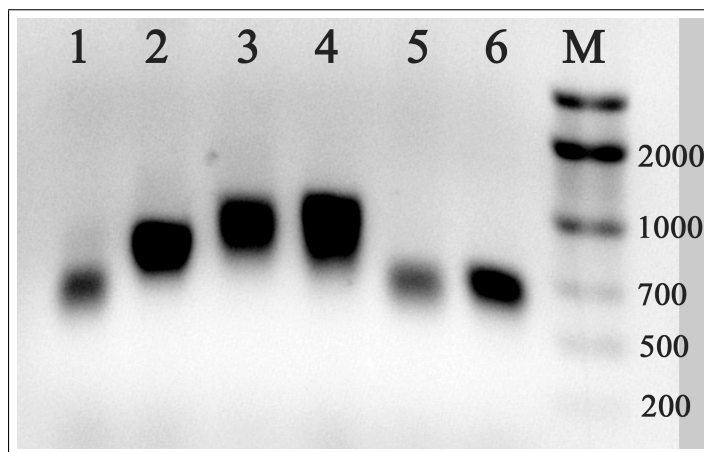


Abb. 66: Vergleich zweier *in vitro* RNA Spezies mit und ohne DIG-Markierung.

Abbildung 66 zeigt das Ergebnis der Agarosegelelektrophorese. Ein RNA-Längenstandard ist auf Spur M aufgetragen, neben diesem befindet sich auf Spur 5 das DIG-markierte GFP-

spezifische RNA-Transkript und auf Spur 6 die unmarkierte GFP RNA-Sonde. Ein Größenunterschied ist nicht erkennbar und der Vergleich mit dem Längenstandard zeigt, dass beide RNA-Spezies eine Länge von ca. 700 Basen aufweisen. Da das gesamte GFP-Gen mit 719 bp in den Transkriptionsvektor kloniert wurde, sollte das linearisierte Plasmid Transkripte von eben dieser Länge generieren. Die RNA-Bande, die aus dem DIG-Transkriptionsansatz hervorgegangen ist, weist eine geringere Intensität auf, als dies für die unmarkierte RNA der Fall ist. Der DIG-Einbau und/oder der veränderte Ribonukleotidmix scheint die Effizienz der *in vitro* Transkription zu senken. Die Spuren 3 und 4 visualisieren DIG-markierte RNA-Sonden, die spezifisch für GlyA (Spur 3) und deren unmarkiertes Analogon (Spur 4) sind. Die Größe für das unmarkierte RNA-Transkript von ca. 1000 Basen passt gut mit der theoretischen Länge von 961 Basen überein. Wiederum wird kein Gelshift beobachtet. Zur Validierung der DIG-Inkorporation während der *in vitro* Transkription wurde daher ein *dot blot* durchgeführt.

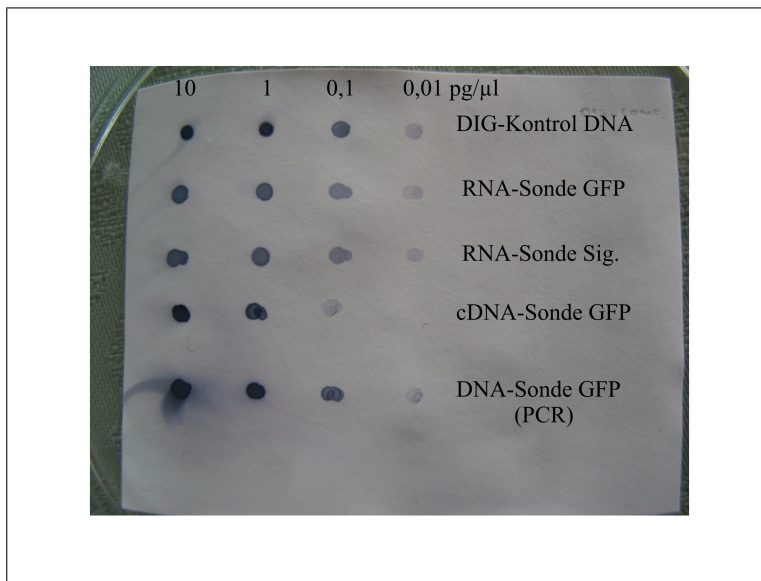


Abb. 67: *Dot-Blot*-Analyse der GFP-Expression. Gleiche Stoffmengen verschiedener DIG-markierter DNA- und RNA-Sonden wurden in einer 1:10 Verdünnungsreihe auf eine Nylonmembran *gespottet*. Die Detektion erfolgte über einen DIG-AP-Antikörper und kolorimetrischer Detektion mittels NBT/BCIP.

Zur Durchführung des *dot-blots* wurden die generierten RNA- und DNA-Sonden initial so verdünnt, dass gleiche Stoffmengen Nukleinsäure auf die Membran appliziert wurden. Aus dieser initialen Verdünnung wurde jeweils eine 1:10 Verdünnungsreihe hergestellt und 1 μ l des jeweiligen Ansatzes auf die Nylon Hybond N(+)-Membran (Amersham) *gespottet*, welche nach dem Eintrocknen der Spots für 30 Minuten bei 120°C gebacken wurde. Anschließend wurde die Membran mittels der im Anhang angeführten Versuchsvorschrift prozessiert und die DIG-Funktionalität nach Inkubation mit einem AP-konjugierten DIG-Antikörper kolorimetrisch mit BCIP/NBT nachgewiesen. Eine Photographie des *dot blots* ist in Abbildung 67 wiedergegeben. Die *dots* der DIG-Kontroll-DNA werden mit ihrer Konzentration in der Abbildung spezifiziert. In Reihe 2 und 3 sind die mittels *in vitro* Transkription generierten RNA-Sonden für GFP und GlyA *gespottet*, in Reihe 4 die mittels RT generierte GFP-spezifische cDNA-Sonde und schließlich in Reihe 5 die mittels PCR generierte GFP-DNA-Sonde. Die

RNA-Sonden wurden initial am stärksten verdünnt, da die *in vitro* Transkription die höchsten Sonden-Konzentrationen generiert. Ein Vergleich der verschiedenen Sonden zeigt, dass die mittels RT-Reaktion generierte cDNA-Sonde die intensitätsschwächsten *dots* aufwies und somit auch die geringste Stoffmenge an inkorporiertem DIG-Nukleotiden. Da die gleiche Stoffmengen an Nukleinsäuren *gespottet* wurden, entspricht dieses auch der niedrigsten Inkorporationsfrequenz. Diese ist für die RNA-Sonden und für die mittels PCR generierten DNA-Sonden in etwa gleich. Die Berechnung der Inkorporationsfrequenz erfolgte nach folgender Formel:

$$a = \frac{n(DIG - d(r)UTP)}{n(\text{Basen in der Nukleinsaeurekette})} = \frac{\frac{m_C(DIG-d(r)UTP)}{M(DIG-d(r)UTP)}}{\frac{m_C(\text{Nukleinsaeure})}{M(\text{Base})}}$$

Die Formel beschreibt das Stoffmengenverhältnis von DIG-11-dUTP bzw. DIG-11-rUTP zur Stoffmenge der Nukleinsäure. Dabei bezeichnet n die Stoffmenge, m die Masse und M die molare Masse der Komponente. Das resultierende Ergebnis muss anschließend noch durch die Länge der Nukleinsäure dividiert und mit dem Faktor 1000 multipliziert werden, um die Anzahl der inkorporierten DIG-Moleküle pro 1000 Basen zu erhalten. Es ergaben sich folgende Inkorporationsfrequenzen für die verschiedenen Sonden:

1. GFP *in vitro* Sonde: 36,13
2. glyA *in vitro* Sonde: 24,97
3. GFP PCR-generierte Sonde: 17,79

Für die mittels Reverser Transkription generierte cDNA-DIG-Sonde konnte die Inkorporationsfrequenz nicht berechnet werden, da eine genaue Bestimmung der DNA-Konzentration nicht möglich war. Wie im *dot blot* gezeigt, war die Inkorporationsfrequenz für DIG-dUTP geringer als für die anderen Sonden. Die höchsten Inkorporationsfrequenzen ergaben sich für die mittels *in vitro* Transkription generierten DIG-Sonden, so dass diese für den nachfolgend durchgeführten *Northern Blot* zum Einsatz kamen. Zusätzlich zur höheren Inkorporationsfrequenz weisen diese auch eine thermodynamisch stabileren Hybridisierung mit der Zielsequenz auf, was in einer deutlich höheren Sensitivität der Analyse resultiert.

Zur Evaluierung der optimalen Sondenkonzentration für die Hybridisierungslösung des *Nothern Blots* wurde eine Mock-Hybridisierung durchgeführt. Dabei wurde die Hybond N(+)-Membran mit unterschiedlichen Sondenkonzentrationen hybridisiert und anschließend die kolorimetrische Detektion einem AP-konjugierten DIG-Antikörper und dem Substrat NBT/BCIP durchgeführt. Es wurden dabei identische Bedingungen zu denen des *Northern Blots* eingehalten. Zuerst wurde ein Teil der Nylonmembran in vier gleich große Abschnitte zerteilt und je ein Abschnitt über Nacht mit 250, 100, 50 und 25 ng/ml GFP-RNA-Sonde in 10 ml Prähybridisierungslösung bei 68°C hybridisiert. Nach dem Waschen erfolgte die Detektion mittels NBT/BCIP völlig analog der eingangs beschriebenen Prozedur. Das Ergebnis der Mock-Hybridisierung ist in Abbildung 68 als Photographie der Membranabschnitte gezeigt. Die Membranabschnitte, welche mit einer Sondenkonzentration von 250 und 100ng/ml hybridisiert wurden, weisen eine deutlich dunklere Färbung und damit eine deutlich größere Hintergrundintensität auf, als dies für die anderen beiden Membranabschnitte der Fall ist. Die Mock-Hybridisierung läßt abschätzen, welche Sondenkonzentration welchen Hintergrund im Experiment generieren wird. Die höchste Sondenkonzentration, die noch

4.7 Konstruktion eines Sekretions-Reportersystems für Genexpressionsstudien

einen akzeptablen Hintergrund generiert, wurde als Arbeitskonzentration für den nachfolgend durchgeführten *Northern Blot* gewählt. Wie aus der Abbildung 68 ersichtlich ist, trifft dieses für eine Sondenkonzentration von 50 ng/ml zu.



Abb. 68: Mock-Hybridisierung mit vier verschiedenen Sondenkonzentrationen.

Nachdem die optimale Sondenkonzentration evaluiert war, wurden die RNA-Proben der Bioreaktorfermentation des Sekretionsstammes mit dem Plasmid pBADtorAGFPSsrA auf ein denaturierendes 1,5 prozentiges Agarosegel aufgetragen. Pro Geltasche wurden jeweils 10 μg RNA appliziert und diese für 1,5 Stunden elektrophoretisch aufgetrennt. Dabei fand ein Vergleich von Proben statt, welche 15, 30, 45, 60, 120 und 180 Minuten nach Induktion entnommen wurden. Als Referenz fungierte eine direkt vor der Induktion aus dem Bioreaktor entnommene Probe. Abbildung 69 zeigt das Gelbild der denaturierenden Agarosegelelektrophorese. Deutlich zu erkennen sind die beiden ribosomalen Banden der 16S- und 23S-rRNA. Dies indiziert die molekulare Integrität der RNA.

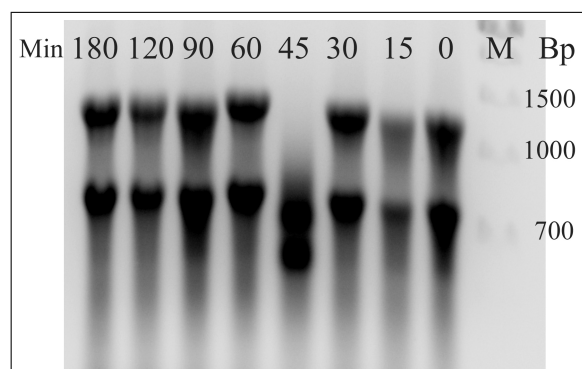


Abb. 69: Gelelektrophorese von 10 μg total-RNA verschiedener Proben der Bioreaktorkultivierung des Stammes XL1-Blue pBADtorAGFPSsrA.

Nachfolgend wurde die RNA mittels *Elektroblotting* auf eine Nylonmembran transferiert. Nach dem *Blotten* und der Fixierung auf der Membran war die RNA noch immer mit dem in-

terkalierende Ethidiumbromid assoziiert, wie ein photographisches Bild der trockenen Membran unter UV-Exposition beweist (Abbildung 70). Deutlich sind die 16S- und 23S-Banden der ribosomalen RNA erkennbar. Somit konnte der erfolgreiche Transfer auf die Membran direkt überprüft werden. Durch eine UV-Exposition des Geles nach dem *Blot*vorgang konnte keine Fluoreszenz mehr detektiert werden, weshalb von einem quantitativen Transfer auszugehen ist. Der bei 120°C getrocknete *Blot* wurde anschließend nach einstündiger Inkubation in Prähybridisierungslösung über Nacht bei 68°C mit 10 ml GFP-RNA-Sonde (50 ng/ml) hybridisiert. Die weitere Behandlung entsprach dem im Anhang wiedergegebenen Protokoll. Nach ca. fünfminütiger Inkubation des *Blots* in NBT/BCIP-Lösung konnten intensive Banden beobachtet werden, und die Reaktion wurde durch Schwenken des *Blots* in destilliertem Wasser beendet. Abbildung 70 zeigt eine Photographie des *Blots* und einen Vergleich mit der UV-exponierten Membran nach der Fixierung der RNA.

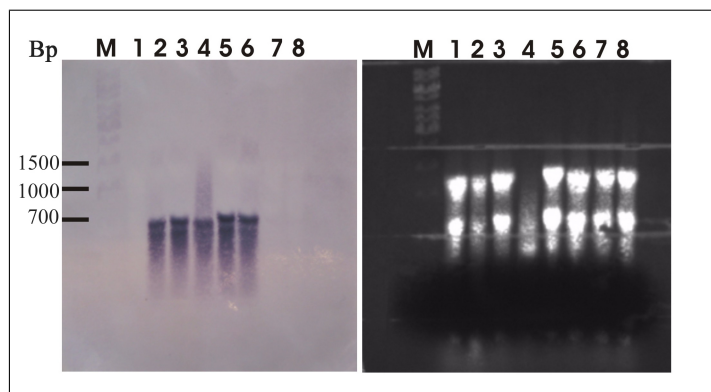


Abb. 70: *Northern Blot*-Analyse der GFP-Expression. Links ist der mit einer *gfp*-spezifischen DIG-Sonde hybridisierte *Northern Blot* zu sehen, während rechts ein Scan der gebackenen Membran (unter UV-Exposition) nach dem *Blot*vorgang gezeigt ist.

Wie aus der Abbildung 70 ersichtlich, ist bei der Probe vor der Induktion keine Bande zu sehen, es wurden daher keine oder nur geringe Mengen an GFP-mRNA gebildet. Dies stimmt sehr gut mit der beschriebenen, sehr geringen basalen Transkription der stromabwärts des Arabinosepromoters lokalisierten Gene überein (Guzman *et al.* (1995)). Die Spuren 3-6 zeigen deutlich eine einzelne Bande im Bereich um 700 bp. Densitometrisch waren keine Unterschiede in den Bandenintensitäten zu messen. Bei der RNA, welche aus der Probe 180 Minuten nach Induktion gewonnen wurde, konnte keine Sondenhybridisierung detektiert werden (Abbildung 70 Spur 7). Das Fehlen von GFP-mRNA 180 Minuten nach Induktion kann durch die metabolischen Degradation der *L*-Arabinose durch die Zellen erklärt werden. Der Induktor muß zwischen 120 und 180 Minuten nach Induktion vollständig metabolisiert sein. Dieser Befund koinzidiert sehr gut mit den Ergebnissen der *low-density* Microarrays, welche für viele nach der Induktion regulierte Gene ab diesem Zeitpunkt ein Zusammenbrechen der Regulation detektierten.

4.7.15 Zusammenfassung

Vorangehend wurde die erfolgreiche Etablierung eines GFP-basierten Sekretionsreportersystems zur Evaluierung der Tat-spezifischen Sekretion ins Periplasma gezeigt. Mittels dieses Systems kann durch Messung der Gesamtzellfluoreszenz auf die periplasmatisch lokalisierte

GFP-Menge geschlossen werden. Zur näheren Charakterisierung des Systems wurden weitere, auf der GFP-Variante basierende, Plasmide konstruiert. Die erfolgreiche Sezernierung des GFP-Konstruktes ins Periplasma wurde mittels durchflußzytometrischer Analyse sowie einer Zellfraktionierung mit nachfolgender SDS-Page-Analyse verifiziert. Eine Kultivierung des Reporterstamms im Bioreaktor lieferte das Probenmaterial für eine transkriptionelle Analyse mittels sekretionsspezifischer *low-density* Microarrays. Es konnte dabei die differentielle Expression verschiedener Chaperone, Komponenten der Translokase und von Genen, die in der Aufrechterhaltung der Triebkraft der Translokation fungieren, festgestellt werden. Die Daten wurden mittels quantitativer *real-time* PCR verifiziert. Zusätzlich wurde die Anwendung von *Northern Blots* für Genexpressionsanalysen anhand der Analyse der GFP-Expression gezeigt.

4.8 Etablierung verschiedener arabinoseinduzierbarer RFP *E. coli* Stämme

Das Emissionsspektrum von GFP interferiert signifikant mit denen verschiedener Metabolitprodukte (z.B. Flavine). Eine *Online*-Analyse der GFP-Fluoreszenz im Bioreaktor ist aus diesem Grund nur schwer realisierbar. Im roten Spektralbereich liegen die Verhältnisse weit günstiger, es existieren signifikant weniger interferierende Substanzen und somit sollte eine Quantifizierung *online* möglich sein. Am Institut für Technische Chemie steht hierfür das 2D-Fluoreszenzspektrometer Bioview zur Verfügung. Motiviert durch diese Applikation wurde die Etablierung eines rot-fluoreszierenden Protein (RFP)-Reportersystems für Genexpressionsstudien angestrebt. RFP eignet sich zudem in idealer Weise für duale Versuche mit GFP (Hakkila *et al.* (2003), Bevis *et al.* (2002)). So ließe sich damit z.B. simultan die Sekretion zweier Substrate erfassen oder die Sekretion über zwei Sekretionspfade miteinander vergleichen (z.B. im Rahmen einer kompetitiven Situation). RFP wurde in der Koralle *Discosoma* entdeckt und zeigt eine orange-rote Fluoreszenz mit einem Emissionsmaximum von 583 nm. In Abbildung 71 ist die photographische Aufnahme der Zelllysate dreier selbst generierter RFP-Stämme zur Verdeutlichung gezeigt.

Unter dem Namen DsRed1 ist das codierende Plasmid kommerziell erhältlich. Die optimale Anregungswellenlänge für das von DsRed1 codierte RFP liegt bei 558 nm, jedoch ist auch eine Anregung bei 488 nm mittels Standard-Lasern möglich. Besonders für die Laserbasierte konfokale Fluoreszenzmikroskopie und durchflußzytometrische Analysen ist diese Tatsache von Bedeutung (Bevis *et al.* (2003)). Das RFP liegt in Form eines Tetramers vor, jedes Monomer ist dabei strukturell GFP sehr ähnlich. Bei der Reifung entwickelt DsRed daher zuerst eine schwache grünliche Fluoreszenz, erst eine sekundären Oxidationsreaktion generiert das rote Chromophor. Die spektrale Verschiebung im Vergleich zu GFP in den längerwelligen Bereich entsteht durch ein ausgedehnteres konjugiertes Π -Elektronensystem. RFP zeigt eine hohe Quantenausbeute und ist photostabil. Neben diesen vielversprechenden Eigenschaften weist RFP allerdings einige Nachteile für die Verwendung als Fluoreszenzreporter auf. Neben der bereits erwähnten initialen grünen Fluoreszenz, (welche bei simultanen Analysen mit GFP interferiert), zeigt das Tetramer eine hohe Tendenz zur Bildung höherer Aggregate. Die Tetramerisierung kann zudem bei Fusionen an andere Proteine deren Funktion und Lokalisierung in der Zelle stören. Die größten Nachteile bei der Verwendung von Wildtyp-RFP zeigen sich jedoch durch die langsame Reifung des Chromophors bei Raumtemperatur mit Halbwertszeiten, die mehr als 24 Stunden betragen. Motiviert durch dieses Faktum wurde mittels einer Kombination aus *random* und *site-directed mutagenesis* eine

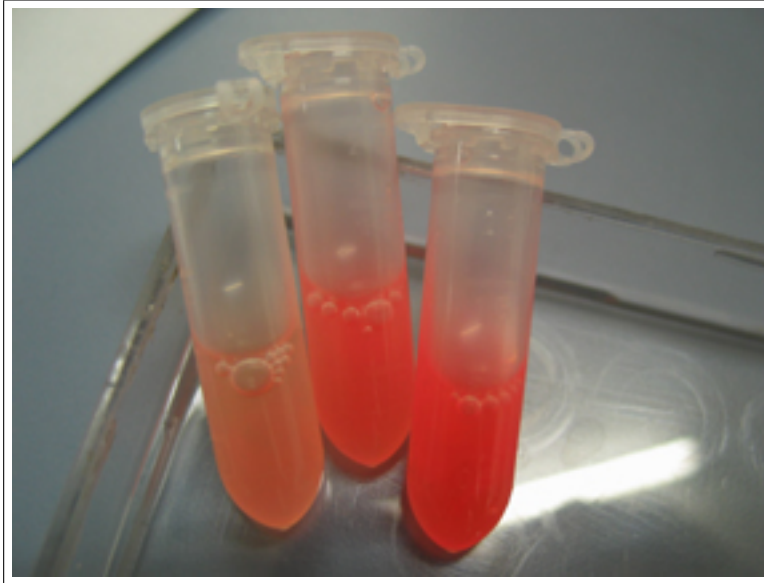


Abb. 71: Zellysate dreier induzierter, RFP-codierender Stämme. Von links nach rechts: TOP10 pBADtorARFPSsrA, TOP10 pBADtorARFPHis, TOP10 pBADtorARFP. Sehr deutlich ist die intensive rote Fluoreszenz des RFP-Express-Derivat auszumachen.

schnell reifende Variante erzeugt. Diese weist eine signifikant reduzierte Emission im grünen Spektralbereich und eine drastisch verminderte Tendenz zur Selbstassoziation auf. Dieses RFP-Derivat ist unter dem Namen DsRedExpress kommerziell (Clontech) erhältlich und dient als Grundlage der vorgestellten Versuche (Bevis *et al.* (2003)).

Für die Konstruktion der RFP-Plasmide wurde eine völlig analoge Vorgehensweise zu den GFP-Klonierungen gewählt (Abschnitt 4.8.2). Ziel war die Konstruktion des Sekretionsplasmids pBADtorARFPSsrA zur Analyse der Tat-spezifischen Sekretion ins Periplasma. Durch die SsrA-Fusion sollte, wie beim GFP-Konstrukt (Abschnitt 4.8.1) angemerkt, nur periplasmatisch lokalisiertes RFP ein Fluoreszenzsignal generieren. Zur näheren Charakterisierung des Systems wurden weitere RFP-codierende Plasmide kloniert. Eine Übersicht ist in Abbildung 72 gegeben.

Die mittels *Pfu*-Polymerase generierten Inserts wurden *blunt-end* in den Klonierungsvektor pGEM5Zf(+) subkloniert. Über die in die Vektorsequenz eingefügten Restriktionsschnittstellen wurden die *Inserts* ausgeschnitten, aus dem Gel aufgereinigt und im molaren Verhältnis von 3:1 (*Insert*: Vektor) über kohäsive Enden in den analog prozessierten Zielvektor pBAD33 kloniert. Nachdem die entsprechenden RFP-codierenden Plasmide konstruiert waren, wurden diese in XL1-Blue Zellen transformiert und die jeweiligen Stämme im Schüttelkolben kultiviert. Durch Supplementierung von *L*-Arabinose wurde die RFP-Expression induziert. Erstaunlicherweise erfolgte keine Bildung eines roten Fluorophors, wie sie zuvor durch Klonierung der RFP-Konstrukte in pGEM5Zf(+) auf der Agarplatte schon beobachtet wurde. Da XL1-Blue nur sehr geringe Mengen rekombinantes Protein zu exprimieren vermag und RFP mit 126 kDa ein sehr großes Protein darstellt, wurden mit KRX (Promega) und TOP10 (Invitrogen GmbH) zwei andere kommerzielle Expressionsstämme für diesen Zweck evaluiert. Zusätzlich wurde vermutet, dass die schnelle Metabolisierung des Induktors *L*-Arabinose dazu führte, dass nicht genügend RFP gebildet wurde. Die Verwendung von KRX-Zellen erwies sich nicht als zielführend, wohingegen *E. coli* TOP10 funktionales RFP

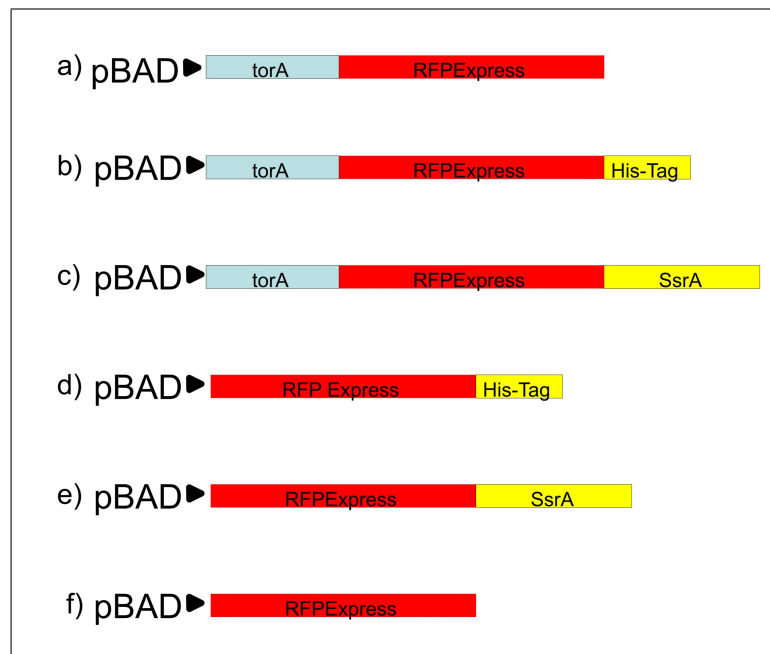


Abb. 72: Verschiedene RFP-Konstrukte zur näheren Charakterisierung und Verifizierung des etablierten Sekretionsrepotersystemes.

zu exprimieren vermochte. Ein Photo der induzierten Schüttelkolbenkultivierungen von drei verschiedenen RFP-Stämmen ist in Abbildung 73 gezeigt. Deutlich ist die rote Fluoreszenz des RFPs visuell zu erfassen.

Der TOP10-Stamm besitzt den Genotyp $araD139\Delta(ara-leu)7697$, was eine Deletion im *araD*-Gen bezeichnet. Daher vermag dieser Stamm keine Arabinose zu metabolisieren. Es kann aus diesem Grund vermutet werden, dass die Expression signifikanter RFP-Mengen mehr Zeit beansprucht als dies für GFP der Fall ist. Bei einem Stamm ohne defizienten Arabinosemetabolismus ist zu diesem Zeitpunkt schon eine zu große Menge des Induktors metabolisiert. Untermuert wird diese These durch die Bildung roter Kolonien auf der Agarplatte von RFP-Konstrukten, welche in pGEM5Zf(+) kloniert und in XL1-Blue-Zellen transformiert wurden. Das *lacZ*-Fragment, in welches RFP kloniert wurde (MCS), steht unter Kontrolle eines konstitutiven Promoters, so dass RFP gebildet werden kann. Die Gründe für die gescheiterte RFP-Expression in XL1-Blue liegen daher ausschließlich am schnellen Abklingen der Induktion und nicht an den Expressionseigenschaften des Stammes. Somit erfolgte die Transformation der verschiedenen RFP-Plasmide in *E. coli* TOP10. Transformierte Zellen, welche die Plasmide pBADtorARFPsSrA, pBADtorARFPHis und pBADtorARFP enthielten, wurden im Schüttelkolben kultiviert und nach Erreichen einer OD₆₀₀ von 0,8 mit 0,2 Prozent *L*-Arabinose induziert. Nach ca. 2 Stunden war bei den Stämmen welche die Plasmide pBADtorARFPHis und pBADtorARFP enthielten eine intensive Rotfärbung der Kultur zu erkennen. Bei dem Stamm mit dem Plasmid pBADtorARFPsSrA erfolgte eine visualisierbare Rotfärbung erst ca. 6 Stunden nach Induktion. Die Betrachtung unter dem Fluoreszenzmikroskop bestätigte das optische Bild. Exemplarisch ist die Aufnahme eines induzierten Stammes mit dem Plasmid pBADtorARFP in Abbildung 74 gezeigt. Als Vergleich fungierte eine uninduzierte Probe des gleichen Stammes, welche keine rote Fluoreszenz zeigte.

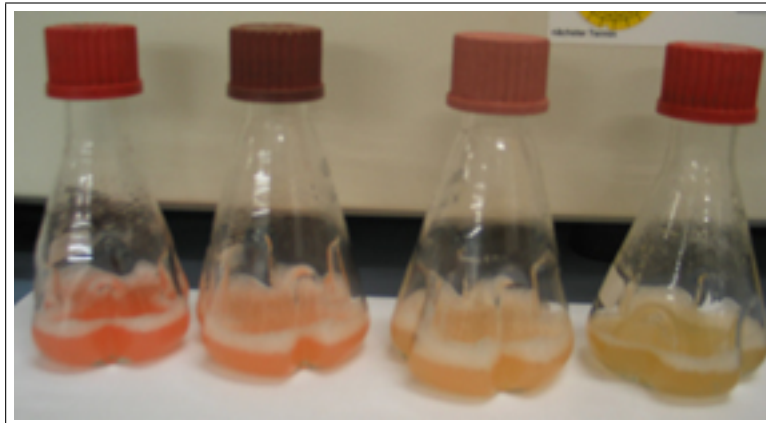


Abb. 73: Photographie der induzierten Schüttelkolbenkultivierungen verschiedener RFP-Stämme in LB-Medium. Von links nach rechts: TOP10 pBADtorARFP, TOP10 pBADtorARFPHIS, TOP10 pBADtorARFPSSrA. Zum Vergleich ist ganz rechts ein uninduzierter TOP 10 pBADRFP-Stamm zu sehen.

Zusätzlich wurden Proben von drei RFP exprimierenden Sekretionsstämmen (pBADtorARFP, pBADtorARFPHIS und pBADtorARFPSSrA) einer durchflußzytometrischen Analyse unterzogen. Der Zeitpunkt der Probennahme erfolgte 6 Stunden nach Induktion, um die rote Fluoreszenz auch bei dem Stamm TOP10 pBADtorARFPSSrA zu detektieren.

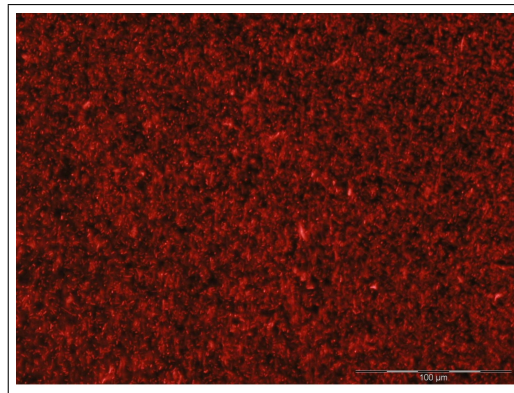


Abb. 74: Fluoreszenzmikroskopische Aufnahmen einer induzierten Kultur des Stammes TOP10 pBADtorARFPSSrA. Deutlich ist die intensive rote Fluoreszenz sichtbar.

Wie in Abbildung 75 gezeigt, weisen alle Stämme nach Induktion mit *L*-Arabinose eine deutliche Fluoreszenz im roten Spektralbereich auf. Der Mittelwert der Fluoreszenzintensität ist dabei für die Stämme TOP10 pBADtorARFPHIS und TOP10 pBADtorARFP nahezu gleich; letzterer weist eine unwesentlich höhere Fluoreszenzintensität auf. Darin könnte sich ein potenzieller Einfluß des Hexahistidin-*tags* widerspiegeln. Im Wesentlichen scheint dieser aber vernachlässigbar. Der Stamm mit dem Plasmid pBADtorARFPSSrA zeigt im Vergleich mit der uninduzierten Referenz eine deutliche Fluoreszenz, allerdings ist diese signifikant geringer als bei den anderen beiden fluoreszierenden Stämmen. Hierin spiegelt sich der Einfluß des Degradations-*tags* wieder.

Es muss im Folgenden evaluiert werden, ob die rote Fluoreszenz von cytoplasmatisch

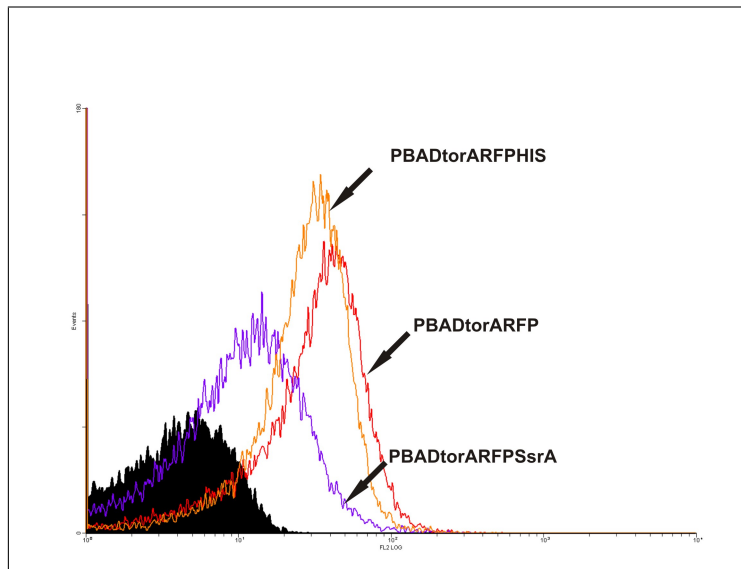


Abb. 75: Durchflußzytometrische Analyse dreier RFP-codierender Stämme. Zum Vergleich ist (schwarz hinterlegt) eine uninduzierte Vergleichsprobe gezeigt. Deutlich ist die Verschiebung der Histogramme zu höheren Fluoreszenzintensitäten im roten Spektralbereich zu sehen.

lokalisiertem RFP herrührt, welches durch Überlastung des Proteasesystems entsteht oder tatsächlich von ins Periplasma sezernierten RFP-Konstrukt. Da der Induktor *L*-Arabinose im Top10-Stamm nicht metabolisiert wird, kann nicht mehr davon ausgegangen werden, dass die gewählte Induktionsstärke das Proteasesystem nicht überlastet, wie dies im XL1-Blue Stamm für das GFP-Konstrukt (Abschnitt 4.8.1) der Fall ist. Die erst nach deutlich längerer Dauer nach Induktion auftretende rote Fluoreszenz von pBADtorARFPsSrA kann sowohl von einem Einfluß des SsrA-*tags* auf die Reifung des RFPs herrühren als auch durch eine sukzessive Akkumulation von cytoplasmatischen RFP entstehen; bedingt durch eine Überlastung des Proteasesystems. Die Kombination aus der langen Zeitperiode bis eine Rotfluoreszenz detektierbar war und die nicht mehr erfüllte Grundvoraussetzung der Nicht-Überlastung des Clp-Proteasesystems lassen den Stamm TOP10 pBADtorARFPsSrA als Sekretionsreporter ungeeignet erscheinen. Zudem konnte keine signifikante Sekretion des RFPExpress ins Periplasma nachgewiesen werden. Ein Grund dafür kann allerdings auch der verwendete TOP10 Stamm sein. Sommer *et al.*(2008) konnten bei den Versuch diesen Stamm als Träger verschiedener Sekretionsplasmide zu etablieren ebenfalls keine Sekretion realisieren. In anderen Stämmen (DH5 α , Ara 1655) waren die plasmid-codierten Sekretionsreporter jedoch funktional. RFP weist mit 126 kDa Molekulargewicht zudem eine grenzwertige Größe für die Translokation mittels des Tat-Pfades auf. Das Molekulargewicht des größten bisher nach dem Tat-Mechanismus sekretierten Proteins betrug 142 kDa (Brüser *et al.* (2007)).

Es konnte gezeigt werden, dass die bei der GFPmut2 Variante erfolgreiche Klonierungsstrategie (Abschnitt 4.8) auch im Falle des RFP-Derivats zum Erfolg führte. Aufgrund der spektralen Eigenschaften und der Möglichkeit, duale Versuche mit GFP durchzuführen, erscheint die Evaluierung weiterer Stämme für eine erfolgreiche Sekretion von RFPExpress äußerst lohnend. Eine zielführende Möglichkeit könnte sich in dem *Screening* verschiedener anderer RFP-Varianten für die gewünschte Applikation ergeben. Besonders die Verwendung

einer monomeren und damit sterisch weniger anspruchsvollen RFP-Variante, wie sie Campell *et. al* (2002) entwickelten, scheint im diesem Kontext sehr erfolgsversprechend.

4.9 Optimierungsansätze für Microarray-Applikationen

Im Rahmen dieser Arbeit wurde die Methode der Genexpressionsanalyse über Microarrays nicht nur eingesetzt, sondern auch selbst optimiert. Zu diesem Zweck wurde ein neues, innovatives Detektionssystem für Microarray-Applikationen mittels *Quantum Dots* etabliert und die Möglichkeit einer chipbasierten absoluten Quantifizierung geschaffen.

4.10 *Quantum Dots*

Mit der Terminologie *Quantum Dots* werden Nanopartikel bezeichnet, welche als Fluorophore Verwendung finden und einen dreischichtigen Aufbau aufweisen. Der Kern besteht aus einem kristallinen Atomcluster eines halbleitenden Materials, welches von einem weiteren Halbleiter umhüllt ist. Als dritte Schicht ist diese Anordnung von einem funktionalisierten Polymer umschlossen, welches zur Kopplung der *Quantum Dots* an Biomoleküle verwendet werden kann (Biju *et al.* (2008)). Durch verschiedenartige Funktionalisierungen der *Dots* wurden Anwendungen zum Nachweis von Immunokomplexen oder Nukleinsäurehybridisierungen realisiert. Ebenso erfolgreich wurden *Quantum Dots* als Reporter für biokatalytische Transformationen eingesetzt. *Quantum Dots* weisen einige herausragende optische Eigenschaften auf, welche sie für eine Anwendung als Reporterfluorophore in Microarray Applikationen prädestiniert erscheinen lassen. Neben sehr hohen Quantenausbeuten und Fluoreszenzintensitäten, die von keinem konventionellen Fluorophor erreicht werden, lässt sich die Wellenlänge des Emissionsmaximums der *Quantum Dots* mittels einer Größenvariation der Partikel in einem weitem Bereich maßschneidern. Die Wellenlänge des Emissionsmaximums nimmt mit wachsender Größe der Partikel zu. Dabei kann die Exzitation über einen weiten Wellenlängenbereich erfolgen, die Extinktionskoeffizienten verhalten sich reziprok zur anregenden Wellenlänge. Eine besonders starke Anregung der *Quantum Dots* lässt sich deshalb im kurzwelligen Bereich des Spektrums realisieren. Im Gegensatz zu konventionellen Fluorophoren existiert ein sehr großer Stokescher *Shift*. Dieses Faktum ist besonders für die Reduzierung des Hintergrundsignals von Bedeutung, da die große Wellenlängendifferenz zwischen Anregungs- und Emissionswellenlänge eine Verwendung von Detektionsfiltern erlaubt, welche die Wellenlängenbereiche der Hintergrundfluoreszenz fast völlig ausschließen. Die Bedeutung der Reduktion des Hintergrundsignals zur signifikanten Erhöhung des dynamischen Bereichs der Analysenmethode wurde bereits im theoretischen Teil der Arbeit (Abschnitt 3.3.6) erläutert. Das breite Anregungsspektrum der *Quantum Dots*, welches sich über einen enormen Wellenlängenbereich erstreckt, weist auch einen signifikanten Vorteil für *Multiplexing*-Versuche auf. Unterschiedliche Proben, welche mit verschiedenen *Quantum Dots* markiert wurden, können simultan mit einer Wellenlänge und damit nur einem Laser angeregt und analysiert werden. Dies senkt den apparativen Aufwand und stellt eine Zeitreduktion des Scanvorganges dar. Die Emissionsspektren der *Quantum Dots* sind zudem, im Gegensatz zum Exzitationsspektrum, eng und symmetrisch über einen kleinen Wellenlängenbereich verteilt, so dass ein *Multiplexing* mit mehreren Markierungen möglich ist sofern entsprechende Emissionsfilter vorhanden sind. Zudem weisen *Quantum Dots* nahezu kein *Photobleaching* auf. Bis dato stehen diese vielversprechenden Fluorophore allerdings erst mit wenigen Modifizierungen zur Verfügung, so dass die mangelnde Biokompatibilität den bislang limitierende Faktor für vielfältige Applikationsmöglichkeiten darstellt.

Die Anwendbarkeit von *Quantum Dots* als Reporterfluorophore für Microarray-Analysen sollte im Rahmen dieser Arbeit evaluiert werden und ein Protokoll zur routinemäßigen Applikation der *Dots* etabliert werden. Es existieren bereits kommerziell erhältliche, an Streptavidin gekoppelte *Quantum Dot*-Partikel (*Q-Dots*, Invitrogen). Aufgrund der außerordentlich starken Streptavidin-Biotin Wechselwirkung und der Existenz von Biotin-markiertem dCTP zur Inkorporation in die cDNA wurde die Verwendung dieses Konjugates favorisiert. Zudem existierte bereits ein großes *know-how* hinsichtlich der Generierung und dem Nachweis von Biotin-markierter cDNA, da diese im routinemäßig durchgeführten TSA-Signalamplifikationskit Anwendung findet. Die verwendeten Streptavidin-funktionalisierten *Quantum Dots* bestehen aus einem kristallinen Kern aus Cadmiumselenid, welcher von einer Hülle aus Zinksulfid umgeben ist. Über eine Polymerbeschichtung ist jeder Partikel von ca. 5-10 Streptavidinmolekülen umgeben, welche direkt über eine Esterbrücke an die Polymerbeschichtung gebunden sind. Für die Etablierung eines Protokolls wurde ein Streptavidinkonjugierter *Quantum Dot* mit einem Emissionsmaximum von 605 nm verwendet. Dieser lässt sich mittels des im konfokalen Laserscanner implementierten Cy3-Lasers anregen und ebenso mittels des Filtersets für das Cy3-Fluorophor detektieren.

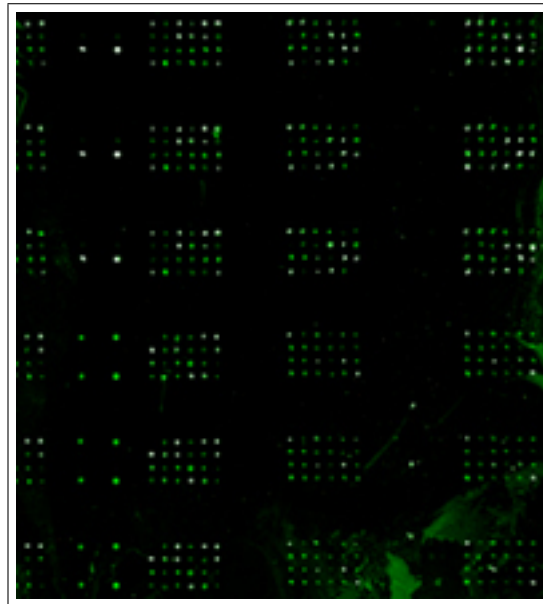


Abb. 76: Scan eines sekretionsspezifischen DNA-Microarrays welcher mit biotinfunktionalisierter cDNA aus 100 μg Gesamt-RNA hybridisiert wurde. Der Nachweis erfolgte über Streptavidin funktionalisierte *Quantum Dots*.

In einem ersten Versuch wurde evaluiert, welches die optimale RNA-Menge für ein *Quantum Dot* Experiment mit dem gewählten System darstellt. Für die Generierung Biotin-markierter cDNA existierte bereits ein etabliertes Protokoll. Allerdings musste eine Modifikation desselben erfolgen, da sich dieses System eines Signalamplifikationsschrittes bedient. Somit war zu erwarten, dass signifikant größere RNA-Mengen für ein Microarray-Experiment benötigt wurden. Für die meisten Microarray-Applikationen, welche auf der Verwendung von Gesamt-RNA basieren und keine Amplifikation des Signales oder der initialen RNA-Menge durchführen, werden 25-100 μg Gesamt-RNA eingesetzt (Skena et al. (1995)). Um diesen Bereich zu erfassen, wurden initiale RNA-Mengen von 10, 50 und 100

μg als Ausgangsmaterial für die cDNA-Synthese evaluiert. Da die *Quantum Dot* Partikel sehr hohe Extinktionskoeffizienten und Quantenausbeuten aufweisen, wurde mit $10\ \mu\text{g}$ auch eine RNA-Menge unterhalb dieses Bereiches evaluiert. Die Gesamt-RNA für dieses Experimente stammt aus einer Probe einer reproduzierbaren Bioreaktorkultivierung des *E. coli* Stammes MG1655 pBADtorAmalE, welche in der mittleren exponentiellen Wachstumsphase entnommen wurde. Die RNA wurde mittels des RNA Isolierungskits der Firma Analytik Jena aufgereinigt und anschließend einem 30 minütigen DNase I-Verdau unterzogen, um noch vorhandene DNA-Kontaminationen zu eliminieren. Nachfolgend wurde das RNA-Isolat mit einer Phenol-Chloroform Extraktion gereinigt, gefällt und mit 75 procentigem Ethanol gewaschen. Das erhaltene RNA-Pellet wurde in nukleasefreiem Wasser resuspendiert und photometrisch Konzentration und Reinheit bestimmt.

Nach der Verifikation der molekularen Integrität durch gelelektrophoretische Analyse erfolgte die Verwendung der RNA für die *Quantum Dot* Experimente. Das Protokoll zur Generierung biotin-markierter cDNA kann dem Anhang entnommen werden (A.1.6). Ein ungefähres Dotierungsverhältnis von Biotin-dCTP zu unmarkierten dCTP wurde aus den etablierten Protokollen für die direkte Markierung von cDNA mittels Cyanin-modifizierten dCTP adaptiert. Es wurde davon ausgegangen, dass der sterische Anspruch des Biotin-Moleküles in etwa dem des Cyanin-Farbstoffes entspricht und somit eine cDNA-Synthese in ähnlichen Inkorporationsfrequenzen resultiert.

Die cDNA wurde in Hybridisierungspuffer resuspendiert und auf sekretionsspezifische *low-density* Microarrays über Nacht appliziert. Die Hybridisierungsparameter entsprachen denen konventioneller Microarrayexperimente. Nach dem Waschen und der Blockierung der Chipoberfläche mit TNB-G Puffer wurde die *Quantum Dot*-Lösung für drei Stunden auf dem Chip inkubiert. Der letzte Waschschrift vor der Applikation der *Quantum Dot*-Lösung erfolgte entgegen dem TSA-Protokoll (A.1.5) mit PBS-Puffer, um die Bindungsparameter für die Streptavidin *Quantum Dot* Konjugate zu optimieren. Nach drei weiteren Waschsritten wurden die Chips getrocknet und gescannt. Der Scan eines generierten Microarrays ist in Abbildung 76 gezeigt.

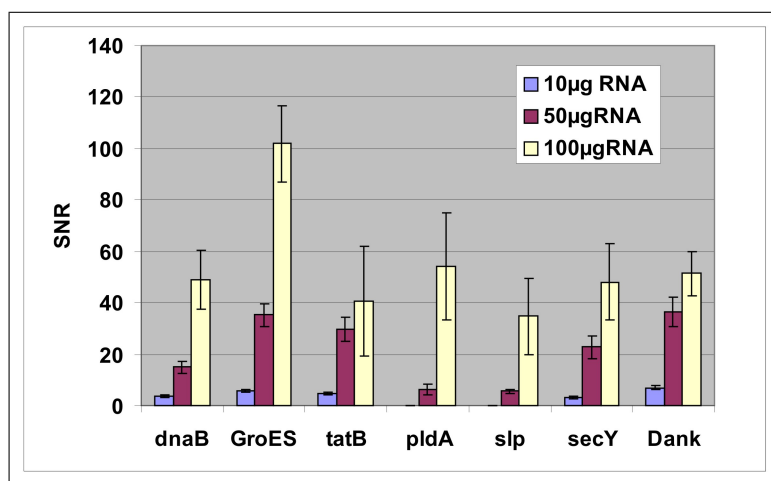


Abb. 77: Unterschiedliche initiale RNA-Mengen wurden für die Synthese Biotin-markierter cDNA eingesetzt und diese mittels Streptavidin-funktionalisierten *Quantum Dots* in einer sekretionsspezifischen *low-density* Microarray-Applikation nachgewiesen. Dargestellt ist das resultierende SNR von sieben repräsentativen Genen des *low-density* Microarrays.

Wie aus der Abbildung 76 ersichtlich ist, wurden intensive Signale erhalten, die sich durch Verwendung entsprechend hoher Scanparameter in die Sättigung bringen ließen (weiße Spots in der Abbildung). Ein erster Blick auf den Chip zeigt ein regelmäßiges Spotmuster der Replikate, was auf eine homogene Hybridisierung und hohe Reproduzierbarkeit der Analyse hindeutet. Eine nähere Analyse erfolgte mittels der quantifizierten Daten des Scans. Die Primäranalyse des Scans erfolgte analog zu der von anderen Microarray-Applikationen (siehe Kapitel 4.2). Aus den ermittelten Daten wurde das SNR als Gütekriterium des Experimentes bestimmt. Ein hohes SNR ist für die Sensitivität und den dynamischen Bereich der Analyse von entscheidender Bedeutung. Es fließen nicht nur die hintergrundbereinigten Signalintensitäten mit in den Wert ein, sondern auch die Standardabweichung des Hintergrundes. Dieser Parameter ist ein Maß für die Homogenität und Signalstärke des Hintergrundes und gibt wichtige Hinweise auf die Qualität des Microarray-Experimentes (Stekel D.: *Microarray Bioinformatics* (2004)).

Sieben Gene wurden stichprobenartig aus der Gesamtheit der 109 auf dem sekretionsspezifischen *low-density* Chip repräsentierten Gene ausgewählt. Die jeweiligen resultierenden SNRs für die drei verschiedenen initialen Ausgangsmengen an RNA sind in Abbildung 77 in einem Vergleich für jedes Gen gezeigt. Aus dem Diagramm ist ersichtlich, dass eine initiale RNA-Menge von 100 μg konsistent für jedes Gen die größten Werte des SNR generiert. Die SNR-Werte, die aus der Verwendung von 50 μg RNA resultieren, liegen wiederum deutlich über denen von 10 μg RNA als Ausgangsmaterial. Für einige Gene wie *pldA* und *slp* fällt das SNR bei der Verwendung von 10 μg RNA unter die Detektionsgrenze ($\text{SNR} = 3$). Daher konnte für diese Gene kein Wert angegeben werden. Die SNR-Werte für 100 μg RNA liegen im Bereich derer von konventionellen Microarrays, die auf der Verwendung von Cyaninfarbstoffen als Reporterfluorophor beruhen (Schadt *et al.* (2000)). Aus den Daten geht eindeutig hervor, dass die Verwendung von 100 μg Gesamt-RNA im evaluierten Bereich die optimalen SNR-Werte generiert. Die Verwendung größerer RNA-Mengen zur cDNA-Synthese ist aufgrund der Limitierung des Probenmaterials und der sinkenden Effizienz der RT-Reaktion bei größeren RNA-Mengen nicht anzustreben.

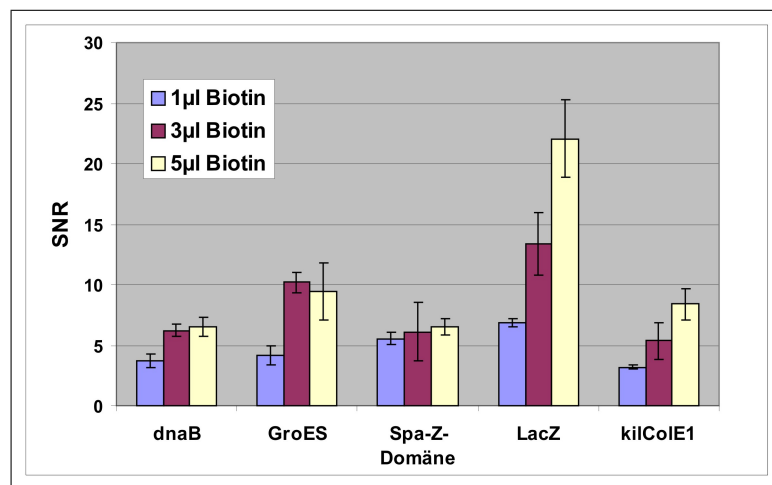


Abb. 78: Die resultierenden SNRs von fünf repräsentativen Genen auf dem sekretionsspezifischen *low-density* Microarray für den Nachweis Biotin-markierter cDNA mittels Streptavidin-funktionalisierter *Quantum Dots* in drei verschiedenen Dotierungsverhältnissen.

In einem nachfolgenden Versuch wurden drei verschiedene Dotierungsverhältnisse von Biotin-dCTP zu dCTP evaluiert, um eine Optimierung der Markierungsdichte der cDNA mit Biotin zu erreichen. Ein zu großes Dotierungsverhältnis kann zu *Quenching*-Effekten zwischen den gebundenen *Quantum Dot* Partikeln führen, wohingegen eine zu geringe Dotierung der cDNA mit Biotin-dCTP zu schwache Signal generiert. Zur exakten Evaluierung des optimalen Dotierungsverhältnisses wurden drei identische Transkriptionsansätzen zur Generierung der cDNA 1, 3 und 5 μl Biotin-dCTP (25 nmol/ μl) zugesetzt. Die daraus resultierende cDNA wurde aufgereinigt und auf sekretionsspezifische *low-density* Microarrays hybridisiert. Nach dem Waschen wurde die Microarrays mit gleichen Konzentrationen an Streptavidin gebundenen *Quantum Dot*-Partikeln inkubiert und nach weiteren Waschstufen getrocknet und gescannt. Aus der Primäranalyse der Daten wurden die SNR-Werte der einzelnen Gene, die auf dem Microarray repräsentiert sind, erhalten und für die drei verschiedenen Dotierungsverhältnisse der cDNA miteinander verglichen. Um eine in Bezug auf die Fragestellung aussagekräftige Auswahl an Genen zu treffen, wurden die fünf Gene gewählt, welche das höchste SNR auf dem Chip mit dem niedrigsten Dotierungsverhältnis Biotin-dCTP zu dCTP aufwiesen. Dieser Chip generierte sehr niedrige SNR-Werte, die für viele Gene unterhalb der Detektionsgrenze lagen. Um einen repräsentativen Vergleich durchzuführen, wurden daher die signifikantesten *Spots* dieses Chips als Auswahlkriterium gewählt. Es handelte sich dabei um die Gene *dnaB*, *GroES*, *Spa-Z-Domäne*, *LacZ* und *kilColE1*. Analog wurde das SNR dieser Gene aus der Datenanalyse der beiden weiteren Microarrays ermittelt. Die Graphik 78 zeigt den Vergleich der SNR-Werte für die drei verschiedenen Chips, welche unterschiedliche Dotierungsverhältnisse repräsentieren.

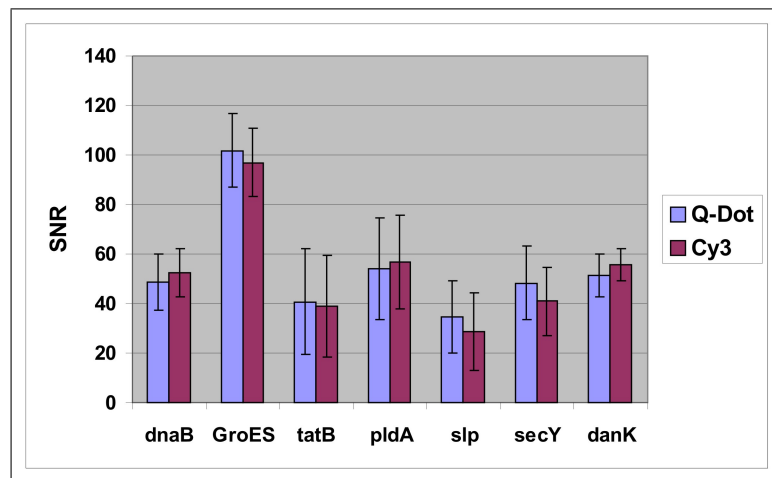


Abb. 79: Vergleich zweier Microarray-Experimente, welche mit Cy3- und *Quantum Dot*-markierter cDNA durchgeführt wurden. Die resultierenden SNRs von sieben repräsentativen Genen auf dem sekretionsspezifischen *low-density* Microarray sind in der Graphik nebeneinandergestellt.

Mit steigendem Dotierungsverhältnis von Biotin-dCTP steigen die generierten SNR-Werte für fast alle Gene konsistent an. Lediglich für das Gen *GroES* wird bei dem mittleren Dotierungsverhältnis das größte SNR beobachtet. Aufgrund der relativ großen Standardabweichung des SNR zwischen den einzelnen Genreplikaten und den marginalen SNR-Unterschied zwischen dem größten und mittleren Dotierungsverhältnis bleibt die allgemeine Tendenz jedoch bestehen. Aus ökonomischen Gründen wurde auf die Evaluierung noch

höherer Dotierungsverhältnisse verzichtet, zumal zu große Inkorporationsfrequenzen zu sterischen Behinderungen der Streptavidin-Bindung an Biotin und zu potenziellen *Quenching*-Effekten führen. Aus den ermittelten Daten wurde ein Dotierungsverhältnis von 2,32: 1 (dCTP: Biotin-11-dCTP) als optimal für die Detektion Biotin-markierter cDNA mittels Streptavidin funktionalisierter *Quantum Dots* erachtet.

In der Abbildung 79 ist der Vergleich der resultierenden SNRs eines konventionellen Cy3-markiertem und einem nach dem etablierten *Quantum Dot* Protokoll prozessierten Chip für sechs verschiedene Gene gezeigt. Beide Chips wurden mit cDNA hybridisiert, welche sich aus der gleichen Menge RNA sowie der gleichen RNA-Probe herleitete. Wie aus der Abbildung ersichtlich, ist das SNR über alle evaluierten Gene vergleichbar. Gleiches gilt für die Standardabweichungen der Genreplikate.

Um einen ersten Schritt in Richtung möglicher *Multiplexing*-Versuche zu unternehmen, wurde ein zweiter Streptavidin-funktionalisierter *Quantum Dot* Partikel mit einer Emissionswellenlänge von 655 nm auf seine Verwendbarkeit für Microarray-Applikationen getestet. Wie aus den evaluierten SNR-Werten von sechs verschiedenen Genen ersichtlich (Abbildung 80), eignen sich auch diese *Quantum Dots* hervorragend als Detektionssystem für Microarrays. Die SNR-Werte liegen noch höher als dies für die *Quantum Dot* Partikel mit einer Emissionswellenlänge von 605 nm der Fall war.

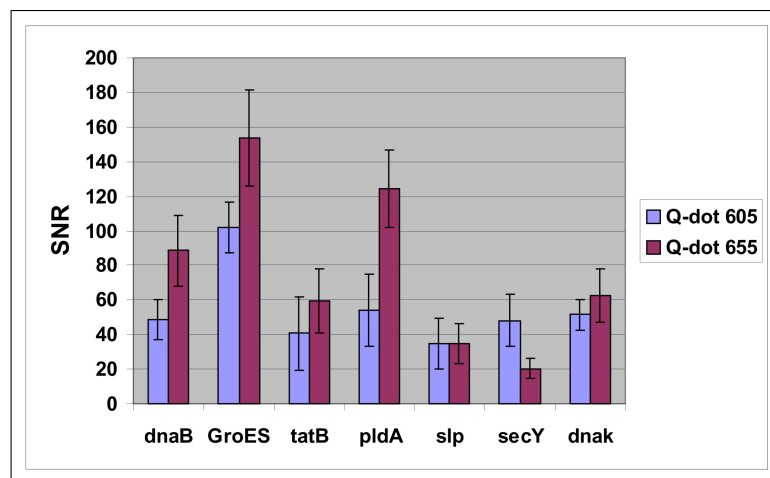


Abb. 80: Die resultierenden SNRs von sieben repräsentativen Genen auf den sekretionsspezifischen *low-density* Microarray für den Nachweis Biotin-markierter cDNA mit Streptavidin-funktionalisierten *Quantum Dot* Partikeln *Q-Dot* 655.

Es konnte durch die beschriebenen Versuche belegt werden, dass sich Streptavidin-funktionalisierte *Quantum Dots* als Reporterfluorophore für Microarray-Applikationen hervorragend eignen. Das Konstatierte gilt um so mehr, da die Exzitation und Detektion der *Quantum Dots* aufgrund der fehlenden Instrumentarisierung unter suboptimalen Bedingungen erfolgte. Die Anregung der *Dots* erfolgte mit einem konventionellen Cy3-Laser, der eine Emissionswellenlänge von 532 nm aufweist. Da die Extinktionskoeffizienten im kürzerwelligen Bereich des Spektrums um mehrere Größenordnungen höher liegen, sollte sich die Signalintensität der *Quantum Dots* drastisch erhöhen lassen. Zudem konnte mittels der Anregungswellenlänge nicht der Vorteil eines großen Stokeschen *Shifts* zur Reduktion der Autofluoreszenz des Hintergrundes evaluiert werden. Weiterhin erschwerend kommt hinzu, dass der Detektionsfilter des Cy3-Kanals des Microarray Scanners eine Bandbreite von 550-

590 nm aufweist. Das Emissionsmaximum des verwendeten *Quantum Dots* liegt jedoch bei 605 nm, so dass nur in der Flanke des Emissionsspektrums des Dots gemessen wird. Dies fällt um so mehr ins Gewicht, da die *Quantum Dots* sehr schmale und symmetrische Emissionsspektren aufweisen, als Folge davon wird nur ein Bruchteil der realen Fluoreszenzintensität detektiert. Äquivalentes gilt für die zweiten verwendeten *Quantum Dot*-Partikel mit einer Emissionswellenlänge von 655 nm. Zieht man diese Fakten mit in Betracht, so erscheint das generierte SNR, welches mit einem konventionellen Reporterfluorophor vergleichbar ist, als sehr guter Wert. Dies zeigt das enorme Potenzial der *Quantum Dots* auf, sofern eine entsprechende Instrumentarisierung auf dem Markt erhältlich ist.

4.11 Etablierung eines Systems zur absoluten Quantifizierung von DNA-Microarrays

Bis dato wird die DNA-Microarray-Technologie fast ausschließlich zur relativen Datenenaquisition zweier oder mehrerer zu vergleichender Zustände verwendet (Skena *et al.* (1995)). Für die Applikation der differentiellen Genexpressionsanalyse wird z.B. das Transkriptionsniveau eines Gens relativ zu einem Referenzzustand (z.B. Phänotyp negativ, Zeitpunkt vor Exposition usw.) angegeben und kann nicht absolut, als Kopienzahl der mRNA, bestimmt werden. Eine absolute Quantifizierung stellt aufgrund der hohen Variabilitäten der einzelnen Prozessschritte eines Microarrayexperimentes nach wie vor eine außerordentliche Herausforderung dar und wurde als chipbasiertes System noch nicht gelöst. Die zur Zeit genauesten Quantifizierungsmethoden (Reassoziationskinetiken, Ribonukleaseprotektions-Assay, quantitative *real-time* PCR) detektieren die zu analysierende Nukleinsäurespezies in homogener Lösung über die Hybridisierung mit einer sequenzspezifischen Sonde. Die Hybridisierung eines Nukleinsäurezielmoleküles mit einem immobilisierten Oligonukleotid an einer heterogenen Phasengrenzfläche fest/flüssig unterscheidet sich signifikant von der Hybridisierung in homogener Lösung befindlicher, komplementärer Nukleinsäuren. So existieren z.B. Studien, welche belegen, dass das Transkriptionsniveau sehr hoch exprimierter und regulierter Gene mittels chipbasierter Analysen nicht richtig erfasst wird. Die ermittelten Genexpressionsquotienten sind niedriger als die Werte, welche z.B. mittels quantitativer *real-time* PCR generiert werden (Pfaffl *et al.* (2007)). Hierin spiegelt sich das unterschiedliche Hybridisierungsverhalten komplementärer Nukleinsäuremoleküle in homogener Phase von dem an einer heterogenen Phasengrenzfläche wieder. Eine mögliche Erklärung liegt in einer schlechteren Zugänglichkeit der immobilisierten Sonde für das *Target* gegenüber der in homogener Lösung erfolgender Reaktion sowie potenziellen *Quenching*-Effekten. Viele bereits gebundene *Targets* bewirken so eine sterische und elektrostatische Abstoßung (negativ geladenes Zuckerphosphat-Rückgrat) auf weitere in räumlicher Nähe befindliche, komplementäre Zielnukleinsäuren. Als Folge dessen wird für hochexprimierte *Targets* eine niedrigere als die real existierende Konzentration bestimmt. Dieses Phänomen wurde auch in der hier vorliegenden Studie anhand des Expressionsniveaus des plasmidcodierten GFP-Genes beobachtet (Abschnitt 4.8.9).

Aufgrund des hohen Parallelisierungsgrades und der Flexibilität des Analysenformates wäre eine chipbasierte Quantifizierung eine äußerst wünschenswerte Applikation, die zu einer enormen ökonomischen Einsparung gegenüber den aufwendigen transkriptionellen Quantifizierungen über *real-time* PCR oder *Northern Blots* führen würde. Kommerziell ist für diese Zwecke ein System des Anbieters Stratagene erhältlich, welches auf einer Quantifizierung über fluoreszenzmarkierte Oligonukleotide basiert. Diese Oligonukleotide sind komplemen-

Tabelle 9: Ausgewählte Gene bzw. Genabschnitte für die *in vitro* Transkription.

Genname	Amplifikatlänge (bp)	Genabschnitt
glyA	961	284-1244
rpoA	473	452-924
dnaK	1720	157-1876
dnaA	1073	89-1161
yidC	1105	95-1199
B0421	663	208-870
tatD	561	10-570
lacA	481	7-487
lacY	786	23-808
ftsH	826	63-888
gfpmut2	719	1-719
rfpExpress	689	1-689

tär zu Kontroll-Oligonukleotiden auf dem Microarray und werden der Hybridisierungslösung zugesetzt. Das Prinzip der Quantifizierung ist allerdings aus mehreren Gründen als kritisch zu bewerten. Die größten Variabilitäten bei Microarray-Experimenten werden dem Schritt der reversen Transkription zugeschrieben (Eickhoff *et al.* (1999)). Eine Quantifizierungsmethode sollte daher obligat diese Schwankungen zu erfassen vermögen. Dies heisst als Konsequenz, dass ein Quantifizierungsstandard möglichst analog der mRNA behandelt werden müsste und diese schwankungsintensiven Schritte simultan mit der zu analysierenden Probe durchläuft. Zudem kann die Signalstärke und Inkorporationsfrequenz sehr stark von der Sequenz des jeweiligen *Templates* und der Länge desselben abhängen. Für zwei unterschiedlich lange mRNAs gleicher Kopienzahl kann das resultierende Signal unterschiedlich ausfallen. Deshalb ist eine Quantifizierung über einen oder wenige Standards als statistisch wenig signifikant zu bewerten. Aus den angeführten Gründen wurde ein System entwickelt, dass auf der Verwendung von mehreren, artifiziell hergestellten RNA-Spezies basiert. Diese können dem Zellpellet vor der RNA-Isolation oder der isolierten RNA zugesetzt werden (*spiken*) und durchlaufen so analog alle Schritte mit denen auch die RNA-Zielmoleküle behandelt werden. Variabilitäten dieser Schritte können so durch das initiale *spiken* gleicher Mengen an RNA für jede analysierte Probe durch Veränderungen im Signalwert der Standards sofort detektiert werden.

Die Darstellung der RNA-Spezies erfolgte dabei mittels der sehr effektiven *in vitro* Transkription. Es wurden Gene bzw. Genabschnitte im Längenbereich von 473 bis 1720 bp gewählt, um ein möglichst großes Spektrum an Transkriptlängen zu realisieren und so der durchschnittlichen Länge und Verteilung von Transkriptlängen der mRNA natürlicher Zellen möglichst nahe zu kommen. Die entsprechenden Genabschnitte wurden so gewählt, dass keine bei der Klonierung bzw. weiteren Prozessierung verwendeten Restriktionsschnittstellen in der Zielsequenz lagen. Tabelle 9 gibt einen Überblick über die ausgewählten Genabschnitte.

Die entsprechenden Zielabschnitte wurden dazu aus Plasmid-DNA bzw. aus cDNA, die sich von der Gesamt-RNA ableitet, mittels PCR hergestellt. Zur Evaluierung und Charakterisierung des Systems wurde ein spezifischer *low-density* Chip mit 13 verschiedenen Oligonukleotiden in jeweils 19 Replikaten entworfen. Die Replikate verteilen sich auf verschiedene Positionen der Chipoberfläche, so dass Positionseffekten Rechnung getragen wird. Zudem ist der Genabschnitt der Serin-Hydroxymethyltransferase (GlyA) doppelt repräsentiert, um einen möglichen Einfluß der Replikatanzahl auf die generierte Signalintensität zu evaluieren. Das *in silico* Design der 50-mer Oligonukleotide erfolgte gemäß der Kriterien von Kane *et al.* (2000). Ein C6-Spacer mit terminaler Aminomodifikation am 5'-Ende der Primer fungierte zur Kopplung der Oligonukleotide mit der aldehydfunktionalisierten Glasoberfläche des Chips. Die Oligonukleotidsynthese erfolgte durch die Firma MWG. Für das *Spotting* der Chips wurden die Oligonukleotide in nukleasefreiem Wasser gelöst und auf eine Konzentration von 100 pmol/ μ l eingestellt. Nachfolgend wurden diese in einer 96 Loch-Platte 1:1 mit *Spotting Buffer* verdünnt und mittels des Affymetrix Arrayers 417 auf der Chipoberfläche deponiert. Die Nachbehandlung der fertig prozessierten Chips kann dem Anhang entnommen werden (A.1.1). Grundsätzlich lassen sich *in vitro* RNA-Transkripte mittels Matrizen aus linearisierter Plasmid-DNA oder PCR-Produkten herstellen, welche über einen Promoter für eine phagenspezifische RNA-Polymerase (z.B. T7, SP6, T3) verfügen.

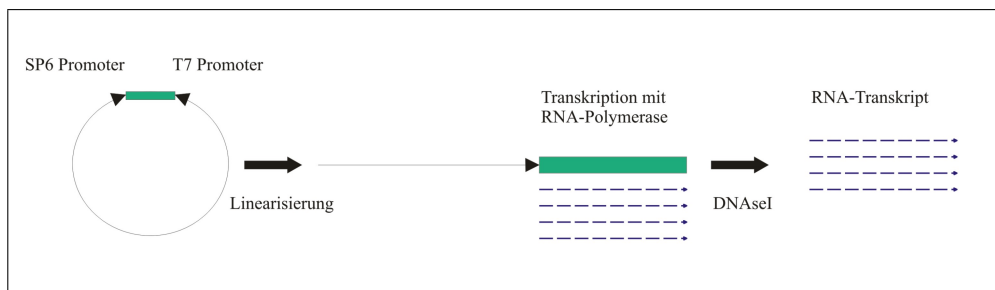


Abb. 81: Synthese von *in vitro* Transkripten. Der Transkriptionsvektor mit der Zielsequenz und den phagenspezifischen Promoterregionen wird linearisiert. Mittels einer phagenspezifischen RNA-Polymerase wird die stromabwärts des Promoters befindliche Zielsequenz in RNA-Transkripte umgeschrieben. Durch eine DNase I-Behandlung wird das DNA-Template entfernt.

Es existieren verschiedene Transkriptionsvektoren für diese Applikation, welche über eine Multi-Klonierungsstelle (MCS) stromabwärts eines solchen phagenspezifischen Promoters verfügen. Stromabwärts der MCS ist zumeist ein Terminator für die jeweilige Polymerase lokalisiert. Das zu transkribierende *Insert* wird in die MCS kloniert und kann nach Linearisierung an einer Stelle stromabwärts des 3'-Endes seiner Sequenz in RNA konvertiert werden (Melton *et al.* (1984)). Das Prinzip der *in vitro* Transkription zur Generierung einzelsträngiger RNA-Transkripte ist in Abbildung 81 dargestellt.

Daneben besteht die Möglichkeit mittels PCR die phagenspezifischen Promotoren über das 5'-Ende der Primer in die Amplifikate zu inkorporieren und so lineare *Templates* für die *in vitro* Transkription zu generieren (Logel *et al.* (1992), Bales *et al.* (1993)). Der experimentelle Aufwand ist bei dieser Methode etwas geringer, allerdings generieren linearisierte Plasmide signifikant größere Mengen an RNA-Transkript als *Templates*, welche aus einer PCR-Reaktion hervorgegangen sind. Zudem resultieren aus den langen 5'-Überhängen der Primer oftmals schwierige und suboptimale Amplifikationsbedingungen. Aus diesem Grunde wurde für die Generierung der *in vitro* RNA-Transkripte die Klonierungsvariante gewählt.

Für die Klonierung der Zielsequenzen in den Transkriptionsvektor wurden drei verschiedene Strategien evaluiert, um so die effizienteste Möglichkeit für die simultane Generierung vieler Transkriptionsvektoren zu bestimmen:

1. direktionale Klonierung
2. *blunt-end*-Klonierung
3. *cut-ligation*-Klonierung

Die Abbildung 82 gibt einen Überblick über die verwendeten Strategien und die notwendigen Prozessschritte. Auf der linken Seite ist die Klonierungsstrategie über kohäsive Enden aufgezeigt. Für diese direktionale Klonierung werden die benötigten Restriktionsschnittstellen über das 5'-Ende der PCR-Primer eingefügt. Die PCR-Amplifikation erfolgt mittels *Taq*-Polymerase, wohingegen für die Klonierungen der Strategien 2 und 3 *Pfu*-Polymerase verwendet wird, um keinen unspezifischen Adenosin-Überhang an den 3'-Enden des Amplifikats zu generieren. Zur Realisierung von Strategie 1 werden Vektor und die zu insertierende, amplifizierte Sequenz mit den gleichen Restriktionsendonukleasen prozessiert (in diesem Fall *SalI* und *NcoI*) und der Vektor durch alkalische Phosphatasebehandlung dephosphoryliert. Nach einer Aufreinigung der gewünschten Fragmente aus dem Gel werden diese mittels T4-DNA-Ligase miteinander ligiert.

Für die Klonierung über glatte Enden wird der Klonierungsvektor mit einem Restriktionsenzym geschnitten, welches keinen Überhang eines Stranges erzeugt (im konkreten Fall *EcoRV*), dephosphoryliert und mit der PCR-generierten Zielsequenz ligiert. Die *cut-ligation* Methode, welche hier nicht gesondert dargestellt ist, fasst die Prozessschritte des Restriktionsverdau und der Ligation in einem Schritt zusammen. Zudem ist eine Dephosphorylierung des Vektors überflüssig. Der Vektor und die Zielsequenz werden durch Zugabe von Ligase und einer *blunt-end* schneidenden Restriktionsendonuklease direkt in einem Schritt miteinander ligiert.

Für die Linearisierung der Transkriptionsplasmide sollten Restriktionsschnittstellen verwendet werden, welche einen 5'-Überhang generieren. Ein 3'-Überhang kann als Primer für die RNA-Polymerase fungieren und so zu langen *run-off* Transkripten des falschen Stranges führen (Schenborn *et al.* (1985)). Da in langen Transkripten signifikante Mengen an rNTPS inkorporiert werden, kann diese Reaktion, selbst wenn sie nur mit geringer Frequenz erfolgt, eine drastische Reduzierung der Transkriptionseffizienz bedeuten. Aus dem gleichen Grund sollten auch Restriktionsschnittstellen gewählt werden, welche von der entsprechenden Restriktionsendonuklease vollständig prozessiert werden. Ein geringer Anteil an nicht-linearisiertem *Template* führt zu einer Transkription der gesamten Vektorsequenz und so zu drastisch reduzierten Ausbeuten des gewünschten Transkripts. Gründe dafür liegen in der erhöhten Transkriptionseffizienz ringförmiger Plasmid-DNA gegenüber derer linearisierten Plasmide. Für die Auswahl des Transkriptionsvektors muss daher bedacht werden, dass sich stromaufwärts und stromabwärts einer *blunt-end* Restriktionsschnittstelle (zur Realisierung der *blunt-end* und *cut-ligation* Strategien) eine Schnittstelle für eine Restriktionsendonuklease befindet, welche einen 5'-Überhang generiert und eine möglichst vollständige Linearisierung gewährleistet. Die Flankierung der *blunt-end* Schnittstelle ist notwendig, damit eine Transkription beider Stränge realisiert werden kann (sowohl *sense* als *anti-sense* Orientierung). Zusätzlich muss beachtet werden, dass die Restriktionsschnittstellen nicht in

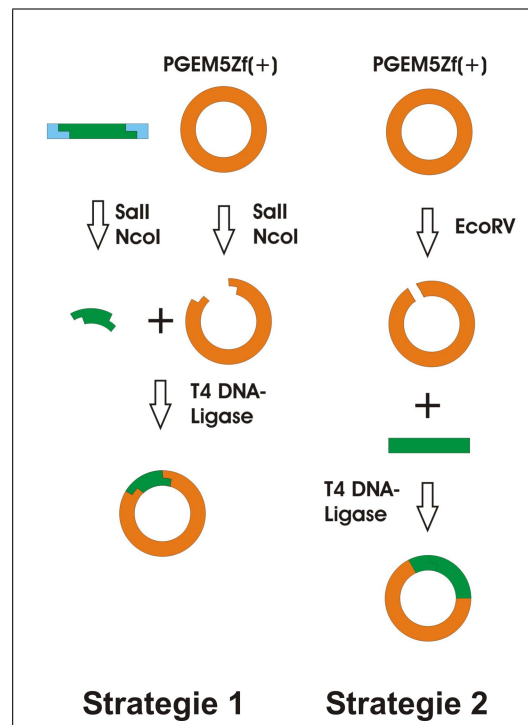


Abb. 82: Zwei Klonierungsstrategien zur Generierung der *in vitro* Transkriptionsplasmide, auf der linken Seite ist eine direktionale Klonierung über kohäsive Enden gezeigt, wohingegen die Strategie auf der rechten Seite auf einer *blunt-end* Klonierung basiert.

den Zielsequenzen, welche zur Transkription eingesetzt werden, enthalten sind. Des weiteren dürfen in den Zielsequenzen keine Terminationscodons für die RNA-Polymerase existieren.

Ein Vektor, der diese Voraussetzungen erfüllt, ist der Transkriptionsvektor pGEM5ZF(+), welcher kommerziell von der Firma Promega vertrieben wird. In der Mitte der MCS lokalisiert befindet sich die *blunt-end* Restriktionsschnittstelle für EcoRV. Stromab und stromaufwärts wird diese flankiert von den Schnittstellen NdeI, NcoI und Sall, welche einen 5'-Überhang generieren, und von den Restriktionsendonukleasen vollständig prozessiert werden. Auf Basis dieser theoretischen Überlegungen erfolgte die Konstruktion der Transkriptionsplasmide.

Die amplifizierten Sequenzen wurden in die MCS des Vektors pGEM5Zf(+) kloniert, so dass sie von beiden Seiten vom SP6- bzw. T7-Promoter flankiert werden. Somit lässt sich sowohl die Synthese des *sense*- als auch des *anti-sense*-Stranges realisieren. Besonders für die Nutzung des Systems für weitere Applikationen wie z.B. *Northern Blots* ist diese Möglichkeit von besonderer Bedeutung. Aufgrund der Konversion der mRNA (*sense*) in cDNA (*anti-sense*), wie sie für die meisten chipbasierten Systeme zur differentiellen Genexpressionanalyse verwendet wird, liegen die Oligonukleotide vieler kommerzieller Chips im *sense-Design* und damit komplementär zur generierten cDNA vor (Frederiksen *et al.* (2003)). Die für *Northern Blots* benötigten RNA-Sonden müssen allerdings *anti-sense* Orientierung aufweisen, da sie für die direkte Detektion der *sense*-mRNA verwendet werden.

Die drei verschiedenen Klonierungsstrategien wurden anhand der Insertion eines Sequenzabschnittes aus dem tatD-Gen in den Transkriptionsvektor pGEM5Zf(+) hinsichtlich ihre Effektivität getestet. Für die Klonierung über kohäsive Enden wurden den 5'-Enden der

verwendeten Primer die Sequenzen für die Restriktionsschnittstellen SallI und NcoI angefügt. Zusätzlich wurden die Primer um weitere 10 bp am 5'-Ende verlängert, um sicherzustellen, dass die Schnittstellen von der entsprechenden Restriktionsendonuklease prozessiert werden. Terminal lokalisierte Restriktionsschnittstellen werden häufig nicht oder nur mit stark verringerter Effizienz geschnitten, da die Restriktionsendonuklease einen Überhang zur Bindung benötigt. Die generierten PCR-Amplifikate wurden mittels des *PCR-Purification Kits* (Quiagen) aufgereinigt und für 3 Stunden wurde ein Restriktionsdoppelverdau mit SallI und NcoI durchgeführt. Parallel dazu wurde der Klonierungsvektor einer identischen Behandlung unterzogen, zusätzlich wurden 0,01 Units Alkalische Phosphatase pro Pikomol terminaler Phosphatgruppen zugesetzt. Die Restriktionsansätze wurden nach der Hitzeinaktivierung auf ein Agarosegel aufgetragen und die gewünschten DNA-Fragmente mittels des Wizard SV Gelisolutionskits (Promega) aus dem Gel aufgereinigt. 112 ng *Insert* und 200 ng Vektor, entsprechend einem molaren Verhältnis von 3:1, wurden mittels T4-DNA-Ligase für 4 Stunden bei 16°C ligiert und der Ligationsansatz in chemisch kompetente XL1-Blue Zellen transformiert (Hitzeschock, 42°C). Für die beiden Strategien, welche auf einer *blunt-end* Klonierung in den Zielvektors pGEM5Zf(+) basieren, kann das PCR-Amplifikat nach der Aufreinigung über das *PCR-Purifikation Kit* direkt für die Klonierung eingesetzt werden. Für die Klonierungsstrategie mittels konventioneller *blunt-end* Klonierung wird der Klonierungsvektor mit EcoRV geschnitten und anschließend mit alkalischer Phosphatase für eine Stunde bei 37°C behandelt. Dieser Prozessschritt ist obligat durchzuführen, da die Religationstendenz eines einfach geschnittenen Vektors sehr groß ist und die geringe Effizienz der *blunt-end* Klonierung sonst kaum ausreichend ist, um eine hinreichend hohe Anzahl an Insertionsreaktionen zu ermöglichen. Zur Eliminierung der verwendeten Enzyme wurde der Vektor vor der Ligation nochmals über das *PCR Purification Kit* aufgereinigt. Die Ligation mit dem mittels PCR generierten *Inserts* erfolgte bei 4°C über Nacht und in einem höheren molaren Vektor-*Insert* Verhältnis (5:1), um der geringeren Effizienz der Ligation Rechnung zu tragen. Zudem wurden mit fünf Units je Ligationsansatz die fünffache der üblichen Menge an T4-DNA-Ligase pro Reaktionsansatz zugesetzt. Zur Realisierung der *cut-ligation* Methode wurde das *tatD-Insert* in einem Massenverhältnis von 2:1 (400 ng *Insert* mit 200 ng Vektor) mit dem Klonierungsvektor vermischt und neben T4-DNA-Ligase ein Überschuss an EcoRV (10 Units) zugesetzt. Diese Methode nutzt die Tatsache, dass Religationen des Vektors die *blunt-end* Restriktionsschnittstelle rekonstituieren, nicht jedoch die Ligation des *Inserts*. Durch den Überschuss an appliziertem Restriktionsenzym werden religierte Vektorsequenzen erneut prozessiert. Ausschließlich die Ligation des *Inserts* in die Vektorsequenz zerstört die *blunt-end* Schnittstelle und verhindert deren Prozessierung durch die Restriktionsendonuklease. Die inhärent niedrige Effektivität der Ligation glatter Enden wird so umgangen. Zusätzlich zum Ligationsansatz appliziertes PEG 6000 dient durch seine hygroskopischen Eigenschaft dazu, die lokale Konzentration an Reaktanden zu erhöhen. Zudem wird bedingt durch die Erhöhung der Viskosität die Brownsche Molekularbewegung erniedrigt, so dass der Anteil der reaktiven Stöße steigt. Eine Evaluierung ergab, dass sich als Puffersystem für beide Enzyme der Tango-Universalpuffer (Fermentas) der Restriktionsendonuklease eignet. Dieser generiert Bedingungen, unter denen auch die T4-DNA-Ligase ohne Aktivitätseinbußen katalytisch aktiv ist; ein Zusatz von ATP muss in diesem Fall separat erfolgen. Für die *cut-ligation* Methode wurde analog zur konventionellen *blunt-end* Klonierung fünf Units T4-DNA-Ligase pro Ligationsansatz verwendet. Der Ligationsansatz wurde für drei Stunden bei Raumtemperatur gehalten und anschließend bei 4°C über Nacht inkubiert.

Tabelle 10: Evaluierung der drei Klonierungsstrategien zur Generierung eines tatD-spezifischen Transkriptionsvektor hinsichtlich Anzahl und prozentualen Anteil an positiven Kolonien sowie dem experimentellen Aufwand.

Klonierungsmethode	Kohäsive Enden	<i>blunt-end</i>	<i>cut-ligation</i>
Positive Kolonien	54 (+/-)19	25 (+/-)7	183 (+/-) 52
Prozentualer Anteil positiver Kolonien	85	96	81
Arbeitsaufwand	hoch	mittel	gering

Für alle drei verschiedenen Ligationsansätze erfolgte die Transformation in chemisch kompetenten XL1-Blue Zellen mittels Hitzeschock bei 42°C. 150 µl der Transformationsansätze wurden auf ampicillinhaltigen Agarplatten über Nacht im Brutschrank bei 37°C inkubiert. Nachfolgend wurde die Anzahl der entstandenen Kolonien und der prozentuale Anteil positiver Kolonien ermittelt. Positive Kolonien sind aufgrund des Blau-Weiss-*Screenings* leicht anhand ihrer Farblosigkeit zu identifizieren. Der Versuch wurde als Triplikat wiederholt. Die positiven Kolonien wurden gepickt, über Nacht vermehrt und eine Kolonie-PCR sowie nach Plasmidisolierung ein analytischer Restriktionsverdau zur Verifikation des Klonierungserfolges unternommen.

Tabelle 10 zeigt die Ergebnisse der Klonierung. Es sind die Anzahl der erhaltenen Kolonien, der prozentuale Anteil an positiven Klonen und eine Bewertung hinsichtlich des experimentellen Aufwandes angegeben. Jede Strategie kann zielführend zur Generierung der gewünschten Transkriptionsplasmide verwendet werden. Allerdings unterscheiden sich die Strategien signifikant hinsichtlich ihres Zeit- und Arbeitsaufwandes. Die *cut-ligation* Methode generiert hohe prozentuale Anteile an positiven Klonen und beinhaltet den Vorteil, die Prozessschritte des Restriktionsverdau und der Ligation in einem Schritt durchzuführen. Zudem muss so keine aufwendige Isolierung von *Insert*- und Vektor-DNA aus dem Agarosegel vorgenommen werden und die Behandlung des Vektors mit alkalischer Phosphatase entfällt. Die Klonierung über kohäsive Enden stellt in dieser Hinsicht den größten Arbeitsaufwand dar, beinhaltet aber den Vorteil einer direktionalen Klonierung. Durch das aufwendigere Primerdesign und schwierigere Amplifikationsbedingungen der PCR wird dieser Vorteil allerdings überkompensiert, zumal eine Verifikation des generierten Transkriptionsvektors mittels Sequenzierung ohnehin obligat zu erfolgen hat. Die konventionelle *blunt-end* Klonierung zeigt keine Vorteile gegenüber der *cut-ligation* Methode, beinhaltet aber einen größeren experimentellen Aufwand. Gerade unter Berücksichtigung der großen Anzahl an benötigten Transkriptionsvektoren ist damit eindeutig die *cut-ligation* Methode zu favorisieren. Mittels einer Eintopfreaktion lässt sich so mit hoher Effizienz eine *blunt-end* Klonierung realisieren. Sämtliche Transkriptionsplasmide wurden daher mittels dieser Strategie generiert.

Zuerst wurden dazu die verschiedenen Zielsequenzen mittels *Pfu*-Polymerase in einer PCR amplifiziert (Abbildung 83). Für die Auswahl der amplifizierten Genabschnitte war entscheidend, dass innerhalb dieses Bereichs keine Schnittstellen für die zur Linearisierung verwendenden Restriktionsendonukleasen enthalten waren. Die *Pfu*-Polymerase weist eine

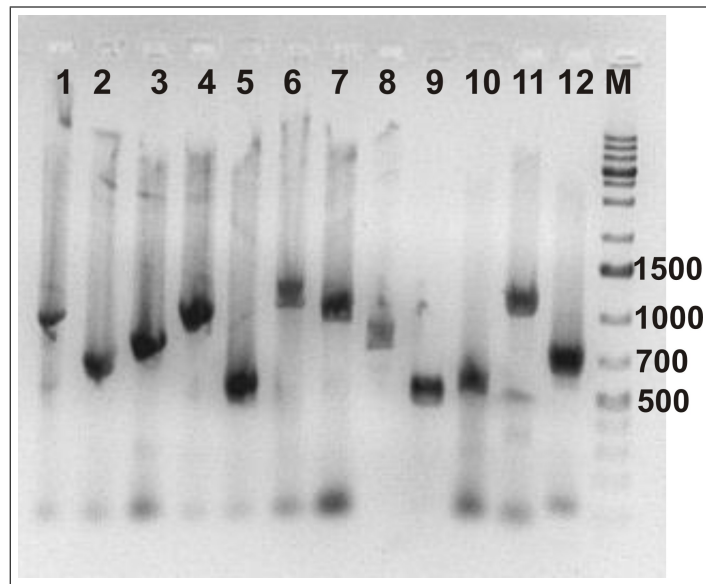


Abb. 83: *Inserts* für die *in vitro* Transkription auf einem 1,5 prozentigem Agarosegel. Der Reihe nach (Spuren 1-12) sind folgende PCR-generierten *Inserts* aufgetragen: *ftsH*, *tatD*, signalpeptidase 2, *secD*, *rpoA*, *yidC*, *glyA*, *lacY*, *lacA*, *dnaK*, *dnaA* und GFP. Für die Zuordnung der Produktlängen wurde ein 1 kb Standard auf das Gel appliziert, die Spezifizierung der relevanten Markerbanden kann dem Gelbild entnommen werden.

geringere Toleranz gegenüber Abweichungen von der optimalen *Annealing*-Temperatur auf als dies z.B. für die *Taq*-Polymerase der Fall ist (Sambrook: Molecular Cloning (2003)). Deshalb wurde zuerst für alle verwendeten Primerpaare die *Annealing*-Temperatur optimiert. Die verwendeten Amplifikationsbedingungen sind dem Anhang (A.1.20) zu entnehmen. Beispielhaft ist dieser Optimierungsschritt und das enge Temperaturoptimum für das *dnaK*-*Insert* in Abbildung 84 gezeigt. Die so generierten Amplifikate wurden auf einem 1,5 prozentigen Agarosegel analysiert, Abbildung 83 zeigt das Ergebnis.

Die unterschiedlichen Produktlängen der Amplifikate sind deutlich zu erkennen, es handelt sich dabei um einzelne, diskrete Banden. Die Größe der Amplifikatbanden stimmen mit den erwarteten theoretischen Produktlängen überein. Deshalb kann von einer spezifischen Amplifikation ausgegangen werden. Lediglich für das längste Produkt *dnaK* konnte kein Amplifikat in der erwarteten Länge erhalten werden. Deshalb wurde diese PCR mit einer längeren Primerverlängerungsphase wiederholt. Erst eine Primer-Elongationsdauer von 4 Minuten je Zyklus führte zu einem Amplifikat der gewünschten Produktlänge. Die Optimierung der PCR hinsichtlich der *Annealing*-Temperatur für *dnaK*-spezifische Primer ist in Abbildung 84 gezeigt.

Die Amplifikate wurden mittels des *PCR Purification Kits* aufgereinigt und mit 50 μ L EB-Puffer von der Säule eluiert. Vor dem Waschschrift wurde der PB-Puffer für 5 Minuten auf der Säule belassen. Dieser zusätzliche Inkubationsschritt dient dazu, verschieden Salze von der Säule zu waschen, welche mit der nachfolgenden, salzsensitiven *blunt-end* Klonierung interferieren. Nach der Elution wurde spektrophotometrisch die Konzentration und Reinheit der Amplifikate bestimmt. Die so präparierten Amplifikate wurden direkt in der *cut-ligation* Reaktion verwendet. Der Ligationsansatz wurde für drei Stunden bei Raumtemperatur und anschließend bei 4°C über Nacht inkubiert. 3 μ L Ligationsansatz wurden

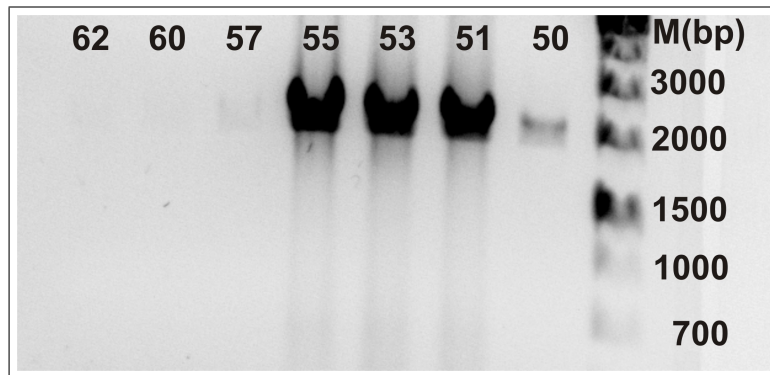


Abb. 84: Optimierung der *Annealing*-Temperatur für die PCR zur Generierung des *dnaK*-*Inserts*. Gezeigt ist ein 1,5 prozentiges Agarosegel mit den aufgetragenen Amplifikationsansätzen. Die jeweiligen *Annealing*-Temperaturen (in Grad Celsius) sind für jede Spur angegeben, gleiches gilt für die Banden des 1 kb Standards.

in chemisch kompetente XL1-Blue Zellen mittels Hitzschock transformiert, auf Ampicillin-haltigen Agarplatten ausplattiert und über Nacht bei 37°C im Brutschrank inkubiert. Weiße Kolonien wurden gepickt und in LB-Medium vermehrt. Eine Verifikation der Klonierung erfolgte über Kolonie-PCR und nach der Plasmidpräparation durch analytischen Restriktionsverdau. Die gelelektrophoretische Analyse der generierten Transkriptionsplasmide und deren Linearisierung ist in Abbildung 85 gezeigt.

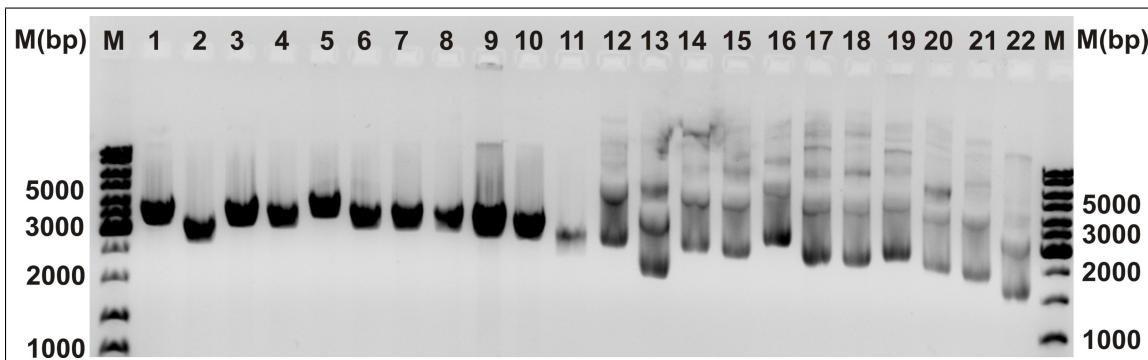


Abb. 85: Gelelektrophoretische Analyse der generierten Transkriptionsplasmide und deren Linearisierung durch die Restriktionsendonuklease *SalI* auf einem 1,5 prozentigem Agarosegel. Auf den Spuren 1-10 sind die durch Verdau mit *SalI* linearisierten Plasmide aufgetragen, als Referenz ist auf Spur 11 der mit *EcoRV* linearisierte Klonierungsvektor *pGEM5Zf(+)* gezeigt. Von links (Spur 1) nach rechts (Spur 10) sind die linearisierten Plasmide mit den *Inserts*: *yidC*, *rpoA*, *glyA*, *secD*, *dnaA*, *dnaK*, *ftsH*, *lacY*, *sig2*, *tatD* und *lacA* gezeigt. Die angegebene Reihenfolge gilt auch für die Spuren 12-21, auf denen die jeweiligen Plasmidpräparationen der generierten Transkriptionsplasmide aufgetragen sind. Spur 22 zeigt den Klonierungsvektor *pGEM5Zf(+)*.

Wiederum fungiert der Klonierungsvektor (Spur 22) als Referenz. Die neu generierten Plasmide unterscheiden sich deutlich in ihrer Größe von dem Ursprungsvektor, eine Verschiebung der Banden in den höhermolekularen Bereich ist auszumachen. Durch die Linearisierung mit dem Restriktionsenzym *SalI* werden einzelne, diskrete DNA-Banden erhalten, die

sich in ihrer Länge von den verschiedenen Transkriptionsplasmiden signifikant unterscheiden und eine größere Länge aufweisen als der linearisierte Klonierungsvektor. Die Linearisierung erfolgt quantitativ, wie aus der Existenz nur einer einzelnen Bande zu erkennen ist. Die theoretische Größenordnung und Verteilung der einzelnen Transkriptionsplasmide wird in der Gelanalyse wiedergefunden. Eine Sequenzierung der Plasmidpräparationen bestätigte das positive Ergebnis der gelelektrophoretischen Analyse.

Da das *Insert* in zwei Orientierungen in den Vektor gelangen kann und diese Orientierung für die nachfolgende Transkription des richtigen Stranges von essentieller Bedeutung ist, wurden die positiven Plasmidpräparationen sequenziert. Die Oligonukleotide auf dem applikationsspezifischen *low-density* Chip wurden in *sense*-Orientierung designed, deswegen erfolgte die Generierung der *in vitro* RNA-Transkripte ebenfalls in der *sense*-Orientierung, um so beim Umschreiben in cDNA eine komplementäre Zielsequenz zu den Chipoligonukleotiden zu erhalten. War der *sense*-Strang des klonierten *Inserts* in der gleichen Orientierung wie die Transkriptionsrichtung des T7-Promoters, so erfolgte die Linearisierung mit Sall und T7 RNA-Polymerase wurde für die Transkription verwendet. Ein in der anderen Orientierung kloniertes *Insert* wurde mittels SP6 RNA-Polymerase transkribiert, und die vorhergehende Linearisierung erfolgte durch Inkubation mit NcoI (Abbildung 86).

Die Linearisierung der Templates erfolgte mit einer Unit Restriktionsendonuklease pro μg DNA für mindestens drei Stunden. Nachfolgend wurden die Restriktionsendonukleasen durch 20 minütige Inkubation bei 65°C hitzeinaktiviert. Eine vollständige Prozessierung, erkennbar an einer einzelnen diskreten Bande, wurde durch Auftragung des Restriktionsansatzes auf einem Agarosegel visualisiert. Mittels einer Phenol-Chloroform Extraktion wurden die Restriktionsenzyme entfernt und Nuklease-freie Bedingungen für die *in vitro* Transkription geschaffen. Bei den Plasmidpräparationen zur Gewinnung der *in vitro* Transkriptionsplasmide wird RNase A zum Verdau der zelleigenen RNA verwendet. Diese muss obligat und vollständig aus dem Transkriptionsansatz entfernt werden. Die so aufgereinigten, linearisierten Plasmide werden mit 5 M NaCl in eiskaltem Ethanol über Nacht quantitativ gefällt und durch Zentrifugation pelletiert. Nach einem Waschschritt mit 75 prozentigem Ethanol wurden die Plasmide bei 37°C getrocknet und in $20\ \mu\text{L}$ nukleasefreiem Wasser gelöst. Es wurde photometrisch die Konzentration und Reinheit am Spektralphotometer Nanodrop ND 1000 bestimmt.

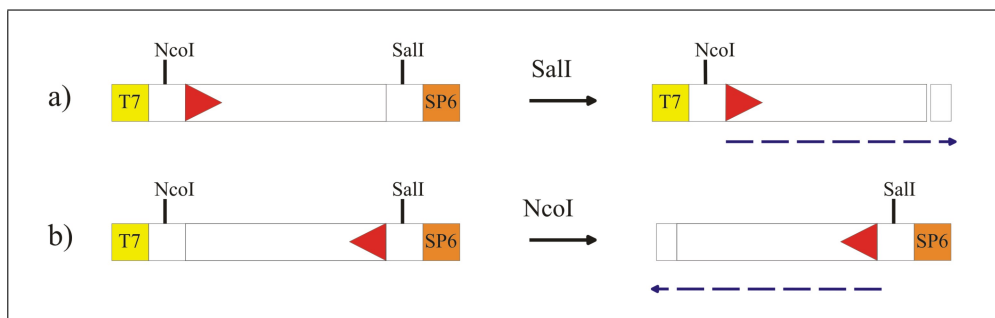


Abb. 86: Synthese von sense-RNA aus den generierten Transkriptionsplasmiden. Anhängig von der 5'-3' Orientierung des *sense* Stranges des Inserts im Vektor, angegeben durch einen roten Pfeil, erfolgt die Linearisierung mit Sall oder NcoI. Nachfolgend wird die *in vitro* Transkription mit T7 oder SP6 RNA-Polymerase durchgeführt.

Die nachfolgende Transkription erfolgte mit $1\ \mu\text{g}$ Plasmid als Matrize, das dazu verwen-

dete Transkriptionsprotokoll kann dem Anhang entnommen werden. Um die genaue Konzentration der generierten RNA aus den Transkriptionsansätzen bestimmen zu können, wurde dieser nach erfolgter Reaktion einem 20 minütigen DNase I-Verdau unterzogen, mittels einer Phenol-Chloroform Extraktion aufgereinigt und mit 5 M NaCl in Ethanol gefällt. Durch diese Prozessierung werden die Ribonukleotide und die DNase I entfernt. Gleiches gilt auch für das im Transkriptionspuffer enthaltene Spermidin, welches eine Visualisierung der RNA auf dem Agarosegel unter denaturierenden Bedingungen verhindert. Spermidin reagiert mit dem Denaturierungsmittel Formaldehyd unter Bildung eines Polymeres, welches im Gel als heller, hochmolekularer Schmier erscheint und in der Abbildung 87 wiedergegeben ist.

Die Elektrophorese der *in vitro* hergestellten RNA sollte unter denaturierenden Bedingungen stattfinden, da die RNA aufgrund ihrer Einzelsträngigkeit und der großen thermodynamischen Stabilität der RNA-Duplexe in höherem Maße intramolekulare und intermolekulare Sekundärstrukturen ausbildet. Das größte Auflösungsvermögen bietet dabei ein denaturierendes Agarosegel, dem Formaldehyd sowohl im Gel als auch im Auftragspuffer zugegeben wird. Hinreichend hohe Auflösungen werden jedoch auch durch Zugabe des Formaldehyds zum Probenpuffer und eine 10 minütige Inkubation bei 70°C erreicht. Abbildung 87 zeigt auf den Spuren 1 und 2 eine mittels *in vitro* Transkription generierte RNA-Spezies, welche sowohl mit denaturierenden (Spur 2) als auch mit nicht-denaturierenden Probenpuffer (Spur 1) gelelektrophoretisch analysiert wurde. Es handelt sich dabei um den 663 bp langen Sequenzabschnitt aus dem Gen der Signalpeptidase 2, welcher in den pGEM-Vektor kloniert und anschließend transkribiert wurde. Unter denaturierenden Bedingungen ist eine diskrete Bande etwas unterhalb 700 bp zu sehen. Unter nicht-denaturierenden Bedingungen hingegen sind mindestens vier verschiedene Banden zu erkennen, welche Hinweise auf die Anzahl und Art der verschiedenen Sekundärstrukturen geben.

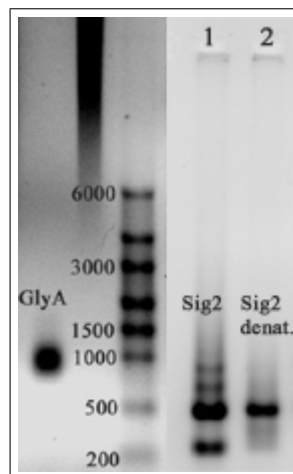


Abb. 87: Gelelektrophoretische Analyse des Transkriptionsansatzes einer glyA RNA-Spezies (linke Spuren) vor und nach Entfernung des Spermidins auf einem 1,5 prozentigem Agarosegel. Spermidin bildet mit dem im Denaturierungspuffer enthaltenen Formaldehyd ein hochmolekulares Polymer, welches als Schmier auf dem Gel erscheint. Die diskrete Bande wurde nach Entfernung des Spermidins mittels Phenol-Chloroform-Extraktion erhalten. Auf der rechten Seite ist eine *in vitro* RNA Spezies, welche spezifisch für einen 663 bp langen Sequenzabschnitt aus dem Gen der Signalpeptidase 2 ist, aufgetragen. Spur 1 zeigt die RNA ohne vorherige Denaturierung, Spur 2 dieselbe unter denaturierenden Bedingungen mit Formaldehyd-Zusatz.

Tabelle 11: cDNA-Konzentrationen, Konzentrationen an inkorporiertem Cyanin5-Farbstoff und sich daraus berechnende Inkorporationsfrequenzen für vier verschiedene *in vitro* Transkripte.

Transkript	RNA-Menge (μg)	cDNA-Konz. $\text{ng}/\mu\text{l}$	FOI	Konz. Cy5 $\text{pmol}/\mu\text{l}$
rpoA	18,58	58,42	16,97	3,44
lacA	19,94	81,97	12,39	3,59
yidC	18,71	62,89	16,83	3,71
glyA	22,04	53,43	16,37	3,07

Für die angegebene Sequenz lassen sich z.B. mittels des Programmes MFold die thermodynamisch stabilsten Sekundärstrukturen berechnen. So läßt sich aus einem Vergleich der Bandenintensitäten der einzelnen Spezies ermitteln, um welche es sich handeln könnte. Zu Bedenken ist hierbei allerdings, dass auch kinetische Parameter in die Bildung der verschiedenen Sekundärstrukturen mit einfließen. RNA-Banden, welche unterhalb der eigentlichen Transkriptlänge erscheinen, sind sehr wahrscheinlich das Resultat intramolekularer Sekundärstrukturen. Durch die etwas kompaktere Form migrieren diese schneller als das lineare *Template*. Dahingegen sollten Banden die oberhalb der eigentliche Transkriptlänge erscheinen das Resultat von intramolekularen Sekundärstrukturen (Oligomere) sein.

In einem ersten Testversuch wurden die vier unterschiedlichen *Templates* glyA, rpoA, lacA und yidC in RNA umgeschrieben und aufgereinigt. Es wurden zwischen 18 und 22 μg RNA aus einem μg linearisiertem *Template* erhalten. Die *in vitro* Transkription verläuft mit großer Effektivität, und in vielen aufeinanderfolgenden Transkriptionsrunden entstehen aus einem *Template*-Molekül mehrere RNA-Transkripte. Im konkreten Fall bedeutet dies für das glyA-*Template*, dass aus einem Molekül DNA-*Template* 181 *in vitro* RNA-Transkripte entstehen. Dadurch kann ein Eindruck hinsichtlich der Effektivität von phagen-spezifischen Transkriptionssystemen gewonnen werden. 10 μg der RNA wurden nachfolgend revers transkribiert und durch Dotierung mit Cy5-dCTP fluoreszenzmarkiert. Eine Adaption des etablierten Microarray-Protokolls (A.1.4) für das direkte *Labeling* stellte eine maximale Vergleichbarkeit mit einem realen Microarray-Experiment sicher. Für die Verwendung eines RNA-Standard ist es zudem wünschenswert, dass er simultan und unter gleichen Bedingungen prozessiert wird wie die zu analysierende RNA-Probe. Auch die dem RT-Schritt nachfolgenden Prozessschritte wurden völlig analog zu denen eines Microarray-Experiment durchgeführt. Die generierten Konzentrationen an cDNA und Cy5-Farbstoff sind der Tabelle 11 zu entnehmen.

Die cDNA-Konzentrationen schwanken von 58 $\text{ng}/\mu\text{l}$ bis zu 82 $\text{ng}/\mu\text{l}$, wohingegen die Farbstoffkonzentrationen einander sehr ähnlich sind (von 3,44 bis 3,71 $\text{pmol}/\mu\text{l}$). Aus den Absorptionskoeffizienten bei 260 und 650 nm wurden die Inkorporationsfrequenzen der vier verschiedenen cDNA-Spezies berechnet. Für die Spezies glyA, yidC und rpoA lagen die Werte zwischen 16,37 und 16,97 sehr eng beieinander. Lediglich die Inkorporationsfrequenz des lacA Transkriptes lag mit 12,39 etwas von diesen Werten entfernt. Für die Cy5-cDNA-Synthese von Proben, welche sich aus der total-RNA einer Zelle ableiten, wurden sehr ähnliche Inkorporationsfrequenzen gefunden. Im Rahmen der Evaluierung von Variabilitäten

von Chipexperimenten wurde in einem Parallelexperiment mit vier revers transkribierten RNA-Ansätzen der gleichen biologischen Probe ein Mittelwert von 16,86 gefunden (Abschnitt 4.1).

Um zu evaluieren, welche Auswirkungen die initial eingesetzte RNA-Menge auf die Effizienz der RT-Reaktion hat, und ob die Ergebnisse vergleichbar mit denen einer realen RNA-Probe sind, wurden vier verschiedene Mengen RNA in cDNA mittels des etablierten Protokolls (A.1.4) konvertiert. Dabei wurde eine mittels *in vitro* Transkription generierte RNA verwendet, welche für ein 663 bp langes Fragment aus dem Gen der Signalpeptidase 2 spezifisch ist. 10, 7, 3,5 und 1 μg RNA wurden in cDNA umgeschrieben und diese mittels Cy5-dCTP-Inkorporation fluoreszenzmarkiert. Der beschriebene Versuch wurde in einem Duplikat ausgeführt. Anschließend wurde die cDNA- und die Farbstoffkonzentration bestimmt, um so auf die Inkorporationsfrequenz und den prozentualen Anteil an umgeschriebener RNA zu schließen.

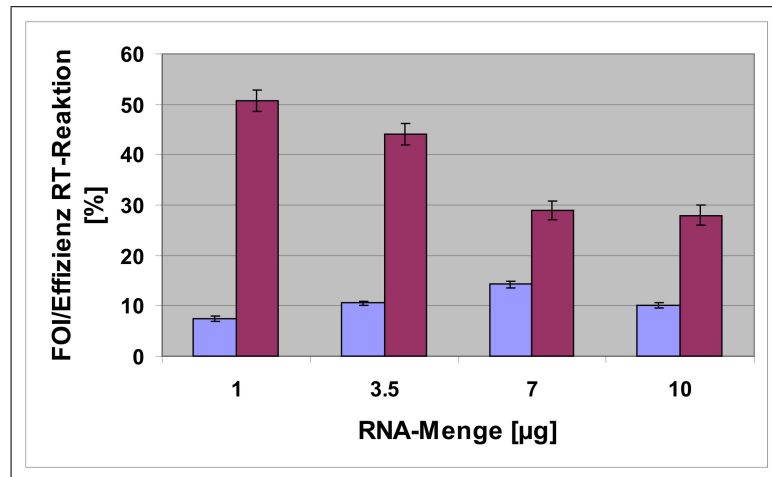


Abb. 88: Inkorporationsfrequenzen von Cy5-dCTP (blaue Säulen) und Reaktionseffizienzen der cDNA-Synthese (violette Säulen) verschiedener initialer RNA-Mengen.

Aus den Ergebnissen (Abbildung 88) ist ersichtlich, dass die Effizienz der RT-Reaktion bei steigender RNA-Menge abnimmt. Diese erreicht ein Effizienzmaximum von 50 Prozent bei einer initial vorliegenden RNA-Menge von 1 μg und nimmt auf 28 Prozent bei der RT-Reaktion von 10 μg *in vitro* RNA ab. Für die Inkorporationsfrequenz der fluoreszenzmarkierten Oligonukleotide in die cDNA ergibt sich aber ein gegenteiliger Verlauf. Diese weist für das verwendete Protokoll ein Optimum von 14,2% bei 7 μg revers transkribierter RNA auf. Eine möglichst hohe spezifische Fluoreszenz der cDNA-Probe ist für die Generierung einer hohen Sensitivität der Analyse der entscheidende Faktor. Nimmt man zudem die Inkorporationsfrequenzen realer RNA-Proben als Vergleichsmaßstab, so werden zumeist Werte gefunden, die sich in einem sehr ähnlichen Bereich befinden. So beträgt zum Beispiel der Mittelwert der Inkorporationsfrequenz für Cy5-dCTP in einem Parallelexperiment zur Evaluierung der Reproduzierbarkeit von Microarray-Experimenten 16,86 (Abschnitt 4.1). In Koizidenz wurde ein sehr ähnlicher Wert auch bei der RT-Reaktion von *glyA*, *rpoA*, *lacA* und *yidC* gefunden. Zusammenfassend wird daher eine initiale RNA-Menge von ca. 7 μg als Optimum für die Generierung von fluoreszenzmarkierter cDNA aus *in vitro* Transkriptionsansätzen betrachtet.

Die Anzahl der mRNA-Transkripte verschiedener Gene, welche in einer Zelle vorliegen, unterscheidet sich stark voneinander. In *E. coli* liegen häufige Transkripte mit ca. 10^5 Kopien pro Zelle neben sehr seltenen Transkripten vor, die nur in wenigen Kopien pro Zelle existieren (Southern *et al.* (1992)). Methoden zur Genexpressionsanalytik müssen diesen enormen dynamischen Bereich von 4-5 Zehnerpotenzen zu erfassen vermögen. Zur Evaluierung des dynamischen Bereichs der hier präsentierten Quantifizierung wurde eine Verdünnungsreihe einer fluoreszenzmarkierten cDNA-Spezies erstellt, welche spezifisch für einen 662 bp langen Bereich des Genes der Signalpeptidase 2 ist. Die cDNA wurde in fünf verschiedenen Verdünnungen aus dem cDNA-Ansatz auf die Chipoberfläche appliziert (1:10; 1:100; 1:1000; 1:10000 und 1:100000) und über Nacht hybridisiert. Die Hybridisierungsparameter entsprachen den Standardvorschriften für Microarray-Experimente (42°C , 650 rpm). Nach dem Waschvorgang wurden die Chips mit verschiedenen PMT- und Lasereinstellungen gescannt. In der nachfolgenden Primäranalyse wurden die SNR-Werte der verschiedenen Verdünnungen bei gleichen Scanparametern ermittelt, um die Detektionsgrenze der Analyse zu erfassen. Es wurde jeweils der Mittelwert aus allen 19 Replikaten des Chips gebildet und die Standardabweichung errechnet. Die Graphik zeigt das Ergebnis der Analyse. Die SNR-Werte befinden sich auch bei einer Verdünnung von 1:100000 bezogen auf den initial vorliegenden cDNA-Ansatz mit 12,28 noch deutlich über der Detektionsgrenze, welche als 3 definiert ist. Erwartungsgemäß steigt das SNR mit steigender cDNA-Konzentration an und erreicht sein Maximum bei der größten cDNA-Konzentration mit 381,52. Abbildung 89 fasst die Ergebnisse zusammen.

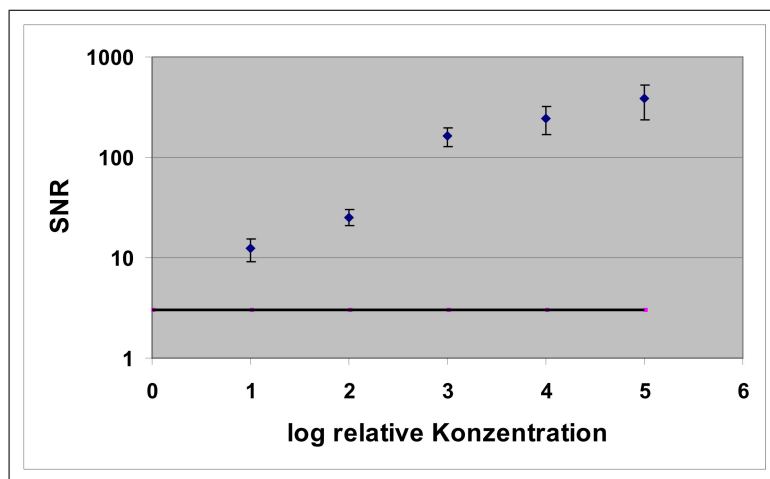


Abb. 89: Das resultierende SNR für verschiedene Verdünnungen einer fluoreszenzmarkierten Signalpeptidase 2-cDNA. Die schwarze Linie stellt die Nachweisgrenze entsprechend eines SNR von 3 dar.

Generell wird eine relativ große Standardabweichung für das SNR zwischen den verschiedenen Genreplikaten beobachtet. Eine genauere Betrachtung der Daten der Primäranalyse offenbart, dass diese von relativ großen Schwankungen zwischen den Signalintensitäten der einzelnen Replikate verursacht wird. Hingegen sind die Signalintensitäten und die Standardabweichung des Hintergrundes sehr gering. Dies gibt Hinweise darauf, dass eine möglichst hohe Anzahl an Genreplikaten für eine statistisch signifikante Analyse von Vorteil ist. Die Gründe für die Variabilitäten zwischen den Genreplikaten können in einer ungleichmässigen Verteilung des *Targets*, aufgrund der diffusionskontrollierten und damit langsamen Hybri-

disierung liegen. Eine bessere Reproduzierbarkeit könnte auch durch die Verwendung einer Hybridisierungsstation erzielt werden, wodurch eine konvektive Verteilung des *Targets* sicher gestellt wird. Initial ist die cDNA aus 5 μg RNA hervorgegangen. Dies entspricht für die niedrigste nachgewiesene Konzentration des cDNA-Ansatzes einer Molekülanzahl von $1,378 \cdot 10^8$, die mindestens als Detektionsgrenze angegeben werden kann. Die RNA, welche für ein Microarray-Experiment benötigt wird, stammt aus ca. 10^8 Zellen (Schena *et al.* (1995)). Selbst ein sehr seltenes Transkript, welches nur in einer Kopie pro Zelle vorkommt, sollte somit detektierbar sein und ließe sich anhand des Standards quantifizieren. Der dynamische Bereich der Methode umfasst eine Spanne von mindestens vier Magnituden und ist damit sehr gut geeignet, um den gesamten Konzentrationbereichs der mRNA einer realen Zelle zu detektieren.

Zur Evaluierung des linearen Bereichs der Analysenmethode wurden die Mediane der hintergrundbereinigten Signalintensitäten aus allen Genreplikaten des Signalpeptidase 2-Transkriptes für jeden Chip berechnet. Dazu fungierten Primäranalysedaten, welche aus Scans generiert wurden, die für alle Chips die gleichen Parametereinstellungen aufwiesen. Lediglich die Daten für die stärkste Verdünnung konnten so nicht berechnet werden, weil die resultierende SNR unter der Nachweissgrenze lag. Das Resultat ist in Abbildung 90 gezeigt.

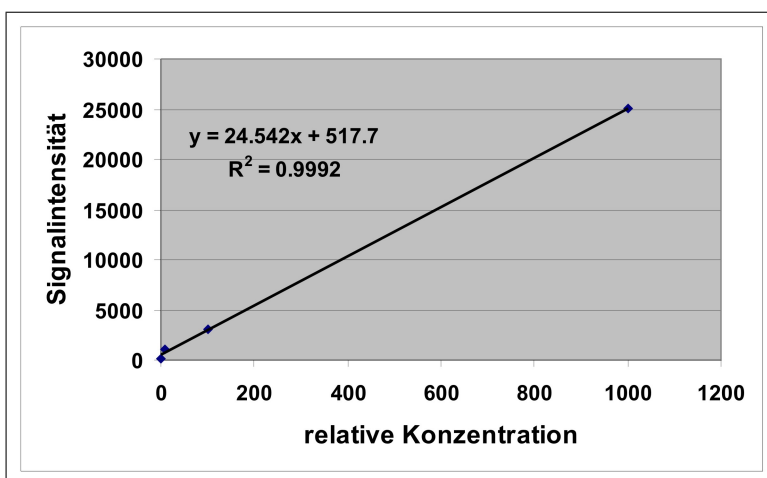


Abb. 90: Auftragung der hintergrundbereinigten Signalwerte gegen die relative cDNA Konzentration einer Verdünnungsreihe der Signapeptidase 2-Spezies. Durch die Datenpunkte wurde eine lineare Regressionsgrade gelegt, die Gradengleichung und der dazugehörige Korrelationskoeffizient sind in der Abbildung angegeben.

Die resultierenden Signalintensitäten für die verschiedenen Verdünnungen wurden gegen die relative Konzentration der cDNA aufgetragen. Die hohe lineare Korrelation der Daten ist aus dem Diagramm (Abbildung 90) ersichtlich und wird durch den fast idealen Korrelationskoeffizienten von 0,9992 bestätigt. Der lineare Bereich der Analysenmethode beträgt also mindestens drei Zehnerpotenzen und ist daher bestens geeignet für Applikationen im Bereich der Genexpressionsanalyse.

Durch die Evaluierung des linearen Bereiches der Analysenmethode wurden die Voraussetzungen für ein *Multiplexing*-Experiment geschaffen. Anhand eines *Multiplexing*-Versuches wurde der Einfluß verschiedener Transkriptlängen und der Anzahl der Cytosin-Positionen (mögliche Orte für den Einbau der Cy-markierten dCTPs) in der cDNA evaluiert. Mittels dieser Daten wurde die lineare Korrelation zwischen Signalintensität und Kopienzahl der

initialen RNA für die einzelnen Spezies ermittelt. Aus dem Mittelwert soll eine Kalibrationsgrade erstellt werden, anhand derer eine Spezies unbekannter Konzentration quantifiziert werden kann.

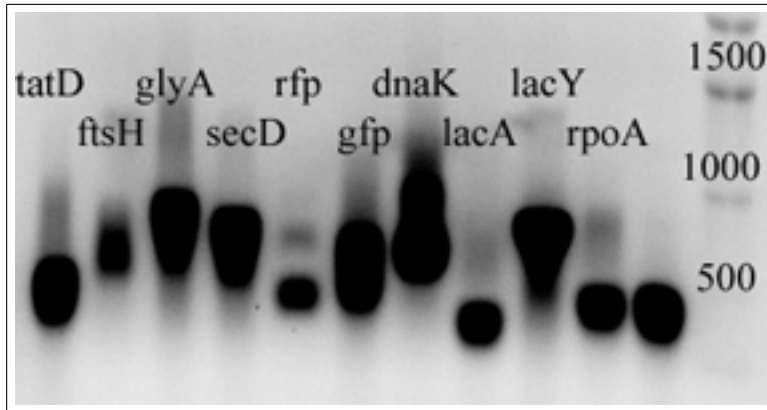


Abb. 91: Visualisierung von zehn verschiedenen *in vitro* RNA-Transkripten auf einem 1,5 prozentigem Agarosegel unter denaturierenden Bedingungen. Die Spezifizierung der RNA-Transkripte kann der Abbildung entnommen werden.

Dazu wurden jeweils 7 μg von zehn verschiedenen RNA-Spezies (Abbildung 91) revers transkribiert, Cy5-fluoreszenzmarkiert und der generierte cDNA-Ansatz in vier verschiedenen Verdünnungen (1:10) auf die spezifischen *low-density* Chips appliziert. Das Experiment erfolgte in zwei Wiederholungen, so dass drei identische Versuchsreihen zur Datenanalyse zur Verfügung standen. Die Absorptionsspektren der fluoreszenzmarkierten cDNAs sind in Abbildung 92 gezeigt.

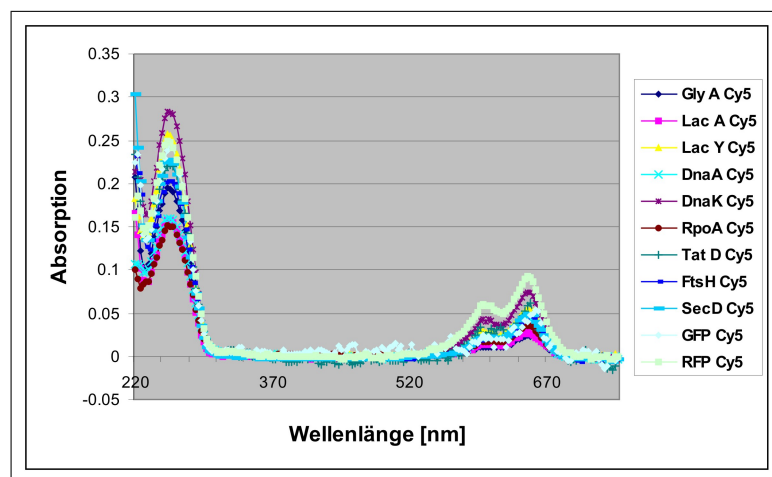


Abb. 92: Absorptionsspektren von zehn verschiedenen, Cy5-markierten cDNA-Spezies. Die cDNA- und Cy5-Konzentrationen sowie die sich daraus berechnenden Inkorporationsfrequenzen sind in der folgenden Tabelle 12 angegeben.

Die fluoreszenzmarkierte cDNA wurde jeweils spektrophotometrisch analysiert und die Inkorporationsfrequenz berechnet. Die Ergebnisse sind in der Tabelle 12 gezeigt. Die Inkorporationsfrequenzen der meisten Transkripte liegen im Bereich zwischen 7 und 9. Rfp

Tabelle 12: cDNA- und Farbstoffkonzentrationen sowie die sich daraus ergebenden Inkorporationsfrequenzen für einen *Multiplexing*-Ansatz mit zehn verschiedenen RNA-Spezies.

	A260	A650	FOI
glyA Cy5	0,193	0,024	4,55
lacA Cy5	0,157	0,029	6,48
lacY Cy5	0,255	0,055	7,57
dnaA Cy5	0,159	0,044	9,71
dnaK Cy5	0,282	0,074	9,71
rpoA Cy5	0,150	0,036	8,42
tatD Cy5	0,221	0,059	9,37
ftsH Cy5	0,203	0,046	7,95
secD Cy5	0,227	0,049	7,58
gfp Cy5	0,244	0,040	5,75
rfp Cy5	0,247	0,093	13,22

repräsentiert mit 13,22 einen Ausreißer zu höheren Werten, wohingegen das Gegenteilige für glyA mit einer Inkorporationsfrequenz von 4,55 gilt.

Die Daten der Primäranalyse wurden einer linearen Regression unterzogen, normalisiert und der Mittelwert der hintergrundbereinigten Signalintensität aus allen Replikaten und den drei Versuchsreihen gebildet. Die erhaltenen, normalisierten Signalwerte der einzelnen Gene wurden anschließend durch die Anzahl der Cytosin-Positionen in der cDNA-Sequenz dividiert um eine maximale Vergleichbarkeit zwischen den verschiedenen Spezies zu erhalten. Das detektierte Fluoreszenzsignal resultiert aus dem Einbau fluoreszenzmarkierter dCTPs in die cDNA-Kette. Die Gesamtanzahl an Cytosin-Positionen in dem Molekül in Kombination mit der Inkorporationsfrequenz bestimmt daher die Anzahl inkorporierter Reportergruppen in der cDNA. Eine große Anzahl an Cytosin-Positionen innerhalb der cDNA generiert daher ein stärkeres Fluoreszenzsignal als dies für eine cDNA gleicher Länge mit einem geringeren C-Gehalt der Fall ist.

In der Abbildung 93 ist die Auftragung dieser Signalwerte gegen die Anzahl der RNA-Moleküle aufgetragen, die in diesem Volumen repräsentiert sind. In dem betrachteten Konzentrationsbereich weisen die Signalwerte der einzelnen Spezies eine sehr hohe lineare Korrelation zur Kopienanzahl der RNA auf; dies belegt der hohe Korrelationskoeffizient der Regressionsgraden von 0,985. Auffällig ist, dass für die kleinsten Transkriptlängen (RpoA und LacA) die größten Abweichungen von der Regressionsgraden gefunden werden. Diese beobachtete Längenabhängigkeit der Signalwerte kann durch den Prozessschritt der reversen Transkription erklärt werden. Die Effizienz der reversen Transkriptase ist u.a. abhängig von der Länge der RNA-Matrize. Längere Matrizen führen wahrscheinlicher zu einem vorzeitigen Kettenabbruch oder Fehlern in der Sequenz, was die Effizienz der Reaktion senkt. Gleiche Anzahlen an initialen RNA-Molekülen generieren daher für längere Transkripte eine geringere Anzahl an korrekt revers transkribierter RNA-Molekülen, entsprechend geringer fällt der Signalwert aus. Begünstigt wird dies durch die Tatsache, dass es sich bei Reversen Transkriptasen um Enzyme handelt, welche eine hohe Fehlerquote aufweisen, was sich z.B.

auch in hohen Mutagenitätsraten von Retroviren äußert.

Aufgrund eines ähnlichen längenabhängigen Effektes werden für die quantitative *real-time* PCR-Analysen Primer gewählt, welche sehr kurze Amplifikate generieren (Pfaffl *et al.* (2001), Wilhelm *et al.* (2003)). Für die Darstellung fluoreszenzmarkierter cDNA liegen die Verhältnisse allerdings anders. Es lässt sich die Transkriptlänge nicht beliebig senken, da die Inkorporationsfrequenz zu gering ist, um eine genügende Anzahl Reporterfluorophore in die cDNA sehr kurzer Transkripte zu inkorporieren. Für eine sehr genaue quantitative Analyse der Genexpression sollte daher die Transkriptlänge mit in Betracht gezogen werden und die Quantifizierung über mehrere Standards erfolgen, die eine ähnliche Länge wie der Analyt aufweisen. Durch die Verwendung von Standards mit einem breitem Spektrum verschiedener Längen können allerdings die meisten mRNA-Transkripte einer realen Probe hinreichend genau erfasst werden.

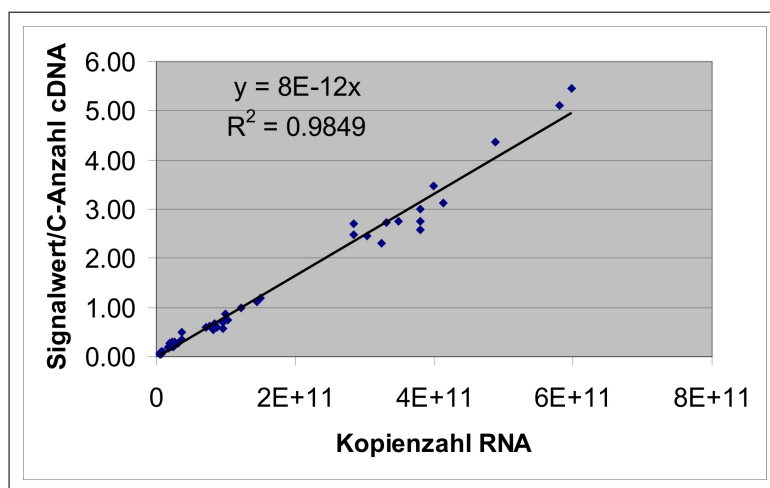


Abb. 93: Regressionsgrade zur Bestimmung der RNA-Kopien aus dem normalisierten Signalwert.

Anhand der generierten Kalibrationsgraden ist es möglich, die Kopienzahl eines Transkripts aus dem normierten Signalwert der Microarray-Analyse zu berechnen. Es konnte gezeigt werden, dass die Inkorporationsfrequenzen denen von RNA-Proben natürlichen Ursprungs entsprachen und der erforderliche dynamische Bereich für Genexpressionsanalysen vollständig durch die hier etablierte Quantifizierungsmethode abgedeckt wird. Die Detektionsgrenze von weniger als 10^8 Spezies RNA belegt die hohe Sensitivität der Analysenmethode. Verglichen mit anderen Methoden wie z.B. dem Ribonuklease-Protektionsassay oder der quantitativen *real-time* PCR sind diese allerdings noch sensitiver. Es lässt sich aber erwarten, dass die Sensitivität zukünftig durch neuartige Reporterfluorophore (z.B. *Quantum Dots*), effektivere Reverse Transkriptasen und verbesserte Chipoberflächen noch wesentlich zu steigern ist. Besonders *Quantum Dots* zeigen aufgrund der signifikant reduzierten Hintergrundfluoreszenz und hohen Extinktionskoeffizienten großes Potenzial für diese Applikation. Zudem wurde im Rahmen dieser Studie keine gesonderte Optimierung der Chipoberfläche vorgenommen. Zusätzlich lässt sich das etablierte Quantifizierungsprinzip auch auf Signalamplifikationssysteme übertragen (Dendrimer-basiert, TSA) und so sollten sich deutlich höhere Sensitivitäten realisieren lassen. Die Quantifizierungsstandards können zudem dazu verwendet werden, Schwankungen in den Prozessschritten eines Microarrayexperimentes zu

detektieren und bilden eine weitere Möglichkeit für Normalisierungs-Applikationen. Besonders für *low-density* Microarrays besitzt dieses Relevanz, da die konventionellen Normalisierungen auf Annahmen basieren die in einem *low-density* Experiment nicht gegeben sind. Als Beispiel sei hier auf die Grundannahme der Nicht-Regulation der meisten auf dem Chip repräsentierten Gene hingewiesen.

5 Zusammenfassung und Ausblick

Der universellen Verwendung von *E. coli* zur industriellen Produktion rekombinanter Proteine steht die Tatsache entgegen, dass *E. coli* von Natur aus Proteine nur sehr limitiert sekretiert. Während eine optimierte Sekretion Ausbeuten erhöhen und signifikante Vorteile im *downstream processing* bieten könnte, ergaben sich bisherige Verbesserungen konzeptionell unstrukturiert durch zufällige Evaluationen. Im Rahmen dieser Arbeit soll zum ersten Mal systematisch auf Transkriptomebene analysiert werden, welche regulatorischen Vorgänge während der Sekretion rekombinanter Proteine ablaufen. Bei der extrazellulären Proteinsekretion in gram-negativen Mikroorganismen handelt es sich um einen zweistufigen Prozess. Die Nutzung des Sec- oder des Tat-Pathways dirigiert Proteine ins Periplasma. Das periplasmatische Intermediat wird nachfolgend z.B. durch Coexpression sogenannter *bacteriocin release proteins* ins Medium ausgeschleust.

Für die durchgeführten Analysen wurden daher modifizierte *E. coli*-Stämme eingesetzt, mit denen sich die einzelnen Phasen der Sekretion separat analysieren lassen. Zusätzlich erfolgte eine Diskriminierung zwischen der Sec- und Tat-spezifischen Sekretion. Die Abbildung 94 gibt einen Überblick über die verwendeten Sekretionsstämme, welche teils selbst generiert und teils von der AG Fermentationstechnik der Universität Bielefeld zur Verfügung gestellt wurden.

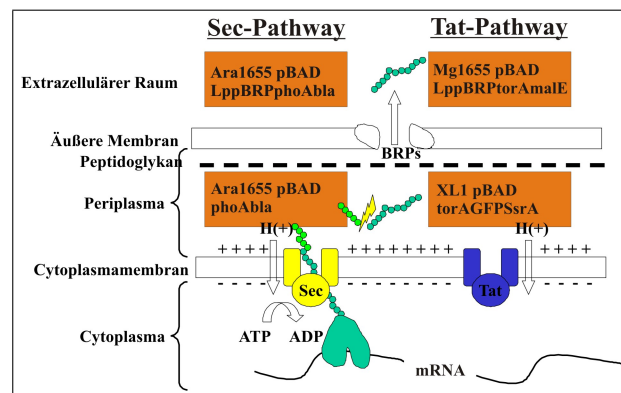


Abb. 94: Proteinsekretions-Pathways in *E. coli*. Proteine mit einer N-terminalen Signalsequenz (hier gezeigt für ein Substrat des Sec-Pathways) werden mittels des Sec- oder Tat-Pathways durch die Cytoplasmamembran ins Periplasma translokiert. Dort wird die Signalsequenz abgespalten und das reife Protein entsteht. Mittels der transkriptionellen Analyse der Stämme XL1 pBADtorAGFPSsrA (Tat-Pathway) und Ara1655 pBADphoAbla (Sec-Pathway) wird dieser Schritt der zweistufigen Sekretion analysiert. Durch Coexpression von BRPs gelangen periplasmatisch lokalisierte Proteine in den extrazellulären Raum. Dieser Vorgang wird mittels der transkriptionellen Analysen der Stämme Ara1655 pBADLppBRPphoAbla (Initial über Sec-Pathway) und MG1655 pBADLppBRPtorAmalE (Initial über Tat-Pathway) untersucht.

Die Ergebnisse der so durchgeführten Genexpressionsanalysen wurden mit kommerziellen *whole-genome* Microarrays (alle *E. coli*-Gene enthalten) als auch mit selbst entwickelten und hergestellten sekretionsspezifischen *low-density* Microarrays generiert. Diese Ergebnisse zeigen konsistent, dass die differentielle Genexpression der sekretionsspezifischen Gene innerhalb weniger Minuten stattfindet. Die BRP-Expression führt zu einer signifikanten positiven Regulation vieler Membranstressgene und Operons (*pspABCDE*, *rpoE*, *spy*, *lolAB*) sowie verschiedener Stressmarker (*rpoH*, *soxS*, *lexA*). Eine Initiation der zellulären SOS-Antwort konnte detektiert und die dadurch resultierende Regulationskaskade experimentell beobachtet werden. Des Weiteren wurde die differentielle Expression von Komponenten der Lipopolysaccharid- und Peptidoglykan-Biosynthese beobachtet. Bedingt durch die Sec- und Tat-spezifische Sekretion trat die Regulation verschiedener Chaperon-Gene (z.B. *fkpA*, *dnaK*, *groES*, *groEL*, *torD*), Komponenten der Translokase (z.B. *tatAD*, *secYEG*) und von Genen, die in der Energiebereitstellung des Sekretionsprozesses fungieren (z.B. *nuo*-Operon, *atp*-Operon), auf. Des Weiteren wurden differentiell exprimierte Gene identifiziert, die mechanistisch in den Sekretionsprozess involviert sind (z.B. *imp*, *mltCDE*, *B1571*, *pIdA*, *B0421*) und zur Aufrechterhaltung der Membranintegritäten fungieren. Die mittels Microarray-Analyse evaluierte Genexpressionsdaten wurden mittels quantitativer *real-time* RT-PCR verifiziert und können daher als signifikant angesehen werden. Um Aussagen über das absolute Maß der Genexpression treffen zu können wurde zusätzlich ein System zur absoluten Quantifizierung der mRNA-Kopienzahl etabliert, dessen dynamischer Bereich mindestens 4 Größenordnungen umfasst. Darüber hinaus konnte durch die Verwendung sogenannter *Quantum Dots* die Microarraytechnologie selbst weiterentwickelt werden.

Mit den Ergebnissen dieser Arbeit sollte es möglich sein zukünftig die Proteinsekretion in *E. coli* zu optimieren. Zur Etablierung von *E. coli*-Stämmen mit einer erhöhten Sekretionseffizienz sollte sowohl eine Optimierung der BRP-Expression als auch der Sec- bzw. Tat-spezifischen Sekretion erfolgen. Um die exprimierende Zelle besser vor der degradativen Wirkung der BRP-Aktivität zu schützen ergeben sich nachfolgend angeführte Optimierungsansätze, die aufgrund der positiven Regulation dieser Komponenten getroffen wurden. Das Gencluster *lolAB* codiert für die Komponenten, welche für die korrekte Translokation der Lipoproteine der äußeren Membran verantwortlich sind und stellt somit ein höchst interessantes Zieloperon dar. In Koinzidenz dazu könnte eine Coexpression des Gens *rpoE* erfolgen, welches für den Membranstress-Sigmafaktor der RNA-Polymerase codiert und die vermehrte Transkription von Genen ermöglicht, die in der Aufrechterhaltung der Membranintegrität fungieren. In diesem Kontext ist auch der Transkriptionsaktivator der allgemeinen Stressantwort *SoxS* erwähnenswert. De Lisa *et al.* (2004) konnten durch Coexpression von *pspA* eine signifikante Sekretionserhöhung detektieren. Insofern stellt das Phagenschockoperon ebenfalls ein konkretes Coexpressionsziel dar. Weiterhin könnte die membranschädigende Wirkung der BRP-Expression durch Coexpression von LPS- und Peptidoglykankomponenten gemildert werden. Als Beispiel hierfür sei die Hauptkomponente *Lpp* der äußeren Membran angeführt. Zur Erhöhung der Sec- und Tat-spezifischen Sekretion sind, basierend auf den experimentellen Daten, vornehmlich (1) Komponenten der Translokase, (2) Chaperone, (3) Stressgene zur Wahrung der Membranintegrität und (4) Gene, die für die Triebkraft der Translokation essentiell sind, zu nennen. Die Coexpression von Chaperonen hat sich in vielen Fällen (Perez-Rodriguez *et al.* (2007), Li *et al.* (2003)) als erfolgreich erwiesen. Durch die schnelle Saturierung der Translokase, besonders im Falle des Tat-Pathways (De Lisa *et al.* (2003)), wird die Sekretionskapazität limitiert. Insofern konnten verschiedene Forschergruppen (Barett *et al.* (2003), Pugsley *et al.* (1993)) zeigen, dass die Coexpression

von Translokase-Komponenten zu einer Steigerung der Sekretion führt. Allerdings lässt sich eine Coexpression von Translokase-Komponenten nicht *ad absurdum* durchführen. Bedingt durch die vermehrte Membranpassage der Sekretionssubstrate tritt ein transmembraner Protonenfluß ins Cytoplasma auf, welcher die Triebfeder der Sekretion, den Protonengradienten, reduziert. In diesem Kontext hat sich die Coexpression von Genen als erfolgreich erwiesen, deren Produkte die Membranintegrität erhöhen (De Lisa *et al.* (2004)). In besonderem Maße hat sich die Coexpression von *pspA* als äußerst erfolgreich erwiesen. Brüser *et al.* (2007) postulierten eine mechanistische Erklärung für die *pspA*-vermittelten Schutzwirkung. Durch die Ausbildung einer netzartigen Struktur, die sich über die beim Sekretionsvorgang gebildet Öffnungen in der Cytoplasmamembran legt, wird eine Zellyse und ein vermehrter Transmembran-Protonenfluss vermieden. Genprodukte, welche die Membranintegrität erhöhen, reduzieren somit nicht nur die BRP-induzierten Wirkungen, sondern fungieren auch zur Aufrechterhaltung der Triebkraft des Translokationsprozesses. Somit stellt die Coexpression von Genprodukten wie *PspA* in zweifacher Hinsicht einen erfolgsversprechende Optimierungsansatz dar. Das *nuo*-Operon (Protonenpumpe gemäss Stolpe *et al.* (2004)) und das *atp*-Operon sind ebenfalls Kandidatengene um die Triebkraft der Sekretion zu erhöhen.

Zusammenfassend sollte eine Coexpression dieser angeführten Gene bzw. Kombinationen aus ihnen zu einer signifikanten Erhöhung der Menge an sekretiertem rekombinanten Protein führen. Neben Coexpression und *knock-out*-Strategien sei auch auf die Möglichkeit einer genetischen Modifikation der Translokasekomponenten-erfolgreich bei den *prl*-Mutatenstämmen (Manting *et al.* (2000))-sowie der verwendeten BRPs (van der Waal *et al.* (1995)) hingewiesen. Weiterhin üben auch die Kultivierungsparameter und Medienzusätze einen signifikanten Einfluß auf den Sekretionsprozess bzw. auf die BRP-Expression aus. Ein Zusatz zweiwertiger Ionen zum Kulturmedium erlaubt z.B. eine erhöhte Expressionsrate der BRPs, ohne die negativen Folgen der Quasi-Lyse zu verursachen. Zur Generierung eines optimierten Sekretionsstammes sollten mithin Kombinationen aus genetischer Modifikation, Fermentationsoptimierung und Coexpressionsstrategien angewandt werden. Wie anhand der erfolgsversprechenden Ergebnisse gezeigt, stellt die Microarray-Technologie dabei ein *Screening-Tool* von unschätzbarem Wert dar.

6 Literatur

- [1] Abrahmsen L., Moks T., Nilsson B., Uhlen M. Secretion of heterologous gene products to the culture medium of *Escherichia coli*. *Nucleic Acids Res* 1986;14:7487-7500.
- [2] Alami M., Luke I., Deitermann S., Eisner G., Koch HG., Brunner J., *et al.* Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* 2003;12:937-946.
- [3] Alba BM., Leeds JA., Onufryk C., Lu CZ., Gross CA. DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigmaE-dependent extracytoplasmic stress response. *Genes Dev* 2002;16(16):2156-2168.
- [4] Bailey MJ., Hughes C., Koronakis U. Increased distal gene transcription by the elongation factor *rfaH*, a specialized homologue NusG. *Mol Microbiol* 1996;22(4):729-737.

-
- [5] Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol* 1999;10:411-421.
- [6] Beck K., Wu LF., Brunner J., Muller M. Discrimination between a SRP- and SecA/SecB-dependent substrate involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J* 2000;19:134-143.
- [7] Berks BC., Sargent F., Palmer T., The Tat protein export pathway. *Mol Microbiol* 2000;35:260-274.
- [8] Bevis BJ., Glick BS. Rapidly maturing variants of the discosoma red fluorescent protein (DsRed). *Nature Biotechnol* 2002;20:83-87.
- [9] Beshay U., Miksch G., Friehs K., Flaschel E. Integrated bioprocess for the production and purification of recombinant proteins by affinity chromatography in *Escherichia coli*. *Bioprocess Biosyst Eng.* 2008 May 15.
- [10] Binet R., Letoffe S., Ghigo JM., Delepelaire P., Wandersman C. Protein secretion by gram-negative bacterial ABC exporters-a review. *Gene* 1997;192:7-11.
- [11] Biju V., Itoh T., Anas A., Sujith A., Ishikawa M. Semiconductor quantum dots and metal nanoparticles: synthesis, optical properties, and biological applications. *Anal Bioanal Chem* 2008;391:2469-2495.
- [12] Blaudeck N., Sprenger GA., Freudl R., Wiegert T. Specificity of signal peptide recognition in tat-dependent bacterial protein translocation. *J Bacteriol* 2001;183:604-610.
- [13] Blight MA., Holland IB. Heterologous protein secretion and the versatile *Escherichia coli* haemolysin translocator. *Trends Biotechnol* 1994;12:450-455.
- [14] Bongaerts RJM., Hautefort I. Sidebotham JM., Hinton YCD. Green fluorescent protein as a marker for conditional gene expression in bacterial cells. *Methods Enzymol* 2002;358:43-66.
- [15] Breyton C., Haase W., Rapoport TA., Kuhlbrandt W., Collinson I. Three-dimensional structure of the bacterial protein-translocation complex SecYEG. *Nature* 2002;418:662-665.
- [16] Brill JA., Quinlon-Walshe C., Gottesman S. Fine structure mapping and identification of two regulations of capsule synthesis in *Escherichia coli* K12. *J Bacteriol* 1988;170(6):2599-2611.
- [17] Brisette JL., Weiner L., Ripmaster TL., Model P. Characterisation and sequence of the *Escherichia coli* stress-induced psp-Operon. *J Mol Biol* 1991;220(1):35-48.
- [18] Brown T. Analysis of RNA by northern and slot-blot hybridisation. *Curr Protoc Immunol* 2001;10:10.12.

-
- [19] Brüser T. The twin arginine translocation system and its capability for protein secretion in biotechnological protein production. *Appl Microbiol Biotechnol* 2007;76:35-45.
- [20] Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000;25(2):169-193.
- [21] Campell RE., Tour O., Palmer AE., Steinbach PA., Baird GS., Zacharias DA., Tsien RY. A monomeric red fluorescent protein. *Biochemistry* 2002;12:7877-7882.
- [22] Carter B., Wu G., Woodward MJ., Anjum MF. A process for analysis of microarray comparative genomics hybridisation studies for bacterial genomes. *BMC Genomics* 2008;9:53.
- [23] Choi JH., Lee SY. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* 2004;64:625-635.
- [24] Choi JH., Jeong KJ., Kim SC., Lee SY. Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant *Escherichia coli* using bacillus sp endoxylanase signal sequence. *Appl Microbiol Biotechnol* 2000;53:640-645.
- [25] Cramer A., Whitehorn EA., Tate E., Stemmer WPC. Improved green fluorescent protein by molecular evolution using dna shuffling. *Nature Biotechnol* 1995;14:315-319.
- [26] Cristobal S., de Gier JW., Nielsen H., von Heijne G. Competition between Sec- and Tat-dependent protein translocation in *Escherichia coli*. *EMBO J.* 1999;18:2982-2990.
- [27] Crut A., Geron-Landre B., Bonnet I., Bonneau S., Desbiolles P., Escude C. Detection of single DNA molecules by multicolor quantum-dot end-labeling. *Nucleic Acids Research* 2005;33:11.
- [28] Danese PN., Silhavy TJ. Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. *Annu Rev Genet* 1998;32:59-94.
- [29] Darling DC., Brickell PM. 1996: Nucleinsäure-Blotting, Spektrum Akademischer Verlag, Heidelberg.
- [30] Dekker C., de Kruijff B., Gross P. Crystal structure of SecB from *Escherichia coli*. *J Struct Biol.* 2003;144:313-319.
- [31] Dekker N., Tommassen J., Verheij HM. Bacteriocin release protein triggers dimerization of outer membrane phospholipase A *in vivo*. *J Bacteriol* 1999;181(10):3281-3283.
- [32] De Lisa MP., Samuelson P., Palmer T., Georgiou G. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* 2002;322:1135-1146.
- [33] De Lisa MP., Tullman D., Georgiou G. Folding quality control in the export of proteins by the bacterial twin arginine translocation pathway. *Proc Natl Acad Sci USA* 2003;100:6115-6120.

-
- [34] De Lisa MP., Lee P., Palmer T., Georgiou G. Phage shock protein *pspA* of *Escherichia coli* relieves saturation of protein export via the Tat pathway. *J Bacteriol* 2004;186:366-373.
- [35] Derouiche R., Benedetti H., Lazzaroni JC., Lazdunski C., Lloubes R. Protein complex within *Escherichia coli* inner membrane TolA N-terminal domain interacts with TolQ and TolR proteins. *J Biol Chem* 1995;270(19):11078-11084.
- [36] Dijkstra AJ., Keck W. Peptidoglycan as a barrier to transenvelope transport. *J Bacteriol* 1996;178(19):5555-5562.
- [37] Driessen AJ. SecB, a molecular chaperone with two faces. *Trends Microbiol* 2001;9:193-206.
- [38] Driessen AJ., Fekkes P., van der Wolk JP. The Sec system. *Curr Opin Microbiol* 1998;1:216-222.
- [39] Driessen AJ., Manting EH., van der Does C. The structural basis of protein targeting and translocation in bacteria. *Nat Struct Biol* 2001;8:492-508.
- [40] Economou A. Following the leader: bacterial protein export through the Sec pathway. *Trends Microbiol* 1999;7:315-320.
- [41] Eickhoff B., Korn B., Schick M., Poustka A., Van der Bosch J. Normalisation of array hybridisation experiments in differential gene expression analysis. *Nucleic Acids Research* 1999;27:No.22.
- [42] Eser M., Ehrmann M., SecA-dependent quality control of intracellular protein localisation. *Proc Natl Acad Sci USA* 2003;100:13231-13234.
- [43] Fath MJ., Mahanty HK., Kolter R. Characterisation of a *purF* operon mutation which affects colicin V production. *J Bacteriol* 1989;171(6):3158-3161.
- [44] Fernandez LA., Berenguer J. Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiol Rev* 2000;24:21-44.
- [45] Fernandez LA., Sola I., Enjuanes L., de Lorenzo V. Specific secretion of active single chain Fv antibodies using into the supernatants of *Escherichia coli* cultures by use of the hemolysin system. *Appl Environ Microbiol* 2000;66:5024-5029.
- [46] Feilmeier BJ., Iseminger G., Schroeder D., Webber H., Phillips GJ. Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J Bacteriol* 2000;182:4068-4076.
- [47] Flynn JM., Levchenko I., *et al.* Overlapping recognition determinants within the *ssrA* degradation tag allow modulation of proteolysis. *PNAS* 2001;19:10584-10589.

-
- [48] Frederiksen CM., Aaboe M., Dyrskjøt L., Laurberg S., Wolf H., Ørntoft TF., Kruhøffer M. Technical evaluation of cDNA microarrays. *APMIS Suppl* 2003;109:96-101.
- [49] Fu Z., Hamid SB., Razak CN., Basri M., Salleh AB., Rahman RN. Secretory expression in *Escherichia coli*. and single step purification of a heat-stable alkaline protease. *Protein Expr Purif* 2003;28:63-78.
- [50] Gao H., Wang Y., Liu X., Yan T., Wu L., Alm E., Arkin A., Thompson DK., Zhou J. Global transcriptome analysis of the heat shock response of *Shewanella oneidensis*. *J Bacteriol* 2004;186(22):7796-7803.
- [51] Gentschev I., Hess J., Goebel W. Change in the cellular localisation of alkaline phosphatase by alteration of its carboxy-terminal sequence. *Mol Gen Genet* 1990;222:211-216.
- [52] Gentschev I., Hess J., Goebel W. The *E. coli* alpha-hemolysin secretion system and its use in vaccine development. *Trends Microbiol* 2002;10:39-45.
- [53] Glick BR., Pasternak JJ. (1995) *Molekulare Biotechnologie*, Spektrum Akademischer Verlag, Heidelberg.
- [54] Gotfredsen M., Gerdes K. The *Escherichia coli* relBE genes belong to a new toxin-antitoxin gene family. *Mol Microbiol* 1998;29(4):1065-1076.
- [55] Gottesman S., Trisler P., Torres-Labassa A. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12: characterisation of three regulatory genes. *J Bacteriol* 1985;162(3):1111-1119.
- [56] Goulding CW., Jeanne Perry L. Protein production in *Escherichia coli* for structural studies by X-ray crystallography. *J Struct Biol* 2003;142:133-143.
- [57] Grant WD., Sutherland IW., Wilkinson JF. Expolysaccharide colanic acid and its occurrence in the Enterobacteriaceae. *J Bacteriol* 1969;100(3):1187-1193.
- [58] Gumpert J., Hoischen C. Use of cell wall-less bacteria (L-forms) for efficient expression and secretion of heterologous gene products. *Curr Opin Biotechnol* 1998;9:506-509.
- [59] Guzman LM., Belin D., Carson MJ., Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD Promoter. *J Bacteriol* 1995;14:4121-4130.
- [60] Hakkila K., Maksimow M., Rosengren A., Karp M., Virta M. Monitoring promoter activity in a single bacterial cell by using green and red fluorescent proteins. *J Microbiol Methods* 2003;54:75-79.
- [61] Harcum SW., Haddadin FT. Global transcriptome response of recombinant *Escherichia coli* to heat-shock and dual heat-shock recombinant protein induction. *J Ind Microbiol Biotechnol* 2006;33(10):801-814.

-
- [62] Held KG., Postle K. ExbB and ExbD do not function independently in TonB-dependent energy transduction. *J Bacteriol* 2002;184(18):5170-5173.
- [63] Holland IB., Kenny B., Steipe P., Pluckthun A. Secretion of heterologous in *Escherichia coli*. *Methods Enzymol* 1990;182:132-143.
- [64] Holzapfel E., Eisner G., *et al.* The entire N-terminal half of *tatC* is involved in twin arginine precursor binding. *Biochemistry* 2007;46:2892-2898.
- [65] Hölte JV., Mivelman D., Sharon N., Schwarz U. Novel type of murein transglycosylase in *Escherichia coli*. *J Bacteriol* 1975;124(3):1067-1076.
- [66] Huggett J., Dheda K., Bustin S., Zumla A. Real-time RT-PCR normalisation: strategies and considerations. *Geneg Immun* 2005;6(4):279-284.
- [67] Jeong KJ., Lee SY. Secretory production of human granulocyte colony stimulating factor in *Escherichia coli*. *Protein Expr Purif* 2001;23:311-318.
- [68] Jones SE., Lloyd LJ., Tan, KK. and Buck, M. (2003) Secretion defects that activate the phage shock response of *Escherichia coli*. *J Bacteriol* 2003;185(22):6707-6711.
- [69] Yung JU., Gutierrez C., Martin F., Ardourel M., Villarejo M. Transcription of *osmB*, a gene coding an *Escherichia coli* lipoprotein, is regulated by dual signal osmotic stress and stationary phase. *J Biol Chem* 1990;265(18):10574-10581.
- [70] Karp PD., Keseler IM., Shearer A., Latendresse M., Krummenacker M., Paley SM., Paulsen I., Collado-Vides J., Gama-Castro S., Peralta-Gil M., Santos-Zavaleta A., Peñaloza-Spínola MI., Bonavides-Martinez C., Ingraham, J. Multidimensional annotation of the *Escherichia coli* K-12 genome. *Nucleic Acids Res* 2007;35(22):7577-7590.
- [71] Kebir MO., Kendall DA. *SecA*-specificity for different signal peptides. *Biochemistry* 2002;41:5573-5580.
- [72] Kleist S., Miksch G., Hitzmann B., Arndt M., Friehs K., Flaschl E. Optimization of the extracellular production of a bacterial phytase with *Escherichia coli* by using different fed-batch fermentation strategies. *Appl Microbiol Biotechnol* 2003;61:456-462.
- [73] Kraft AR., Templin MF., Hölte JV. Membrane bound lytic endotransglycosylases in *Escherichia coli*. *J Bacteriol* 1998;180(13):3441-3447.
- [74] Kuzimov A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol Mol Biol Rev* 1999;63(4):751-813.
- [75] Lewin B.(1995) *Gene*; Wiley-VCH, Weinheim.
- [76] Li SY., Chang BY., Lin SC. Coexpression of *TorD* enhances the transport of GFP

via the Tat pathway. *J Biotechnol* 2006;122:412-421.

[77] Li S.J., Cronan J.E. jr. The genes encoding the two carboxyltransferase subunits of *Escherichia coli* acetyl-CoA carboxylase. *J Biol Chem* 1992;267(24):16841-16847.

[78] Liang R-Q., Li W., Li Y., Tan C-Y., Li J-X., Jin Y-X., Ruan K-C. An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe. *Nucleic Acids Research* 2005;33(2):e17.

[79] Lin E.C. Glycerol dissimilation and its regulation in bacteria. *Annu Rev Microbiol* 1976;30:535-578.

[80] Lin J.J., Samcar A. A new mechanism for repairing oxidative damage to DNA (A)BC exonuclease removes AP sites and thymine glycols from DNA. *J Biochemistry* 1989;28(20):7979-7984.

[81] Loo T., Patchett M.L., Norris G.E., Lott J.S. Using secretion to solve a solubility problem: high-yield expression in *Escherichia coli* and purification of the bacterial glycoamidase PN-Gase F. *Protein Expr Purif* 2002;24:90-98.

[82] Lottspeich F., Zorbas H. 1998: *Bioanalytik*, Spektrum Akademischer Verlag, Heidelberg.

[83] Luirink J., Duim B., de Gier J.W., Oudega B. Functioning of the stable signal peptide of the pCloDF13-encoded bacteriocin release protein. *Mol Microbiol* 1991;5(2):393-399.

[84] Maier R., Eckert B., Scholz C., Lilie H., Schmid F.X. Interaction of trigger factor with the ribosome. *J Mol Biol* 2003;326:585-592.

[85] Manting E.H., Driessen A.J. *Escherichia coli* translocase: the unrevelling of a molecular machine. *Mol Microbiol* 2000;37:226-238.

[86] Manting E.H., van der Does C., Remigy H., Engel A., Driessen A.J. SecYEG assembles into a tetramer to form the active protein translocation channel. *EMBO J* 2000;19:852-861.

[87] Matsujama S., Tajima T., Tokuda H. A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. *EMBO J* 1995;14(14):3365-3372.

[88] Mergulaho F.J.M., Summers D.K., Monteiro G.A. Recombinant protein secretion in *Escherichia coli*. *Biotechnol Advances* 2005;23:177-202.

[89] Miksch G., Kleist S., Friehs K., Flaschl E. Overexpression of the phytase from *Escherichia coli* and its extracellular production in bioreactors. *Appl Microbiol Biotechnol* 2002;59:685-694.

[90] Miksch G., Neitzel R., Fiedler E., Friehs K., Flaschl E. Extracellular production of a hybrid beta-glucanase from *Bacillus* by *Escherichia coli* under different cultivation con-

ditions in shaking cultures and bioreactors. *Appl Microbiol Biotechnol* 1997;47:120-127.

[91] Missiakas D., Raina S. Signal transducing pathway in response to protein misfolding in the extracytoplasmic compartments of *Escherichia coli*: role of two new phosphoprotein phosphatases PrpA and PrpB. *EMBO J* 1997;16(7):1670-1685.

[92] Mori H., Ito K. The Sec protein-translocation pathway. *Trends Microbiol* 2001;9:494-500.

[93] Mori H., Ito K. Biochemical characterisation of a mutationally altered protein translocase: proton motive force stimulation of the initiation phase of translocation. *J Bacteriol* 2003;185:405-412.

[94] Mühlhardt C. (2003): *Der Experimentator: Molekularbiologie/Genomics*. Spektrum Akademischer Verlag, Heidelberg.

[95] Müller H-J., Röder T. (2004): *Der Experimentator: Microarrays*, Elsevier Spektrum Verlag, Heidelberg.

[96] Nagai K., Oubridge C., Kuglstatter A., Menichelli E., Isel C., Jovine L. Structure, function and evolution of the signal recognition particle. *EMBO J* 2003;22:3479-3485.

[97] Nikaido H., Rosenberg EY. Civ and Fiu proteins in the outer membrane of *Escherichia coli* catalyse transport of monomeric catechols: study with beta-lactam antibiotics containing catechol and analogous groups. *J Bacteriol* 1990;172(3):1361-1367.

[98] Nakatogawa H., Murakami A., Ito K. Control of SecA and SecM translation by protein secretion. *Curr Opin Microbiol* 2004;7:145-150.

[99] Nishiyama K., Fukuda A., Morita K., Tokuda H. Membrane deinsertion of SecA underlying proton motive force-dependent stimulation of protein translocation. *EMBO J* 1999;18:1049-1058.

[100] Nouwen N., Driessen AJ., SecDFyajC forms a heterotetrameric complex to with YidC. *Mol Microbiol* 2002;44:1397-1405.

[101] Otto K., Silhavy TJ. Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. *Proc Natl Acad Sci USA* 2002;99(4):2287-2292.

[102] Oudega B., Mol O., van Ulsen P., Stegehuis F., van der Wal FJ., Luirink J. *Escherichia coli* SecB, SecA, and SecY proteins are required for expression and membrane insertion of the bacteriocin release protein, a small lipoprotein. *J Bacteriol* 1993;175(5):1543-1547.

[103] Pallen MJ., Chaudhuri RR., Henderson IR. Genomic analysis of secretion systems. *Curr Opin Microbiol* 2003;6:519-527.

[104] Palmer T., Berks BC. Moving folded proteins across the bacterial cell membrane. *Microbiology* 2003;149:547-556.

-
- [105] Palmer T., Sargent F., Berks BC. Light traffic: photo-crosslinking a novel transport system. *Trends Biochem Sci* 2004;29:55-57.
- [106] Pauli G., Overath P. Ato-operon: a highly inducible system for acetoacetate and butyrate degradation in *Escherichia coli*. *Eur J Biochem* 1972;29(3):553-562.
- [107] Pedelacq JD., Piltch E., Liong EC., Berendzen J., Kim CY., Rho BS., *et al.* Engineering soluble proteins for structural genomics. *Nat Biotechnol* 2002;20:927-932.
- [108] Peterson JH., Woolhead CA., Bernstein HD. Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. *J Biol Chem* 2003;278:46155-46162.
- [109] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* 2001;9:2002-2007.
- [110] Poranen MM., Ravantti JJ., Grahn AM., Gupta R., Auvinen P., Bamford, DH. Global changes in cellular gene expression during bacteriophage PRD1 infection. *J Virol* 2006;80(16):8081-8088.
- [111] Pugsley AP. The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* 1993;57:50-108.
- [112] Pugsley AP., Francetic O. Protein secretion in *Escherichia coli* K12: dead or alive? *Cell Mol Life Sci* 1998;54:347-352.
- [113] Pugsley AP., Francetic O., Possot OM., Sauvonnnet N., Hardie K. Recent progress and future directions in studies of the main terminal branch of the general secretory pathway in Gram-negative bacteria-a review. *Gene* 1997;192:13-29.
- [114] Pugsley AP., Francetic O., Driessen AJ., de Lorenzo V. Getting out: protein traffic in prokaryotes. *Mol Microbiol* 2004;52:3-11.
- [115] Qui J., Schwartz JR, Georgiou G., Expression of active human tissue type plasminogen activator in *Escherichia coli*. *Appl Environ Microbiol* 1998;64:4891-4896.
- [116] Radonic A., Thulke S. Guideline for reference gene selection for quantitative real-time PCR. *Biochem Biophys Res Commun* 2004;313(4):856-862.
- [117] Rahman RN., Leow TC., Basri M., Salleh AB. Secretory expression of thermostable T1 lipase through bacteriocin release protein. *Protein Expr Purif* 2005;40(2):411-416.
- [118] Ren D., Bedzyk LA., Thomas SM., Ye RW., Wood TK. Gene expression in *Escherichia coli* biofilms. *Appl Microbiol Biotechnol* 2004;64(4): 515-524.
- [119] Rinas U., Hoffmann F. Selective leakage of host-cell proteins during high-cell-density

cultivation of recombinant and non recombinant *Escherichia coli*. Biotechnol Prog 2004;20:679-687.

[120] Perez-Rodriguez R., Fisher AC., *et al.* An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. J Mol Biol 2007;367:715-730.

[121] Robinson C., Bolhuis A. Tat-dependent protein targeting in prokaryotes and chloroplasts. Biochimica et Biophysica Acta 2004;1694:135-147.

[122] Rowe DC., Summers DK. The quiescent-cell expression system for protein synthesis in *Escherichia coli*. Appl Environ Microbiol 1999;65:2710-2715.

[123] Rusch SL., Kendall DA. Interactions that drive Sec-dependent bacterial protein transport. Biochemistry 2007;46:No34.

[124] Sambrook J., Fritsch EF., Maniatis T. (2002) Molecular Cloning: A laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour NY.

[125] Sanders C., Wethkamp N., Lill H. Transport of cytochrom c derivatives by the bacterial tat protein translocation system. Mol Microbiol 2001;41:241-246.

[126] Sandkvist M. Biology of type II secretion. Mol Microbiol 2001;40:271-283.

[127] Sandkvist M., Bagdasarian M. Secretion of recombinant proteins by Gram-negative bacteria. Curr Opin Biotechnol 1996;7:505-511.

[128] Santini CL., Ize B., Chanal A., Muller M., Giordano G., Wu LF. A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. EMBO J 1998;17:101-112.

[129] Sargent F., Gohlke U., De Leeuw E., Stanley NR., Palmer T., Saibil HR., *et al.* Purified components of the *Escherichia coli* Tat protein transport system form a double layered ring structure. Eur J Biochem 2001;268:3361-3367.

[130] Sarker S., Oliver D. Critical regions of secM that control its translation and secretion and promote secretion-specific secA regulation. J Bacteriol 2002;184:2360-2369.

[131] Saurin W., Hofnung M., Dassa E. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette ABC transporters. J Mol Evol 1999;48(1):22-41.

[132] Schadt EE., Li C., Su C., Wong WH. Analyzing high-density oligonucleotide gene expression array data. J Cell Biochem 2000;80(2):192-202.

[133] Schäfferling M., Nagel S. Optical technologies for the read out and quality control of DNA and protein microarrays. Anal Bioanal Chem 2006;385:500-517.

-
- [134] Schaffner J., Winter J., Rudolph R., Schwarz E. Cosecretion of chaperones and low-molecular-size medium additives increase the yield of recombinant disulfide-bridged proteins. *Appl Environ Microbiol* 2001;67:3994-4000.
- [135] Schena M., Shalon D., Davis RW., Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467-470.
- [136] Schenborn ET., Mierendorf RC. A novel transcription property of SP6 and T7 RNA polymerases: Dependence on template structure. *Nucl Acids Res* 1985;13:6223-6236.
- [137] Schirch V., Hopkins S., Villar E., Angelaccio S. Serine hydroxymethyltransferase from *Escherichia coli*: purification and properties. *J Bacteriol* 1985;163(1):1-7.
- [138] Scholz O., Thiel A., Hillen W., Niederweis M. Quantitative analysis of gene expression with an improved green fluorescent protein. *J Biochem* 2000;267:1565-1570.
- [139] Semighini CP., Marins M., Goldman MHS., Goldman GH. Quantitative Analysis of the relative transcript levels of ABC transporter *atr* genes in *Aspergillus nidulans* by real-time reverse transcription-PCR assay. *Appl Environ Microbiol* 2002;3:1351-1357
- [140] Shokri A., Sanden AM., Larsson G. Cell and process design for targeting of recombinant proteins into the culture medium of *Escherichia coli*. *Appl Microbiol Biotechnol* 2003;278:4654-4659.
- [141] Simmons LC., Yansura DG. Translational level is a critical factor for the secretion of heterologous proteins in *Escherichia coli*. *Nat Biotechnol* 1996;14:629-634.
- [142] Singer M., Berg P. (1992) *Gene und Genome*, Spektrum Akademischer Verlag, Heidelberg.
- [143] Sommer B., Schmidt B., Friehs K., Flaschel E. Extrazelluläre Produktion und Affinitätsreinigung einer rekombinanten Ribonuklease mit *Escherichia coli*. *Chemie Ingenieur Technik* 2008;80:No6.
- [144] Sommer B. New strategies for extracellular production of recombinant proteins with *E. coli*. 2008 PhD-thesis, University of Bielefeld, Department of engineering.
- [145] Southern EM., Maskos U., Elder JK. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics* 1992;13:1008-1017.
- [146] Stanley NR., Palmer T., Berks BC. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* 2000;275:11591-11596.
- [147] Stegehuis F., van der Waal FJ., Luirink J., Oudega B. Expression of the pCloDF13 encoded bacteriocin release protein or its stable signal peptide causes early effects on protein

biosynthesis and Mg^{2+} transport. *Antonie van Leeuwenhoek* 1995;67:255-260.

[148] Stekel D. *Microarray Bioinformatics* 2004; Oxford Press, Oxford.

[149] Stolpe S., Friedrich T. The *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I) is a primary proton pump but may be capable of secondary sodium antiport. *J Biol Chem* 2004;279(18):18377-18383.

[150] Strom AR., Kassen I. Trehalose metabolism in *Escherichia coli*: stress protection and stress regulation of gene expression. *Mol Microbiol* 1993;8(2):205-210.

[151] Sugai R., Shimizu H., Nishiyama K., Tokudu H. Overexpression of *yccL(gnsA)* and *ydY(gnsB)* increase the levels of unsaturated fatty acids and suppresses both the temperature-sensitive *FabA6* mutation and cold-sensitive *SecG* null mutation of *Escherichia coli*. *J Bacteriol* 2001;183(19):5523-5528.

[152] Thomas JD., Daniel RA., Errington J., Robinson C. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway. *Mol Microbiol* 2001;39:47-53.

[153] Tian H., Boyd D. Beckwith J. A mutant hunt for defects in membrane protein assembly yields mutations effecting the bacterial signal recognition particle and Sec machinery. *Proc Natl Acad Sci USA* 2000;97:4730-4735.

[154] Tullman-Ercek D., De Lisa MP., Kawarasaki Y., Iranpour P., Ribnicky B., Palmer T., Georgiou G. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. *J Biol Chem* 2007;11:8309-8316.

[155] Tziatzios C., Schubert D., Lotz M, Gundogan D., Betz H., Schagger H., *et al.* The bacterial protein-translocation complex: SecYEG dimers associate with one or two SecA molecules. *J Mol Biol* 2004;340:513-524.

[156] Ullers RS., Luirink J., Harms N., Schwager F., Georgopoulos C., Genevaux P. SecB is a bona fide generalized chaperone in V. *Proc Natl Acad Sci USA* 2004;101:7583-7588.

[157] Van der Waal FJ., Luirink J., Oudega B. Bacteriocin release proteins: mode of action, structure and biotechnological application. *FEMS Microbiol Rev* 1995a;17:381-399.

[158] Van der Waal FJ., ten Hagen CM., Oudega B., Luirink J. The stable bacteriocin release protein signal sequence, expressed as a separate entity, functions in the release of cloacin DF13. *FEMS Microbiol Lett* 1995b;131:173-177.

[159] Van der Waal FJ., ten Hagen-Jongman CM., Oudega B., Luirink J. Optimization of bacteriocin-release-protein-induced protein release by *Escherichia coli*: extracellular production of the periplasmic molecular chaperone FaeE. *Appl Microbiol Biotechnol* 1995c;44:459-465.

-
- [160] Van der Waal FJ., Koningstein G., ten-Hagen CM., Oudega B., Luirink J. Optimisation of bacteriocin release protein (BRP)-mediated protein release by *Escherichia coli*: random mutagenesis of the pCloDF13-derived BRP gene to uncouple lethality and quasi-lysis from protein release. *Appl Environ Microbiol* 1998;64:392-398.
- [161] van der Wal FJ., Valent QA., ten Hagen-Jongman CM., de Graaf FK., Oudega B., Luirink J. Stability and function of the signal peptide of the pCloDF13-derived bacteriocin release protein. *Microbiology* 1994;140(Pt 2):369-378.
- [162] van der Wal FJ., Oudega B., Kater MM., ten Hagen-Jongman CM., de Graaf FK., Luirink J. The stable BRP signal peptide causes lethality but is unable to provoke the translocation of cloacin DF13 across the cytoplasmic membrane of *Escherichia coli*. *Mol Microbiol* 1992;6(16):2309-2318.
- [163] Van der Wolk JP., Fekkes P., Boorsma A., Huie JL., Silhavy TJ., Driessen AJ. PrlA4 prevents the rejection of signal sequence defective preproteins by stabilizing the SecA-SecY interaction during the initiation of translocation. *EMBO J* 1998;17:3631-3639.
- [164] Vogelstein B., Gillespie D. Preparative and analytical purification of DNA from agarose. *Proc Nat Acad Sci USA* 1979;76:615-619.
- [165] Wachi M., Doi M., Okada Y., Matsushashi M. New mreC and mreD genes responsible for formation of the rod-shape of *Escherichia coli*. *J Bacteriol* 1989;171(12):6511-6516.
- [166] Wang X., Preston JF 3rd., Romeo T. The pgaABCD locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J Bacteriol* 2004;186(9):2724-2734.
- [167] Wilhelm J., Pingoud A. Real-time polymerase chain reaction. *Chem Bio Chem* 2003;4:1120-1128.
- [168] Wu L., Lin XM., Peng XX. From proteome to genome for functional characterisation of pH-dependent outer membrane proteins in *Escherichia coli*. *J Proteome Res* 2009.
- [169] Wu T., Mc Candlish AC., *et al.* Identification of a protein complex that assembles lipopolysaccharides in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci USA* 2006;103(31):11574-11579.
- [170] Xu R., Du P., Fan JJ., Zhang Q., Li TP., Gan RB. High-level expression and secretion of recombinant mouse endostatin by *Escherichia coli*. *Prot Expr Purif* 2002;24:453-459.
- [171] Yahr TL., Wickner WT. Functional reconstitution of bacterial Tat translocation in vitro. *EMBO J* 2001;20:2472-2479.
- [172] Yahr TL., Wickner WT. Evaluating the oligomeric state of SecYEG in preprotein translocase. *EMBO J* 2000;19:4393-4401.

[173] Yamabhai M., Emrat S., *et al.* Secretion of recombinant *Bacillus* hydrolytic enzymes using *Escherichia coli* expression systems. J Biotechnol 2008;133:50-57.

[174] Yang YH., Dudoit S., Luu P., Lin DM., Peng V., Ngai J., Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 2002;30(4):e15.

[175] Yokota N., Kuroda T., Matsuyama S., Tokuda H. Characterisation of the LolA-LolB-System as the general lipoprotein localization mechanism of *Escherichia coli*. J Biol Chem 1999;274(43):30995-30999.

[178] Yu P., San KY. Protein release in recombinant *Escherichia coli* using bacteriocin release protein. Biotechnol Prog 1992;8(1):25-29.

[179] Zhao K., Liv M., Burgess RR. The global transcriptional response of *Escherichia coli* to induced sigma 32 protein involves the sigma 32 regulon activation followed by inactivation and degradation of sigma 32 in vivo. J Biol Chem 2005;280(18):11578-11568.

A Anhang

A.1 Protokolle

A.1.1 *Spotten* von Oligonukleotiden auf Aldehyd-funktionalisierte *Glasslides*

Mit dem Affymetrix 417 Arrayer werden die 1:1 mit *Spotting*-Puffer versetzten Oligonukleotide (Endkonzentration 50 pmol/ μ l) auf die Glasoberfläche Aldehyd-funktionalisierter *Glasslides* (Super Aldehyd, Telechem) transferriert. Die Chips werden 3 Min bei 254 nm einer UV-Bestrahlung ausgesetzt, für 2 Stunden bei 80°C inkubiert und nachfolgend 24 Stunden bei Raumtemperatur im Dunkeln gelagert. Anschließend werden die Chips dem stichwortartig wiedergegebenen Prozedere unterzogen:

- 2 Min mit 0,2% SDS waschen
- 2x1 Min mit *ddH₂O* waschen
- 5 Min mit *NaBH₄*-Lösung (300ml 1xPBS, 100ml Ethanol, 1g *NaBH₄*) reduzieren
- 1 Min mit eiskaltem *ddH₂O* waschen
- 1 Min mit 0,2% SDS waschen
- 2x1 Min mit *ddH₂O* waschen
- Slides im Hybridisierungssofen bei 42°C für 45 Minuten in Blockierungslösung (6xSSC, 0,1% SDS und 1% BSA) inkubieren
- 5x 1 Min mit *ddH₂O* waschen
- Slides für 2 Min bei 1200 rpm trocken zentrifugieren

A.1.2 RNA-Isolation aus *E. coli* mittels Phenol-Chloroform Extraktion

Wenn nicht anders beschrieben werden alle experimentellen Schritte auf Eis durchgeführt. Alle verwendeten Geräte und Lösungen wurden autoklaviert und die Lösungen mit DEPC-behandelten Wasser angesetzt. 10^9 *E. coli* Zellen werden in 900 μ l Lysozym-Lösung (1 mg/ml in TE-Puffer) resuspendiert und mit dem gleichen Volumen an mit Wasser gesättigter Phenollösung versetzt. Es folgt eine Inkubation für 6 Minuten bei 64°C; alle 20 Sekunden wird das Reaktionsgefäß invertiert. Die Ansätze werden kurz auf Eis inkubiert (2 Min) und nachfolgend bei 4°C und 13.000 rpm für 15 Minuten zentrifugiert. Die obere wässrige Phase wird in ein neues Reaktionsgefäß überführt und mit dem gleichen Volumen Chloroform versetzt. Das Reaktionsgefäß wird mehrmals invertiert. Alternativ kann auch für 1 Minute gevortext werden. Es folgt eine 10 minütige Zentrifugation bei 4°C und 13.000 rpm. Erneut wird die obere wässrige Phase in ein neues Gefäß überführt und 1/10 des Volumens an 5 M NaCl sowie das doppelte Volumen an absolutem Ethanol (eiskalt) zugesetzt. Die Lösung wird für mindestens 20 Minuten bei -80°C inkubiert und nachfolgend 15 Minuten bei 4°C und 13.000 rpm zentrifugiert. Der Überstand wird verworfen und das Pellet mit 1 ml 75 prozentigem Ethanol (in DEPC-Wasser) gewaschen. Nachfolgend wird 5 Minuten bei 13000 rpm und 4°C zentrifugiert und das resultierende Pellet getrocknet und in 50-200 μ l nukleasefreiem Wasser resuspendiert. Photometrisch wird die Konzentration und Reinheit der RNA-Lösung bestimmt sowie die Integrität der RNA gelelektrophoretisch ermittelt.

A.1.3 RNA-Isolation aus *E. coli* mittels des *Innuprep*TM RNA Mini Kits (Analytik Jena)

Ca. 10^9 *E. coli* Zellen werden in 500 μ l Lysozym/TE-Lösung (1mg/ml) resuspendiert und 5 Minuten bei Raumtemperatur inkubiert. Nachfolgend werden 300 μ l Lysepuffer RL zugesetzt und die Lösung durch vorsichtiges Pipettieren homogenisiert. Es schließt sich eine weitere Inkubationszeit von 3 Minuten bei Raumtemperatur an, bevor die Lösung auf dem Spin-Filter D gegeben und 1 Minute bei 12000 rpm zentrifugiert wird. Der Durchlauf wird mit dem gleichen Volumen an 75 prozentigem Ethanol versetzt und der Spin-Filter R sukzessiv durch wiederholte Zentrifugationschritte für 2 Minuten bei 12000 rpm damit beladen. Der Filter wird mit 500 μ L Waschpuffer HS beladen und bei 12000 rpm für 1 Minute zentrifugiert. Der Durchlauf wird verworfen und es schließt sich ein äquivalenter Waschschrift mit 700 μ l Waschpuffer LS an. Nachfolgend wird die Säule für 3 Minuten bei 12000 rpm von ethanolischen Resten befreit und 30-100 μ l nukleasefreies Wasser auf die Säulenmitte appliziert. Nach Inkubation für 2 Minuten bei Raumtemperatur wird für 1 Minute bei 8000 rpm zentrifugiert (steriles RNase freies Reaktionsgefäß zum Auffangen der RNA-Lösung verwenden).

A.1.4 Direktes *Labeling* mit der Superscript III Reversen Transkriptase

50 μ g total RNA werden mit nukleasefreiem Wasser auf ein Volumen von 25,5 μ L eingestellt und mit 5 μ L Random Hexamerprimer sowie 3 μ L Primer-Mix (10 mM dATP, dGTP, dTTP; 1,6 mM dCTP) versetzt und bei 65°C für 5 Minuten inkubiert. Anschließend erfolgt die Zugabe von 10 μ L 5xRT-Puffer, 3 μ L DTT (100 mM), 1,5 μ L Cy3/Cy5-dCTP (1mM) sowie von 2 μ L Superscript III Reverser Transkriptase. Der Reaktionsansatz wird für 2 Stunden bei 37°C im Dunklen inkubiert. Nach Zugabe von 10 μ L 1 M NaOH erfolgt ein Inkubationsschritt für 10 Minuten bei 65°C um das RNA-Template zu verdauen und

nachfolgend eine Neutralisation mit 10 μL 1 M HCl. Dem Reaktionsansatz werden 350 μL Bindungspuffer PBI zugefügt und nachfolgend erfolgt die Aufreinigung über *Spin Columns* mittels des *PCR-Purification Kits* (Qiagen, Hilden). Die Elution von der Säule erfolgt in zwei nachfolgenden Schritten mit jeweils 25 μL 1:10 verdünntem EB-Puffer (pH 8,5). Zur Bestimmung der cDNA-Konzentration und der Inkorporationsfrequenz der fluoreszenzmarkierten Nukleotide werden 1,5 μL des Eluats im Spektralphotometer Nanodrop ND1000 vermessen. Als Nullreferenz fungiert der 1:10 verdünnte EB-Puffer. In der *speed-vac* wird das wässrige Lösungsmittel abgezogen und die Probe bis zur Trockene einrotiert. Für einen 50 μL Reaktionsansatz sind dabei ca. 45 Minuten bei Raumtemperatur zu veranschlagen. Die Probe wird in 9 μL Hybridisierungspuffer resuspendiert und nach Vereinigung mit der Vergleichsprobe für 3 Minuten bei 95°C inkubiert. Nach einer einminütigen Inkubation im Eis werden 2 μL Top-Block Lösung (20 mg/ml) zugesetzt und der Ansatz auf die Chipoberfläche transferiert. Als Verdunstungsschutz wird ein Deckgläschen verwendet, und die Ränder desselbigen durch Fixogum-Kleber verschlossen. Die so preparierten *Slides* werden über Nacht bei 42°C im Eppendorf-Schüttler bei 650 rpm inkubiert. Nach Entfernung des Deckgläschens mitsamt aller Klebereste werden die Chips in 3 sukzessiven Waschschritten mit Waschpuffer 1-3 gewaschen. Um ein Ausfallen des im Waschpuffer 1 enthaltenen SDS zu verhindern werden die Waschpuffer leicht erwärmt. Durch 3 minütige Zentrifugation bei 1200 rpm werden die Chips getrocknet und im konfokalen Laserscanner Gene Pix 4000B (Axon) bei verschiedenen Photomultipliersensitivitäten und Laserstärken vermessen. Dies dient dazu um den gesamten dynamischen Bereich des Chips zu erfassen.

A.1.5 Indirektes *Labeling* mit dem *MicromaxTMTSATM Labeling Kit*

6 μg der RNA-Präparation werden mit 1 μL Biotin- bzw. Fluorescein-dCTP, 2 μL Random-Hexamerprimern und 2 μL dNTPs versetzt und mit nukleasefreiem Wasser auf 14 μL aufgefüllt. Nachfolgend wird der Ansatz für 10 Minuten bei 65 °C inkubiert und nach dem Abkühlen auf Eis mit folgenden Substanzen versetzt:

- 5 μL First Strand Buffer
- 3 μL DTT
- 2 μL Superscript III Reverse Transkriptase
- 1 μL RNase Inhibitor (falkultative Zugabe)

Der Ansatz wird lichtdicht abgedeckt und für 2 Stunden bei 42°C inkubiert. Anschließend wird durch Zugabe von 2,5 μL 1 M NaOH und 2,5 μL 0,5 M EDTA-Lösung (pH 8,0) und Inkubation für 10 Minuten bei 65°C ein alkalischer RNA-Verdau durchgeführt. Durch Zugabe von 6,5 μL 1 M Tris-HCL Lösung (pH 7,5) wird die Lösung neutralisiert und mit 250 μL Bindungspuffer PBI versetzt. Es erfolgt die bereits beschriebene Aufreinigung mittels des *PCR-Purification Kits*. Als erster und zusätzlicher Waschschriff werden, abweichend vom beschriebenen Protokoll, 750 μL 35%iges Guanidinhydrochlorid auf die Säule gegeben und 1 Minute bei 13.000 rpm zentrifugiert. Nachfolgend wird weiter nach dem Standardprotokoll verfahren und mit PE-Puffer gewaschen. Die Inkorporationsfrequenz von Biotin- und Fluorescein-dCTP kann leider nicht spektrophotometrisch bestimmt werden, ebenso fällt die generierte Menge an cDNA unter die Nachweisgrenze. Deshalb wird von dem cDNA-Ansatz direkt in der *speed-vac* das wässrige Lösungsmittel abgezogen. Die getrocknete cDNA wird in

9 μL Hybridisierungspuffer resuspendiert und mit der unterschiedlich *gelabelten* Vergleichsprobe vereinigt und für 3 Minuten bei 95°C inkubiert. Nach einer Inkubation auf Eis für eine Minute wird dem Ansatz 2 μL Top Block zugefügt und die Lösung auf die bespottete Chipoberfläche transferriert. Analog zum *Direktlabeling* fungiert ein Deckgläschen, welches mit Fixogumm abgedichtet wurde, als Verdunstungsschutz. Die Hybridisierung findet in einem Eppendorf-Schüttler mit *Slide*-Aufsatz bei 42°C und 650rpm über Nacht statt. Das Deckgläschen wird vom Chip entfernt und die drei ersten Waschschrte analog zum *Direktlabeling*-Protokoll durchgeführt. Nachfolgend wird 1 Minute mit TNT-Puffer gewaschen und die *Spot*-Fläche mit einem immunologischen Markerstift auf Fettbasis umrandet. In diese umrandete Fläche werden 300 μL TNB-G-Puffer appliziert und so verteilt, dass die gesamte Fläche mit Flüssigkeit benetzt ist. Es folgt eine 10 minütige Inkubation bei Raumtemperatur. Das nachfolgende Protokoll wird stichwortartig aufgeführt, alle Inkubations- und Waschrte werden aufgrund der Photosensitivität von Fluorescein und den Cyanin-Farbstoffen unter Lichtausschluss durchgeführt. Alle Inkubationsschrte werden für 10 Minuten durchgeführt.

- Waschrte in TNT-Puffer für 1 Minute
- Inkubation mit 200 μL Antifluorescein-Antikörperlösung für 10 Minuten (2 μL Stammlösung des Antikörpers in 200 μL TNB-G-Puffer gelöst)
- 3 Waschrte für 1 Minute in TNT-Puffer
- Inkubation mit 250 μL Streptavidin-Cy3-Lösung (0,5 μL der Streptavidin Cy3-Stammlösung in 250 μL Amplification Diluent (S) gelöst.
- 3 Waschrte mit TNT-Puffer für jeweils 5 Minuten
- Inkubation mit 300 μL HRP-Inaktivierungslösung (10 μL 3M Natriumacetatlösung werden zu 290 μL Lösung T gegeben) für 10 Minuten.
- 3 Waschrte für jeweils 1 Minute mit TNT-Puffer
- Inkubation mit 200 μL Anti-Biotin-Antikörperlösung (2 μL der Antikörperstammlösung werden in 200 μL TNB-G-Puffer gelöst)
- 3 Waschrte für jeweils 1 Minute in TNT-Puffer
- Inkubation mit 250 μL Cy5-Streptavidinlösung (0,5 μL Cy5-Streptavidin werden in 250 μL Amplification Diluent (S) gelöst
- 3 Waschrte in TNT-Puffer für jeweils 5 Minuten.
- 1 Waschrte mit 1:10 verdünntem Waschpuffer 3 (0,1xSSC) für 1 Minute
- Zentrifugation der *Slides* für 3 Minuten bei 1200 rpm.

Nachfolgend kann das *Scannen* der *Slides* bei variierenden PMT- und Lasereinstellungen erfolgen.

A.1.6 Indirekte Markierung der cDNA mit Streptavidin funktionalisierten *Quantum Dot*-Partikeln

50 μg total RNA werden mit nukleasefreiem Wasser auf ein Volumen von 23 μl eingestellt und folgende Substanzen der Lösung zugesetzt:

- 3 μl Biotin-dCTP
- 5 μl Random Hexamer
- 3 μl dNTPs (10 mM dATP, dGTP, dTTP und 1,6 mM dCTP)

Dieser Ansatz wird für 5 Minuten bei 65°C inkubiert und anschließend kurz in einem Eisbad abgekühlt. Nachfolgend werden dem RT-Ansatz gelistete Reagenzien in der angegebenen Reihenfolge zugesetzt:

- 10 μl 5x RT-Puffer Invitrogen (First Strand Puffer)
- 3 μl DTT
- 3 μl Superscript III Reverse Transkriptase

Der Reaktionsansatz wird für 120 Minuten bei 42°C inkubiert und anschließend durch Zugabe von 10 μl 1 M NaOH und Inkubation bei 65°C für 10 Minuten ein alkalischer RNA-Verdau durchgeführt. Die Neutralisation erfolgt durch Zugabe von 10 μl 1 M HCl. Die so gewonnene, Biotin-markierte cDNA wird mittels des PCR-Purification Kits aufgereinigt, mit 2x 25 μl EB-Puffer (1:10 verdünnt; pH 8,5) von der Säule eluiert und in der speed-vac bis zur Trockene einrotiert. Das cDNA-Pellet wird in 18 μl Hybridisierungspuffer aufgenommen, für 3 Minuten auf 95°C erhitzt und 1 Minute im Eisbad abgekühlt. Es folgt die Zugabe von 2 μl wässriger Top Block-Lösung (20 mg/ml) und der Transfer der so generierte Hybridisierungslösung auf die Chipoberfläche. Ein Deckgläschen fungiert für die nachfolgende Hybridisierung über Nacht in einem Eppendorf-Schüttler (42°C, 650 rpm) als Verdunstungsschutz. Nach erfolgter Hybridisierung wird das Deckgläschen entfernt und die Chips unter Lichtausschluß je 5 Minuten mit WP1, WP2 und WP3 gewaschen.

Mittels eines Immunologischen Fettstiftes wird die Spotfläche auf dem Chip umrahmt, 300 μl TNB-G Puffer appliziert und für 10 Minuten inkubiert. Der Chip wird anschließend für 5 Minuten mit PBS-Puffer (pH 7,4) gewaschen und 200 μl *Quantum Dot*-Lösung auf die umrahmte Fläche gegeben. Zur Herstellung der *Quantum Dot*-Lösung wird die bei 4°C gelagerte *Quantum Dot*-Stammlösung für 3 Minuten bei 5000 g zentrifugiert und 10 μl des Überstandes zu 200 μl Inkubationspuffer (6%BSA in 1xPBS-Puffer pH 7,4) gegeben. Die *Quantum Dot*-Lösungen sollten nicht gevortext werden. Die Inkubation auf der Chipoberfläche wird für mindestens 3 Stunden durchgeführt und der Chip nachfolgend für 5 Minuten in 1xPBS-Puffer gewaschen. Weitere Waschschrte mit 1xPBS können gegebenenfalls das SNR erhöhen. Durch Zentrifugation für 3 Minuten bei 1200 rpm wird der Chip getrocknet und anschließend bei unterschiedlichen PMT-Empfindlichkeiten und Laserstärken gescannt.

A.1.7 Plasmidisolierung mit dem *Peq-Lab Miniprep Kit II*

a) *high-copy* Plasmide

10ml einer Übernacht-Kultur des jeweiligen Bakterienstammes in LB-Medium werden in ein *Falcon-tube* überführt und bei 4°C und 4000rpm für 15 Minuten zentrifugiert. Das resultierende Pellet wird in 500 µl Lösung 1 resuspendiert. Anschließend werden 500 µl Lösung 2 appliziert, 10-15 mal invertiert sowie 5 Minuten bei Raumtemperatur inkubiert. Nun erfolgt die Zugabe von 700 µl Lösung 3 und eine erneute Invertierung für 10-15 mal. Um die Präzipitation möglichst quantitativ durchzuführen wird die Plasmidpräparation 2 Minuten auf Eis inkubiert und für 10 Minuten bei 13.000 rpm zentrifugiert. Der Überstand wird auf die im Kit enthaltenden Säulen appliziert und 1 Minute bei 13.000 rpm zentrifugiert. Dieser Schritt wird sukzessiv wiederholt, bis der gesamte Überstand aufgetragen wurde. Es schließt sich ein Waschschriff mit 500 µl Waschpuffer 1 und nachfolgender Zentrifugation bei 13.000 rpm für 1 Minuten an. Analog wird mit 700 µl Waschpuffer 2 verfahren. Zur vollständigen Entfernung ethanolischer Reste wird 2 Minuten bei 13.000 rpm zentrifugiert. Um die Plasmidlösung von der Säule zu eluieren werden 75 µl nukleasefreies Wasser auf die Säule appliziert und 1 Minute bei Raumtemperatur inkubiert. Zum Schluß wird 1 Minute bei 13.000 rpm zentrifugiert und die Plasmidlösung in einem sterilen 1,5 ml Reagiergefäß aufgefangen. Die Konzentration und Reinheit der Mini-Präparation wird photometrisch bestimmt.

b) *low-copy* plasmide

Für die Isolierung von *low-copy* Plasmiden werden 20 ml bakterielle LB-Übernachtskultur verwandt und 15 Minuten bei 4000 rpm und 4°C zentrifugiert. Das Pellet wird in 1500 µl Lösung 1 resuspendiert und respective mit 1500 µl Lösung 2 versetzt. Nach der 5 minütigen Inkubation bei Raumtemperatur wird mit 2100 µl Lösung 3 neutralisiert und anschließend weiter wie bei den *high-copy* Plasmiden beschrieben verfahren.

A.1.8 Herstellung von LB-Agarplatten

- 10 g Trypton
- 10 g NaCl
- 5 g Hefeextrakt
- 15 g Agar

werden mit 800 ml *ddH₂O* versetzt, der pH-Wert auf 7 eingestellt und mit weiterem *ddH₂O* ein Gesamtvolumen von 1000 ml eingestellt. Die Lösung wird 20 Min bei 121°C autoklaviert und auf ca. 50°C abgekühlt. Nun erfolgt die Zugabe des entsprechenden Antibiotikums:

- Ampicillin: 4 ml Ampicillin-Stammlösung(25 mg/ml)
- Chloramphenicol: 3 ml Chloramphenicol-Stammlösung(10 mg/ml)
- Ampicillin + X-Gal + IPTG:
 - 3 ml Ampicillin-Stammlösung
 - 40 mg X-Gal(gelöst in N,N-Dimethylformamid)
 - 120 mg IPTG (gelöst in *ddH₂O*)

A.1.9 Herstellung chemisch kompetenter Zellen mittels TSS-Methode

100 ml LB-Medium wird in 500 ml Schüttelkolben mit 500 μ l einer Vorkultur des gewünschten Bakterienstammes inokuliert und im Schüttelschrank bei 37°C und 150 rpm kultiviert. Beim Erreichen einer OD600 von 0,5-0,8 wird die Kultur in Falcon-Tubes überführt und für 15 Minuten zentrifugiert (4°C; 4000 rpm). Der Überstand wird möglichst quantitativ entfernt und das Pellet in 10 ml eiskalter TSS-Lösung resuspendiert. Nach Aliquotierung in 1,5 ml Reaktionsgefäße werden die Zellen in flüssigem Stickstoff schockgefroren und bei -80°C gelagert.

A.1.10 Herstellung chemisch kompetenter Zellen mittels *CaCl*₂-Methode

240 ml SOB-Medium werden im 500 ml Schüttelkolben mit 1000 μ l einer über Nacht gewachsenen Vorkultur des jeweiligen Bakterienstammes inokuliert. Die Kultur wird bei 18°C für 3-4 Tage unter intensivem Schütteln (200-250 rpm) im Schüttelschrank inkubiert bis die optische Dichte bei 600 nm 0,8 beträgt. Die Kultur wird in Zentrifugationsgefäße überführt und bei 4°C und 4000 rpm für 15 Minuten zentrifugiert. Der Überstand wird möglichst quantitativ entfernt und das resultierende Pellet in 80 ml eiskalter TB-Lösung vorsichtig resuspendiert und für 10 Minuten auf Eis inkubiert. Es schließt sich ein erneuter Zentrifugationsschritt bei 4000 rpm und 4°C für 10 Minuten an. Das resultierende Pellet wird nun mit 20 ml TB-Lösung und 1,5 ml DMSO versetzt und vorsichtig resuspendiert. Die Lösung wird in PCR-Tubes aliquotiert und sofort in flüssigem Stickstoff schockgefroren. Die nachfolgende Lagerung erfolgt bei -80°C. Die so behandelten Zellen sind mindestens 2 Monate ohne Kompetenzverlust lagerbar.

A.1.11 Transformation chemisch kompetenter Zellen mittels Hitzeschock

100 μ L der bei -80°C gelagerten, chemisch kompetenten Zellen werden auf Eis inkubiert. Sofort nach dem Auftauen werden 1-10 μ L des Transformationsansatzes zugefügt und durch schnelles Auf- und Abpipettieren suspendiert. Anschließend erfolgt eine 15 minütige Inkubation auf Eis bis die Zellen einem zweiminütigen Hitzeschock bei 42°C ausgesetzt werden. Sofort danach werden die Zellen zwei Minuten auf Eis inkubiert und 900 μ L LB- oder SOC-Medium zugefügt. Die Verwendung von SOC-Medium generiert dabei höhere Kompetenzen und ist in kritischen Fällen der Verwendung von LB-Medium vorzuziehen. Der Transformationsansatz wird für 60 Minuten bei 37°C und 300 rpm geschüttelt. Dieser Schritt dient dazu das Plasmid in der Zelle zu etablieren und die Antibiotikaresistenz auszubilden. Bei der Verwendung von bakteriostatischen Antibiotika als Selektionsdruck (z.B. Ampicillin) ist eine 15 minütige Inkubation ausreichend. Die Zellen werden nachfolgend für 3 Minuten bei 3000 rpm sedimentiert und der Überstand bis auf ca. 150 μ L verworfen. In dem Restvolumen werden die Zellen resuspendiert und auf den Agarplatten mit dem entsprechenden selektiven Antibiotikum ausplattiert.

A.1.12 Ligation mittels *cut-ligation* Methode

Für die Ligation von 200 ng des Vektors welcher eine *blunt-end* Schnittstelle enthält werden zugesetzt:

- 400 ng Insert
- 2 μL PEG 4000
- 4 μL Tango 10xPuffer
- 1 μL EcoRV (1u/ μL) bzw. *blunt-end* schneidendes Enzym
- 1 μL 10 mM ATP
- 1 μL T4-DNA-Ligase (5 u/ μL)

Der Ligationsansatz wird mit nukleasefreiem Wasser auf 20 μL aufgefüllt. Nachfolgend wird 2 Stunden bei Raumtemperatur und anschließend über Nacht im Kühlschrank bei 4°C inkubiert.

A.1.13 Ligation mittels kohäsiver Enden

200 ng des restriktionsverdauten Vektors werden im molaren Verhältnis von 3:1 bis 5:1 mit dem Insert versetzt. Dem Ligationsansatz werden 1 μL T4-DNA-Ligase (5 u/ μL) und 2 μL 10x Ligationspuffer (ATP enthalten) zugefügt und mit nukleasefreiem Wasser auf 20 μL aufgefüllt. Die Ligation erfolgt 4 Stunden bei 16°C oder in einer großen, mit Wasser von Raumtemperatur gefüllten Wanne über Nacht im Kühlraum (Gradientenligation).

A.1.14 Restriktionsverdau

1 μg der zu schneidenden DNA wird mit 1 μL Restriktionsenzym und 2 μL Restriktionspuffer versetzt und mit nukleasefreiem Wasser auf 20 μL aufgefüllt. Für eine vollständige Prozessierung wurde 1 Stunde bei 37°C inkubiert und anschließend eine 20 minütige Hitzeinaktivierung der Restriktionsendonuklease bei 65°C durchgeführt. Für größere Mengen DNA wurde das Protokoll entsprechend erweitert. Für den Restriktionsverdau von PCR-Fragmenten wurden Inkubationszeiten von mindestens 4 Stunden gewählt. Der Doppelverdau von DNA wurde im Universalpuffer der Firma Fermentas (Tango-Puffer) durchgeführt. Je nach verwendeter Enzymkombination wurde der Tango-Puffer in einfacher oder doppelter Arbeitskonzentration verwendet.

A.1.15 Aufreinigung mittels *PCR-Purification Kit*

Diese Aufreinigungsmethode basiert auf der pH-abhängigen Adsoption von DNA auf einer Silikamatrix. Bei pH-Werten unter 7 sind die Silanolgruppen der Silikamatix undissoziiert und damit ungeladen. In diesem Zustand lässt sich die Membran mit der zu reinigenden DNA beladen. Nach verschiedenen Waschschritten zur Elution der Verunreinigungen wie z.B. Proteinkontaminationen, kann durch Verwendung eines Elutionspuffers mit basischem pH-Wert die aufgereinigte DNA von der Säule eluiert werden. Die Erhöhung des pH-Wertes führt zu einer Dissoziation und damit negativen Ladung der Silanolgruppen. Die durch das

Zucker-Phosphatrückgrat ebenfalls negativ geladene DNA erfährt eine elektrostatische Abstoßung und eluiert somit von der Säule. Zur Aufreinigung von PCR-Produkten oder cDNA-Ansätzen für Microarray-Applikationen wurde das kommerziell erhältliche *PCR-Purification Kit* der Firma Qiagen verwendet, welches auf dem beschriebenen Prinzip basiert. Die aufzureinigende DNA wird dabei mit dem fünffachen Volumen an Puffer PBI versetzt und auf die Säule transferiert. Ein zugesetzter pH-Indikator zeigt den korrekten pH-Wert für eine Bindung auf der Säule an. Es erfolgt ein Zentrifugationsschritt für 1 Minute bei RT und 13.000 rpm. Der Durchfluß wird verworfen und nachfolgend 750 μL Waschpuffer auf die Säule überführt. Bei salzsensitiven Applikationen wird 5 Minuten inkubiert, bevor eine erneute Zentrifugation mit gleichen Bedingungen wie im ersten Zentrifugationsschritt erfolgt. Es folgt ein Trocknungsschritt durch Zentrifugation für 1 Minute bei 13.000 rpm um Ethanolreste von der Säule zu entfernen. Diese stammen aus dem Ethanol-basierten Waschpuffer und interferieren mit vielen *downstream*-Applikationen. Durch zweimaliges Beladen der Säule mit 25 μL 1:10 verdünnten Elutionspuffer EB (pH 8,5) und nachfolgender Zentrifugation von 1 Minute bei 13.000 rpm wird die DNA von der Säule eluiert.

A.1.16 Gelaufreinigung

Für die Isolation und Aufreinigung von DNA-Banden aus Agarosegelen wurde das kommerziell erhältliche *Wizard SV Gel and PCR-Purification Kit* der Firma Promega verwendet. Das Agarosegel wurde mit Ethidiumbromid für 15 Minuten angefärbt und anschließend 5 Minuten in *ddH₂O* gewaschen. Die Visualisierung der DNA-Banden erfolgte auf einem UV-Tisch und die gewünschten Banden wurden mittels eines Skalpell aus der Gelmatrix ausgeschnitten, in ein Eppendorfgefäß transferiert, und ihre Masse bestimmt. Die UV-Exposition des Geles wurde so kurz wie möglich gehalten, um die Einführung von Mutationen in die zu isolierende DNA zu minimieren. Pro 10 mg Gelmatrix wurden 10 μL Membran- und Bindungspuffer zugesetzt und 10 Minuten bei 50 °C inkubiert. Dabei wurde alle 2 Minuten intensiv gevortext. Nach dieser Zeit waren visuell keine Gelpartikel mehr zu erkennen, und der gesamte Ansatz wurde nach dem Abkühlen auf die Säule überführt. Dort erfolgte eine Inkubation von einer Minute bei Raumtemperatur. Die Bindung auf der Säule erfolgte durch eine Zentrifugation für 1 Minute bei 13.000 rpm. Die weitere Probenprozessierung erfolgte gemäss dem folgenden Protokoll:

- Zugabe von 700 μL Waschpuffer und Zentrifugation für 1 Minute bei 13.000 rpm
- Zugabe von 500 μL Waschpuffer und Zentrifugation für 5 Minuten bei 13000 rpm
- Trocknen der Membran durch Zentrifugation für 1 Minute bei 13000 rpm, bei diesem Schritt bleibt der innere deckel der Zentrifuge geöffnet
- Zugabe von 40 μL nukleasefreiem Wasser auf die Membran und Inkubation für mindestens 1 Minute
- Elutionsschritt durch Zentrifugation bei 8000 rpm für 1 Minute

A.1.17 Generierung von cDNA für PCR-Applikationen

2 μg total RNA werden mit

- 3 μl dNTPs (2 mM)
- 3 μl Random Hexamerprimer (100 mM)

versetzt, mit *ddH₂O* auf 21 μl aufgefüllt und für 5 Minuten bei 65°C inkubiert. Nach einer kurzen Inkubation auf Eis werden

- 8 μl 5xTranskriptionspuffer
- 1 μl Improm Reverse Transkriptase 10 Units/ μl (Promega)

A.1.18 Real-time PCR

Zur Generierung der Mastermixe werden folgende Substanzen verwendet:

- 10 μl IQSybrGreenMastermix
- 4,3 μl *ddH₂O*
- 0,7 μl Primer Forward+Reverse(10 pmol/ μl)

Der so präparierte Mastermix wird in die entsprechenden PCR-Platten überführt und final mit 5 μl der 1:20 verdünnten cDNA versetzt. Eine PCR-Folie wird über die Platte geklebt und diese in den PCR-Cycler überführt.

Für die *real-time* Experimente wurde standardmäßig folgendes PCR-Programm verwendet:

- Initiale Denaturierung 98°C 1 Min
- 40 Amplifikationszyklen je:
Denaturierung: 15 Sec 95°C
Primer-Annealing: 30 Sec variierende Temperatur
Primer-Elongation: 30 Sec 72°C
- Denaturierung und Reassoziaton zur Vorbereitung für die Schmelzkurvenanalyse:
30 Sec 95°C
30 Sec 60°C
- Schmelzkurvenanalyse:
60°C Starttemperatur, 10 Sec je Schritt, Gradient 0,5°C pro Schritt, 95°C Endtemperatur.

Die Fluoreszenzmessung erfolgte in der Primer-Elongationsphase

A.1.19 PCR mit *Taq*-Polymerase

Für die PCR mittels OLS *Taq*-Polymerase wurde folgender Mastermix verwendet:

- 36,25 μL nukleasefreies Wasser
- 4 μL dNTPS, jeweils 2 mM dATP, dCTP, dGTP, dTTP
- 0,25 μL *Taq*-Polymerase
- 5 μL 10xPuffer OLS
- 4 μL Primer (forward + reverse, je 10 pM)

Dazu werden 0,5 μL des jeweiligen cDNA-Templates getan, bei den Negativkontrollen wird die cDNA durch das gleiche Volumina nukleasefreiem Wasser substituiert. Für eine PCR unter Standardbedingungen wird folgendes Programm gewählt:

- Initiale Denaturierung für 3 Minuten bei 95°C
- 30 Zyklen je:
Denaturierung 30 Sec 95°C
Primer-Annealing 30 Sec bei 55-70°C, variable Temperatur, abhängig vom verwendeten Primerpaar
Primer-Elongation 45 Sec 72°C (für Amplifikate bis 1 kb Länge)
- Finale Primer-Elongation 72°C 5 Min.

A.1.20 PCR mittels *Pfu*-Polymerase

Folgendes Protokoll wurde für die Amplifikation von DNA mittels der *Pfu*-Polymerase verwendet:

- 35 μL nukleasefreies Wasser
- 5 μL dNTPS, jeweils 2 mM dATP, dCTP, dGTP, dTTP
- 0,50 μL *Pfu*-Polymerase
- 5 μL 10xPuffer *Pfu* inkl. Magnesiumsulfat
- 4 μL Primer (forward + reverse, je 10 pM)

Dazu werden 0,5 μL des jeweiligen cDNA-Templates oder Plasmides getan, bei den Negativkontrollen wird die cDNA durch das gleiche Volumina nukleasefreiem Wasser substituiert. Es wird das gleiche Standardprotokoll wie bei der *Taq*-Polymerase verwendet, lediglich die Elongationszeiten der Primer werden erhöht. Für Amplifikate bis 1 kb wird 1 Minute elongiert, für längere Amplifikate wird die Dauer bis zu 4 Minuten ausgedehnt.

A.1.21 Kolonie-PCR

Die Kolonien auf der Agarplatte wurden mit einer sterilen, DNA-freien Spitze berührt und die dabei anhaftenden Zellen in ein mit endotoxinfreiem Wasser (10 μL) gefülltes Eppendorfgefäß überführt. Die Zellsuspension wurde für 5 Minuten bei 95 °C inkubiert um eine Zellyse herbeizuführen. 5 μL dieser Zellsuspension fungierten als Target in der nachfolgenden PCR-Reaktion. In dieser wurden auch immer Negativkontrollen parallel durchgeführt. Dazu werden 0,5 μL des jeweiligen cDNA-Templates getan, bei den Negativkontrollen wird die cDNA durch das gleiche Volumina nukleasefreiem Wasser substituiert.

A.1.22 Generierung DIG-markierter DNA-Sonden mittels PCR

Die Inkorporation von DIG-11-dUTP kann mittels PCR nach dem angegebenen Standard-Protokoll erfolgen. Dazu wird folgender Nukleotidmix anstelle der regulären dNTPs verwendet:

- 2 mM dATP, dCTP, dGTP
- 1,3 mM dTTP
- 0,7 mM DIG-11-dUTP

Die Visualisierung der DIG-Inkorporation kann durch einen Gelshift auf dem Agrosegel nachgewiesen werden. Zur Abschätzung der inkorporierten DIG-Menge kann die Sonde zusätzlich auf eine Membran *gespottet* werden und nach kolorimetrischer Detektion mittels NBT/BCIP mit einer *gespotteten* Verdünnungsreihe bekannter Konzentration verglichen werden. Dazu wird jeweils 1 μl einer Verdünnungsreihe der generierten Sonde auf die Nylonmembran *gespottet* und diese nachfolgend für 30 Minuten bei 120°C inkubiert. Nach kurzem Schwenken der Membran in Waschpuffer wird die Membran für 30 Minuten in Blockierungslösung inkubiert und mit 10 ml DIG-Antikörperlösung (2 μL Anti-DIG-AP in 10 ml Blockierungslösung verdünnt) für 30 Minuten versetzt. Nach zweimaligen Waschen für jeweils 15 Minuten in Waschpuffer wird die Membran für 2 Minuten in Detektionspuffer äquilibriert. Zum Schluß erfolgt der Zusatz von 10 ml NBT/BCIP-Lösung (100 μL NBT/BCIP in 10 ml Detektionspuffer) bis sich die Spotintensität voll entwickelt hat. Die Reaktion wird durch Zusatz von destilliertem Wasser gestoppt.

A.1.23 Generierung DIG-markierter cDNA mittels Reverser Transkription

50 μg total RNA werden mittels Superscript III Reverser Transkriptase in DIG-cDNA konvertiert. Dazu wird die RNA-Lösung mit nukleasefreiem Wasser auf 25 μL aufgefüllt und folgender Ansatz bereitet:

- 5 μL genspezifische Primer, 10 mM
- 3 μL dNTPs(10 mM dATP, 10 mM dCTP, 10 mM dGTP, 1,6 mM dTTP)
- 2 μL DIG-11-dUTP, 1 mM
- 25 μL RNA-Lösung

Der Ansatz wird für 5 Minuten bei 65°C inkubiert und mit nachfolgend aufgeführten Substanzen versetzt.

- 10 μL 5xRT-Puffer
- 3 μL DTT
- 2 μL Superscript III Reverse Transkriptase

Die RT-Reaktion wird durch Inkubation für 2 Stunden bei 42°C durchgeführt.

A.1.24 Genierung von DIG-markierten RNA-Sonden mittels *in vitro* Transkription

Gegenüber der Generierung unmarkierter RNA mittels *in vitro* Transkription wird ein leicht abgewandeltes Protokoll verwendet. Der Ribonukleotidmix wird durch folgende Mischung der Ribonukleotide ersetzt:

- 2,5 mM rATP, rCTP, rGTP
- 1,3 mM rUTP
- 0,7 mM DIG-11-rUTP

1 μl nukleasefreies Wasser wird durch 1 μl RNA-Polymerase substituiert. Die übrige Versuchsvorschrift wird vollständig analog durchgeführt.

A.1.25 SDS-PAGE

Trenn- und Sammelgel werden gemäss der Anleitung gegossen und in die vorbereitete Elektrophoreseapparatur überführt. Die destinierten Proben für die PAGE-Analytik werden mit dem gleichen Volumen Lämmli-Puffer (frisch hergestellt aus 900 μl 2xSDS-Probenpuffer und 100 μl β -Mercaptoethanol) versetzt. Der Probenpuffer sollte dabei seine blaue Färbung beibehalten, gegebenenfalls wird bei einem Farbumschlag ins Gelbe mit 1-2 μl 2 M Tris-Base Lösung (pH 9,5) neutralisiert. Anschließend wird die Probe bei 95°C für 5 Min inkubiert und kurz zentrifugiert (10000 rpm; 1 Min). Die Probe wird in die Geltaschen des Sammelgels appliziert und bei 100 V solange elektrophoretisch getrennt bis die Probenfront das Trenngel erreicht hat. Nun wird die Spannung auf 200 V erhöht und die Elektrophorese fortgesetzt bis die Bromphenolblau-Front die untere Gelkante erreicht hat. Die Elektrophorese wird beendet und die Apparatur auseinandergenommen. Dabei wird das Gel vorsichtig vom Glaträger getrennt und in die entsprechende Färbelösung transferriert.

Coomassie-Färbung

Das Geld wird für mindestens 120 Minuten in Coomassie-Lösung geschwenkt (50 U/min). Anschließend wird die Färbelösung entfernt und das Gel in Entfärber/Fixierer inkubiert, bis Banden sichtbar werden und der Hintergrund keine nennenswerte Blaufärbung mehr aufweist. Durch sukzessiven Austausch der Entfärber/Fixierer-Lösung kann der Entfärbungsvorgang beschleunigt werden.

A.1.26 *Nothern Blotting*

I Denaturierende Agarosegelelektrophorese

Im Folgenden wird die Herstellung eines denaturierenden 1,5%igen Agarosegels beschrieben, welches für die Auftrennung der meisten RNA-Spezies geeignet ist. Die nachfolgend aufgeführten Reagenzien werden homogen suspendiert und bei 600 Watt in der Mikrowelle aufgekocht.

- 1,2 g Agarose
- 4 ml 20xMOPS
- 61,7 ml DEPC-H₂O

Nach dem Abkühlen auf ca. 50°C werden 14,3 ml 12,3 M Formaldehydlösung zugesetzt und das Gel gegossen. 5 µL der RNA-Lösung werden mit 12 µL Probenpuffer und 3 µL Ladepuffer versetzt. Pro Geltasche werden 10 µg Gesamt-RNA aufgetragen und das Gel für 5 Min bei 50V laufen gelassen. Anschließend wird die Spannung auf 125 Volt hochgeregelt und die Elektrophorese für ca. 1,5 Stunden durchgeführt. Vom dem so gewonnene Gel wird eine photographische Aufnahme unter UV-Exposition gemacht.

II Blotting

Die vorbereitete Membran (Nylon Hybond N(+), Amersham) wird 5 Minuten in 1xTBE-Puffer äquilibriert und der *Electroblotter* (Biorad) nach Instruktionsanweisung aufgebaut. Die Membran und das Gel werden zwischen zwei Lagen Whatman-Papier deponiert und die gesamte Anordnung im Schlitten des *Blotters* gemäss der Anleitung fixiert. Der Transfer der RNA vom Gel auf die Membran mittels *Electroblotting* findet im eisgekühlten und gerührtem 1xTBE Puffer für 60 Minuten bei 50 Volt statt. Anschließend wird die Membran aus der Apparatur genommen und für 30 Minuten bei 120°C inkubiert.

III Detektion

Die generierte RNA-Sonde wird für 10 Minuten bei 95°C inkubiert um mögliche Sekundärstrukturen aufzubrechen und in Prähybridisierungslösung verdünnt. Die Konzentration der RNA-Sonde sollte 50-100 ng/ml betragen, es werden 10-20 ml Sondenlösung für die Detektion eines *Blots* benötigt. Die *geblottete* Membran wird bei 68 °C für 1 Stunde in Prähybridisierungslösung inkubiert. Anschließend wird die überstehende Lösung verworfen und die präparierte Sondenlösung zugefügt. Die Hybridisierung des *Blots* erfolgt über Nacht bei 68°C unter leichtem Schwenken im Hybridisierungssofen. Nach beendeter Hybridisierung wird die Sondenlösung aus dem Hybridisierungsgefäß zurückgewonnen und bei -80°C gelagert. Die Membran wird zweimal für 15 Minuten mit Waschpuffer 1 bei Raumtemperatur und nachfolgend zweimal für gleiche Zeitspanne mit Waschpuffer 4 bei 68°C gewaschen. Nachfolgend wird die kolorimetrische Detektion des *Blots* mit NBT und BCIP durchgeführt. Dazu wird die Membran für 1 Minute in Waschpuffer äquilibriert und für 60 Minuten in Blockierungslösung inkubiert. Die Stammlösung des mit alkalischer Phosphatase markiertem Digoxigenin-Antikörper wird zentrifugiert, 1:5000 mit Blockierungslösung verdünnt (2 µl Anti-DIG-AP in 10 ml Blockierungslösung) und für 30 Minuten mit der Membran inkubiert. Anschließend werden zwei Waschschritte mit dem Waschpuffer für jeweils

15 Minuten durchgeführt. 100 μL BCIP/NBT-Lösung werden in 10 ml Detektionspuffer gelöst und nach Äquilibrierung der Membran für 2 Minuten in Detektionspuffer auf die Membran gegeben. Das Reaktionsgefäß wird lichtdicht verschlossen und der Fortgang der Reaktion fortwährend beobachtet. Bei der benötigten Bandenintensität wird die Reaktion durch Zugabe von Wasser gestoppt, um eine Überentwicklung der Banden zu verhindern. Der so generierte *Northern Blot* wird photographiert und die Bandenintensität kann mittels densitometrischer Analyse quantifiziert werden.

A.1.27 Generierung von RNA mittels *in vitro* Transkription

1 μg des linearisierten und mittels Phenol-Chloroform Extraktion aufgereinigten *Templates* wird mit nukleasefreiem Wasser auf ein Volumen von 22,75 μl aufgefüllt und mit folgenden Substanzen in der angegebenen Reihenfolge versetzt:

- 10 μl 5x Transkriptionspuffer
- 10 μl Ribonukleotid-Mix (rATP, rCTP, rGTP und rUTP; je 2,5 mM)
- 5 μl DTT 100 mM
- 1,25 μl RNase Inhibitor (40 Units/ μl)
- 1 μl T7 oder SP6 RNA-Polymerase(20 Units/ μl)

Es ist darauf zu achten, dass der Transkriptionspuffer, die Ribonukleotide und das DTT Raumtemperatur aufweisen. Das im Transkriptionspuffer enthaltene Spermidin kann andernfalls zur Präzipitation des *Templates* führen. Der Reaktionsansatz wird für 120 Minuten bei 37°C inkubiert und anschließend mit 1 μL DNase I (10 Units/ μl) versetzt und für weitere 30 Minuten bei 37°C inkubiert. Dieser Schritt dient zur quantitativen Eliminierung des DNA-Templates. Der so erhaltene Transkriptionsansatz wird auf 500 μl mit nukleasefreiem Wasser aufgefüllt und mit dem gleichen Volumen Phenol/Chloroform/Isoamylalkohol-Mischung (25:24:1) versetzt. Nachfolgend wird eine Phenol-Chloroform Extraktion durchgeführt und die erhaltene wässrige Phase mit 5 M NaCl in eiskaltem Ethanol über Nacht gefällt. Nach einer Zentrifugation von 20 Minuten bei 4°C und 13.000 rpm wird das erhaltene RNA-Pellet mit 750 μl 75%igem Ethanol gewaschen und erneut für 10 Minuten bei 4°C und 13.000 rpm zentrifugiert. Das RNA-Pellet wird getrocknet und in 25 μl nukleasefreiem Wasser aufgenommen. Nachfolgend wird photometrisch die Reinheit und Konzentration bestimmt. Die Visualisierung des Transkriptes kann auf einen Standard-Agarosegel (1,5%) mit einem denaturierenden Probenpuffer erfolgen. 5 μl der RNA-Lösung werden dabei zu 15 μl RNA-Probenpuffer gegeben und mit 2-5 μl RNA-Ladepuffer versetzt. Die so behandelte Probe wird für 5-10 Minuten bei 65-70°C inkubiert und anschließend auf das Gel aufgetragen.

A.1.28 Affinitätschromatographische Aufreinigung über Vivapure Metall-Chelat *Spin Columns*

Die *Spin Columns* werden mit 400 μl destilliertem Wasser befeuchtet und für 1 Minute bei 5000 g zentrifugiert. Der Durchlauf wird verworfen. 400 μl 0,5 M CoCl_2 werden auf die Säule appliziert und 1 Minute bei 5000 g zentrifugiert. Dieser Schritt wird wiederholt. Anschließend wird die Säule mit 400 μl destilliertem Wasser gewaschen (Zentrifugation jeweils 1 Minute

bei 5000 g). Die Membran wird mit 400 μl Äquilibrationpuffer (50 mM NaH_2PO_4 , 300 mM NaCl, 10 mM Imidazol, pH 8,0) beladen und 1 Minute bei 3000 g zentrifugiert. Der Durchlauf wird verworfen und der Schritt wiederholt. Nachfolgend wird auf die Säule 400 μl der Probenlösung appliziert und 3 Min bei 3000 g zentrifugiert. Nun erfolgen drei Waschrunde mit jeweils 400 μl Waschpuffer (50 mM NaH_2PO_4 , 300 mM NaCl, 20 mM Imidazol, pH 8,0) und einer Zentrifugation von 3 Min bei 3000 g. Zur Elution der Probe werden 200 μl Elutionspuffer (50 mM NaH_2PO_4 , 300 mM NaCl, 250 mM Imidazol, pH 8,0) auf die Säule gegeben und 3 Min bei 1500 g zentrifugiert. Der Elutionsschritt wird wiederholt.

A.1.29 Phenol-Chloroform Aufreinigung

Die zu reinigende Nukleinsäurelösung wird mit dem gleichen Volumen Phenol/Chloroform/Isoamylalkohol-Mischung versetzt und 1 Minute gevortext. Es schließt sich ein Zentrifugationsschritt von 5 Min bei 4°C und 13.000 rpm an. Die obere, wässrige Phase wird abgenommen, in ein neues Reagiergefäß überführt und mit dem gleichen Volumen Chloroform versetzt. Es wird analog zum ersten Extraktionsschritt vorgegangen. Nach Zusatz von einem Zehntel Volumen 5 M NaCl und dem doppelten Volumen eiskaltem, absolutem Ethanol wird die RNA über Nacht bei -80°C inkubiert und so nach Zentrifugation für 20 Minuten bei 13.000 rpm quantitativ gefällt und mit 75%igem Ethanol gewaschen. Durch Inkubation im Thermoblock bei 37°C mit geöffnetem Deckel des Reaktionsgefäßes wurde das RNA-Pellet quantitativ von anhaftenden ethanolischen Resten befreit und getrocknet sowie nachfolgend in nukleasefreiem Wasser resuspendiert.

A.1.30 Zellfraktionierung

2 ml der Zellsuspension wird bei 13.000 g für 3 Minuten zentrifugiert. Der Überstand entspricht der Fraktion Medium und wird abgetrennt. Das Zellpellet wird in 1 ml Lösung 1 (20% Saccharose; 0,1 M EDTA; 0,2 Tris pH 8,0) resuspendiert und 20 min bei Raumtemperatur inkubiert. nachfolgend wird 15 Minuten bei 13.000 g zentrifugiert und der Überstand abgetrennt. Das Zellpellet wird in 1 ml der Lösung 2 (10 mM Tris und 5 mM Magnesiumsulfat, pH 8,0) resuspendiert. Die Suspension wird 10 Minuten auf Eis geschüttelt (700 rpm) und für 10 Minuten bei 5000 g zentrifugiert. Der Überstand entspricht der Fraktion Periplasma. Das Zellpellet wird in 2 ml destilliertem Wasser resuspendiert und durch eine Ultraschallbehandlung aufgeschloßen (Stufe 2, 2x 45 Sec auf Eis). Es schließt sich eine Zentrifugation für 5 Min bei 13.000 g an, der Überstand beinhaltet die Fraktion Cytoplasma.

B Puffer und Lösungen

B.1 Puffer und Lösungen

Puffer und Lösungen

RNA-Probenpuffer

10 ml deionisiertes Wasser
3,5 ml 37%ige Formaldehydlösung
2 ml MOPS-Puffer

5xMOPS-Puffer

0,2 M MOPS(pH 7.0)
50 mM Natriumacetat
5 mM EDTA(pH 8.0)

LB-Medium

10 g NaCl
10 g Trypton
5 g Hefeextrakt
800 ml *ddH₂O*
pH 7,0 einstellen, auf 1 l auffüllen
Autoklavieren 20 Min 121°C

SOC-Medium

20 g Trypton
5 g Hefeextrakt
10 mM NaCl
2,5 mM KCl
10 mM *MgCl₂*
10 mM *MgSO₄*
20 mM Glucose (filtersterilisiert)
800 ml *ddH₂O*
pH 7,0 einstellen, auf 1 l auffüllen
Autoklavieren 20 Min 121°C

TN-Puffer

0,1 M TrisHCl
0,15 M NaCl
800 ml *ddH₂O*
pH 7,5 einstellen, auf 1 l auffüllen.

TNB-Puffer

100 ml TN-Puffer
0,5 g *blocking reagent*
gelöst bei 60°C
sterilfiltriert (220 nm Filter)

TNB-G-Puffer

100 ml TNB-Puffer
10% Ziegenserum

TNT-Puffer

0,1 M TrisHCl
0,15 M NaCl
0,05% Tween 20
800 ml *ddH₂O*
pH 7,5 einstellen, auf 1 l auffüllen.

RNA-Ladepuffer

50% Glycerin
1 mM EDTA
0,4% Bromphenolblau
1 mg/ml Ethidiumbromid

20xSSC

175,3 g NaCl
88,3 g Trinatriumcitrat-Dihydrat
800 ml *ddH₂O*
pH 7,0 einstellen, auf 1 l auffüllen.

SOB-Medium

20 g Trypton
5 g Hefeextrakt
10 mM NaCl
2,5 mM KCl
10 mM *MgCl₂*
10 mM *MgSO₄*
pH 6,7 einstellen (mit KOH), auf 1 l auffüllen
Autoklavieren 20 Min 121°C

CgXIIFS-Medium

15 g/l Ammoniumsulfat
5 g/l Harnstoff
3,92 g/l $K_2HPO_4 \cdot 2H_2O$
4,87 g/l $NaH_2PO_4 \cdot 2H_2O$
0,25 g/l $MgSO_4 \cdot 7H_2O$
pH 7,0 einstellen
Autoklavieren 20 Min 121°C
Zusatzzugabe nach dem Autoklavieren

Zusätze CgXIIFS-Medium

1 ml $CaCl_2$ -Lösung 1% (w/v)
1 ml Biotin-Lösung 1% (w/v)
1 ml Thiamin-Lösung 1% (w/v)
1 ml Natrium-3,4-dihydroxybenzoat 3% (w/v)
40 ml Glucose-Lösung 50% (w/v)
4 ml Chloramphenicol (10 mg/ml)

Spurensalze CgXIIFS-Medium

100 mg $ZnSO_4 \cdot 7H_2O$
1 g $FeSO_4 \cdot 7H_2O$
1 g $MnSO_4 \cdot 7H_2O$
20 mg $CuSO_4$
2 mg $NiCl_2 \cdot 6H_2O$

Waschpuffer für Microarrays

Waschpuffer 1: 2xSSC, 0,1% SDS
Waschpuffer 2: 1xSSC
Waschpuffer 3: 0,5xSSC

2xSDS-Probenpuffer

20 mM Tris-HCl pH 8,0
2 mM EDTA (Na-Salz)
5% SDS
0,02% Bromphenolblau

Coomassie-Färbelösung

1 Coomassie-Tablette
120 ml Methanol
80 ml ddH_2O , auflösen+filtrieren

Trenngel 12%

2 ml Acrylamid/Bisacrylamidmischung (40%, 37,5:1)
1,25 ml Tris-HCL 1,5M pH 8,8
2,15 ml ddH_2O
50 μ l SDS 10%
5 μ l TEMED
50 μ l APS(10%)

Sammelgel 6%

0,25 ml Acrylamid/Bisacrylamidmischung (40%, 37,5:1)
0,26 ml Tris-HCL 1M pH 6,8
1,45 ml ddH_2O
20 μ l SDS (10%)
2 μ l TEMED
20 μ l APS(10%)

Prähybridisierungslösung

5xSSC
50% Formamid (deionisiert)
0,1% Natriumsarkosinat
0,02% SDS
2% *blocking reagent*(Roche)

Waschpuffer für Northern Blots

10 mM Maleinsäure
15 mM NaCl
0,03% Tween 20
pH-Wert auf 7,5 einstellen

Maleinsäurepuffer

10 mM Maleinsäure
15 mM NaCl
pH-Wert auf 7,5 einstellen

Detektionspuffer

10 mM TrisHCl
10 mM NaCl, pH 9,5 einstellen

TE-Puffer

10 mM Tris-HCl
1 mM EDTA
pH-Wert auf 8,0 einstellen

10xStammlösung Blockierungsreagenz

10 g *blocking reagent*
100 ml Maleinsäurepuffer
unter leichtem Erwärmen (40-50°C) lösen
10xStammlösung bei -20°C lagern

C PCR-Primer

Angegeben ist die 5' zu 3' Orientierung:

C.0.1 Absolute Quantifizierung

1) dnaA

Forward: AGGCGGAACTGAGCGATAACAC

Reverse: GTAGTACTCCGCCACCGTCTTCTG

2)ftsH

Forward: CTTTGGGCCCAGCGAGTCTAATG

Reverse: CGCGGCGATAACGATGATACCTT

3)glyA

Forward: CGCACTCCGGCTCCCAGGCTAACT

Reverse: ACCGGGTAACGTGCGCAGATGTCTG

4)lacY

Forward: ACTTTTGGATGTTTCGGTTTATTC

Reverse: ATTGCCCCATTGTCGTTAC

5) lacA

Forward: ATGCCAATGACCGAAAGAATAAGA

Reverse: TACTACCCGCGCCAATAACAGAAT

6) rpoA

Forward: GTTATGTGCCGGCTTCTACCC

Reverse: AGCCAGCACGTCTTTAATCTCAGT

7)secD

Forward: CCCGCGAAGGTCAGCCAGTT

Reverse: GCCGCCATATAGCAGGTTTACGAT

8) sig2

Forward: GCTGCCGCCGTTGAGTGTATCCA
Reverse: TAGCGCTTCCAGTGCCGAGGTATC

9) yidC

Forward: AGGCCCAACAGACCACGCAGACA
Reverse: TCATCGCCCAGACGCTCACG

10) dnaK

Forward: CCGGCTAAACGTCAGGCAGTG
Reverse: CATCGTCATCTTTCGCGTTGTTTG

11) tatD (blunt end)

Forward: ATCGGCGTTAATTTGACCAGTTCG
Reverse: CAAAAGTTCCCGCAGCTCCAGTCC

12) tatD (direktionale Klonierung)

Forward: CAGAGATGCAGTCGACATCGGCGTTAATTTGACCAGTTCG
Reverse: CCAAGCCTTCCCATGGCAAAGTTCCCGCAGCTCCAGTCC

C.0.2 GFP-Klonierung

GFP_{HindIII}: GCGATGAAGCTTCTATTTGTATAGTTCATCCAT

GFP_{SsrA}: GCGATGAAGCTTGCATGCTTAAGCTGCTAAAGCGTAGTTTTTCGTCGTTT-
GCTGCGTCGACTTTGTATAGTTCATCCATGCC

GFP_{HexaHis}: GCGAAGAAGCTTCTAATGATGGTGATGGTGATGTTTGTATAGTTCAT-
CCAT

GFP_{XbaI}: GCGATGTCTAGAAGTAAAGGAGAAGAAGCTTTTCACT

GFP_{SacI}mitRBS: GAGCTCTTAAAGAGGAGAAAGGTCATGAGTAAAGGAGAAGAAGCTTTT-
CACT

tor_{ASacI}: GCGATGTCTAGAAGCGTCAGTCGCCGCTTGCGCCGC

tor_{AXbaI}: GCGATGGAATTCGAGCTCTTAAAGAGGAGAAAGGTCATGAACAATAA-
CGATCTCTTTCAG

13) gst

Forward: AGGCGTAGCGATTATGCAGT
Reverse: GGCGATGTAATTCAGCCATT

Geräte

Scanner Genepix 4000B	Axon Instruments
Vakuumentrifuge Eppendorf Concentrator 5301	Eppendorf AG Wesseling-Berzdorf
Mikrozentrifuge 5415 R	Eppendorf AG Wesseling-Berzdorf
Slide Schüttler Thermomixer comfort	Eppendorf AG Wesseling-Berzdorf
<i>Arrayer417TM</i>	Affymetrix
<i>CriterionTM</i> Blotter	Bio Rad München
Hybridisierungssofen Thermo Hybaid	Thermo Scientific
IQ5 real-time PCR-Cycler	Biorad München
Bioanalyzer 2100	Agilent Technologies GmbH Deutschland, Böblingen
Transferpette	Brandt GmbH+Co KG, Wertheim
Schüttler MTS 4	IKA Werke GmbH, Staufen
Thermo EC 105 Gelelektrophorese	Thermo Scientific Waltham USA
<i>MiniProteanTM</i> Tetra Cell	Biorad München
Schüttelschrank Certomat BS1	B. Braun Melsungen
Bioreaktor Biostat C	B. Braun Melsungen
Plattenschweissgerät <i>ALPSTM</i> 50V	Thermo Scientific Waltham USA
Spektralphotometer Nanodrop ND1000	PEQ-Lab Erlangen

Software

Gene Pix Pro 6.0	Axon Instruments Segundo USA
Lasergene 7.0	DNA-Star Madison USA

Verbrauchsmaterialien

Mikrotiterplatte	Sarstedt AG, Nürnberg
Mikrotiterplatte f. <i>real-time</i> PCR	Biorad, München
Adhäsive optische Folie f. Mikrotiterplatten	Biorad, München
Reagiergefäße 2ml	Sarstedt AG, Nürnberg
Reagiergefäße 1,5ml	Sarstedt AG, Nürnberg
Reagiergefäße 0,1ml	Sarstedt AG, Nürnberg
Spritzenvorsatzfilter (0,22µm)	Sartorius Stedim, Göttingen
<i>falcon tubes</i> 50ml	Sarstedt AG, Nürnberg
<i>falcon tubes</i> 15ml	Sarstedt AG, Nürnberg
Aledehyd-funktionalisierte Chips (<i>VSS Silylated Slides</i>)	Cell Associates Inc., USA

Chemikalien

Random Hexamer Primer 100mM	Fermentas, St. Leon-Rot
SacI 10u/ μ l	Fermentas, St. Leon-Rot
SalI 10u/ μ l	Fermentas, St. Leon-Rot
NcoI 10u/ μ l	Fermentas, St. Leon-Rot
HindIII 10u/ μ l	Fermentas, St. Leon-Rot
XbaI 10u/ μ l	Fermentas, St. Leon-Rot
EcoRV 10u/ μ l	Fermentas, St. Leon-Rot
NdeI 10u/ μ l	Fermentas, St. Leon-Rot
ATP 100mM	Fermentas, St. Leon-Rot
<i>Pfu</i> DNA-Polymerase	Fermentas, St. Leon-Rot
Omni- <i>Taq</i> DNA-Polymerase	OLS GmbH, Hamburg
T7 RNA-Polymerase	Promega, Mannheim
SP6 RNA-Polymerase	Promega, Mannheim
rNTPS 10mM	Promega, Mannheim
dNTPs	Fermentas, St. Leon-Rot
RNasin RNase-Inhibitor	Promega, Mannheim
5xTranskriptionspuffer	Promega, Mannheim
DTT 100mM	Promega, Mannheim
DNaseI 10U/ μ l	Roche Diagnostics GmbH, Mannheim
Alkalische Phosphatase 1U/ μ l	Promega, Mannheim
BPC-Wasser (endotoxinfrei)	Sigma-Aldrich Chemie GmbH, Steinheim
DIG-11-dUTP 1mM	Roche Diagnostics GmbH, Mannheim
DIG-rNTP Labeling-Mix	Roche Diagnostics GmbH, Mannheim
BCIP/NBT-Mix	Roche Diagnostics GmbH, Mannheim
DIG Kontroll-DNA	Roche Diagnostics GmbH, Mannheim
pGEM5Zf(+) Vektor-DNA	Promega, Mannheim
PEG 4000	Fermentas, St. Leon-Rot
T4-DNA-Ligase 5u/ μ l	Fermentas, St. Leon-Rot
Superscript III Reverse Transkriptase	Invitrogen GmbH Karlsruhe
5xFirst Strand Buffer	Invitrogen GmbH Karlsruhe
Cy3/Cy5-dCTP 1mM	Perkin-Elmer Life Sciences USA
Top Block	Sigma-Aldrich Chemie GmbH, Steinheim
<i>PCR Purification Kit</i>	Qiagen, Hilden
<i>TSATM Labeling and Detection Kit</i>	Perkin-Elmer Life Sciences USA
<i>Wizard SV Gel and PCR Purification Kit</i>	Promega, Mannheim
Agarose NEEO Ultra Qualität	Carl Roth GmbH, Karlsruhe
DEPC	Carl Roth GmbH, Karlsruhe
Primer	MWG Biotech AG, Ebersberg
L-Arabinose	Fluka Chemie AG, Seelze
X-Gal	Appllichem GmbH Darmstadt
IPTG	Appllichem GmbH Darmstadt
<i>RotiphenolTM</i>	Carl Roth GmbH, Karlsruhe
Chloroform	Sigma-Aldrich Chemie GmbH, Steinheim

Chemikalien (Fortsetzung)

Improm Reverse Transkriptase 5u/ μ l	Promega, Mannheim
<i>Total RNA Isolation Kit</i>	Analytik Jena GmbH, Jena
IQ SybrGreen Mastermix	Biorad GmbH, München
Plasmid gold Mini Kit	PeqLab GmbH, Erlangen
Trypton Plus	Fluka Chemie AG, Seelze
Natriumchlorid	Fluka Chemie AG, Seelze
Hefeextrakt	Fluka Chemie AG, Seelze
Ethanol, absolut	Carl Roth GmbH, Karlsruhe
Trinatriumcitrat-Dihydrat	Fluka Chemie AG, Seelze
Tween 20	Sigma-Aldrich Chemie GmbH, Steinheim
DMSO	Sigma-Aldrich Chemie GmbH, Steinheim
Ziegenserum	Spenderin: Fabienne Anton
Kaliumchlorid	Sigma-Aldrich Chemie GmbH, Steinheim
Magnesiumchlorid Dihydrat	Sigma-Aldrich Chemie GmbH, Steinheim
Magesiumsulfat	Sigma-Aldrich Chemie GmbH, Steinheim
HEPES	Sigma-Aldrich Chemie GmbH, Steinheim
Manganchlorid	Sigma-Aldrich Chemie GmbH, Steinheim
Kaliumhydroxid	Fluka Chemie AG, Seelze
SDS	Sigma-Aldrich Chemie GmbH, Steinheim
Natriumborhydrid	Sigma-Aldrich Chemie GmbH, Steinheim
BSA	Sigma-Aldrich Chemie GmbH, Steinheim
MOPS	Sigma-Aldrich Chemie GmbH, Steinheim
Formaldehyd 37%ig	Sigma-Aldrich Chemie GmbH, Steinheim
Formamid	Sigma-Aldrich Chemie GmbH, Steinheim
N,N-Dimethylformamid	Sigma-Aldrich Chemie GmbH, Steinheim
EDTA 0,5M pH8,0	Sigma-Aldrich Chemie GmbH, Steinheim
Natruimacetat	Carl Roth GmbH, Karlsruhe
Natriumborat	Carl Roth GmbH, Karlsruhe
Tris-HCl	Carl Roth GmbH, Karlsruhe
Tris-Base	Carl Roth GmbH, Karlsruhe
Betain	Carl Roth GmbH, Karlsruhe
Agar	Carl Roth GmbH, Karlsruhe
<i>DIG Nucleic Acid Detection Kit</i>	Roche Diagnostics GmbH, Mannheim
<i>DIG Wash and Block buffer set</i>	Roche Diagnostics GmbH, Mannheim
Bromphenolblau	Fluka Chemie AG, Seelze
Ethidiumbromid	Fluka Chemie AG, Seelze
Sarkosinat	Fluka Chemie AG, Seelze
Trichloressigsäure	Fluka Chemie AG, Seelze
Natriumdihydrogenphosphat	Fluka Chemie AG, Seelze
Imidazol	Fluka Chemie AG, Seelze

D Mathematische Formeln und Statistik

D.1 Berechnung von Inkorporationsfrequenzen

Für die Berechnung der Inkorporationsfrequenzen von Cy3- bzw. Cy5-dNTPs in die cDNA wurde folgende Formel verwendet:

a) Cy3-Inkorporation: $58,5 \cdot \frac{A_{550}}{A_{260}}$

b) Cy5-Inkorporation: $35,1 \cdot \frac{A_{650}}{A_{260}}$

mit A_{260} = Absorption bei 260 nm; A_{550} = Absorption bei 550 nm und A_{650} = Absorption bei 650 nm.

Die Faktoren 58,5 und 35,1 leiten sich aus den Absorptionskoeffizienten von Cy3/Cy5, dem mittleren molaren Gewicht der dNTPs, der Konzentration der *Cy-gelabelten* ssDNA die 1 Absorptionseinheit von elektromagnetischer Strahlung der Wellenlänge 260 nm absorbiert her. Das Ergebnis wird in der Anzahl von inkorporierten Cy-dCTP pro 1000 Nukleotiden angegeben.

D.2 Statistik

Für Ergebnisse aus mehreren Einzelexperimenten wurden Mittelwert und Standardfehler berechnet. Um Unterschiede zwischen einzelnen Versuchsgruppen statistisch zu verifizieren, wurden die Daten auf Normalverteilung und gleiche Varianzen geprüft und im Anschluss ein studentischer t-Test durchgeführt. Die Unterschiede zwischen den Gruppen wurden als statistisch signifikant gewertet ab $p < 0,05$. Waren die Proben nicht normalverteilt oder wiesen ungleiche Varianzen auf, so wurde ein Mann-Whitney-U-Test durchgeführt. Bei mehreren voneinander unabhängigen Eingangsgrößen (x_1, x_2, \dots) mit den zugehörigen Fehlern ($\Delta x_1, \Delta x_2, \dots$) wurde der Fehler des aus den Mittelwerten der Eingangsgrößen errechneten Ergebnisses (Δy) anhand der Fehlerfortpflanzung nach Gauß berechnet:

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial x_1} \cdot \Delta x_1\right)^2 + \left(\frac{\partial y}{\partial x_2} \cdot \Delta x_2\right)^2}$$

Lebenslauf

Persönliche Daten

Name	Michael Reck
Geburtsdatum	7.5.1978
Nationalität	deutsch

Schulbildung

1985-1989	Grundschule Am Castrum, Gehrden
1989-1991	Orientierungsstufe, Gehrden
1991-1998	Matthias-Claudius Gymnasium, Gehrden

Zivildienst

1998-1999	Robert Koch Krankenhaus, Gehrden
-----------	----------------------------------

Studium

1999-2005	Studium der Chemie, Universität Hannover
2005-2008	Doktorand im Institut für Technische Chemie, Universität Hannover

Publikationsliste

Optimisation of a Microarray Sandwich-ELISA against hINF- γ on a modified Nitrocellulose Membrane.

M. Reck, F. Stahl, J.G. Walter, M. Hollas, D. Melzner und T. Scheper.
Biotechnol Prog. 23(6), 1498-1505, 2007.

Genomic response analysis of *Escherichia coli* during coexpression of a bacteriocin release protein using DNA microarray technology.

M. Reck, C. Repenning, F. Stahl, B. Sommer, K. Friehs, E. Flaschel und T. Scheper.
Zur Publikation eingereicht bei BMC Genomics.

Aufgaben und Verfahren der Bioinformatik für die Microarray-Technologie.

C. Repenning, **M. Reck**, F. Stahl, T. Scheper, B. Hitzmann.
Zur Publikation angenommen beim Biospektrum.

Extracellular Production and Affinity Purification of Recombinant Proteins with *Escherichia coli* using the Versatility of the Maltose Binding Protein.

B. Sommer, K. Friehs, E. Flaschel, **M. Reck**, F. Stahl, T. Scheper.

Zur Publikation angenommen beim Journal of Biotechnology, elektronische Version bereits vor dem Druck publiziert.