

**Herstellung einer Biosynthesemutante von *Amycolatopsis  
balhimycina* durch Modulinsertion in die nicht-ribosomale  
Peptidsynthetase BpsB mittels eines kombinatorisch-  
biosynthetischen Ansatzes und deren massenspektrometrische  
Charakterisierung**

vorgelegt von  
Diplom-Biochemikerin  
Diane Butz  
aus Tübingen

Von der Fakultät II – Mathematik und Naturwissenschaften  
der Technischen Universität Berlin  
zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften  
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. rer. nat. Reinhard Schomäcker  
Erster Bericht: Prof. Dr. rer. nat. Roderich D. Süßmuth  
Zweiter Bericht: PD Dr. rer. nat. Joachim Vater

Tag der wissenschaftlichen Aussprache: 18. April 2008

Berlin 2008

D 83



## Zusammenfassung

### **Herstellung einer Biosynthesemutante von *Amycolatopsis balhimycina* durch Modulinsertion in die nicht-ribosomale Peptidsynthetase BpsB mittels eines kombinatorisch-biosynthetischen Ansatzes und deren massenspektrometrische Charakterisierung**

Die vorliegende Arbeit beschäftigt sich mit der kombinatorischen Biosynthese bei nicht-ribosomalen Peptidsynthetasen (NRPS), die aufgrund der Komplexität der Strukturen von nicht-ribosomal synthetisierten Peptiden eine ernstzunehmende Alternative zur chemischen Synthese darstellt, um Strukturderivate zu erzeugen. Ziel dieser Arbeit war die Komplettierung der bisher bei NRPS literaturbekannten Modulmanipulationen „Modulaustausch“ und „Moduldeletion“ durch das noch ausstehende Experiment der „Modulinsertion“. Als Modellverbindung diente das von *A. balhimycina* synthetisierte Vancomycin-Typ Glykopeptidantibiotikum Balhimycin, welches durch sein stark modifiziertes Peptidgrundgerüst (Chlorierung, oxidative Seitenkettenzyklisierung, *N*-Methylierung und Glykosylierung) ein sehr komplexes Beispiel einer nicht-ribosomalen Peptidsynthese darstellt. In der vorliegenden Arbeit ist es gelungen, ein zusätzliches Modul in die NRPS BpsB von *A. balhimycina* über eine *double crossover*-Strategie einzuführen, was sowohl auf genetischer Ebene mittels PCR und *Southern Blot*-Hybridisierung als auch auf Proteinebene mittels Disk-SDS-PAGE und anschließende massenspektrometrische Identifizierung der Gelbande nachgewiesen werden konnte. Daran schloss sich eine eingehende Charakterisierung des Produktionsprofils dieser kombinatorischen Biosynthesemutante mittels HPLC-ESI-MS und -MS/MS an. Diese Analyse erbrachte den Beweis, dass das eingeführte Modul auch enzymatische Aktivität besitzt und somit das Peptidrückgrat des Balhimycins um eine Aminosäure erweitert. Neben dem Beweis der Machbarkeit einer Modulinsertion bei NRPS gewährte dieses kombinatorische Biosynthese-Experiment auch Einblicke in die Flexibilität des gesamten Biosyntheseapparates. Beispielsweise war die Chlorierung der Produkte durch die Halogenase BhaA nicht beeinträchtigt, während *N*-Methylierung und Glykosylierung in allen 17 nachgewiesenen Biosyntheseprodukten fehlten. Bezüglich der oxidativen Seitenkettenzyklisierungen brachte die erste P450-abhängige Monooxygenase OxyB noch eine gewisse Toleranz für das neue Substrat bzw. die veränderte NRPS auf, während OxyA und OxyC keine nachweisbare Aktivität mehr zeigten.



**Die Ergebnisse der vorliegenden Arbeit wurden veröffentlicht:**

Butz, D., Schmiederer, T., Hadatsch, B., Wohlleben, W., Weber, T., Süssmuth, R.D. (2008). Module extension of a non-ribosomal peptide synthetase of the glycopeptide antibiotic balhimycin produced by *Amycolatopsis balhimycina*. ChemBioChem, im Druck.



*Für Timo, Christel, Wolfgang und Jasmin*

## **Danksagung:**

Herrn Prof. Dr. Roderich D. Süßmuth danke ich ganz herzlich für die hervorragenden Arbeitsbedingungen und seine Unterstützung in jeglicher Hinsicht. Besonders danken möchte ich ihm für die Überlassung meines interessanten Promotionsthemas und die Möglichkeit, an dem EU-Projekt „Combig-Top“ teilnehmen zu dürfen.

Vielen Dank an Herrn PD Dr. Joachim Vater und Herrn Prof. Dr. Reinhard Schomäcker für die Übernahme der zweiten Berichterstattung bzw. des Prüfungsvorsitzes.

Großen Dank schulde ich Herrn Prof. Dr. Wolfgang Wohlleben und seiner Arbeitsgruppe für die fruchtbare Kooperation und die Möglichkeit, in seinen Laborräumen arbeiten zu dürfen. Besonders danken möchte ich hierbei Dr. Tilmann Weber für die exzellente Zusammenarbeit bei der Ausarbeitung der Klonierstrategie und Dr. Bianka Hadatsch, die beste Kooperationspartnerin, die ich je hatte. Weiterhin bedanken möchte ich mich bei Dr. Evi Stegmann und Sigrid Stockert für die Unterstützung v.a. bei der Fermentation.

Herzlichen Dank an Dipl.-Ing. Graeme Nicholson für die Durchführung der Aminosäureanalytik und der FT-ICR-MS-Messungen.

Bei Herrn PD Dr. Ullrich Keller möchte ich mich für sein Interesse und viele fruchtbare Diskussionen über unterschiedlichste Themen bedanken.

Danke an Herrn PD Dr. Rainer Zocher für die Überlassung der Enniatinsynthetase als Größenmarker bei der Disk-SDS-PAGE.

Herzlichen Dank an den gesamten Arbeitskreis Süßmuth für die tolle Arbeitsatmosphäre: Louise Aigrain, Dr. Christian Appelt, Mahsuni Atli, Frank Dettner, Dr. Sven Feifel, Elvira Gottardi, Anne Hänchen, Soliman Helali, Maik Henkel, Dr. Simone Keller, Joanna Krawczyk, Wolfgang Müller, Jonny Nachtigall, Stefan Pohle, Marta Ramirez Hernández, Dr. Andreas Reicke, Pierre Loic Saaidi, Georg Sambeth,

Heiko Schadt, Caroline Schlosser, Timo Schmiederer, Dr. Kathrin Schneider, Hanna von Suchodoletz, Dr. Suada Turkanović und Falko Wolter.

Vielen Dank an das Oetkerbau-Umzugsteam Simone, Kathrin und Timo für unvergessliche Erlebnisse. Vielen Dank an Dr. Andreas Reicke für die Fermentation in Tübingen und an Dr. Suada Turkanović für die Synthese des Oktapeptids. Besonders danken möchte ich Kathrin und Falko für ihre Hilfsbereitschaft und ihre Freundschaft weit über die Arbeit hinaus!

Meinen Eltern und meiner Schwester Jasmin danke ich für ihre grenzenlose Unterstützung und ihr Verständnis. Danke, dass ihr immer für mich da seid!

Timo, vielen Dank für alles...aber vor allem für deine unendliche Geduld mit mir!

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	<i>Antibakterielle Wirkstoffe</i>	1
1.2	<i>Die nicht-ribosomale Peptidsynthese</i>	3
1.3	<i>Glykopeptide</i>	6
1.3.1	Struktur der Glykopeptide	6
1.3.2	Biosynthese der Glykopeptide	9
1.3.3	Antimikrobielle Aktivität der Glykopeptide und auftretende Resistenz-mechanismen	14
1.4	<i>Kombinatorische Biosynthese bei nicht-ribosomalen Peptidsynthetasen</i>	18
<b>2</b>	<b>Zielsetzung</b>	<b>22</b>
<b>3</b>	<b>Grundlagen</b>	<b>23</b>
3.1	<i>Chromatographie</i>	23
3.1.1	Reverse-Phase Chromatographie	24
3.1.2	Hochdruckflüssigchromatographie (HPLC)	25
3.1.3	Kapillar-Hochdruckflüssigchromatographie (Kapillar-HPLC)	26
3.2	<i>Massenspektrometrie (MS)</i>	28
3.2.1	ESI-Ionenquelle	29
3.2.2	Triple-Quadrupol-Analysator	30
3.2.3	QTrap-Analysator	36
3.3	<i>Kapillar-HPLC-ESI-QTrap-Kopplung</i>	37
<b>4</b>	<b>Material und Methoden</b>	<b>38</b>
4.1	<i>Material</i>	38
4.1.1	Geräte	38
4.1.2	Chemikalien	39
4.1.3	Bakterienstämme	41
4.1.4	Kits	42
4.1.5	Enzyme	42
4.1.6	Oligonukleotide	42
4.1.7	Medien	43
4.1.8	Antibiotika und andere Medienzusätze	44
4.1.9	Lösungen und Puffer	45
4.1.10	Sonstiges	51
4.2	<i>Methoden</i>	52
4.2.1	Kultivierung und Stammhaltung von <i>E. coli</i>	52
4.2.2	Kultivierung und Stammhaltung von <i>A. balhimycina</i>	52
4.2.3	Polymerase-Kettenreaktion (PCR)	53
4.2.4	Photometrische Konzentrationsbestimmung von DNA	57
4.2.5	Gelelektrophorese von DNA	58
4.2.6	Klonierung von DNA-Fragmenten	59
4.2.7	DNA-Transfer in Bakterien	64
4.2.8	Stressprotokoll zur Forcierung von homologen Rekombinationsereignissen in <i>A. balhimycina</i>	67
4.2.9	Plasmidisolierung aus <i>E. coli</i>	68
4.2.10	Isolierung von Gesamt-DNA aus <i>A. balhimycina</i>	69
4.2.11	<i>Southern Blot</i> -Hybridisierung	71
4.2.12	Isolierung von nicht-ribosomalen Peptidsynthetasen aus <i>A. balhimycina</i>	74
4.2.13	Photometrische Konzentrationsbestimmung von Proteinen nach Bradford	76
4.2.14	Gelelektrophorese von Proteinen	79
4.2.15	Färbung von Proteingelen	80
4.2.16	Massenspektrometrische Analyse von Protein-Banden aus dem Gel	81

4.2.17	Vorbereitung der Fermentationskulturen der <i>A. balhimycina</i> C5A5T4E4-Mutante für massenspektrometrische Untersuchungen	82
4.2.18	Identifizierung der Glykopeptidmetabolite der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels HPLC-ESI-MS	87
4.2.19	Strukturaufklärung der identifizierten Glykopeptidmetabolite der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels HPLC-ESI-MS/MS	90
4.2.20	Summenformelbestätigung der Glykopeptidmetabolite der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels hochaufgelöster ESI-FTICR-MS	92
4.2.21	Chirale Aminosäureanalytik	92
<b>5</b>	<b>Ergebnisse</b>	<b>93</b>
5.1	<i>Ausarbeitung und Durchführung eines molekularbiologischen Konzepts zur Modulinsertion bei nicht-ribosomalen Peptidsynthetasen (NRPS)</i>	93
5.1.1	Auswahl des neu einzufügenden Moduls einer NRPS	93
5.1.2	Klonierstrategie zur Durchführung der ausgewählten Modulinsertion	96
5.2	<i>Molekularbiologische Charakterisierung der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels PCR und Southern Blot-Hybridisierung</i>	114
5.3	<i>Proteinbiochemische Charakterisierung der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels Disk-SDS-PAGE und massenspektrometrischer Analyse</i>	118
5.4	<i>Charakterisierung des Glykopeptidproduktionsprofils der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels HPLC-ESI-MS und -MS/MS</i>	122
5.5	<i>Ergänzende Untersuchungen zum Glykopeptidproduktionsprofil der <i>A. balhimycina</i> C5A5T4E4-Mutante</i>	130
5.5.1	Chirale Aminosäureanalytik des Glykopeptidmetaboliten 6	130
5.5.2	Summenformelbestätigung der Glykopeptidmetabolite 1 – 11 der <i>A. balhimycina</i> C5A5T4E4-Mutante mittels hochaufgelöster ESI-FTICR-MS	131
5.5.3	Abschätzung der relativen Produktionsmengen der Glykopeptide	132
<b>6</b>	<b>Diskussion</b>	<b>134</b>
<b>7</b>	<b>Ausblick</b>	<b>139</b>
<b>8</b>	<b>Anhang</b>	<b>141</b>
8.1	<i>Ergebnisse der massenspektrometrischen Analyse der Proteinbande aus <i>A. balhimycina</i> und deren Identifizierung als nicht-ribosomale Peptid-synthetase BpsB mittels Proteindatenbanksuche</i>	141
8.2	<i>Ergebnisse der massenspektrometrischen Analyse der Proteinbande aus <i>A. balhimycina</i> C5A5T4E4-Mutante und deren Identifizierung als nicht-ribosomale Peptidsynthetase BpsB mittels Proteindatenbanksuche</i>	143
<b>9</b>	<b>Literaturverzeichnis</b>	<b>145</b>

---

## Abkürzungen

AcN	Acetonitril
AL	alkalische Lyse
AP	alkalische Phosphatase
APS	Ammoniumperoxodisulfat
$\beta$ -Hty	$\beta$ -Hydroxytyrosin
BSA	<i>bovine serum albumin</i>
CAD	<i>collisionally activated dissociation</i>
CDA	<i>calcium-dependent antibiotic</i>
CE	<i>collision energy</i>
CES	<i>collision energy spread</i>
CSPD	Dinatrium 3-(4-methoxyspiro{1,2-dioxethan-3,2'-(5'-chlor) tricyclo [3.3.1.1] decan} -4-yl) phenyl phosphat
CUR	<i>Curtain-Gas</i>
Da	Dalton
DIG	Digoxigenin
DMF	Dimethylformamid
DMSO	Dimethylsulfoxid
DP	<i>declustering potential</i>
dNTP	Desoxyribonukleosidtriphosphat
DNA	Desoxyribonukleinsäure
Dpg	3,5-Dihydroxyphenylglycin
DTT	1,4-Dithiothreitol
EDTA	Ethylendiamintetraessigsäure
EMS	<i>Enhanced MS-Analyse</i>
EP	<i>entrance potential</i>

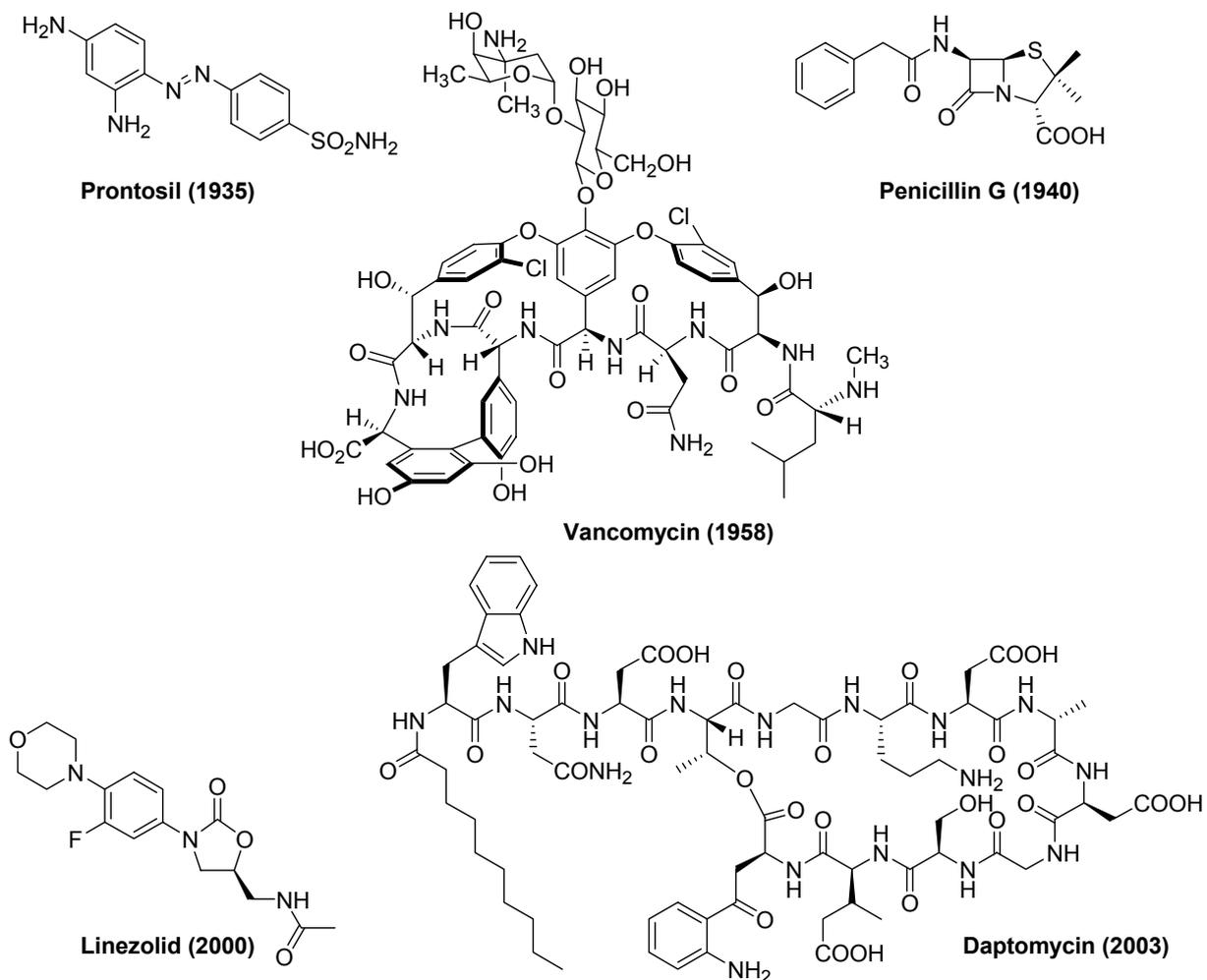
---

ESI	Elektrospray-Ionisation
FTICR	Fourier-Transform-Ionen-Zyklotron
Hpg	4-Hydroxyphenylglycin
HPLC	Hochdruck-/Hochleistungsflüssigchromatographie
IPTG	Isopropyl- $\beta$ -D-thiogalaktopyranosid
IS	<i>ion spray voltage</i>
LIT	<i>linear ion trap</i>
<i>m/z</i>	Molekülmasse-zu-Ladungs-Verhältnis
MeOH	Methanol
MS	Massenspektrometer / Massenspektrometrie
NRPS	nicht-ribosomale Peptidsynthetase(n)
OD	optische Dichte
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase-Kettenreaktion
PEG	Polyethylenglykol
PMSF	Phenylmethylsulfonylfluorid
ppm	<i>parts per million</i>
SDS	<i>sodium dodecyl sulfate</i> / Natriumdodecylsulfat
SSC	<i>sodium chloride sodium citrate</i> / Natriumchlorid-Citrat
TEM	<i>temperature of turbo gas</i>
TEMED	Tetramethylethylendiamin
TES	<i>N</i> -[Tris(hydroxymethyl)-methyl]-2-aminoethansulfonsäure
TFA	Trifluoressigsäure
Tris	Tris-(hydroxymethyl)-aminomethan
w/v	<i>weight/volume</i>
X-Gal	5-Brom-4-chlor-3-indoxyl- $\beta$ -D-galaktopyranosid

# 1 Einleitung

## 1.1 Antibakterielle Wirkstoffe

Eine nicht zu unterschätzende Bedrohung für die menschliche Gesundheit stellen seit jeher bakterielle Infektionen dar. Genauso alt sind deshalb auch die Versuche von Menschen, dieser Bedrohung entgegenzuwirken. Seit Beginn der Industrialisierung und damit der Entwicklung einer pharmazeutischen Industrie wird die Therapie bakterieller Infektionen professionalisiert, indem antibakterielle Wirkstoffe erforscht, entwickelt, produziert und schließlich vermarktet werden (s. Abbildung 1-1).



**Abbildung 1-1: Fünf Beispiele für wichtige antibakterielle Wirkstoffe. Das Jahr der Markteinführung steht jeweils in Klammern hinter der Wirkstoffbezeichnung.**

Die Bedeutung und Wertschätzung der beiden ersten antibakteriellen Wirkstoffe Prontosil, ein Sulfonamid, und Penicillin G,<sup>[1,2]</sup> ein  $\beta$ -Lactam, zeigte sich daran, dass

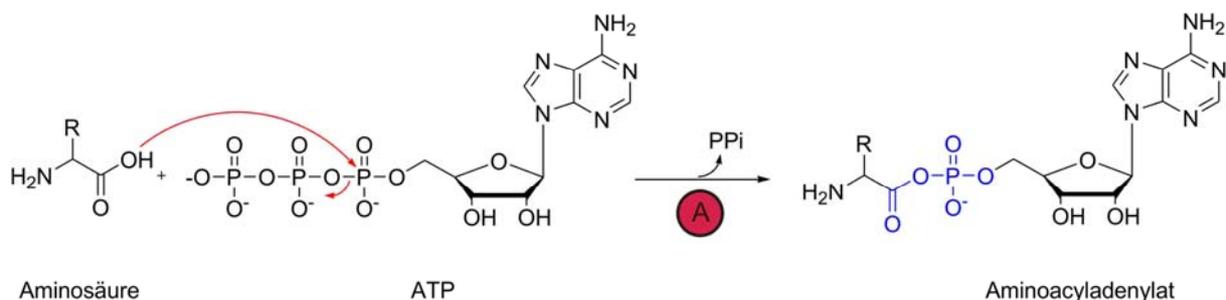
ihren Entdeckern Gerhard Domagk und Alexander Fleming (zusammen mit Ernst B. Chain und Howard W. Florey) 1939 bzw. 1945 der Nobelpreis für Medizin verliehen wurde. Die Folgejahre waren geprägt durch die Einführung weiterer Wirkstoffe bzw. Wirkstoffklassen. Hierunter fällt auch die Einführung der in der vorliegenden Arbeit thematisierten Verbindungsklasse der Glykopeptide mit dem Vancomycin<sup>[3]</sup> im Jahre 1958. Mit diesem Arsenal an antibakteriellen Wirkstoffen fühlte man sich so sicher, dass sich die Pharmaindustrie mehr und mehr den bakteriellen Infektionen ab- und stattdessen anderen Krankheitsbildern zuwandte, vor allem den sehr viel profitableren chronischen Erkrankungen. Aus heutiger Sicht ist diese Entscheidung fragwürdig, da fortschreitende Resistenzentwicklungen humanpathogener Mikroorganismen viele Wirkstoffe der Vergangenheit unwirksam gemacht haben.<sup>[4]</sup> Diese Problematik macht einen kontinuierlichen Nachschub an neuen wirksamen Verbindungen unerlässlich. Momentan befindet man sich deshalb in einer Phase des Umdenkens, die sich nach vier Jahrzehnten Stillstand in den Jahren 2000 und 2003 durch die Einführung der zwei neuen Wirkstoffe bzw. Wirkstoffklassen Linezolid,<sup>[5]</sup> ein Oxazolidinon, und Daptomycin,<sup>[6,7]</sup> ein Lipopeptid bemerkbar macht. Trotz allem sind heute Infektionskrankheiten die zweithäufigste Todesursache weltweit, wovon in den USA die meisten infektionsbedingten Todesfälle auf Bakterien zurückzuführen sind.<sup>[8]</sup> Aus diesem Grund ist und bleibt die einzige Möglichkeit zur Resistenzbekämpfung die ständige Entwicklung neuer antimikrobieller Wirkstoffe. Als sehr ergiebige Quelle haben sich hierbei in der Vergangenheit die Naturstoffe erwiesen, die sich durch eine scheinbar unbegrenzte strukturelle Vielfalt auszeichnen.<sup>[8]</sup> Im engeren Sinne ist der Begriff Naturstoff mit den sog. Sekundärmetaboliten gleichzusetzen, die nicht-essentiellen Biosynthesewegen von Mikroorganismen entstammen. Eine große Subklasse der Sekundärmetabolite stellen die sog. nicht-ribosomal synthetisierten Peptide dar, zu denen auch die in der vorliegenden Arbeit thematisierten Glykopeptide gehören. Im Folgenden soll sowohl die nicht-ribosomale Peptidsynthese im Allgemeinen als auch die Verbindungsklasse der Glykopeptide im Speziellen näher beleuchtet werden.

## 1.2 Die nicht-ribosomale Peptidsynthese

Die nicht-ribosomale Peptidsynthese ist eine alternative Möglichkeit von Lebewesen, Peptidbindungen zu knüpfen. Während die ribosomale Synthese vorwiegend große, ausschließlich aus proteinogenen *L*-Aminosäuren bestehende Proteine herstellt, erweitert die nicht-ribosomale Synthese das Spektrum auf eher kleine, häufig nicht-proteinogene Aminosäuren und Aminosäuren der *D*-Konfiguration enthaltende Peptide. Die Entdeckung der nicht-ribosomalen und damit Nukleinsäure-unabhängigen Peptidsynthese erfolgte in den 1970er Jahren u.a. durch Fritz A. Lipmann, wobei diese Thematik seither eingehend beforscht wird. Im Folgenden soll ein kurzer Überblick über die Abläufe und Gesetzmäßigkeiten der nicht-ribosomalen Peptidsynthese gegeben werden, wobei darüber hinaus auf die in der Literatur erschienenen Übersichtsartikel verwiesen wird.<sup>[9,10,11]</sup>

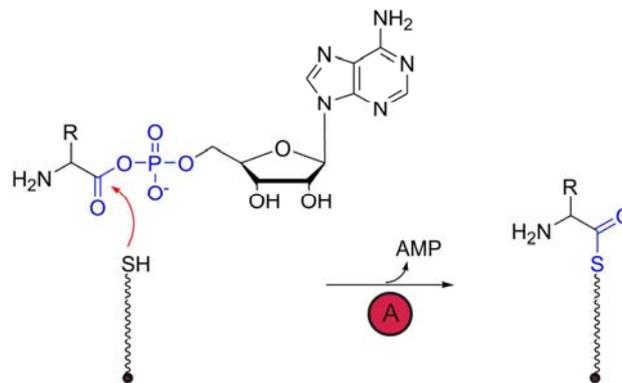
Die für die nicht-ribosomale Peptidsynthese verantwortlichen Enzyme sind die sog. nicht-ribosomalen Peptidsynthetasen (NRPS). Hierbei handelt es sich um meist sehr große, modular aufgebaute Proteine, die in vielen Fällen in einem Multienzym-Komplex zusammenarbeiten. Jedes dieser Module ist für den Einbau einer Aminosäure in das Peptid zuständig, was die zentralen Aufgaben Erkennung und Aktivierung der Aminosäure, Knüpfen der Peptidbindung, aber auch Bindung von Substraten und Biosyntheseintermediaten sowie demgegenüber die Freigabe des fertigen Produkts umfasst. Aus diesem Grund ist ein Modul aus verschiedenen Domänen aufgebaut, die die unterschiedlichen Aufgaben übernehmen.<sup>[12]</sup>

Die Adenylierungs- bzw. A-Domäne ist für die spezifische Erkennung<sup>[13,14,15]</sup> und Aktivierung der entsprechenden Aminosäure mit ATP zu Aminoacyladenylat zuständig (s. Abbildung 1-2).



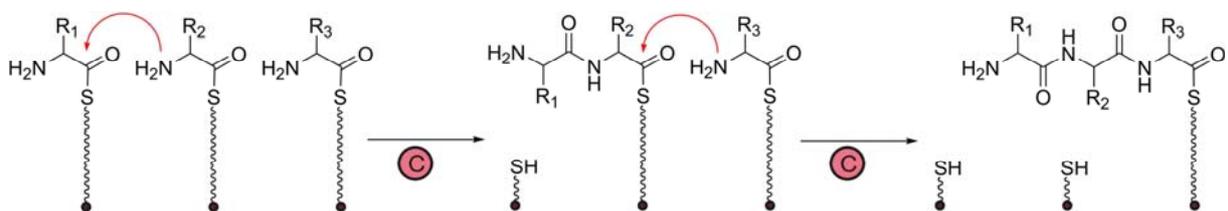
**Abbildung 1-2: Aktivierung einer Aminosäure mit ATP zu Aminoacyladenylat. Dabei handelt es sich um ein gemischtes Säureanhydrid (blau), was einer energiereichen Bindung entspricht. Diese Reaktion wird bei der nicht-ribosomalen Peptidsynthese durch die A-Domäne katalysiert.**

Bei der dadurch gebildeten gemischten Säureanhydridbindung handelt es sich um eine energiereiche Bindung, die die Übertragung der Aminosäure auf die Thiolierungs- bzw. T-Domäne erlaubt. Die T-Domäne trägt einen 4'-Phosphopantetheinyl-Arm als posttranslationale Modifikation, der die Aminosäure in Form eines energiereichen Thioesters binden kann. Diese Übertragung wird ebenfalls durch die A-Domäne katalysiert (s. Abbildung 1-3).



**Abbildung 1-3: Bindung einer Aminosäure an den 4'-Phosphopantetheinyl-Arm der T-Domäne. Dabei wird die energiereiche gemischte Säureanhydridbindung (blau, links) durch eine energiereiche Thioesterbindung ersetzt (blau, rechts). Diese Reaktion wird ebenfalls durch die A-Domäne katalysiert.**

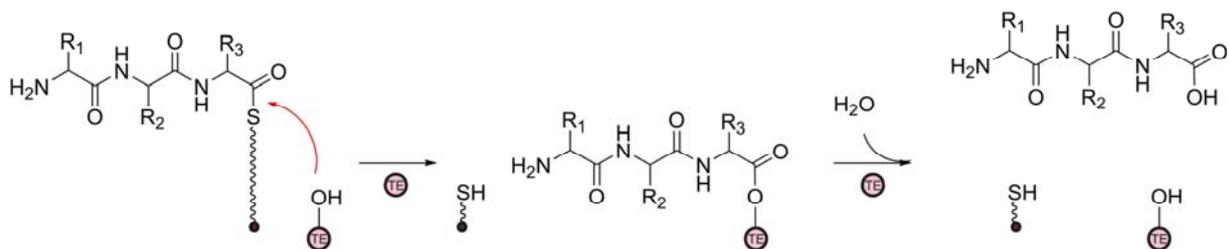
Die so auf den T-Domänen der Module befindlichen Aminosäuren müssen in einem nächsten Schritt durch Peptidbindungen verknüpft werden. Katalysiert wird der dafür notwendige Angriff der Aminogruppe der z.B. zweiten Aminosäure auf die als Thioester vorliegende Carboxygruppe der ersten Aminosäure durch die Kondensations- bzw. C-Domänen, die immer zwischen AT-Domänen lokalisiert sind (s. Abbildung 1-4).



**Abbildung 1-4: Das Knüpfen der Peptidbindungen wird durch C-Domänen katalysiert. Die dafür notwendige Energie liefern die energiereichen Thioesterbindungen.**

Die C-Domänen besitzen dabei eine gewisse Kontrollfunktion, die bezüglich der jeweils neu in die Peptidkette eintretenden Aminosäure relativ ausgeprägt ist. Dies trägt zusammen mit der hohen Substratspezifität der A-Domänen dazu bei, dass fehlerhafte Peptidsynthetaseprodukte vermieden werden. Nach der erfolgreichen

Peptidsynthese muss das fertig prozessierte Produkt vom NRPS-Multienzym-Komplex freigesetzt werden. Zuständig hierfür ist die Thioesterase- bzw. TE-Domäne, die aus naheliegenden Gründen nur im letzten Modul vorkommt und zunächst das als Thioester gebundene Peptid auf einen Serinrest überträgt sowie anschließend die Hydrolyse des Esters und damit die Abspaltung des Peptids katalysiert (s. Abbildung 1-5).



**Abbildung 1-5: Freisetzung des fertigen Peptids durch die TE-Domäne in zwei Stufen. Zunächst wird das als Thioester gebundene Peptid auf einen Serinrest der TE-Domäne übertragen und anschließend katalysiert diese die Hydrolyse des Esters und damit die Abspaltung des Peptids.**

Als alternative Möglichkeit zur Hydrolyse kann der Ester auch durch ein peptidinternes Nukleophil angegriffen werden, was z.B. in der Abspaltung eines zyklisierten Peptids resultiert.

Die bisher beschriebenen Domänen A, T, C und TE kommen in allen NRPS vor und stellen somit das Basisdomänenset dar. Während A- und T-Domänen in jedem einzelnen Modul der NRPS vorkommen, fehlen C-Domänen im ersten sog. Initiationsmodul. TE-Domänen kommen nur im letzten sog. Terminationsmodul vor.

Durch dieses Basisset an Domänen alleine ließe sich allerdings die enorme strukturelle Vielfalt der nicht-ribosomal synthetisierten Peptide nicht erklären. Es existieren zusätzlich eine Reihe von modifizierenden Domänen, die optional in Modulen enthalten sein können. Hierzu zählen beispielsweise die Epimerisierungs- bzw. E-Domänen, die das  $\alpha$ -C-Atom von an T-Domänen gebundenen L-Aminosäuren epimerisieren, wobei die verantwortliche C-Domäne nur die D-Form in das Peptid einbaut. Dies erklärt den hohen Anteil an D-Aminosäuren in nicht-ribosomal synthetisierten Peptiden.

Darüber hinaus gibt es außerdem die Möglichkeit, dass das durch die NRPS synthetisierte Peptidgrundgerüst durch unabhängige *in trans* arbeitende Enzyme weiter modifiziert wird. Diese Enzyme werden meist zusammen mit den NRPS in einem sog. Biosynthesegencluster kodiert. Ein Paradebeispiel hierfür stellt die Biosynthese der Glykopeptide dar (s. Abschnitt 1.3.2).

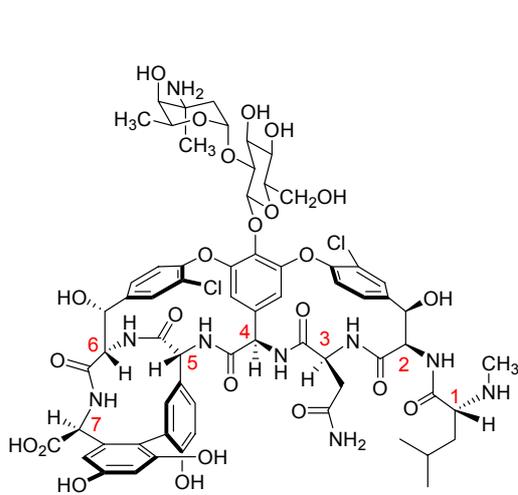
### 1.3 Glykopeptide

Glykopeptide werden als Sekundärmetabolite von Actinomyceten produziert, wobei man derzeit mehrere hundert natürlich vorkommende Vertreter kennt.<sup>[16]</sup> Zählt man die semisynthetisch hergestellten Analoga hinzu, erhöht sich die Zahl auf einige tausend.<sup>[17]</sup> Die folgenden Abschnitte sollen einen Überblick über die strukturellen und biosynthetischen Eigenschaften dieser Verbindungsklasse geben, die in vielerlei Hinsicht außergewöhnlich ist. Darüber hinaus soll die molekulare Grundlage für die antimikrobielle Aktivität der Glykopeptide dargestellt werden, die in der Humanmedizin Anwendung findet sowie die demgegenüberstehenden Resistenzmechanismen, die die Suche nach neuen Antibiotika vor große Herausforderungen stellen.

#### 1.3.1 Struktur der Glykopeptide

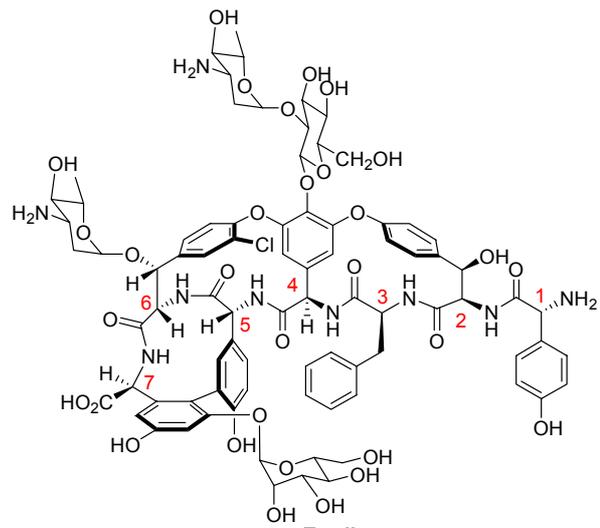
Der erste Vertreter dieser Verbindungsklasse, das Vancomycin, wurde 1956 bei Eli Lilly aus dem Kulturüberstand von *Amycolatopsis orientalis* isoliert.<sup>[3]</sup> Die Strukturaufklärung von Vancomycin nahm aufgrund der Komplexität und der damals begrenzten analytischen Möglichkeiten mehrere Jahrzehnte in Anspruch<sup>[18,19]</sup> und legte die Basis für die Strukturaufklärung nachfolgender Glykopeptide. Im Jahre 1996 folgte schließlich die Röntgenkristallstrukturanalyse von Vancomycin<sup>[20]</sup> und lieferte somit den endgültigen Beweis für die Richtigkeit der vorgeschlagenen Struktur.

Glykopeptide besitzen in der Regel ein Heptapeptidrückgrat, das vorwiegend aus aromatischen, nicht-proteinogenen Aminosäuren wie beispielsweise  $\beta$ -Hydroxytyrosin ( $\beta$ -Hty), 4-Hydroxyphenylglycin (Hpg) und 3,5-Dihydroxyphenylglycin (Dpg) besteht. Diese Aminosäuren können an ihren  $\alpha$ -C-Atomen nicht nur die für ribosomal synthetisierte Proteine typische *L*-Konfiguration aufweisen, sondern auch die *D*-Konfiguration. Die aromatischen Aminosäuren sind häufig in den Seitenketten oxidativ über Biaryl- bzw. Diaryletherbrücken verbrückt, was den Glykopeptiden eine gewisse konformative Starrheit verleiht. Überdies ist das Heptapeptidrückgrat häufig mit Cl-Atomen, *N*-Methylgruppen, verschiedenen Zucker- und/oder Fettsäureresten dekoriert. Anhand dieser genannten Strukturelemente werden die Glykopeptide in fünf Klassen eingeteilt (s. Abbildung 1-6).<sup>[16]</sup>



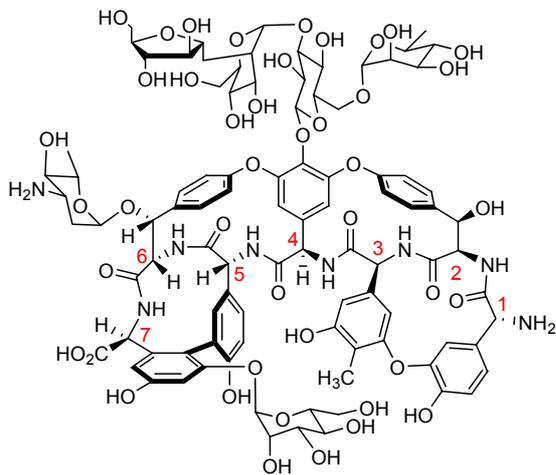
**Typ I**

Beispiel: Vancomycin  
Charakteristikum: aliphatische Aminosäuren  
in Position 1 und 3



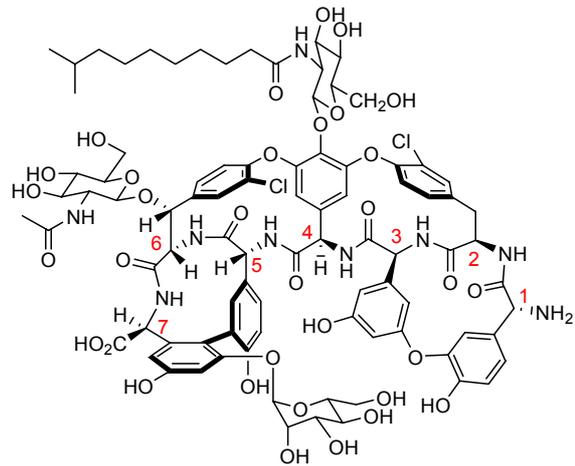
**Typ II**

Beispiel: Actinoidin A  
Charakteristikum: aromatische Aminosäuren  
in Position 1 und 3



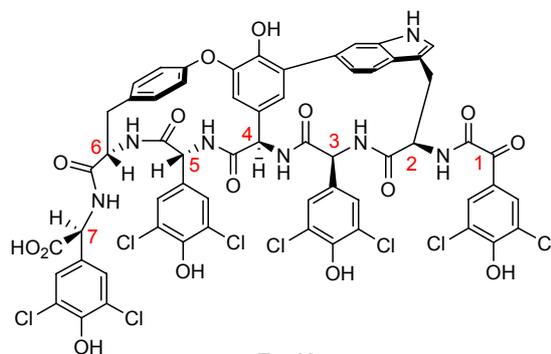
**Typ III**

Beispiel: Ristocetin  
Charakteristikum: aromatische Aminosäuren in Position 1 und 3  
sind durch eine Diaryletherbrücke verknüpft



**Typ IV**

Beispiel: Teicoplanin A<sub>2</sub>-2  
Charakteristikum: analog Typ III, wobei die Aminosucker  
Fettsäurereste tragen



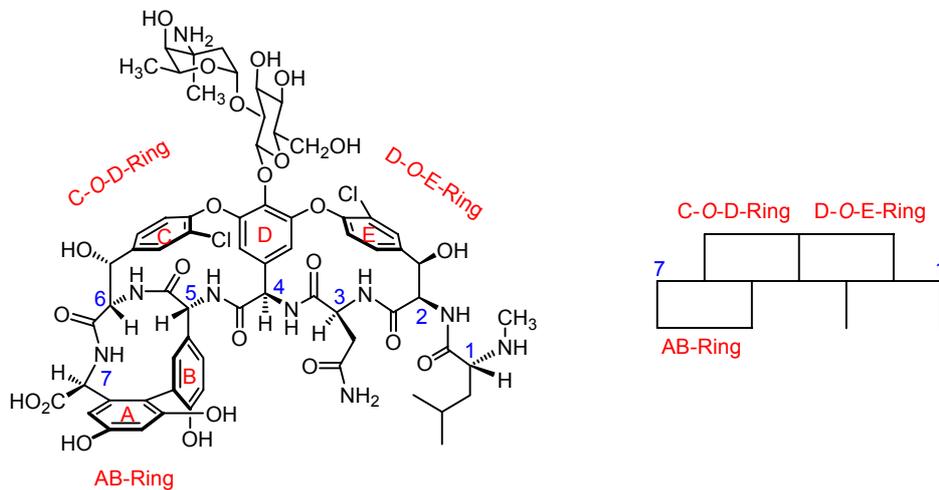
**Typ V**

Beispiel: Complestatin  
Charakteristikum: Trp in Position 2 ist über Biarylbrücke mit  
Aminosäure in Position 4 verknüpft

**Abbildung 1-6: Überblick über die fünf strukturellen Klassen der Glykopeptide.**

Für die Nomenklatur dieser außergewöhnlichen Peptide gelten folgende Regeln (s. Abbildung 1-7):<sup>[16]</sup>

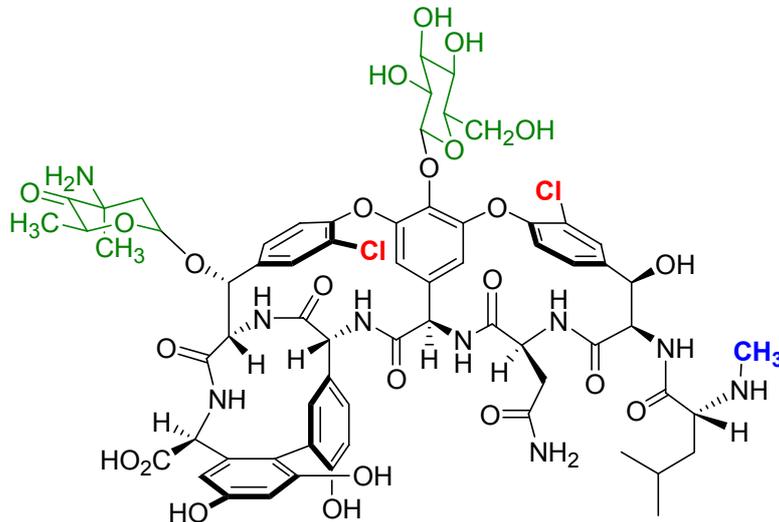
- 1) Die Aminosäuren werden wie bei allen Peptiden vom *N*- zum *C*-Terminus mit eins bis sieben durchnummeriert. Allerdings wird aus historischen Gründen der *C*-Terminus links und der *N*-Terminus rechts gezeichnet.
- 2) Die aromatischen Ringe werden mit Großbuchstaben benannt, beim Vancomycin beispielsweise mit A bis E.
- 3) Die durch die oxidativen Seitenkettenverbrückungen entstehenden großen Ringe werden mit Hilfe der Buchstaben der beteiligten aromatischen Ringe bezeichnet. Bei Vancomycin spricht man deshalb vom AB-, CD- und DE-Ring. Da die beiden letzten über Etherbrücken geschlossen sind, werden sie auch alternativ als C-O-D- und D-O-E-Ring bezeichnet.



**Abbildung 1-7: Nomenklatur der Glykopeptide am Beispiel des Typ I-Glykopeptids Vancomycin. Links: Strukturabbildung. Rechts: Schemazeichnung.**

Das zentrale Glykopeptid der vorliegenden Arbeit ist das Typ I-Glykopeptid Balhimycin, welches von *Amycolatopsis balhimycina* produziert wird<sup>[21,22,23]</sup> und die in Abbildung 1-8 gezeigte Struktur besitzt. Das Heptapeptidrückgrat besteht hier aus der Aminosäure-Sequenz *D*-Leu – *D*- $\beta$ -Hty – *L*-Asn – *D*-Hpg – *D*-Hpg – *L*- $\beta$ -Hty – *L*-Dpg und weist die drei für Typ I-Glykopeptide typischen oxidativen Seitenkettenzyklisierungen AB, C-O-D und D-O-E auf (s. Abbildung 1-7). Darüber hinaus ist das Heptapeptidrückgrat mit einer *N*-Methylgruppe am *N*-Terminus (blau), zwei Chloratomen an den aromatischen Ringen der zwei  $\beta$ -Hty's (rot) sowie zwei Glykosylresten (grün) dekoriert (s. Abbildung 1-8). Bei den Glykosylresten handelt es sich zum einen um eine Glukose, die  $\alpha$ -glykosidisch mit der phenolischen

Hydroxygruppe des <sup>4</sup>D-Hpg verknüpft ist und zum anderen um ein 4-oxo-Vankosamin, das  $\alpha$ -glykosidisch mit der  $\beta$ -Hydroxygruppe des <sup>6</sup>L- $\beta$ -Hty verbunden ist. Damit unterscheidet sich die Struktur des Balhimycins lediglich im Glykosylierungsmuster von der Vancomycinstruktur.



**Abbildung 1-8: Struktur von Balhimycin, einem Typ I-Glykopeptid. Das Heptapeptidrückrat ist mit einer N-Methylgruppe (blau), zwei Chloratomen (rot) sowie den zwei Glykosylresten Glukose und 4-oxo-Vankosamin (grün) dekoriert.**

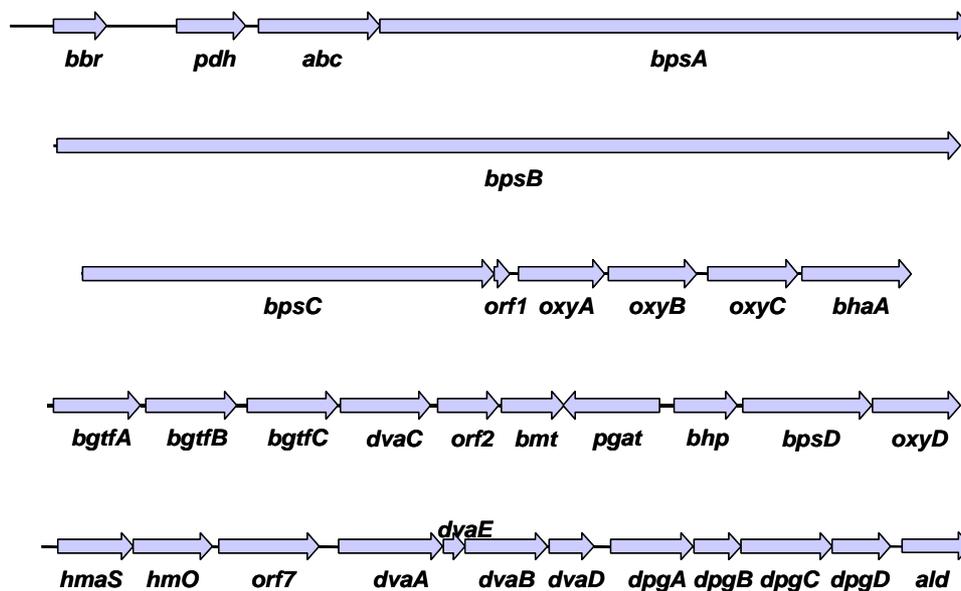
### 1.3.2 Biosynthese der Glykopeptide

Wie die Komplexität der Glykopeptidstrukturen bereits vermuten lässt, handelt es sich bei der Biosynthese der Glykopeptide um einen mehrstufigen Prozess, an dem viele Biosyntheseenzyme beteiligt sind. Die entsprechenden Gene liegen im Genom der Produzentenstämme geclustert vor, was eine gemeinsame Regulierung erlaubt und vor allem für Biosyntheseuntersuchungen sehr hilfreich ist. Das erste Biosynthesegencluster eines Glykopeptids wurde 1998 von van Wageningen et al. publiziert,<sup>[24]</sup> wobei es sich um das Typ I-Glykopeptid Chloroeremomycin handelte. Kurz darauf folgte 1999 das Biosynthesegencluster des Typ I-Glykopeptids Balhimycin von Pelzer et al.,<sup>[25]</sup> dessen Zuständigkeit für die Balhimycinbiosynthese durch Geninaktivierung bewiesen wurde.<sup>[25,26]</sup> Damit war der Startschuss für tiefgreifende Biosyntheseuntersuchungen gefallen,<sup>[27,28]</sup> wobei zwei grundsätzliche Strategien zur Anwendung kamen bzw. kommen. Die erste Strategie umfasst die heterologe Überexpression von Biosyntheseenzymen meist in *E. coli* und deren anschließende Funktionsuntersuchung mit Hilfe von Substraten. Die zweite Strategie sieht eine Inaktivierung der entsprechenden Gene vor, wobei über die

Charakterisierung der von diesen Mutanten produzierten Glykopeptidvarianten Rückschlüsse auf die Funktion der zugehörigen Enzyme gezogen werden können. Mit Hilfe dieser Strategien konnten viele Erkenntnisse über die Biosynthese gewonnen werden,<sup>[27,28]</sup> die aufgrund der strukturellen und genetischen Verwandtschaft meist auf alle Vertreter der Typ I-Glykopeptide und mit Einschränkungen auch auf die anderen Klassen übertragbar sind.

Da in der vorliegenden Arbeit das Glykopeptid Balhimycin im Mittelpunkt steht, soll die Glykopeptidbiosynthese anhand dieses Beispiels kurz beschrieben werden. Eine Übersicht des Balhimycin-Biosynthesegenclusters ist in

Abbildung 1-9 gezeigt, wobei ergänzende Erklärungen der Tabelle 1 zu entnehmen sind.



**Abbildung 1-9: Schematischer Überblick über das Balhimycin-Biosynthesegencluster von *Amycolatopsis balhimycina*.**

**Tabelle 1: Übersicht über die den Balhimycin-Biosynthesegenen zugeordneten Enzym- bzw. Proteinfunktionen. Diese sind teilweise experimentell bestätigt, teilweise auch nur hypothetisch und beruhen somit lediglich auf Sequenzhomologievergleichen.<sup>[29]</sup>**

Gen	Enzym-/Proteinfunktion	übergeordnete Funktion
<i>vanS</i>	Zweikomponentensystem-Sensorkinase <sup>1)</sup>	vermutlich Selbstresistenz
<i>vanR</i>	Zweikomponentensystem-responsiver Regulator <sup>1)</sup>	vermutlich Selbstresistenz
<i>vanY</i>	VanY-Typ Carboxypeptidase <sup>1)</sup>	vermutlich Selbstresistenz
<i>bbr</i>	transkriptioneller Regulator der StrR-Familie	-

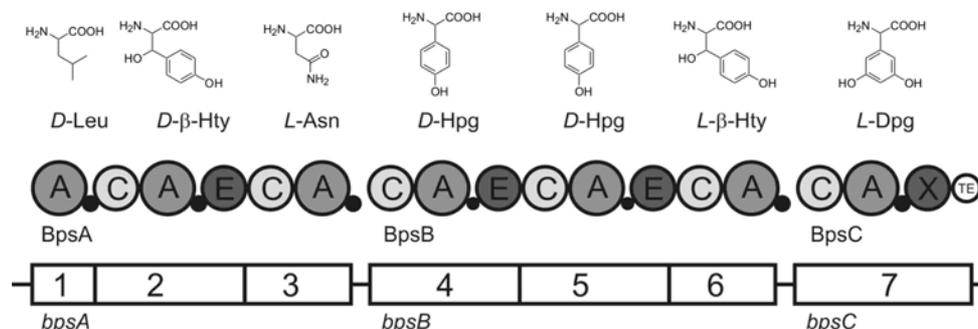
<i>pdh</i>	Prephenatdehydrogenase <sup>1)</sup>	4-Hydroxyphenylglycin-Biosynthese <sup>4)</sup>
<i>abc</i>	ABC-Transporter ATP-bindendes Protein <sup>1)</sup>	vermutlich Exkretion
<i>bpsA</i>	Peptidsynthetase	Aufbau des Heptapeptidrückgrats
<i>bpsB</i>	Peptidsynthetase	Aufbau des Heptapeptidrückgrats
<i>bpsC</i>	Peptidsynthetase	Aufbau des Heptapeptidrückgrats
<i>orf1</i>	unbekannt	-
<i>oxyA</i>	P450-Monooxygenase	Verbrückung der aromatischen Seitenketten
<i>oxyB</i>	P450-Monooxygenase	Verbrückung der aromatischen Seitenketten
<i>oxyC</i>	P450-Monooxygenase	Verbrückung der aromatischen Seitenketten
<i>bhaA</i>	Halogenase	Halogenierung
<i>bgfA</i>	Glykosyltransferase	Glykosylierung
<i>bgfB</i>	Glykosyltransferase	Glykosylierung
<i>bgfC</i>	Glykosyltransferase	Glykosylierung
<i>dvaC</i>	C-3 Methyltransferase <sup>1)</sup>	C-3 Methylierung von Dehydrovankosamin <sup>2)</sup>
<i>orf2</i>	unbekannt	-
<i>bmt</i>	N-Methyltransferase <sup>1)</sup>	Methylierung des D-Leu <sup>2)</sup>
<i>pgat</i>	Phenylglycin-Aminotransferase	4-Hydroxy- und 3,5-Dihydroxyphenylglycin-Biosynthese
<i>bhp</i>	Hydrolase <sup>1)</sup>	$\beta$ -Hydroxytyrosin-Biosynthese
<i>bpsD</i>	Peptidsynthetase	$\beta$ -Hydroxytyrosin-Biosynthese
<i>oxyD</i>	P450-Monooxygenase	$\beta$ -Hydroxytyrosin-Biosynthese <sup>2)</sup>
<i>hmaS</i>	Hydroxyphenylpyruvat-Dioxygenase <sup>1)</sup>	4-Hydroxyphenylglycin-Biosynthese <sup>2)</sup>
<i>hmO</i>	Phenylglykolat-Oxidase <sup>1)</sup>	4-Hydroxyphenylglycin-Biosynthese <sup>2)</sup>
<i>orf7</i>	Antiporter <sup>1)</sup>	-
<i>dvaA</i>	NDP-Hexose-2,3-Dehydratase <sup>1)</sup>	Dehydrovankosamin-Biosynthese <sup>2)</sup>
<i>dvaE</i>	4-Ketoreduktase <sup>3)</sup>	Dehydrovankosamin-Biosynthese <sup>5)</sup>
<i>dvaB</i>	C-3 Aminotransferase <sup>1)</sup>	Dehydrovankosamin-Biosynthese <sup>2)</sup>
<i>dvaD</i>	3,5-Epimerase <sup>1)</sup>	Dehydrovankosamin-Biosynthese <sup>2)</sup>
<i>dpgA</i>	Dihydroxyphenylessigsäure-Synthase	3,5-Dihydroxyphenylglycin-Biosynthese
<i>dpgB</i>	Enoyl-CoA-Hydratase <sup>1)</sup>	3,5-Dihydroxyphenylglycin-Biosynthese
<i>dpgC</i>	Hydroxyacyl-Dehydrogenase	3,5-Dihydroxyphenylglycin-Biosynthese
<i>dpgD</i>	Enoyl-CoA-Isomerase <sup>1)</sup>	3,5-Dihydroxyphenylglycin-Biosynthese
<i>ald</i>	Aldolase <sup>1)</sup>	-
<i>pks</i>	Typ I Polyketidsynthase <sup>1)</sup>	-

- 1) Diese Enzym-/Proteinfunktion wird vermutet aufgrund von Sequenzhomologievergleichen.
- 2) Die Zugehörigkeit zu diesem Biosyntheseweg wird vermutet.
- 3) Eventuell inaktiv aufgrund einer *in-frame* Deletion von 226 Aminosäuren.
- 4) Ergänzung aus [27].
- 5) Ergänzung aus [28].

Allgemein umfasst die Glykopeptidbiosynthese die drei Bereiche „Biosynthese ungewöhnlicher Ausgangsverbindungen“, „nicht-ribosomale Peptidsynthese“ und „Dekorierungsreaktionen“ („tailoring“).

Bei Balhimycin gehören sowohl die nicht-proteinogenen Aminosäuren  $\beta$ -Hty,<sup>[30,31]</sup> Hpg<sup>[32,33,34]</sup> und Dpg<sup>[35,36,37,38]</sup> als auch die Hexose 4-oxo-Vankosamin<sup>[39]</sup> zu den ungewöhnlichen Ausgangsverbindungen, die nicht im Primärstoffwechsel vorkommen und deshalb gesondert synthetisiert werden müssen.

Die Grundzüge der nicht-ribosomalen Peptidsynthese wurden bereits in Abschnitt 1.2 beschrieben. Im Falle des Balhimycins sind die drei nicht-ribosomalen Peptidsynthetasen BpsA, BpsB und BpsC für die Biosynthese des Heptapeptidrückgrats zuständig.<sup>[40]</sup> Dafür sind sieben Module notwendig, die in einer 3:3:1-Verteilung auf die drei NRPS verteilt sind. Die Domänenstruktur der einzelnen Module sowie die Zuordnung der Module zu den entsprechenden Aminosäuren sind in Abbildung 1-10 gezeigt. Erstaunlicherweise kommt in Modul 1 keine Epimerasedomäne vor, obwohl Leu in der *D*-Konfiguration eingebaut vorliegt. Zum gegenwärtigen Zeitpunkt wird daher eine *in trans* arbeitende Epimerase bzw. Racemase postuliert.<sup>[41]</sup>



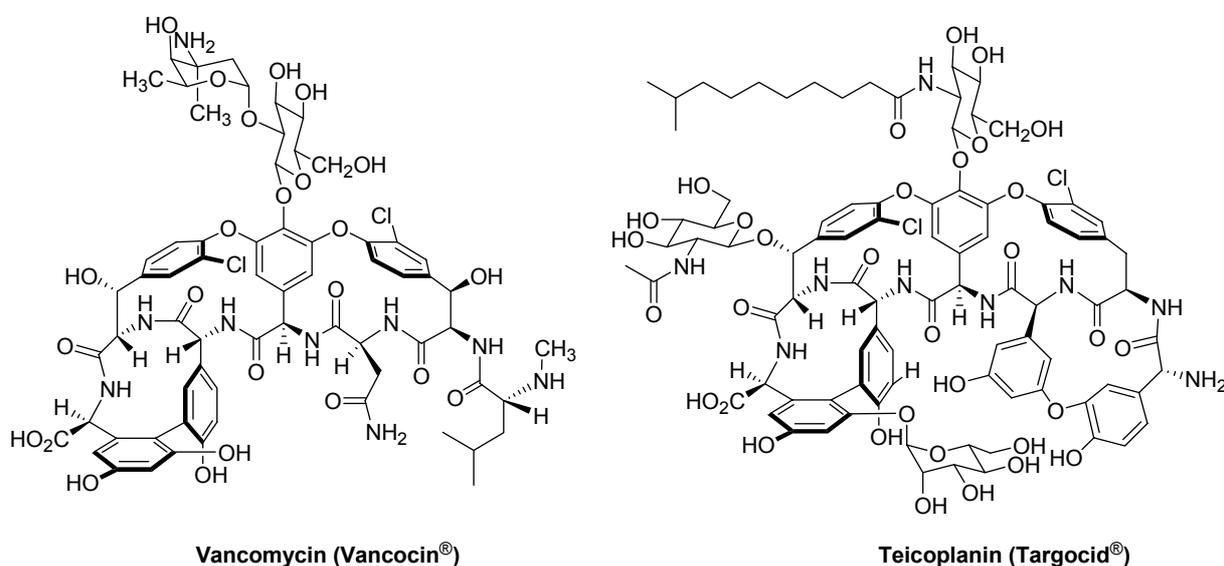
**Abbildung 1-10: Die Gene bzw. Proteine der drei nicht-ribosomalen Peptidsynthetasen von Balhimycin. Zu sehen sind die insgesamt sieben Module, die wiederum aus unterschiedlichen Domänen aufgebaut sind. A = Adenylierungsdomäne, T(•) = Thiolierungsdomäne, C = Kondensationsdomäne, TE = Thioesterasedomäne, E = Epimerisierungsdomäne. X = Domäne mit Homologie zu E bzw. C mit bislang ungeklärter Funktion. Ganz oben sind die zu den einzelnen Modulen gehörigen Aminosäuren abgebildet.**

Die Dekorierungsreaktionen des Heptapeptidrückgrats umfassen im Falle des Balhimycins „Chlorierung“, „oxidative Seitenkettenzyklisierung“, „*N*-Methylierung“ und „Glykosylierung“. Bezüglich der Chlorierung konnten durchgeführte Geninaktivierungsexperimente zum einen die Zuständigkeit der Halogenase BhaA zeigen<sup>[30]</sup> und zum anderen den Chlorierungszeitpunkt auf `während der nicht-ribosomalen Peptidsynthese` einschränken.<sup>[31]</sup> Auch im Falle der oxidativen Seitenkettenzyklisierung wurde der heutige Kenntnisstand vorwiegend mittels Geninaktivierungsexperimenten in *A. balhimycina* ermittelt. Die drei oxidativen Seitenkettenzyklisierungen A-B, C-O-D und D-O-E werden von den drei P450-abhängigen Monooxygenasen OxyA, OxyB und OxyC katalysiert,<sup>[42]</sup> wobei folgende Reihenfolge und folgende Enzym-Ring-Zuordnung gilt: C-O-D (OxyB) -> D-O-E (OxyA) -> A-B (OxyC).<sup>[43]</sup> Außerdem konnte gezeigt werden, dass diese oxidativen Seitenkettenzyklisierungen auf der Stufe der NRPS-gebundenen Hexa- bzw. Heptapeptide ablaufen, was für eine Interaktion zwischen den nicht-ribosomalen Peptidsynthetasen und den P450-abhängigen Monooxygenasen spricht.<sup>[44,45]</sup> Während es sich also bei der Chlorierung und der oxidativen Seitenkettenzyklisierung um inter-NRPS-Dekorierungsreaktionen handelt, sind die *N*-Methylierung<sup>[26,46]</sup> und die Glykosylierung<sup>[26,47,48]</sup> post-NRPS-Dekorierungsreaktionen, die erst ablaufen, wenn das Heptapeptidrückgrat vom Multienzymkomplex der NRPS abhydrolysiert wurde. Das Balhimycin-Biosynthesecluster kodiert dafür das Gen für die *N*-Methyltransferase Bmt, die *S*-Adenosylmethionin als Methylgruppendonor verwendet sowie Gene für die drei Glykosyltransferasen BgtfA, BgtfB<sup>[25]</sup> und BgtfC.

Zusammenfassend handelt es sich also bei der Balhimycin-Biosynthese und damit allgemein bei Glykopeptidbiosynthesen um sehr komplexe Beispiele der nicht-ribosomalen Peptidbiosynthese, die ein koordiniertes Zusammenspiel vieler verschiedener Enzyme erfordern. Ein Eingriff in eine solche Biosynthese muss deshalb sehr durchdacht sein, wie das in der vorliegenden Arbeit durchgeführte Projekt zeigt.

### 1.3.3 Antimikrobielle Aktivität der Glykopeptide und auftretende Resistenzmechanismen

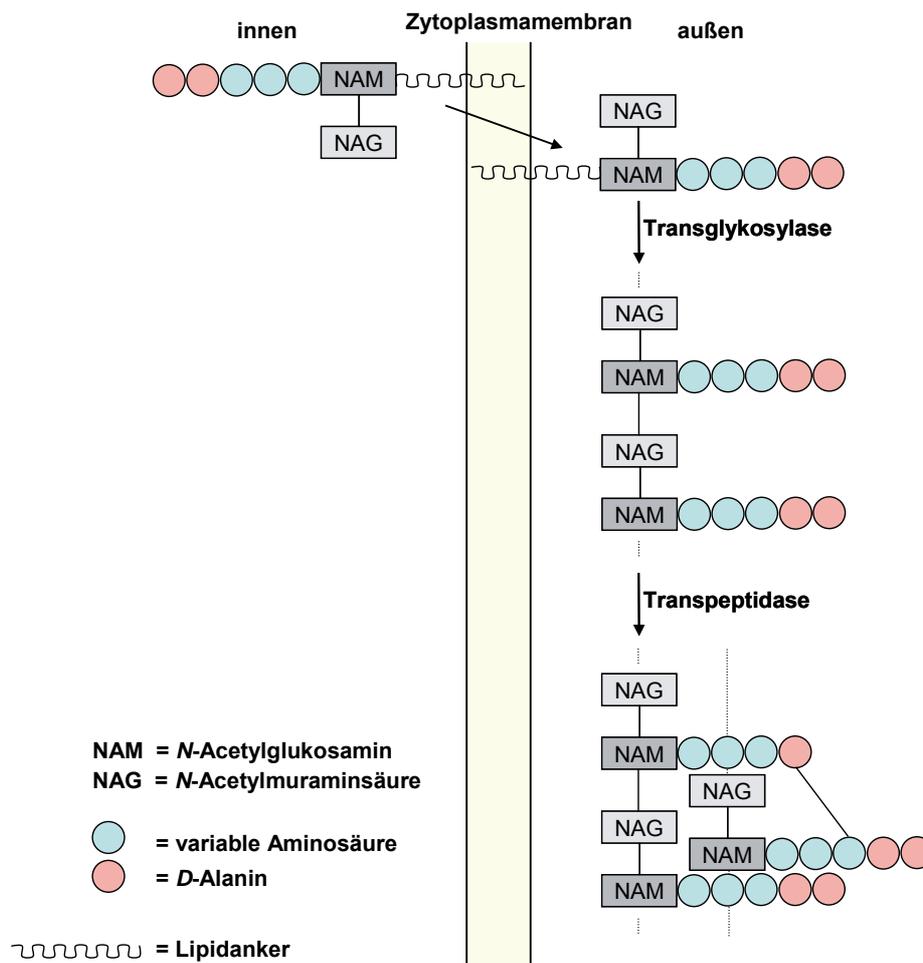
Die meisten Vertreter der Verbindungsklasse der Glykopeptide zeichnen sich durch eine antimikrobielle Aktivität gegen Gram-positive Bakterien aus, weshalb sie auch Glykopeptidantibiotika genannt werden. Neben dem bereits in Abschnitt 1.1 erwähnten Vancomycin, welches 1958 unter dem Handelsnamen Vancocin als Medikament zugelassen wurde, ist auch das von *Actinoplanes teichomyceticus* produzierte Teicoplanin<sup>[49,50]</sup> als Targocid in der humanmedizinischen Anwendung (s. Abbildung 1-11).



**Abbildung 1-11: Die beiden für die klinische Anwendung am Menschen zugelassenen Glykopeptidantibiotika Vancomycin und Teicoplanin. Die Handelsnamen sind in Klammern gesetzt.**

Ihre Indikation sind schwere und lebensbedrohliche Infektionen mit Gram-positiven Bakterien, wobei sie vor allem als Notfallantibiotika gegen Methicillin-resistente Stämme von *S. aureus* (MRSA) große Bedeutung erlangt haben.<sup>[51]</sup> Ihre Wirksamkeit beruht auf der Hemmung der Zellwandbiosynthese. Dies ist ein Stoffwechselweg, der wie z.B. auch die Folsäurebiosynthese nicht zum Biosyntheserepertoire des menschlichen Organismus gehört und damit ein ideales Target für antimikrobielle Wirkstoffe darstellt. Genauer gesagt hemmen Glykopeptidantibiotika die Zellwandbiosynthese beim Aufbau bzw. bei der Quervernetzung der Peptidoglykanschicht. Diese Schicht kommt bei Gram-positiven und Gram-negativen Bakterien außerhalb der Zytoplasmamembran vor und dient der osmotischen

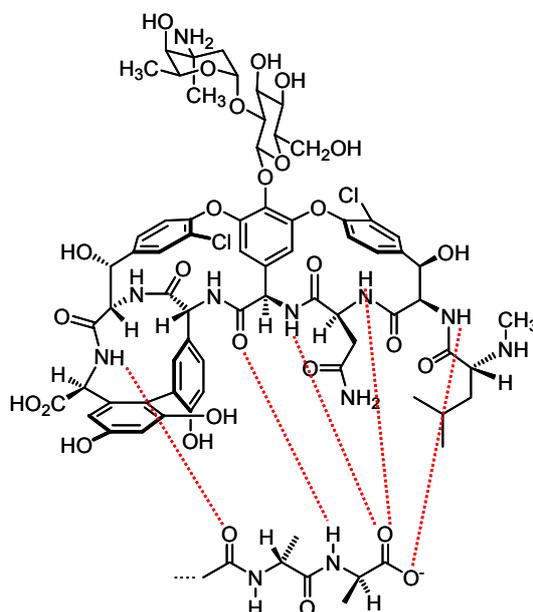
Stabilität. Allerdings besitzen Gram-negative Bakterien zusätzlich eine äußere Membran, die das Eindringen der Glykopeptidantibiotika an ihren Wirkort verhindert und sie somit unschädlich macht. Die Peptidoglykanschicht besteht aus Polysaccharidketten, die abwechselnd aus *N*-Acetylglukosamin (NAG) und *N*-Acetylmuraminsäure (NAM) aufgebaut sind. Die die Festigkeit verleihende Quervernetzung erfolgt über Peptidketten, wobei verschiedene Aminosäuren je nach Bakterienarten verwendet werden können. Die Biosynthese der Peptidoglykanschicht findet außerhalb der Zelle in zwei Hauptschritten statt (s. Abbildung 1-12).



**Abbildung 1-12: Schematische Darstellung der Peptidoglykanbiosynthese.**

Im ersten Schritt werden Disaccharide aus NAG und NAM mit anhängenden Peptiden aus dem Zytoplasma auf die Zellaußenseite transportiert, wo sie durch eine Transglykosylase miteinander verknüpft werden. Der zweite Schritt umfasst die Verknüpfung der Peptidenden durch eine Transpeptidase. Dieses Enzym erkennt die Sequenz *D*-Ala-*D*-Ala am Ende der Peptidkette, spaltet ein Ala ab und verknüpft den Rest mit der Peptidkette einer benachbarten Polysaccharidkette.

Glykopeptidantibiotika wechselwirken über fünf Wasserstoffbrückenbindungen mit dem *D*-Ala-*D*-Ala-Ende der Peptidketten, was in Abbildung 1-13 am Beispiel von Vancomycin gezeigt ist.

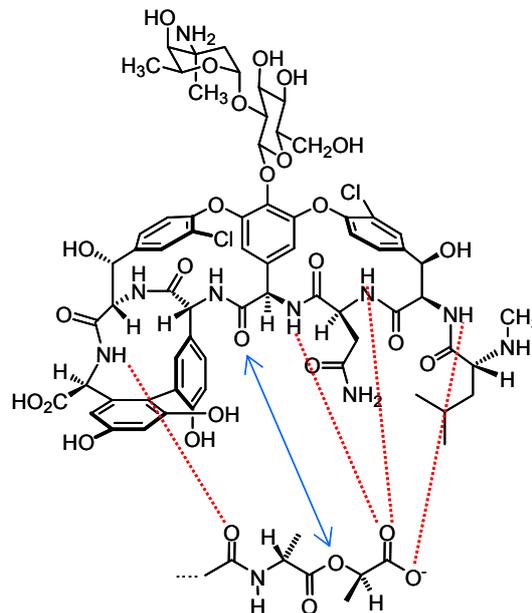


**Abbildung 1-13: Wechselwirkung der Peptidoglykan-Zwischenstufe *D*-Ala-*D*-Ala mit Vancomycin. Die fünf Wasserstoffbrückenbindungen sind rot dargestellt.**

Die Abschirmung des Substrats behindert die Arbeit der Transpeptidase und damit indirekt auch die der Transglykosylase, was eine fehlende Quervernetzung der Zellwand und in letzter Konsequenz eine osmotische Lyse der Bakterien bewirkt.<sup>[52]</sup>

Mikroorganismen können sich der schädigenden Wirkung von Antibiotika entziehen, was vor allem auf ihre hohe spontane Mutationsrate sowie ihre Fähigkeit, Resistenzgene untereinander austauschen zu können (Gentransfer), zurückzuführen ist. Über 30 Jahre lang wurden keine Resistenzentwicklungen bei der Applikation von Vancomycin beobachtet, bis schließlich 1988 die ersten Vancomycin-resistenten Enterokokkenstämme (VRE)<sup>[53]</sup> und 1997 die ersten Staphylokokkenstämme mit verminderter Empfindlichkeit gegenüber Vancomycin (VISA, *vancomycin-intermediate S. aureus*)<sup>[54]</sup> entdeckt wurden. Im Folgenden sollen kurz die typischen Resistenzmechanismen bei Glykopeptiden dargestellt werden.<sup>[27]</sup> Eine häufig genutzte Möglichkeit der Resistenz besteht darin, statt des *D*-Ala-*D*-Ala-Motivs ein *D*-Ala-*D*-Lac-Motiv zu verwenden. Während die Transpeptidase dieses Strukturmotiv ebenso als Substrat akzeptiert, bindet Vancomycin mit 1000-fach geringerer Affinität. Grund hierfür ist zum einen der Wegfall einer Wasserstoffbrückenbindung und zum anderen eine hinzukommende elektrostatische Abstoßung zwischen dem

Carbonylsauerstoffatom des Glykopeptids und dem Estersauerstoffatom der *D*-Ala-*D*-Lac-Bindung (s. Abbildung 1-14).<sup>[55,56]</sup>

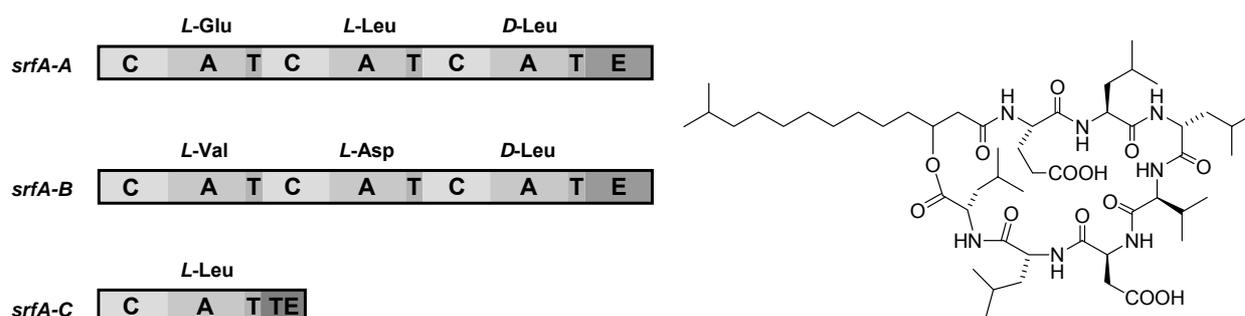


**Abbildung 1-14:** Ein typischer Resistenzmechanismus bei Glykopeptidantibiotika ist die Verwendung des *D*-Ala-*D*-Lac-Motivs anstelle des *D*-Ala-*D*-Ala-Motivs. Dadurch ändert sich die Wechselwirkung mit Vancomycin. Es treten nur noch vier Wasserstoffbrückenbindungen auf (rot) und außerdem eine elektrostatische Abstoßung (blau).

Als Alternative dazu wurde auch das Auftreten eines *D*-Ala-*D*-Ser-Motivs beobachtet.<sup>[57]</sup> Der Austausch der Methylgruppe des *D*-Ala gegen die sterisch anspruchsvollere Hydroxymethylgruppe des *D*-Ser verringert die Bindungsaffinität des Vancomycins um eine Größenordnung. Eine ganz andere Resistenzvariante wird durch eine vermehrte Peptidoglykan-Biosynthese erreicht, die außerdem mit einer verminderten Quervernetzung einhergeht. Die Folge ist, dass durch die vermehrte Biosynthese der *D*-Ala-*D*-Ala-Motive das Vancomycin sozusagen „abgefangen“ wird.<sup>[58]</sup> Dies zeigt eindringlich, dass dem „Einfallsreichtum“ der Mikroorganismen nahezu keine Grenzen gesetzt sind und dass es deshalb eine der wichtigsten Aufgaben der Pharmaindustrie sein muss, sich dieser Herausforderung zu stellen.

## 1.4 Kombinatorische Biosynthese bei nicht-ribosomalen Peptidsynthetasen

Der Begriff kombinatorische Biosynthese beschreibt im weitesten Sinne „die genetische Manipulation eines Biosynthesewegs mit dem Ziel, neue Biosyntheseprodukte zu erhalten“.<sup>[59,60,61,62]</sup> Die Voraussetzungen für solche Manipulationen sind die Kenntnis der entsprechenden Biosynthesegencluster sowie die eingehende Charakterisierung der Biosyntheseabläufe. Bezüglich der nicht-ribosomalen Peptidsynthese sind diese Voraussetzungen in den vergangenen Jahrzehnten erfüllt worden, so dass heute zum einen viele nicht-ribosomale Biosynthesegencluster bekannt und deren Sequenzen in Datenbanken hinterlegt sind und zum anderen viele Details der biosynthetischen Abläufe bekannt sind (s. Abschnitt 1.2). Besonders wichtig war in diesem Zusammenhang die Entdeckung der Kolinearitätsregel,<sup>[11]</sup> die für viele nicht-ribosomale Peptidsynthetasen gilt und die den Zusammenhang zwischen den Substratspezifitäten aufeinanderfolgender Module und der Aminosäure-Sequenz im Produkt beschreibt. Auf dieser Grundlage aufbauend erschienen Mitte der 1990er Jahre die ersten Publikationen über genetische Manipulationen der nicht-ribosomalen Peptidsynthese. Besonders bedeutsam war hierbei die von Marahiel et al. 1995 in Science erschienene Arbeit<sup>[63]</sup> über das von *Bacillus subtilis* produzierte nicht-ribosomale Modellpeptid Surfactin. Die Gene der drei NRPS von Surfactin *srfA-A*, *srfA-B* und *srfA-C* sowie die Struktur von Surfactin sind in Abbildung 1-15 gezeigt.

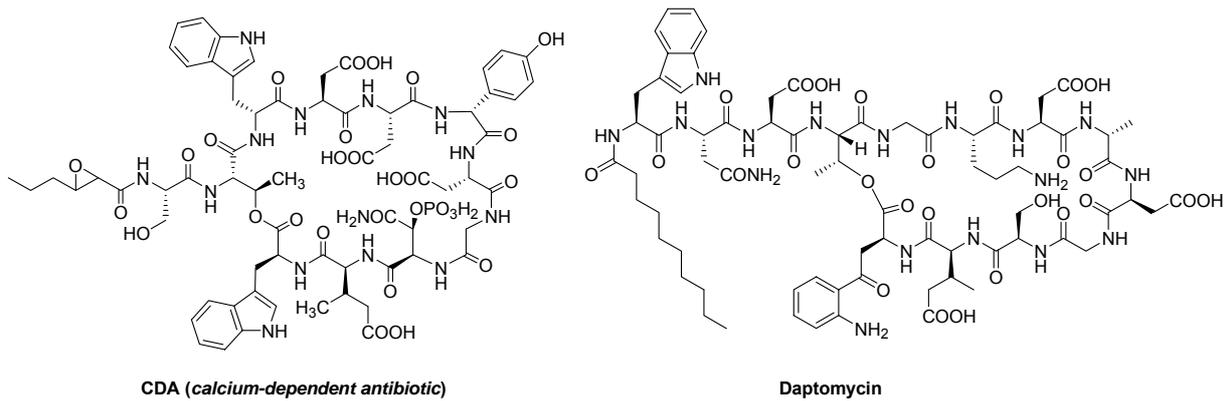


**Abbildung 1-15: Links: Gene der drei NRPS von Surfactin *srfA-A*, *srfA-B* und *srfA-C*. Zu sehen ist der modulare Aufbau, wobei jedes Modul wiederum aus mehreren Domänen besteht. Rechts: Struktur von Surfactin.**

Ziel dieser Arbeit war der Austausch der siebten Aminosäure *L*-Leu im Surfactinprodukt gegen andere Aminosäuren durch genetische Manipulation des dafür zuständigen siebten Moduls in *srfA-C*. Diese Manipulation erfolgte *in vivo*, also im *Bacillus subtilis*-Produzenten, durch den Austausch der AT-Domäne im siebten

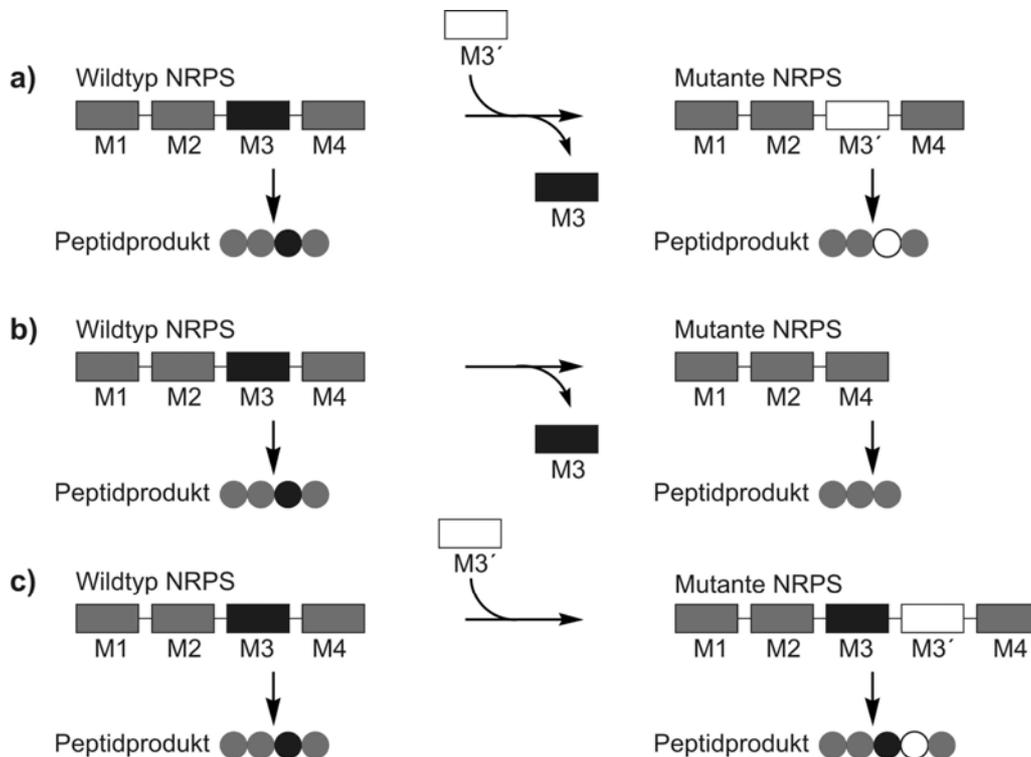
Modul gegen verschiedene AT-Domänen aus anderen Produzentenstämmen. Tatsächlich konnten in diesen *Bacillus subtilis*-Mutanten die erwarteten Surfactinderivate nachgewiesen werden, was als Durchbruch auf dem Gebiet der kombinatorischen Biosynthese bewertet werden kann. Allerdings deutete die geringe Produktivität der Mutanten von 0,1 bis 0,5 % der Wildtyp-Produktion darauf hin, dass die vorgenommenen genetischen Manipulationen nicht ohne weiteres von dem komplexen System der nicht-ribosomalen Peptidsynthese toleriert wurden. Ein Hauptproblem war in diesem Zusammenhang die bereits in Abschnitt 1.2 erwähnte Substratspezifität der C-Domäne für die neu in die Peptidkette eintretende Aminosäure. Durch den Austausch der AT-Domäne durch AT-Domänen mit anderen Substratspezifitäten war die Kompatibilität zwischen C- und A-Domäne nicht mehr gegeben. Ein weiterer kritischer Punkt war die Gestaltung der Klonierungsstrategie, die bevorzugt in den nicht-enzymatisch-aktiven Linkerregionen zwischen den Domänen bzw. Modulen<sup>[12]</sup> stattfinden sollte, um so keine für die Enzymaktivität essentiellen Aminosäuren zu verlieren.

Diese Problematik führte in der Folgezeit zu einer Vielzahl weiterer Arbeiten, die von unterschiedlichen Strategien für Aminosäureaustausche im Produkt mit mehr oder weniger großem Erfolg berichten.<sup>[64,65,66]</sup> Neben Domänen- bzw. Modulaustauschen (s. Abbildung 1-17a)) wurde hierbei auch die Strategie beschrieben, die Substratspezifität einer A-Domäne durch Punktmutation zu ändern, um so einen Aminosäureaustausch im Produkt herbeizuführen.<sup>[66]</sup> Diese Strategie wurde im Jahr 2004 zum ersten Mal von *Bacillus* auf die in der genetischen Handhabung bedeutend anspruchsvolleren Actinomyceten übertragen. Hierbei handelte es sich um einen durch zwei Punktmutationen in einer A-Domäne herbeigeführten Asp->Asn Austausch im Lipopeptid CDA (*calcium-dependent antibiotic*), welches von *Streptomyces coelicolor* produziert wird (s. Abbildung 1-16).<sup>[67]</sup> Im Jahr 2006 erschienen außerdem beeindruckende Arbeiten von Baltz et al.,<sup>[68,69]</sup> die durch Kombination diverser Modulaustausche eine große Zahl an Derivaten des in der medizinischen Anwendung (s. Abschnitt 1.1) befindlichen Lipopeptids Daptomycin herstellten, welches von *Streptomyces roseosporus* produziert wird (s. Abbildung 1-16). Neben einzelnen Modulen wurden hier auch mehrere Module und sogar ganze NRPS ausgetauscht.



**Abbildung 1-16: Strukturen von CDA (*calcium-dependent antibiotic*) und Daptomycin.**

Als fundamentale Weiterentwicklung der gerade mit diversen Beispielen belegten kombinatorischen Biosynthesestrategie der Aminosäureaustausche im Produkt ist die 2002 von Marahiel et al. publizierte Arbeit zu sehen.<sup>[70]</sup> Statt eines Aminosäureaustausches im Produkt wurde hier durch eine *in-frame* Moduldeletion (s. Abbildung 1-17b)) eine Aminosäuredeletion im Modellpeptid Surfactin (s. Abbildung 1-15) erzeugt, so dass anstelle des zyklischen Heptapeptids nur noch ein zyklisches Hexapeptid produziert wurde. Die Produktivität belief sich in diesem Fall immerhin auf 10 % des Wildtyps.



**Abbildung 1-17: Schematische Übersicht über die drei verschiedenen Möglichkeiten der Modulmanipulation bei nicht-ribosomalen Peptidsynthetasen. a) Modulaustausch führt zu einem Aminosäureaustausch im Peptidprodukt. b) Moduldeletion führt zu einer Aminosäuredeletion im Peptidprodukt. c) Modulinsertion führt zu einer Aminosäureinsertion im Peptidprodukt.**

Betrachtet man diese beiden grundsätzlichen Möglichkeiten der Modulmanipulation bei nicht-ribosomalen Peptidsynthetasen so fällt auf, dass eine dritte Möglichkeit, nämlich die Modulinsertion (s. Abbildung 1-17c)), noch nicht realisiert wurde. Dies sollte theoretisch in der Einführung einer zusätzlichen Aminosäure in das Produkt resultieren. Diese Art von Experiment wurde bisher in der Literatur noch nicht beschrieben und ist deshalb Ziel der vorliegenden Arbeit.

## 2 Zielsetzung

Das grundlegende Ziel der vorliegenden Arbeit war die Komplettierung der bisher bei nicht-ribosomalen Peptidsynthetasen literaturbekannten Modulmanipulationen „Modulaustausch“ und „Moduldeletion“ durch das noch ausstehende Experiment der „Modulinsertion“ (s. Abbildung 1-17). Während die ersten beiden genetischen Manipulationen nachgewiesenermaßen zum Aminosäureaustausch bzw. zur Aminosäuredeletion im Peptidprodukt führten, sollte letzteres theoretisch eine Aminosäureinsertion im Peptidprodukt bewirken.

Da das nicht-ribosomal synthetisierte Glykopeptid Balhimycin und damit sein Produzenten-Stamm *Amycolatopsis balhimycina* seit Jahren in den Arbeitskreisen der Professoren Süßmuth und Wohlleben bearbeitet wird und dementsprechend sowohl die Sequenz des Biosyntheseegenclusters als auch viele Details der Biosynthese bekannt waren, bot sich dieses System geradezu für die geplanten Experimente an. Dazu kommt noch die besondere Herausforderung, dass es sich bei der Biosynthese von Balhimycin um ein sehr komplexes Beispiel einer nicht-ribosomalen Peptidsynthese handelt, vor allem aufgrund der vielen Modifikationen des Peptidgrundgerüsts wie Chlorierung, oxidative Seitenkettenzyklisierung, *N*-Methylierung und Glykosylierung. Ein derart komplexes System wurde bisher in der Literatur noch nie in Bezug auf genetische Manipulation der nicht-ribosomalen Peptidsynthese beschrieben.

Konkret war das Ziel der vorliegenden Arbeit zunächst die Ausarbeitung eines geeigneten molekularbiologischen Konzepts für eine Modulinsertion in einer nicht-ribosomalen Peptidsynthetase von Balhimycin in *Amycolatopsis balhimycina*. Dieses Konzept galt es anschließend praktisch umzusetzen sowie die erfolgreiche Durchführung mit molekularbiologischen und proteinbiochemischen Methoden nachzuweisen. Die zweite Hauptaufgabe umfasste die massenspektrometrische Charakterisierung der erzeugten *Amycolatopsis balhimycina*-Mutante mittels HPLC-ESI-MS und -MS/MS.

## 3 Grundlagen

### 3.1 Chromatographie

Das Prinzip der Chromatographie wurde 1903 erstmals von dem russischen Botaniker Mikhail S. Tswett angewendet, der farbige Petroletherextrakte von Pflanzen über  $\text{CaCO}_3$ -Säulen in verschiedene Farbstoffe trennte. Aufgrund der Tatsache, dass Russland zu dieser Zeit eher wissenschaftlicher Nebenschauplatz war, erlebte die Chromatographie weltweit erst in den 40er Jahren des 20. Jahrhunderts ihren Durchbruch. Ganz allgemein versteht man unter Chromatographie eine Trennmethode, bei der eine Substanzmischung mit Hilfe eines Gas- oder Flüssigkeitsstromes, der mobilen Phase, über eine stationäre Phase geleitet und dabei in ihre einzelnen Bestandteile aufgetrennt wird. Die Trennung basiert auf den individuellen Wechselwirkungen der Mischungskomponenten mit der stationären Phase, die zu unterschiedlichen Retentionen und damit zu der gewünschten Trennung führen.

Aus diesem Grundprinzip hat sich eine Vielzahl von Methoden entwickelt, die sich nach verschiedenen Kriterien einteilen lassen. Das gebräuchlichste Einteilungskriterium ist der Retentionsmechanismus. Grundsätzlich können so fünf verschiedene Chromatographiegrundtypen unterschieden werden. Bei der Adsorptionschromatographie ergibt sich eine Trennung dadurch, dass Verbindungen unterschiedlich stark an die feste, stationäre Phase adsorbiert und wieder desorbiert werden. Die Verteilungschromatographie nutzt dagegen eine flüssige, stationäre Phase, die sich nicht mit der mobilen Phase mischt. Die Verbindungen verteilen sich nun gemäß ihrer Eigenschaften zwischen diesen beiden Phasen und wandern umso langsamer, je besser sie sich in der stationären Phase lösen. Bei der Ionenaustauschchromatographie besitzt die stationäre Phase geladene Gruppen, die mit den zu trennenden Verbindungen Coulomb-Wechselwirkungen eingehen können. Die Affinitätschromatographie gründet auf spezifischen Wechselwirkungen, da sich Bestandteile der stationären Phase und die zu trennenden Verbindungen wie Ligand und Rezeptor verhalten. Die Retention bei der Größenausschlusschromatographie erklärt sich dadurch, dass die Verbindungen je nach Größe und Form unterschiedlich stark in die Poren der stationären Phase eindringen und sich so unterschiedlich lange der Elution durch die mobile Phase entziehen.

Die letztlich in der Praxis angewendeten Chromatographiearten lassen sich teilweise gut durch einen der eben aufgezeigten Grundtypen beschreiben, teilweise liegen aber auch zwei oder mehr Prinzipien gleichzeitig zugrunde.

In der vorliegenden Arbeit wurden zwei Chromatographiearten verwendet, zum einen die XAD-Adsorptionschromatographie (s. Abschnitt 4.2.17.1) und zum anderen die reverse-Phase Hochdruckflüssigchromatographie. Letztere soll im Folgenden ausführlich erläutert werden.

### **3.1.1 Reverse-Phase Chromatographie**

In den Anfängen der Chromatographie wurden vornehmlich polare stationäre Phasen wie z.B. Papier, Cellulose oder Kieselgel verwendet, deren Oberflächen sich durch die Anwesenheit von Hydroxygruppen auszeichnen. Das zu trennende Substanzgemisch wird durch eine mobile Phase über diese Materialien geleitet, die typischerweise aus einer der folgenden Lösungsmittelkombinationen besteht: Hexan/Ethylacetat, Chloroform/Methanol oder Toluol/Aceton. Die Zusammensetzung dieser Lösungsmittelkombinationen variiert dabei entweder mit der Zeit von unpolar zu polar, dann spricht man von einer Gradiententrennung, oder sie bleibt konstant, dann handelt es sich um eine isokratische Trennung. In jedem Fall eluieren in den frühen Fraktionen unpolare und in den späten Fraktionen polare Verbindungen, da polare Wechselwirkungen mit der stationären Phase eine Retention verursachen. Diese Art der Chromatographie wird als Normalphasen-chromatographie bezeichnet. Im Laufe der Zeit wurde allerdings festgestellt, dass bei manchen Trennproblemen die hydrophilen Hydroxygruppen zu starke Wechselwirkungen mit den Analyten eingingen, weshalb man diese mit langen Kohlenwasserstoffen veretherte und so unpolare stationäre Phasen erhielt. Dies war die Geburtsstunde der Umkehrphase- oder Reverse-Phase Chromatographie. Heutzutage gibt es stationäre Phasen mit unterschiedlichem Hydrophobizitätsgrad, deren Kohlenstoffketten zwischen C1 und C30 variieren können. Typischerweise verwendet man für Proteine C4- bzw. C8-Phasen und für Peptide C8- bzw. C18-Phasen. Die Elution erfolgt mit Wasser/Methanol- oder Wasser/Acetonitril-Gradienten, deren organischer Anteil mit der Zeit ansteigt, so dass hydrophile Moleküle in den frühen, vorwiegend wässrigen und hydrophobe Moleküle in den späten, vorwiegend organischen Fraktionen enthalten sind. Eine Besonderheit der Reverse-Phase Chromatographie ist der optionale Einsatz von Ionenpaar-Reagentien, die mit ihrem geladenen Molekülteil an

die ionischen Seitengruppen der zu trennenden Verbindungen binden und durch ihren ungeladenen Molekülteil deren Hydrophobizität ändern. Die entstandenen Ionenpaar-Komplexe zeigen ein anderes chromatographisches Verhalten, wodurch die Trennung verbessert werden kann. Bei Peptiden werden routinemäßig Trifluoressigsäure (TFA) oder Ameisensäure eingesetzt.

Die Zuordnung zu einem der oben genannten chromatographischen Grundtypen fällt bei der Normalphasenchromatographie leicht, da sie die typischen Merkmale der Adsorptionschromatographie zeigt. Im Gegensatz dazu ist die Zuordnung bei der Reverse-Phase Chromatographie vor allem bei besonders hydrophoben stationären Phasen nicht mehr eindeutig. Einerseits kann man die Interaktion mit den Kohlenstoffketten durch hydrophobe Wechselwirkungen beschreiben, andererseits existiert aber auch das Bild des „Eintauchens“ der Verbindungen zwischen die ausgestreckten Kohlenstoffreste, was vor allem bei langen Ketten gut vorstellbar ist. Folglich handelt es sich bei der Reverse-Phase Chromatographie um eine Mischform aus Adsorptions- und Verteilungschromatographie.

### **3.1.2 Hochdruckflüssigchromatographie (HPLC)**

Die Bezeichnung Hochdruckflüssigchromatographie beschreibt kein Chromatographieprinzip, sondern die instrumentelle Umsetzung der Chromatographie. Im Falle der Hochdruckflüssigchromatographie werden kompakte, kommerziell erhältliche Chromatographiestationen eingesetzt, um durch erhöhten Druck schnelle Trennungen mit hoher Auflösung zu erreichen. Aus diesem Grund spricht man auch oft von Hochleistungsflüssigchromatographie.

In der vorliegenden Arbeit wurden zwei verschiedene HPLC-Anlagen verwendet, zum einen eine präparative HPLC-Anlage (s. Abschnitt 4.2.17.2) und zum anderen eine Kapillar-HPLC-Anlage. Letztere wurde in Kopplung mit einem Massenspektrometer betrieben und soll deshalb im Folgenden kurz beschrieben werden.

### 3.1.3 Kapillar-Hochdruckflüssigchromatographie (Kapillar-HPLC)

Die in der vorliegenden Arbeit verwendete Kapillar-HPLC-Anlage von Agilent Technologies (1100 Serie) umfasst die folgenden fünf Module, die hauptsächlich über *fused silica*-Kapillaren (100 bzw. 50  $\mu\text{m}$  Durchmesser) verbunden sind:

- Stellplatz für Lösungsmittelflaschen
- Mikro-Vakuumentgaser
- binäre hochdruckseitige Kapillarpumpe mit vier Zugängen und 100  $\mu\text{l}$  Flusssensor
- Mikro-Wellplate-Probengeber mit 8  $\mu\text{l}$  Injektionsschleife und Probenhalter für 100x2 ml Gefäße
- Dioden-Array-Detektor mit 500 nl Flusszelle

Anhand des in Abbildung 3-1 gezeigten Flussdiagrammes soll die Funktionsweise dieser Kapillar-HPLC kurz beschrieben werden.

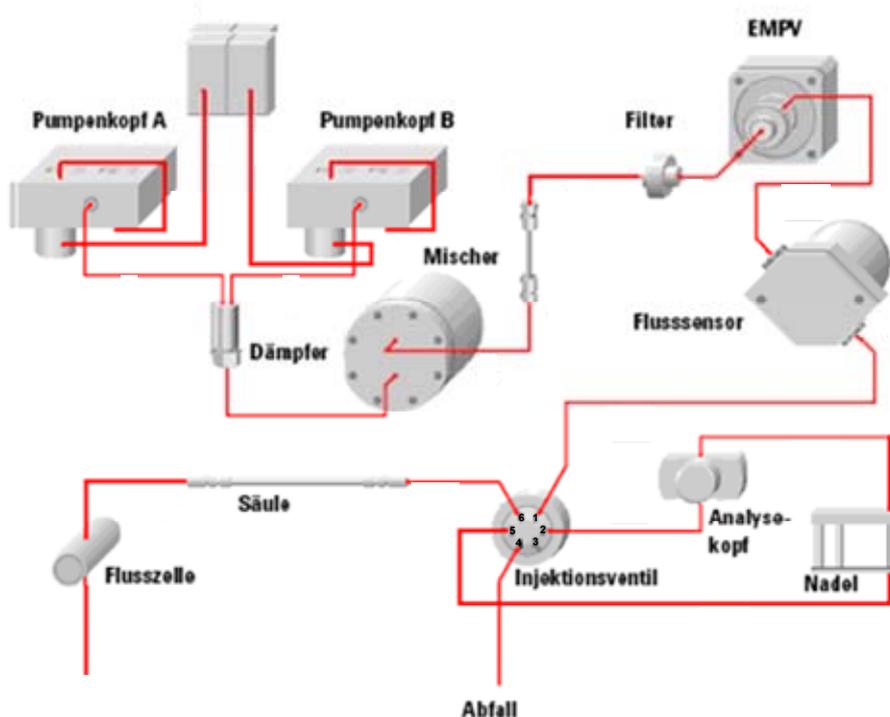


Abbildung 3-1: Flussdiagramm der Kapillar-HPLC.<sup>[71]</sup>

Die beiden für die Erstellung des mobile-Phase-Gradienten notwendigen Lösungsmittel A und B werden über zwei getrennte Pumpen A und B zunächst durch den Mikro-Vakuumentgaser gefördert und anschließend gemischt („hochdruckseitiges“ System). Neben dem Normalmodus, bei dem die Flussrate durch die Kolbenleistung bestimmt wird und deshalb nur Flussraten  $> 100 \mu\text{l}/\text{min}$  möglich sind, kann die Kapillarpumpe auch im sog. Mikromodus betrieben werden. In

diesem Modus wird von den Pumpen ein sog. Primärfluss erzeugt, von dem allerdings nur ein Teil auf die Säule gelangt, während der Rest abgesplittet wird. Wichtig hierfür sind die Baueinheiten EMPV (*electro-magnetic proportioning valve*) und Flusssensor, die das sog. EFC-(*electronic flow control*-)System bilden. Der Flusssensor misst die Wärmeleitfähigkeit und kann daraus sehr genau die Flussrate ableiten. Voraussetzung ist allerdings die Angabe der verwendeten Lösungsmittel (bei Analyst über den „*flow sensor calibration table index*“), da diese unterschiedliche Wärmekapazitäten besitzen. Der Flusssensor meldet sein Ergebnis an den vorgeschalteten EMPV, der dementsprechend den Absplittvorgang sehr fein regulieren kann. Dieses Prinzip erlaubt zum einen eine sehr präzise Förderung niedriger Flussraten ( $< 100 \mu\text{l}/\text{min}$ ) und zum anderen die schnelle Umsetzung von Gradienten. Wenn die Kapillar-HPLC im Normalmodus betrieben werden soll, kann man die Strecke über diese beiden Bauteile überbrücken. Die Probenauftragung erfolgt über ein Injektionsventil, wobei ein Greifarm die im *Autosampler* befindliche Probe abholt, die anschließend von einer Nadel aufgezogen und injiziert wird. Das Injektionsventil muss dafür zwei verschiedene Stellungen besitzen. Während der Probennahme selbst befindet sich das Injektionsventil in der sog. Nebenflussstellung, bei der das Lösungsmittel auf direktem Weg (Anschluss 1 -> 6) zur Säule geleitet wird und so ein Aufziehen der Probe durch die Nadel ermöglicht wird. Anschließend schaltet das Injektionsventil um in die sog. Injektionsstellung. Jetzt durchfließt das Lösungsmittel die Probenschleife (Anschluss 1 -> 2) und transportiert so die Probe zur Säule und anschließend zum Dioden-Array-Detektor. In dieser Stellung verbleibt das Injektionsventil bis zum Ende bzw. bis eine neue Probennahme erfolgt.

### 3.2 Massenspektrometrie (MS)

Die Massenspektrometrie geht auf Arbeiten von J. J. Thomson im Jahr 1910 zurück, in denen er zeigen konnte, dass das Edelgas Neon aus einer Mischung der zwei Isotope Ne-20 und Ne-22 besteht. Das Grundprinzip der Massenspektrometrie umfasst drei in Reihe geschaltete Prozesse. In der Ionenquelle werden aus den zu untersuchenden Verbindungen isolierte Ionen erzeugt, die anschließend im Analysator nach ihren Molekülmasse<sup>[72]</sup>-zu-Ladungs- $(m/z)$ -Verhältnissen getrennt und schließlich im Detektor erfasst werden. Das so generierte Massenspektrum trägt die Intensität der Ionen gegen die entsprechenden  $m/z$ -Verhältnisse auf. Die Anwendungen der Massenspektrometrie sind vielfältig. Aus der Bestimmung der Molekülmassen<sup>[72]</sup> von Verbindungen können bei genügend hoher Auflösung bzw. Massengenauigkeit sogar die Summenformeln abgeleitet werden. Außerdem können über Fragmentierungen der Verbindungen Informationen über deren Struktur gewonnen werden. Neben diesen qualitativen Anwendungen können auch Quantifizierungen mittels MS vorgenommen werden.

Das in der vorliegenden Arbeit verwendete Massenspektrometer besitzt eine ESI-Ionenquelle und einen QTrap-Analysator, der eine Weiterentwicklung des Triple-Quadrupol-Analysators ist. Die zugrunde liegenden Prinzipien dieser Bauteile sollen im Folgenden kurz beschrieben werden.

### 3.2.1 ESI-Ionenquelle

ESI ist die Abkürzung für Elektrospray-Ionisation und beschreibt eine sehr schonende Ionisierungsmethode, die zusammen mit MALDI (*matrix-assisted-laser-desorption/ionization*) in den 80er Jahren die Massenspektrometrie im Bereich biologischer Makromoleküle revolutioniert hat. Aus diesem Grund wurde 2002 der Nobelpreis für Chemie an die beiden Erfinder John B. Fenn und Koichi Tanaka verliehen.

Der Begriff Elektrospray kommt daher, dass mit Hilfe eines elektrostatischen Feldes eine Flüssigkeit in viele kleine, geladene Tröpfchen versprüht wird (s. Abbildung 3-2).

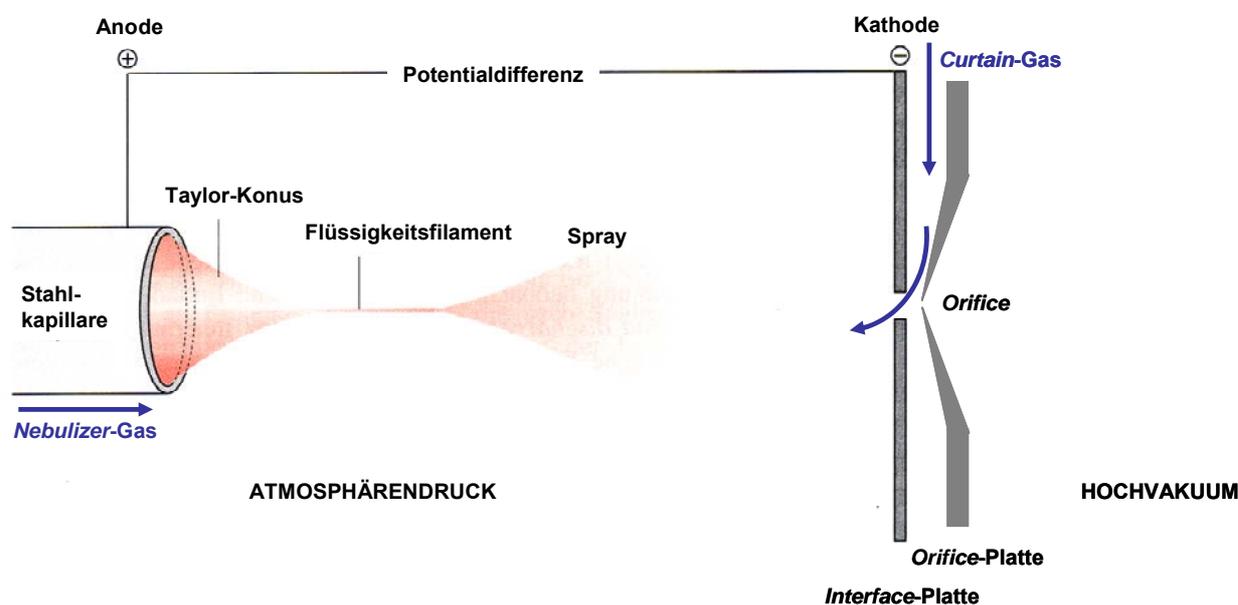


Abbildung 3-2: ESI-Prozess im Positivionen-Modus.<sup>[73]</sup>

Im Gegensatz zu den meisten anderen Ionisierungstechniken handelt es sich um eine Ionisation unter Atmosphärendruck. Die Probe wird dazu in einem verdampfbaren Lösungsmittel mit einem die Ionisierung unterstützenden Additiv z.B. Methanol mit Ameisensäure (0,1 %) und als kontinuierlicher Flüssigkeitsstrom in eine Stahlkapillare eingeführt. Zwischen der Spitze dieser Stahlkapillare und dem Eingang zum Analysator des Massenspektrometers liegt eine Spannung an. Je nach Spannungspolung werden so Tröpfchen mit positiven oder negativen Ladungen auf der Tröpfchenoberfläche erzeugt und in Richtung des Analysators beschleunigt. Unterstützt wird dieser Sprayprozess durch koaxial zur Stahlkapillare ausströmenden Stickstoff (*Nebulizer-Gas*), wobei Scherkräfte zur effizienten Zerstäubung beitragen. Die allgemeine Vorstellung zur Entstehung freier Ionen aus geladenen Tröpfchen ist, dass das Lösungsmittel nach und nach aus den Tröpfchen verdunstet und diese so

eine kritische Größe erreichen, bei der die abstoßenden Coulombkräfte zwischen den Oberflächenladungen nicht mehr durch die Oberflächenspannung kompensiert werden können. Die Folge ist eine Reihe von sog. Coulomb-Explosionen, die zu immer kleineren Tröpfchen und schließlich zu freien Ionen führen, die abhängig von Größe und Art der Verbindung auch mehrfach geladen sein können. Der Prozess ist so schonend, dass nahezu keine Fragmentierungen auftreten und daher große Biomoleküle ohne Probleme als Molekül-Ionen detektiert werden können. Wichtig bei dieser Art der Ionisation ist die Realisierung des Übergangs zwischen Atmosphärendruck in der Ionenquelle und Hochvakuum (Größenordnung:  $10^{-5}$  torr bzw. mbar) im Analysator. Dieser Übergang wird auch als Interface bezeichnet und umfasst zum einen eine Mikroöffnung (*orifice*) und zum anderen das sog. *Curtain-Gas* (Stickstoff), welches zwischen der *orifice*-Platte und der davor liegenden *Interface*-Platte in den Ionisierungsraum strömt. Dadurch werden Kollisionen mit dem Elektrospray provoziert, was zum einen das ungewollte Eintreten von Neutralteilchen in den Analysator verhindert und zum anderen die Desolvatisierung der Ionen fördert. Das in dieser Arbeit verwendete Massenspektrometer besitzt eine sog. *TurbolonSpray*-Quelle, die als Besonderheit neben dem *Nebulizer*-Gas und dem *Curtain*-Gas einen weiteren Gasstrom besitzt, der im  $90^\circ$ -Winkel auf den Flüssigkeitsstrom trifft und so die Desolvatisierung der Ionen auch bei relativ hohen Flussraten ermöglicht.

### 3.2.2 Triple-Quadrupol-Analysator

Ein Quadrupol-Analysator besteht aus vier parallelen stabförmigen Metallelektroden, die kreisförmig um eine z-Achse angeordnet sind und durch Anlegen eines kombinierten Wechsel- und Gleichspannungsfeldes als sog. Massenfilter dienen. Dies bedeutet, dass sie zu einem bestimmten Zeitpunkt nur Ionen mit einem bestimmten  $m/z$ -Verhältnis auf einer stabilen oszillierenden Bahn zum Detektor passieren lassen. Dieses Prinzip wurde 1953 von W. Paul und H. Steinwedel entwickelt und soll im Folgenden beschrieben werden.

Die Spannungsfelder liegen so am Quadrupol an, dass gegenüberliegende Stäbe die gleiche Polarität der Gleichspannung ( $U$ ) und die gleiche Phase der Wechselspannung ( $V \cdot \cos(2\pi ft)$ ) besitzen, während nebeneinander liegende Stäbe entgegengesetzte Polarität der Gleichspannung ( $-U$ ) und eine um  $180^\circ$  versetzte

Phase der Wechselspannung ( $-V \cdot \cos(2\pi ft)$ ) besitzen. Dadurch entsteht in der Nähe der z-Achse folgendes elektrisches Potential:

$$\phi(x, y, t) = (U + V \cdot \cos(2\pi ft)) \cdot \frac{x^2 - y^2}{r^2}$$

Mithilfe einer geringen Beschleunigungsspannung werden die Ionen in Richtung der z-Achse in das elektrische Feld des Quadrupols geleitet. Dadurch werden Bewegungen in der xy-Ebene hervorgerufen, die durch die folgenden Bewegungsgleichungen beschrieben werden können:

$$F_x = ma = m \frac{d^2x}{dt^2} = ze \frac{\partial \Phi}{\partial x} = \frac{z \cdot e}{r^2} (U + V \cos(2\pi ft)) \cdot x$$

$$F_y = ma = m \frac{d^2y}{dt^2} = ze \frac{\partial \Phi}{\partial y} = \frac{z \cdot e}{r^2} (U + V \cos(2\pi ft)) \cdot y$$

$m$  = Molekülmasse des Ions

$e$  = Elementarladung

$z$  = Anzahl der Ladungen

Diese Bewegungsgleichungen können als Mathiesche lineare Differentialgleichungen aufgefasst werden, wenn man folgende Substitutionen vornimmt:

$$a = \frac{2zeU}{m(\pi fr)^2} \quad \text{und} \quad q = \frac{zeV}{m(\pi fr)^2}$$

Die Parameter  $a$  und  $q$  beschreiben die Beziehung zwischen dem Ion mit den Eigenschaften  $m$ ,  $z$  und  $e$  und dem Quadrupol mit den Eigenschaften  $r$ ,  $U$ ,  $V$  und  $f$ . Damit sehen die Bewegungsgleichungen wie folgt aus:

$$\frac{d^2x}{d(\pi ft)^2} + (a + 2q \cos 2\pi ft)x = 0$$

$$\frac{d^2y}{d(\pi ft)^2} - (a + 2q \cos 2\pi ft)y = 0$$

Es gibt zwei Arten von Lösungen für die Mathieschen Gleichungen, wobei einmal eine stabile Bewegung durch den Quadrupol mit einer endlichen Amplitude beschrieben wird und einmal eine Bewegung mit in x- und/oder y-Richtung exponentiell anwachsender Amplitude. Für Ionen mit einem bestimmten  $m/z$ -Verhältnis gibt es bestimmte Werte für die Parameter  $a$  und  $q$ , bei denen stabile Bewegungen möglich sind. Dies lässt sich in einem sog. Stabilitätsdiagramm darstellen, wie es in Abbildung 3-3 gezeigt ist.

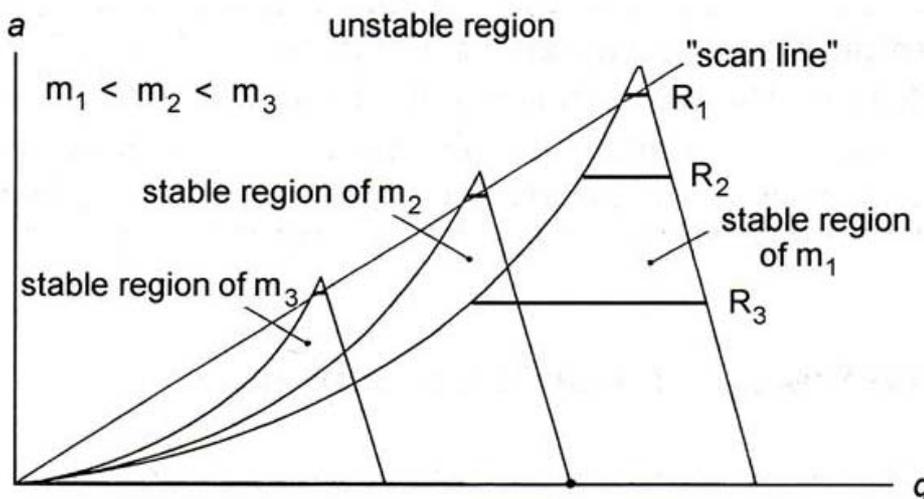


Abbildung 3-3: Stabilitätsdiagramm für das zweidimensionale Quadrupolfeld in x- und y-Richtung.<sup>[74]</sup>

Aus den oben genannten Gleichungen für  $a$  und  $q$  ergibt sich, dass das Verhältnis  $a/q$  immer gleich  $2U/V$  ist:

$$\frac{a}{q} = \frac{2zeU}{m(\pi fr)^2} \cdot \frac{m(\pi fr)^2}{zeV} = \frac{2U}{V}$$

In der Praxis werden zum *Scannen* eines Massenbereichs die Gleichspannung  $U$  und die Wechselspannungsamplitude  $V$  so erhöht, dass ihr Verhältnis  $U/V$  und damit  $a/q$  konstant bleibt. Dies entspricht einem *Scannen* entlang der im Stabilitätsdiagramm gezeigten Arbeitsgeraden bzw. „scan line“, wobei die Auflösung  $m/\Delta m$  von der Steigung dieser Geraden abhängt. Ist die Steigung Null, d.h. wird nur eine Wechselspannung angelegt, können alle  $m/z$ -Verhältnisse stabil den Quadrupol durchqueren. Theoretisch ist die Auflösung maximal, wenn die Arbeitsgerade die Spitze der Bereiche stabiler Oszillationen berühren würde. Allerdings muss in der Praxis von Anfangsgeschwindigkeiten der Ionen in x- und y-Richtung sowie von

Abweichungen von der idealen z-Richtung beim Eintritt in den Quadrupol ausgegangen werden, was die Empfindlichkeit extrem herabsetzt. Aus diesem Grund wählt man die Arbeitsgerade etwas flacher, um neben einer ansprechenden Auflösung auch eine akzeptable Empfindlichkeit zu erreichen.

Neben dieser mathematischen Betrachtungsweise gibt es auch noch eine etwas anschaulichere Erklärung für die Massenfilterfunktion eines Quadrupols. Durch Anlegen der Wechselspannungen an jeweils zwei gegenüberliegende Stäbe bauen sich abwechselnd positive und negative Felder relativ zur Mittelachse auf. Im Positivionen-Modus werden die Ionen deshalb während der positiven Phase zur Mittelachse und während der negativen zu den Stäben beschleunigt. An eines der beiden Paare gegenüberliegender Quadrupolstäbe wird zusätzlich eine positive Gleichspannung angelegt, an das andere eine negative Gleichspannung. Bei schweren Ionen überwiegt der Einfluss der Gleichspannung, so dass das Stabpaar mit der negativen Gleichspannung zum Aussortieren zu großer Massen dient. Diese werden von der negativen Gleichspannung so stark zu den Stäben abgelenkt, dass das überlagerte positive Feld der Wechselspannung nicht ausreicht, um sie wieder Richtung Mittelachse zu beschleunigen. Bei leichten Ionen sind die Verhältnisse umgekehrt, so dass das Stabpaar mit der positiven Gleichspannung zum Aussortieren zu kleiner Massen dient. Diese werden vom negativen Feld der Wechselspannung so stark zu den Stäben abgelenkt, dass die überlagerte positive Gleichspannung nicht ausreicht, um sie wieder Richtung Mittelachse zu beschleunigen.

Ein Triple-Quadrupol-Analysator zeigt den in Abbildung 3-4 dargestellten Aufbau.

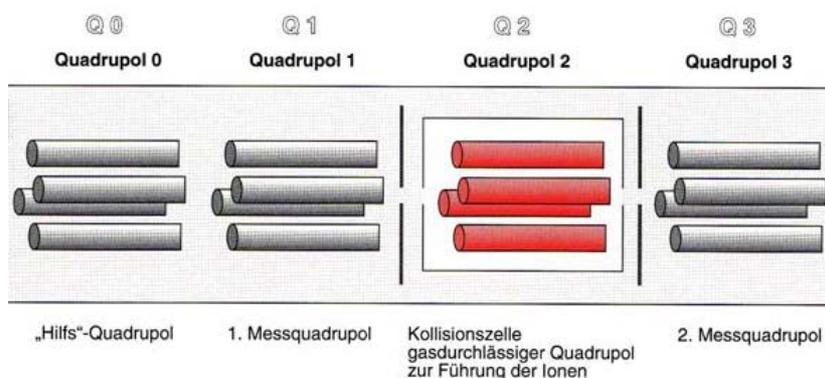


Abbildung 3-4: Schematischer Aufbau eines Triple-Quadrupol-Analysators.<sup>[75]</sup>

Dieser Analysatortyp besteht aus vier Quadrupolen, wobei nur Q1 und Q3 als Messquadrupole verwendet werden. Q0 ist ein sog. „Hilfs“-Quadrupol, der nur mit Wechselspannung betrieben wird und deshalb von allen  $m/z$ -Verhältnissen stabil durchquert werden kann. Seine Aufgabe ist die Fokussierung der aus der Ionenquelle kommenden Ionen in die Mitte des Ionenpfades. Q2 ist ein mit  $N_2$  gefüllter Quadrupol, der je nach Experiment als Kollisionszelle oder in Analogie zu Q0 als „fokussierender Transportquadrupol“ fungiert. Bei seiner Funktion als Kollisionszelle findet eine Fragmentierung der Ionen aufgrund von Stößen mit inerten Gasatomen statt. Die Fragmentierungseffizienz kann durch Parameter wie kinetische Energie der Ionen und Gasdruck in der Kollisionskammer gesteuert werden.

Es gibt drei typische Experimente, die mit einem Triple-Quadrupol-Analysator durchgeführt werden können (s. Abbildung 3-5).

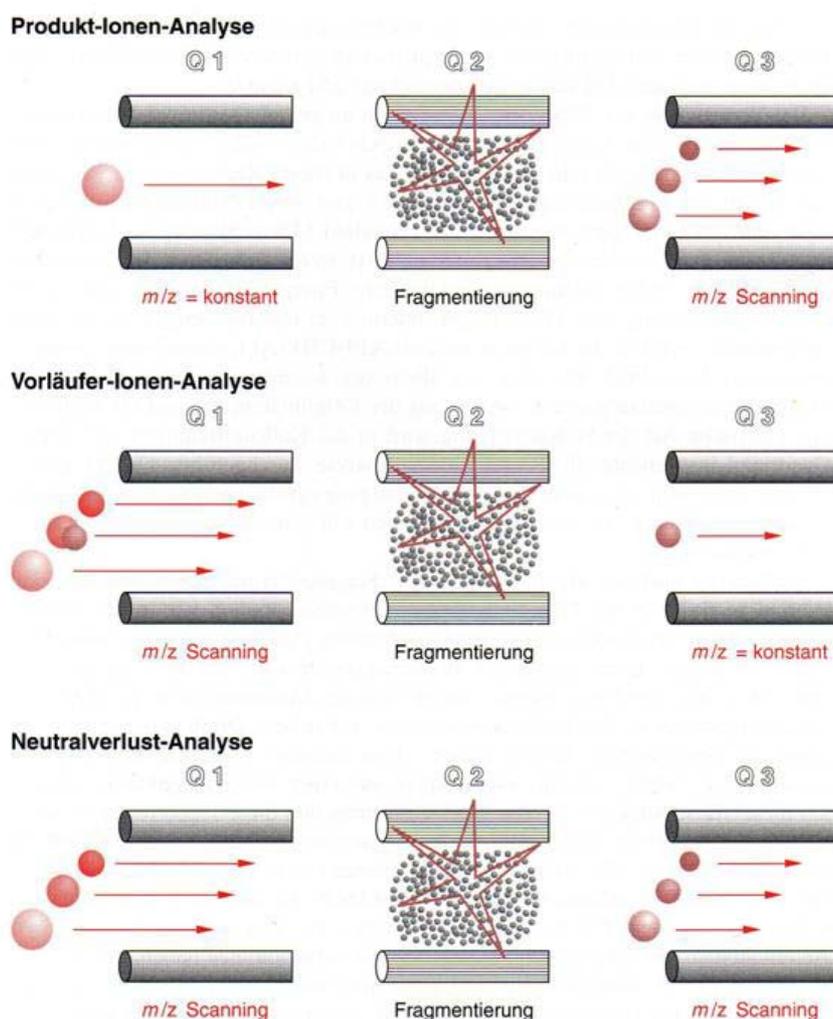


Abbildung 3-5: Drei typische Experimente eines Triple-Quadrupol-Analysators.<sup>[76]</sup>

Bei der Produkt-Ionen-Analyse wird in Q1 ein bestimmtes  $m/z$ -Verhältnis selektiert. Dieses wird in Q2 fragmentiert und die dabei entstandenen Fragment-Ionen in Q3 analysiert. Das erhaltene Produkt-Ionen-Spektrum gibt Hinweise für die Strukturaufklärung der in Q1 selektierten Verbindung. Eine andere Aufgabenstellung ist z.B. die Analyse einer Probe mit dem Ziel, Vertreter einer bestimmten Verbindungsklasse zu identifizieren. Dabei nutzt man aus, dass Vertreter einer Verbindungsklasse gemeinsame Struktur motive besitzen, die bei einer Fragmentierung entweder typische Fragment-Ionen oder typische Neutralverluste erzeugen. Im ersten Fall bedient man sich der sog. Vorläufer-Ionen-Analyse, bei der Q1 einen gewissen Massenbereich durchscannt und diese  $m/z$ -Verhältnisse zeitlich versetzt fragmentiert. Q3 ist allerdings nur auf das  $m/z$ -Verhältnis des typischen Fragment-Ions eingestellt, so dass das erhaltene Vorläufer-Ionen-Spektrum nur  $m/z$ -Verhältnisse abbildet, die nach der Fragmentierung dieses typische Fragment-Ion ergeben haben. In Analogie dazu verläuft die Neutralverlust-Analyse, bei der Q1 und Q3 parallel und um den typischen Neutralverlust versetzt scannen, so dass das erhaltene Neutralverlust-Spektrum nur  $m/z$ -Verhältnisse abbildet, die nach der Fragmentierung um den typischen Neutralverlust kleinere Fragment-Ionen ergeben haben.

Neben diesen drei qualitativen Experimenten gibt es noch die sog. MRM-(*multiple-reaction-monitoring*)Analyse, die zur Quantifizierung eingesetzt wird. Hierbei sind die beiden Messquadrupole Q1 und Q3 auf bestimmte  $m/z$ -Verhältnisse eingestellt, die bei Q1 dem  $m/z$  der zu quantifizierenden Verbindung und bei Q3 dem  $m/z$  eines intensiven und möglichst spezifischen Tochter-Ions dieser Verbindung entsprechen. Je nachdem, wie viele Verbindungen gleichzeitig quantifiziert werden sollen, kann man unterschiedlich viele solcher Ionenpaare angeben. Diese Art der Analyse ist extrem empfindlich, da die Messzeit ausschließlich auf das interessierende bzw. die interessierenden  $m/z$ -Verhältnisse verwendet wird. Überdies ist sie auch sehr spezifisch, da zwei Auswahlkriterien herangezogen werden, nämlich  $m/z$  des Molekül-Ions und  $m/z$  eines möglichst spezifischen Tochter-Ions.

### 3.2.3 QTrap-Analysator

Der QTrap-Analysator ist eine Weiterentwicklung des Triple-Quadrupol-Analysators und wurde von der Firma Applied Biosystems patentiert. Die Besonderheit an diesem Analysatortyp ist, dass Q3 auch als sog. lineare Ionenfalle verwendet werden kann, woraus sich einige Vorteile ergeben. Bei einem Triple-Quadrupol hat man für eine MS-Analyse die Möglichkeit einer Q1-Analyse, bei der Q1 *scannt* und Q3 nur mit Wechsellspannung betrieben wird und so Ionen aller  $m/z$ -Verhältnisse passieren lässt oder einer Q3-Analyse, die analog abläuft. Bei der QTrap kann man eine sog. *Enhanced* MS-Analyse (EMS-Analyse) durchführen, bei der Q1 alle Ionen passieren lässt und Q3 als lineare Ionenfalle alle  $m/z$  zunächst sammelt und dann ausliest. Diese MS-Analyse hat eine viel größere Auflösung und Empfindlichkeit als eine Quadrupol-Analyse. Diese Vorteile kann man auch bei der sog. *Enhanced* Produkt-Ionen-Analyse (EPI-Analyse) nutzen, bei der die in Q2 erzeugten Fragment-Ionen im als lineare Ionenfalle arbeitenden Q3 zunächst gesammelt und dann ausgelesen werden. Neben diesen beiden verbesserten Versionen der MS-Analyse und Produkt-Ionen-Analyse gibt es eine Reihe von Experimenten, die in einem gewöhnlichen Triple-Quadrupol-Analysator gar nicht durchführbar sind. Ein Beispiel hierfür ist die  $MS^3$ -Analyse, bei der eines der in Q2 gebildeten Fragmente im als lineare Ionenfalle arbeitenden Q3 isoliert und erneut fragmentiert wird. Die dabei entstehenden Fragmente werden zunächst gesammelt und dann ausgelesen. Diese Funktion ist für Strukturaufklärungszwecke sehr hilfreich. Weitere mögliche Experimente wurden in der vorliegenden Arbeit nicht genutzt und sollen deshalb nur aufgezählt werden: *Enhanced Multiply Charged*-Analyse (EMC-Analyse), *Enhanced Resolution*-Analyse (ER-Analyse) und *Time Delayed Fragmentation*-Analyse (TDF-Analyse).

Neben diesen zusätzlichen Experimenten gibt es zwei weitere Besonderheiten des QTrap-Analysators. Bei der Verwendung von Q3 als lineare Ionenfalle ist die Füllzeit der Ionenfalle ein wichtiger Parameter für die Güte der Messung. Deshalb wurde die Möglichkeit entwickelt, diese Füllzeit dynamisch an den momentanen Ionenstrom aus der Quelle anzupassen (*dynamic fill time*, DFT). Dieser Ionenstrom wird durch eine sehr schnelle Quadrupol-Analyse erfasst und dementsprechend eine passende Füllzeit gewählt. Die zweite Besonderheit wird *Q0-Trapping* genannt und bedeutet, dass die von der Ionenquelle kommenden Ionen während der *Scanzeit* der linearen Ionenfalle nicht wie sonst verloren gehen, sondern von Q0 gesammelt werden. Dies

führt zu einer deutlich verbesserten Empfindlichkeit, was vor allem bei gering konzentrierten Analyten entscheidend ist.

### **3.3 Kapillar-HPLC-ESI-QTrap-Kopplung**

In der vorliegenden Arbeit wurde eine HPLC-MS-Kopplung verwendet, bei der die in Abschnitt 3.1.3 beschriebene Kapillar-HPLC-Anlage von Agilent Technologies (1100 Serie) mit der in den Abschnitten 3.2.1 und 3.2.3 beschriebenen ESI-QTrap2000 von Applied Biosystems verbunden ist. Dieses System wurde über die Analyst-Software des Massenspektrometers angesteuert, wobei die verwendeten Methoden in den Abschnitten 4.2.18 und 4.2.19 beschrieben werden. Die auf einem HPLC-MS-System gemessenen Analysen umfassen eine ganze Reihe von Massenspektren, die über die Zeit aufgenommen wurden. Darstellerisch wird dies in einem Chromatogramm, dem sog. Totalionenstrom, wiedergegeben, der die Intensitäten aller  $m/z$ -Verhältnisse gegen die Zeit aufträgt. Um die Auswertung zu erleichtern, gibt es die Funktion des extrahierten Ionenstromes. Dieses Chromatogramm trägt die Intensität eines gewünschten  $m/z$ -Verhältnisses gegen die Zeit auf, so dass man leicht erkennen kann, zu welchem Zeitpunkt die gesuchte Verbindung von der Säule eluiert.

## 4 Material und Methoden

### 4.1 Material

#### 4.1.1 Geräte

-80°C-Schrank	Thermo Electron GmbH, Dreieich
Brutschrank	BD 53; Binder, Tuttlingen
Eismaschine	AF 80 AS; Scotsman Europe, Mailand, Italien
<i>French Press</i>	SIM-AMINCO; SLM Instruments, Inc. Rochester, New York, USA
<i>French Press Zelle</i>	40 K Cell (40000 psi, 35 ml); SIM-AMINCO; SLM Instruments, Inc. Rochester, New York, USA
Geldokumentationsanlage	GelSystem Flexi 5020 S; biostep GmbH, Jahnsdorf
Heizblock	Thermomixer comfort; Eppendorf AG, Hamburg
Horizontalkammer für Agarose-gelelektrophorese	Mini-Sub Cell GT; Bio-Rad Laboratories GmbH, München
HPLC-Anlagen	1100-Series; Agilent Technologies Deutschland GmbH, Waldbronn
HPLC-Säulen	Für Kapillar-HPLC: Jupiter 4u Proteo 90A (Vorsäule 30x1 mm, Hauptsäule 150x1 mm); Phenomenex, Aschaffenburg Für Präparativ-HPLC: C18 Grom-Sil 300 ODS-5 (250x20 mm); Grom, Herrenberg
Hybridisierungsöfen	Hybaid Maxi 14; Thermo Electron GmbH, Dreieich
Massenspektrometer	QTRAP 2000; Applied Biosystems Applera Deutschland GmbH, Darmstadt
Medical Film Processor QX-150U	Konica
PCR-Maschine	Primus25 advanced und Primus 96 advanced Gradient; PEQLAB Biotechnologie GmbH, Erlangen
Peristaltikpumpe	Reglo-Analog MS-2/6-160; ISMATEC Laboratoriumstechnik GmbH, Wertheim-Mondfeld
pH-Meter	pH/Ion 510; Eutech Instruments Pte Ltd, Singapur
Photometer	Ultrospec 2100 pro Classic; Amersham Pharmacia Biotech, Freiburg
Schüttelinkubator	MULTITRON II; Infors GmbH, Einsbach
Spannungsquellen	PowerPac Basic, PowerPac HC und PowerPac 1000; Bio-Rad Laboratories GmbH, München

Sterilbank	Holten Safe 2010; Thermo Electron, Dreieich
Taumeltisch	Polymax 1040 Neigungswinkel 5°; Thermo Electron, Dreieich
Vakuumbiotapparatur	Pharmacia, Freiburg
Vertikalkammer für Polyacrylamidgelelektrophorese	Mini-Protean 3; Bio-Rad Laboratories GmbH, München
Vortex	Reax top; Heidolph, Schwabach
Waagen	EW 6000-1 M; Kern & Sohn GmbH, Balingen XR 125SM; Precisa Instruments AG, Dietikon, Schweiz
Zentrifugen	Mikroliterzentrifuge Z233MK-2 mit Winkelrotor 220.87 V05; HERMLE GmbH, Wehingen Tischkühlzentrifuge Z383K mit Ausschwingrotor 221.08 V01 und Winkelrotor 220.78 V02; HERMLE GmbH, Wehingen

#### 4.1.2 Chemikalien

1 kB DNA-Leiter, GeneRuler	Fermentas, St. Leon-Rot
2-log DNA-Leiter, TriDye	New England Biolabs GmbH, Frankfurt am Main
Acetonitril	LC-MS grade; Carl Roth GmbH & Co, Karlsruhe
Acrylamid/Bisacrylamid-Lösung (30 % / 0,8 %)	Carl Roth GmbH & Co, Karlsruhe
Agar-Agar	Carl Roth GmbH & Co, Karlsruhe
Agarose Broad Range	Carl Roth GmbH & Co, Karlsruhe
Albumin Fraktion V	Carl Roth GmbH & Co, Karlsruhe
Ameisensäure	Carl Roth GmbH & Co, Karlsruhe
Ammoniumsulfat	Carl Roth GmbH & Co, Karlsruhe
Ampicillin, Natriumsalz	Carl Roth GmbH & Co, Karlsruhe
APS	Carl Roth GmbH & Co, Karlsruhe
Bacto Pepton	Becton, Dickinson and Company, Sparks, USA
Benzamidin	Sigma-Aldrich Chemie GmbH, München
Bromphenolblau, Natriumsalz	Carl Roth GmbH & Co, Karlsruhe
CaCl <sub>2</sub> ·2H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
Casamino Acids	Becton, Dickinson and Company, Sparks, USA

Coomassie Brillantblau G 250	Carl Roth GmbH & Co, Karlsruhe
CsCl	Sigma-Aldrich Chemie GmbH, München
CuCl <sub>2</sub>	Carl Roth GmbH & Co, Karlsruhe
DMSO, für die Molekularbiologie, ≥ 99,9 %	Sigma-Aldrich Chemie GmbH, München
DNA-Längenstandards VII, Digoxigenin-markiert	Roche Diagnostics GmbH, Mannheim
dNTP-Mix	New England Biolabs GmbH, Frankfurt am Main
DTT	Carl Roth GmbH & Co, Karlsruhe
EDTA	Carl Roth GmbH & Co, Karlsruhe
Eisessig	Carl Roth GmbH & Co, Karlsruhe
Erythromycin, Base	Serva, Heidelberg
Ethidiumbromid	Carl Roth GmbH & Co, Karlsruhe
FeCl <sub>3</sub> ·6H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
Ficoll 400	Serva, Heidelberg
Glukose	Carl Roth GmbH & Co, Karlsruhe
Glycin	Carl Roth GmbH & Co, Karlsruhe
HCl	Carl Roth GmbH & Co, Karlsruhe
Hefeextrakt	Carl Roth GmbH & Co, Karlsruhe
IPTG	Carl Roth GmbH & Co, Karlsruhe
Isopropanol	Carl Roth GmbH & Co, Karlsruhe
K <sub>2</sub> HPO <sub>4</sub> ·3H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
K <sub>2</sub> SO <sub>4</sub>	Carl Roth GmbH & Co, Karlsruhe
Kalbsthymus-DNA	Amersham Bioscience, Freiburg
Kaliumacetat	Carl Roth GmbH & Co, Karlsruhe
KH <sub>2</sub> PO <sub>4</sub>	Carl Roth GmbH & Co, Karlsruhe
Maleinsäure	Carl Roth GmbH & Co, Karlsruhe
Mannitol	Carl Roth GmbH & Co, Karlsruhe
Methanol	Hochschullieferung; Biesterfeld Chemiedistribution GmbH, Hamburg
MgCl <sub>2</sub> ·6H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
MnCl <sub>2</sub> ·4H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
Na <sub>2</sub> B <sub>4</sub> O <sub>7</sub> ·10H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
NaCl	Carl Roth GmbH & Co, Karlsruhe
NaOH	Carl Roth GmbH & Co, Karlsruhe
Natriumacetat	Carl Roth GmbH & Co, Karlsruhe
(NH <sub>4</sub> ) <sub>6</sub> Mo <sub>7</sub> O <sub>24</sub> ·4H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
N-Lauroylsarkosin (Na-Salz, 35 % Lösung)	Serva, Heidelberg

Nutrient Broth No. 1	Fluka, Buchs
PEG1000	Carl Roth GmbH & Co, Karlsruhe
Phosphorsäure	Carl Roth GmbH & Co, Karlsruhe
PMSF	Carl Roth GmbH & Co, Karlsruhe
Polymin P	BASF, Ludwigshafen
Prolin (L)	Carl Roth GmbH & Co, Karlsruhe
Protein-Leiter, <i>prestained</i> , PageRuler	Fermentas GmbH, St. Leon-Roth
Saccharose	Südzucker, Mannheim
Sephadex G50, 50-150 µm Partikelgröße <i>pore size 500-10,000 fractionation range: M.W.</i>	Pharmacia Fine Chemicals, Uppsala, Schweden
SDS	Carl Roth GmbH & Co, Karlsruhe
TEMED	Carl Roth GmbH & Co, Karlsruhe
TES	Carl Roth GmbH & Co, Karlsruhe
TFA	Carl Roth GmbH & Co, Karlsruhe
tri-Natriumcitratx2H <sub>2</sub> O	Carl Roth GmbH & Co, Karlsruhe
Tris	Carl Roth GmbH & Co, Karlsruhe
Triton X 100	Carl Roth GmbH & Co, Karlsruhe
Trypton/Pepton aus Casein	Carl Roth GmbH & Co, Karlsruhe
Tween 20	Carl Roth GmbH & Co, Karlsruhe
Wasser	LC-MS grade; Carl Roth GmbH & Co, Karlsruhe
XAD1180	Amberlite; Rohm und Haas Deutschland, Frankfurt
X-Gal	Carl Roth GmbH & Co, Karlsruhe
ZnCl <sub>2</sub>	Carl Roth GmbH & Co, Karlsruhe

### 4.1.3 Bakterienstämme

*E. coli* XL1 Blue<sup>[77]</sup> exprimiert ω-Fragment für Blau-Weiß-Selektion: lacZΔM15  
*E. coli* ET12567<sup>[78]</sup> methylierungsdefizient: *dam*<sup>-</sup> *dcm*<sup>-</sup> *hsdM*  
 Anmerkung: Chloramphenicol bei Kompetenzmachung zugeben

*A. baohimycina*<sup>[21,22]</sup>

*A. baohimycina* C5A5T4E4-Mutante diese Arbeit

#### 4.1.4 Kits

<i>A-Addition-Kit</i>	Qiagen, Hilden
<i>DIG Luminescent Detection-Kit</i>	Roche Diagnostics GmbH, Mannheim
<i>NucleoSpin Tissue-Kit</i>	MACHERY-NAGEL GmbH, Düren
<i>PCR DIG labeling-Mix</i>	Roche Diagnostics GmbH, Mannheim
<i>peqGOLD Plasmid Miniprep-Kit I</i>	PEQLAB Biotechnologie GmbH, Erlangen
<i>pGEM-T-Easy-Vektorsystem</i>	Promega, Mannheim
<i>ProofStart DNA-Polymerase</i>	Qiagen, Hilden
<i>SilverQuest Silver Staining-Kit</i>	Invitrogen, Karlsruhe
<i>Taq DNA-Polymerase</i>	PEQLAB Biotechnologie GmbH, Erlangen
<i>QIAEX II Gelextraktions-Kit</i>	Qiagen, Hilden
<i>QIAGEN PCR Klonierungs-Kit (pDrive)</i>	Qiagen, Hilden
<i>QIAquick Gelextraktions-Kit</i>	Qiagen, Hilden

#### 4.1.5 Enzyme

<i>DNaseI</i>	Sigma-Aldrich Chemie GmbH, München
<i>Lysozym</i>	Serva, Heidelberg
<i>Restriktionsenzyme</i>	Fermentas, St. Leon-Rot
	New England Biolabs GmbH, Frankfurt am Main
<i>RNase A</i>	Fermentas, St. Leon-Rot
<i>T4 DNA-Ligase</i>	New England Biolabs GmbH, Frankfurt am Main

#### 4.1.6 Oligonukleotide

Bezeichnung	Sequenz (5' -> 3')	PCR-Produkt
<i>Insleft5' -MunI BmtI</i>	AACAATTGGCTAGCGAGGAC ACCGCGTGGCAGCCCAAC	<i>Insleft</i>
<i>Insleft3' -HindIII XbaI</i>	AAAGCTTCTAGAGGCCGC CTTCCCGGCGAAC	<i>Insleft</i>
<i>Insright5' -MunI XbaI</i>	AACAATTGTCTAGAGAGCCG GTCACCGAGG	<i>Insright</i>
<i>Insright3' -HindIII BmtI</i>	ATAAGCTTGTAGCGAGGAC AGTCCGGCCGAGGTAG	<i>Insright</i>
<i>Tarleft5' -EcoRI</i>	AAGAATTCGTCGGGTGCG GCACCTCTAC	<i>Tarleft</i>

<i>Tarleft</i> 3'- <i>HindIII</i> <i>BmtI</i>	ATAAGCTTGGCTAGCGAGGAC AGTCCGGCCGAGGTAG	<i>Tarleft</i>
<i>Tarright</i> 5'- <i>MunI</i> <i>BmtI</i>	AACAATTGGCTAGCGAGGAC ACCGCGTGGCAGCCCAAC	<i>Tarright</i>
<i>Tarright</i> 3'- <i>XbaI</i> <i>NcoI</i>	AATCTAGACCATGGCCTTGC GGGTCGCTTCCACGCAGAC	<i>Tarright</i>
ermE links	GCACGCCTGGTCGATGTC	„Erythromycinresistenz“
ermE rechts	CTTCTCCCGCAACGACTT	„Erythromycinresistenz“
Modulininsertion links	ACGGTTACCTGGGCAACAC	„Modulininsertion“
Modulininsertion rechts	GTCAGCGAGATCCCTTC	„Modulininsertion“
DIG-Sonde links	GATCGGCACTGGAATC	DIG-Sonde
DIG-Sonde rechts	GTCAGCGAGATCCCTTC	DIG-Sonde

Alle Oligonukleotide wurden von biomers.net GmbH, Ulm bezogen und vor dem Gebrauch im angegebenen Volumen H<sub>2</sub>O<sub>dd, steril</sub> in einer Konzentration von 100 pmol/μl gelöst.

#### 4.1.7 Medien

Für die Kultivierung von *E. coli*:

TY-Medium:	8 g	Trypton/Pepton aus Casein
	5 g	Hefeextrakt
	5 g	NaCl

Für die Kultivierung von *A. baohimycina*:

R5:	103 g	Saccharose
	10 g	Glukose
	0,25 g	K <sub>2</sub> SO <sub>4</sub>
	10,12 g	MgCl <sub>2</sub> x6H <sub>2</sub> O
	0,1 g	Casamino Acids
	5 g	Hefeextrakt
	5,74 g	TES
	2 ml	Spurenelement-Lösung

mit H<sub>2</sub>O<sub>deion.</sub> auf 955 mL, pH 7,2; autoklavieren  
danach sollen folgende, ebenfalls autoklavierte  
Lösungen steril zugesetzt werden:

20 mL	CaCl <sub>2</sub> ·2H <sub>2</sub> O (1 M)
10 mL	KH <sub>2</sub> PO <sub>4</sub> (0,5 %)
15 mL	L-Prolin (20 %)

Für die direkte Transformation von *A. balhimycina*:

S27M-Medium:	5 g	Bacto Pepton
	3 g	Hefeextrakt
	73 g	Mannitol

frisch herstellen

TSB-Medium:	20 g	Bacto Pepton
	5 g	NaCl
	3,3 g	K <sub>2</sub> HPO <sub>4</sub> ·3H <sub>2</sub> O
	2,5 g	Glukose

Alle Medien wurden falls nicht anders angegeben mit H<sub>2</sub>O<sub>deion</sub> auf 1 L aufgefüllt und autoklaviert. Für die entsprechenden Festmedien wurden 15 g/L Agar zugesetzt.

#### 4.1.8 Antibiotika und andere Medienzusätze

Die angegebenen Konzentrationen beziehen sich, falls nicht anders angegeben, auf die 1000x konzentrierte Stammlösung.

Ampicillin:	100 mg/mL in H <sub>2</sub> O <sub>dd, steril</sub> , sterilfiltrieren
Erythromycin:	50 mg/mL in H <sub>2</sub> O <sub>dd, steril</sub> , zum Lösen mit HCl auf pH 6,0; sterilfiltrieren
IPTG (400x):	2 % in H <sub>2</sub> O <sub>dd, steril</sub> , sterilfiltrieren, vor Licht schützen
X-Gal (400x):	2 % in DMF, sterilfiltrieren, vor Licht schützen

#### 4.1.9 Lösungen und Puffer

Für die Stammhaltung von *E. coli*:

Glycerin-Lösung (20 %), autoklavieren

Für die Kultivierung von *A. balhimycina*:

Spurenelement-Lösung für R5:	200 mg	FeCl <sub>3</sub> x6H <sub>2</sub> O
	40 mg	ZnCl <sub>2</sub>
	10 mg	Na <sub>2</sub> B <sub>4</sub> O <sub>7</sub> x10H <sub>2</sub> O
	10 mg	(NH <sub>4</sub> ) <sub>6</sub> Mo <sub>7</sub> O <sub>24</sub> x4H <sub>2</sub> O
	7,88 mg	CuCl <sub>2</sub>
	10 mg	MnCl <sub>2</sub> x4H <sub>2</sub> O

mit H<sub>2</sub>O<sub>deion.</sub> auf 1 L, autoklavieren

Für die Gelelektrophorese von DNA:

DNA-Probenauftragspuffer (10x):	10 mM	Tris
	5 %	Saccharose
	0,02 %	SDS
	0,01 %	Bromphenolblau
Ethidiumbromidfärbebad:	1 Spatelspitze	Ethidiumbromid/200mL H <sub>2</sub> O <sub>dd</sub>
TAE-Elektrophoresepuffer:	40 mM	Tris/HCl
	20 mM	Natriumacetat
	1 mM	EDTA
	pH 7,8	

Für die Eckhardt-Lyse:

E-1-F-Probenauftragspuffer:	25 g	Saccharose
	3 g	Ficoll 400
	mit TAE-Puffer auf 100 mL und autoklavieren	
	10 mg	RNase A
	0,04 %	Bromphenolblau
	5 mg/mL	Lysozym

Für die Herstellung  $\text{CaCl}_2$ -kompetenter *E. coli*:

$\text{CaCl}_2$ -Lösung (50 mM und 100 mM), autoklavieren

Glycerin-Lösung (60 %), autoklavieren

Für die direkte Transformation von *A. baohimycina*:

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  (3,68 %), autoklavieren

CsCl (4,17 M), autoklavieren, frisch herstellen

Kalbsthymus-DNA (7,5  $\mu\text{g}/\mu\text{l}$ ): Kalbsthymus-DNA in  $\text{H}_2\text{O}_{\text{dd, steril}}$   
im Ultraschallbad lösen (5 min)  
zentrifugieren (15 min, 16000 g)  
Überstand abnehmen und bei 100°C kochen  
(20 min), auf Eis abkühlen  
Lagerung bei -20°C

$\text{MgCl}_2$  (0,2 M), autoklavieren

NB-Weichagar mit Erythromycin: 8 g Nutrient Broth

2,5 g Agar

mit  $\text{H}_2\text{O}_{\text{deion.}}$  auf 500 mL, autoklavieren

Lagerung bei 60°C

vor Benutzung Erythromycin zusetzen (dabei das  
Volumen der zu überschichtenden Platte berücksichtigen und nicht nur das des NB-Weichagars)

PEG1000 (65 %), autoklavieren, frisch herstellen

R2L-Weichagar: 0,01 g Casamino Acids

1 g Glukose

0,5 g Hefeextrakt

0,025 g  $\text{K}_2\text{SO}_4$

7,32 g Mannitol

1,8 g  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$

0,7 g Agarose

mit  $\text{H}_2\text{O}_{\text{deion.}}$  auf 82 mL, autoklavieren

Lagerung bei 60°C

vor Benutzung zusetzen

10 mL TES (0,25 M, pH 7,2)

8 mL  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  (3,68 %)



Für die Isolierung von Gesamt-DNA aus *A. balhimycina*:

Lysepuffer:	20 mM	Tris/HCl
	2 mM	EDTA
	1 %	Triton X 100
	pH 8,0	
	autoklavieren	
	kurz vor Verwendung zugeben	
	20 mg/mL	Lysozym
TE/10 % Saccharose-Lösung:	10 mM	Tris/HCl
	1 mM	EDTA
	10 %	Saccharose
	pH 7,5	
	autoklavieren	

Für die *Southern Blot*-Hybridisierung:

0,1xSSC/0,1 % SDS

2xSSC/0,1 % SDS

20x SSC-Puffer:	3 M	NaCl
	300 mM	tri-Natriumcitratx2H <sub>2</sub> O
	pH 7,0	

*Blocking*-Lösung: 1 % *Blocking*-Reagenz in Maleinsäurepuffer

Denaturierungs-Lösung: 0,5 M NaOH

1,5 M NaCl

Detektionspuffer: 0,1 M Tris/HCl

0,1 M NaCl

0,05 M MgCl<sub>2</sub>x6H<sub>2</sub>O

pH 9,5

Färbe-Lösung: 10 µl CSPD-Lösung/Kit (11,6 mg/mL)

1 mL Detektionspuffer

Hybridisierungs-Lösung: 25 mL SSC (20x)

0,3 mL *N*-Lauroylsarkosin (35 %)

0,2 mL SDS (10 %)

10 mL *Blocking*-Lösung (10 %)

mit H<sub>2</sub>O<sub>deion.</sub> auf 100 mL

Maleinsäurepuffer:	0,1 M	Maleinsäure
	0,15 M	NaCl
	pH 7,5 (mit NaOH Plätzchen)	
Neutralisierungs-Lösung:	0,5 M	Tris/HCl
	1,5 M	NaCl
	pH 7,5	
Waschpuffer:	0,3 %	Tween 20 in Maleinsäurepuffer

Für die Isolierung nicht-ribosomaler Peptidsynthetasen:

Ammoniumsulfat-Lösung (gesättigt):

Ammoniumsulfat

bei 4°C kaltstellen

Polymin P-Lösung: 50 g Polymin P  
in 100 mL H<sub>2</sub>O<sub>dd.</sub>

Dialyse gegen H<sub>2</sub>O<sub>dd.</sub>, um Ethyleniminmonomere zu entfernen

wenn pH außerhalb des Dialyseschlauchs nicht mehr basisch wird, treten keine Monomere mehr aus und die Dialyse kann beendet werden

Volumen und darüber Konzentration bestimmen mit HCl auf pH 7,0 neutralisieren

0,02 % Azid

bei 4°C kaltstellen

Proteinisolierungspuffer: 100 mM Tris/HCl  
pH 8,0  
15 % (w/v) Glycerin  
0,3 M NaCl

entgasen und bei 4°C lagern

kurz vor Verwendung zugeben

4 mM DTT

1 mM Benzamidin

1 mM PMSF

2 mM EDTA

Ultraschall zum Unterstützen des Lösungsvorgangs

Für die Konzentrationsbestimmung nach Bradford:

1,1x Bradford-Stammlösung:	100 mg	Coomassie Brillantblau G 250
	100 mL	Phosphorsäure (85 %)
	50 mL	Ethanol
	mit H <sub>2</sub> O <sub>dd.</sub> auf 1 L	

Für die Gelelektrophorese von Proteinen:

10x Elektrophoresepuffer:	25 mM	Tris
	1,9 M	Glycin
	1 %	SDS
	pH nicht einstellen (ergibt sich zu 8,9)	

APS-Stammlösung: 100 mg/mL

Probenauftragspuffer:	3,75 mL	Tris/HCl (1 M, pH 6,8)
	1,2 g	SDS
	6 mL	Glycerin
	6 mg	Bromphenolblau
	13,5 mL	H <sub>2</sub> O <sub>dd.</sub>

kurz vor Verwendung 1,5 mL DTT (1 M) zusetzen

Sammelgel (3 %):	1,236 mL	Acrylamid/Bisacrylamid-Lösung
	3,884 mL	H <sub>2</sub> O <sub>dd.</sub>
	6,3 mL	Sammelgelpuffer
	120 µl	SDS-Stammlösung
	30 µl	TEMED
	80 µl	APS-Stammlösung

Sammelgelpuffer: 0,25 M Tris/HCl  
pH 6,8

SDS-Stammlösung: 10 % SDS

Trenngel (5 %):	3,8 mL	Acrylamid/Bisacrylamid-Lösung
	12,8 mL	H <sub>2</sub> O <sub>dd.</sub>
	5,74 mL	Trenngelpuffer
	115 µl	SDS-Stammlösung
	57,5 µl	TEMED
	460 µl	APS-Stammlösung



## 4.2 Methoden

### 4.2.1 Kultivierung und Stammhaltung von *E. coli*

Die Kultivierung von *E. coli* zur Plasmidisolierung erfolgte in TY-Medium mit dem entsprechenden Antibiotikum zur Selektion bei 37°C und 150 rpm für 16 h. Die Kultivierungsgefäße richteten sich nach dem jeweiligen Kulturvolumen und damit nach dem Bedarf an Plasmid.

Zur Stammhaltung von *E. coli* wurde eine 16 h alte Flüssigkultur (2 mL) pelletiert, in 20 % Glycerin-Lösung (250 µl) resuspendiert und bei -20°C eingefroren. Idealerweise wurde der Rest der Flüssigkultur zur Plasmidisolierung und anschließender Auftragssequenzierung verwendet.

### 4.2.2 Kultivierung und Stammhaltung von *A. balhimycina*

Die Kultivierung von *A. balhimycina* erfolgte stets in Erlenmeyerkolben (500 mL), die zur besseren Belüftung mit Schikanen und einer Metallspirale ausgestattet waren. Diese Erlenmeyerkolben wurden mit 100 mL Medium befüllt, angeimpft und bei 28 bis 30°C und 220 rpm auf dem Schüttelinkubator inkubiert.

Für die Produktion von Glykopeptiden und für die Isolierung von nicht-ribosomalen Peptidsynthetasen wurde R5 Medium verwendet. Die mit einem Agarblöckchen (1 cm<sup>2</sup>) oder aus einer Kryokultur beimpfte Vorkultur wurde nach zwei Tagen zum Beimpfen der Hauptkultur verwendet (Inokulum 1 %). Die Kultivierungsdauer betrug sechs Tage.

Für die direkte Transformation wurde TSB-Medium verwendet. Die mit einem Agarblöckchen (1 cm<sup>2</sup>) oder aus einer Kryokultur beimpfte Vorkultur wurde nach 24 h zum Beimpfen der Hauptkultur verwendet (Inokulum 5 %). Die Kultivierungsdauer betrug 52 bis 55 h.

Für die Durchführung des Stressprotokolls, ausgehend von den *single crossover*-Klonen, wurde R5-Medium ohne Erythromycin verwendet. Die mit einem Agarblöckchen (1 cm<sup>2</sup>) oder aus einer Kryokultur beimpfte Hauptkultur wurde nach 48 bis 72 h Kultivierung verwendet.

Für die Isolierung von Gesamt-DNA wurde R5-Medium verwendet. Die mit einem Agarblöckchen (1 cm<sup>2</sup>) oder aus einer Kryokultur beimpfte Hauptkultur wurde nach 48 bis 72 h Kultivierung verwendet.

Zur Stammhaltung von *A. balhimycina* wurde eine 48 h alte Flüssigkultur in R5-Medium entweder direkt bei  $-80^{\circ}\text{C}$  eingefroren oder vorher mit 86 % Glycerin (1/6 Volumen der Kultur) versetzt. Idealerweise wurde der Rest der Flüssigkultur zur Isolierung von Gesamt-DNA verwendet, die mittels PCR oder *Southern Blot*-Hybridisierung auf den Genotyp getestet wurde.

Im Allgemeinen wurde jede Kultur vor ihrer Verwendung mikroskopisch auf Kontaminationen untersucht.

### 4.2.3 Polymerase-Kettenreaktion (PCR)

Die PCR ist eine zyklische Methode zur *in vitro*-Amplifikation beliebiger Nukleinsäureabschnitte. Sie wurde im Jahre 1983 von Kary B. Mullis entdeckt, 1986 veröffentlicht<sup>[79]</sup> und 1993 mit dem Nobelpreis für Chemie ausgezeichnet. Aufgrund der zumindest theoretischen Möglichkeit, ausgehend von einem einzelnen DNA-Molekül beliebig viele Kopien herzustellen, ist die PCR auch den derzeit empfindlichsten, routinemäßig einsetzbaren Analyseverfahren hinsichtlich ihrer Sensitivität weit überlegen.

Die Zyklen einer PCR umfassen drei Schritte, die bei unterschiedlichen Temperaturen stattfinden. Im ersten Schritt („Denaturierung“) erfolgt durch Temperaturerhöhung typischerweise auf  $92-98^{\circ}\text{C}$  eine Trennung der DNA-Doppelstränge. An die so erhaltenen Einzelstränge können sich im zweiten Schritt („Annealing“) kurze Oligonukleotidprimer anlagern, wenn der Reaktionsansatz auf eine durch die Primer (GC-Gehalt, Länge) vorgegebene Temperatur abgekühlt wird. Im dritten Schritt („Elongation“) wird die Temperatur auf  $72^{\circ}\text{C}$  erhöht, was optimale Bedingungen für die DNA-Synthese durch thermostabile DNA-Polymerasen darstellt. Je nach Bedarf können DNA-Polymerasen mit unterschiedlichen Eigenschaften verwendet werden. In der vorliegenden Arbeit kamen folgende DNA-Polymerasen zum Einsatz:

Die ProofStart DNA-Polymerase gewährleistet eine sehr hohe Genauigkeit der PCR aufgrund ihrer  $3' \rightarrow 5'$ -Exonukleaseaktivität. Diese Korrekturlesefunktion führt zu PCR-Produkten mit glatten Enden und vor allem zur gewünschten niedrigen Fehlerrate von ca.  $3,0 \times 10^{-6}$  falschen Basen pro Verdopplung des PCR-Produkts.

Der Taq-Polymerase fehlt die  $3' \rightarrow 5'$ -Exonukleaseaktivität, weshalb sie eine etwa um den Faktor 10 größere Fehlerrate besitzt. Allerdings besitzt sie eine sehr hohe Syntheserate von ca. 2800 Nukleotiden pro min, was dazu führt, dass sie viel und

gerne verwendet wird. Neben ihrer templatabhängigen 5'→3'-Polymeraseaktivität besitzt sie auch noch eine templatunabhängige, die für die A-Überhänge am 3'-Ende verantwortlich ist, die in der sog. TA-Klonierung genutzt werden (s. Abschnitt 4.2.6.4).

#### 4.2.3.1 PCR-Ansätze

Der folgende Ansatz mit der ProofStart DNA-Polymerase wurde für die vier PCR-Reaktionen *Tarleft*, *Tarright*, *Insleft* und *Insright* verwendet. Als Templat-DNA diente ein das Balhimycin-Biosyntheseegencluster tragendes Cosmid (Bezeichnung 4.3).

H <sub>2</sub> O <sub>dd, steril</sub>	32,5 µl
MgSO <sub>4</sub> (25 mM)	5 µl
ProofStart Puffer (10x)	5 µl
DMSO	3 µl
dNTP-Mix (je 10 mM)	1 µl
Primer vorwärts (100 pmol/µl)	1 µl
Primer rückwärts (100 pmol/µl)	1 µl
Templat-DNA	1 µl
ProofStart DNA-Polymerase	0,5 µl

Der folgende Ansatz mit der Taq DNA-Polymerase wurde für die KontrollPCRs „Erythromycinresistenz“ und „Modulinserion“ verwendet. Als Templat-DNA diente jeweils die Gesamt-DNA des zu testenden *A. balhimycina* Klons.

H <sub>2</sub> O <sub>dd, steril</sub>	37,75 µl
Reaktionspuffer S (10x)	5 µl
DMSO	3 µl
dNTP-Mix (je 10 mM)	1 µl
Primer vorwärts (100 pmol/µl)	1 µl
Primer rückwärts (100 pmol/µl)	1 µl
Templat-DNA	1 µl
Taq DNA-Polymerase	0,25 µl

Der folgende Ansatz mit der Taq DNA-Polymerase wurde für die Herstellung der DIG-Sonde für die *Southern Blot*-Hybridisierung verwendet. Als Templat-DNA diente das Transformationsplasmid pC5A5T4E4.

H <sub>2</sub> O <sub>dd, steril</sub>	33,75 µl
Reaktionspuffer S (10x)	5 µl
DMSO	3 µl
PCR DIG <i>labeling</i> -Mix	5 µl
(2 mM dA/C/GTP, 1,9 mM dTTP, 0,1mM DIG-dUTP)	
Primer vorwärts (100 pmol/µl)	1 µl
Primer rückwärts (100 pmol/µl)	1 µl
Templat-DNA	1 µl
Taq DNA-Polymerase	0,25 µl

#### 4.2.3.2 Kolonie-PCR

Bei den bisher beschriebenen PCR-Ansätzen lag die eingesetzte Templat-DNA immer gereinigt vor. Wenn man allerdings viele Klone mit einer PCR untersuchen möchte, kann man stattdessen auch ungereinigte Templat-DNA verwenden, die nur mittels Erhitzen (10 min, 100°C) aus den in H<sub>2</sub>O<sub>dd, steril</sub> (100 µl) aufgenommenen Zellen freigesetzt wurde. Vor der Verwendung in der PCR wird dieser Ansatz auf Eis abgekühlt und abzentrifugiert (10 min, 16000 g). Statt wie sonst 1 µl müssen allerdings bei der Kolonie-PCR 10 bis 15 µl Templat-DNA eingesetzt werden.

#### 4.2.3.3 PCR-Programme

Zur Amplifikation von *Tarright* und *Insright*:

Schritt	Reaktion	Temperatur	Zeit
1	Initiale Denaturierung	94°C	2 min
2	Denaturierung	94°C	1 min
3	Annealing	56°C	1,5 min
4	Elongation	72°C	2 min
5	30x Schritt 2 bis 4		
6	Finale Elongation	72°C	7 min

Zur Amplifikation von *Tarleft*:

Schritt	Reaktion	Temperatur	Zeit
1	Initiale Denaturierung	98°C	30 s
2	Denaturierung	98°C	10 s
3	Annealing	68°C	30 s
4	Elongation	72°C	2,5 min
5	30x Schritt 2 bis 4		
6	Finale Elongation	72°C	10 min

Zur Amplifikation von *Insleft*:

Schritt	Reaktion	Temperatur	Zeit
1	Initiale Denaturierung	95°C	2,5 min
2	Denaturierung	95°C	30 s
3	Annealing	70°C	1 min
4	Elongation	72°C	3 min
5	30x Schritt 2 bis 4		
6	Finale Elongation	72°C	7 min

Zur Amplifikation von „Erythromycinresistenz“ und „Modulinserion“:

Schritt	Reaktion	Temperatur	Zeit
1	Initiale Denaturierung	95°C	2 min
2	Denaturierung	95°C	30 s
3	Annealing	56°C	45 s
4	Elongation	72°C	2 min
5	30x Schritt 2 bis 4		
6	Finale Elongation	72°C	7 min

Zur Amplifikation der DIG-Sonde für die *Southern Blot*-Hybridisierung:

Schritt	Reaktion	Temperatur	Zeit
1	Initiale Denaturierung	95°C	2 min
2	Denaturierung	95°C	30 s
3	Annealing	49°C	45 s
4	Elongation	72°C	2 min
5	30x Schritt 2 bis 4		
6	Finale Elongation	72°C	7 min

#### 4.2.4 Photometrische Konzentrationsbestimmung von DNA

Die photometrische Konzentrationsbestimmung von DNA erfolgt in UV-durchlässigen Küvetten mit 1 cm Schichtdicke bei  $\lambda = 260$  nm, dem Absorptionsmaximum der Nukleinsäuren. Eine Lösung, die 50  $\mu\text{g/mL}$  doppelsträngige DNA enthält, besitzt unter diesen Bedingungen eine Absorption bzw. optische Dichte (OD) von 1. Durch Messung der  $\text{OD}_{260}$  können so Konzentrationen unbekannter DNA-Lösungen bestimmt werden, wobei das reine Lösungsmittel zum Nullabgleich verwendet wird. Es gilt folgende Formel zur Berechnung der DNA-Konzentration:

$$\text{Konzentration } [\mu\text{g/mL}] = \text{OD}_{260} \times 50$$

Eine weitere Möglichkeit ist die photometrische Reinheitskontrolle von DNA-Lösungen, die darauf beruht, dass Proteine ein Absorptionsmaximum bei  $\lambda = 280$  nm aufweisen. Eine reine DNA-Lösung besitzt einen  $\text{OD}_{260}/\text{OD}_{280}$ -Wert von 1,8, Proteinkontaminationen reduzieren diesen Wert signifikant. Neben diesem Quotienten dient auch der  $\text{OD}_{260}/\text{OD}_{230}$ -Wert zur Charakterisierung der Reinheit der DNA-Lösung, da neben Peptidbindungen auch Tris, EDTA und andere Salze eine Absorption bei  $\lambda = 230$  nm zeigen.

In der Praxis wird außerdem die  $\text{OD}_{320}$  bestimmt, die zur Hintergrundkompensation verwendet wird. Die Wellenlänge  $\lambda = 320$  nm ist hierfür geeignet, da sie weit von den Absorptionsmaxima von DNA und Protein bei  $\lambda = 260$  nm bzw.  $\lambda = 280$  nm entfernt ist. Es gilt folgende Formel:

$$\text{Konzentration } [\mu\text{g/mL}] = (\text{OD}_{260} - \text{OD}_{320}) \times 50$$

#### 4.2.5 Gelelektrophorese von DNA

Die Gelelektrophorese ist eine sehr wichtige Methode zur Analyse von DNA. Als Trägermaterial wird vorwiegend Agarose, ein Polysaccharid, verwendet. Die Phosphatgruppen der DNA liegen innerhalb eines sehr großen pH-Bereichs negativ geladen vor, so dass DNA eine konstante Ladungsdichte, also ein konstantes Molekülmasse-zu-Ladungs-Verhältnis besitzt. Deshalb ist die Wanderungsgeschwindigkeit im elektrischen Feld in freier Lösung für alle DNA-Moleküle in erster Näherung gleich. In einer festen Gelmatrix hingegen treten Unterschiede in der Wanderungsgeschwindigkeit auf, abhängig von der Molekülgröße und der Form der DNA (superhelikal, relaxiert, doppelsträngig-linear oder einzelsträngig). Bei doppelsträngig-linearer DNA besteht über einen weiten Größenbereich eine empirisch beobachtete lineare Abhängigkeit zwischen dem Logarithmus ( $\log_{10}$ ) der DNA-Fragmentlänge und der relativen Wanderungsdistanz im Gel. Superhelikale DNA wandert in der Regel schneller, relaxierte DNA dagegen wesentlich langsamer als doppelsträngig-lineare DNA. In der Praxis verwendet man sog. Längenstandards oder DNA-Leitern zur Charakterisierung von Fragmentgrößen. Um Diffusionsverluste zu vermeiden, wird die DNA mit Hilfe sog. DNA-Probenauftragspuffer in die Geltaschen aufgetragen, die neben Saccharose oder ähnlichen Verbindungen zur Erhöhung der Dichte auch Farbstoffe wie beispielsweise Bromphenolblau enthalten. Diese wandern während der Elektrophorese in Richtung Anode und bieten somit einen Anhaltspunkt für die Wanderung der DNA im Gel. Zur Detektion der DNA wird in den meisten Fällen Ethidiumbromid verwendet. Dieser Fluoreszenzfarbstoff kann aufgrund seiner planaren Struktur in v. a. doppelsträngige DNA interkalieren, wobei die Wechselwirkung der aromatischen Ringe mit den heteroaromatischen Ringen der DNA-Basen die Fluoreszenz verstärkt.

In der vorliegenden Arbeit wurden zur Gelelektrophorese 1 %-ige Agarosegele in Horizontalkammern verwendet. TAE-Elektrophoresepuffer diente hierbei als Gel- und Laufpuffer. Je nach Trennproblem wurde die Elektrophorese bei verschiedenen Spannungen und unterschiedlichen Laufzeiten durchgeführt, wobei die Kombination 90 V und 30 min häufig verwendet wurde. Aus Laborsicherheitsgründen wurden alle mit dieser Arbeit verbundenen Agarosegele erst nach der Elektrophorese in einem Ethidiumbromidbad angefärbt und unter UV-Licht einer Geldokumentationsanlage analysiert bzw. fotografiert.

#### 4.2.6 Klonierung von DNA-Fragmenten

Unter Klonierung versteht man allgemein die Einführung von DNA-Fragmenten in Vektoren. Dieser Vorgang erfordert und bedingt eine Vielzahl zum Teil enzymatisch katalysierter Schritte, wobei im Folgenden nur die in dieser Arbeit angewendeten Reaktionen bzw. Methoden erwähnt werden sollen.

##### 4.2.6.1 Restriktionsverdau

Restriktionsendonukleasen oder kurz Restriktionsenzyme (Typ II) erkennen je nach Spezifität vier bis acht, normalerweise palindromische Basenpaare in einem DNA-Doppelstrang und schneiden diesen meist innerhalb dieser Erkennungssequenz. Nach dem Restriktionsverdau bleiben an den 5'-Enden der DNA-Fragmente fast immer Phosphatreste zurück, die für eine anschließende Ligation benötigt werden, während die 3'-Enden lediglich Hydroxygruppen tragen. Neben der Erkennungssequenz ist auch die Art der generierten Enden ein wichtiges Charakteristikum eines Restriktionsenzym. Während manche Enzyme glatte Enden (*blunt ends*) produzieren, können andere 5'- bzw. 3'-Überhängenden (*sticky ends*) unterschiedlicher Länge generieren. In der Praxis dienen Restriktionsverdau aber nicht nur als Vorstufe für Ligationsreaktionen, sondern können mit anschließender Agarosegelelektrophorese auch zur Kontrolle von DNA-Konstrukten verwendet werden. Die käuflich erhältlichen Restriktionsenzyme liegen vorwiegend in einer Konzentration von 10 units/ $\mu$ l vor, wobei eine unit definiert ist als die Enzymmenge, die 1  $\mu$ g DNA in 1 h bei der angegebenen Temperatur in einem 50  $\mu$ l Ansatz vollständig verdaut.

In der vorliegenden Arbeit wurde für analytische Zwecke folgender Ansatz verwendet und 1 h bei der vom Hersteller angegebenen Temperatur inkubiert:

DNA-Lösung (s. Abschnitt 4.2.9)	5 bis 7 $\mu$ l
Restriktionspuffer	1 $\mu$ l
evtl. 100x BSA	0,1 $\mu$ l
Restriktionsenzym	1 $\mu$ l
H <sub>2</sub> O <sub>dd, steril</sub>	auf 10 $\mu$ l auffüllen

Für präparative Zwecke, d.h. wenn anschließend mit einem oder mehreren Fragmenten weitergearbeitet werden sollte, wurde ein größerer Ansatz verwendet, der 2 bis 5 h bei der vom Hersteller angegebenen Temperatur inkubiert wurde:

DNA-Lösung (s. Abschnitt 4.2.9)	20 µl
Restriktionspuffer	4 µl
evtl. 100x BSA	0,4 µl
Restriktionsenzym	4 µl
H <sub>2</sub> O <sub>dd, steril</sub>	auf 40 µl auffüllen

Für die *Southern Blot*-Hybridisierungen musste Gesamt-DNA von *A. balhimycina* in präparativem Maßstab verdaut werden. Hierzu wurde folgender Ansatz bei der vom Hersteller angegebenen Temperatur 16 h inkubiert:

DNA-Lösung (s. Abschnitt 4.2.10)	53 µl
Restriktionspuffer	7 µl
evtl. 100x BSA	0,7 µl
Restriktionsenzym	5 µl
H <sub>2</sub> O <sub>dd, steril</sub>	auf 70 µl auffüllen

Ein Restriktionsverdau mit zwei Restriktionsenzymen, ein sog. Doppelverdau, wurde unter Zugabe beider Restriktionsenzyme durchgeführt. Bezüglich des Reaktionspuffers musste in diesen Fällen ein Kompromiss gefunden werden, was in schwierigen Fällen mit Hilfe von Testeinzelverdaus erfolgte.

#### 4.2.6.2 Reinigung von DNA durch Agarosegelelektrophorese und Gelextraktion

Nach PCR-Reaktionen oder Restriktionsverdaus wird das gewünschte PCR-Produkt bzw. Restriktionsfragment häufig mittels Agarosegelelektrophorese von unerwünschter DNA getrennt und anschließend aus dem Gel extrahiert. Um Schäden an der DNA zu vermeiden wird hierbei das Agarosegel nur leicht mit Ethidiumbromid angefärbt und die Zeit unter UV-Licht möglichst kurz gehalten. Nach dem Ausschneiden der gewünschten DNA-Bande aus dem Gel mit einem Skalpell wird die DNA aus dem Gel extrahiert. In einigen Fällen wurde von den die DNA enthaltenden Spuren nur eine gefärbt, dadurch die Höhe der gewünschten DNA-Bande bestimmt und damit die ungefärbte DNA aus den restlichen Spuren ausgeschnitten. Für die Gelextraktion sind verschiedene Kits im Handel erhältlich,

wobei in der vorliegenden Arbeit das *QIAquick Gelextraktions-Kit* sowie das *QIAEX II Gelextraktions-Kit* verwendet wurden. In beiden Fällen muss zunächst das ausgeschnittene Agaroseblöckchen aufgelöst werden. Dies geschieht in Analogie zu der von Vogelstein und Gillespie<sup>[80]</sup> 1979 vorgeschlagenen Methode, Agarose mit Hilfe chaotroper Salze wie z.B. Natriumiodid zu lösen. Dies geschieht durch das Aufbrechen der Wasserstoffbrückenbindungen zwischen den Zuckerresten der Agarose. In der damit hergestellten Gegenwart hoher Konzentrationen chaotroper Salze adsorbiert DNA an Silikagel. Man versucht dieses Phänomen mit einer Modifikation der Struktur des Wassers zu erklären.<sup>[81]</sup> An dieser Stelle unterscheiden sich die beiden Kits bei der praktischen Durchführung, wobei das Silikagel beim *QIAquick Gelextraktions-Kit* als Membran in einer Säule und beim *QIAEX II Gelextraktions-Kit* in partikulärer Form vorliegt. Im ersten Fall können so die Lösungen leicht durch Zentrifugation die Säule passieren und mit ihr in Wechselwirkung treten. Im zweiten Fall sind die Silikagelpartikel direkt in den Lösungen suspendiert und können durch Pelletieren und Abpipettieren leicht wieder von selbigen getrennt werden. Kontaminationen wie z.B. Agarose und Ethidiumbromid treten nicht in Wechselwirkung mit dem Silikagel und können durch Waschschriffe entfernt werden. Die abschließende Elution der DNA erfolgt im Gegensatz zur Adsorption bevorzugt bei niedrigen Salzkonzentrationen und eher basischem pH (Tris-Puffer oder Wasser), wobei Erwärmung des Elutionsmittels auf 60°C die Effizienz erhöht. Die genaue praktische Durchführung in der vorliegenden Arbeit erfolgte jeweils nach den Herstellerangaben.

#### 4.2.6.3 A-Addition

Mit TA-Klonierung bezeichnet man eine häufig verwendete Methode, bei der PCR-Produkte mit A-Überhang leicht in einen Vektor mit T-Überhang kloniert werden können. Während DNA-Polymerasen ohne 3'-> 5'-Exonukleaseaktivität, z.B. die Taq-Polymerase, meist diesen unspezifischen A-Überhang synthetisieren, erzeugen DNA-Polymerasen mit 3'-> 5'-Exonukleaseaktivität ausschließlich glatte Enden. Um bei damit erzeugten PCR-Produkten die Methode der TA-Klonierung auch anwenden zu können, bedient man sich der sog. *A-Addition*. Die Firma Qiagen nutzt hierzu nach eigenen Angaben eine „Art Klenow-Fragment“, welches in Anwesenheit von dATP die gewünschten A-Überhänge generiert. Wie den Herstellerangaben zu entnehmen

ist, sollen 0,15 bis 1,5 pmol PCR-Produkt in die Reaktion eingesetzt werden, wobei sich folgender Ansatz ergab, der 30 min bei 37°C inkubiert wurde:

PCR-Produkt	≤ 8 µl
5x QIAGEN <i>A-Addition Master-Mix</i>	2 µl
H <sub>2</sub> O <sub>dd, steril</sub>	auf 10 µl auffüllen

#### 4.2.6.4 TA-Klonierung

Für das bereits in Abschnitt 4.2.6.3 beschriebene Prinzip der TA-Klonierung wurden in der vorliegenden Arbeit die zwei Klonierungsvektoren pDrive und pGEM-T-Easy verwendet. Die Herstellerempfehlungen raten zu einem fünf- bis zehnfachen molaren Überschuss an PCR-Produkt gegenüber dem Klonierungsvektor pDrive und zu einem molaren Verhältnis von 1:1 beim Klonierungsvektor pGEM-T-Easy. Hieraus ergaben sich folgende in der vorliegenden Arbeit verwendete Ansätze.

pDrive:

PCR-Produkt bzw. <i>A-Addition-Ansatz</i>	1 bis 4 µl
pDrive-Klonierungsvektor (50 ng/µl)	1 µl
2x Ligation Master Mix	5 µl
H <sub>2</sub> O <sub>dd, steril</sub>	auf 10 µl auffüllen

pGEM-T-Easy:

PCR-Produkt bzw. <i>A-Addition-Ansatz</i>	1 bis 3 µl
pGEM-T-Easy-Klonierungsvektor (50 ng/µl)	1 µl
2x <i>Rapid</i> Ligation Puffer, T4 DNA-Ligase	5 µl
T4 DNA-Ligase	1 µl
H <sub>2</sub> O <sub>dd, steril</sub>	auf 10 µl auffüllen

Nach Inkubation (16 h) bei langsamem Abkühlen von 30°C auf 4°C und anschließender Transformation in CaCl<sub>2</sub>-kompetente *E.coli* XL1 Blue (s. Abschnitt 4.2.7.1) wurden die Zellen auf Ampicillin-Platten ausgestrichen. Bei den auf Ampicillin wachsenden Klonen war somit die Existenz eines Klonierungsvektors in *E.coli* XL1 Blue sichergestellt, allerdings nicht die erfolgreiche TA-Klonierung eines PCR-Produkts in den Klonierungsvektor. Zur Selektion auf dieses Ereignis bedient man sich der sog. Blau-Weiß-Selektion. Das Prinzip dieser Selektion beruht darauf, dass

die Klonierungsstelle in einem Gen liegt, welches durch die Insertion eines PCR-Produktes inaktiviert wird. Bei diesem Gen handelt es sich um das lacZ'-Gen, welches das N-terminale  $\alpha$ -Fragment der  $\beta$ -Galaktosidase kodiert, welches zusammen mit dem vom Bakterien-Stamm exprimierten  $\omega$ -Fragment LacZ $\Delta$ M15 eine  $\beta$ -Galaktosidaseaktivität besitzt. Setzt man den TY-Platten Isopropyl- $\beta$ -D-thiogalaktopyranosid (IPTG) zur Induktion und 5-Brom-4-chlor-3-indoxyl- $\beta$ -D-galaktopyranosid (X-Gal) als Substrat zu, so kann die  $\beta$ -Galaktosidaseaktivität X-Gal zu Galaktose und 5-Brom-4-chlor-indoxyl hydrolysieren, wobei letzteres vom Luftsauerstoff zum tiefblauen Farbstoff 5,5'-Dibrom-4,4'-dichlor-indigo oxidiert wird. Wurde das lacZ'-Gen allerdings durch Insertion eines PCR-Produktes inaktiviert, so bleiben die Klone weiß.<sup>[82]</sup>

#### 4.2.6.5 Eckhardt-Lyse zur Untersuchung von Klonen

Hinter dem Namen Eckhardt-Lyse<sup>[83]</sup> verbirgt sich ein Verfahren zur schnellen Untersuchung von vielen Klonen. Statt des normalen Prozederes von Anzucht in Flüssigmedium, nachfolgender Plasmidisolierung, Restriktionsverdau und Gelelektrophorese kann man hier die Klone direkt von Platte gelelektrophoretisch untersuchen. Von jedem Klon werden einige Zellen in 20  $\mu$ l des speziellen Probenauftragspuffers E-1-F aufgenommen und vom darin enthaltenen Lysozym lysiert. Diese Proben werden auf ein Agarosegel aufgetragen, das neben 1 %-Agarose außerdem 0,2 % SDS enthält. In den ersten 15 min wird die Gelelektrophorese nur bei 15 V betrieben. In dieser Zeit wird die Zelllyse durch den elektrophoretischen Transfer des negativ geladenen SDS in die Geltaschen komplettiert. Anschließend erfolgen weitere 40 min bei 80 V. Nach Abwaschen des SDS wird das Gel im Ethidiumbromidbad gefärbt und eine unter UV-Licht analysiert bzw. fotografiert. Typischerweise kann man die leeren Klonierungsvektoren deutlich von den weiter oben laufenden Klonierungsvektoren mit aufgenommenem PCR-Produkt unterscheiden. Die entsprechenden Klone werden dann nach den üblichen Protokollen auf die Identität des aufgenommenen PCR-Produkts hin untersucht.

#### 4.2.6.6 Ligation

Unter Ligation versteht man das Verknüpfen von DNA-Strängen. Als Standardenzym findet hierbei die T4 DNA-Ligase Verwendung, die sowohl glatte als auch überhängende Enden doppelsträngiger DNA unter ATP-Verbrauch miteinander verbindet, indem sie Phosphodiesterbindungen zwischen den 3'-OH- und den 5'-Phosphatenden ausbildet.

In der vorliegenden Arbeit wurde die T4 DNA-Ligase von Biolabs verwendet, die eine Konzentration von 400 units/ $\mu$ l für überhängende Enden aufweist. Die Menge 1 unit für überhängende Enden ist definiert als die Enzymmenge, die 2,4 pmol überhängende Enden in 30 min bei 16°C in einem 20  $\mu$ l Ansatz zu 50 % ligiert. In der Praxis wurde ein Ligationsansatz gewählt, der etwa 500 ng Vektor-DNA sowie den zweifachen bis dreifachen molaren Überschuss an Insert-DNA umfassen sollte. Die eingesetzten Volumina an Vektor-DNA und Insert-DNA wurden je nach Konzentration und Größe von Vektor-DNA und Insert-DNA gewählt:

Vektor DNA	x $\mu$ l
Insert-DNA	x $\mu$ l
T4 DNA-Ligase	1 $\mu$ l
10x T4 DNA-Ligase Reaktionspuffer	1 $\mu$ l
H <sub>2</sub> O <sub>dd, steril</sub>	auf 10 $\mu$ l auffüllen

Dieser Ansatz wurde 3 h bei 25°C inkubiert.

#### 4.2.7 DNA-Transfer in Bakterien

Beim DNA-Transfer in Bakterien unterscheidet man prinzipiell zwischen Transduktion, Konjugation und Transformation. Während Transduktion die Übertragung von DNA von Viren auf Bakterien und Konjugation den Austausch von DNA zwischen Bakterien über eine sich ausbildende Plasmabrücke beschreibt, versteht man unter Transformation die Aufnahme freier DNA in Bakterien. Die Aufnahme freier DNA ist mit wenigen Ausnahmen nur nach einer Vorbehandlung des Empfängerbakteriums möglich. Eine bakterielle Zelle, die zur Transformation befähigt ist, wird auch als kompetente Zelle bezeichnet. Für *E. coli* gibt es mehrere Methoden der Kompetentmachung, wobei in der vorliegenden Arbeit die klassische CaCl<sub>2</sub>-Methode verwendet wurde.<sup>[84]</sup> Die Transformation in *A. baohimycina* erfolgte in der

vorliegenden Arbeit durch die sog. direkte Transformation. Beide Methoden sollen im Folgenden beschrieben werden.

#### 4.2.7.1 Herstellung von und Transformation in CaCl<sub>2</sub>-kompetente *E. coli*

Die Herstellung CaCl<sub>2</sub>-kompetenter *E. coli* basiert auf der Behandlung mit einer hohen Konzentration divalenter Kationen, hier Ca<sup>2+</sup>, bei Temperaturen nahe dem Gefrierpunkt.<sup>[84]</sup> Die Hintergründe der dabei ablaufenden Vorgänge konnten jedoch bis heute nicht vollständig aufgeklärt werden. In der vorliegenden Arbeit wurde zunächst eine Hauptkultur des gewünschten *E. coli*-Stammes in TY-Medium (100 ml) mit einer 16 h alten Vorkultur beimpft (1 % Inokulum) und bei 37°C bis zu einer OD<sub>578</sub> von 0,5 – 0,8 auf dem Schüttelinkubator kultiviert. Die Hauptkultur wurde durch Zentrifugation (10 min; 4.200 g, 4°C) pelletiert und der Überstand quantitativ verworfen. Alle weiteren Schritte erfolgten auf Eis und mit gut vorgekühlten Lösungen. Das Pellet wurde durch vorsichtiges Triturieren mit steriler CaCl<sub>2</sub>-Lösung (5 ml; 100 mM) resuspendiert und erneut pelletiert. Nach Resuspension in eiskalter steriler CaCl<sub>2</sub>-Lösung (40 ml; 50 mM) und erneuter Zentrifugation (10 min; 4.200 g, 4°C) wurde das Pellet in eiskalter steriler CaCl<sub>2</sub>-Lösung (1,85 ml; 100 mM) aufgenommen und für 16 h auf Eis inkubiert. Dieser Schritt sollte die Kompetenz der Zellen verbessern und dadurch die Transformationseffizienz erhöhen. Die kompetenten Zellen wurden anschließend entweder direkt zur Transformation eingesetzt oder nach Zugabe von Glycerin-Lösung (60 %; ¼-Volumen des Ansatzes) zur Lagerung im -80°C-Schrank eingefroren.

Zur Transformation von DNA in CaCl<sub>2</sub>-kompetente *E. coli* wurde in der vorliegenden Arbeit das Protokoll von Sambrook und Russell angewendet.<sup>[85]</sup> Die kompetenten *E. coli*-Zellen (200 µl) wurden auf Eis aufgetaut und mit der zu transformierenden DNA (1 bis 10 µl) versetzt. Nach Inkubation auf Eis (30 min) schloss sich ein Hitzeschock (90 s) bei 42°C an, wobei man davon ausgeht, dass die Aufnahme der DNA bei 0°C stattfindet und durch den Hitzeschock beendet wird. Nach Beendigung des Hitzeschocks wurde sofort TY-Medium (1 ml) zugegeben und 40 min bei 37°C zur Entwicklung der eingebrachten Resistenz inkubiert. Nach einem Zentrifugationsschritt (1 min; 21.000 g) wurde der Überstand nahezu quantitativ entfernt, das Zellpellet im Rücklauf (ca. 100 µl) resuspendiert und auf einer TY-Platte mit dem entsprechenden Antibiotikum mit einer sterilen Glaspipette (1 ml)

ausplattiert. Die Platten wurden anschließend bei 37°C im Brutschrank inkubiert (16 h bzw. bis das Wachstum einzelner Kolonien beobachtet werden konnte).

#### 4.2.7.2 Direkte Transformation von *A. balhimycina*

Die Methode der direkten Transformation wurde zum ersten Mal 1991 für den Stamm *Amycolatopsis (Nocardia) mediterranea*<sup>[86]</sup> beschrieben und seither außerdem auf die Stämme *Nocardia lactamdurans*<sup>[87]</sup> und *Amycolatopsis methanolica*<sup>[88]</sup> angewendet. Die direkte Transformation hat ihren Namen daher, dass Plasmid-DNA direkt in Mycel anstatt wie normalerweise bei Streptomyceten üblich in Protoplasten transformiert wird. Die Anwesenheit von Alkalimetallkationen wie Cs<sup>+</sup> und Polyethylenglykol ist für diese Methode absolut essentiell, wobei die Zugabe von Kalbsthymus-DNA als Carrier die Transformationseffizienz drastisch erhöht. Da bei *A. balhimycina* keine Regenerierung der Protoplasten erreicht werden konnte, kam auch hier die direkte Transformation zum Einsatz. Pelzer et al. publizierten 1997 ein auf *A. balhimycina* optimiertes Protokoll,<sup>[89]</sup> welches auch in der vorliegenden Arbeit verwendet wurde.

- Kultivierung wie in Abschnitt 4.2.2 beschrieben
- 25 mL Kultur pelletieren (10 min, 4200 g)
- Pellet in TE<sub>10,1</sub>-Puffer (25 mL) waschen und dazu mit steriler Glaspipette resuspendieren
- Pellet in TE<sub>10,1</sub>-Puffer (1,0 mL) aufnehmen und Aliquots herstellen (100 µl)
- Zugabe zu jedem Ansatz in dieser Reihenfolge und vorsichtig mischen:
  - 10 µl MgCl<sub>2</sub> (0,2 M)
  - 60 µl CsCl (4,17 M)
  - 4 µl Kalbsthymus-DNA (7,5 µg/µl)
  - 26 µl Transformationsplasmid pC5A5T4E4
  - 200 µl PEG1000 (65 %)
- Inkubation bei 30°C (40 min)
- Zugabe von S27M-Medium (1 mL) und mischen
- Zellen pelletieren (4 min, 6000 g) und Überstand verwerfen
- erneute Zugabe von S27M-Medium (1 mL), mischen
- Zellen pelletieren (4 min, 6000 g) und Überstand verwerfen
- Zellen in S27M-Medium (400 µl) aufnehmen und zusammen mit 2,5 mL R2L-Weichagar auf einer S27M-Platte verteilen, trocknen (15 min)

- Inkubation bei 30°C (14 bis 16 h)
- Platte mit Erythromycin-haltigem NB-Weichagar (3 mL) überschichten
- Inkubation bei 30°C, bis die Klone deutlich sichtbar werden (ca. 8 Tage)

Um sicher zu sein, dass die Klone wirklich Erythromycin-resistent sind, wurden sie noch einmal auf Erythromycin-haltigen R5-Platten ausgestrichen.

#### **4.2.8 Stressprotokoll zur Forcierung von homologen Rekombinationsereignissen in *A. balhimycina***

Da eine homologe Rekombination in *A. balhimycina* relativ selten auftritt, wurde von Puk et al. 2002 ein sog. Stressprotokoll publiziert, welches die Wahrscheinlichkeit für ein solches Ereignis drastisch erhöht.<sup>[30]</sup> Zum Ausüben des nötigen Stresses wird zum einen Hitze (37°C) und zum anderen eine Ultraschallbehandlung angewendet. Anschließend erfolgt ein Protoplastierungsschritt, der zum einen die Zellen vereinzelt und damit für einen eindeutigen Genotyp sorgt und zum anderen selbst als Stress angesehen werden kann. In der vorliegenden Arbeit wurde folgendes Vorgehen gewählt, um *double crossover*-Ereignisse zu forcieren:

- Kultivierung wie in Abschnitt 4.2.2 beschrieben, allerdings erfolgen die letzten 16 h bei erhöhter Temperatur (37°C)
- Ultraschallbehandlung der Kultur (10 min)
- 25 mL Kultur pelletieren (10 min, 25900 g)
- Pellet in P-Puffer (20 mL) mit frisch zugesetztem Lysozym (10 mg/mL) resuspendieren
- Inkubation bei 30°C und 80 rpm auf dem Inkubationsschüttler, regelmäßige Mikroskopkontrolle der Protoplastierung
- bei deutlichem Fortschritt der Protoplastierung (ca. 1 bis 1,5 h), Ansatz 3x mit Glaspipette (20 mL) aufziehen und ohne Druck auslaufen lassen
- Pelletierung der Mycelreste (7 min, 800 g)
- Überstand mit P-Puffer versetzen (5 mL) und vorsichtig mischen
- Pelletierung der Protoplasten (10 min, 4200 g) und Überstand verwerfen
- Protoplastenpellet in P-Puffer waschen (7,5 mL)
- Protoplasten im Rücklauf resuspendieren und in verschiedenen Verdünnungsstufen auf R5-Platten ohne Selektion ausstreichen (unverdünnt, 1:10, 1:100, 1:1000)
- Inkubation bei 30°C, bis die Klone deutlich sichtbar werden (1 bis 2 Wochen)

Die Klone wurden jeweils auf Erythromycin-haltige R5-Platten und auf R5-Platten ohne Selektion überimpft. Da die gewünschten *double crossover*-Klone die Erythromycinresistenz verloren hatten, konnten sie durch ihr Wachstum auf R5-Platten ohne Selektion und durch das Ausbleiben des Wachstums auf Erythromycin-haltigen R5-Platten erkannt werden.

#### 4.2.9 Plasmidisolierung aus *E. coli*

Zur Plasmidisolierung aus *E. coli* wurden in der vorliegenden Arbeit zwei Protokolle verwendet. Während Protokoll 1 mit ausschließlich selbst hergestellten Lösungen die kostengünstigere Variante darstellte, erbrachte Protokoll 2 unter Verwendung des *peqGOLD Plasmid Miniprep-Kits I* Plasmid-DNA von höherer Reinheit. Beide Protokolle führen den bakteriellen Zellaufschluss sowie die Abtrennung genomischer DNA über die sog. alkalische Lyse herbei, die auf Arbeiten von Birnboim und Doly<sup>[90]</sup> zurückgeht. Die so erhaltene Plasmid-DNA-Lösung wird im Fall von Protokoll 1 über eine Ethanolfällung aufkonzentriert, während Protokoll 2 durch die Verwendung von Silikagelsäulen (Prinzip s. Abschnitt 4.2.6.2) neben der Konzentrierung auch einen weiteren Reinigungsschritt beinhaltet. Die Durchführung von Protokoll 2 erfolgte gemäß den Herstellerangaben. Die Durchführung von Protokoll 1 sowie die in den einzelnen Schritten ablaufenden Prozesse sollen im Folgenden kurz beschrieben werden.

- Abzentrifugation (1 min, 16000 g) einer 16 h Flüssigkultur (2 ml)
- Resuspension des Pellets in AL1 (100 µl)
- Zugabe von AL2 (200 µl) und Umdrehen des Reaktionsgefäßes (5x)

Durch das in AL2 vorhandene alkalische Milieu sowie die Anwesenheit des Detergens SDS wird die Lyse der Bakterienzellen herbeigeführt. Bei diesen alkalischen Bedingungen kommt es zur Denaturierung sowohl der genomischen DNA als auch der Plasmid-DNA, wobei sich die beiden Stränge eines zirkulären Plasmids aufgrund ihrer Topologie nicht vollständig voneinander trennen können. Parallel dazu baut die in AL1 enthaltene *RNase A* die freigesetzte RNA ab.

- Zugabe von AL 3 (150µl), Umdrehen des Reaktionsgefäßes (3x) und Inkubation auf Eis (3 bis 5 min)

Durch Zugabe der sauren Kaliumacetat-Lösung AL3 wird die Lösung neutralisiert, wodurch es zur Renaturierung der DNA kommt. Während sich die

Plasmid-DNA vollständig renaturieren kann, kommt es bei der genomischen DNA zur unspezifischen Verknüpfung mehrerer Stränge und damit zu deren Präzipitation. Wegen der hohen Kaliumacetat-Konzentration fallen außerdem Protein-SDS-Komplexe aus.

- Zentrifugation (5 min, 16000 g) zur Präzipitatabtrennung
- Überführung des Überstandes in ein neues Reaktionsgefäß, Zugabe von 100 % Ethanol (2 Volumen) und Vortexen

Das Prinzip der Ethanolfällung beruht darauf, dass DNA in Gegenwart monovalenter Kationen in Ethanol einen unlöslichen Niederschlag bildet.

- Zentrifugation (5 min, 16000 g) zur Fällung des Niederschlags
- Waschen der DNA in 70 % Ethanol (1 ml), um mitgefälltes Salz zu entfernen
- Trocknen der DNA bei 60°C
- Lösen der DNA in H<sub>2</sub>O<sub>dd, steril</sub> (50 µl)

#### **4.2.10 Isolierung von Gesamt-DNA aus *A. balhimycina***

Bei der Isolierung von Gesamt-DNA müssen hauptsächlich zwei Aufgaben erfüllt werden: Aufschluss der Zelle und Reinigung der Gesamt-DNA bzw. Abtrennung von vor allem Protein.

Der Zellaufschluß erfolgt typischerweise durch Zugabe von Detergenzien wie z.B. SDS, wobei bei Bakterien außerdem Lysozym zum Abbau der Zellwand zugegeben werden muß. Für die Abtrennung der Proteine von der DNA gibt es unterschiedliche Möglichkeiten, z.B. das Ausschütteln mit gepuffertem Phenol oder mit gepuffertem Phenol und Chloroform/Isoamylalkohol (24:1). Durch die Pufferung löst sich die DNA nicht in der Phenolphase, sondern verbleibt in der wässrigen Phase. Die Proteine denaturieren jedoch durch das Phenol und sammeln sich entweder in der sog. Interphase zwischen der wässrigen DNA-Lösung und der Phenolphase an oder sie lösen sich in der Phenolphase. Diese Phenolextraktion reicht aber nicht für die Abtrennung des kompletten Zellproteins, weshalb man sich des Enzyms Proteinase K bedient. Diese Methode kann als Ergänzung oder auch geruchsfreie Alternative zur Phenolextraktion verwendet werden. Proteinase K ist eine Subtilisin-verwandte Serinprotease, die Proteine proteolytisch bevorzugt C-terminal von aromatischen, aliphatischen oder hydrophoben Aminosäuren spaltet. Bei langen Inkubationszeiten und hohem Überschuss kann der Verdau bis zu den freien Aminosäuren führen. Eine dritte Möglichkeit ist die Verwendung von Guanidinium-Hydrochlorid, welches zum

einen den Zellaufschluß unterstützt und zum anderen Proteine vollständig denaturiert.

In der vorliegenden Arbeit erfolgte die Isolierung von Gesamt-DNA aus *A. balhimycina* ausschließlich mit Hilfe des *NucleoSpin Tissue-Kits* von MACHEREY-NAGEL. Dieses Kit verwendet alle oben beschriebenen Prinzipien außer der Phenolextraktion. Stattdessen kommen Silikagelsäulchen zum Einsatz. Das Prinzip der DNA-Bindung an Silikagel wurde ausführlich in Abschnitt 4.2.6.2 beschrieben. Durch die bei diesem Reinigungsschritt auftretenden Scherkräfte wird die hochmolekulare genomische DNA zwar zum Teil in kleinere Bruchstücke zerlegt, die erhaltene Qualität der genomischen DNA reicht aber für Anforderungen wie PCR und *Southern Blot*-Hybridisierung aus. Das in der vorliegenden Arbeit verwendete Protokoll soll im Folgenden beschrieben werden:

- Kultivierung wie in Abschnitt 4.2.2 beschrieben
- Pelletieren von 2 mL Kultur (5 min, 19000 g)
- Pellet in TE/10 % Saccharose-Lösung (500 µl) waschen
- Pellet in Lysepuffer (170 µl, je nach Pelletgröße eventuell mehr) resuspendieren
- Inkubation bei 37°C (50 min)
- Zugabe von Proteinase K (25 µl)
- Inkubation bei 56°C (2 bis 3 h) im schüttelnden Heizblock (800 rpm)
- Zugabe von RNase A (40 mg)
- Inkubation bei 22°C (5 min)
- Zugabe von Puffer B3 (200 µL) und mehrmaliges Umdrehen
- Inkubation bei 70°C (10 min) unter mehrmaligem Umdrehen
- Partikel pelletieren (5 min, 11000 g)
- Überstand mit Ethanol versetzen (210 µL), sofort umdrehen
- Auftragen auf Silikagelsäulchen und durchzentrifugieren (1 bis 5 min, 11000 g)
- Waschen mit Puffer BW (500 µl, 1 min, 11000 g)
- Waschen mit Puffer B5 (600 µl, 1 min, 11000 g)
- Trocknen (3 min, 11000 g)
- Zugabe von 70°C warmem Puffer BE (100 µl)
- Inkubation bei 22°C (1 min)
- Elution (3 min, 11000g) in ein neues Reaktionsgefäß

#### 4.2.11 Southern Blot-Hybridisierung

Die Methode der *Southern Blot*-Hybridisierung geht auf Versuche von E. Southern aus dem Jahre 1975 zurück<sup>[91]</sup> und umfasst mehrere Schritte. Zunächst wird die zu untersuchende DNA mit Hilfe von Restriktionsenzymen verdaut und die dadurch entstehenden DNA-Fragmente gelelektrophoretisch aufgetrennt. Das eigentliche *Blotting* erfolgt im nächsten Schritt und beschreibt den Transfer der DNA-Fragmente aus dem Gel auf eine Membran. Vorher wird das Gel in einer alkalischen Lösung inkubiert, um eine Denaturierung der DNA in Einzelstränge zu erreichen, was für die Hybridisierung nach dem *Blotting* essentiell ist. Nach einem Neutralisierungsschritt ist die Vorbehandlung des Gels abgeschlossen und der DNA-Transfer auf die Membran kann erfolgen. Die treibende Kraft für den Transfer können neben einem elektrischen Feld oder einem angelegten Vakuum auch Kapillarkräfte sein, wobei in der vorliegenden Arbeit ein Vakuumblot durchgeführt wurde. Die am meisten verwendeten Membranen sind Nylonmembranen, deren Aminogruppen auf der Oberfläche durch UV-Bestrahlung kovalente Bindungen zu den Thymidinresten der DNA-Fragmente eingehen und sie so fest fixieren. Die so auf Membran geblotteten DNA-Fragmente werden im nächsten Schritt der Hybridisierung mit einer spezifischen DNA-Sonde unterzogen. Ein sich anschließender Waschschrift soll unspezifisch an der Membran gebundene DNA-Sonde entfernen, um eine hohe Hintergrundfärbung zu vermeiden. Für die Detektion der Hybridisierungsereignisse gibt es eine Vielzahl an Möglichkeiten, wobei in der vorliegenden Arbeit die Digoxigenin-Strategie verwendet wurde. Die dafür notwendige DNA-Sonde wird über eine PCR-Reaktion hergestellt, der neben den gewöhnlichen Nukleotiden auch Digoxigenin markierte dUTPs sog. DIG-dUTPs zugesetzt werden. So entsteht eine DIG-markierte DNA-Sonde, die nach der Hybridisierung durch einen DIG-spezifischen Antikörper mit gekoppeltem Enzym, welches eine detektierbare Reaktion katalysiert, lokalisiert werden kann. Um eine unspezifische Bindung des Antikörpers an die Membran zu vermeiden, wird die Membran vor Zugabe des Antikörpers mit einer *Blocking*-Lösung inkubiert. In der vorliegenden Arbeit wurde zur Detektion das *DIG Luminescent Detection-Kit* verwendet. Dieses Kit enthält den gegen Digoxigenin gerichteten Antikörper mit einer gekoppelten alkalischen Phosphatase (Anti-DIG-AP), *Blocking*-Reagenz und CSPD als Substrat. Die alkalische Phosphatase katalysiert die Dephosphorylierung von CSPD zu einer metastabilen Verbindung, die unter Chemilumineszenz ( $\lambda = 477 \text{ nm}$ ) zerfällt. Diese

kann zur Detektion verwendet werden, z.B. durch Auflegen und anschließende Entwicklung eines Röntgenfilms. Die praktische Durchführung der *Southern Blot*-Hybridisierung erfolgte in der vorliegenden Arbeit wie im Folgenden beschrieben:

- Isolierung von Gesamt-DNA gemäß Abschnitt 4.2.10
- Restriktionsverdau von Gesamt-DNA gemäß Abschnitt 4.2.6.1
- gelelektrophoretische Auftrennung unter Verwendung DIG-gelabelter DNA-Leitern
- Gel kurz in Ethidiumbromidbad färben und fotografieren
- Gel mit Denaturierungs-Lösung (250 mL) auf Taumeltisch inkubieren (2x20 min)
- Gel mit Neutralisierungs-Lösung (250 mL) auf Taumeltisch inkubieren (2x20 min)
- *Vakuumblotting*:
  - Blot gemäß den Herstellerangaben der Vakuumblotapparatur aufbauen
  - Gel mit 20x SSC bedecken und Vakuum anlegen (50 mbar)
  - nach 50 min Blot abbauen und Orientierung der Membran markieren
- Membran an der Luft trocknen
- UV-Bestrahlung der Membran (2 min)
- Membran mit der DNA nach innen in Hybridisierungsglasgefäß geben, so dass die folgenden Schritte im Hybridisierungsofen unter ständiger Rotation durchgeführt werden können
- Membran bei 68°C mit Hybridisierungs-Lösung (20 mL) vorinkubieren (1,5 bis 2 h)
- die gelgereinigte DIG-Sonde (Herstellung s. Abschnitt 4.2.3) bei 100°C denaturieren (10 min) und auf Eis abkühlen
- Membran bei 68°C mit Hybridisierungs-Lösung (20 mL) und DIG-Sonde (80 µl) inkubieren (16 h)
- Membran bei 22°C mit 2xSSC/0,1 %SDS (50 mL) waschen (2x10 min)
- Membran bei 68°C mit 0,1xSSC/0,1 %SDS (50 mL) waschen (2x15 min)
- Membran bei 22°C mit Waschpuffer (50 mL) waschen (5 min)
- Membran bei 22°C mit *Blocking*-Lösung (20 mL) inkubieren (30 min)
- Membran bei 22°C mit *Blocking*-Lösung (20 mL) und Anti-DIG-AP-Lösung (1 µl) inkubieren (30 min)
- Membran bei 22°C mit Waschpuffer (50 mL) waschen (2x15 min)

- Membran mit Detektionspuffer (50 mL) äquilibrieren (3 min)
- Membran mit Detektionspuffer (1 mL) und CSPD-Lösung (10 µl) benetzen, zwischen zwei Folien legen, mit Alufolie vor Licht schützen und bei 22°C inkubieren (5 min)
- überschüssige Lösung ausstreichen, Folien versiegeln
- Inkubation bei 37°C (5 min)
- Röntgenfilm auf die Membran auflegen und unter Lichtausschluss inkubieren (45 min)
- Röntgenfilm entwickeln

#### **4.2.12 Isolierung von nicht-ribosomalen Peptidsynthetasen aus *A. balhimycina***

Da die nicht-ribosomalen Peptidsynthetasen intrazellulär in *A. balhimycina* vorliegen, mussten die Zellen zunächst aufgeschlossen werden. Bei Gram-positiven Bakterien sind die dafür gängigsten Methoden die enzymatische Lyse der Zellwand mittels Lysozym und der Zellaufschluss mittels *French Press*, wobei letztere bei *A. balhimycina* verwendet wurde. Die Durchführung erfolgte in einem sog. Proteinisolationpuffer, dessen Zusätze für die Integrität der Proteine sorgten. Neben dem reduzierenden Agens DTT, welches die Oxidation der Proteine durch Luftsauerstoff verhinderte, waren mehrere Proteaseinhibitoren enthalten. Aufgrund der hochmolekularen DNA wies der so erhaltene Zellaufschluss eine sehr zähflüssige Konsistenz auf. Die Abtrennung der DNA erfolgte durch enzymatische Zerkleinerung der DNA mittels *DNase* I und anschließende Fällung mittels Polymin P. Diese stark basische Substanz aggregiert aufgrund ihrer Ladung mit der entgegengesetzt geladenen DNA und führt so zu deren Fällung. Allerdings darf die Konzentration an Polymin P nicht zu hoch gewählt werden, da sonst auch Proteine mitgefällt werden. Im nächsten Schritt wurden die Proteine mit Ammoniumsulfat gefällt, was auch als „Aussalzen“ bezeichnet wird. Ammoniumsulfat eignet sich dafür besonders, da es als antichaotropes Salz hydrophobe Effekte und somit die Proteinaggregation fördert. Um das so erhaltene Proteinpellet zum einen vom Ammoniumsulfat-Salz zu befreien und zum anderen in seine verschiedenen Bestandteile aufzutrennen, wurde eine Größenausschlusschromatographie mit Sephadex G50-Material angeschlossen. Das zugrunde liegende Prinzip beruht darauf, dass kleine Moleküle in die Poren des Materials eindringen können und dadurch mehr Zeit für den Durchtritt durch die Säule brauchen als große Moleküle, die an den Poren vorbei direkt von der Säule eluiert werden können. Diese Strategie eignete sich besonders für die nicht-ribosomalen Peptidsynthetasen, da sie zu den größten Proteinen einer Zelle gehören und somit vor anderen Molekülen in den ersten Fraktionen von der Säule eluiert werden.

Die praktische Durchführung dieser Schritte erfolgte in der vorliegenden Arbeit nach folgendem Protokoll, wobei von jeder Stufe Proben genommen wurden, die mittels Bradford-Konzentrationsbestimmung den Aufreinigungsverlauf dokumentieren sollten:

- Vorbereitung der Sephadex G50-Säule:
    - 4 g Säulenmaterial in H<sub>2</sub>O<sub>dd.</sub> (100 mL) quellen lassen (16 h)
    - entgasen am Wasserstrahlvakuum
    - in Glassäule gießen und langsam absetzen lassen
    - vor Verwendung mit Proteinisolierungspuffer äquilibrieren
  - Kultivierung von *A. balhimycina* wie in Abschnitt 4.2.2 beschrieben
- von diesem Schritt an sollte möglichst immer bei 4°C bzw. auf Eis gearbeitet werden
- Ernten der Zellen durch Abzentrifugieren (4200 g, 30 min), sofort weiterverarbeiten oder bei -80°C einfrieren
  - 4 g Zellen (Naßgewicht) zweimal in Proteinisolierungspuffer waschen
  - Zellen in Proteinisolierungspuffer (15 mL) resuspendieren
  - Zellaufschluss in der vorgekühlten *French Press* (10000psi)
  - Spatelspitze DNaseI und MgCl<sub>2</sub> (Endkonzentration 10 mM) zugeben
  - Inkubation auf Taumeltisch, bis zähflüssiger Charakter der Lösung verschwunden ist (0,5 h)
  - Zentrifugation (27000 g, 1 h)
  - DNA aus dem Überstand mit Polymin P (Endkonzentration 0,05 %) fällen (1 h)
  - Zentrifugation (27000 g, 0,5 h)
  - Proteine aus dem Überstand mit Ammoniumsulfat (Endkonzentration 66 % Sättigung) fällen (2,5 h)
  - Zentrifugation (27000 g, 0,5 h)
  - Proteinpellet in Proteinisolierungspuffer (2 mL) resuspendieren
  - Proteinsuspension auf eine Sephadex G50-Säule auftragen und mit Proteinisolierungspuffer eluieren
  - Fraktionen sammeln: Totvolumen, danach immer 1 mL Fraktionen

#### 4.2.13 Photometrische Konzentrationsbestimmung von Proteinen nach Bradford

Bei der quantitativen Proteinbestimmung nach Bradford handelt es sich um die empfindlichste, auf einer Färbereaktion beruhende Methode, die allerdings für verschiedene Proteine sehr unterschiedliche Ergebnisse liefern kann.<sup>[92]</sup> Das Prinzip beruht darauf, dass der blaue Säurefarbstoff Coomassie Brilliantblau G 250 in Gegenwart von Proteinen und in saurem Milieu sein Absorptionsmaximum von  $\lambda = 465$  nach  $\lambda = 595$  nm verschiebt. Der Grund hierfür ist wahrscheinlich die Stabilisierung des Farbstoffs in seiner unprotonierten, anionischen Sulfonatform durch Komplexbildung mit Proteinen. Der Farbstoff bindet dabei recht unspezifisch an kationische und nichtpolare, hydrophobe Seitenketten der Proteine, vor allem an Arginin. Als Maß für die Protein-Konzentration kann deshalb die Absorption bei einer Wellenlänge zwischen  $\lambda = 570$  und  $\lambda = 610$  nm herangezogen werden. Als Referenz wird die Absorption einer Leerprobe des verwendeten Puffers ohne Protein gemessen und auf Null gesetzt. Um die Genauigkeit der Ergebnisse zu erhöhen, werden in der Praxis Mehrfachbestimmungen durchgeführt.

In der vorliegenden Arbeit wurde mit einer Bradford-Stammlösung gearbeitet, die den zu untersuchenden Proben zugegeben wurde. Nach Inkubation (10 min) bei 22°C wurde die Absorption dieser Ansätze bei  $\lambda = 570$  nm gemessen.

Um eine grobe Abschätzung der Absolutkonzentrationen vornehmen zu können, wurde eine Kalibriergerade mit Albumin erstellt. Dazu wurden ausgehend von zwei unabhängig voneinander hergestellten Albumin-Stammlösungen (1 mg/mL) in Proteinisolierungspuffer zwei Konzentrationsreihen erstellt. Das zugehörige Pipettierschema ist in Tabelle 2 abgebildet.

**Tabelle 2: Herstellung einer Albumin-Konzentrationsreihe aus einer 1 mg/mL Stammlösung.**

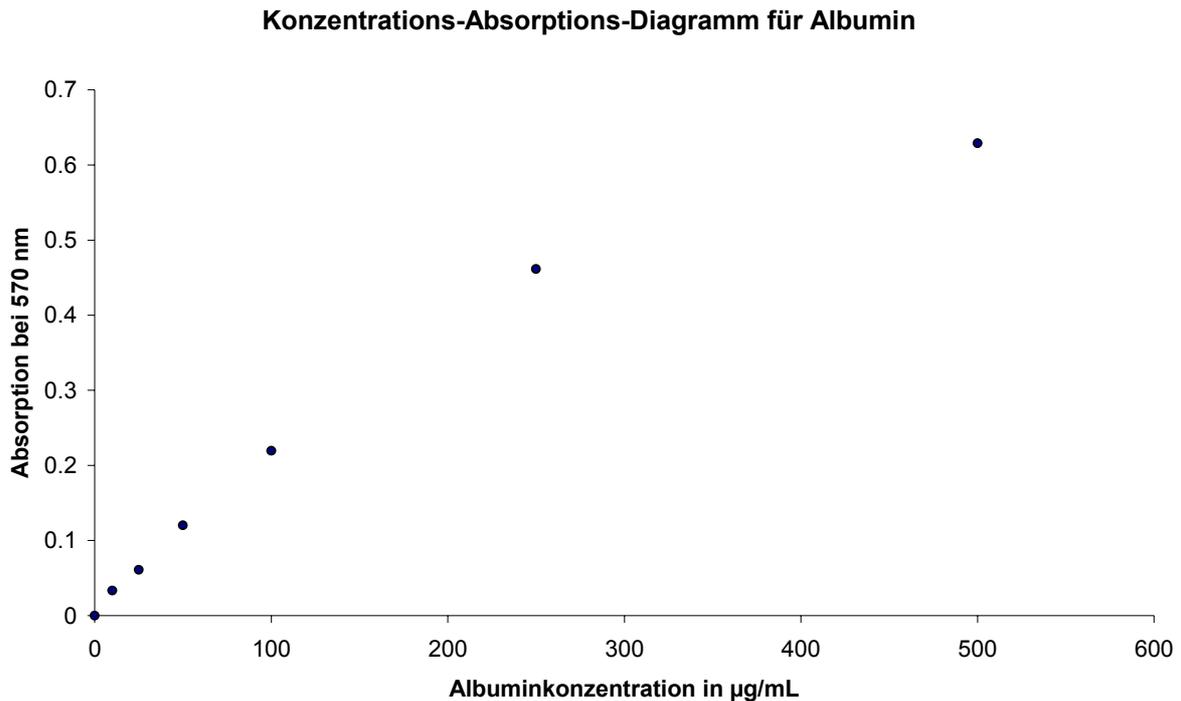
Albumin-Konzentration [ $\mu\text{g/mL}$ ]	Zugabe Stammlösung [ $\mu\text{l}$ ]	Zugabe Proteinisolierungspuffer [ $\mu\text{l}$ ]
10	10	990
25	25	975
50	50	950
100	100	900
250	250	750
500	500	500

Die Albumin-Lösungen (10 µl) wurden mit 1,1x Bradford-Stammlösung (100 µl) versetzt und bei 22°C inkubiert (10 min). Anschließend wurde von allen Ansätzen die Absorption bei  $\lambda = 570$  nm im Photometer mittels Dreifachbestimmung gemessen. Die Messwerte sind in Tabelle 3 dargestellt.

**Tabelle 3: Bestimmung der Absorptionen der zwei Albumin-Konzentrationsreihen bei  $\lambda = 570$  nm. Jede Probe wurde dreimal gemessen und anschließend der Mittelwert ermittelt. Abschließend wurden die Mittelwerte der beiden Messreihen gemittelt.**

Albuminkonz. [µg/mL]	Messreihe 1		Messreihe 2		Mittelwert aus Messreihe 1 und 2
	Dreifachbest.	Mittelwert	Dreifachbest.	Mittelwert	
0 (Referenz)	0,000	0,000	0,000	0,000	0,000
10	0,023	0,035	0,011	0,032	0,034
	0,039		0,034		
	0,043		0,051		
25	0,041	0,051	0,062	0,071	0,061
	0,051		0,075		
	0,062		0,075		
50	0,116	0,118	0,122	0,122	0,120
	0,113		0,118		
	0,125		0,127		
100	0,220	0,219	0,225	0,220	0,219
	0,218		0,216		
	0,219		0,218		
250	0,446	0,464	0,452	0,459	0,461
	0,477		0,453		
	0,468		0,473		
500	0,593	0,625	0,645	0,633	0,629
	0,633		0,636		
	0,649		0,617		

Die Mittelwerte der beiden Messreihen wurden gemittelt und gegen die entsprechenden Albumin-Konzentrationen aufgetragen. Wie man in Abbildung 4-1 sieht, sinkt die Linearität zwischen Konzentration und Absorption bei steigenden Werten.



**Abbildung 4-1: Konzentrations-Absorptions-Diagramm für Albumin. Es sind die ermittelten Absorptionswerte bei  $\lambda = 570$  nm gegen die entsprechende Albumin-Konzentration in  $\mu\text{g/mL}$  aufgetragen. Die Werte sind aus zwei unabhängigen Messreihen gemittelt, wobei jede einzelne Probe dreimal vermessen wurde.**

Deshalb wurden zur Bestimmung der Kalibriergeraden durch den Ursprung auch nur die Albumin-Konzentrationen bis einschließlich  $250 \mu\text{g/mL}$  berücksichtigt. Daraus ergab sich folgende Geradengleichung, die zur näherungsweisen Bestimmung der Protein-Konzentration während der Isolierung der nicht-ribosomalen Peptidsynthetasen herangezogen wurde:

$$\text{Absorption}_{570\text{nm}} = 0,0019 \times \text{Albumin-Konzentration } [\mu\text{g/mL}]$$

$$\text{Albumin-Konzentration } [\mu\text{g/mL}] = \text{Absorption}_{570\text{nm}} : 0,0019$$

#### 4.2.14 Gelelektrophorese von Proteinen

Die Gelelektrophorese ist nicht nur eine wichtige Methode zur Analyse von DNA, sondern auch von Proteinen. Allerdings gibt es einige grundlegende Unterschiede zwischen diesen beiden Gelelektrophoresearten, die auf der Verschiedenheit der beiden Verbindungsklassen beruhen. Für die meisten Proteine sind Agarosegele zu großporig, weshalb man hier typischerweise vertikale Polyacrylamid-Gele verwendet. Die radikalische Kopolymerisation zwischen Acrylamid und dem Quervernetzer *N,N'*-Methylenbisacrylamid wird hierbei durch den Radikalstarter Ammoniumperoxodisulfat (APS) und den Radikalüberträger Tetramethylethyldiamin (TEMED) gesteuert. Die in der vorliegenden Arbeit verwendete Elektrophoresemethode wird als diskontinuierliche SDS-Polyacrylamid-Gelelektrophorese (Disk-SDS-PAGE) bezeichnet, was im Folgenden kurz erläutert werden soll: Da Proteine im Gegensatz zu DNA ganz unterschiedliche Nettoladungen aufweisen können, bedient man sich des anionischen Detergens SDS, welches sich über hydrophobe Wechselwirkungen an die Proteine anlagert und so für ein konstantes Molekülmasse-zu-Ladungs-Verhältnis sorgt. Dadurch ist es möglich, alle Proteine in Laufrichtung Anode aufzutrennen. Weiterhin charakteristisch ist, dass mit zwei verschiedenen Gelzonen gearbeitet wird, dem sog. Sammel- und Trenngel. Im Sammelgel gilt das Elektrophoreseprinzip der Isotachophorese, die sich durch ein diskontinuierliches Puffersystem aus Tris/Chlorid (pH 6,8) im Sammelgel und Tris/Glycin (pH 8,9) im Laufpuffer auszeichnet. Nach dem Auftragen der Proben auf das Sammelgel und dem Anlegen der Spannung trifft Glycin auf den pH des Sammelgels, welcher seinem isoelektrischen Punkt sehr nahe kommt. Dadurch hat Glycin eine niedrige elektrophoretische Mobilität und wird so zum Folge-Ion. Im Gegensatz dazu haben die Chlorid-Ionen eine hohe elektrophoretische Mobilität und werden dementsprechend als Leit-Ionen bezeichnet. Die in der Probe enthaltenen Proteine liegen sowohl örtlich als auch von ihrer elektrophoretischen Mobilität her zwischen dem Leit- und dem Folge-Ion. Es bildet sich ein Feldstärkegradient über die nach ihrer elektrophoretischen Mobilität sortierten Analyten als molekulare Stapel aus und alle Ionen laufen mit der gleichen Geschwindigkeit in Richtung Anode. Dieses System regelt sich selbst, da zu schnelle oder zu langsame Ionen durch die dort herrschende kleinere oder größere Feldstärke wieder an ihren Platz befördert werden. Treffen die Proteine schließlich auf das engmaschigere Trenngel, erfahren sie plötzlich einen großen Reibungswiderstand, der zu einer Aufkonzentrierung führt.

Das Folge-Ion Glycin ist aufgrund seiner Größe davon nicht betroffen, überholt die Proteine und beendet damit das Elektrophoreseprinzip der Isotachophorese. Im Trenngel (pH 8,8) liegt ein homogenes Puffersystem und damit eine konstante Feldstärke vor, so dass jetzt eine Auftrennung der Proteine nach dem Prinzip der Zonelektrophorese erfolgt. Durch die Verwendung von SDS besteht über einen weiten Größenbereich eine lineare Beziehung zwischen dem Logarithmus ( $\log_{10}$ ) der Molekülmasse und der relativen Wanderungsdistanz im Gel. In der Praxis verwendet man analog zur Gelelektrophorese von DNA Längenstandards oder Protein-Leitern zur Größenbestimmung. Als Ergänzung muss noch erwähnt werden, dass der Probenauftragspuffer neben dem bereits erwähnten SDS und den schon bei der DNA verwendeten Zusätzen Glycerin und Bromphenolblau auch das reduzierende Agens DTT enthält.

Die praktische Durchführung erfolgte in der vorliegenden Arbeit nach folgendem Protokoll:

- Reinigung der Glasplatten mit Ethanol
- Zusammenbau der Gelkammer nach den Herstellerangaben
- Gießen des Trenngels bis etwa 1 cm unterhalb der Unterkante des Kammes und Überschichten mit Isopropanol
- nach Beendigung der Polymerisation wird Isopropanol entfernt, Kamm positionieren und Sammelgel gießen
- Proben (34  $\mu$ l) mit Probenauftragspuffer (16  $\mu$ l) versetzen und auf 95°C erhitzen (5 min), auf Eis abkühlen, auf das Gel auftragen
- Elektrophoresebedingungen: 70 V / 0,5 h und 120 V / 1,5 h
- Gel aus der Gelkammer nehmen und Orientierung markieren
- Färbung gemäß Abschnitt 4.2.15

#### **4.2.15 Färbung von Proteingelen**

Die häufigste Methode zur Färbung von Proteingelen ist wohl die mit Coomassie Brillantblau G 250, welches auch schon zur Konzentrationsbestimmung von Proteinen nach Bradford verwendet wurde. Diese Methode ist einfach in der Durchführung und hat eine Nachweisgrenze zwischen 100 ng und 1  $\mu$ g in Abhängigkeit des nachzuweisenden Proteins. Da in der vorliegenden Arbeit die Empfindlichkeit der Färbung aufgrund der erwartungsgemäß geringen Konzentration der nicht-ribosomalen Peptidsynthetasen eine entscheidende Rolle spielte, wurde

hier die Silberfärbung mit einer Nachweisgrenze im Subnanogrammbereich verwendet. Das Prinzip der Silberfärbung beruht darauf, dass nach der Fixierung der Proteine im Gel ein Reduktionsschritt mit Natriumdithionit ( $\text{Na}_2\text{S}_2\text{O}_4$ ) erfolgt, welcher die Disulfidbrücken der Proteine zu Thiolgruppen reduziert. Diese Thiolgruppen komplexieren Silber-Ionen, die initiiert von den funktionellen Gruppen und den Peptidbindungen der Proteine zu Silber reduziert werden. Die Zugabe des starken Reduktionsmittels Natriumthiosulfat ( $\text{Na}_2\text{S}_2\text{O}_3$ ) in alkalischem Milieu setzt diesen Prozess fort, zunächst vor allem in der Nähe der bereits gebildeten Silberkeime. Eine pH-Änderung stoppt diesen Prozess rechtzeitig bevor sich überall auf dem Gel metallisches Silber ablagert.

In der vorliegenden Arbeit wurde dazu das *SilverQuest Silver Staining-Kit* verwendet, welches die anschließende massenspektrometrische Analyse der Protein-Banden erlaubt.<sup>[93,94]</sup> Das gewöhnliche Protokoll der Silberfärbung arbeitet zur Steigerung der Empfindlichkeit mit Formaldehyd oder Glutaraldehyd. Während die Vorgänge bei Formaldehyd noch nicht ganz verstanden sind, bindet Glutaraldehyd beispielsweise an Lysinreste und generiert so reduzierende Aldehydgruppen, die Silber-Ionen reduzieren können. Dadurch wird ein anschließender Trypsin-Verdau der Proteine als Vorbereitung für die massenspektrometrische Analyse verhindert. Außerdem geht man davon aus, dass Lysinreste verknüpft werden, was wiederum die Extraktion der Proteine aus dem Gel erschwert. Das verwendete Kit nutzt eine andere, allerdings nicht angegebene Möglichkeit, um dieselbe Empfindlichkeit der Silberfärbung zu erreichen, ohne eben erwähnte Folgeerscheinungen nach sich zu ziehen. Die Durchführung erfolgte nach Herstellerangaben, wobei sehr auf Sauberkeit vor allem im Hinblick auf Hautpartikel geachtet wurde, da sich diese in der anschließenden massenspektrometrischen Analyse der Proteine in Form eines Keratinhintergrunds zeigen.

#### **4.2.16 Massenspektrometrische Analyse von Protein-Banden aus dem Gel**

Die massenspektrometrische Analyse von Protein-Banden aus dem Gel wurde extern von der Firma Proteome Factory AG in Berlin durchgeführt. Die Vorgehensweise war dabei in etwa folgendermaßen: Ausschneiden der silbergefärbten Bande – Entfärbung – Trocknen des Gelstücks – Trypsin-Lösung zugeben – Trypsin-Verdau – Extraktion der Peptid-Fragmente – Aufkonzentrierung in der *Speedvac* – nano-HPLC-ESI-MS/MS – NCBI-Inr-Proteindatenbanksuche.

#### **4.2.17 Vorbereitung der Fermentationskulturen der *A. balhimycina* C5A5T4E4-Mutante für massenspektrometrische Untersuchungen**

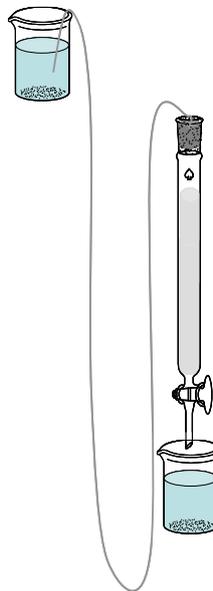
Um eine valide massenspektrometrische Analytik von Fermentationskulturen durchführen zu können, sind diverse Vorbereitungsmaßnahmen erforderlich. Da Glykopeptidmetabolite erfahrungsgemäß von den Zellen nach außen abgegeben werden, werden in einem ersten Schritt die Zellen vom flüssigen Teil der Kultur, dem sog. Kulturüberstand abgetrennt. Für kleinere Volumina eignet sich dafür die Zentrifugation (4200 g, 10 min) am besten, während sich für größere Volumina im Bereich mehrerer Liter die Filtration in einer Mehrschichten-Filterpresse unter Zugabe von 2 % Cellit als Methode der Wahl erwiesen hat. Der somit erhaltene Kulturüberstand ist in seiner Zusammensetzung immer noch sehr komplex, da er neben den gesuchten Glykopeptidmetaboliten auch Medienbestandteile, andere zelluläre Ausscheidungsprodukte usw. enthält. Dieser molekulare Hintergrund verhindert eine empfindliche massenspektrometrische Analyse, so dass lediglich hoch konzentrierte Glykopeptidmetabolite detektiert werden können. Da die Produktivität von Mutanten, die über ein kombinatorisches Biosyntheseexperiment hergestellt wurden, erfahrungsgemäß schlechter ist als die des zugehörigen Wildtyp-Stammes und außerdem neben den Hauptmetaboliten immer auch ebenso interessante Nebenmetabolite geringerer Konzentration produziert werden, ist eine massenspektrometrische Analytik direkt aus dem Kulturüberstand nicht geeignet. Um die Komplexität des Kulturüberstands zu verringern, wurden in der vorliegenden Arbeit zunächst eine Adsorptionschromatographie mit XAD1180-Material und darauf folgend eine präparative Hochdruckflüssigchromatographie (HPLC) durchgeführt.

##### **4.2.17.1 Adsorptionschromatographie mit XAD1180-Material**

Ganz allgemein kann man sagen, dass sich die Trennung bei der Adsorptionschromatographie dadurch ergibt, dass Verbindungen unterschiedlich stark an eine feste stationäre Phase adsorbiert und wieder desorbiert werden. In der vorliegenden Arbeit wurde als stationäre Phase XAD1180 verwendet. Hierbei handelt es sich um ein polymeres Material, welches keine polaren funktionellen Gruppen, sondern aromatische Seitenketten besitzt. Dadurch werden adsorptive Wechselwirkungen mit Molekülen hydrophoben und aromatischen Charakters möglich. Üblicherweise verwendet man als mobile Phase einen Wasser-Methanol-Stufengradient, der bei 100 % Wasser beginnt und dabei zunächst alle nicht

adsorbierten, hydrophilen Verbindungen eluiert. In der vorliegenden Arbeit wird so z.B. der hohe Anteil an Saccharose aus dem Kulturüberstand entfernt. Je höher der Methanolanteil wird, desto eher eluieren die adsorbierten hydrophoben, aromatischen Moleküle. Die apparative Umsetzung und Durchführung der XAD1180-Adsorptionschromatographie erfolgte in der vorliegenden Arbeit wie im Folgenden beschrieben:

Üblicherweise wurde ein Fünftel des Volumens des Kulturüberstands an stationärer Phase XAD1180-Material für die Chromatographie verwendet (Bettvolumen), mit H<sub>2</sub>O am Wasserstrahlvakuum entgast und in eine Glassäule gegossen. Das Betreiben der Säule erfolgte nach dem in Abbildung 4-2 gezeigten Schema, welches ein adäquates Nachfördern des Lösungsmittels gewährleistet.



**Abbildung 4-2: Aufbau der XAD1180-Adsorptionschromatographie.**

Vor dem Erstgebrauch sowie zur Regeneration des XAD1180-Materials wurde die in Tabelle 4 beschriebene Prozedur durchgeführt.

**Tabelle 4: Protokoll für den Erstgebrauch bzw. zur Regeneration des XAD1180-Materials.**

	Lösungsmittel	Volumen
1	1 N Salzsäure/Methanol (1:10)	4 Bettvolumen
2	Aceton	4 Bettvolumen
3	1 N Natronlauge	4 Bettvolumen
4	H <sub>2</sub> O	bis pH neutral

Nach diesen Vorbereitungen war die Säule gebrauchsfertig und konnte mit einer Flussrate von typischerweise 5 Bettvolumen/h betrieben werden. Standardmäßig wird für die XAD1180-Adsorptionschromatographie das in Tabelle 5 beschriebene Protokoll verwendet.

**Tabelle 5: Standardprotokoll für die XAD1180-Adsorptionschromatographie.**

Schritt	Lösungsmittel	Volumen
1 Probe auftragen	Kulturüberstand	5 Bettvolumen
2 Elution	0 % Methanol	4 Bettvolumen
3 „	20 % Methanol	„
4 „	40 % Methanol	„
5 „	60 % Methanol	„
6 „	80 % Methanol	„
7 „	100 % Methanol	„

Allerdings haben Erfahrung und Vorversuche gezeigt, dass sich bei der Anwendung auf Glykopeptide die gesuchten Metabolite vor allem in den Elutionsfraktionen 40 % und 60 % befinden, was zu dem in Tabelle 6 gezeigten abgewandelten Protokoll für Glykopeptide geführt hat, bei dem dreimal mit 40 % Methanol und dreimal mit 60 % Methanol eluiert wird.

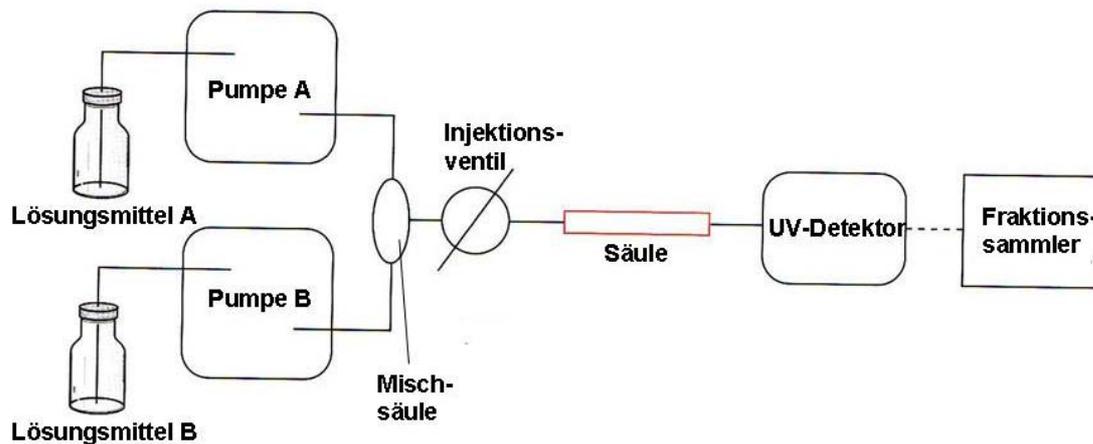
**Tabelle 6: Glykopeptid-Protokoll für die XAD1180-Adsorptionschromatographie.**

Schritt	Lösungsmittel	Volumen	Bezeichnung der Fraktionen
1 Probe auftragen	Kulturüberstand	5 Bettvolumen	Durchlauf
2 Elution	0 % Methanol	4 Bettvolumen	0 %
3 „	20 % Methanol	„	20 %
4 3x Elution	40 % Methanol	3x4 Bettvolumen	40 % (1) bis (3)
5 „	60 % Methanol	„	60 % (1) bis (3)

Die erhaltenen Fraktionen wurden am Vakuum eingedampft und anschließend gefriergetrocknet. An dieser Stelle soll angemerkt werden, dass sich sowohl zur Regeneration der Säule als auch zur Elution das Leitungswasser aus der Hausleitung des Oetker-Gebäudes als am besten erwiesen hat, da es mit pH 7,4 eher einem neutralen pH entsprach als  $H_2O_{deion.}$  oder  $H_2O_{dd.}$ , die beide eher im basischen Bereich lagen. Außerdem wurde der pH des Kulturfiltrats mit NaOH (1 N) von pH 5,6 bis 6,6 auf ca. 7,0 eingestellt.

## 4.2.17.2 Präparative Hochdruckflüssigchromatographie (Präp-HPLC)

Für die Weiterverarbeitung der aus der XAD1180-Adsorptionschromatographie erhaltenen und getrockneten Fraktionen wurde in der vorliegenden Arbeit die Methode der präparativen HPLC verwendet. Der Aufbau und damit die Funktionsweise der hier verwendeten präparativen HPLC-Anlage soll im Folgenden anhand von Abbildung 4-3 erläutert werden.



**Abbildung 4-3: Schemazeichnung der präparativen HPLC-Anlage.**

Die beiden für die Erstellung des mobile-Phase-Gradienten notwendigen Lösungsmittel A und B werden über zwei getrennte Pumpen A und B gefördert und in der Mischsäule gemischt. Diese Anordnung bezeichnet man als „hochdruckseitig“ im Gegensatz zum „niederdruckseitigen“ System, welches nur eine Pumpe nach erfolgter Mischung benötigt, dafür aber an Reproduzierbarkeit einbüßt. Die Probenauftragung erfolgt über ein Injektionsventil, wobei ein Greifarm die im *Autosampler* befindliche Probe abholt, die anschließend von einer Nadel aufgezogen und injiziert wird. Die aufgetragene Probe wird unter hohem Druck und mittels der mobilen Phase über eine Säule geleitet. Hinter der Säule schließen sich ein UV-Detektor, der die Elution der Verbindungen dokumentiert sowie ein Fraktionssammler an, der die Proben zeit- oder UV-getriggert sammelt. Alle Angaben zu Lösungsmittelgradient, Injektionsvolumen, UV-Detektion und Sammelmodus werden in einer Methode abgespeichert und vollautomatisch ausgeführt.

In der vorliegenden Arbeit wurde die Methode der präparativen HPLC zunächst verwendet, um die drei XAD-Fractionen mit 40 % Methanol (40 % (1) bis (3)) und die drei mit 60 % Methanol (60 % (1) bis (3)) weiter aufzutrennen. Für jede XAD-Fraktion mussten dazu mehrere präparative HPLC-Läufe durchgeführt werden, wobei die

analogen Fraktionen vereint und anschließend lyophilisiert wurden. Die verwendeten Parameter sind im Folgenden aufgelistet:

Lösungsmittel A:	H <sub>2</sub> O <sub>dd</sub> + 0,1 % Ameisensäure	
Lösungsmittel B:	Acetonitril + 0,1 % Ameisensäure	
Gradient:	0 min	5 % B
	6 min	15,5 % B
	12 min	60 % B
	13 min	60 % B
	14 min	100 %
Flussrate:	15 mL/min	
Injektionsvolumen:	900 µl (20 – 30 mg)	
Sammelmodus:	zeitgetriggert in 0,2 min-Schritten von 6. Minute bis 14.	
	Minute $\hat{=}$ 40 Fraktionen	
Säule:	C18 Grom-Sil 300 ODS-5	
	Partikelgröße: 10 µm	
	Säulenmaße: 250 mm x 20 mm I. D.	

Die so erhaltenen 240 Fraktionen (6 XAD-Fraktionen ergaben jeweils 40 präp. HPLC-Fraktionen) waren in der Komplexität ihrer Zusammensetzung deutlich geringer als der ursprüngliche Kulturüberstand und konnten nun für die massenspektrometrische Analytik eingesetzt werden.

Zusätzlich wurde die Methode der präparativen HPLC genutzt, um Metabolite weiter, d.h. möglichst bis zur Analysenreinheit zu isolieren. In diesem Zusammenhang wurden auch andere Bedingungen als die oben beschriebenen verwendet, wobei vor allem der Gradient und der Sammelmodus verändert wurden.

#### 4.2.18 Identifizierung der Glykopeptidmetabolite der *A. balhimycina* C5A5T4E4-Mutante mittels HPLC-ESI-MS

Das Prinzip des in der vorliegenden Arbeit verwendeten HPLC-ESI-MS/MS-Systems wurde in Abschnitt 3.1 ausführlich beschrieben. Die praktische Umsetzung, bezogen auf die Identifizierung von Glykopeptidmetaboliten, verlangte die Optimierung von chromatographischen und massenspektrometrischen Parametern. Bezüglich der Chromatographie lag die Herausforderung vor allem darin, Glykopeptide sehr ähnlicher Struktur voneinander zu trennen, was letztlich zu folgendem Chromatographiesystem führte:

Lösungsmittel A:	H <sub>2</sub> O <sub>dd</sub> + 0,1 % Ameisensäure	
Lösungsmittel B:	Acetonitril + 0,1 % Ameisensäure	
Gradient:	0 min	5 % B
	10 min	20 % B
	13 min	50 % B
	14 min	100 % B
	22 min	100 % B
	22,1 min	5 % B
	30 min	5 % B
Flussrate:	50 µl/min (Mikromodus)	
Injektionsvolumen:	5 µl	
Säulensystem:	Jupiter 4u Proteo 90A	
	Vorsäule: 30x1 mm; Hauptsäule: 150x1 mm	

Die Optimierung der massenspektrometrischen Parameter erbrachte folgende Werte, die die höchste Empfindlichkeit bzw. Auflösung für Glykopeptide ermöglichen und deshalb in jeder Methode verwendet wurden:

Ionenquellen-abhängige Parameter:

CUR:	30
IS:	5500
<i>nebulizer</i> gas:	70
turbo gas:	70
TEM:	300
Ionenmodus:	positiv
Quellenposition:	3,5 vertikal / 1 horizontal

Analyt-abhängige Parameter:

DP:	30
EP:	12
CE:	10
Q3 <i>entry barrier</i> :	12

sonstige Parameter:

LIT <i>scan rate</i> :	1000 amu/s
LIT <i>fill time</i> :	<i>dynamic</i>

Die mittels der präparativen HPLC generierten und lyophilisierten Fraktionen wurden in 1 mL 40 % Methanol gelöst, zentrifugiert, um nicht lösliche Bestandteile abzutrennen und schließlich analysiert. Zur Identifizierung der Glykopeptidmetabolite der *A. balhimycina* C5A5T4E4-Mutante wurde der Scanmodus EMS in einem Massenbereich von 250 bis 1000 amu gewählt. Die so generierten Messungen wurden über zwei unterschiedliche Strategien ausgewertet. Die erste Strategie setzte die Berechnung der Molekülmassen der erwarteten Glykopeptidmetabolite voraus, wobei aufgrund der Kenntnisse der Balhimycin-Biosynthese (s. Abschnitt 1.3.2) eine vollständige Chlorierung, aber eine fehlende *N*-Methylierung und Glykosylierung am wahrscheinlichsten waren (s. Tabelle 7). Dadurch konnte über die Funktion des extrahierten Ionenstroms sehr empfindlich nach vermuteten Glykopeptidmetaboliten der *A. balhimycina* C5A5T4E4-Mutante gesucht werden.

**Tabelle 7: Molekülmassen der vermuteten Glykopeptidmetabolite der *A. balhimycina* C5A5T4E4-Mutante. Alle hier aufgeführten Glykopeptidmetabolite sind vollständig chloriert, aber nicht methyliert und nicht glykosyliert. Die Di- bis Heptapeptide sind C-terminal verkürzt und stellen somit hypothetische Biosynthesezwischenstufen dar.**

	Grundgerüst	Anzahl der Ringe	Molekülmasse
1	Dipeptid	0	344,1
2	Tripeptid	0	458,2
3	Tetrapeptid	0	607,2
4	Pentapeptid	0	756,3
5	Hexapeptid	0	905,3
6	Heptapeptid	0	1118,3
7	Heptapeptid	1	1116,3
8	Heptapeptid	2	1114,3
9	Oktapeptid	0	1283,4
10	Oktapeptid	1	1281,3
11	Oktapeptid	2	1279,3
12	Oktapeptid	3	1277,3

Die zweite Strategie zur Auffindung der Glykopeptidmetabolite der *A. balhimycina* C5A5T4E4-Mutante war weniger empfindlich, aber auch weniger selektiv, so dass die Möglichkeit bestand, auch vorher nicht erwartete Metabolite zu finden. Das Suchkriterium war in diesem Fall das Chlorisotopenmuster, was durch die Isotopenverteilung von Cl-35 (75 %) und Cl-37 (25 %) ein sehr charakteristisches Aussehen hat. Für diese Suche wurden Massenspektren über ein Zeitfenster von 2 min aufsummiert und nach dem charakteristischen Chlorisotopenmuster durchsucht (vor allem im Massenbereich > 1000 amu). Eine solche Summation wurde für jede HPLC-ESI-MS-Messung 10 x durchgeführt, so dass der Zeitraum der zu erwartenden Retentionszeiten von 0 bis 20 min abgedeckt war. Die so gefundenen Molekülmassen wurden anschließend erneut über den extrahierten Ionenstrom analysiert und auf ihre Wahrscheinlichkeit, einem Glykopeptid zu entsprechen, bewertet. Hierbei wurde vor allem auf die Retentionszeit und das Isotopenmuster geachtet.

Um von den so gefundenen Molekülmassen Spektren mit besserer Intensität und Auflösung zu erhalten, wurden weitere Messungen durchgeführt. Hierbei wurden folgende Parameter verändert:

<i>Scanbereich:</i>	eingeschränkt
<i>LIT scan rate:</i>	250 amu/s
<i>LIT fill time:</i>	<i>fixed</i> (100 bis 500 ms)
<i>Q0 Trapping:</i>	eingeschaltet

#### 4.2.19 Strukturaufklärung der identifizierten Glykopeptidmetabolite der *A. balhimycina* C5A5T4E4-Mutante mittels HPLC-ESI-MS/MS

Alle Molekülmassen, die in Abschnitt 4.2.18 über eine der beiden Suchstrategien identifiziert werden konnten, sollten mit Hilfe von HPLC-ESI-MS/MS-Messungen weiter untersucht werden. Zum einen sollte so geklärt werden, ob es sich wirklich um Glykopeptidmetabolite handelte und zum anderen sollte damit deren Struktur aufgeklärt werden. Hierzu eignete sich der *Scan*modus EPI, bei dem es sich wie in Abschnitt 3.1 beschrieben um einen Tochterionen-*Scan* handelt. Die Optimierung massenspektrometrischer Parameter erbrachte für den Tochterionen-*Scan* folgende zusätzliche bzw. alternative Werte:

Q1 <i>resolution</i> :	<i>unit</i> (in Einzelfällen <i>open</i> )
LIT <i>fill time</i> :	<i>fixed</i> 500 ms (in Einzelfällen 750 ms)
CAD gas:	<i>high</i>
CE:	30
CES:	20
Q0 <i>Trapping</i> :	eingeschaltet

Bei der Auswertung der Tochterionen-Spektren musste das für Peptide bzw. Proteine typische Fragmentierungsverhalten berücksichtigt werden, welches eine bevorzugte Fragmentierung entlang der Peptid- bzw. Proteinkette vorsieht (s. Abbildung 4-4). Bei jeder Spaltung entstehen aus dem Vorläufer-Ion ein geladenes und ein neutrales Fragment, wobei nur ersteres detektiert werden kann und nach der Nomenklatur von Roepstorff und Fohlman bezeichnet wird.<sup>[95]</sup> Verbleibt bei der Spaltung die Ladung an den Fragmenten der *N*-terminalen Serie, so werden die entsprechenden Fragmente mit a, b und c bezeichnet. Handelt es sich dagegen um geladene Fragmente der *C*-terminalen Serie, lautet die Benennung x, y und z. Die Zahl der im jeweiligen Fragment enthaltenen Aminosäurereste wird als Index angehängt. Bei höheren Kollisionsenergien können zusätzlich Fragmente auftreten, die durch weitere Fragmentierungen in den Seitenketten der Aminosäuren entstanden sind.

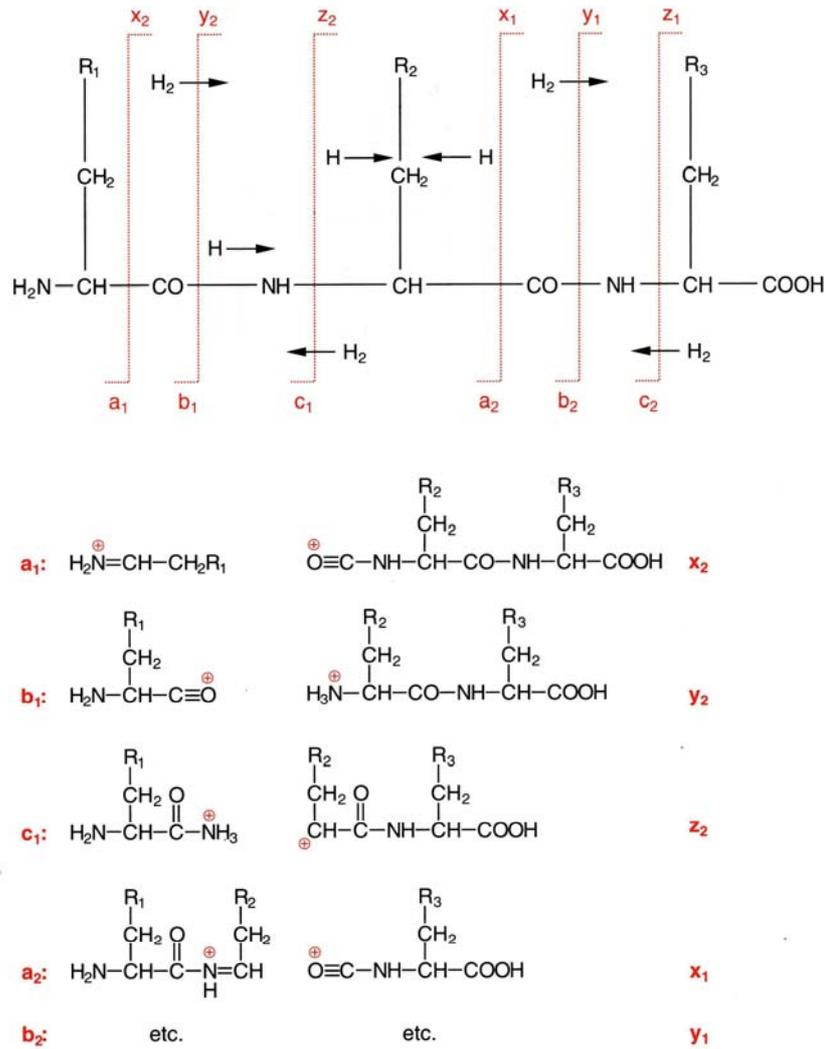


Abbildung 4-4: Typisches Fragmentierungsverhalten von Peptiden bzw. Proteinen entlang der Peptid- bzw. Proteinkette. Die dadurch entstehenden geladenen Fragmente werden nach der Nomenklatur von Roepstorff und Fohlman benannt.<sup>[95]</sup>

#### **4.2.20 Summenformelbestätigung der Glykopeptidmetabolite der *A. balhimy-cina* C5A5T4E4-Mutante mittels hochaufgelöster ESI-FTICR-MS**

Die hochaufgelösten ESI-FTICR-MS-Messungen wurde von Dipl.-Ing. Graeme Nicholson an der Eberhard-Karls-Universität Tübingen durchgeführt. Bei dem verwendeten Gerät handelte es sich um ein APEX II FTICR-Massenspektrometer mit 4,7 T der Firma Bruker Daltonics, Bremen. Um die bestmögliche Auflösung zu erhalten, wurde eine interne Kalibrierung vorgenommen.

#### **4.2.21 Chirale Aminosäurenanalytik**

Die chirale Aminosäurenanalytik wurde von Dipl.-Ing. Graeme Nicholson an der Eberhard-Karls-Universität Tübingen durchgeführt. Dabei wurde die Reinsubstanz zunächst in 6 M deuterierter Salzsäure (DCI) bei 110°C hydrolysiert (17 h). Das getrocknete Hydrolysat wurde anschließend zu den deuterierten *N*-(*O*-)TFA/Ethylestern derivatisiert und zusammen mit nicht-deuterierten internen Standards über chirale Gaschromatographie-Massenspektrometrie (GC-MS) analysiert. Die hierbei verwendete chirale Säule war eine Chirasil *L*-Val Kapillarsäule (20 m x 0,25 mm).

## 5 Ergebnisse

### 5.1 Ausarbeitung und Durchführung eines molekularbiologischen Konzepts zur Modulinsertion bei nicht-ribosomalen Peptidsynthetasen (NRPS)

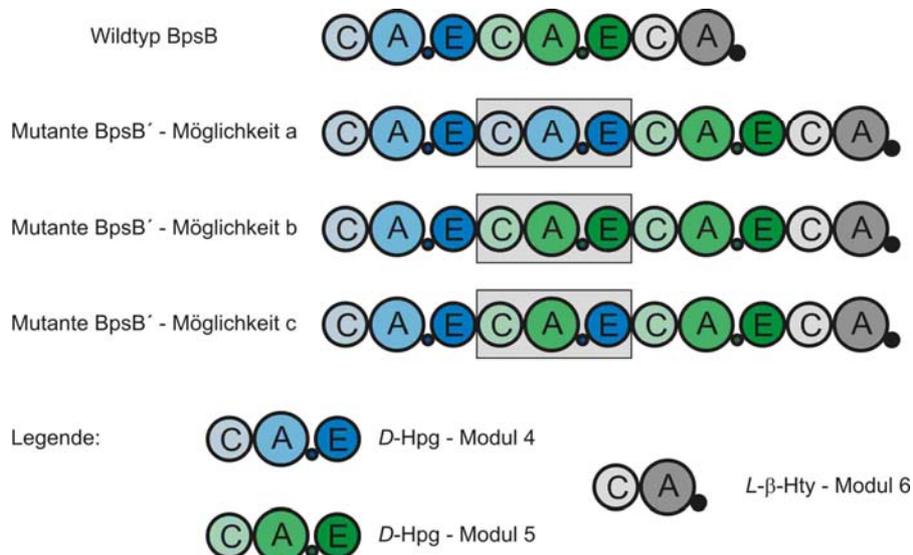
#### 5.1.1 Auswahl des neu einzufügenden Moduls einer NRPS

Bei der Auswahl des neu einzufügenden Moduls mussten mehrere Faktoren berücksichtigt werden:

- Verfügbarkeit der Substrat-Aminosäure
- Kompatibilität der Substratspezifitäten angrenzender Module bzw. Domänen
- natürliche Übergänge zwischen Nachbarmodulen
- sinnvolle Veränderung des Glykopeptidproduktes

Die Umsetzung all dieser Faktoren erforderte ein gründliches Abwägen verschiedener Möglichkeiten, wobei am Ende die Entscheidung auf ein für *D*-Hpg kodierendes Modul fiel, welches zwischen die zwei bereits vorhandenen *D*-Hpg - Module in BpsB eingefügt werden sollte (s. Abschnitt 1.3.2). Die Biosynthesegene der nicht-proteinogenen Aminosäure Hpg sind integraler Bestandteil des Balhimycin-Biosynthesegenclusters (s. Abschnitt 1.3.2), so dass eine ausreichende Verfügbarkeit von Hpg erwartet werden konnte. Auch bezüglich der Kompatibilität der Substratspezifitäten angrenzender Module bzw. Domänen schien diese Strategie wenig störungsanfällig. Während T-, E- und TE-Domänen eine geringe bis keine Substratspezifität aufweisen, sind A- und C-Domänen mäßig bis sehr substratspezifisch (s. Abschnitt 1.2). So muss die Aminosäure, die von einer A-Domäne ausgewählt wird, mit der Substratspezifität der *acceptor site* der stromaufwärts gelegenen C-Domäne zusammenpassen. Im Gegensatz dazu ist die Spezifität der *donor site* einer C-Domäne für die von stromaufwärts kommende Aminosäure bzw. Peptidkette vergleichsweise gering und beschränkt sich vornehmlich auf die richtige Stereochemie. Die ausgesuchte Strategie berücksichtigt diese Erfordernisse. Des Weiteren wurde die Einführung eines *D*-Hpg zwischen zwei bereits vorhandene *D*-Hpg's als eine sinnvolle Veränderung des Glykopeptidprodukts betrachtet, da die Aminosäure Hpg einen zentralen Baustein darstellt und in sog. „tailoring“-Reaktionen wie oxidative Seitenkettencyklisierung und Glykosylierung involviert ist. Die Erweiterung des Glykopeptids um ein zusätzliches Hpg bietet somit die Möglichkeit, diese Enzymreaktionen auf Funktionalität und Spezifität hin zu untersuchen.

Nachdem der Ort der Modulinsertion und die Art des einzuführenden Moduls feststanden, musste noch über den Ursprung bzw. die Zusammensetzung dieses Moduls entschieden werden. Hierbei bestand die Herausforderung vor allem darin, möglichst natürliche Übergänge zwischen Nachbarmodulen zu schaffen. Zunächst galt es deshalb zwischen den beiden Möglichkeiten zu entscheiden, die DNA-Sequenz eines der beiden *D*-Hpg-Module aus dem zu modifizierenden Gencluster selbst oder aber die DNA-Sequenz eines *D*-Hpg-Moduls aus einem anderen Gencluster bzw. heterologen Wirt zu verwenden. Da Eingriffe in komplexe Biosynthesysteme bekanntermaßen dazu neigen, zum Teil drastische Produktionseinbrüche zur Folge zu haben (s. Abschnitt 1.4), fiel die Entscheidung auf ein *D*-Hpg-Modul des Balhimycin-Biosynthesegenclusters, da somit nur bereits vorhandene DNA- bzw. Protein-Sequenzen verdoppelt werden. Das Balhimycin-Biosynthesegencluster besitzt zwei aufeinanderfolgende Module, die für *D*-Hpg kodieren: *D*-Hpg-Modul 4 und *D*-Hpg-Modul 5 von BpsB. Für die Modulinsertion ergaben sich somit drei Möglichkeiten (s. Abbildung 5-1): Modul 4 wird dupliziert (Möglichkeit a), Modul 5 wird dupliziert (Möglichkeit b) oder es wird eine Kombination aus der C-A – Didomäne von Modul 5 und der T-E – Didomäne von Modul 4 eingeführt (Möglichkeit c).



**Abbildung 5-1: Darstellung der drei Möglichkeiten der *D*-Hpg Modulinsertion zwischen die beiden bereits existierenden *D*-Hpg-Module 4 und 5 im Wildtyp-Protein BpsB, welches dadurch zum Mutanten-Protein BpsB' wird: Modul 4 wird dupliziert (Möglichkeit a), Modul 5 wird dupliziert (Möglichkeit b) oder es wird eine Kombination aus der C-A – Didomäne von Modul 5 und der T-E – Didomäne von Modul 4 eingeführt (Möglichkeit c). Das zusätzlich eingebrachte Modul ist jeweils mit einem grauen Kasten unterlegt.**

Dieses im Folgenden als C5A5T4E4-Modul abgekürzte Hybridmodul aus Möglichkeit c hatte den entscheidenden Vorteil, dass alle Modulübergänge in ihrer natürlichen Form erhalten bleiben und nur *ein* nicht-natürlicher Übergang erzeugt wird und zwar innerhalb des C5A5T4E4-Hybridmoduls zwischen den Domänen A5 und T4. Wie in Abschnitt 1.4 bereits beschrieben sind diese sog. Linkerregionen zwischen Domänen bzw. Modulen prädestiniert für die bei kombinatorischen Biosyntheseexperimenten erforderlichen Eingriffe, da sie nicht zur enzymatischen Aktivität beitragen.<sup>[12]</sup> Um sicher zu gehen, sollten aber auch hier die vorgenommenen Änderungen nicht allzu groß sein, weshalb bei der in Abschnitt 5.1.2 beschriebenen Klonierstrategie auf eine möglichst „natürliche“ Linkerregion zwischen A5 und T4 geachtet wurde. In Abbildung 5-2 ist ein Proteinalignment der Wildtyp-Linkerregionen A4 – T4 und A5 – T5 zwischen den Domänen A4 und T4 bzw. A5 und T5 mit der künstlichen Linkerregion A5 – T4 zwischen den Domänen A5 und T4 gezeigt.

```

----T PRGEHLIGYVVAEAGHDADPVRLREQLAGTLPEFMVPAAVLVLDELPLTVNGKVDRRALPEPDFAAKSA GREPVTEAER A4 - T4
AREGQPG EKRLIAYVTAEAGSALESAAVRAHLATRLPEFMVPSVVVVLESFPLTLNGKIDRAALPAPEFAGKAA SREPVTEAER A5 - T4
AREGQPG EKRLIAYVTAEAGSALESAAVRAHLATRLPEFMVPSVVVVLESFPLTLNGKIDRAALPAPEFAGKAA GREPRTEAER A5 - T5

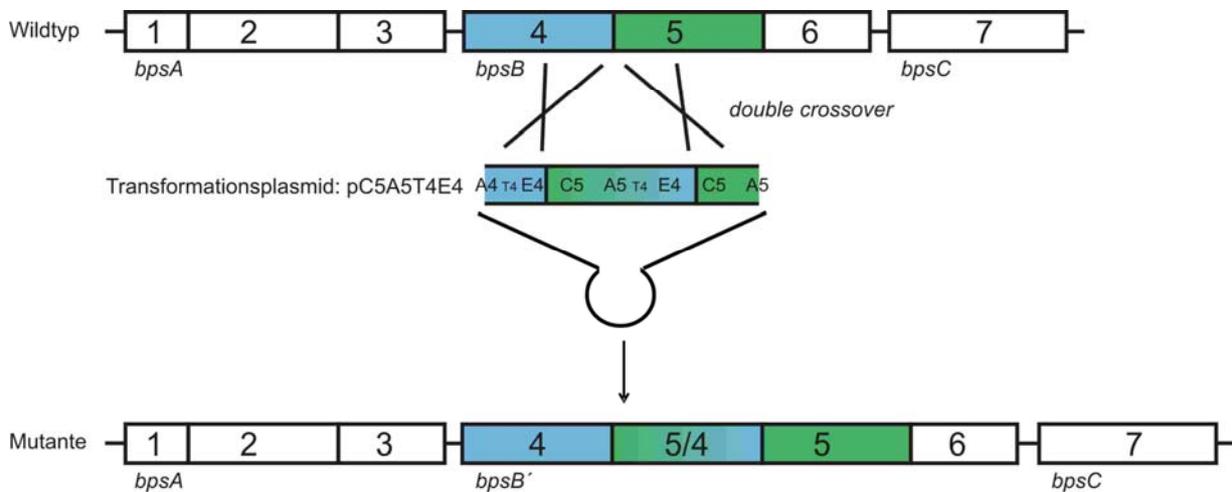
```

**Abbildung 5-2: Proteinalignment der Wildtyp-Linkerregionen A4 – T4 und A5 – T5 zwischen den Domänen A4 und T4 bzw. A5 und T5 mit der künstlichen Linkerregion A5 – T4 zwischen den Domänen A5 und T4 im zusätzlich eingebrachten C5A5T4E4-Hybridmodul.**

Wie man sieht besteht die künstliche Linkerregion A5 – T4 aus der Anfangssequenz der A5 – T5-Linkerregion und der Endsequenz der A4 – T4-Linkerregion, wobei dazwischen lediglich ein Gly -> Ser Aminosäureaustausch auftritt. Diese hohe Homologie zu den Wildtyp-Linkerregionen erhöhte die Wahrscheinlichkeit, dass der künstliche Eingriff in das Balhimycin-Biosynthesesystem in Form einer Modulinsertion akzeptiert und umgesetzt wird.

### 5.1.2 Klonierstrategie zur Durchführung der ausgewählten Modulinsertion

Zur Umsetzung der in Abschnitt 5.1.1 beschriebenen Modulinsertion musste eine Klonierstrategie entwickelt werden. Diese Strategie basierte auf der Konstruktion des Transformationsplasmids pC5A5T4E4, welches das einzubringende C5A5T4E4-Hybridmodul zwischen den flankierenden Regionen der gewünschten Insertionsstelle im Genom von *A. balhimycina* Wildtyp enthielt. Ein *double crossover* zwischen pC5A5T4E4 und der genomischen DNA von *A. balhimycina* Wildtyp führte dementsprechend zur gewünschten *A. balhimycina* C5A5T4E4-Mutante (s. Abbildung 5-3).



**Abbildung 5-3:** Das nicht-ribosomale Peptidsynthetasegen *bpsB* der Balhimycin-Biosynthese wird durch *double crossover* mit dem Transformationsplasmid pC5A5T4E4 zu dem um ein Modul erweiterten *bpsB'*-Gen modifiziert.

#### 5.1.2.1 Definition der Domänengrenzen

Essentielle Voraussetzung für die Entwicklung einer Klonierstrategie war die Festlegung der Domänengrenzen in Modul 4 und 5 von BpsB. Zwischen den Domänen befinden sich typischerweise kurze, nicht funktionale Linkerbereiche<sup>[12]</sup> aus vor allem kleinen und hydrophilen Aminosäuren, deren Sequenz wenig konserviert und daher relativ flexibel ist.<sup>[9]</sup> Die Domänen wurden primär über Hits gegen die Sequenzmotiv-Datenbank Pfam<sup>[96]</sup> bestimmt und die Grenzen dann über ein Proteinalignment bzw. Vergleiche mit publizierten Alignments verifiziert.

Das hierfür verwendete Proteinalignment von BpsB mit den homologen nicht-ribosomalen Peptidsynthetasen anderer Glykopeptide (TeiC von Teicoplanin, StaC von A47934, Dbv17 von Dalbavancin, CepB von Chloroeremomycin und ComC von Complestatin)<sup>[97]</sup> ist in Abbildung 5-4 gezeigt.

```

      *      20      *      40      *      60      *
teiC : VTQARIEDTWPLSPLOECLLFHTVYDDEGPDVYVGHWILDLDCPVDAARLRAWEALLARHIALRACFRQ : 70
staC : MAQSRIBEDTWPLSPLOECLLFHTVYDDEGPHVYVGHWILELDCPVDPARLRTAWETLLTRHIALRACFRQ : 70
dbv17 : MAQSRIBEDTWPLSPLOECLLFHTVYDDEGPDVYVGHWILNLCNPVEADPLRAWEALLARHIALRACFRQ : 70
cepB : MTQSRIBEDTWPLSPLOECLLFHAWYDCEGPDVYVGHWILDLDCPVDAARLRAWEALLTRHIALRACFRQ : 70
bpsB : MSQSRIBEDTWPLSPLOECLLFHAWYDCEGPDVYVGHWILDLDCPVDAAGLRAWEALLARHIALRACFRQ : 70
comC : MTRSQVEDWPLSPLOECLLFHAAFDSDGPDVYVQQRPLDLDCPVDDLLRSTSEALLTRHPVLRACFRQ : 70
      6 qsr6E WPLSPLOECLLFH 5D gP 6Y GhwIL L GPVda LR aWE LL RHa LRACFRQ
    
```

```

      80      *      100      *      120      *      140
teiC : RRTGETVQVACRVELPWRVVDLHLLDPEHAVR--ELADEDRLRPFVAKPPLLRLTLIRLADDERLRLV : 138
staC : RRTGETVQVACRVELPWRVVDLHLLDPEHAVR--ELAEEDRLRPFVAKPPLLRLTLIRLADDERLRLV : 138
dbv17 : RKSCETVQLIARQVELPWRVVDLHLLDPEHAVR--AVAEEDRTRRFDLAKPPLLRLTLIRLAGDDERLV : 138
cepB : RKSCETVQLIARQVELPWRVVDLHLLDPEHAVR--ELAEQDRTTRFDLAKPPLLRLTLIRLADDERLRLV : 138
bpsB : RKSCETVQLIARQVELPWRVVDLHLLDPEHAVR--ELAEQDRTTRFDLAKPPLLRLTLIRLAGDDERLV : 138
comC : RKSCETVQLIARQVELPWRVVDLHLLDPEHAVR--ELAEQDRTTRFDLAKPPLLRLTLIRLAGDDERLV : 140
      R 3GtVQ66A VeLPWR vDLshLd p AvR e6A 2DR rF16a PLLRLtL6RL hRL6
    
```

C4

```

      *      160      *      180      *      200      *
teiC : VTCHHAVMDGWSLPVIVDELTAIYAAAGCPAQLPAVTPYREYLAWLSRQDRKPSALAAGAAELRGAQEPFL : 208
staC : VTCHHAVMDGWSLPVIVDELTAIYAAAGCPDLEVAASYRDYLAWLNRQDKRALAAGAAELRGAQEPFL : 208
dbv17 : MTCHHAVMDGWSMPVIMDELTAIYAAAGCPDLEVAASYRDYLAWLNRQDKRALAAGAAELRGAQEPFL : 208
cepB : VTCHHTIMDGSWLPVIVDELTAIYAAAGCPDLEVAASYREYLAWLSRQDRKRALAAGAAELRGAQEPFL : 208
bpsB : VTCHHTIMDGSWLPVIVDELTAIYAAAGCPDLEVAASYREYLAWLSRQDRKRALAAGAAELRGAQEPFL : 208
comC : VTCHHTIMDGSWMPVLLHETAIYAAAGCPDLEVAASYRDYLAWLNRQDKRALAAGAAELRGAQEPFL : 210
      6TcHH 6MDGWS6P66 dEL 6Y Ag LP v sYR YLaWL Rqlk ral AW aEL G 2ePFL
    
```

```

      220      *      240      *      260      *      280
teiC : VAPADPCRAPGMPESITVEVPEHLTRSEDLAPSNCLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 278
staC : VAPADPCRAPGMPESITVEVPEHLTRSEDLAPSNCLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 278
dbv17 : VAPADANRAPAMPENITVELPEDLTRSELARHGLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 278
cepB : VAPADPCRAPGMPESITVEVPEHLTRSEDLAPSNCLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 278
bpsB : VAPADPCRAPGMPESITVEVPEHLTRSEDLAPSNCLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 278
comC : VAPADPCRAPGMPESITVEVPEHLTRSEDLAPSNCLTLNTVVQCAWALVLAQLSCRIDVVVFCAAVSARP : 280
      VapADpgrRaP PEs6 6pe LTr L LAR GLT6NT66QCAWA66LA L GRTDVVVFC VsaRP
    
```

```

      *      300      *      320      *      340      *
teiC : AHLPGVECHVGLFLNTVPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 348
staC : ADLPGVECHVGLFLNTVPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 348
dbv17 : PDLPGVECHVGLFLNTVPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 348
cepB : PDLPGVECHVGLFLNTVPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 348
bpsB : PDLPGVECHVGLFLNTVPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 348
comC : AGLPGVESMIGHFINTLPRVRLRCSAPVVEHLAEQLRQSAIIPDQVGLADIQQAAGPAAVFDTLVVF : 350
      LpGVE M6G6F6NT6PVRvL C Pv6e L LQ RQSALIp Q 6CLa IQ aG A6FDTL6VF
    
```

```

      360      *      380      *      400      *      420
teiC : ENFPRDLRPSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 418
staC : ENFPRDLRPSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 418
dbv17 : QNFPRDLRPSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 418
cepB : EKFFHGPACSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 418
bpsB : EKFFHGPACSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 418
comC : ENFPRDLRPSDASQTFGICQVWAAHYPLTLVAVPCESMYKLDYLDLFDQDTEAIVERFTGVLRLQ : 420
      2 5 F 6 qGr aAHYPLTLVAVP M kLD t LFD A 6 eRf 6Lr
    
```

```

      *      440      *      460      *      480      *
teiC : LTDAGELTVAAIDVTSAAERDRA-RWGAAVGARPDRALDLFARQVQRPEVAVDQCDRUMSFGELAE : 487
staC : LAADHETVADLAVTTAAERNLVS-AMGTAPCAEPCDLPVLDLDRQAERRCDETAIDCDRITTSFRDLAC : 487
dbv17 : LACAGDLTVAEVDVTSAAERALVWNAASAAPRVAPGELAPDLDRQV ERGDRVAVVECKRAVSFGLAE : 488
cepB : LTDAGDLTVADVDTSAERALVWDEGAAASQATPSMLALDLRCSQKDRRDEVAVDQCDRUMSFGELAE : 488
bpsB : LTDAGELTVAGVETTTAAERALVWDEGAAASQATPSMLALDLRCSQKDRRDEVAVDQCDRUMSFGELAE : 488
comC : LAADPLQVGRVDAIGEAERTVTDHONSTATDRPAESVLELFAHQVVRTPHATAISDCERQITTYAQLAV : 490
      L tVa 6 vt aAER V W P 1 LF Qv d A6 dG R 35 LA
    
```

A4

```

      500      *      520      *      540      *      560
teiC : RAERLACHTMRCVRCGRVAVVHERSCHELIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 557
staC : YADRLACHTMRCVRCGRVAVVHERSTDLIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 557
dbv17 : HAERLACHTMRCVRCGRVAVVHERSCHELIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 558
cepB : HAERLACHTMRCVRCGRVAVVDRSLDLIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 558
bpsB : RAERLACHTMRCVRCGRVAVVDRSLDLIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 558
comC : ESCQLARHLDRCVRSAGDRVAVVHERSADLLIATLLAVWAGAAVFPVDPAYPAERVKMLLTDAPVAVVC : 560
      a rLAg L RGV rGDRVAVV6 RS L6aTLAVW AGAAVFPVDPaYP ER f6L D PaavVc
    
```

```

*      580      *      600      *      620      *
teiC : TAAFRAAVLDGGLLEAIVVDDPCTWPAVPCP--PVPTCPDLDAYVMTSGSTCTPKCVAVSHCNVAALWG : 625
staC : RAANFAAVLDGGLDPIVLDLPEIRRAVRECARPSVQVSAADDLAYVMTSGSTCTPKCVAVSHCNVAALWG : 627
dbv17 : ERACQAAVPAAGGLDPIVLDLPTLRRAVREHARLSAGAHADDLAYVMTSGSTCTPKCVAVSHCNVAALAG : 628
cepB : AEAFTDAVLDVGLDPIVLDLPTLRRAVREHARLSAGAHADDLAYVMTSGSTCTPKCVAVSHCNVAALWG : 628
bpsB : AEAFTDAVLDVGLDPIVLDLPTLRRAVREHARLSAGAHADDLAYVMTSGSTCTPKCVAVSHCNVAALAG : 628
comC : SERTFVVVPPEDVPALV---PLDEASILDCEPLSLPVAGGDIAYVMTSGSTCTPKCVAVSHCNVAALWG : 627
a r aV      p6v ddP      a6a      s      D6AYVMTSGSTC PKCVAVSHC VAAL G
    
```

```

*      640      *      660      *      680      *      700
teiC : DPEGRTCPDPTVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 694
staC : EPRGRTCPDPTVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 696
dbv17 : EPGGGLCPDPAVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 697
cepB : EPGGGLCPDPAVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 697
bpsB : EPGGGLCPDPAVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 697
comC : EPGGGLCPDPAVLMHASHAFDISLFEIUVPPLSGARVVIAGPCAV-DEALAAQVAAGVTAHLTACAFR : 697
ePgW GP D VL HAsHAFDISLFEIUVPPLSGAR6 6A pgaV Dg A6a y6A GVT aH6TAC FR
    
```



```

*      720      *      760      *
teiC : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 764
staC : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 766
dbv17 : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 767
cepB : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 767
bpsB : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 767
comC : VLAESPEAVAGLREVLTGGDAVPLAAVERVRACPDVVRVHLYCPTETTLTCATWLLLEPGDETGSVLPV : 767
VLA ESPE v GIREVLTGGD VP aAV RVR cp 6RVRHLYCPT E 3LCATWLL2PG e g VLP6
    
```

A4

```

*      780      *      800      *      820      *      840
teiC : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 834
staC : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 836
dbv17 : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 837
cepB : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 837
bpsB : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 837
comC : GRPLAGRRVWVLDALFLRPVPGTCELYVAGACVAGCYLGRPALTAERFVAAPFAPCRMYRTGDLAYWT : 837
GRPLagRrV VLDASLrPpPGv GELY6AGAC6A CYL R aL3aERFVA Pf g RMYRTGDLa WT
    
```

```

*      860      *      880      *      900      *
teiC : EQCALAFAGCRADDQVKIRGYRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 899
staC : DEGTLVVFCGRADDQVKIRGYRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 901
dbv17 : DQCALAFAGCRADDQVKIRGYRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 902
cepB : HQCALAFAGCRADDQVKIRGYRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 902
bpsB : EQCALAFAGCRADDQVKIRGYRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 902
comC : SDGELVVFGRADDQVKVRCGRVEPCEIEVLAAGLPCVCGAVVSRDEH----LIGYVVAEACQALDPAF : 906
C L F GRADDQVK6RGSRVEpGE6E vLaglpGVgqAVV R e LICYV6ae G d Dp
    
```



```

*      920      *      940      *      980
teiC : LREQLAETLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 969
staC : LREQLAETLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 971
dbv17 : IREQLAETLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 972
cepB : LREQLAETLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 972
bpsB : LREQLAETLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 972
comC : LREQLAKLPEFMVPAAVLVLDLPLTNGKVDRLPALPEPDFAAKSTGRAPTAERLCCGVFAEVLGCLD : 976
6Re LA t6PESMVPA66VLD LPLT NGKVDr ALP PDFA gR P T aER LC 6FA VLG1
    
```

T4

```

*      1000      *      1020      *      1040      *
teiC : RVGTEDDFFELGCDSSISMQVAARARREGIALTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1039
staC : RVGTEACFFELGCDSSISMQVAARARREGIALTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1038
dbv17 : RVGTEADFFELGCDSSISMQVAARARREGIPLTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1040
cepB : RVGTEDDFFELGCDSSISMQVAARARREGISLTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1040
bpsB : RVGTEDDFFELGCDSSISMQVAARARREGISLTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1040
comC : RVGTEDDFFELGCDSSISMQVAARARREGIPLTPROVFEHRTPERLAQAQAQCSARPDRAADAGLGEI : 1044
rVG d FF LCGDSSISMQVAARAr g6 LTPRqVFe RTPERLA A a g6GE6
    
```

```

*      1060      *      1080      *      1100      *      1120
teiC : PWTVMRALCDDAQRPCFAQARVVVAPPGLDPAALTCALRAVLDTHDQLRGRVSDG-RLIVPERGAVAA : 1108
staC : PWTVMRALCDDAQRPCFAQARVVVAPPGLDRAVLAALQAVLDTHDQLRGRVSDG-RLIVPERGAVAA : 1107
dbv17 : PWTVMRALCDDAQRPCFAQARVVVAPPGLDPAALVTALQAVLDVHDLLRTRVSDG-RLIVPERGAVAA : 1109
cepB : PWTVMRALCDDAQRPCFAQARVVVAPPGLDPAALVSAQAVLDCHDQLRGRVSDG-RLIVPERGAVAA : 1109
bpsB : PWTVMRALCDDAQRPCFAQARVVVAPPGLDPAALVSAQAVLDVHDLLRTRVSDG-RLIVPERGAVAA : 1109
comC : PWTVMRALCDDAQRPCFAQARVVVAPPGLDPAALVSAQAVLDVHDLLRTRVSDG-RLIVPERGAVAA : 1114
PwTVMRA6G dA RPCFaQ rVVv P g aL AL a6LD HD LR Rve D rL6V erG V A
    
```

E4

```

*      1140      *      1160      *      1180      *
teiC : ADLLTRVWVSD-CHDARAEREATAAAGTLDPSAGIMRPAVWLDAGDAEPGRLLAVVHHLVWDVAVSWIL : 1177
staC : AGLVTRVAATA-NLDEIARREACTAACVLDPTCTGTMRVWVWIDAGDAEPGRLLAVVHHLVWDVAVSWIL : 1176
dbv17 : AGLVTRVAAGNG-NLARREREAPTAAGTLDPSAGIMRPAVWVDAGDAEPGRLLAVVHHLVWDVAVSWIL : 1178
cepB : TGLVTRVAVGTG-DLDEIARREVTATCTLNPSAGIMARLVWIDAGDAEPGRLLAVVHHLVWDVAVSWIL : 1178
bpsB : AGLLTRVAAGTG-CLDEIARREVTATCTLDPSAGIMARVWVWIDAGDAEPGRLLAVVHHLVWDVAVSWIL : 1178
comC : AALVARVELRTDEDPEVVERAAAREADCLDPACVMVRLVWVDAGPERACRLVWVHHLVWDVAVSWIL : 1184
a L6tRva      de aeRe      tA g LiPs G M R VW6DAGdaepGRLa V HHL VDaVSW IL
    
```

```

*      1200      *      1220      *      1240      *      1260
teiC : LPDLQAAYQEVLAGATPALEPAATSYRQWARRLTEQASPSLAEIDHWVTVLDAEPPLA-----E : 1239
staC : LPDLLETANBOVISCGTASLEPATTSFROWARRLTEQASERTVAEIDHWVGVLDGADFPPE-----E : 1238
dbv17 : LSDLRAAYDEAVSCGTPVLEPAVTSYRQWARRLAGAALSESTVAEACHQACVLEGGDLPLE-----R : 1240
cepB : LPDLQTANDEVISCGTPALEPPVTSYRQWARRLTEQALSESTVAEIDQWVAVLDGANPPE---QDTC-Q : 1244
bpsB : LPDLRAAYDEVISCGTPALEPPVTSYRQWARRLTARALSESTVAELEKQAAVVEGAEPAFP---QDTC-Q : 1244
comC : LPDLRVACEAAVACRAADDPVCTPFPRWAQLEAEAREDEKRIAEIDPWRELLGCGDTLIGARALDPAD : 1254
LpDL Ay 6 G t LeP Ts5RqWArrL A se t6AE W 66
    
```

```

*      1280      *      1300      *      1320      *
teiC : HHCQAHSWATLSCAVACHLVSRMPCAFHCCIGEVLLAGLAAAVARRR-CDDAGVLVDVEGCHGRHA-ADG : 1307
staC : HACRPHSWATLSCSAARSVLVCGQAPCAFHCCGVQVLLAGLACAVARRR-GCDAGVLVDLEGCHGRHA-ADG : 1306
dbv17 : HPCQASASVSRLLSDAQARNLVARVPAAFHCCGVQVLLAGLACAVARRR-CADAGILVDVEGCHGRHA-ADG : 1308
cepB : HTCQSHSWSTRLSCTEVRLVAHLPCAFHCCIGDVLVLLAGLACAVARRR-GTCALLVDVEGCHGREA-ADG : 1312
bpsB : HTCQSHSWSTRLSCTEVRLVLTVLPCAFHCCIGDVLVLLAGLACAVARRR-GSCALLVDVEGCHGREA-ADG : 1312
comC : TAATLRHRSWATLSCAPAHDLICRPSAAPHCCVHEALLATLAPAVARSRSCADSAVLVDVEGCHGREPLASD : 1324
h g swS L8g L6 p AFHCG6q vLLAgLA AVAR R G a 6LVD6EGHGR a Adg
    
```

```

*      1340      *      1360      *      1380      *      1400
teiC : EDLLRTVCGWFTSVHPVRLDVSQVDLAAAGADAAAGLLKAVKQVRAAPCGDCFCGCLLRHLNPDTAERL : 1377
staC : EDLLRTVCGWFTSVHPVRLDVSGLDLAAAGADAAAGLLKAVKQVRAAPCGDCFCYCLLRHLNPDTCARL : 1376
dbv17 : EDLLRTVCGWFTSVHPVRLDVSQV-----CPCAAAGAGLLKAVKQVRAAPCGDCFCYCLLRHLNPDTCARL : 1373
cepB : EDLLRTVCGWFTSVHPVRLDVSQVDLAAAGADAAAGLLKAVKQVRAAPCGDCFCYCLLRHLNPDTCARL : 1382
bpsB : EDLLRTVCGWFTSVHPVRLDVSQVDLAAAGADAAAGLLKAVKQVRAAPCGDCFCYCLLRHLNPDTCARL : 1382
comC : VDLRSRTVCGWADAHVRLDADIDVADASAGCAAGALLKAVKQVRAAPCGDCFCYCLLRHLNPDTSAL : 1394
eDLLRTVCGWFTsvHPVRL s 6d a a G a AG LLKaVKeQ RAvPGDC G5GLLRhLNP T rL
    
```

```

*      1440      *      1460      *
teiC : AALPAPQICFNFLCRSGVAAEAVPQVRCGSLCAGEAPPDLVLAHPLBAGADVWDTDPGCLLRLTLTD--G : 1445
staC : AELPSPRIGFNFLCRSAAASPDTPQVQVVEGPLCAGADAPDLVLAHPLBAGADVWDTDPGCPQVRLTVD--G : 1444
dbv17 : AELPSAQICFNFLCRSGVASEDTAQVQVCEGALCQAAQPDVQSHALBAGADVWDTDPACPRRLRLAID--G : 1441
cepB : AELPSAQICFNFLCRVLAQONTAQO-QPVAECPGCGPDVLAHPLBAGADVWDTDPACPRRLCLAID--T : 1449
bpsB : AELPSAQICFNFLCRVLAQONTAQO--CPLCGCPDMLVAHPLBAGADVWDTDPACPRRLCLAID--T : 1448
comC : AQLPAPQVCFNFLCRFPAPTAA--ACPSAPC-GSIAPDTPVLAHPLBAGADVWDTDPGCPRLTLTLNWAG : 1461
A LP q6CFNYLCR Bmtf PD v1 H 6E A 6 DTP GP 6 L 61
    
```

```

*      1480      *      1500      *      1520      *      1540
teiC : RDLAPVTVELLGEAWLEMLTGLATHAGDPACGHTPADFDLWVLTQRDVTLEAAAPCLTDIWPVLSPLQR : 1515
staC : RDLDPDTVERLGEAWLEMLTGLAAHADPCACGHTPSDFLDVMSQRDVALEAAAPCLTDIWPVLSPLQE : 1514
dbv17 : RDLDPAAVERLGEAWLDTLGLAALADTPCAGGHTPSDFELVWVQRDVDLEAVAPCLTDVWPVLSPLQE : 1511
cepB : RDFDLATVERLGEAWLEMLTGLAAHACNFCACGHTPSDFLDVLTQRDVTLEAAAPCLTDIWPVLSPLQE : 1519
bpsB : RDFDLATVERLGEAWLEMLTGLAAVRCGACGHTPADFDLWVLTQRDVALEAAAPCLTDIWPVLSPLQE : 1518
comC : NILDAAEDARLCGTWLAALLTGLATHLTPSACGHTPSDFPLHALTRSDVVELETTVPCLTDVWPVLSPLQE : 1531
rd d vErLGeaWL LtGLA a p AGGHTP DF Lv 6 qrDV eLEa aPCLTD6WVLSPLQE
    
```

```

*      1560      *      1580      *      1600      *
teiC : GMLFERALDEDCQDVYQSRILDLDCPLDADRLEAAWRTLVARHESLRTSFHQLESGETVQVWVQADIG : 1585
staC : GMLFERALDEDCQDVYQSRILDLDCPLDPERLRAAWQPLVARHESLRTSFHQLESGETVQVWVQADIP : 1584
dbv17 : GILFERALDEDCQDVYQSRILDLDCPLDAQRHAAWQSVDRHETLRTCFHQLESGETVQVWVQAEVIL : 1581
cepB : GMLFERALDEDCQDVYQSRILDLDCPLDAQRLTAAWQVVRHESLRTSFHQLESGETVQVWVQAREADIP : 1589
bpsB : GMLFERALDEDCQDVYQSRILDLDCPLDEPRLEAAWQVLAARHESLRTCFHQLESGETVQVWVQAREADIP : 1588
comC : GLLRHTATDDCQPDVYETQRVLDLTCPLDADRHRSSWNTLVARHOMLRVWFHELADGRTVQAVTGAEPF : 1601
G6LFerA De GvDvy23QR6LdLdGPLD RL aW 66aRH Lrt FH2L sG TVQvV A
    
```

```

*      1620      *      1640      *      1660      *      1680
teiC : WRVADVSHLRAEADAAAEVCRLLAEDQAQRFDVTPAPLLRLLVRLGADRHRLVVTSHHVLVDCWSTPIIL : 1655
staC : WRVADVSHLGEADAAAEVERLLAEDQAERFDVTPAPLLRLLVRLGADRHRLVVTSHHVLVDCWSTPIIL : 1654
dbv17 : WRVADLSHLDEABAAAEVERLLAEDQAERFDVTPAPLLRLLVRLGADRHRLVVTSHHVLVDCWSTPIIL : 1649
cepB : WRVADLSHLDEABAAAEVERLLAEDQAERFDVTPAPLLRLLVRLGADRHKLVTSHHVLVDCWSTPLIV : 1659
bpsB : WRVADLSHLDEABAAAEVERLLAEDQAERFDVTPAPLLRLLVRLGADRHRLVVTSHHVLVDCWSTPLIV : 1658
comC : WRVADLSDRPEPEALAEALAEAREPREFDLTPAPLLRLLVRLGADRHRLVVTSHHVLVDCWSTPIAL : 1671
WR aD6S e A aZverLia dqa RFD6 r PLLRLlL6RLG H4LV6TSHH666DGWSTP6 6
    
```

E4

C5

```

*      1700      *      1720      *      1740      *
teiC : GEMSAAYAGAPSTA---TAPSPRDYLAWLNRQDAATRAAWPSELACDEPTLV-DADAKRTMVMVPGHEHC : 1721
staC : GEMSSGYAGGPGPS---TSPSYRDYLLWLSRQEQDATLSAWRSEFAGSDEPTVV-DADAGRAMVMPQHEHC : 1720
dbv17 : GEMLTAYADGRVSP---APPSYRDYVAWLSRQEDAAASAWRSELACDEPTVV-GLDAGKAPVMPDGHGA : 1715
cepB : GEMSAGYAGGPSAS---KSPSYRDYLAWLSRQEEETRSAWRSELACDEPTLV-DADAKRTLVMVPMNRS : 1725
bpsB : GEMSDGYAGGRSSS---KPPSYRDYLAWLSRQAEATRSAWRAELACADEPTLV-DADAKRTLVMVPEDEHA : 1724
comC : DEASRIYAADCDASGLCPAPSF RDHLAWLARQKD TAHRAWRAELACADEPTLLIAPTSAGCGPAATEGDS : 1741
gRms YA PS Dy6aWL RQd AWR ELAC DEPT66 dAg4 vmp
    
```

```

*      1760      *      1780      *      1800      *      1820
teiC : EGLSEQATRNLTETPARGHVTLSITLQCAWALVLRARLACRTDVVFGTVVSGRPPE-VPDVERMVGMIINT : 1790
staC : AWLSEESTRTLTDPARGHGLTSLTIVQCAWALVLRARLACRTDVVFGTVVSGRPAD-VPDVERMVGMIINT : 1789
dbv17 : EWLSEATRTALTCFARGHGLTSLTVVQCAWALVLRARLARTDVVFGTVVSGRPADALPDVERMVGMIINT : 1785
cepB : ELLSEATRTLTADPARGHGLTSLTVVQCAWALVLRARLACRTDVVFGSVVSGRPAD-VPDVERMVGMIINT : 1794
bpsB : EWLSEPATRALACFARGHGLTVSTIVLCAWALVLRARLACRTDVVFGSVVSGRPAD-VPDVERMVGMIINT : 1793
comC : LLLSEELTRALTSVAREHGLTLLTFLAQCAWALVLSRLAPRTDVVFGISVAGRPPE-LPGVESMVGMIINT : 1810
Lse TR L fARgHg6T6sT6 GAWALVLaRLA RTDVvFG vVsGRPa 6PdVerMVGMIINT
    
```

C5

```

*      1840      *      1860      *      1880      *
teiC : VPRVRRLDCGRVLDLLHQLQARQSALT EHQHGLCLPDIQRAACFCALFDTIMMVENYPHDAAGLGC-CEGC : 1859
staC : VPRVRRLDCGLRVLLELLQLQRRQAALT EHQYMGLSEIQKAVCPGAVFDTIMMVENYPLDAAGLGC-DDGC : 1858
dbv17 : VPRVRRLDCAVPVDLLQLQRRQSSLT EHQHGLCLPEIQKAVCPGSIFFDITIMMVNYPLDADGLD--DGC : 1853
cepB : VPRVRRLDCRVPVVEMLKQLQCRQSALT EHQYMGLPEIQKAVCTGAIFFDITIVMIENYPHDAAGLGC-DDGC : 1863
bpsB : VPRVRRLDCRRVLEMLLEQLQARQAALT EHQYVGLPEIQKAVCTGAIFFDITIVMVENYPHDAAGLGC-DDGC : 1862
comC : LPRVRRLDCAESAVRMLAPLQDRQSALLAHQHVGLAEIQSLGCACAVFDTHLMFENYPCGCPHAAATDE : 1880
6P RVRRLDC 6 6L dLQ RQ alTeHQ 6GL eIQ C Ga6FDT66M eNYP a gL dgg
    
```

```

*      1900      *      1920      *      1940      *      1960
teiC : VRISSVLTTRTCTSYPLTMSASPCEERLQIPVSYRPDRERETAABAAGQVVRVIERIVARPSLAVGRLCVT : 1929
staC : VTISAVTRTCTCYPLTMSASPCEERLQIPVSYRPDRIDERTAEVACQVVRVIERIVARPSLAVGRLCVT : 1928
dbv17 : VAISIRTRTCTTYPLSVSIPCARLQIQLDYRPDWDGGDLAAEITCQVVRVILARMVAREPSLAVGRLAVT : 1923
cepB : VVISVATRTCTSYPLTMMVSLCERLQIPVSYRPDRIDDEETAEVAVQVVRVLEQVVAQPLAVGRLCVT : 1933
bpsB : VAISVVTTRTCTSYPLTMMVSLCERLQIPVSYRPDRIDDATAEVARQVVRVLEQVVAEPSLAVGRLCVT : 1932
comC : VTTERONTTRACTSYPLAVCIMPDRMRVHTYRPDLDFRDEALHVARQWERVLEQLVADPAAVVGRVQV : 1950
V 6 s TrcGT YPL 6 C R6 6 6 SRPD i A e a QVvRV6e 6vA P lpVGR6 Vt
    
```

```

*      1980      *      2000      *      2020      *
teiC : GESVRRERVVERWNTGVAAGGSSVLTGFRRRARSSPDAAVAVSGGRTLSYAALDRESDRLACHLACRGVR : 1999
staC : SDARTSEVERWNTGVAAGGASVVELFRRRSAGTPDAAVAVDGRILSYGDLRESDRLACHLAEITGVR : 1998
dbv17 : SRSTRGCVTERWNTGVAAGGSSVPELFRQDAAPDATAVIGDGRILSYAGLDRESDRLACHLARRGVR : 1993
cepB : SEPARASVVERWNTGVAANRTSVLELFRQQADASPDAAVAVDAARTLSYADLDRESDRLACHLAAAGVR : 2003
bpsB : SEPTRAAVVERWNTGVAEAETSVLELFRQQAGASPDAAVAVAGERTLSYADLDRESDRLACHLACTGVR : 2002
comC : GPLERGLVVEGWNTRTRTLGSSVLELFRARVAQAPDAAVAVDGERQVSYAELDAESNLAAHLRERCVG : 2020
R vvErWn Tg aa SV eIFR PDavAV Rt6SYa LDrESrlAg La GV
    
```

```

*      2040      *      2060      *      2080
teiC : RGDVGVVMEHRCDFVALLLAVKAGAAVPPVNDYPPDRIERMLTDACVSAVAVCVQATSGAVPDGCLAPV : 2069
staC : RGDVGVVMEHRCADLVAVLAVKAGAAVPPVHDYPPVRIERMLDAGCVTVAVCAECTRNAVPDGCLAPV : 2068
dbv17 : RGDVGVVMEHRCADLIVALLAVKAGAAQVPPVNDYPAERIERMLDAGCASVAVCAGATRHAVPDGCLBPV : 2063
cepB : RGDVGVVMEHRCDFVALLLAVKAGAAQVPPVNDYPAERIERMLDAGCASVAVCLEATRKAVPDGCLBPV : 2073
bpsB : RGDVGVVMEHRCADLVALLGVKAGAAQVPPVNDYPAERIERMLDAGCASVAVCVVEATRKAVPDGCLBPV : 2072
comC : RGDVAVVRLERSTDLIAALLGVKAGAAVPPV GAYPAERVAVFVLEDSSPAVTH-----DAVVKC---- : 2080
RGDrVgV 6eRg DL vALL 6wKAGAA VPV dYP R6er6L D g Va6c t AVpdG pv
    
```



```

*      2120      *      2140      *      2160      *
teiC : VMDSPAIAAAPSEAPPITVCAHDLAYVMYTSGSTGPKGVAVPHGSVAALACDPGWSQPCDGVLMHAPH : 2139
staC : PVDAPWAGETRHEPT--VTARDAYVMYTSGSTGCEPKGIWPHGSVAALACDPGWDALDADDGVLMHAPH : 2136
dbv17 : VMDAPATEARPEHAPPLAVCAHDLAYVMYTSGSTGPKGVAVPHGSVAALACDPGWSQPCDGVLMHAPH : 2133
cepB : VMDVPAIDCVRHEAPQVTVCAHDLAYVMYTSGSTGPKGVAVPHGSVAALASDPGWSQPCDGVLMHAPH : 2143
bpsB : VVDLVIGVGRPEAPPITVCAHDLAYVMYTSGSTGPKGVAVPHGSVAALASDPGWSQPCDGVLMHAPH : 2142
comC : -----GGEAPSVAVTEGD LAYVMYTSGSTGCTPKGVAVPHGSVAALVGETGGLCPGDSVLFHAPH : 2140
d p EaP V a D AYVMYTSGSTG PKG6aVPHGSvAALA dpGW g D VL HA H
    
```

A5

```

*      2180      *      2200      *      2220      *      2240
teiC : AFDASLLEIWWPLSCAVHVVADPG-AVDAQLRREAVDRCVTVHLTAGSFRFLAEESSPDARGLREVLT : 2208
staC : AFDASLFEIWWPLVRCARVHVAEPG-AVDTRQLREAVDRCVTVHLTAGSFRFLAEESSPGSFDGLREVLT : 2205
dbv17 : AFDASLLEIWWPLSCAVHVAEPG-AVDAQLRQVIAVCATVHLTAGTFRFLAEESSPDSFGLREVLT : 2202
cepB : AFDASLVEIWWPLVNGSFWHVAEPG-AVDAERLREAVSICVTVHLTACAFRAVAEESSPDSFTGLREVLT : 2212
bpsB : AFDASLVEIWWPLSCAVHVAEPG-TVDAERLREAVSICVTVHLTACAFRAVAEESSPDSFTGLREVLT : 2211
comC : AFDASLFEVWWPLASCAVVIAPGAVDAARVRRHTAACVTVHVTACLFRFLADEBPECFMCAHREVLT : 2210
AFDASL E6WvPL Ga V66AePG a6Da r6R 6 rCvTtVH6TAG FR 6AeEsP F G1r6LT
    
```

```

*      2260      *      2280      *      2300      *
teiC : GGDVWPLASVWVRLRETCEPIRVVRELYGPTETTLCAITWHLIEPCVATGDTLPICRPLAGRRAYVLDADFLLQP : 2278
staC : GGDVWPLASVWVRLRETCEPIRVVRELYGPTETTLCAITWHLIEPCVATGDTLPICRPLAGRRAYVLDADFLLQP : 2275
dbv17 : GGDVWPLASVWVRRACPEVVRVRELYGPTETTLCAITWHLIEPHTETGDTLPICRPLAGRRQVYVLDADFLLQP : 2272
cepB : GGDVWPLASVWVRRACPEVVRVRELYGPTETTLCAITWHLIEPCVATGDTLPICRPLAGRRQVYVLDADFLLQP : 2282
bpsB : GGDVWPLASVWVRRACPEVVRVRELYGPTETTLCAITWHLIEPCVATGDTLPICRPLAGRRQVYVLDADFLLQP : 2281
comC : GGDVWPLASVWVRRACPEVVRVRELYGPTETTLCAITWHLIEPCVATGDTLPICRPLAGRRQVYVLDADFLLQP : 2280
GGD VPL sV r6R aCP 6RVR LYGPTE 3LCAITWh66EPg   tgd LP6GrPLAgR   YVLDADFLLQP
    
```

```

*      2320      *      2340      *      2360      *      2380
teiC : VAPNVTGELYLAGAGLARGYLCAAAATSERFVA-----DPPFACGERMYRTGDLARWTDQCELLFAGRADA : 2343
staC : VAPNVTGELYLAGAGLARGYLCARCATSERFVA-----DPPVPGERMYRTGDLARRNDQCELLFAGRADA : 2340
dbv17 : VAPNVTGELYLAGAGLARGYLCAPARTSERFIAVPAASVWPAASGERMYRTGDLARWTDQCELLFAGRADS : 2342
cepB : VAPNVTGELYIACAGLARGYLCENNGSTSERFIA-----NPFASGERMYRTGDLARWTDQCELLFAGRADS : 2347
bpsB : VAPNVTGELYLAGAGLARGYLCNTAARTSERFVA-----NPFSCGGERMYRTGDLARWTDQCELLFAGRADS : 2346
comC : VPPFVAGELYIACAGLARGYLCRAGLSERFVA-----CPEFDGERMYRTGDLARWTDQCELLFAGRADA : 2345
VaPrVtGELY6AGaGLA GYLg   3   ERF6A   Pf   GeRMYRTGDLARWtd GEL FAGRAD
    
```

A5

```

*      2400      *      2420      *      2440      *
teiC : QVKIRGCVREVEPCEIETVAEAPQVAQTVVAREDCPCEKRLIAYVTAEAGQCPDPAVREHLGERLPEFHM : 2413
staC : QVKIRGCVREVEPTEIETVLAEPQVAQTVVAREDCPCEKRLIAYVAIEPDQVLDPEALREHLARLPEFHM : 2410
dbv17 : QVKIRGCVREVEPCEIETVAEAPQVAQTVVAREDCPCEKRLIAYVTAEAGSGLDPAVREHLGERLPEFHM : 2412
cepB : QVKIRGCVREVEPCEIETVAEAPQVAQTVVAREDCPCEKRLIAYVTAEAGPALAADAVREHLAARMPEFHM : 2417
bpsB : QVKIRGCVREVEPCEIETVAEAPQVAQTVVAREDCPCEKRLIAYVTAEAGSALESAAVRAHLARLPEFHM : 2416
comC : QVKIRGCVREVEPCEIETVAEAPQVAQTVVAREDCPCEKRLIAYVTAEAGSGLDPAVREHLARLPEFHM : 2414
QVKIRG5RVEp E6E aL   P   VAQaVVVARE   PGeKRL6aY   a   a6R hLa r6PE5M
    
```

```

*      2460      *      2500      *      2520
teiC : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2483
staC : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2480
dbv17 : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2482
cepB : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2487
bpsB : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2486
comC : VPAWVVLDFPFLTINGKIDRAALPAPEPSCRPAGREPRTETSERVLCCLFADLLCLDRUGADDSFFELGG : 2483
VPA 666Ld   PLT   NGR6Dr   ALPaPEf   k   AgREPRTe   E466C   LFA   6Lg6   rvGa   DsFFELGG
    
```

T5

```

*      2540      *      2560      *      2580      *
teiC : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--ACPDCCVGEVANTPVMRGLDRDP : 2551
staC : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--AELADDVGEVANTPVMRGLDRDP : 2548
dbv17 : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--ASLDDCVGEVANTPVMRGLDRDP : 2550
cepB : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--VEDAECVGEVANTPVMRGLDRDP : 2555
bpsB : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--ACVADCVGEVANTPVMRGLDRDP : 2555
comC : DSILSMQLAARARREGVVFCAQVFEQRTPAATAAVALRGGQRP--ACVADCVGEVANTPVMRGLDRDP : 2553
DSI6SM LAArArR   VFgAk   VFE   TPA   IAAVA   rgg   d   vGEV   wTPV6Ra6   RdP
    
```

```

*      2600      *      2620      *      2640      *      2660
teiC : GTMNRSAQQQVTVGAPDLSRDVLAACLGAVLDAHMLRSRVVAAGTEPEPLVVAAPRCVHRAAGLVDRV : 2621
staC : GALTGCALTQVVTVGAPDLSVDALAACLGAVICTHDLRSRIVTAAGEPEPQVVSARSADVDAAGLVDRV : 2618
dbv17 : DAMTFCALSQQVTVGAPDLSVDVLAACLGAVLDAHMLRSRIVRTGAAQPLVVAAPRCVHRAAGLVDRV : 2620
cepB : IGLTRCAMAQQVTVGAPDLSVTALAACLGAVLDAHMLRSRIVDTGAEPEPLVVAAPRCVHRAAGLVDRV : 2625
bpsB : AGLTRCAMAQQVTVGAPDLSVTALAACLGAVLDAHMLRSRIVDESEPEPLVVAAPRCVHRAAGLVDRV : 2625
comC : RAVETCGLAQQVTVGAPDLSVTALAACLGAVLDAHMLRSRIV--EPGAEPEPLVVAAPRCVHRAAGLVTRI : 2621
6   ga6   QWv   v   P   Ls   LaaGLCAV6d   HD6LRsR6v   2Pr6vVa   Rg   VdAa   LV   R6
    
```

```

*      2680      *      2700      *      2720      *
teiC : EAGACEVDDLTERRSAAEAATRLDPGACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2691
staC : DADTCDIDELIADRCAREAAARLDPAACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2688
dbv17 : EACTCDVDIADRCARDAAARLDPAACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2690
cepB : EAGADDVDEIADRCANDAAARLDPAACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2695
bpsB : EAGCDLADIADRCADHTAARLDPAACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2695
comC : TATVCAINELIADRCARDVTRLDPAACVMVRAVVVDACPCRVCRVLVVAHHLVVDVAVSWRILLEDLQAC : 2690
A   g   6   e6adR   A   aa   RLDP   AGV6   rAVW6DAGPgr   GR6V6   aHHLvvd   VSW   LL   DL   aAC
    
```

E5

```

*      2740      *      2760      *      2780      *      2800
teiC : EAVVAGCPALDPDVSFRRWARTLDAQA--TRTAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2757
staC : EAVSAGAEALDPDVSFRRWARTLDAQA--VRAAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2754
dbv17 : EAVVAGCPALDPDVSFRRWARTLDAQA--TRTAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2756
cepB : EAVVAGCPALDPDVSFRRWARTLDAQA--TRTAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2761
bpsB : EAVVAGCPALDPDVSFRRWARTLDAQA--TRTAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2761
comC : EDTDAGCPALDPDVSFRRWARTLDAQA--TRTAELETWTEILDGARSRLCE--LDPKRDVSTAGRK : 2760
Eav   AG   P   ldp   6sFRRWartL   2A   R   EL   WT   6L   a   r   g   LDP   RDT   sTagR
    
```

```

*      2820      *      2840      *      2860      *
teiC : TWTVPQDRAGVLEKVTSAFHCCVHEVLLATLACAVARWR-----GCAAVVVDVECHGRRLCEMDLSR : 2821
staC : SWTLPDRDAGVLEKVTSAFHCCVHEVLLATLACAVTWR-----GCDVVVVDVECHGRRLPVGELDLSR : 2818
dbv17 : SWTLPDRDAGVLEKVTSAFHCCVHEVLLATLACAVARWR-----GCTAVVVDVECHGRRLPELDLSR : 2820
cepB : SWTVPQKHAGVLEKVTSAFHCCVHEVLLATLACAVASWR-----GCTVVVVDVECHGRRLPEMDLSR : 2825
bpsB : SWTVPREHAGVLEKVTSAFHCCVHEVLLATLACAVACWR-----GCTAVVVDVECHGRRLPELDLSR : 2825
comC : TWTLEPAQRELWNRAPNAFHCCVHEVLLATLTCGIRQWRERTAEEDAMVLLDIECHCREALDGVLDLSR : 2830
3WT6P AgvLV sASHCCVHEVLLATLaCa6a WR Gg V66D6EGHGR p6 e6DLSR

*      2880      *      2900      *      2920      *      2940
teiC : TVGWFTDVEHPLRLDVTGIDTAEVTIAGCCAAAGLLKTKVKNVRAVVDGCLCYGMLRHLNAETCPVLAAPQA : 2891
staC : TVGWFTDIEHPVRLDVTGIDTDEVTIAGCCAAAGLLKAVKNVRAVVDGCLCYGVLRLHLNAETAPVLAALPK : 2888
dbv17 : TVGWFTDVEHPLRLDVTGIDPAEVTIAGCCAAAGLLKQVKNVRAVVDGCLCYGILRYLNAETCPVLAALPK : 2890
cepB : TLGWFTDVEHPLRLDVTGVDTAEVTIAGCCAAAGLLKQVKNVRAVVDGCLCYGMLRYLNAETCPVLAALPK : 2895
bpsB : TLGWFTDVEHPLRLDVTGVDPAEAVAGCCAAAGLLKQVKNVRAVVDGCLCYGMLRYLNAETCPVLAALPK : 2895
comC : TVGWFTSVPVRLDVTAGVDLADALAGCPAAGALLKAVKQARLPGDCLGFEELRHLNPETAPALAEIPV : 2900
T6GWFTd6HP6RLdvtG6D ae 6AGG AAG LLK VKEnvR 6PdgGLG5g6LR LNaeT LAa P

*      2960      *      2980      *      3000      *
teiC : AEICFNLYCRFSECPGCDVPPWQLCTIGCTAEQDTPMPLRHAVEIDAAVLDAAGCPRLLELTVTWAGRLVCA : 2961
staC : AEICFNLYCRFSSRPECDEPWPQMTGTTCGAAEHEHAPLRHAVEVDAVWLDSPDGPPELRLTLTWAGRLICE : 2958
dbv17 : PEICFNLYCRFPSSRACAPPEWQLCTIGCTAEQDTPALRHAVEIDAAVLDAAGCPBELSLVTWAGRLICE : 2960
cepB : AEICFNLYCRFSSRPDCEAQPWQIKGIVGCAAEQDTPLRHVVEIDAAVLDAAGCPPEFTLTVTWAGRLVGD : 2965
bpsB : AEICFNLYCRFSSACSGEAQPWQITGIVGCAAEQDTPLRHVVEIDAAVLDAAGCPPEFTLTVTWAGRLLED : 2965
comC : PQVGFNLYCRSTAPTGTGVEPWQATACIGCSVDPATPVKHVIBACAARQDTPDGPPELTLSLWAGEVLDT : 2970
26GFNLYCRf g 2PWQ g 6GG ae p64H 6E dA V D dGPe L363WAGr66g

*      3020      *      3040      *      3060      *      3080
teiC : AEAEELGRAWQDVLACLAAHAGDVRAAGCHTPSPDPLTALSGQEWAVQIEAAVVDLDDIUPLSPLQEGCLLFH : 3031
staC : AEAEELGRAWLDDLSCLAHAA-LCAGCHTPCDFSLTALTQDDVEEIEAAVVDLADIUPLSPLQEGCLLFH : 3027
dbv17 : AEAEELAQAWLMDLGLAAHVGGCCAGCHTPSPDPLISLTLQDDVAEVEAAVVDLDDIUPLSPLQEGCLLFH : 3030
cepB : AEAEELFANAWLMDLSCLAHVEACGAGCHTPSPDPLTLTQREVAEIEAAVVDLDDIUPLSPLQEGCLLFH : 3035
bpsB : AEAEELAQAWLMDLACLAAHVAAAGCPGCHTPSPDPLTALTQREVAEIEAAVVDLDDIUPLSPLQEGCLLFH : 3035
comC : TDAEALGRAWLMDLGLAAHVADPCAGCHTPSPDPLLDLQREVAEIEAAVVSALADVVPLSPLQEGCLLFH : 3040
aeAEs1 AW1 6L GLA H gaGCHTPsDFpL L3q Va2 eAAVp L D6WPLSPLQEGCLLFH

*      3100      *      3120      *      3140      *
teiC : AV-DEGCPDYYASRNTLQDGLDVARFRASWQAVLDRHAALRASFHQLESQAVQIAREVTLPPQETD : 3100
staC : AV-DEGCPDYYAGNTLQDGLDVARFRASWQTLDRHAALRASFHQLESQAVQIAREVTLPPQETD : 3096
dbv17 : AA-DEGCPDYYAGRKLQDGLDVARFRASWQALLDRHPALRASFHQLCSQAVQIAREVTLPPQETD : 3099
cepB : AA-DEGCPDYYASRNTLQDGLDVARFRASWQTLDRHAALRASFHQVESCTAVQIAREVTLPPQETD : 3104
bpsB : AA-DDGCPDYYASRNTLQDGLDVARFRASWQVLLDRHPALRASFHQLESQAVQIAREVTLPPQETD : 3104
comC : AAYDEGCPDYYEGQRWLELQDGLDVARFRASWQAVVARRHETLRAGFRQLESQAVQIAREVTLPPQETD : 3110
A De GPDVYa R L 6dGPLDVaRfRAsWq 66dRH aLRAsFhQ6 SG aVQ 6Ar 6 W etD

*      3160      *      3180      *      3200      *      3220
teiC : LSDLPEDEALAEEDRLAELRLDREFDLTRAPQLRLHLVRLGERRHRLAFASHHISDCGWSLPIVSTVMA : 3170
staC : LSHLPDEALAEABRLABELHADRFELDPAPLRLHLRLRISERSHRLALASHHISDCGWSLPIVSGDVTA : 3166
dbv17 : LSHLPDEALAEEDRLAELRLDREFDLTRAPQLRLHLVRLGERRHRLVLTSHHIVADGWSLPLITDVLIT : 3169
cepB : LSHLPDEEAVAEEDRLAELRLDREFDPTQAPLRLHLVRFCEERHRLIFTSHHIAADGWSLPLICVDVLA : 3174
bpsB : LSHLPDESEALAEEDRLAELRLDREFDLTRAPQLRLHLVRLCDREYRLIFTSHHIVADGWSLPLILVDVLIT : 3174
comC : VSELPEAEALAEEDRLAELDQEHRFDLTAHAPLRLVLLVRIARERHRLAFTSHHIASDCGWSLPMVWSEWAR : 3180
6S LPE eA6Aef RLA Rfd AP LR6hL6R eR hRL SHHI DCWWSLP66 V

*      3240      *      3260      *      3280      *
teiC : AYE----GRGLPAPTYSYRDLWLRQDKAAARQAWRRELACLDEATEVVPPEITITAPAPQVRVFCLEDD : 3236
staC : AYESGCDGRSLPAPTYSYRDLWLRQDKAAARQAWRRELACLDEPTHEVVPPEAVTAPVEPERLRFDMDE : 3236
dbv17 : VYESGCDGRALPAPTYSYRDLWLRQDKAAARQAWRRELACLDEATEVVPPEITITPLEPERVGFELDE : 3239
cepB : AYEAGCDGRTLPAPTYSYRDLWLRQDKAAARQAWRRELACLDEATEVVPPEISIRTPVPERVRFELCD : 3244
bpsB : AYEAGCDGRTLPAPTYSYRDLWLRQDKAAARQAWRRELACLDEATEVVPPEISITPLEPERVAFELDD : 3244
comC : LYEAGCDPRALPAPTYSYRDLWLRQDKAAARQAWRRELAPVEPTLVVPPADRVITAPAEPRVITYELDE : 3250
YE ggdGR LPa TSYR 5L W6aRQDK AAr AWR ELAg6dE T VVPp 6 P P2R6 5 6d

*      3300      *      3320      *      3340      *      3360
teiC : ELSRRLVETTRRHCVTMTLFCQIWSLLLRLTCRQDVVFGSAVACRPAEIPCVESWVGLFHMNLPVVR : 3306
staC : DLSRRLVETTRRHCVTMTLFCQIWSLLLRLTCRQDVVFGSAVACRPEEIPCVESWVGLFHMNLPVVR : 3306
dbv17 : ALSRRVETTRRHCVTMTLFCQIWSLLLRLTCRQDVVFGSAVACRPEEIPCVESWVGLFHMNLPVVR : 3309
cepB : EPRSRLVETTRRHCVTMTLFCQVWALHLRLTCRQDVVFGSAVACRPEEIPCVESWVGLFHMNLPVVR : 3314
bpsB : ETSKRLVETTRRHCVTMTLFCQVWALHLRLTCRQDVVFGSAVACRPEEIPCVESWVGLFHMNLPVVR : 3314
comC : ELGRRLVAGLGCESLTLNLLQCVWALHLRLTCRQDVVFGTIVACRPEELPCVESWVGLFHMNLPVVR : 3320
s4R6V f rrHg6T NTLfQG6Wa6 LARLTCR DVVFG aVACRP E6PGVeS 6GLF6N LPVR r

```

E5

```

*          3380          *          3400          *          3420          *
teiC : LDCAEPFLDMLTDLQRRQVAMHMQHGLCLSEIKQVACPGAAFDTLVVFENYPRPPRPSDDPDALVVRPAG : 3376
staC : LDGGEFFLDMLTRLQRRQVAMHMQHGLCLSEIKQVACPGAAFDTIIVVFENYPRPPLPSEDDPDALVVRPAG : 3376
dbv17 : LACAEFFLDMLTDLQRRQVACMPHQHGLCLSEINQVACPGAAFDTIIVVFENYPRPPRPEGPDALVVRPAG : 3379
cepB : LTCAEPVVDMLRDLQRRQVAMHMQHGLCLSEIKQVACPGAAFDTIIVVFENYPRPPRPFDDPDALVVRPAG : 3384
bpsB : LTCAEPVVDMLRDLQRRQVAMHMQHGLCLSEIKQVACPGAAFDTIIVVFENYPRPAPPSEDDPDALVVRPAG : 3384
comC : LDCAEPFLDMLTDLQRRQVAMHMQHGLCLTEIQRLAGPAAFDTLVVVYENYPRPPLCSGDDPDALSIRPAG : 3390
L Ga2P 6dML dLQeRQ6A M HQH6GL EI q aGpgAaFDT6VV5ENYP pP          PDA6v6RPag

*          3440          *          3460          *          3480          *          3500
teiC : IPEDSGHYPPVSMRASHD-CGVHCFEYFRPDADFRAEACMLAAITBALRQVVAEPCIAVGPICFLVGAAR : 3445
staC : IPNDSGHYPISMRASHD-CGVHCFEYFRPDVDFRTRAEEMLAARVVRALAEQVADPWTLVGRVCLIGPTEP : 3445
dbv17 : IPNDTCHYPLSMRASHV-CRVHCFEYFRPDVDFRAEAEMLASILRALRQVVAEPRVFPVGPVCLIGPTEP : 3448
cepB : TPNDTCHYPLSIRASVAACPVQCFEYFRPDVDFRAEAEVVAALLRALRQVADPPRMPVGVQVCLIGPQQR : 3454
bpsB : IPNDTCHYPLSMRASHVAAACPVQCFEYFRPDVDFRTEAGEMVAAILRALRQVVAEPWTLPVGVQVCLIGPTEP : 3454
comC : TPEDAGHYPLTIVAVLEDERVQCDLVYRPFDAFERAEVDELITLALVLTLEQLVADPSAPVGPVCLVTCAR : 3460
P D GHYP636rAs g V Gef65RPD dR ea e66aa66raLeQ6 A P          VG 6g6 G 2R

*          3520          *          3540          *          3560          *
teiC : RQVVEEWNEDAPLTAQP-LPMLFRPQARSPDAVAVRDAARSLSTGALLCEAEGLARLRLWASGVRRRETR : 3514
staC : CLVVEEWNATDAFFVPEP-LPVLFRPRAELSRDAVAVEDADRKLSYGCLLEBAEDLARLLAGICVRRRETR : 3514
dbv17 : PLVVEEWNRTCVPPAEP-VPMLFRPQVSPDAVAVVDAARSLSTSGELDEAEELARLLVGLCVRRRETR : 3517
cepB : PLVVEEWNQTDAPLAAET-LPULFREQAQSPDAVAVQCVRSLSSTGCLDEVEBALARLLVGLCVRRRETR : 3523
bpsB : PLVVEEWNRTDVFLLAET-LPULFRPQAEPSDAVAVEDCARSLTFGCLLCEVEALARLLVGCAGVRRRETR : 3523
comC : DMVLGEWNTATADRPRAAALPELFAAQAAHTPGNTALIGADRELTVAELREPSGRLATHLRYLCVAAETP : 3530
V6 EWN T p          6P LFr qa 3pda A6 d R L35 L1 E e LAr L          GVrrE R

*          3580          *          3600          *          3620          *          3640
teiC : WVLVRSRSELVWALLCVSLAGCVFVPVDDPYPCDRIALMLLDAAPQVLVCTARTRSVWVPCDFEAGVLLAL : 3584
staC : WGVLCRSSELVWALLCVSLSGCVFVPVDPEYPACRIELMLTDSAPVLLCTKATRTAVPDCFAGVVWVLL : 3584
dbv17 : WGVLCRSSELVWALLCVSSAGCVFVPMDPDYPRERISFILEDSAPVLLCTSETRQAVPPEEFAAVWVLL : 3587
cepB : WGVLCRSSELVWALLCVSFAAGCVFVPVDDPYPSERIEMLANSAPVLLVCTKTTTRAVVPAQFTGTVLVLL : 3593
bpsB : WGVLCRSSELVWALLCVSFAAGCVFVPVDDPYPRERIEMLANSAPGVVVCVCTKTTTRAVVPAQFTGTVLVLL : 3593
comC : WALLVRSSELVWALLCVSRAAGCVFVPVDDPAHPADRIAYLLEDAAPPVLLCTARTRAVVPADYPCRNWVLL : 3600
V 6LV RS E6 V 66gVS aGvFVP6DP yP R6 6L 1 AP V66CT TR VP 5 G v6 L

*          3660          *          3680          *          3700          *
teiC : DEPS----AACPQVSLPRVAARDGAYVIYTSGSTCWPKCVLVTHACLNLASAQIERFCVTSSSRVLQFA : 3650
staC : DERS----AADPAAALPAWAVCDGAYVIYTSGSTCTPKCVLVTHACLNLACAQIERFCVTSSSRVLQFA : 3650
dbv17 : DAPL----AADPRTALPRWEACD GAYVIYTSGSTCWPKCVLVPHACLNLASAQIERFCVTSSSRVLQFA : 3653
cepB : DELP----AADPGVELPDVWPEDAAYVIYTSGSTCWPKCVLVTHSCLVNLGYAHIERMAVTSSSRVLQLS : 3659
bpsB : DELP----AADPDVELPPWPEDAAYVIYTSGSTCWPKCVLVTHSCLANLGYAHIERMAVTSSSRVLQLS : 3659
comC : DELDSSALPAPFFATP-WPPDQAAYVIYTSGSTCAPKCVLVPHACLNLAAAQIDREAVGPDARVLQFA : 3669
De          aA P 1P V d AYVIYTSGSTC PKCV V H CL NL A IeR Vts sR6LQ

*          3720          *          3740          *          3760          *          3780
teiC : ALGFDAAISEVCMALLSGCSTVLDAEAMPPRVSLCDAVRRWDITHVTVPSPVLAVEDDLPPDL-TLIVA : 3719
staC : SLGFDAAISELCHALLSGCTVVLADPDSMPPRVSLCDAVRRWRITHVTVPSPVLAVEDDLPSLETLVVA : 3720
dbv17 : ALGFDAAVSELCMALLSGCTVVLDAEAMPPRVSLCDAVRRWGITHTVTVPSPVAVEDDLPSLETLVVA : 3723
cepB : ALGFDAAVSELMALLAGATLVLPDSANMPPRVTLGDAIRKAGITHLTVSPSVLASEDDLPDLTQVTLTG : 3729
bpsB : ALGFDAAVSELMALLAGATLVLPDAAMPPRVTLGEAIRKAGITHLTVSPSVLASEDDLPDLTQVTLTG : 3729
comC : SLGFDAAVSELMALLSCAAAVVAPAEATLPLVSLTEALRHWDVTHVTVPSPALATADELPPCLRTLVA : 3739
GFDA 6SE6 MALL G 66 d 6PPrV3Lg A6R 6TH6TV Psv1A eddLPD L T66

*          3800          *          3820          *          3840          *
teiC : GEACPPALVDRWSPGRRMINGYCPTEITTVCAHMSPLSPCRGAVPICCPIANMRTYVLDaFLQVPPPGMT : 3789
staC : GEACPPALVDRWSPGRRMINGYCPTEITTVCVTHMSAPLSPCRDAPPICRPIANMRAYVLDaFLQVPPPCVA : 3790
dbv17 : GEACPPALVDRWSPGRRMINDYCPTEITTVCAHMSPLSPCRDVPVIGRPITCLRAYVLDaFLQVPPPGVT : 3793
cepB : GEALPVALVDRWSPGRRVIQAYCPTETTICSTMSAPLSPCHQVPLGCPINVRHYVLDaFLQVPPPCVW : 3799
bpsB : GEALPVALVDRWSPGRRVIQAYCPTETTICSTMSAPLSPCHQVPLGCPINVRHYVLDaFLQVPPPCVW : 3799
comC : GEACPPALADRWAGRRMINDYCPTEITTVCAHMSPLAPCADVVPVGRPIANMRTYVLDaFLQVPPPGVT : 3809
GEA PpALvDRWs gRR6I YCPTEIT6C 3MS PLsPG d vP6G PI n R YVLDaFLqVPPPG6

*          3860          *          3880          *          3900          *          3920
teiC : GELYVTCAGLARGYLGRPGLTAERFVA-----DPESSGGRMYRTGDRARWTCDEGLVFAGRVDAQVKVRC : 3854
staC : GELYVTCAGLARGYLGRPGLTAERFVA-----DPESSGGRMYRTGDRARWTRREGELVFWGRADAQVKVRC : 3855
dbv17 : GELYVTCAGLARGYLGRPGLTAERFVAVPASVSPARFGEMRYRTGNRARWTRDCELVFTGRADAQVKVRC : 3863
cepB : GELYVTCVGLARGYLGRPGLTAERFVA-----SPEAFGERMYRSGLDFRWTRREGQLMFAGRVDAQVKVRC : 3864
bpsB : GELYVTCVGLARGYLGRPGLTAERFVA-----SPEAFGERMYRSGLDFRWTRREGQLLFAGRVDAQVKVRC : 3864
comC : GELYVTCAGLARGYLGRPGLTAERFVA-----DPEITGGERMYRTGDLARWTRDGLLVFAGRADAQVKVRC : 3874
GELY6 G CLARGYLGRPGLTAERFVa Pf pG RMYR3G1 RWT G L6F GR DAQVKVRC

```

```

*      3940      *      3960      *      3980      *
teiC : YRIEPCIEAVLAEHPGVAQVAVVAREDCPGERQLVAYVVPAAACP----- : 3899
staC : YRVEPCIEAVLSDHPCAAQVAVVAREDCPGERQLVAYVVPAAAQ----- : 3900
dbv17: YRIEPCIEAVLADHPCVAQVAVVAREDCPGERQLVAYVVPAAAEQ----- : 3908
cepB : YRVEPCIEAVLIEHPGVAQVAVVAREDCPGERQLVAYVVPASADA----- : 3909
bpsB : YRVEPCIEAVLAEHPGVAQVAVVAREDCPGERQLVAYVVPASADA----- : 3909
comC : HRIEPCIEAVLAEHPGVAQVAVVAREDCPGERQLVAYVVPASADA----- : 3944
      yR6EP E6EAVL HPgvaQv 6 R DgPG 4 L6aYVVP A

      4000      *      4020      *      4040      *      4060
teiC : ----- : -
staC : ----- : -
dbv17: ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : HTEAFPGENFACWSSYDQPIPLAEMREWERTVDRIRALRPRRVLEIGVGTGLLLSRLAGCEEYWCT : 4014

*      4080      *      4100      *      4120      *
teiC : ----- : -
staC : ----- : -
dbv17: ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : DFSAPVVAELRRHVERDPELAAKVRLRVQAADETDGLPEGHFDTVVLNSVVQYFPNADYLRGVLEWALRL : 4084

      4140      *      4160      *      4180      *      4200
teiC : ----- : -
staC : ----- : -
dbv17: ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : VAPCGAVFVGDVRNHRLRPFATAVETARTADQDVFTLRRAVEQNLTIEKELLVDPEFFAALGTELPDAA : 4154

*      4220      *      4240      *      4260      *
teiC : ----- : -
staC : ----- : -
dbv17: ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : GADIRLKRASHHNELSRHYDVVLRKTCITTLTVADAPRLTWTEAGDVTGLAQHLTSAHPERLRLTCVPM : 4224

      4280      *      4300      *      4320      *      4340
teiC : ----- : -
staC : ----- : -
dbv17: ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : SRVAHEVALARALDEGVPLPAEEAVTGPDVEELYELGDRLCYVVCVTSASAPDEVDAVLVRTGLTASAA : 4294

*      4360      *      4380      *      4400      *
teiC : -----TAEASTLISALREAAARLPEHMPPAFVPLDAMPLTPNGKVDHRAALQA : 3948
staC : -----PADE--LIALRETAERLPEHMPPAFVPLDAVPLTPNGKVDHRAALQA : 3947
dbv17: -----VAGAPSEACQDGCALISALRESAAGRLPEHMPPAFVPLDTMPLTPNGKVDHRAALRA : 3964
cepB : -----ACENGTLSALRELAERLPEYMPPAFVSLERKMLTPNGKLDHKAALQA : 3958
bpsB : -----AAENGTLSALRELAERLPEYMPPAFVSLERKMLTPNGKLDHRAALQA : 3958
comC : MVDVYTPAPVPPAAVTNDPAACRGTGALLAELEHVERRLPAYMPPAFVPMDRPLTPNGKLDHRAALPA : 4364
      L aLRE aa RLPe M P AFV 6 6PLTPNGK6Dh4AL A

      4420      *      4440      *      4460      *      4480
teiC : PDFAGMSACRDPRTALEARLCELFAEVLCLERVCAGDSDFFELGGDSITSMQLSTRARPEGLELTPWQVFD : 4018
staC : PDLAGNSSCRCPRTAVEELLCAVFAEILCVERVCGDSDFFELGGDSISSMQLSARLRPEGLELTPWQVFE : 4017
dbv17: PDFARSSCRDPPTASAMEARLCELFAEVLCLERVCAGDSDFFELGGDSITSMQLSALAREKGLDLPWQVFD : 4034
cepB : PDFAAANSTKRDPSTPMEARLCELFAEVLCLERVCAGDSDFFELGGDSITSMQLSARAREGLELTPWQVFD : 4028
bpsB : PDFAGMSKRAPPTPMEARLCLAFADVLGLDQVCPDSDFFELGGDSITSMQLSARARPTGLELTPWQVFD : 4028
comC : PDFACRAGCRDPPTPAEALLCRLFAEVLCLERVCAGDSDFFELGGDSITSMQLSARAREGWLTPWQVVF : 4434
      PDFa s R Pr3 Ea LC 6FA 6LG6e VC dSFFe6GGDSI3SMQLs raRr G6 6TPWQVF

```

```

      *      4500      *      4520      *      4540      *
teiT : EKTPERLAALVKEIPADGGAT----- : 4039
staC : EKTPERLAALAKELPAEGEES----- : 4038
dbv17 : EKTAERLAALVVKELPADGEGT----- : 4055
cepB : EKTPERLAALVKEIQAEETA----- : 4049
bpsB : EKTPERLAVIQELAAEGGTT----- : 4049
comC : EKTPERLALLAEAAAEEAATGEVWPWTAMRAVPADGTVPAQWVELAVPAGLDPAALVLRALRAVLDTHAM : 4504
      EKTpERLA 6 e A

      4560      *      4580      *      4600      *      4620
teiT : ----- : -
staC : ----- : -
dbv17 : ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : LRCRICEELVVGEPGAPDAAGLIDRVDADTPLDDAVADAVSRLLDPAGVLRLLWVDAGAESRLVLVVHE : 4574

      *      4640      *      4660      *      4680      *
teiT : -----GAPR : 4043
staC : -----SEPE : 4042
dbv17 : -----GEPE : 4059
cepB : -----AVPK : 4053
bpsB : -----PAPE : 4053
comC : LVVDDSWRILVSDLRATAYEGAAAGQEPVLEPESTSYRRWARLLAEQATADARVEELDGWAAQSEAGEPS : 4644
      P

      4700      *      4720      *      4740      *      4760
teiT : PAAGT----- : 4048
staC : SDDCH----- : 4047
dbv17 : PPAGT----- : 4064
cepB : PAAGT----- : 4058
bpsB : PEGGT----- : 4058
comC : PAAGTETRATWSLTPEITAVLTHRTPETFHCEPSEVVLAALVGAVAEDGTATVMDLEISGRWPFDGADPS : 4714
      p Gt

      *      4780      *      4800      *      4820      *
teiT : ----- : -
staC : ----- : -
dbv17 : ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : RTVGRFAHDI PVRVELAEVDLGSARAGCAAAGELLKTVKEQVRAVPGDDGLCYGLLRHLNPRTAARLAAL : 4784

      4840      *      4860      *      4880      *      4900
teiT : ----- : -
staC : ----- : -
dbv17 : ----- : -
cepB : ----- : -
bpsB : ----- : -
comC : APPTLAFGYHGATPDTGEPWPLARSAVRCSAHALRATAELRDTPNGPGLILELRGRTEELDRLGRAWQ : 4854

      *      4920      *      4940
teiT : -----LVVLSPEQMDQLEAGPGE- : 4067
staC : -----LLVLSPDQMDQLQSCPTDG- : 4066
dbv17 : -----LVVLSPDQMDQLEAGPGE- : 4083
cepB : -----LLVLSADQMDQLEAGLAGE- : 4077
bpsB : -----LVVLSPDQMDQLEAGLAGE- : 4077
comC : ELLDGLAAHTEDPAAGGHTPSDFSLLDLAQSQIEEELAQLEDDQ : 4898
      L6 Ls Q6 L2ag a

```

Abbildung 5-4: Proteinalignment der NRPS BpsB von Balhimycin mit den homologen NRPS anderer Glykopeptide: TeiC (Teicoplanin),<sup>[98,99]</sup> StaC (A47934),<sup>[100]</sup> Dbv17 (Dalbavancin),<sup>[101]</sup> CepB (Chloroeremomycin)<sup>[24]</sup> und ComC (Complestatin).<sup>[102]</sup> Die Domänenstruktur sowie die Annealingstellen der Primer, die für die Amplifikation der vier PCR-Produkte *Tarleft/Tarright/Insleft/Insright* (s. Abschnitt 5.1.2.2) verwendet wurden, sind eingezeichnet. Die beiden rot markierten Bereiche ergaben in der Auftragssequenzierung von *Tarright* und *Insleft* Unterschiede zu den publizierten Sequenzen.

### 5.1.2.2 Konstruktion des Transformationsplasmids pC5A5T4E4

Das Transformationsplasmid pC5A5T4E4 enthält das einzubringende C5A5T4E4-Hybridmodul, im Bezug auf molekularbiologisches Arbeiten im Folgenden auch „Insert“ genannt, zwischen den flankierenden Regionen der gewünschten Insertionsstelle im Genom von *A. baohimycina* Wildtyp, in Analogie im Folgenden auch „Target“ genannt. Die einzelnen Schritte der Konstruktion des Transformationsplasmids pC5A5T4E4 werden im Folgenden beschrieben und sind außerdem in Abbildung 5-5 graphisch dargestellt.

Die flankierenden Regionen der gewünschten Insertionsstelle („Target“) sollten in Form von zwei PCR-Produkten aneinandergesetzt werden. Die zwei mit Hilfe der ProofStart DNA-Polymerase amplifizierte, gelgereinigte und durch Restriktionsverdau geprüfte PCR-Produkte *Tarleft* (2021 bp) und *Tarright* (2012 bp) wurden nach *A-Addition* unabhängig voneinander in pDrive-Klonierungsvektoren kloniert und in CaCl<sub>2</sub>-kompetente *E. coli* XL1 Blue transformiert. Die unter Selektion auf Ampicillin erhaltenen Klone wurden zunächst mittels Blau-Weiß-Selektion und Eckhardt-Lyse auf die Insertion eines PCR-Produktes in pDrive untersucht. Von den hierbei positiv aufgefallenen Klonen wurden die Plasmide isoliert und deren Identität zunächst mittels Restriktionsverdau und anschließend mittels Auftragssequenzierung des gesamten integrierten PCR-Produkts geprüft. Das Ergebnis dieser Auftragssequenzierung erbrachte bei *Tarleft* eine 100 %ige Übereinstimmung der Sequenz mit der in Datenbanken publizierte Sequenz. Bei *Tarright* ergaben sich drei Unterschiede zwischen den Ergebnissen der Auftragssequenzierung und den publizierten Sequenzen, wobei die betroffenen Stellen in Abbildung 5-4 rot gekennzeichnet sind. Die beiden ersten Unterschiede befinden sich im Basenpaarbereich 231 bis 239 (Nummerierung bezieht sich auf das PCR-Produkt). Während die publizierte Version hierfür die Sequenz CGCGGATCC vorsieht, ergab die Auftragssequenzierung die Sequenz GCGGATCCC, was formal durch die Positionsveränderung eines C zustande kommt. Auf Aminosäureebene bedeutet dies eine verhältnismäßig drastische Veränderung von Arg-Gly-Ser zu Ala-Asp-Pro. Allerdings sprachen zwei Argumente dafür, dass die publizierten Sequenzen an dieser Stelle falsch sind. Erstens ergab die unabhängig davon durchgeführte PCR *Insleft* und nachfolgende Sequenzierung über diesen Bereich ebenfalls diese beiden Unterschiede und zweitens passt die neue Aminosäure-Sequenz viel besser zu den in dem Alignment von Abbildung 5-4 verglichenen Aminosäure-Sequenzen der

anderen Glykopeptide. Der dritte Unterschied befindet sich in Basenpaarposition 1224 und entspricht einer T->G bzw. einer Leu->Val Mutation. Da sich auch diese Mutation im PCR-Produkt *Insleft* wiederholt und auch keine drastische Veränderung auf Aminosäureebene nach sich zieht, wurde dieses Plasmid zur weiteren Klonierung verwendet. Das 1992 bp große *Tarright*-Fragment wurde über einen *BmtI/XbaI*-Doppelverdau aus dem Klonierungsvektor ausgeschnitten, gelgereinigt und in das ebenfalls mit *BmtI/XbaI* geschnittene, 5824 bp große *Tarleft*-pDrive-Konstrukt mit Hilfe der T4 DNA-Ligase inkloniert. Dieser Ligationsansatz wurde anschließend in  $\text{CaCl}_2$ -kompetente *E. coli* transformiert. Hierbei war zu beachten, dass neben den über die PCR-Primer eingeführten *BmtI*- und *XbaI*-Restriktionsschnittstellen auch jeweils eine *BmtI*- und *XbaI*-Restriktionsschnittstelle in der *multiple cloning site* des pDrive-Klonierungsvektors vorkommen. Während diese Tatsache das Ausschneiden von *Tarright* nicht beeinträchtigte, musste bei der Auswahl des *Tarleft*-pDrive-Konstrukts auf die geeignete Orientierung des PCR-Produktes geachtet werden, da in einem der beiden möglichen Fälle beim Öffnen des Konstruktes *Tarleft* ausgeschnitten wird. Die aus diesem Schritt unter Selektion auf Ampicillin erhaltenen Klone wurden mittels Eckhardt-Lyse und Plasmidisolierung mit anschließendem Restriktionsverdau auf ihre Identität geprüft. Um ganz sicher zu gehen, wurde außerdem der Übergang zwischen *Tarleft* und *Tarright* mittels Auftragssequenzierung geprüft. Das damit erhaltene *Tarleft-Tarright*-pDrive-Konstrukt enthielt somit die mit „Target“ bezeichneten flankierenden Regionen der gewünschten Insertionsstelle. Diese „Target“-Sequenzen wurden im nächsten Schritt in den Vektor pSP1 umkloniert, der von Dr. Stefan Pelzer eigens für die Transformation in *A. balhimycina* konstruiert worden war und eine Ampicillin-Resistenz zur Selektion in *E. coli* sowie eine Erythromycin-Resistenz zur Selektion in *A. balhimycina* enthält.<sup>[89]</sup> Die hierfür gewählte Vorgehensweise umfasste zunächst die Plasmidisolierung von *Tarleft-Tarright*-pDrive und pSP1 sowie deren *EcoRI/XbaI*-Doppelverdau in präparativem Maßstab. Da im Fall von *Tarleft-Tarright*-pDrive neben dem gewünschten 4001 bp-Fragment ein unerwünschtes 3801 bp-Fragment entstand, welches gelelektrophoretisch nicht befriedigend abgetrennt werden konnte, wurde stattdessen ein Dreifachverdau *EcoRI/XbaI/ClaI* durchgeführt. Das zusätzliche Restriktionsenzym *ClaI* schneidet das unerwünschte Fragment in zwei etwa gleich große Teile. Die Zielfragmente aus beiden Verdau (4001 bp großes „Target“-Fragment und 4501 bp großer pSP1-Vektor) wurden mittels

Agarosegelelektrophorese und Gelextraktion gereinigt sowie anschließend miteinander ligiert und in  $\text{CaCl}_2$ -kompetente *E. coli* transformiert. Die unter Selektion auf Ampicillin erhaltenen Klone wurden durch Plasmidisolierung mit Restriktionsverdau getestet und mittels Auftragssequenzierung der Vektor-Insert-Übergänge verifiziert.

Das sog. „Insert“ wurde ebenfalls in zwei getrennten PCR-Reaktionen *Insleft* (3172 bp) und *Insright* (1375 bp) mit Hilfe der ProofStart DNA-Polymerase amplifiziert, gelgereinigt und durch Restriktionsverdau auf Identität geprüft. Nach *A-Addition* wurden diese beiden PCR-Produkte unabhängig voneinander in pGEM-T-Easy-Klonierungsvektoren kloniert und in  $\text{CaCl}_2$ -kompetente *E. coli* XL1 Blue transformiert. Die unter Selektion auf Ampicillin erhaltenen Klone wurden zunächst mittels Blau-Weiß-Selektion und Eckhardt-Lyse auf die Insertion eines PCR-Produktes in pGEM-T-Easy untersucht. Von den hierbei positiv aufgefallenen Klonen wurden die Plasmide isoliert und ihre Identität zunächst mittels Restriktionsverdau und anschließend mittels Auftragssequenzierung des gesamten integrierten PCR-Produkts geprüft. Das Ergebnis dieser Auftragssequenzierung erbrachte bei *Insright* eine 100 %ige Übereinstimmung der Sequenz mit der in Datenbanken publizierte Sequenz und bei *Insleft* dieselben drei bereits bei *Tarright* beschriebenen Unterschiede zu den publizierten Sequenzen. Das 1359 bp große *Insright*-Fragment wurde über einen *HindIII/XbaI*-Doppelverdau aus dem Klonierungsvektor ausgeschnitten, gelgereinigt und in das ebenfalls mit *HindIII/XbaI* geschnittene, 6183 bp große, gelgereinigte *Insleft*-pGEM-T-Easy-Konstrukt mit Hilfe der T4 DNA-Ligase inkloniert. Dieser Ligationsansatz wurde anschließend in  $\text{CaCl}_2$ -kompetente *E. coli* transformiert. Die aus diesem Schritt unter Selektion auf Ampicillin erhaltenen Klone wurden mittels Eckhardt-Lyse und Plasmidisolierung mit anschließendem Restriktionsverdau auf ihre Identität geprüft. Das damit erhaltene *Insleft-Insright*-pGEM-T-Easy-Konstrukt enthielt somit das mit „Insert“ bezeichnete C5A5T4E4-Hybridmodul.

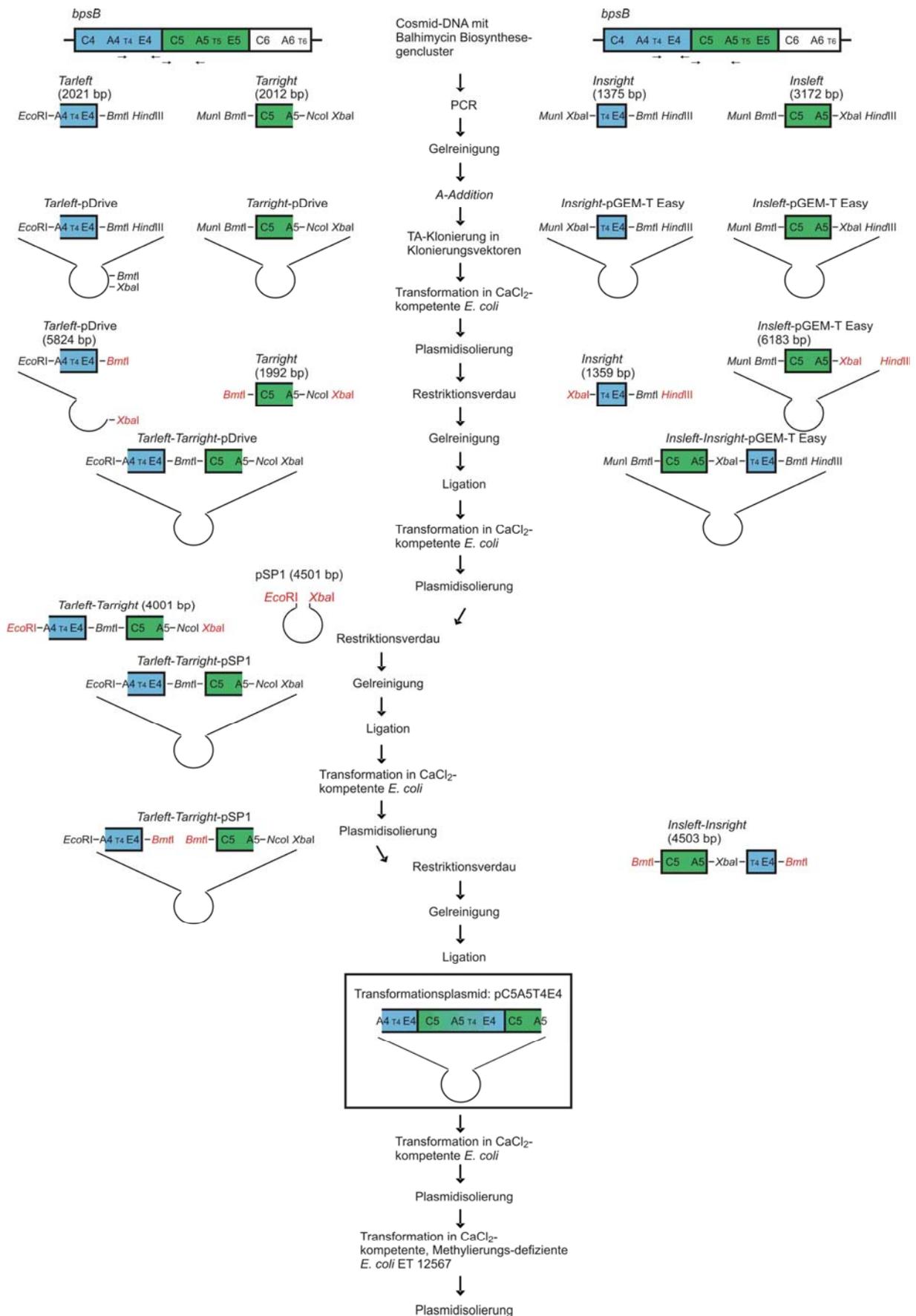


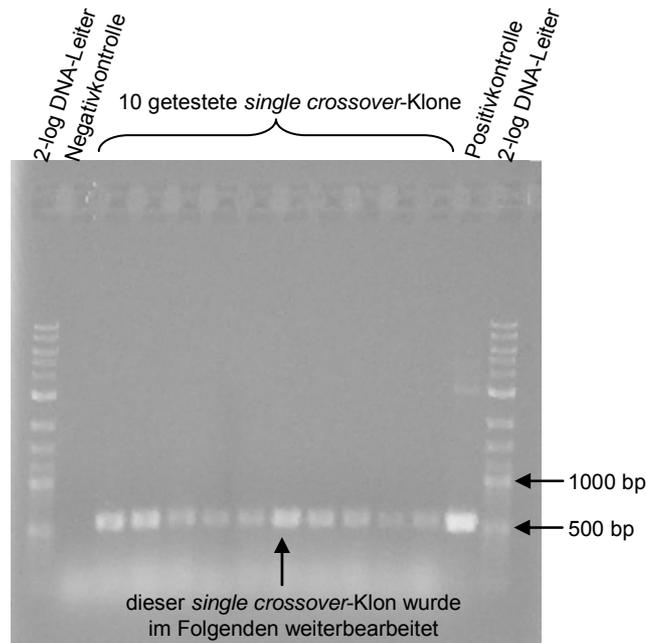
Abbildung 5-5: Graphische Darstellung der einzelnen Schritte der Konstruktion des Transformationsplasmids pC5A5T4E4.

Um aus „Insert“-pGEM-T-Easy und „Target“-pSP1 nun das gewünschte Transformationsplasmid pC5A5T4E4 zu erhalten, war noch ein letzter Klonierungsschritt notwendig. Beide Plasmide wurden isoliert und mit *BmtI* verdaut, was im Fall von „Target“-pSP1 zu einer Linearisierung genau an der gewünschten Insertionsstelle und im Fall von „Insert“-pGEM-T-Easy zu einem 4503 bp großen „Insert“-Fragment führte. Diese wurden gelgereinigt, miteinander ligiert und in CaCl<sub>2</sub>-kompetente *E. coli* transformiert. Auch in diesem Fall erfolgte die Kontrolle der unter Selektion auf Ampicillin erhaltenen Klone über Eckhardt-Lyse und Plasmidisolierung mit anschließendem Restriktionsverdau. Hierbei war es wichtig, auf die richtige Orientierung des Inserts zu prüfen. Das so erhaltene Transformationsplasmid pC5A5T4E4 musste vor der Transformation in *A. balhimycina* in den Methylierungsdefizienten *E. coli* ET12567 transformiert und daraus isoliert werden.

#### 5.1.2.3 Direkte Transformation von pC5A5T4E4 in *A. balhimycina* und Selektion auf *single crossover*-Ereignisse

Da es sich bei pSP1 um einen nicht-replikativen Vektor handelt, konnte das mittels direkter Transformation in *A. balhimycina* eingebrachte Transformationsplasmid pC5A5T4E4 nur durch Integration ins Genom mittels homologer Rekombination (*single crossover*) in der Zelle verbleiben.

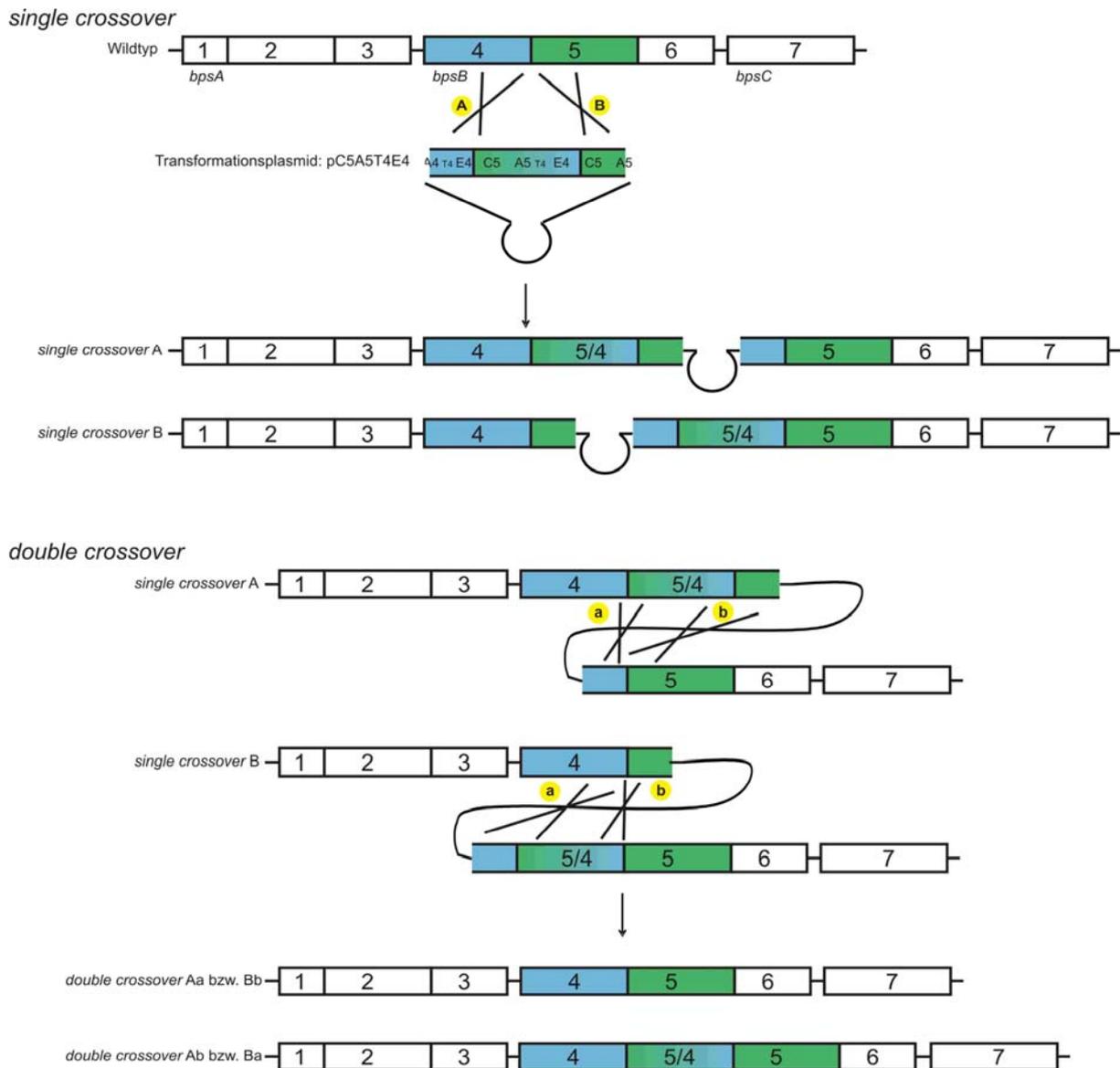
Die Durchführung der direkten Transformation erfolgte wie in Abschnitt 4.2.7.2 beschrieben und erbrachte aus 25 mL Kultur 149 Erythromycin-resistente Klone. Das dafür eingesetzte Transformationsplasmid pC5A5T4E4 (10x26 µl) lag hierbei in einer Konzentration von 193 ng/µl vor. Neben dem Wachsen auf Erythromycin-haltigen Platten sollte ein weiterer Beweis für die Integration des kompletten Transformationsplasmids pC5A5T4E4 über homologe Rekombination in das Genom von *A. balhimycina* erbracht werden (*single crossover*-Stadium). Hierfür wurde eine Kontroll-PCR gewählt, die ein 561 bp-Fragment aus dem *ermE*-Gen amplifiziert. Dieses Gen ist für die Resistenz gegen Erythromycin verantwortlich, da es für eine Methyltransferase kodiert, die die 23S rRNA modifiziert. Die Kontroll-PCR wurde der Einfachheit halber als Kolonie-PCR durchgeführt und erbrachte bei den zehn getesteten Klonen durchweg positive Ergebnisse (s. Abbildung 5-6). Als Positivkontrolle wurde der Vektor pSP1 verwendet und als Negativkontrolle DNA des *A. balhimycina* Wildtyps.



**Abbildung 5-6: Kontroll-Kolonie-PCR zum Nachweis des *single crossover*-Stadiums über Amplifikation eines 561 bp-Fragments aus dem *ermE*-Gen. Die 10 getesteten Klone waren alle positiv. Als Positivkontrolle wurde der Vektor pSP1 verwendet und als Negativkontrolle DNA des *A. balhimycina* Wildtyps.**

#### 5.1.2.4 Durchführung eines Stressprotokolls zur Forcierung von *double crossover*-Ereignissen in *A. balhimycina*

Um die gewünschte Modulinsertion im Genom von *A. balhimycina* zu erhalten, musste ausgehend vom *single crossover*-Stadium, bei dem das komplette Transformationsplasmid pC5A5T4E4 im Genom integriert vorliegt, eine zweite homologe Rekombination erfolgen. Dieser sog. *double crossover* kann über einen der beiden flankierenden Bereiche der gewünschten Modulinsertionsstelle erfolgen, die sowohl im Genom als auch in pC5A5T4E4 vorhanden sind. Erfolgt die zweite homologe Rekombination (*double crossover*) über denselben flankierenden Bereich wie bei der ersten homologen Rekombination (*single crossover*), so liegt wieder der Genotyp des Wildtyps vor. Nur wenn die beiden homologen Rekombinationsereignisse über einen jeweils anderen flankierenden Bereich stattfinden, wird der gewünschte Genotyp mit inseriertem Modul erhalten. Um letztere Variante so wahrscheinlich wie möglich zu gestalten, wurde bei der Klonierungsstrategie darauf geachtet, dass die beiden flankierenden Bereiche etwa dieselbe Größe besitzen. Eine graphische Darstellung dieser Sachverhalte enthält Abbildung 5-7.



**Abbildung 5-7:** Um die gewünschte Modulinsertion im Genom von *A. balhimycina* zu erhalten, müssen ausgehend vom Wildtyp zwei homologe Rekombinationen mit dem Transformationsplasmid pC5A5T4E4 stattfinden und zwar über einen jeweils anderen flankierenden Bereich der Insertionsstelle. Findet die erste homologe Rekombination (*single crossover*) über den vorderen flankierenden Bereich statt (A), muss die zweite homologe Rekombination (*double crossover*) über den hinteren flankierenden Bereich stattfinden (b) oder umgekehrt (Ba). Zwei homologe Rekombinationen nach dem Schema Aa oder Bb führen wieder zum Wildtyp-Genom.

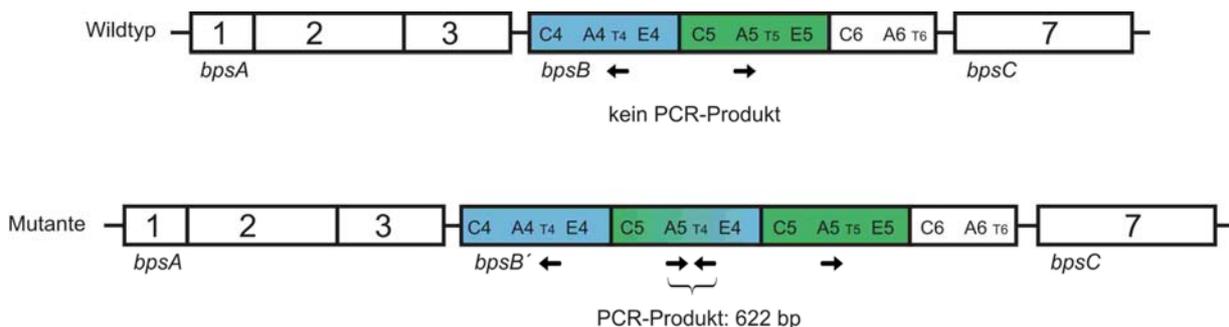
Um das in *A. balhimycina* relativ seltene Ereignis der homologen Rekombination zu forcieren, wurde das in Abschnitt 4.2.8 beschriebene Stressprotokoll angewendet. Von den aus 25 mL Kultur erhaltenen 187 Klonen identifizierten sich zwölf durch ihre wiedererlangte Erythromycin-Sensitivität als *double crossover*-Klone. Diese zwölf *double crossover*-Klone wurden mittels PCR und *Southern Blot*-Hybridisierung weiter untersucht, wobei sich zeigte, dass fünf Klone im Genom die gewünschte

Modulinsertion enthielten und somit den Genotyp der *A. balhimycina* C5A5T4E4-Mutante aufwiesen. Die restlichen sieben Klone waren zum Genom des *A. balhimycina* Wildtyps revertiert. Die ausführliche Beschreibung der molekularbiologischen Charakterisierung der *double crossover*-Klone und damit insbesondere der *A. balhimycina* C5A5T4E4-Mutante mittels PCR und *Southern Blot*-Hybridisierung erfolgt in Abschnitt 5.2.

## 5.2 Molekularbiologische Charakterisierung der *A. balhimycina* C5A5T4E4-Mutante mittels PCR und *Southern Blot*-Hybridisierung

Um aus den zwölf Erythromycin-sensitiven *double crossover*-Klonen diejenigen herauszufinden, bei denen die zweite homologe Rekombination zum Genotyp der gewünschten *A. balhimycina* C5A5T4E4-Mutante geführt hatte, wurden in der vorliegenden Arbeit die beiden molekularbiologischen Methoden PCR und *Southern Blot*-Hybridisierung verwendet. Die dafür notwendige Gesamt-DNA wurde wie in Abschnitt 4.2.10 beschrieben isoliert.

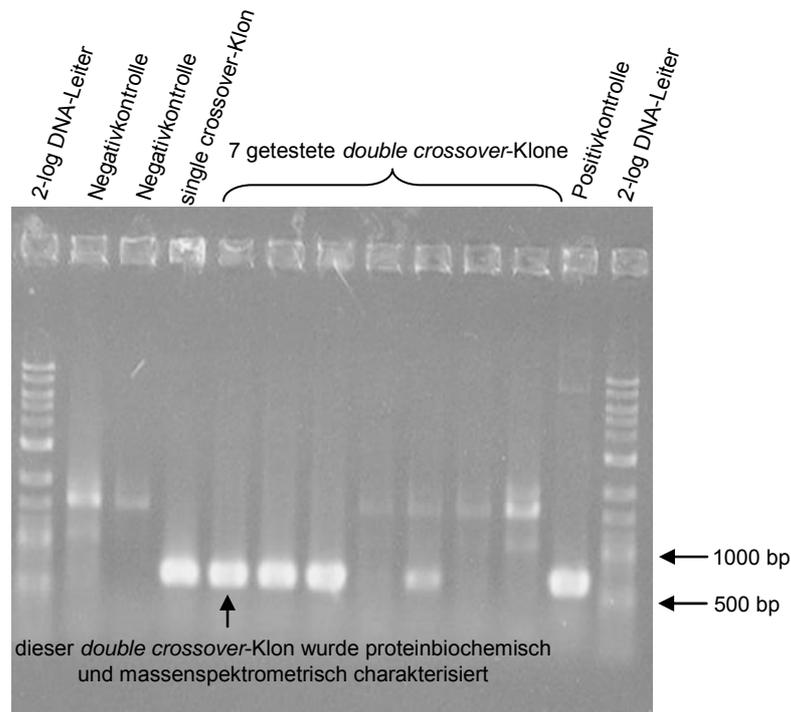
Der Nachweis der Modulinsertion auf genomischer Ebene mittels PCR gestaltete sich insofern schwierig, als dass die neu eingeführten DNA-Sequenzen bereits im Genom des *A. balhimycina* Wildtyps vorhanden waren. Die einzige Möglichkeit der Unterscheidung zwischen Wildtyp und Mutante ergab sich demnach aus dem artifizialen Übergang zwischen den Domänen A5 und T4 des eingeführten C5A5T4E4-Hybridmoduls in der Mutante. Für die Kontroll-PCR „Modulinserion“ wurde deshalb ein Vorwärtsprimer abgeleitet, der in der Domäne A5 bindet und ein Rückwärtsprimer mit Bindung an die Domäne T4. Diese Primer binden sowohl an die DNA der Mutante als auch an die DNA des Wildtyps, allerdings ergibt sich nur im Fall der Mutante ein PCR-Produkt (622 bp). Eine graphische Darstellung dieses Sachverhalts zeigt Abbildung 5-8.



**Abbildung 5-8:** Die Primer der Kontroll-PCR „Modulinserion“ wurden so gewählt, dass der Vorwärtsprimer in der Domäne A5 und der Rückwärtsprimer in der Domäne T4 bindet. Diese Primer binden sowohl an die DNA der Mutante als auch an die DNA des Wildtyps, allerdings ergibt sich nur im Fall der Mutante ein PCR-Produkt (622 bp).

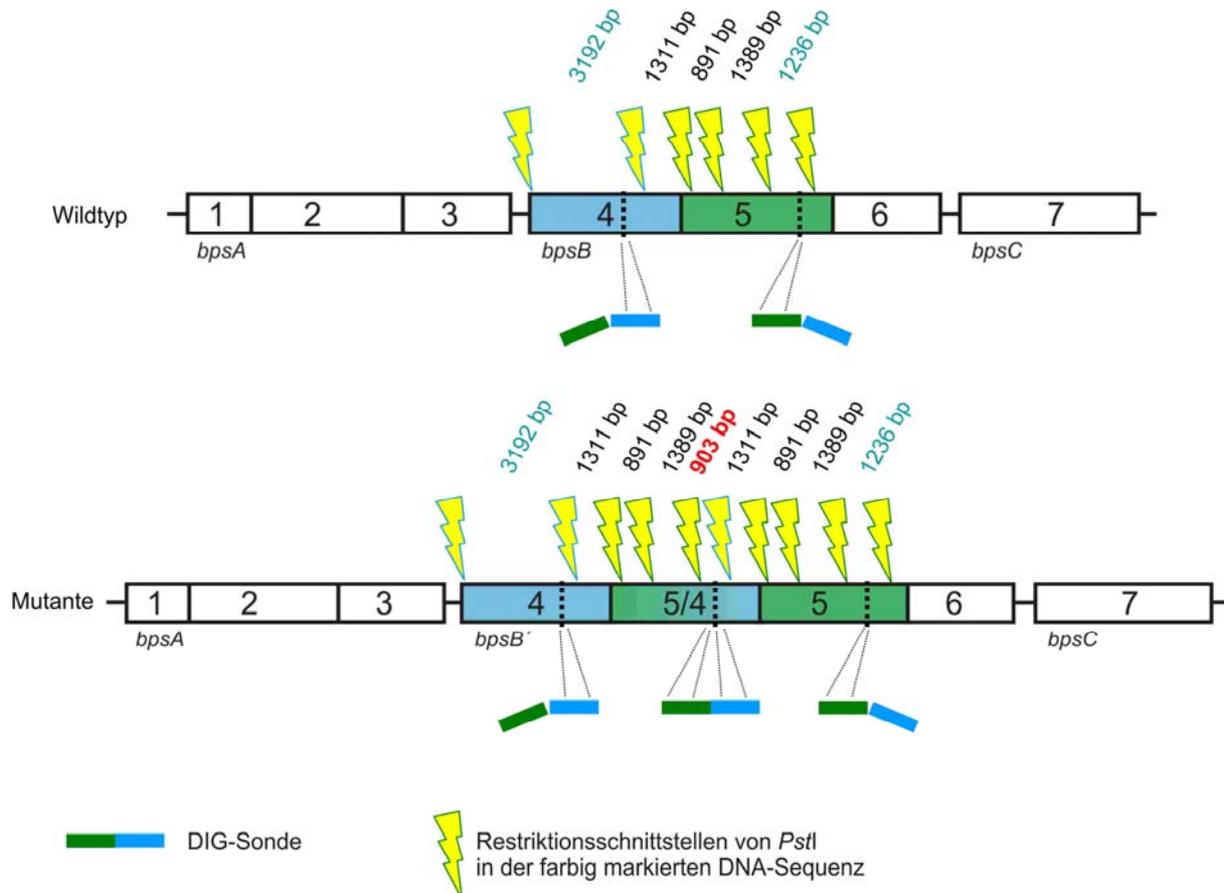
Die Durchführung der Kontroll-PCR „Modulinserion“ und deren Sichtbarmachung mittels Gelelektrophorese erbrachte bei fünf der zwölf *double crossover*-Klonen die charakteristische Bande bei 622 bp. Das Gel in Abbildung 5-9 zeigt exemplarisch die Untersuchung von sieben der zwölf *double crossover*-Klone. Als Positivkontrolle

wurde das Transformationsplasmid pC5A5T4E4 verwendet und als Negativkontrolle DNA des *A. balhimyina* Wildtyps.



**Abbildung 5-9: Gelelektrophoretische Auftrennung der Kontroll-PCR „Modulininsertion“.** Hier ist exemplarisch die Untersuchung von sieben der zwölf *double crossover*-Klonen dargestellt, wobei vier die charakteristische 622 bp-Bande aufwiesen. Als Positivkontrolle wurde das Transformationsplasmid pC5A5T4E4 verwendet und als Negativkontrolle DNA des *A. balhimyina* Wildtyps.

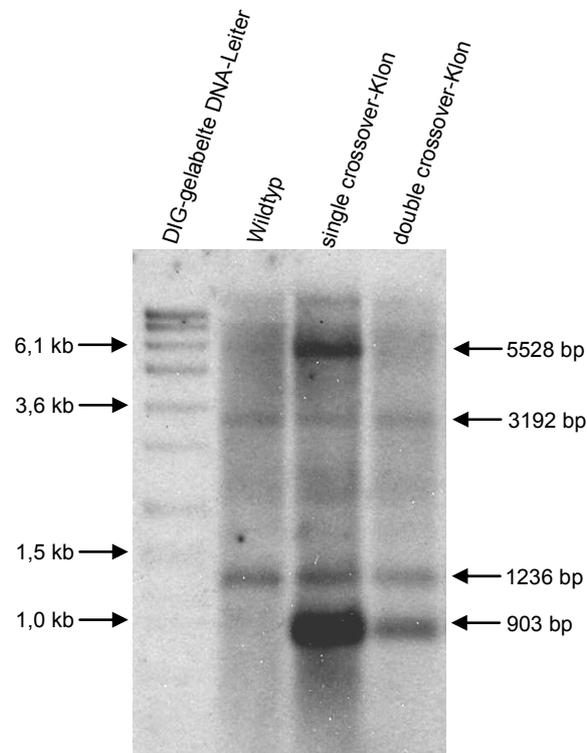
Neben der Kontroll-PCR wurde außerdem eine *Southern Blot*-Hybridisierung durchgeführt, um die Modulinsertion auf genomischer Ebene nachzuweisen. Das Prinzip sowie die praktische Durchführung sind in Abschnitt 4.2.11 beschrieben. Bei der Auswahl einer geeigneten Sonde trat das gleiche Problem auf wie bei der Kontroll-PCR, nämlich dass die neu eingeführten DNA-Sequenzen bereits im Genom des *A. balhimycina* Wildtyps vorhanden waren. So gab es also keine Möglichkeit für eine spezifische Sonde, die nur in der Mutante und nicht im Wildtyp hybridisiert. Trotzdem konnte eine Strategie ausgearbeitet werden, die unter Verwendung des Restriktionsenzym *Pst*I einen Nachweis der Modulinsertion über *Southern Blot*-Hybridisierung zuließ. Die Visualisierung der folgenden Erläuterungen ist in Abbildung 5-10 gezeigt.



**Abbildung 5-10: Visualisierung der Strategie zum Nachweis der Modulinsertion mittels Southern Blot-Hybridisierung.**

Die gewählte DNA-Sonde wurde über eine PCR-Reaktion hergestellt und bestand aus einer 331 bp-Sequenz über den artifiziellen DNA-Sequenzübergang zwischen den Domänen A5 und T4. Wie Abbildung 5-10 zeigt, hybridisierte diese Sonde nicht nur im eingeführten C5A5T4E4-Hybridmodul in der Mutante, sondern auch mit ihrem vorderen bzw. hinteren Teil in den Domänen A5 und T4 allein. Somit ergaben sich in der DNA des Wildtyps zwei Hybridisierungsstellen für die Sonde und in der *A. balhimycina* C5A5T4E4-Mutante drei. Durch den Restriktionsverdau mit *PstI* lagen diese Hybridisierungsstellen auf unterschiedlichen DNA-Fragmenten, deren Größen aufgrund der bekannten Sequenz des Balhimycin-Biosynthesegencluster bestimmt werden konnten. Die beiden Hybridisierungsstellen, die gleichermaßen in Wildtyp und Mutante vorkamen, lagen auf DNA-Fragmenten mit den Größen 3192 bp und 1236 bp. Die Hybridisierung der kompletten Sonde mit der DNA der Mutante fand auf einem Fragment der Größe 903 bp statt, so dass dieses Fragment als charakteristisch für die Unterscheidung zwischen Wildtyp und Mutante erachtet werden konnte. Der in Abbildung 5-11 gezeigte Southern Blot spiegelte diese

erwarteten Fragmente wieder und bestätigte somit die Integrität der *A. balhimycina* C5A5T4E4-Mutante auf genomischer Ebene. Der ebenfalls dargestellte *single crossover*-Klon zeigte eine zusätzliche Bande bei 5528 bp, da in diesem Stadium noch das komplette Transformationsplasmid pC5A5T4E4 im Genom integriert vorlag.



**Abbildung 5-11: Southern Blot-Ergebnisse des *A. balhimycina* Wildtyps sowie beider *crossover*-Stadien zur *A. balhimycina* C5A5T4E4-Mutante. Während der Wildtyp und beide *crossover*-Stadien DNA-Fragmente der Größe 3192 bp und 1236 bp zeigten, fehlte dem Wildtyp das für die *crossover*-Stadien charakteristische Fragment der Größe 903 bp. Beim *single crossover*-Klon trat außerdem ein Fragment der Größe 5528 bp auf, da in diesem Stadium noch das komplette Transformationsplasmid pC5A5T4E4 im Genom integriert vorlag. Der hier gezeigte *double crossover*-Klon wurde anschließend proteinbiochemisch und massenspektrometrisch charakterisiert.**

### 5.3 Proteinbiochemische Charakterisierung der *A. balhimycina* C5A5T4E4-Mutante mittels Disk-SDS-PAGE und massenspektrometrischer Analyse

Nachdem die Modulinsertion der *A. balhimycina* C5A5T4E4-Mutante auf genetischer Ebene eindeutig mittels PCR und *Southern Blot*-Hybridisierung nachgewiesen war, sollte anschließend eine proteinbiochemische Charakterisierung erfolgen. Während die beiden nicht-ribosomalen Peptidsynthetasen BpsA und BpsC mit 334 kDa und 199 kDa Größe nicht von den Veränderungen betroffen waren, sollte BpsB mit einer Größe von 432 kDa im Wildtyp durch die Modulinsertion (BpsB') auf eine theoretische Größe von 591 kDa angewachsen sein. Ziel der proteinbiochemischen Charakterisierung der *A. balhimycina* C5A5T4E4-Mutante war die Gegenüberstellung mit dem Wildtyp-Protein, vor allem die Darstellung der Größenunterschiede von BpsB und BpsB'. Aus diesem Grund wurde zunächst eine Isolierung der nicht-ribosomalen Peptidsynthetasen aus *A. balhimycina* Wildtyp und *A. balhimycina* C5A5T4E4-Mutante nach dem in Abschnitt 4.2.12 beschriebenen Protokoll vorgenommen. Dieses Protokoll wurde im Zuge einiger, anfänglich nicht erfolgreicher Isolierungsversuche optimiert, wobei sich vor allem folgende Punkte als essentiell für das Gelingen erwiesen haben:

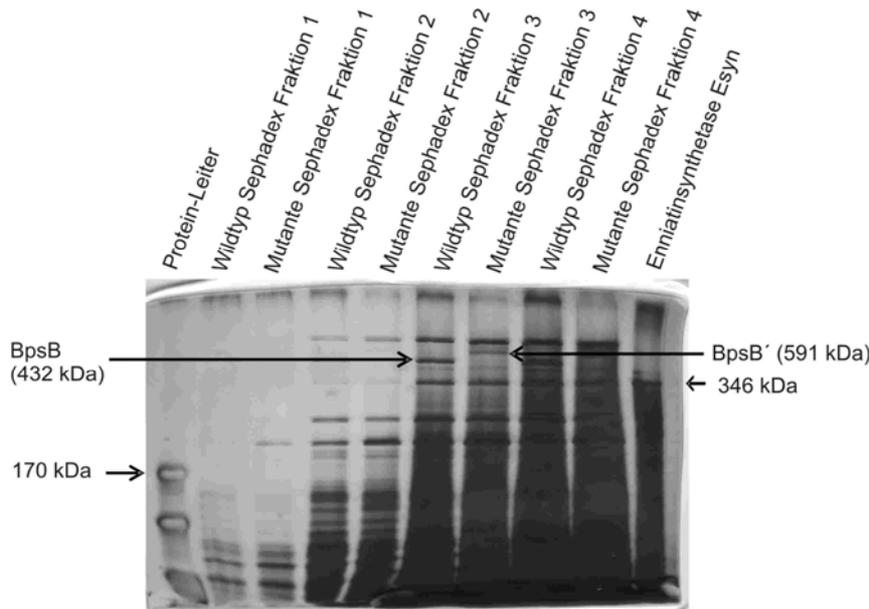
- Arbeit mit kleinen, handhabbaren Zellmengen (4 g Nassgewicht)
- ausreichende Länge des *DNase*I-Verdau
- Polymin P Fällung
- Durchführung des Isolierungsprotokolls bis zur Größenausschlusschromatographie an einem Tag

Die Dokumentation des Reinigungsverlaufs mittels Bradford-Protein-Konzentrationsbestimmung (s. Abschnitt 4.2.13) lieferte die in Tabelle 8 aufgeführten Ergebnisse. Zu erwähnen ist hierbei, dass die Überstandsproben immer nach dem entsprechenden Zentrifugationsschritt entnommen wurden und dass alle Proben vor der Konzentrationsbestimmung um den Faktor 1:200 mit Proteinisolierungspuffer verdünnt wurden, um in den linearen Bereich des Konzentrations-Absorptions-Diagramms zu kommen. Nach Errechnen der absoluten Konzentrationswerte aus den gemessenen Absorptionsmittelwerten über die in Abschnitt 4.2.13 bestimmte Kalibriergerade von Albumin wurde diese Verdünnung wieder zurückgerechnet.

**Tabelle 8: Dokumentation des Reinigungsverlaufs der nicht-ribosomalen Peptidsynthetasen mittels Bradford-Protein-Konzentrationsbestimmung. Grundlage der absoluten Konzentrationswerte ist die Kalibriergerade von Albumin (s. Abschnitt 4.2.13).**

Probe (1:200)	Wildtyp			Mutante		
	Dreifachbest.	Mittelwert	Proteinkonz. [µg/mL]	Dreifachbest.	Mittelwert	Proteinkonz. [µg/mL]
<i>DNase</i> Überstand	0,075 0,072 0,102	0,083	8737	0,084 0,088 0,093	0,088	9298
Polymix P Überstand	0,030 0,030 0,031	0,030	3193	0,045 0,049 0,048	0,047	4982
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> Überstand	0,003 -0,006 0,011	0,003	281	0,001 0,007 0,012	0,007	737
vor Sephadex	0,218 0,238 0,245	0,234	24596	0,176 0,175 0,172	0,174	18351
Sephadex Totvolumen	-0,014 -0,016 0,014	-0,005	-	0,014 0,094 0,087	0,065	6842
Sephadex Fraktion 1	-0,006 -0,003 0,018	0,003	316	-0,016 0,001 0,020	0,002	175
Sephadex Fraktion 2	-0,007 -0,009 -0,001	-0,006	-	0,022 0,048 0,045	0,038	4035
Sephadex Fraktion 3	0,001 0,003 0,035	0,013	1368	0,072 0,073 0,081	0,075	7930
Sephadex Fraktion 4	0,061 0,062 0,077	0,067	7018	0,041 0,044 0,044	0,043	4526
Sephadex Fraktion 5	0,110 0,119 0,159	0,129	13614	0,131 0,128 0,148	0,136	14281
Sephadex Fraktion 6	0,073 0,092 0,092	0,086	9018	0,091 0,092 0,092	0,092	9649
Sephadex Fraktion 7	0,062 0,060 0,080	0,067	7088	0,056 0,055 0,080	0,064	6702

Die über diese Isolierungsstrategie generierten Sephadex-Fraktionen wurden wie in Abschnitt 4.2.14 und 4.2.15 beschrieben gelelektrophoretisch aufgetrennt und mit Silberfärbung angefärbt (s. Abbildung 5-12).



**Abbildung 5-12: Silbergefärbtes Disk-SDS-PAGE-Gel der Sephadex-Fractionen von *A. balhimycina* Wildtyp und *A. balhimycina* C5A5T4E4-Mutante. Als Größenmarker wurden die größte käuflich erhältliche Protein-Leiter sowie die Enniatinsynthetase (ESyn) mit 346 kDa verwendet.**

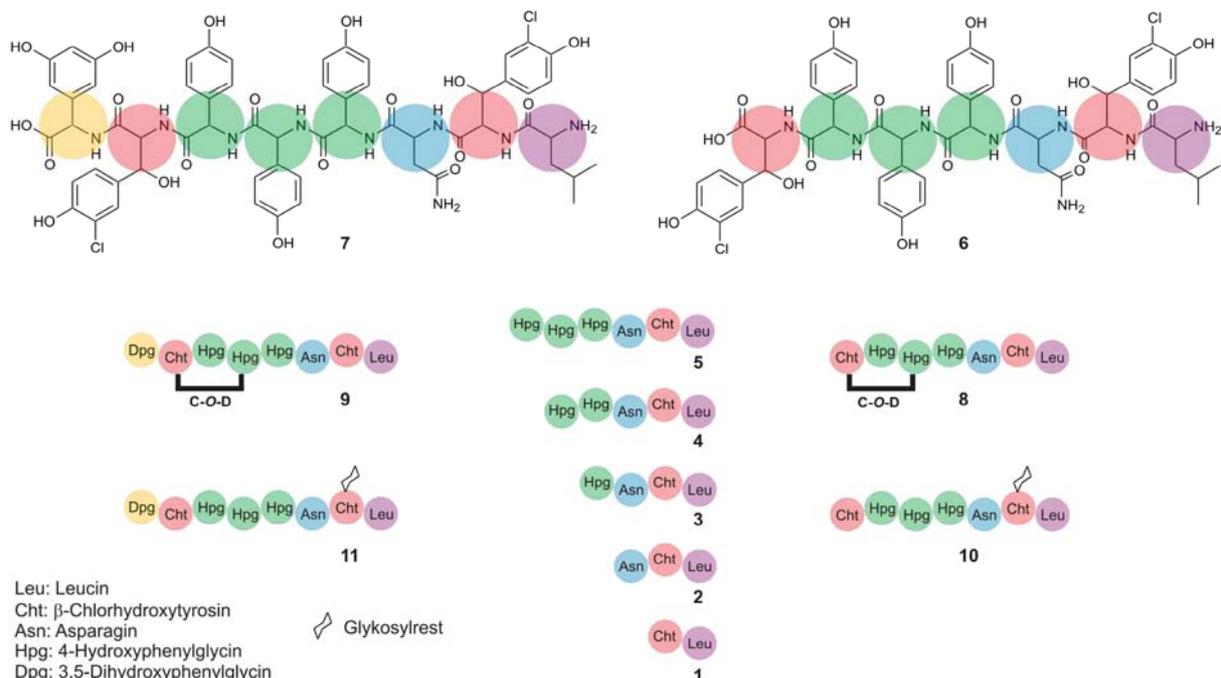
Vergleicht man die beiden Spuren der Sephadex-Fractionen 3 von Wildtyp und Mutante, so beobachtet man jeweils eine Bande, die in der anderen Spur nicht auftritt. Berücksichtigt man die Protein-Banden bekannter Größe in der Protein-Leiter-Spur bei 170 kDa und in der Enniatinsynthetase-Spur bei 346 kDa, so war es sehr wahrscheinlich, dass es sich bei der Bande in der Wildtyp-Spur um BpsB mit 432 kDa und bei der schwächeren Bande in der Mutanten-Spur um das modulerweiterte BpsB' bei 591 kDa handelte. Um diese Vermutung zu bestätigen, wurde eine massenspektrometrische Analyse dieser beiden Protein-Banden bei der Firma Proteome Factory AG in Auftrag gegeben. Da die erste Analyse des in Abbildung 5-12 gezeigten Gels aufgrund zu geringer Konzentrationen zu keinem Ergebnis führte, wurden die Reste von Wildtyp Sephadex-Fraktion 3 und Mutanten Sephadex-Fraktion 3 in mehreren Spuren auf jeweils einem Gel gelelektrophoretisch getrennt. Die Anfärbung dieser Gele übernahm die Proteome Factory AG mit einem firmeneigenen Kit, welches dem in der vorliegenden Arbeit verwendeten Kit hinsichtlich MS-Kompatibilität überlegen sein soll. Die damit durchgeführte massenspektrometrische Analyse und Proteindatenbanksuche erbrachte für beide Banden eine eindeutige Zuordnung zur nicht-ribosomalen Peptidsynthetase BpsB und bestätigte somit die oben angestellte Vermutung (s. Anhang). Die Tatsache, dass auch BpsB' der Mutante zu BpsB in der Datenbank zugeordnet wurde, erklärt

sich dadurch, dass die Datenbank BpsB' nicht enthält und BpsB' nur aus Sequenzen von BpsB besteht.

Die beiden anderen nicht-ribosomalen Peptidsynthetasen BpsA und BpsC lagen aufgrund ihrer geringeren Größe in einem Bereich des Gels, in dem die Spuren eine Vielzahl an Proteinkontamination zeigten. Da der Fokus der proteinbiochemischen Charakterisierung eindeutig auf BpsB und BpsB' lag, wurden auch keine weiteren Aufreinigungsversuche unternommen, um auch noch BpsA und BpsC identifizieren zu können.

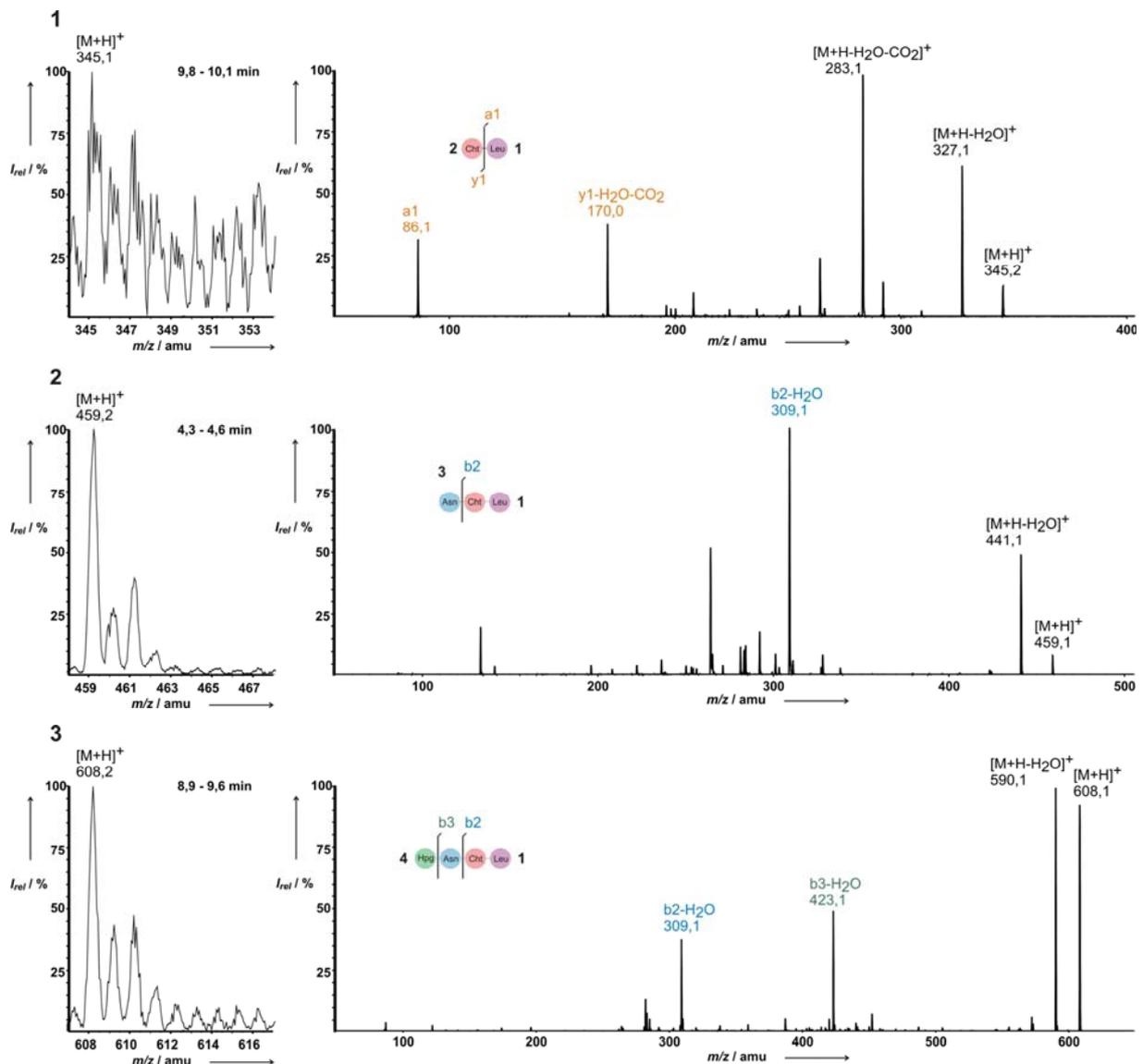
#### 5.4 Charakterisierung des Glykopeptidproduktionsprofils der *A. balhimycina* C5A5T4E4-Mutante mittels HPLC-ESI-MS und -MS/MS

Für die massenspektrometrische Charakterisierung des Glykopeptidproduktionsprofils der *A. balhimycina* C5A5T4E4-Mutante wurde mehrere Male wie in Abschnitt 4.2.2 beschrieben fermentiert. Für die Vorbereitung der Fermentationskulturen wurden unterschiedliche Prozeduren getestet, wobei sich die in Abschnitt 4.2.17 beschriebene Vorgehensweise am Ende als die beste erwiesen hat. Die massenspektrometrischen Untersuchungen umfassten zunächst eine Identifizierung von potentiellen Glykopeptidmetaboliten mittels HPLC-ESI-MS (s. Abschnitt 4.2.18) und eine anschließende Verifizierung des Glykopeptidcharakters und gleichzeitige Strukturaufklärung mittels HPLC-ESI-MS/MS (s. Abschnitt 4.2.19). Auf diesem Weg konnte aus einer enormen Datenmenge letztlich das Glykopeptidproduktionsprofil der *A. balhimycina* C5A5T4E4-Mutante aufgeklärt werden. Dieses zeigte erfreulicherweise einige Glykopeptidmetabolite mit einem um ein Hpg erweiterten Peptidgrundgerüst, was den erfolgreichen Verlauf der kombinatorischen Biosynthese auf Produktions- und damit Enzymaktivitätsebene bestätigte. Neben den erwarteten Oktapeptiden traten auch C-terminal verkürzte Metabolite auf, bei denen es sich höchstwahrscheinlich um Biosynthesestufen handelte. Eine Übersicht über die strukturaufgeklärten Glykopeptidmetabolite 1 - 11 ist in Abbildung 5-13 gezeigt.



**Abbildung 5-13: Übersicht über die strukturaufgeklärten Glykopeptidmetabolite 1 - 11. Als Symbolik wurden hier für gleiche Aminosäuren Kugeln gleicher Farbe gewählt. (Anmerkung: N-Terminus ist gegen die offizielle Definition rechts gesetzt).**

Im Folgenden soll die Strukturaufklärung der Metabolite **1** - **11** näher beleuchtet und jeweils das MS- und MS/MS-Spektrum gezeigt werden. Am eindeutigsten war die Strukturaufklärung der linearen, vollständig chlorierten, aber nicht-glykosylierten und nicht *N*-methylierten Di- bis Oktapeptide **1** - **7**, da mit Hilfe der Roepstorff-Fragmente die Aminosäure-Sequenz genau nachvollzogen werden konnte. Als Sekundärfragmentierungen traten neben den Peptidbindungsbrüchen häufig  $\text{H}_2\text{O}$ -Abspaltungen, die typisch für  $\beta$ -Hydroxysäuren sind, und  $\text{CO}_2$ -Abspaltungen, die typisch für den C-Terminus sind, auf. Die zugehörigen Spektren dieser sieben Metabolite sind in Abbildung 5-14 gezeigt.



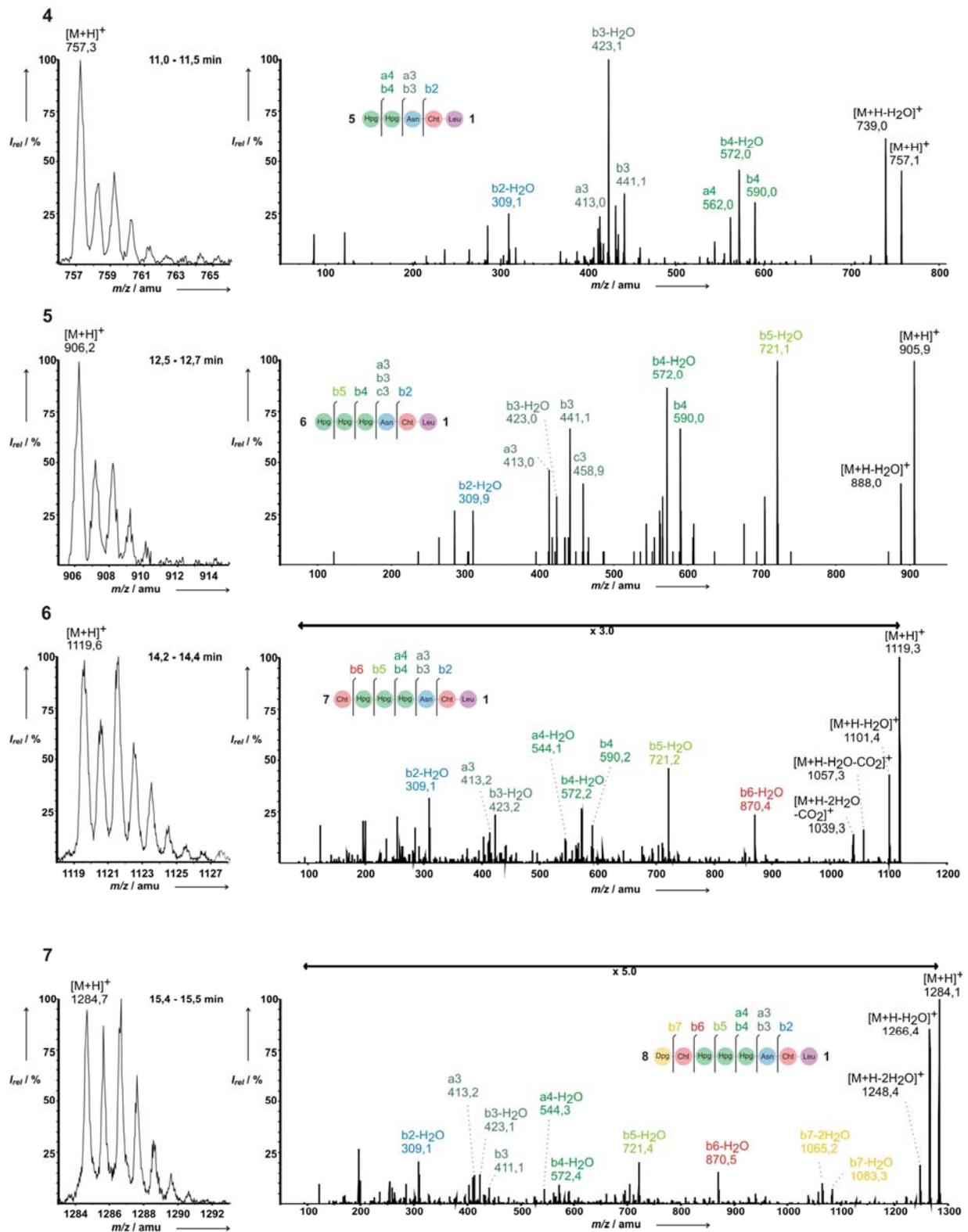
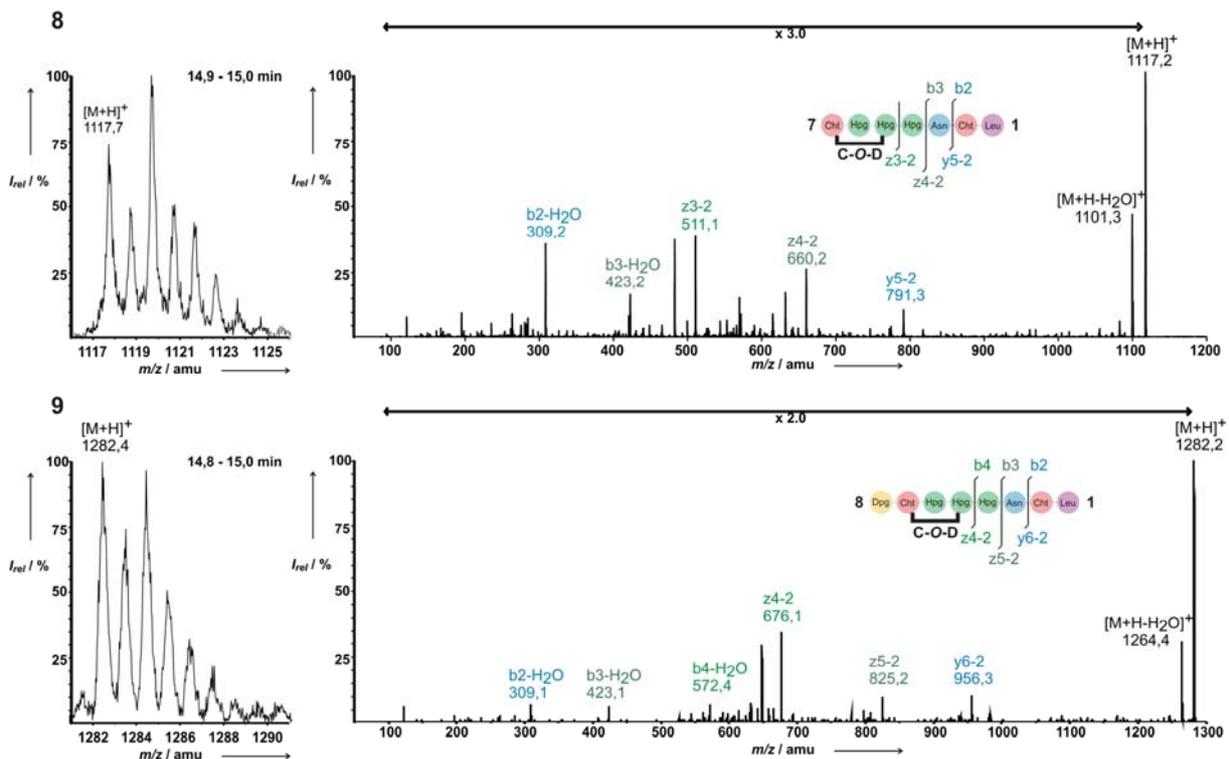


Abbildung 5-14: MS- und MS/MS-Spektren der sieben linearen, vollständig chlorierten, aber nicht glykosylierten und nicht *N*-methylierten Di- bis Oktapeptide 1 – 7. (Anmerkung: *N*-Terminus ist gegen die offizielle Definition rechts gesetzt).

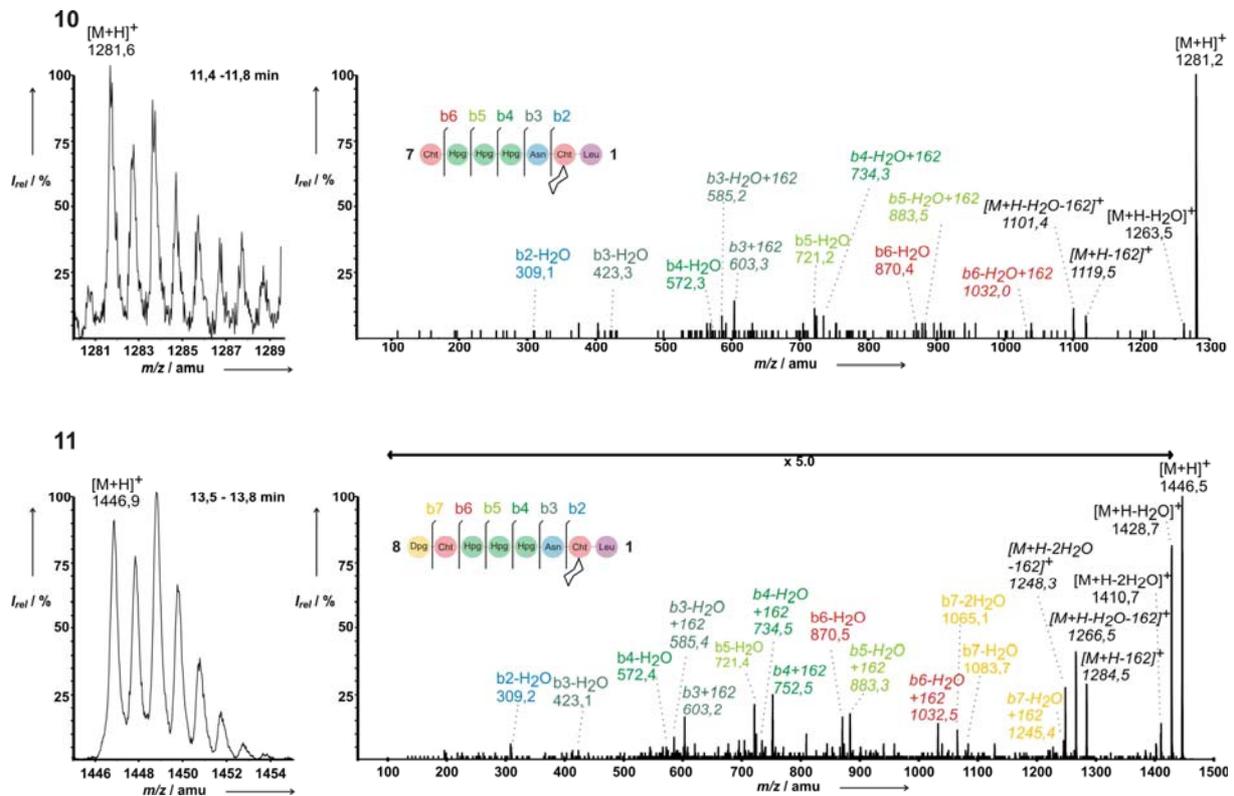
Neben diesen Metaboliten traten die mit **6** und **7** verwandten Metabolite **8** und **9** auf, die jeweils einen Massenshift von  $\Delta m = 2$  amu zu niedrigeren Molekülmassen aufwiesen, was stark auf das Vorkommen einer oxidativen Seitenkettencyklisierung hindeutete. Mit Hilfe der in Abbildung 5-15 gezeigten MS/MS-Spektren konnte in beiden Fällen die oxidative Seitenkettencyklisierung eindeutig zwischen  $^5\text{Hpg}$  und  $^7\text{Cht}$  lokalisiert werden.



**Abbildung 5-15: MS- und MS/MS-Spektren der beiden monozyklischen, vollständig chlorierten, aber nicht glykosylierten und nicht N-methylierten Hepta- und Oktapeptide **8** und **9**. (Anmerkung: N-Terminus ist gegen die offizielle Definition rechts gesetzt).**

Zwei weitere Metabolite **10** und **11** zeigten einen Massenshift von  $\Delta m = 162$  amu zu höheren Molekülmassen im Vergleich zu **6** und **7**, was typischerweise einer Glykosylierung entspricht. Mit Hilfe der in Abbildung 5-16 gezeigten MS/MS-Spektren konnten in beiden Fällen Hinweise über den Ort der Glykosylierung gewonnen werden. Die Roepstorff-Fragmente  $[\text{b}_6-\text{H}_2\text{O}+162]$  bis  $[\text{b}_3-\text{H}_2\text{O}+162]$  bei **10** bzw.  $[\text{b}_7-\text{H}_2\text{O}+162]$  bis  $[\text{b}_3-\text{H}_2\text{O}+162]$  bei **11** schränkten die Glykosylierungsstelle auf die drei N-terminalen Aminosäuren  $^1\text{Leu}$ ,  $^2\beta\text{-Hty}$  und  $^3\text{Asn}$  ein. Mit Hilfe von MS<sup>3</sup>-Spektren konnte bei **11** außerdem  $^3\text{Asn}$  ausgeschlossen werden. Bei Betrachtung der verbleibenden Aminosäuren  $^1\text{Leu}$  und  $^2\beta\text{-Hty}$  konnten die  $\beta\text{-OH}$ - bzw. die phenolische OH-Gruppe von  $^2\beta\text{-Hty}$  als die wahrscheinlichsten Anknüpfungspunkte für den

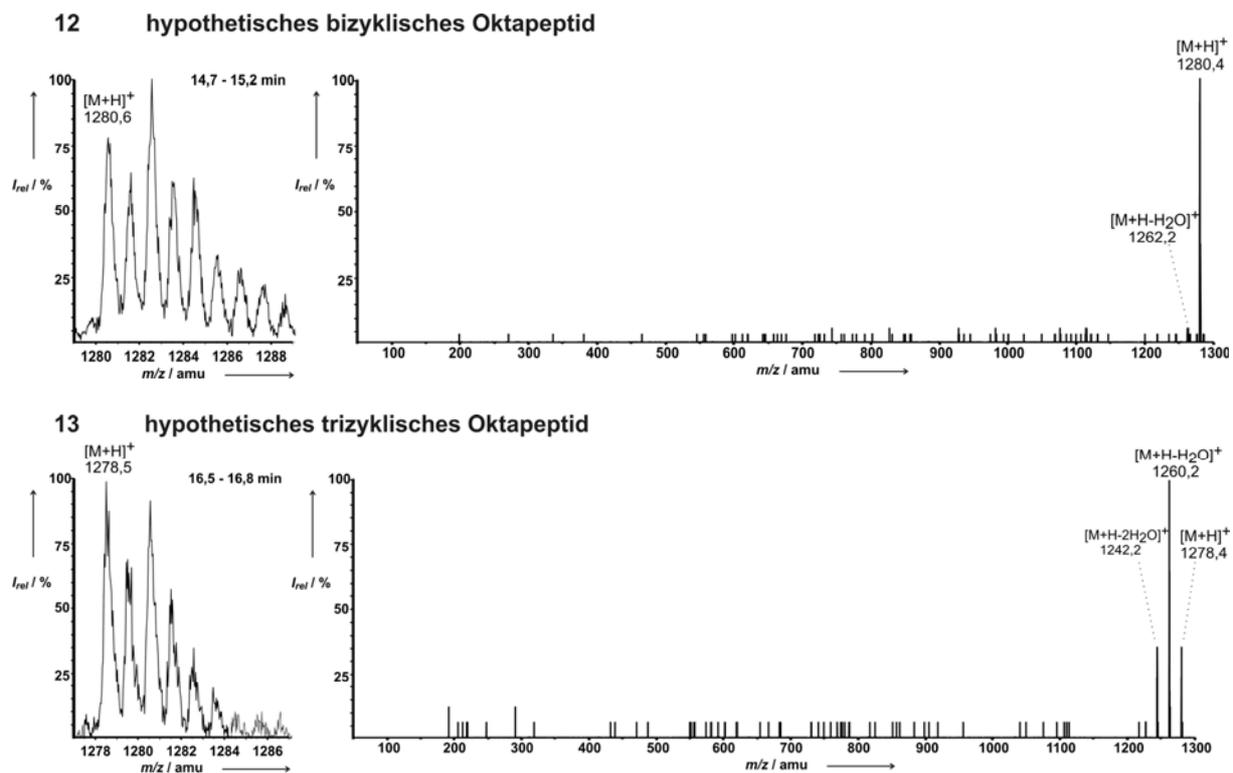
Glykosylrest ausgemacht werden. Eine weitere Charakterisierung der Glykosylierungsstelle war experimentell nicht möglich.



**Abbildung 5-16:** MS- und MS/MS-Spektren der beiden linearen, vollständig chlorierten, monoglykosylierten und nicht *N*-methylierten Hepta- und Oktapeptide 10 und 11. (Anmerkung: *N*-Terminus ist gegen die offizielle Definition rechts gesetzt).

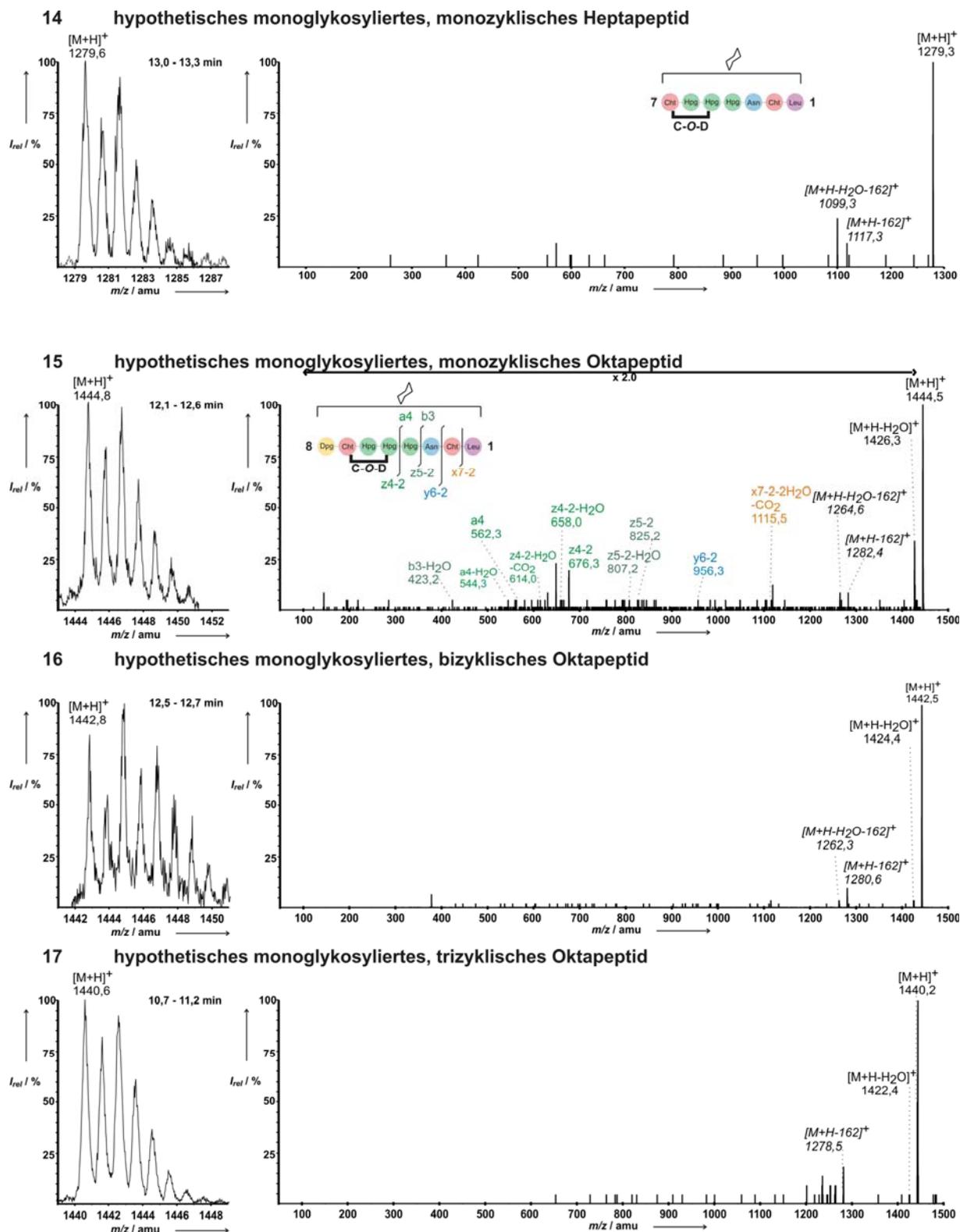
Neben den soeben beschriebenen elf strukturaufgeklärten Glykopeptidmetaboliten **1** – **11** traten sechs weitere Metabolite **12** – **17** auf, deren Strukturen allerdings aufgrund zu geringer Konzentration bzw. aufgrund zu starker Verbrückung in den Seitenketten nicht mittels HPLC-ESI-MS/MS aufgeklärt werden konnten.

Dies sind die Metabolite **12** und **13**, bei denen es sich aufgrund ihrer Molekülmasse um das bicyclische und das tricyclische, vollständig chlorierte, aber nicht glykosylierte und nicht *N*-methylierte Oktapeptid handeln könnte. Die MS/MS-Spektren sind in beiden Fällen sehr fragmentarm, was gut mit dem hohen Grad der oxidativen Seitenkettenzyklisierung zusammenpassen würde. Leider waren die Konzentrationen für eine Aufreinigung und anschließende NMR-Analyse dieser Verbindungen zu gering, was die einzige Möglichkeit zur Aufklärung dieser wahrscheinlich mehrfach verbrückten Strukturen gewesen wäre. Die MS- und MS/MS-Spektren zu diesen beiden Metaboliten sind in Abbildung 5-17 gezeigt.



**Abbildung 5-17: MS- und MS/MS-Spektren der beiden hypothetischen bi- bzw. trizyklischen, vollständig chlorierten, aber nicht glykosylierten und nicht *N*-methylierten Oktapeptide 12 und 13.**

Neben den bereits beschriebenen linearen monoglykosylierten Hepta- bzw. Oktapeptiden **10** und **11** wurden auch vier Metabolite **14** - **17** gefunden, deren Molekülmassen zu den monoglykosylierten Derivaten des monozyklischen Heptapeptids **8** bzw. der mono- bis trizyklischen Oktapeptide **9**, **12** und **13** passen würden. Während die MS/MS-Spektren von **14**, **16** und **17** sehr fragmentarm waren und nur die Abspaltung des Glykosylrests zeigten, gab das MS/MS-Spektrum von **15** mehr Auskunft über die Struktur. Allerdings konnte auch hier nichts über die Lokalisation des Glykosylrests herausgefunden werden. Alle Spektren zu den Metaboliten **14** - **17** sind in Abbildung 5-18 gezeigt.



**Abbildung 5-18: MS- und MS/MS-Spektren des hypothetischen monoglykosylierten, monozyklischen Heptapeptids 14 und der hypothetischen monoglykosylierten, mono- bis trizyklischen Oktapeptide 15 – 17, die jeweils vollständig chloriert, aber nicht *N*-methyliert vorliegen. Die Klammer bei 14 und 15 soll andeuten, dass die Lokalisation des Glykosylrests unbestimmt ist. (Anmerkung: *N*-Terminus ist gegen die offizielle Definition rechts gesetzt).**

Neben diesen 17 Metaboliten produziert die Mutante auch Balhimycin, das Glykopeptid des Wildtyps, sowie die ebenso im Wildtyp vorkommenden typischen mikroheterogenen Derivate. Die möglichen Gründe für diese Tatsache werden in Abschnitt 6 diskutiert.

## 5.5 Ergänzende Untersuchungen zum Glykopeptidproduktionsprofil der *A. balhimycina* C5A5T4E4-Mutante

### 5.5.1 Chirale Aminosäurenanalytik des Glykopeptidmetaboliten 6

Stellvertretend für die anderen Glykopeptidmetabolite wurde **6** mit Hilfe einer chiralen Aminosäurenanalytik untersucht, was eine vorherige Aufreinigung mittels präparativer HPLC zu sehr hoher Reinheit erforderte. Das so erhaltene Chromatogramm ist in Abbildung 5-19 dargestellt.

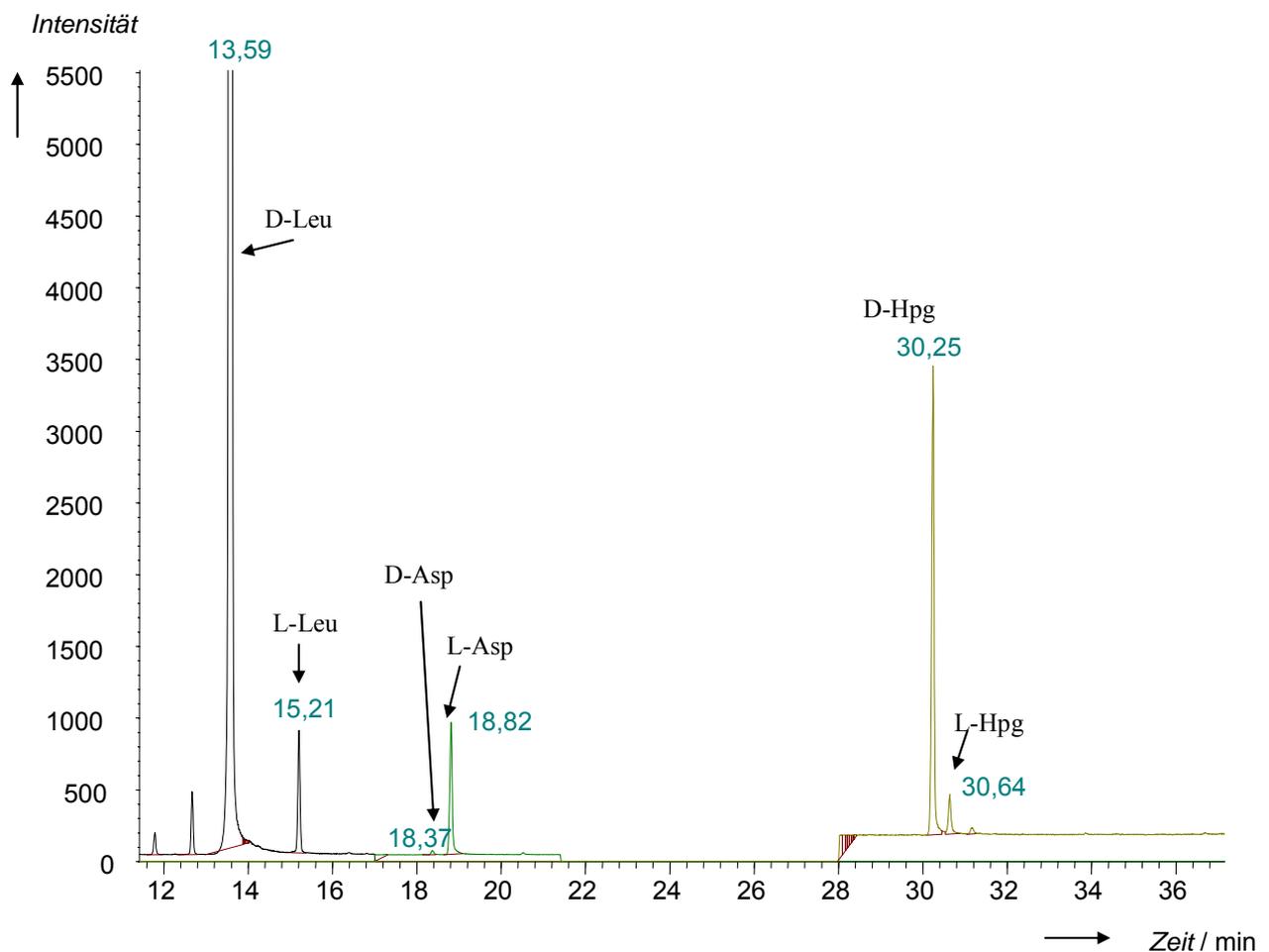


Abbildung 5-19: GC-MS-Chromatogramm zur chiralen Aminosäurenanalytik des Glykopeptidmetaboliten 6.

Die daraus abgeleiteten Ergebnisse sind in Tabelle 9 zusammengefasst.

**Tabelle 9: Ergebnisse der chiralen Aminosäurenanalytik des Glykopeptidmetaboliten 6.**

erwartet	gefunden
<i>D</i> -Leu	<i>D</i> -Leu (98,6 %, ee = 97,2 %)
<i>L</i> -Asn	<i>L</i> -Asp (96,9 %, ee = 93,8 %)
<i>D</i> -Hpg (3x)	<i>D</i> -Hpg (91,2 %, ee = 82,4 %)

Diese Ergebnisse zeigten vor allem, dass im Glykopeptidmetaboliten **6** nur *D*-Hydroxyphenylglycine vorhanden sind und daher die Epimerasedomäne des eingebrachten Moduls funktional ist. Dieses Ergebnis kann mit großer Wahrscheinlichkeit auf die anderen Glykopeptidmetabolite übertragen werden. Bezüglich der anderen Aminosäuren sind die Ergebnisse wie erwartet. *D*-Leu konnte direkt und *L*-Asn indirekt als *L*-Asp nachgewiesen werden. Diese Umwandlung von Asn zu Asp ist typisch und tritt inhärent unter den verwendeten Hydrolysebedingungen auf.  $\beta$ -Hty wird unter diesen Bedingungen sogar ganz zerstört und konnte deshalb nicht nachgewiesen werden.

### 5.5.2 Summenformelbestätigung der Glykopeptidmetabolite 1 – 11 der *A. balhimycina* C5A5T4E4-Mutante mittels hochaufgelöster ESI-FTICR-MS

Die Ergebnisse dieser Messungen sind in Tabelle 10 aufgelistet. Für die meisten Metabolite konnten die Summenformeln, die den in Abschnitt 5.4 gezeigten Strukturen entsprechen, mit einer relativen Massengenauigkeit  $\leq 1,0$  ppm bestätigt werden. Die schlechten relativen Massengenauigkeiten von Metabolit **5** mit 2,6 ppm und Metabolit **11** mit 15,6 ppm erklären sich durch die geringen Konzentrationen dieser Metabolite und die dadurch erzeugten schlechten Intensitäten im Massenspektrum. Bei Metabolit **6** waren die Konzentrationen so gering, dass ohne vorherige HPLC-Kopplung nichts im Massenspektrum detektiert werden konnte.

**Tabelle 10: Ergebnisse der hochaufgelösten ESI-FTICR-MS-Messungen für die Glykopeptidmetabolite 1 – 11.**

Metabolit	Summenformel des Molekül-Ions [M+H] <sup>+</sup>	Theoretisch berechnete Molekülmasse des Molekül-Ions [M+H] <sup>+</sup>	Experimentell bestimmte Molekülmasse des Molekül-Ions [M+H] <sup>+</sup>	relative Massengenauigkeit [ppm]
1	C <sub>15</sub> H <sub>22</sub> ClN <sub>2</sub> O <sub>5</sub>	345,12118	345,12144	0,75
2	C <sub>19</sub> H <sub>28</sub> ClN <sub>4</sub> O <sub>7</sub>	459,16410	459,16445	0,76
3	C <sub>27</sub> H <sub>35</sub> ClN <sub>5</sub> O <sub>9</sub>	608,21178	608,21215	0,61
4	C <sub>35</sub> H <sub>42</sub> ClN <sub>6</sub> O <sub>11</sub>	757,25946	757,25981	0,46
5	C <sub>43</sub> H <sub>49</sub> ClN <sub>7</sub> O <sub>13</sub>	906,30714	906,30946	2,6
6	C <sub>52</sub> H <sub>57</sub> Cl <sub>2</sub> N <sub>8</sub> O <sub>16</sub>	1119,32641	1119,32632	0,08
7	C <sub>60</sub> H <sub>64</sub> Cl <sub>2</sub> N <sub>9</sub> O <sub>19</sub>	1284,36900	1284,36766	1,0
8	C <sub>52</sub> H <sub>55</sub> Cl <sub>2</sub> N <sub>8</sub> O <sub>16</sub>	1117,31076	1117,30965	0,99
9	C <sub>60</sub> H <sub>62</sub> Cl <sub>2</sub> N <sub>9</sub> O <sub>19</sub>	1282,35335	1282,35222	0,88
10	C <sub>58</sub> H <sub>67</sub> Cl <sub>2</sub> N <sub>8</sub> O <sub>21</sub>	1281,37924	n. b.	n. b.
11	C <sub>66</sub> H <sub>74</sub> Cl <sub>2</sub> N <sub>9</sub> O <sub>24</sub>	1446,42183	1446,44445	15,6

### 5.5.3 Abschätzung der relativen Produktionsmengen der Glykopeptide

Die Bestimmung der absoluten Produktionsmengen der Glykopeptide war aufgrund fehlender Standardverbindungen nicht durchführbar, weshalb eine Abschätzung der relativen Produktionsmengen durchgeführt wurde. Um quantitative Aussagen treffen zu können, eignet sich normalerweise ein Q1-Scan im Quadrupol besser als ein EMS-Scan in der linearen Ionenfalle. Allerdings erforderte die Abschätzung der relativen Produktionsmengen aus dem Kulturüberstand, der die Metabolite in ihren ursprünglichen Verhältnissen enthielt, eine Mindestempfindlichkeit, die beim Q1-Scan nicht gegeben war. So blieb also als einzige Möglichkeit die näherungsweise Abschätzung aus den Peakflächen der extrahierten Ionenströme einer HPLC-ESI-MS-Messung mit EMS-Scan. Bei glykosylierten Metaboliten musste die typischerweise auftretende *in source*-Fragmentierung berücksichtigt werden. Diese führt dazu, dass nicht nur der glykosylierte Metabolit, sondern vor allem das daraus entstandene nicht-glykosylierte Fragment im Massenspektrum erscheint. Aus diesem Grund wurden in diesen Fällen die Peakflächen von beiden extrahierten Ionenströmen addiert. Die Ergebnisse dieser Abschätzung zeigten, dass **6** als Hauptmetabolit der *A. balhimycina* C5A5T4E4-Mutante produziert wird und **7** und **8** in fünffach bzw. dreifach niedrigeren Konzentrationen vorkommen. Alle anderen

Glykopeptidmetabolite werden in Konzentrationen <10 % des Hauptmetaboliten **6** produziert. Die Gesamtproduktivität der Mutante konnte auf etwa 10 % der Gesamtproduktivität des Wildtyp-Stammes geschätzt werden. An dieser Stelle muss noch ergänzt werden, dass die Mutante Balhimycin in etwa der gleichen Größenordnung produziert wie Hauptmetabolit **6**.

## 6 Diskussion

Im Ergebnisteil der vorliegenden Arbeit erfolgte eine ausführliche Beschreibung der Herstellungsstrategie der mittels der Prinzipien der kombinatorischen Biosynthese erzeugten *A. balhimycina* C5A5T4E4-Mutante. Die erfolgreiche Umsetzung dieser Strategie wurde zunächst auf genetischer Ebene und anschließend auch auf Proteinebene nachgewiesen. Ob die somit konstruierte, um ein komplettes Modul von 159 kDa erweiterte nicht-ribosomale Peptidsynthetase BpsB' die Biosynthese des hypothetischen Produktes mit einem um ein 4-Hydroxyphenylglycin erweiterten Balhimycin-Peptidrückgrat katalysiert, wurde durch die anschließende massenspektrometrische Charakterisierung des Metabolitspektrums untersucht. Dieses Produktionsprofil wurde in den Abschnitten 5.4 und 5.5 ausführlich beschrieben, wobei von den 17 gefundenen Metaboliten 13 ein Peptidrückgrat mit der zusätzlichen Aminosäure 4-Hydroxyphenylglycin an Position 5 besaßen und somit die Funktionalität des neu generierten Proteins BpsB' eindeutig bewiesen. Bei den übrigen vier Metaboliten handelte es sich um C-terminal verkürzte Peptide, in denen die Manipulation der Biosynthese nicht zum Tragen kam. Die Funktionalität der im eingeführten Modul enthaltenen E-Domäne wurde durch die chirale Aminosäurenanalytik von Metabolit **6** in Abschnitt 5.5.1 exemplarisch gezeigt. Im Folgenden soll der „Dekorierungszustand“ dieser 17 Metabolite durch sog. „tailoring“-Reaktionen ausführlicher diskutiert werden. Darunter fallen neben den post-NRPS-Dekorierungen „N-Methylierung“ und „Glykosylierung“ auch die während der nicht-ribosomalen Peptidsynthese ablaufende inter-NRPS-Dekorierung „Chlorierung“ und die „oxidative Seitenkettenzyklisierung“.

Zum Chlorierungszustand der Metabolite ist zu bemerken, dass ausschließlich vollständig chlorierte Metabolite nachgewiesen wurden. Dies spricht für die Tatsache, dass zum einen die Funktionalität der Halogenase BhaA durch den genetischen Eingriff nicht beeinträchtigt wurde und zum anderen die für die Chlorierung erforderliche Protein-Protein-Interaktion der Halogenase mit den NRPS auch trotz der Modulinsertion in BpsB weiterhin möglich war. Die Substrate der Halogenase BhaA,<sup>[31]</sup> also die zwei an die T-Domänen der NRPS gebundenen  $\beta$ -Hydroxytyrosin-Aminosäuren, waren von der Modulinsertion nicht betroffen, so dass somit auch kein Kompatibilitätsproblem bezüglich der Substratspezifität zu erwarten war.

Anders liegen die Verhältnisse bei der oxidativen Seitenkettenzyklisierung. Die drei oxidativen Seitenkettenzyklisierungen A-B, C-O-D und D-O-E werden in *A. balhimycina* Wildtyp von den drei P450-abhängigen Monooxygenasen OxyA, OxyB und OxyC katalysiert, wobei folgende Reihenfolge und folgende Enzym-Ring-Zuordnung gilt: C-O-D (OxyB) -> D-O-E (OxyA) -> A-B (OxyC).<sup>[43]</sup> Außerdem konnte in der Vergangenheit gezeigt werden, dass diese oxidativen Seitenkettenzyklisierungen auf der Stufe der NRPS-gebundenen Hexa- bzw. Heptapeptide ablaufen, was für eine Protein-Protein-Interaktion zwischen den nicht-ribosomalen Peptidsynthetasen und den P450-abhängigen Monooxygenasen spricht.<sup>[44,45]</sup> Das Produktionsprofil der *A. balhimycina* C5A5T4E4-Mutante zeigte neben den linearen Metaboliten **1-7** sowie **10** und **11** die zwei monozyklischen Hepta- bzw. Oktapeptide **8** und **9**, bei denen mit Hilfe der aufgenommenen MS/MS-Spektren eindeutig die C-O-D-Seitenkettenzyklisierung nachgewiesen werden konnte (s. Abbildung 5-15). Daneben traten zwei weitere Metabolite **14** und **15** auf, bei denen sowohl die Molekülmassen als auch die MS/MS-Spektren darauf hindeuteten, dass es sich um die monoglykosylierten Derivate von **8** und **9** handelt. Somit kann man schließen, dass OxyB, das als erstes Enzym für die Knüpfung des C-O-D-Rings verantwortlich ist, eine gewisse Toleranz und damit Aktivität gegenüber dem neuen Peptidrückgrat aufbringen kann. Bei dem monozyklischen Heptapeptid **8** handelt es sich immerhin um den Metaboliten der quantitativ am zweitstärksten vertreten ist. Außerdem scheint die eben erwähnte Protein-Protein-Interaktion von OxyB mit den NRPS durch die Modulinsertion in BpsB nicht wesentlich gestört zu werden. Erstaunlicherweise traten vier weitere Molekülmassen auf, die auf ein bizyklisches bzw. trizyklisches Oktapeptid (**12**, **13**) sowie deren monoglykosylierte Analoga (**16**, **17**) hindeuteten. Ein so hoher Grad an oxidativer Seitenkettenzyklisierung verhinderte eine Strukturaufklärung anhand von MS/MS-Spektren, so dass eine NMR-Analyse notwendig war. Leider war die Produktionsrate dieser Metabolite so gering, dass keine Aufreinigung ausreichender Mengen erfolgen konnte, um den bizyklischen bzw. trizyklischen Charakter der Metabolite nachzuweisen. Dies deutete darauf hin, dass die zwei P450-abhängigen Monooxygenasen OxyA und OxyC das neue Peptidrückgrat, wenn überhaupt, nur in vernachlässigbarem Ausmaß als Substrat tolerieren können. Neben der nicht vorhandenen Substraterkennung durch die Enzyme könnte aber auch die Struktur des veränderten Peptidrückgrats selbst dazu führen, dass vor allem die beiden letzten oxidativen Seitenkettenzyklisierungen

beispielsweise aus sterischen Gründen nicht mehr oder sehr stark verringert möglich sind. Ein dritter möglicher Grund für die nicht nachweisbaren oxidativen Seitenkettenzyklisierungen durch OxyA und OxyC könnte auch eine Störung der Protein-Protein-Interaktion mit BpsB' sein.

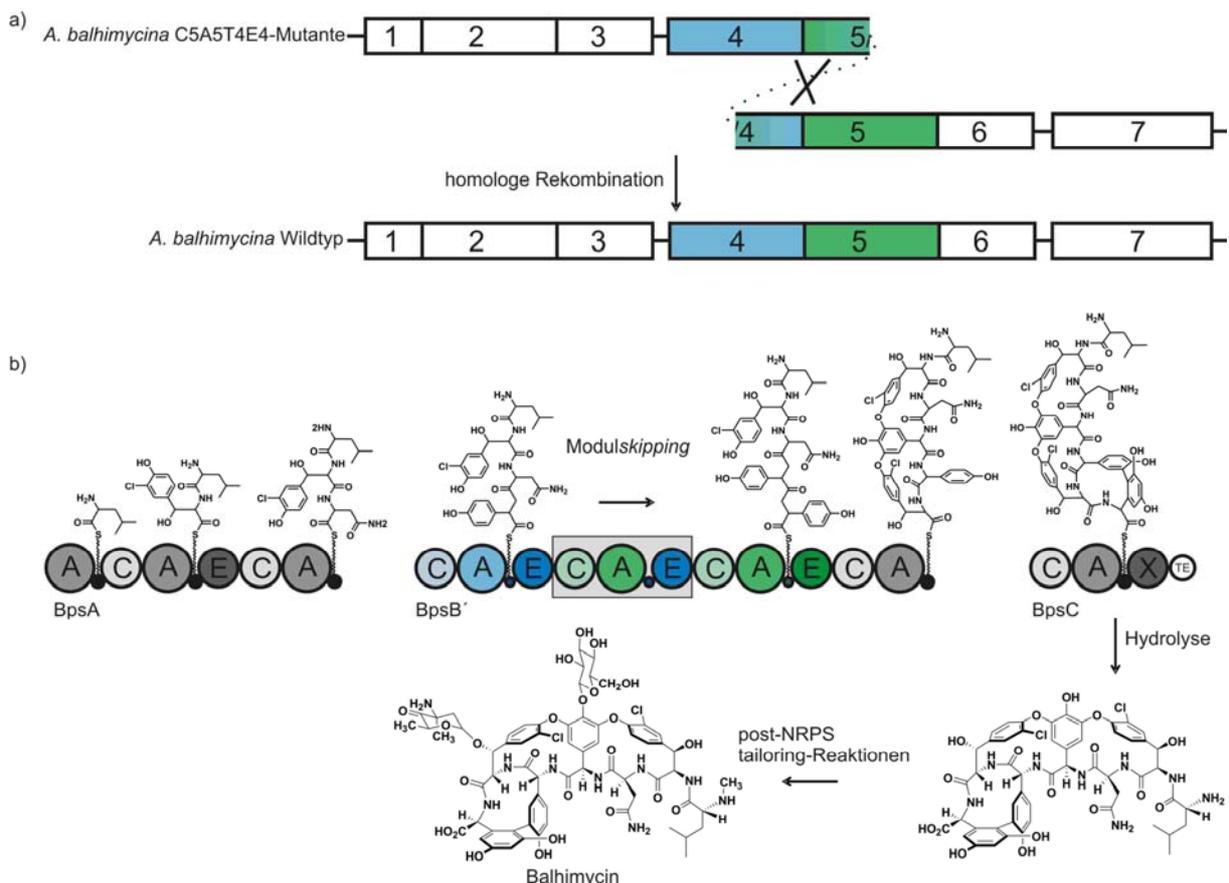
Die dritte Dekorierungsreaktion ist die *N*-Methylierung, die in allen Metaboliten des Produktionsprofils fehlte. Dies war nicht weiter verwunderlich, da es sich bei der *N*-Methylierung um eine post-NRPS-Reaktion handelt und im Wildtyp ein trizyklisches Heptapeptid als Substrat der *N*-Methyltransferase erkannt wird.<sup>[26,46]</sup> Die fehlenden bzw. unvollständigen oxidativen Seitenkettenzyklisierungen bei den Metaboliten der *A. balhimycina* C5A5T4E4-Mutante werden offensichtlich von der *N*-Methyltransferase nicht toleriert.

Auch bei der Glykosylierung geht man davon aus, dass sie als post-NRPS-Reaktion ein trizyklisches Heptapeptid als Substrat besitzt.<sup>[26,47,48]</sup> Aus diesem Grund fehlten bei allen gefundenen Metaboliten die für Balhimycin typischen Glykosylierungen, d.h. es fehlten die Glukose an 4-Hydroxyphenylglycin und das 4-oxo-Vankosamin an  $\beta$ -Hydroxytyrosin. Allerdings zeigten die MS/MS-Spektren der sechs Metabolite **10**, **11**, **14** bis **17** eine Abspaltung eines 162 amu Neutralfragments, was typisch für die Anwesenheit von Glukose oder verwandten isobaren Hexosen ist. Aus diesem Befund wurde geschlossen, dass es sich wahrscheinlich um die monoglykosylierten Analoga der unterschiedlich stark zyklisierten Hepta- bzw. Oktapeptide **6** bis **9**, **12** und **13** handelte. Die Lokalisierung der mutmaßlichen Hexose an die Aminosäure  $\beta$ -Hydroxytyrosin in Position zwei konnte anhand von MS/MS-Spektren der linearen Metabolite **10** und **11** zugeordnet werden (s. Abbildung 5-16). Dies entsprach zwar einer Balhimycin-typischen Glykosylierungsposition, an der aber üblicherweise 4-oxo-Vankosamin kovalent gebunden ist, welches die Abspaltung eines 143 amu Neutralfragments ergeben hätte. Aus diesem Grund muss von einer unspezifischen Glykosylierung ausgegangen werden. Unter Umständen enthält das Genom von *A. balhimycina* eine unspezifisch wirkende Glykosyltransferase, die das neue Peptidrückgrat als Substrat akzeptierte und glykosylierte.

Ein weiterer wichtiger Punkt, der Anlaß für Diskussionen liefert, ist das Vorkommen der C-terminal verkürzten Biosynthesevorläufer **1** bis **6** in relativ hohen Konzentrationen. Es gibt einige Beispiele in der Literatur, die das Vorkommen von Biosynthesevorläufern in Wildtyp-Produzenten<sup>[103,104]</sup> oder Mutanten<sup>[44]</sup> beschreiben. Hierbei wird vermutet, dass diese Intermediate vor Beendigung der

Biosynthesesequenz vom Multienzymkomplex abhydrolysiert werden und deshalb im Kulturüberstand nachweisbar sind. Dabei wurde die Beobachtung gemacht, dass Manipulationen der nicht-ribosomalen Peptidsynthese zu einer „kinetischen Blockade“ und damit zu einer verstärkten Hydrolyse der Biosyntheseintermediate führen können.<sup>[67]</sup> Dies könnte auch bei der *A. balhimycina* C5A5T4E4-Mutante eine Rolle spielen, da die nicht-ribosomale Peptidsynthese durch das Einbringen eines zusätzlichen Moduls vermutlich nicht mehr mit derselben Geschwindigkeit ablaufen kann wie im *A. balhimycina* Wildtyp. Einen wesentlichen Beitrag zur „kinetischen Blockade“ der *A. balhimycina* C5A5T4E4-Mutante liefert sehr wahrscheinlich die oben diskutierte reduzierte Aktivität der P450-abhängigen Monooxygenasen.

Der letzte Punkt, der in dieser Diskussion angesprochen werden soll, ist die in der *A. balhimycina* C5A5T4E4-Mutante beobachtete Produktion des Wildtyp-Glykopeptids Balhimycin in etwa derselben Menge wie die des Hauptmetaboliten **6**. Dazu sind prinzipiell zwei Erklärungen denkbar (s. Abbildung 6-1).

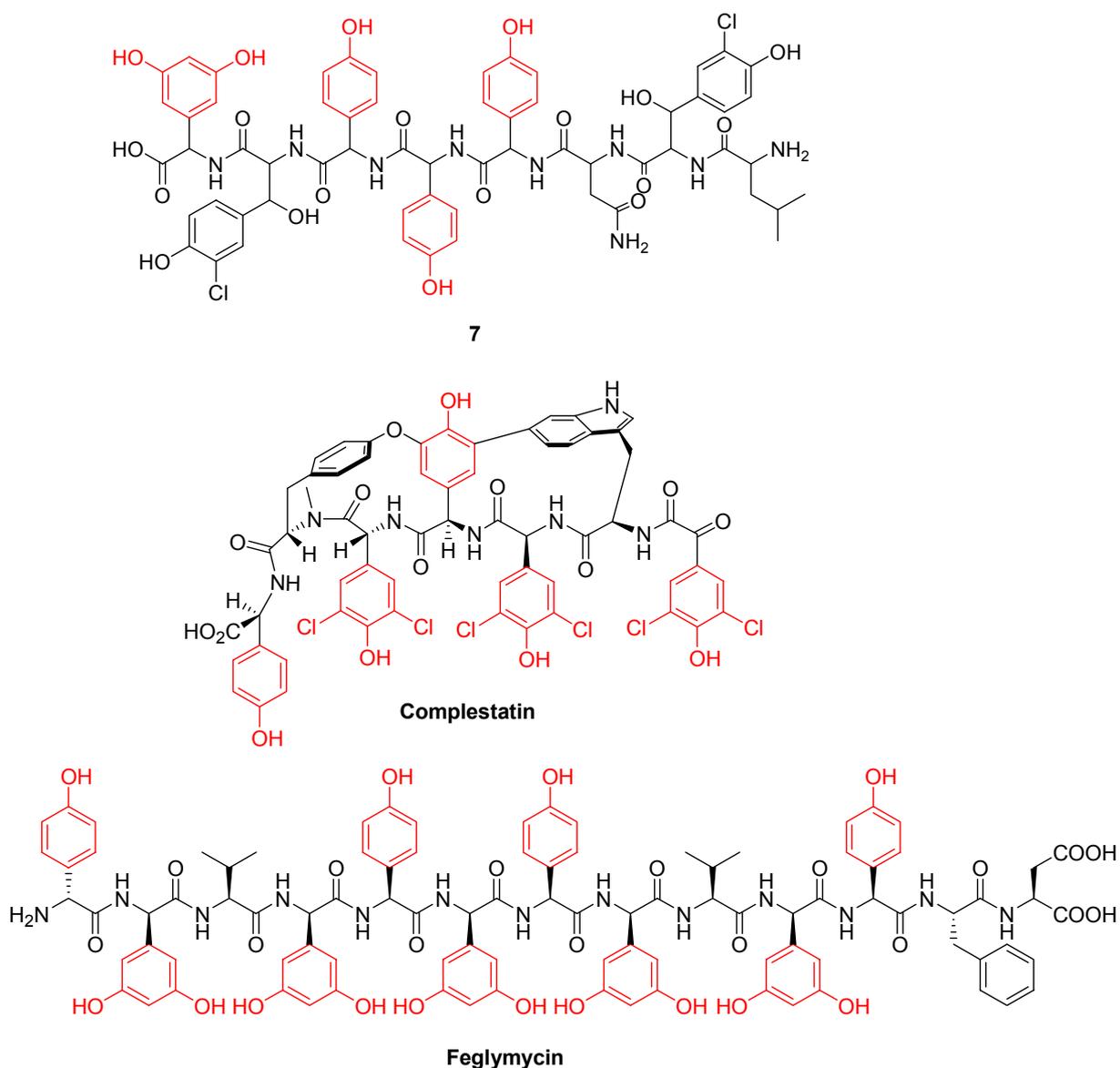


**Abbildung 6-1:** Zwei mögliche Erklärungen für die Produktion des Wildtyp-Glykopeptids Balhimycin in der *A. balhimycina* C5A5T4E4-Mutante. a) Die Instabilität des *A. balhimycina* C5A5T4E4-Genotyps ermöglicht eine erneute homologe Rekombination und damit Rückmutation zum Genotyp des *A. balhimycina* Wildtyps. b) In der *A. balhimycina* C5A5T4E4-Mutante tritt ein sog. *Modulskippping* des neu eingebrachten Hybridmoduls (grau unterlegt) auf.

Die erste Erklärung geht davon aus, dass neben dem *A. balhimycina* C5A5T4E4-Genotyp auch der Genotyp des *A. balhimycina* Wildtyps vorkommt. In anderen Worten bedeutet dies, dass es sich bei der fermentierten Kultur um eine Mischkultur gehandelt haben könnte. Es kann allerdings ausgeschlossen werden, dass es sich bei der Herstellung der Mutante von Anfang an um eine Mischkultur gehandelt hat, da das Stressprotokoll eine Protoplastierung und dementsprechend eine Vereinzelung der Zellen vorsah. In dem Zusammenhang muss darauf hingewiesen werden, dass bei den mycelbildenden Actinomyceten selbst keine klare Begrenzung der Zellen vorhanden ist. Man kann also nicht von vielen Einzelzellen mit jeweils einem Genom ausgehen, sondern eher von Zellclustern mit mehreren Genomen. Zur Herstellung von Mutanten wird deshalb der Weg über die Protoplasten gewählt. Aufgrund der Beschaffenheit des Genoms der Mutante muss aber von einer gewissen Instabilität der Mutante ausgegangen werden, die in der Verdopplung von Sequenzbereichen begründet ist. Diese Situation ermöglicht, zumindest theoretisch, eine erneute homologe Rekombination, die wieder zum Genotyp des Wildtyps führt (s. Abbildung 6-1a)). Der Nachweis des Wildtyp-Genoms in Gegenwart des Genoms der Mutante mit PCR oder *Southern Blot* ist allerdings nicht möglich, da jeder Sequenzabschnitt des Wildtyps auch in der Mutante vorkommt. Auf Proteinebene konnte diese teilweise Reversion der Mutante zum Wildtyp nicht beobachtet werden, wie Abbildung 5-12 zeigt. In der Spur des Proteingels der Mutante ist nämlich neben BpsB' kein Wildtyp BpsB sichtbar. Allerdings ist denkbar, dass diese Reversion zu BpsB nur zu einem sehr kleinen Prozentsatz auftritt, was mit Silberfärbung des Gels nicht detektiert werden kann. Es ist allerdings anzunehmen, dass die Biosynthesegeschwindigkeit von BpsB deutlich über der von BpsB' liegt, so dass auch eine geringe Menge an BpsB zu einer deutlich nachweisbaren Balhimycinproduktion führen könnte. Eine zweite mögliche Erklärung für das Vorhandensein von Balhimycin kann durch den Begriff „Modulskipping“ umschrieben werden. Dies beschreibt die Situation, dass ein Modul zwar auf genetischer Ebene und auf Proteinebene vorhanden ist, bei der Biosynthese des Produktes aber übersprungen wird (s. Abbildung 6-1b)). Dieses Phänomen wurde vorwiegend für Polyketidbiosynthesen beschrieben und ist für nicht-ribosomale Biosynthesen bisher kaum bekannt.<sup>[105]</sup> Aus diesem Grund wird im vorliegenden Fall die Reversion der Mutante zum Wildtyp als wahrscheinlichere Erklärung erachtet.

## 7 Ausblick

Im Arbeitskreis von Professor Süßmuth wird momentan von Frau Dr. Suada Turkanović an der Totalsynthese eines der Hauptprodukte der *A. balhimycina* C5A5T4E4-Mutante gearbeitet und zwar des linearen Oktapeptids **7** (s. Abbildung 7-1).



**Abbildung 7-1:** Das lineare Oktapeptid **7**, eines der Hauptmetabolite der *A. balhimycina* C5A5T4E4-Mutante, zeigt eine strukturelle Ähnlichkeit zu Complestatin und Feglymycin. Die häufig vorkommenden Aminosäuren 4-Hydroxyphenylglycin und 3,5-Dihydroxyphenylglycin sind rot dargestellt.

Das Ziel dieser Totalsynthese ist die Bestätigung der in der vorliegenden Arbeit vorgenommenen Strukturaufklärung dieser Verbindung. Darüber hinaus soll auf diese Weise eine genügend große Menge in Reinstform gewonnen werden, die für

das *Screening* nach biologischer Aktivität verwendet werden kann. Eine Aktivität analog der des Balhimycins bzw. Vancomycins ist aber nicht zu erwarten, da Struktur-Aktivitäts-Studien der Vergangenheit zeigten, dass die oxidativen Seitenkettenzyklisierungen und die damit bewirkte konformationelle Starrheit der Struktur essentiell sind für die hier vorhandene Form der antimikrobiellen Aktivität. So waren beispielsweise die von den OxyA- bzw. OxyC-Deletionsmutanten von *A. balhimycina* produzierten mono- bzw. bizyklischen Balhimycinderivate komplett inaktiv im Hemmhoftest mit dem Indikatorstamm *Bacillus subtilis*.<sup>[106]</sup> Allerdings besteht eine gewisse strukturelle Ähnlichkeit von **7** zu Complestatin<sup>[107]</sup> bzw. zu Feglymycin,<sup>[108,109]</sup> deren Strukturen in Abbildung 7-1 gezeigt sind. Alle drei Peptide besitzen eine außergewöhnlich hohe Zahl an 4-Hydroxyphenylglycinen und 3,5-Dihydroxyphenylglycinen. Bemerkenswerterweise weisen sowohl Complestatin als auch Feglymycin eine antivirale Aktivität gegenüber dem HI-Virus auf, die durchaus mit der eben genannten strukturellen Gemeinsamkeit zusammenhängen könnte. Aus diesem Grund ist bei **7** ein *Screening* nach biologischer Aktivität vor allem in Hinblick auf antivirale Eigenschaften sehr interessant und durchaus vielversprechend.

## 8 Anhang

### 8.1 Ergebnisse der massenspektrometrischen Analyse der Proteinbande aus *A. balhimycina* und deren Identifizierung als nicht-ribosomale Peptid-synthetase BpsB mittels Proteindatenbanksuche

#### *{MATRIX}* Mascot Search Results

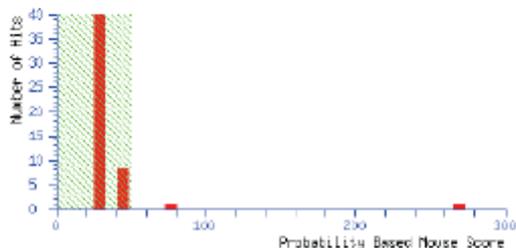
*{SCIENCE}*

```

User          :
Email         :
Search title  : D:\Data\070802\BUTZ1.D\Analysis.mgf
MS data file  : D:\Data\070802\BUTZ1.D\Analysis.mgf
Database      : NCBIInr 070122 (4462937 sequences; 1534242322 residues)
Taxonomy      : Bacteria (Eubacteria) (2173970 sequences)
Timestamp     : 6 Aug 2007 at 11:40:20 GMT
Significant hits: gi|15131493          peptide synthetase [Amycolatopsis balhimycina]
                  gi|21911414          StaC [Streptomyces toyocaensis]
  
```

#### Probability Based Mowse Score

Score is  $-10 \times \log(P)$ , where P is the probability that the observed match is a random event. Individual ions scores > 50 indicate identity or extensive homology ( $p < 0.05$ ).



#### *{MATRIX}* Mascot Search Results

##### Protein View

Match to: [gi|15131493](#); Score: 269  
 peptide synthetase [Amycolatopsis balhimycina]  
 Found in search of D:\Data\070802\BUTZ1.D\Analysis.mgf

Nominal mass ( $M_r$ ): 432163; Calculated pI value: 5.03  
 NCBI BLAST search of [gi|15131493](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Amycolatopsis balhimycina](#)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: 4%

Matched peptides shown in **Bold Red**

```

1 MSQSRIIBIW PLSPLQAGLL PHAVYDGBGP DVIYIHWILD LAGPVDAAGL
51 RAAWETLLAR HAPLRACFRQ RKSGETVQII ARQVELPWRE VDLSHLDDPE
101 BAVRRLABQD RTTRFDLAQA PLLRLTLIRL GADARLVVT CHFTIMDGWS
151 LPIVIDELSV LYFAGGDASA LFDVPSYREY LAWLSRQDKE RALSAWTAEI
201 SGAERPTLVV PADPGRAPAE PESVEAHLPE HLTRSLAELA RRHGLTLNTV
251 VQGAWALVLA QLAGRPDVVF GAAVSARPPD LFGVEGMWGL FLNIVPVVRE
301 LRgSTPVVEL LAELQXRQGA LIPDQPVGLA DIQQAAGPAA VPDTLVPEK
351 FHHGPAGSDS AGTFRIHVNQ GRVAAHYPLT LVAVPGBSMY LKLDYLTBLF
401 DRETAFALIE RFTGVLRQLT GAGELTVAGV EVTTAAERAL VAGEWGAETS
451 APPSLPALDI FGHQVAHRRD EFAVVDGDRT VSYGELAERA ERLAGYLNGR
501 GVRGRDRVAV VLDRSPDLIA TLLAVWKAGA AYVPVDPAYP VERRKFMAD
551 SGPAAVVCAR AYRAAVPDTG PEPIVLDDPR TRQAVRESPR LSAGTSADDL
  
```

601 AYVMTYSGST GTPKGVAVSH GNVAALAGEP GWRVGGDAV LLHASHAFDI  
 651 SLPEMNVFLL SGARVVLAGP GAVDGAALAA YVAGGVTAAH LTAGAPR**VLA**  
 701 **DES**PRAVAGL **REV**LTTGGDAV FLAAVBRVRC RVRNVRVRL YGPTREATLCA  
 751 TWMLLEPGDE TGSVLPFGRP LAGRRVHVD AFLRFVPEGV AGELYVAGAG  
 801 VAQGYSSRPA LTABRFVADP SGSGARMYRT GDLAYWTEQG ALAFAGRADD  
 851 QVKIRGYRVE PGELEVLAG LFGVGGAVVT PRGEHLI GYV VAENGHDADP  
 901 VRLREQLAGT LPEFMVPAAV LVLDLPLTV NGKVDRRALP EPDPAKSAG  
 951 REPVTRARRV LCGVPADVLG LDHVGVDDSF FELGGDS ISS MQVAARARRE  
 1001 GISLTPRLVF EHRTPERLAA LAQEGATPR AEVVTGVGEI FWTVPVRLAG  
 1051 DDAMRPGFAQ VVVVVTPAGV NPDALVSALQ AVLDAHDLR ARVEPDGRLI  
 1101 VPERGAVAAA GLLTR**VAA**GT **GGLDEIAERE** VRTAT**GTLD**P **SAGIMARVVV**  
 1151 **IDA**GDABPGR LAFVAHLSV DAVSWGILLP DLRAAYDEVI SGGTPALEPP  
 1201 VTSYRQWARR LTRALSBEST VAELEKMAAV VEGABPALPQ DTGQHTQOSH  
 1251 SWSTSLSGTE VRDLVTVLPG APHCGIQDVL LAGLAGAVAR VRGSGAALLV  
 1301 DVBEHGRRAA DGEDLLRTVG WFTSVHPVRL **ELSDVDLAGA** **ADGERPAGQL**  
 1351 **LK**AVBEQLRA VPGDGGYGL LRHLNPGTGA RLAEPLSAQI GFNYLGRITV  
 1401 APEDTAWQFN GGGPLGGGPD MVLAAHVEVS AELQDTPAGP RLGLAIDTRD  
 1451 PDLATVERLQ EAWLEMLTGL AAVARGSGAG GHTPADPALV DLTQDVAEL  
 1501 BAAAPGLTDI WPLSPLQEGM LPERAFDEBG VDVYQTRIL DLDGPLDEPR  
 1551 LRAAWNQVLA RHASLRTPGH QLGGGATVQV VVREADI PWR VADLSHLDA  
 1601 BAAAEVERLL AEDQGRFPDV TRPFLRLLL IRLGADHRL VVTSHHVLLD  
 1651 GWSTPLVGB MSDGYAGRS SSKPPSYQDY LAWLSRQDAB ATRSAWRABL  
 1701 AGADEPTLVD ADAGKTLVMP DEHAEWLEP ATRALAGFAR GHGLTVSTIV  
 1751 LGAMALVLR LAGRTDVVPG SVVSGRPADV PDVERMVGMP INTVPARVRL  
 1801 DGRAPLEML DLQAR**QAAL** **TEHQYLGLPR** **IQK**VAGTGAI FDTIVMVENY  
 1851 PHDAAGLGGD GGVAISSVVT **RTGTSYPLTM** **NVSLGDR**LRI TVSYRPRID  
 1901 DATAEVARQ VVRVLERVVA EPSLPGRLG VTSSEPTRAAV VERWNSTGEA  
 1951 AAEVSVLRFP RRQAGASDA VAVVAGERTL SYADLDRESD RLAGHLAGIG  
 2001 VGRGDRVGVV MTRGADLPVA LLGVWKAGAA QVPVNVDPYA ERIBRMLADV  
 2051 GASVAVCVBA TRKAVFDGVE FVVVDLPIG GVREPAPEVT VGADHDVAYVM  
 2101 YTSGSTGVPK AVAVPHGSVA ALASDPCWSQ GPGDCVLLHA SHAPDASLVE  
 2151 IWVPLVSGAR VLVAEPGTVD AERLREAVSR GVTTVHLTAG AFRAVAEESP  
 2201 DSFTGLR**RI**L **TGODAVPLAS** **VVR**MRQACPD VVRVQLYQPT EITLCATMLV  
 2251 LEPGAATGDV LPIGRPLAGR QAYVLDAPLQ PVAPNVTGEL YLAGAGLAHG  
  
 2301 YLGNTAATSE RFVANPFGG GRMYRTGDLA RWTQGRVLF AGRADSQVKI  
 2351 RGYRVEPGEV EVALTEVPHV AQAVVVAREG QPGEKRLIAY VTAEAGSALE  
 2401 SAAVRAHLAT RLPRPMVPSV VVULESPFLT LNKIKDRAAL PAPEFAGKAA  
 2451 GRBPRTRABR VLQGLFABIL GLBRVGADDG FPELGGDSIL SMRLAARARR  
 2501 BNFVFGAKQV FEQKTPAGIA AVAERGGQSR PAGVADGUGE VFWTPVREAL  
 2551 LERDPAGLTR GAMAQWVVA APRDLSVTAL VAGLGAVIDT HDMLRSRIVE  
 2601 SEGVPEPLVV AGRGTVDAAA LVERVEAGDG DLARIADROA HDTAARLDPV  
 2651 AGVLVRAVWV DAGPGRAGRL VVAHHLLVVD VVSWRTLLPD LQAACBAVVA  
 2701 GGGPALDFPD VSFRRWSRTL DGBAAIRTGE LAVWTEILDG AQSRGLGELDP  
 2751 RRDVSTAGR RSWTVPREHA GVLVEQVTSF FHCQVHRELL ATLAGAVAGV  
 2801 RCGTAVVVVD EHGGRQLGE LDLSRTLWPF TDVHPLRLDV TGVDPARAVA  
 2851 GGDAAAGLLK QVKENVRAV DGGLYGMLR YLNAETGPVL AALPKABIGF  
 2901 NYLGRPSAGE GGEAQPWQIT GIVGGAABQD TPLRHVVVID AVVVDGPDGP  
 2951 EFTLTVTWAG RMLGDABAES LAQAWLDMLA GLAAHVAAAG PGHSTPSDFP  
 3001 LTALTQREVA EFEEAVPGLL DIWFPLSPLQE GLLFHAADDR GPDVYAGMET  
 3051 **LAI**DGPLDVA **RFR**ASWTVLL DRHPALRASF HQLESGBAVQ VIARDVPPDW  
 3101 RETDLSGLPE SEALAEFDRL AARMHAERPD LTKAPQLRLH LVRLGDRRYR  
 3151 LIPTSHHIVA DGWSLPLILV DVLTAEBAGG DGRTPAATS YRDFLAWVDR  
 3201 QDKGAAGQAM RTELAGLDEA THVVPPGSI TPLEPERVAF ELDDETSREL  
 3251 VEPTRRHCVT ANTLPQGVWA LHLARLAGRN DVVFGAAVAG RPEIPGVES  
 3301 AVGLFMMMLP VRARLTGAEP VVDMLKDLQE RQVAMMAHQH IGLPEIKQLT  
 3351 GFGAAPDTIV VFENYPPAPP RSDDFDALVI RPVGIPNDTG HYPLEMRASV  
 3401 AAGPVRGEFI YRPDVVDRTB AGENVAAILR ALBQVVABPW TPVQVGLIG  
 3451 PEQRRLVVDE WNRDVPPLAA ETLVPLFRQ AERSPDAAV EDGARSLTFG  
 3501 GLLGEVBALA RLLVGAGVRR EHRVGVVLR SARLAVTMA VSPAGGVFVP  
 3551 VDPDYPREV EFMANSAPG VMVCTKTTRA AVPAEFAGTV LVLDLPAAD  
 3601 PDVELPPVAP EDAAVVIYTS GSTGVPEGLV VTHSGLANLG YAHIERMAYT  
 3651 SSSRVLQLSA TGFDAIVSEL YMALLAGATL VLPDAAAMPP RVTLSBAIRR  
 3701 AGITHLTVSP SVLASEDDL DTLR**TVLTGC** **RALPPALVDR** WSPGRVVIQA  
 3751 YGPTETTICS TMSAPLSPGH DQVPLGGPIH NVRHYVLDAP LQVPPGVVVG  
 3801 HLYITGVGLA RGYLGRPGLT AERPVAAPPA PGERMYRSGD LPRWTRBQL  
 3851 LFAGRVDAAQV KVRGYRVEPA EIEAVLAEHP WVGQVAVSVR RDGPGDKQLV  
 3901 AYVVPASADAA AENGTASAL RELAAERLPE YMPAAAFVSL EQMPLTPNGK  
 3951 LDHRLQAPD FAGMSSKRAP RTPWEARLCA LPADVGLDQ VGPDDSPFEL  
  
 4001 GGDSITSNQL SARARPTGLE LTPWQVDEK TPERLAVIQ ELAABGGTTP  
 4051 APEPGEGLV ALSPDQMDLL EAGLAGE

## 8.2 Ergebnisse der massenspektrometrischen Analyse der Proteinbande aus *A. balhimycina* C5A5T4E4-Mutante und deren Identifizierung als nicht-ribosomale Peptidsynthetase BpsB mittels Proteindatenbanksuche

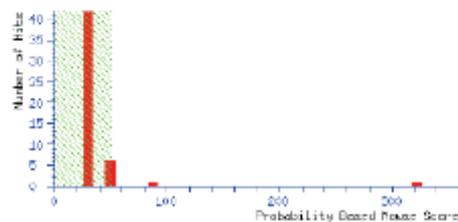
### **(MATRIX) SCIENCE Mascot Search Results**

User :  
 Email :  
 Search title : D:\Data\070802\BUTZ2.D\Analysis.mgf  
 MS data file : D:\Data\070802\BUTZ2.D\Analysis.mgf  
 Database : NCBI nr 070122 (4462937 sequences; 1534242322 residues)  
 Taxonomy : Bacteria (Subacteria) (2173970 sequences)  
 Timestamp : 6 Aug 2007 at 11:40:20 GMT  
 Significant hits: [gi|15131493](#) peptide synthetase [Amycolatopsis balhimycina]  
                   [gi|21911414](#) StaC [Streptomyces toyocaensis]  
                   [gi|115372740](#) ATPase, AAA family [Stigmatella aurantiaca]

DW4/3-1]

#### Probability Based Mowse Score

Score is  $-10 \times \log(P)$ , where P is the probability that the observed match is a random event.  
 Individual ions scores > 50 indicate identity or extensive homology ( $p < 0.05$ ).



### **(MATRIX) SCIENCE Mascot Search Results**

#### Protein View

Match to: [gi|15131493](#); Score: 323  
 peptide synthetase [Amycolatopsis balhimycina]  
 Found in search of D:\Data\070802\BUTZ2.D\Analysis.mgf  
 Nominal mass ( $M_r$ ): 432163; Calculated pI value: 5.03  
 NCBI BLAST search of [gi|15131493](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Amycolatopsis balhimycina](#)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is D  
 Sequence Coverage: 3%

Matched peptides shown in **Bold Red**

```

1  MQQRISSEIW PLSPLQAGLL PBAVYDGGCP DVTYIGHWILD LAGPVDAAGL
51  RAAMETLLAR HAPLRACFRQ RESGETVQII ARQVELPWRE VDLSHLDDPE
101  EAVRELAEQD RTRPDLAQA PLLRLTLRL GADAHRLVVT CHTTIMDQWS
151  LPIVIDELSV LYPAGGDASA LPDVPSEYREY LAWLSRQDKE RALSAWTAEI
201  SGAREPTLVV PADPGRAPAE PESVEAHLEP HLTRSLABEL ERHGLTLNTV
251  VQGANALNLA QLAGRFDVVF GRAVSRPFD LKGVGQVUGL FLNTVVRVR
301  LRGSTPVVEL LABLQERQSA LIPDQPVGLA DIQQANGPAA VFDTLVPEK
351  PHESPAGSDS AGTFRIHVNQ GRVAHYPLT LVAVPQEGMY LKLDYLTLEP
401  DRETAFAILE RPTCVLRQLT CAELTVACV EYTTAEREL VACEWCASTS
451  ADPSLDALDL FCHQVAHRED EDAVVDQDT VSYCLAEBA ERKACYLEKR
501  CVERCDRVAV VLRSFDLIA TLLAVWQCA AYVDVDDAYD VERREKMLAD
551  SCPAAVVCAB AYRAAVDQC DEDIVLDDP TRQAVAESDR LSACTSADDL
  
```

601 AYVMTYSGST GTPEGVAVSH GNVAAALGEP GNRVGPCEAV LLHASHAFDI  
 651 SLPEMVPVLL SGAEVVLACP GAUVDGAALAA YVAGGVTAAR LTACAPRYLA  
 701 DESPREAVAGL REVLTGGDAV PLAAVBRVRC EYRNVVREHL YGPTREATCA  
 751 TWMLLEGGDE TGSVLPICRP LAGREVVHLD AFLR PVPPGV AGLYVAGAG  
 801 VAQYYSRPA LTAERFVADP SGGARMYRT GDLAYWTEQG ALATAGRADD  
 851 QYKIRGTRVE PGEIEVVLAC LPGVQAVVT ERGEHLIGYV VARAGHDADP  
 901 VRLREQLAGT LPEFMVPAAY LVLDLPLTV NGKVDRRALP EPDPAKASG  
 951 REPVTBARKV LOGVFPADV LGDHVGVDDSE FELGGDSISS NQVAARARRE  
 1001 CISLTPRLVE EERTPERLAA LAQERAGATPR AEVVTGVBGI EWTVWNRALG  
 1051 DDAMRPGPAQ VEVVVTGAV NFDALVGLQ AVLDADHLR ARVEPDGRLI  
 1101 VFERGAVAAA CLLTRVAAGT GGLDEIAERS VRATGTLLDP SAGINARVWV  
 1151 **LDACDAEPCR** LAFVAHHLV DAVSWCILLD DLRAAYDEVI SCOTIALEDD  
 1201 VTSYRQWARR LTRALSSST **VABLEKWAAY** VECABDALQ DTCQHTQSH  
 1251 SWSTSLGCTE VEDLVTVLFC AFHCCIQDVL LACLACAVAR VRGSCAALLV  
 1301 DVEGCHREAA DCEDLLRTVC WFTSVHVPRL **ELSDVLLAGA** **ADGERFAGQL**  
 1351 **LKAVKQJIRA** VPCDGSYVCL LRHLNDCTGA RLARLPSAQI GPNYLGRVTL  
 1401 APEDTARQIN CCCCLOCCTD MVLAAHAEVS AELQDTDAGD RLCLAIDTRD  
 1451 PDLATVERLC EAWLEMLTCL AAVARCSGAC CHTDADPALV DLTQSDVABL  
 1501 EAMADCLTDI WFLSDLQBCM LFERAFDEDC VDVYQTRIL **DLDCPLDEPR**  
 1551 LRAANNQVLA RHASLETGFH QLGSGATVQV VVREADIPWE VADLSHLDA  
 1601 AAAAEVRELL AEDQGERFV TRPDLLELLL IRLGADHRL VVTSHHVLLD  
 1651 GWTPLVVGK MSDCYAGGRS SSKPPSYQDY LAWLSEQOAE ATRSAMRAEL  
 1701 AGADEPTLVD ADAGKTLVMP DEHAHWLPEP ATRALAGFAE GHGLTVSTIV  
 1751 LGAALVLRAR LAGRTDVVFG SVTSGRPADV EDERVMGVGE INTVPRVRL  
 1801 DGRRLPLEML EDLQARQAL TEHQYLGLEP IQEVAGTGAI FDTIVWENVY  
 1851 PHDAAGLQSD GGVALSIVVT RTGTSYPLTM NVSLGDLRLI TVSYRPRRID  
 1901 DATAAEVARQ VVRLERVVA EPLSPVGRLG VTSEPTRAAV VERMNSTGSA  
 1951 AABTSVLELP RRQAGASPA VAVVAGERTL SYADLDRBSD RLACHLAGIC  
 2001 VGRGDRVGVV MTRGADLPA LIGVWKGAGA **QVPPVNDYPA** **ER**IERMLADV  
 2051 GASVAVCVEA TRKAVPDGVE PVVVDLPAIC GVRPEADPVT VGHADVAIVM  
 2101 YTSGGTGVPE AVAVEHGSVA ALASDFQWQO GPGDCVLLHA SHAFDASLVE  
 2151 IWDVLSGAR VLVARDCTVD AERLREAVSR CVTTVHLTAC AFRAVABESD  
 2201 DSPICLRIL **TGCDAVPLAS** **VVRMRQACDD** VVRVQLYCDT EITLCATNLV  
 2251 LEDCAATGCV LPICRPLAGR QAYVLDAPLQ PVARNVTCEL YLACACLAHC  
  
 2301 YLGNAAATSE RFVANPFGG GMYRTGDLA RWIDQCELVF AGRADSQVKI  
 2351 EYRVERGEVE EVALIDEVPHV AQAUVVAREG QDGEKRLIAY VTARAGSALE  
 2401 SAARVAHLAT RLPEPMVPSV VVVLESPLLT LNKKIDRAAL PAFRFAGKAA  
 2451 GREPTBAER VLICLPAEIL GLEEVGADDG FPRLEGGDSIL SMRLAARARR  
 2501 ENFVGEAKQV PEQKTPAGIA AVAERGGQSR PAGVADGVGE VPWTFVVRAL  
 2551 LEKDAQLTR GAMAQWVSYA APRDLSTAL VAGLGAVIDT HEMLRGRIVE  
 2601 SBGVEPRLVV AGRCTVDAAA LVERVEAGDG DLAEIADRCA HDTAARLDPV  
 2651 AGVLVRAVWV DAGPCGRAGL VVAHHHLVVD VVSWRTLLPD LQAACRABVA  
 2701 GQQAALDPPD VSPREWSRL DGBAAIRTGE LAWVTEILDG AQSELGELPD  
 2751 EEDTVSTAGE RSWTVREHA GVLVEQVTSA FHCGVHEVLL ATLAGAVAGW  
 2801 RGGTAVVVVV BSHGRQPLGE LDLSRTLGWF TDVHPLRLDV TGVDPAEAVA  
 2851 GGDANGRLK QYKENVRAVP DGLGQYGMRL YLNAETGPVL AALPKAEIGF  
 2901 NYLGRFSAGS GGBAQPNQIT GIVGGAABQD TPLRHVVVID AVVVDGDDP  
 2951 EPTLTVTWAG RMLGDABAES LAQAWLMLA GLAAHVAAQO PGHPTPSDFP  
 3001 LTALTQREVA EFBAAVPGLL DIWFLSPLQE GLLFHAADDR GPDVYASMT  
 3051 **LAIDQFLDVA** **RFRASWTVLL** DRHDALRAS FOLESCERAVQ VIARDVDDW  
 3101 RETDLGLFE SEALAEFDRL AARMHAERPD LTKAPQLRLH LVRLGDRRYR  
 3151 LIFTSHHIVA DQWLSPLILV DVLTAEBAGS DGRTLPAATS YRDPLAWDR  
 3201 QDKGAAQAM RTELAGLDEA THVVPFGSII TPLEPERRAP ELDETSKRL  
 3251 VEFTREHGVY ANTLPOGVMA LHLARLAGRN DVVFGAIVAG RPEEIPGVES  
 3301 AVGLFMHMLD VRARLTCAED VVDMKLDLQE KQVMMMAHQH ICLDEIKQLT  
 3351 GCAAFDITV VFENYDDADP RSDODDALVI RDVCIHNDTC HYDLSMRASV  
 3401 AACDVRCEFI YRDVVDRTE AGEMVAAILR ALBQVVAEIV TPVQVGLIC  
 3451 EBQRLLVWDE WNRIDVGLAA ETLVLPFRQ AERSHDAVAV EDCARSLTFC  
 3501 GLLGRVREAL RLLVGCQVRE RHRVGVLERE SARLAVTMA VSPAGGVFVD  
 3551 VDFDYPRREY HFMLANSAPG VMVCTRTTRA AVPAEFAGTV LVLDLPLAAD  
 3601 DDVELDPVAD EDAAYVITS CSTCVVKCVL VTHECLAMLC YAHIERMAVT  
 3651 SSSRVLQLSA TCFDAIVSEL YMALLACATL VLPDAASHMP RVTLCBAIRR  
 3701 ACITHLTVSD SVLASEDDL DTLRTVLTCC EALDHALVDR WSDCRVIGA  
 3751 YGPTETTICS TMSADLSGCH DQVPLGGPIH NVRHVYLDAP LQVVDGQVVG  
 3801 ELYITGVGLA RGYLGRPGLT AEFVVASPPA DGRMNVRESD LPEWTRRQJL  
 3851 LPAGRVDAQV KYRGTYRVEPA EIRAVLAESP WVGQVAVSVR EDGPGDKQLV  
 3901 AVVUPSADAA AENCTLAGAL RELAAERLPE YMMPAAFVSL EQMPLTPNKK  
 3951 LDRALQAPD FAGMSSKRAP RTWMBARLCA LPADVLELDQ VGPDDSFREL  
  
 4001 GCDSTSMQL SARARPTGLR LTPNQVFDK TPRRLAVIVQ RLAAEGCTTP  
 4051 APBEGCTIV ALSPDQMDLL RAGLAGE

## 9 Literaturverzeichnis

- [1] Fleming, A. (1929). On the antibacterial action of cultures of a penicillium with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* *10*, 226-236.
- [2] Lerner, P.I. (2004). Producing penicillin. *N. Engl. J. Med.* *351*, 524.
- [3] McCormick, M.H., McGuire, J.M., Pittenger, G.E., Pittenger, R.C., Stark, W.M. (1955/1956). Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiot. Annu.* *3*, 606-611.
- [4] Walsh, C. (2000). Molecular mechanisms that confer antibacterial drug resistance. *Nature*, *406*, 775-781.
- [5] a) Barbachyn, M.R., Ford, C.W. (2003). Struktur-Wirkungs-Beziehungen von Oxazolidinonen: Grundlage der Entwicklung von Linezolid. *Angew. Chem.* *115*, 2056-2070.  
b) Barbachyn, M.R., Ford, C.W. (2003). Oxazolidinone structure-activity relationships leading to linezolid. *Angew. Chem. Int. Ed.* *42*, 2010-2023.
- [6] Kirkpatrick, P., Raja, A., LaBonte, J., Lebbos, J. (2003). Daptomycin. *Nat. Rev. Drug Discov.* *2*, 943-944.
- [7] Alder, J.D. (2005). Daptomycin: a new drug class for the treatment of Gram-positive infections. *Drugs Today*, *41*, 81-90.
- [8] a) von Nussbaum, F., Brands, M., Hinzen, B., Weigand, S., Häbich, D. (2006). Antibakterielle Naturstoffe in der medizinischen Chemie – Exodus oder Renaissance? *Angew. Chem.* *118*, 5194-5254.  
b) von Nussbaum, F., Brands, M., Hinzen, B., Weigand, S., Häbich, D. (2006). Antibacterial natural products in medicinal chemistry – exodus or revival? *Angew. Chem. Int. Ed.* *45*, 5072-5129.
- [9] Sieber, S.A., Marahiel, M.A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem. Rev.* *105*, 715-738.
- [10] Schwarzer, D., Finking, R., Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* *20*, 275-287.
- [11] Mootz, H.D., Schwarzer, D., Marahiel, M.A. (2002). Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBioChem*, *3*, 490-504.
- [12] Stachelhaus, T., Marahiel, M.A. (1995). Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* *270*, 6163-6169.

- [13] Stachelhaus, T., Mootz, H.D., Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6, 493-505.
- [14] Challis, G.L., Ravel, J., Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* 7, 211-224.
- [15] Eppelmann, K., Stachelhaus, T., Marahiel, M.A. (2002). Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry*, 41, 9718-9726.
- [16] a) Nicolaou, K.C., Boddy, C.N.C., Bräse, S., Winssinger, N. (1999). *Chemie, Biologie und medizinische Anwendungen der Glycopeptid-Antibiotika*. *Angew. Chem.* 111, 2230-2287.  
b) Nicolaou, K.C., Boddy, C.N.C., Bräse, S., Winssinger, N. (1999). Chemistry, biology and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed.* 38, 2096-2152.
- [17] Pace, J.L., Yang, G. (2006). Glycopeptides: update on an old successful antibiotic class. *Biochem. Pharmacol.* 71, 968-980.
- [18] Sheldrick, G.M., Jones, P.G., Kennard, O., Williams, D.H., Smith, G.A. (1978). Structure of vancomycin and its complex with acetyl-D-alanyl-D-alanine. *Nature*, 271, 223-225.
- [19] Harris, C.M., Kopecka, H., Harris, T.M. (1983). Vancomycin: structure and transformation to CDP-1. *J. Am. Chem. Soc.* 105, 6915-6922.
- [20] Schäfer, M., Schneider, T.R., Sheldrick, G.M. (1996). Crystal structure of vancomycin. *Structure*, 4, 1509-1515.
- [21] Nadkarni, S.R., Patel, M.V., Chatterjee, S., Vijayakumar, E.K.S., Desikan, K.R., Blumbach, J., Ganguli, B.N. (1994). Balhimycin, a new glycopeptide antibiotic produced by *Amycolatopsis* sp. Y-86,21022. Taxonomy, production, isolation and biological activity. *J. Antibiot.*, 47, 334-341.
- [22] Chatterjee, S., Vijayakumar, E.K.S., Nadkarni, S.R., Patel, M.V., Blumbach, J., Ganguli, B.N. (1994). Balhimycin, a new glycopeptide antibiotic with an unusual hydrated 3-Amino-4-oxoaldopyranose sugar moiety. *J. Org. Chem.* 59, 3480-3484.
- [23] Wink, J.M., Kroppenstedt, R.M., Ganguli, B.N., Nadkarni, S.R., Schumann, P., Seibert, G., Stackebrandt, E. (2003). Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimycina* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. *System. Appl. Microbiol.* 26, 38-46.

- 
- [24] van Wageningen, A.M.A., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J.M., Solenber, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of vancomycin group antibiotic. *Chem. Biol.* 5, 155-162.
- [25] Pelzer, S., Süßmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G., Wohlleben, W. (1999). Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. *Antimicrob. Agents Chemother.* 43, 1565-1573.
- [26] a) Süßmuth, R.D., Pelzer, S., Nicholson, G., Walk, T., Wohlleben, W., Jung, G. (1999). Neues zur Biosynthese von Glycopeptid-Antibiotika des Vancomycin-Typs aus *Amycolatopsis mediterranei*. *Angew. Chem.* 111, 2096-2099.  
b) Süßmuth, R.D., Pelzer, S., Nicholson, G., Walk, T., Wohlleben, W., Jung, G. (1999). New advances in the biosynthesis of glycopeptide antibiotics of the vancomycin type from *Amycolatopsis mediterranei*. *Angew. Chem. Int. Ed.* 38, 1976-1979.
- [27] a) Hubbard, B.K., Walsh, C.T. (2003). Der Aufbau von Vancomycin: so macht es die Natur. *Angew. Chem.* 115, 752-789.  
b) Hubbard, B.K., Walsh, C.T. (2003). Vancomycin assembly: nature's way. *Angew. Chem. Int. Ed.* 42, 730-765.
- [28] Süßmuth, R.D., Wohlleben, W. (2004). The biosynthesis of glycopeptide antibiotics – a model for complex, non-ribosomally synthesized, peptidic secondary metabolites. *Appl. Microbiol. Biotechnol.* 63, 344-350.
- [29] EMBL/GenBank/DDBJ, accession number Y16952, letzte Aktualisierung am 12.06.2006.
- [30] Puk, O., Huber, P., Bischoff, D., Recktenwald, J., Jung, G., Süßmuth, R.D., van Peé, K.-H., Wohlleben, W., Pelzer, S. (2002). Glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908: function of a halogenase and a haloperoxidase/perhydrolase. *Chem. Biol.* 9, 225-235.
- [31] Puk, O., Bischoff, D., Kittel, C., Pelzer, S., Weist, S., Stegmann, E., Süßmuth, R.D., Wohlleben, W. (2004). Biosynthesis of chloro- $\beta$ -hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. *J. Bacteriol.* 186, 6093-6100.
- [32] Choroba, O.W., Williams, D.H., Spencer, J.B. (2000). Biosynthesis of the vancomycin group of antibiotics: involvement of an unusual dioxygenase in the pathway to (S)-4-hydroxyphenylglycine. *J. Am. Chem. Soc.* 122, 5389-5390.
- [33] Hubbard, B.K., Thomas, M.G., Walsh, C.T. (2000). Biosynthesis of L-p-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chem. Biol.* 42, 1-12.
-

- 
- [34] Li, T.L., Choroba, O.W., Charles, E.H., Sandercock, A.M., Williams, D.H., Spencer, J.B. (2001). Characterisation of a hydroxymandelate oxidase involved in the biosynthesis of two unusual amino acids occurring in the vancomycin group of antibiotics. *J. Chem. Soc. Chem. Commun.* 1752-1753.
- [35] Pfeifer, V., Nicholson, G.J., Ries, J., Recktenwald, J., Schefer, A.B., Shawky, R.M., Schröder, J., Wohlleben, W., Pelzer, S. (2001). A polyketide synthase in glycopeptide biosynthesis: the biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. *J. Biol. Chem.* 276, 38370-38377.
- [36] Chen, H., Tseng, C.C., Hubbard, B.K., Walsh, C.T. (2001). Glycopeptide antibiotic biosynthesis: enzymatic assembly of the dedicated amino acid monomer (S)-3,5-dihydroxyphenylglycine. *Proc. Natl. Acad. Sci. USA*, 98, 14901-14906.
- [37] Li, T.L., Choroba, O.W., Hong, H., Williams, D.H., Spencer, J.B. (2001). Biosynthesis of the vancomycin group of antibiotics: characterization of a type III polyketide synthase in the pathway to (S)-3,5-dihydroxyphenylglycine. *J. Chem. Soc. Chem. Commun.* 2156-2157.
- [38] Sandercock, A.M., Charles, E.H., Scaife, W., Kirkpatrick, P.N., O'Brien, S.W., Papageorgiou, E.A., Spencer, J.B., Williams, D.H. (2001). Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. *J. Chem. Soc. Chem. Commun.* 14, 1252-1253.
- [39] Chen, H., Thomas, M.G., Hubbard, B.K., Losey, H.C., Walsh, C.T., Burkart, M.D. (2000). Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. *Proc. Natl. Acad. Sci. USA*, 97, 11942-11947.
- [40] Recktenwald, J., Shawky, R., Puk, O., Pfennig, F., Keller, U., Wohlleben, W., Pelzer, S. (2002). Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. *Microbiol.* 148, 1105-1118.
- [41] Trauger, J.W., Walsh, C.T. (2000). Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic. *Proc. Natl. Acad. Sci. USA*, 97, 3112-3117.
- [42] Stegmann, E., Pelzer, S., Bischoff, D., Puk, O., Stockert, S., Butz, D., Zerbe, K., Robinson, J., Süssmuth, R.D., Wohlleben, W. (2006). Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. *J. Biotechnol.* 124, 640-653.

- [43] a) Bischoff, D., Pelzer, S., Bister, B., Nicholson, G.J., Stockert, S., Schirle, M., Wohlleben, W., Jung, G., Süssmuth, R.D. (2001). Biosynthese von Glycopeptid-Antibiotika vom Vancomycin-Typ: die Reihenfolge der Ringschlüsse. *Angew. Chem.* *113*, 4824-4827.  
b) Bischoff, D., Pelzer, S., Bister, B., Nicholson, G.J., Stockert, S., Schirle, M., Wohlleben, W., Jung, G., Süssmuth, R.D. (2001). The biosynthesis of vancomycin-type glycopeptide antibiotics - the order of the cyclization steps *Angew. Chem. Int. Ed.* *40*, 4688-4691.
- [44] Bischoff, D., Bister, B., Bertazzo, M., Pfeifer, V., Stegmann, E., Nicholson, G.J., Keller, S., Pelzer, S., Wohlleben, W., Süssmuth, R.D. (2005). The biosynthesis of vancomycin-type glycopeptide antibiotics – a model for oxidative side-chain cross-linking by oxygenases coupled to the action of peptide synthetases. *ChemBioChem*, *6*, 267-272.
- [45] a) Zerbe, K., Woithe, K., Li, D.B., Vitali, F., Bigler, L., Robinson, J.A. (2004). An oxidative phenol coupling reaction catalyzed by OxyB, a cytochrome P450 from the vancomycin-producing microorganism. *Angew. Chem.* *116*, 6877-6881.  
b) Zerbe, K., Woithe, K., Li, D.B., Vitali, F., Bigler, L., Robinson, J.A. (2004). An oxidative phenol coupling reaction catalyzed by OxyB, a cytochrome P450 from the vancomycin-producing microorganism. *Angew. Chem. Int. Ed.* *43*, 6709-6713.
- [46] O'Brien, D.P., Kirkpatrick, P.N., O'Brien, S.W., Staroske, T., Richardson, T.I., Evans, D.A., Hopkinson, A., Spencer, J.B., Williams, D.H. (2000). Expression and assay of an N-methyltransferase involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Commun.* 103-104.
- [47] Solenberg, P.J., Matsushima, P., Stack, D.R., Wilkie, S.C., Thompson, R.C., Baltz, R.H. (1997). Production of hybrid glycopeptide antibiotics in vitro and in *Streptomyces toyocaensis*. *Chem. Biol.* *4*, 195-202.
- [48] Losey, H.C., Peczuh, M.W., Chen, Z., Eggert, U.S., Dong, S.D., Pelczar, I., Kahne, D., Walsh, C.T. (2001). Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. *Biochemistry*, *40*, 4745-4755.
- [49] Bardone, M.R., Paternoster, M., Coronelli, C. (1978). Teichomycins, new antibiotics from *Actinoplanes teichomyceticus*-nov-sp. II. Extraction and chemical characterization. *J. Antibiot.* *31*, 170-177.
- [50] Barna, J.C.J., Williams, D.H., Stone, D.J.M., Leung, T.W.C., Doddrell, D.M. (1984). Structure elucidation of the teicoplanin antibiotics. *J. Am. Chem. Soc.* *106*, 4895-4902.
- [51] Levine, J.F. (1987). Vancomycin: a review. *Med. Clin. N. Am.* *71*, 1135-1145.

- [52] a) Williams, D.H., Bardsley, B. (1999). Die Vancomycin-Antibiotika und der Kampf gegen resistente Bakterien. *Angew. Chem.* *111*, 1264-1286.  
b) Williams, D.H., Bardsley, B. (1999). Die vancomycin group of antibiotics and the fight against resistant bacteria. *Angew. Chem. Int. Ed.* *38*, 1173-1193.
- [53] Uttley, A.H., Collins, C.H., Naidoo, J., George, R.C. (1988). Vancomycin-resistant enterococci. *Lancet*, *1*, 57-58.
- [54] Hiramatsu, K., Hanaki, H., Ino, T., Yabuta, K., Oguri, T., Tenover, F.C. (1997). Methicillin-resistant *Staphylococcus aureus* clinical strain with reduced vancomycin susceptibility. *J. Antimicrob. Chemother.* *40*, 135-136.
- [55] McComas, C.C., Crowley, B.M., Boger, D.L. (2003). Partitioning the loss in vancomycin binding affinity for D-Ala-D-Lac into lost H-bond and repulsive lone pair contributions. *J. Am. Chem. Soc.* *125*, 9314-9315.
- [56] Crowley, B.M., Boger, D.L. (2006). Total synthesis and evaluation of [ $\Psi$ [CH<sub>2</sub>NH]Tpg<sup>4</sup>]Vancomycin aglycon: reengineering Vancomycin for dual D-Ala-D-Ala and D-Ala-D-Lac binding. *J. Am. Chem. Soc.* *128*, 2885-2892.
- [57] Park, I.-S., Lin, C.-H., Walsh, C.T. (1997). Bacterial resistance to vancomycin: overproduction, purification, and characterization of VanC2 from *Enterococcus casseliflavus* as a D-Ala-D-Ser ligase. *Proc. Natl. Acad. Sci. USA*, *94*, 10040-10044.
- [58] Lowy, F.D. (2003). Antimicrobial resistance. the example of *Staphylococcus aureus*. *J. Clin. Invest.* *111*, 1265-1273.
- [59] Floss, H.G. (2006). Combinatorial biosynthesis – potential and problems. *J. Biotechnol.* *124*, 242-257.
- [60] Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: challenges and opportunities. *ChemBioChem*, *3*, 125-134.
- [61] Keller, U., Schauwecker, F. (2003). Combinatorial biosynthesis of non-ribosomal peptides. *Comb. Chem. High Throughput Screen.* *6*, 527-540.
- [62] Baltz, R.H. (2006). Molecular engineering approaches to peptide, polyketide and other antibiotics. *Nat. Biotechnol.* *24*, 1533-1540.
- [63] Stachelhaus, T., Schneider, A., Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science*, *269*, 69-72.
- [64] Schneider, A., Stachelhaus, T., Marahiel, M.A. (1998). Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Mol. Gen. Genet.* *257*, 308-318.

- 
- [65] Yakimov, M.M., Giuliano, L., Timmis, K.N., Golyshin, O.N. (2000). Recombinant acylheptapeptide lichenysin: high level of production by *Bacillus subtilis* cells. *J. Mol. Microbiol. Biotechnol.* 2, 217-224.
- [66] Eppelmann, K., Stachelhaus, T., Marahiel, M.A. (2002). Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry*, 41, 9718-9726.
- [67] Uguru, G.C., Milne, C., Borg, M., Flett, F., Smith, C.P., Micklefield, J. (2004). Active-site modifications of adenylation domains lead to hydrolysis of upstream nonribosomal peptidyl thioester intermediates. *J. Am. Chem. Soc.* 126, 5032-5033.
- [68] Baltz, R.H., Brian, P., Miao, V., Wrigley, S.K. (2006). Combinatorial biosynthesis of lipopeptide antibiotics in *Streptomyces roseosporus*. *J. Ind. Microbiol. Biotechnol.* 33, 66-74.
- [69] Nguyen, K.T., Ritz, D., Gu, J.Q., Alexander, D., Chu, M., Miao, V., Brian, P., Baltz, R.H. (2006). Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proc. Natl. Acad. Sci. USA*, 103, 17462-17467.
- [70] Mootz, H.D., Kessler, N., Linne, U., Eppelmann, K., Schwarzer, D., Marahiel, M.A. (2002). Decreasing the ring size of a cyclic nonribosomal peptide antibiotic by in-frame module deletion in the biosynthetic genes. *J. Am. Chem. Soc.* 124, 10980-10981.
- [71] Agilent 1100-Serie Kapillar-LC-System Systemhandbuch, 137.
- [72] Mit Molekülmasse ist im Bezug auf die Massenspektrometrie immer die monoisotopische Masse gemeint.
- [73] Bild modifiziert ausgehend von: Lottspeich, F., Engels, J.W. (2006). *Bioanalytik*, 352.
- [74] Gross, J. H. (2004). *Mass Spectrometry*, Springer Verlag, 149.
- [75] Bild modifiziert ausgehend von: Lottspeich, F., Engels, J.W. (2006). *Bioanalytik*, 366.
- [76] Lottspeich, F., Engels, J.W. (2006). *Bioanalytik*, 367.
- [77] Bullock, W.O., Fernandez, J.M., Short, J.M. (1987). XL1-Blue: a high-efficiency plasmid transforming *recA Escherichia coli* strain with  $\beta$ -galactosidase selection. *Biotechniques*, 5, 376-378.
- [78] MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene*, 111, 61-68.

- 
- [79] Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51, 263-273.
- [80] Vogelstein, B., Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. U. S. A.* 76, 615-619.
- [81] Hamaguchi, K., Geiduschek, P. (1962). The effect of electrolytes on the stability of the deoxyribonucleate helix. *J. Am. Chem. Soc.* 84, 1329-1338.
- [82] Ullmann, A., Perrin, D. (1970). Complementation in  $\beta$ -galactosidase. Beckwith, J.R., Zipser, D. (Hrsg.) *The lactose operon*. Verlag Cold Spring Harbor Laboratory Press, New York, 143-172.
- [83] a) Eckhardt, T. (1978). A rapid method for the identification of plasmid deoxyribonucleic acid in bacteria. *Plasmid*, 1, 584-588.  
b) Labes, G., Simon, R., Wohlleben, W. (1990). A rapid method for the analysis of plasmid content and copy number in various *Streptomyces* grown on agar plates. *Nucleic Acids Res.* 18, 2197.
- [84] Dagert, M., Ehrlich, S.D. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene*, 6, 23-28.
- [85] Sambrook, J., Russell, D.W. (2001). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press.
- [86] Madoń, J., Hütter, R. (1991). Transformation system for *Amycolatopsis (Nocardia) mediterranei*: direct transformation of mycelium with plasmid DNA. *J. Bacteriol.* 173, 6325-6331.
- [87] Kumar, C.V., Coque, J.-J.R., Martín, J.F. (1994). Efficient transformation of the cephamycin C producer *Nocardia lactamdurans* and development of shuttle and promoter-probe cloning vectors. *Appl. Environ. Microbiol.* 60, 4086-4093.
- [88] Vrijbloed, J.W., Madoń, J., Dijkhuizen, L. (1995). Transformation of the methylotrophic actinomycete *Amycolatopsis methanolica* with plasmid DNA: stimulatory effect of a pMEA300-encoded gene. *Plasmid*, 34, 96-104.
- [89] Pelzer, S., Reichert, W., Huppert, M., Heckmann, D., Wohlleben, W. (1997). Cloning and analysis of a peptide synthetase gene of the balhimycin producer *Amycolatopsis mediterranei* DSM5908 and development of a gene disruption/replacement system. *J. Biotechnol.* 56, 115-128.
- [90] Birnboim, H.C., Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7, 1513-1523.
- [91] Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503-517.

- 
- [92] Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* *72*, 248-254.
- [93] Rabilloud, T. (1990). Mechanisms of protein silver staining in polyacrylamide gels: a 10-year synthesis. *Electrophoresis*, *11*, 785-794.
- [94] Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S., Mische, S.M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis*, *20*, 601-605.
- [95] Roepstorff, P., Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* *11*, 601.
- [96] Finn, R.D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L.L., Bateman, A. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res.* *34*, 247-251.
- [97] Donadio, S., Sosio, M., Stegmann, E., Weber, T., Wohlleben, W. (2005). Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. *Mol. Gen. Genomics*, *274*, 40-50.
- [98] Li, T.-L., Huang, F., Haydock, S.F., Mironenko, T., Leadlay, P.F., Spencer, J.B. (2004). Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* *11*, 107-119.
- [99] Sosio, M., Kloosterman, H., Bianchi, A., de Vreugd, P., Dijkhuizen, L., Donadio, S. (2004). Organization of the teicoplanin gene cluster in *Actinoplanes teichomyeticus*. *Microbiology*, *150*, 95-102.
- [100] Pootoolal, J., Thomas, M.G., Marshall, C.G., Neu, J.M., Hubbard, B.K., Walsh, C.T., Wright, G.D. (2002). Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci. USA*, *99*, 8962-8967.
- [101] Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., Donadio, S. (2003). The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. *Chem. Biol.* *10*, 541-549.
- [102] Chiu, H.T., Hubbard, B.K., Shah, A.N., Eide, J., Fredenburg, R.A., Walsh, C.T., Khoshla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. USA*, *98*, 8548-8553.

- 
- [103] Kammerer, B., Kahlich, R., Laufer, S., Li, S.M., Heide, L., Gleiter, C.H. (2004). Mass spectrometric pathway monitoring of secondary metabolites: systematic analysis of culture extracts of *Streptomyces* species. *Anal. Biochem.* 335, 17-29.
- [104] Hadatsch, B., Butz, D., Schmiederer, T., Steudle, J., Wohlleben, W., Süßmuth, R.D., Stegmann, E. (2007). The biosynthesis of teicoplanin-type glycopeptide antibiotics: assignment of P450 mono-oxygenases to side chain cyclizations of glycopeptide A47934. *Chem. Biol.* 14, 1078-1089.
- [105] Wenzel, S.C., Kunze, B., Höfle, G., Silakowski, B., Scharfe, M., Blöcker, H., Müller, R. (2005). Structure and biosynthesis of myxochromides S<sub>1-3</sub> in *Stigmatella aurantiaca*: evidence for an iterative bacterial type I polyketide synthase and for module skipping in nonribosomal peptide biosynthesis. *ChemBioChem*, 6, 375-385.
- [106] Dissertation Daniel Bischoff (2003). Zur Biosynthese der Vancomycin-Typ Glykopeptid-Antibiotika – Isolierung und Strukturaufklärung von Biosynthese-Metaboliten aus *Amycolatopsis balhimycina*.
- [107] Momota, K., Kaneko, I., Kimura, S., Mitamura, K., Shimada, K. (1991). Inhibition of human immunodeficiency virus type-1-induced syncytium formation and cytopathicity by complestatin. *Biochem. Biophys. Res. Commun.* 179, 243-250.
- [108] Vértesy, L., Aretz, W., Knauf, M., Markus, A., Vogel, M., Wink, J. (1999). Feglymycin, a novel inhibitor of the replication of the human immunodeficiency virus. Fermentation, isolation and structure elucidation. *J. Antibiot.* 52, 374-382.
- [109] a) Bunkóczi, G., Vértesy, L., Sheldrick, G.M. (2005). The antiviral antibiotic feglymycin: first direct-methods solution of a 1000+ equal-atom structure. *Angew. Chem.* 117, 1364-1366.  
b) Bunkóczi, G., Vértesy, L., Sheldrick, G.M. (2005). The antiviral antibiotic feglymycin: first direct-methods solution of a 1000+ equal-atom structure. *Angew. Chem. Int. Ed.* 44, 1340-1342.