

**Das Proteom eines halophilen Archaeons und seine  
Antwort auf Änderung der Lebensbedingungen**

—

**Inventarisierung, Quantifizierung und posttranslationale  
Modifikationen**

**Dissertation**

**Andreas Tebbe**



München 2005



Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

**Das Proteom eines halophilen Archaeons und seine Antwort auf  
Änderung der Lebensbedingungen –  
Inventarisierung, Quantifizierung und posttranslationale  
Modifikationen**

von

Andreas Tebbe

aus

Dortmund

München 2005



## **Erklärung**

Diese Dissertation wurde im Sinne von § 13 (3) der Promotionsordnung vom 29. Januar 1998 von Prof. Dr. D. Oesterhelt betreut.

### **Ehrenwörtliche Versicherung:**

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe angefertigt.

München, am 28.11.2005

.....  
Andreas Tebbe

Dissertation eingereicht am 29.11.2005

1. Berichterstatter: Prof. Dr. D. Oesterhelt
2. Berichterstatter: P.D. Dr. C. Eckerskorn

Tag der mündlichen Prüfung: 13.04.2006



# Inhaltsverzeichnis

<b>1. Einleitung</b> .....	<b>1</b>
1.1 Halophile Archaea .....	1
1.2 Zur Biologie von <i>Halobacterium salinarum</i> .....	2
1.3 Posttranslationale Modifikationen (in <i>H. salinarum</i> ).....	4
1.4 Das Genom von <i>Halobacterium salinarum</i> .....	5
1.5 Das Proteom - Proteomik.....	6
1.6 Trennmethode in der Proteomik.....	7
1.6.1 Elektrophoretische Trennmethode .....	7
1.6.2 Chromatographische Trennmethode.....	8
1.7 Massenspektrometrie in der Proteomik .....	9
1.7.1 Die Ionisation .....	11
1.7.2 Die Massenanalytoren .....	12
1.7.3 Vergleich der Komponenten und Kombinationsmöglichkeiten .....	15
1.8 Quantifizierung in der Proteomik.....	16
1.9 Phosphoproteom Analyse .....	18
<b>2. Ziel der Promotion</b> .....	<b>20</b>
<b>3. Material und Methoden</b> .....	<b>21</b>
<b>3.1 Material und Geräte</b> .....	<b>21</b>
3.1.1 Materialien / Chemikalien .....	21
3.1.2 Geräte .....	22
3.1.3 Programme .....	23
3.1.4 Medien und Lösungen .....	23
3.1.4 Abkürzungen.....	23
<b>3.2 Methoden</b> .....	<b>25</b>
3.2.1 Anzucht von <i>Halobacterium salinarum</i> .....	25
3.2.1.1 Anzucht der <i>Halobacterium</i> -Stammkultur.....	25
3.2.1.2 Anzucht der <i>Halobacterium</i> -Gebrauchskultur.....	25
3.2.2 Zellaufschlüsse .....	25
3.2.2.1 Zellaufschluss und Proteinpräparation mittels Taurodeoxycholat.....	25
3.2.2.2 Zellaufschluss mittels Sonifizierung in Basalsalz .....	26
3.2.2.3 Zellaufschluss mittels Sonifizierung in Harnstofflösung .....	26
3.2.3 Proteinbestimmungen.....	27
3.2.3.1 Proteinbestimmung nach Bradford .....	27
3.2.3.2 Proteinbestimmung mittels BCA.....	27
3.2.4 <i>In-vivo</i> Markierung mit <sup>32</sup> / <sup>33</sup> -P.....	27

3.2.5	ICPL Markierung von Proteinen.....	28
3.2.6	Gelelektrophorese von Proteinen .....	29
3.2.6.1	SDS-PAGE im Mini- und MEDI-Format.....	29
3.2.6.2	Zweidimensionale Gelelektrophorese (2-DE).....	30
3.2.6.3	IEF mittels IPG-Phor.....	31
3.2.6.4	Gießen von Gelen für die 2-DE .....	32
3.2.7	Proteinvisualisierung in SDS-Gelen .....	32
3.2.7.1	Coomassie Färbung .....	32
3.2.7.2	Silberfärbung nach Vorum.....	33
3.2.7.3	Immunoblotten von SDS-Gelen.....	33
3.2.8	Proteinverdau.....	34
3.2.8.1	Tryptischer In-Gel Verdau von Proteinen aus 2-D Gelen.....	34
3.2.8.2	Tryptischer Verdau aus eindimensionalen SDS-Gelen.....	35
3.2.8.3	In-Lösungsverdau von <i>H. salinarum</i> Proteinen .....	36
3.2.8.4	In-Lösungsverdau von BSA und $\beta$ -Casein .....	36
3.2.9	Entsalzung von Peptiden mittels Reversed Phase Material.....	37
3.2.9.1	Entsalzung in GeLoader Tips mittels $C_{18}$ Material.....	37
3.2.9.2	Entsalzung mittels- $C_{18}$ Kartuschen.....	37
3.2.9.3	Entsalzung in GeLoader Tips mittels Oligo R3-Material.....	37
3.2.10	Anreicherung von Phosphopeptiden .....	38
3.2.10.1	Anreicherung mittels SCX-Chromatographie .....	38
3.2.10.2	Anreicherung mittels $Fe^{3+}$ IMAC in GeLoader Tips.....	39
3.2.10.3	Anreicherung mittels $TiO_2$ Material.....	40
3.2.11	Massenspektrometrische Methoden.....	41
3.2.11.1	Automatische Dried-Droplet Probenpräparation .....	41
3.2.11.2	Automatische Probenpräparationmittels GyroLab <sup>TM</sup> .....	41
3.2.11.3	Dünnschichtpräparationauf Anker-Targets.....	42
3.2.11.4	MALDI-TOF PMF-Messung.....	42
3.2.11.5	Reinigung der MALDI-Targets.....	43
3.2.11.6	Packen von Nano-LC Säulen .....	43
3.2.11.7	Nano-LC ESI-MS/MS.....	44
3.2.11.8	Off-line ESI-MS/MS .....	44
3.2.11.9	Nano-LC MALDI MS/MS der ICPL Proben.....	45
3.2.12	Datenprozessierung.....	45
3.2.12.1	Mascot Recherche der PMF-Daten .....	45
3.2.12.2	Mascot Recherche der ICPL Datenund manuelle Nachanalyse .....	46
3.2.12.3	Mascot Recherche der LC ESI-MS/MS Daten .....	47
3.2.12.4	Mascot Recherche der LTQ FT-ICR Daten.....	47
3.2.12.5	Korrelationsanalyse der 2-DE Identifizierungen .....	48
3.2.12.6	Korrelationsanalyse der 2-DE Identifizierungen .....	48
3.2.12.7	Integration der Proteomik-Daten in das HALOLEX System.....	48

3.2.12.8 2-D Gel Analyse mittels Proteomweaver.....	49
3.2.12.9 Proteinstruktur Modellierung.....	49

## **4. Ergebnisse ..... 50**

### **4.1 Das zweidimensionale Referenzgel ..... 50**

4.1.1 Die cytosolische Standardprobe.....	51
4.1.2 Zweidimensionale Auftrennung des cytosolischen Proteoms .....	52
4.1.3 Massenspektrometrische Analyse der 2-D Referenzgele .....	54
4.1.4 GyroLab™ als alternative Probenvorbereitung.....	59
4.1.5 Genomannotation anhand von MS-Daten .....	61
4.1.6 Genomannotation anhand von MS-Daten und Korrelationsanalysen .....	63
4.1.7 Genominterpretation anhand von Proteom- und Genomdaten .....	66

### **4.2 Quantitative Proteomik ..... 69**

4.2.1 Vergleich von aerobem mit phototrophem Wachstum mittels 2-DE.....	71
4.2.2 Vergleich des Wachstums in Komplex- bzw. synthetischem Medium mittels 2-DE.....	72
4.2.3 Vergleich des Wachstums in Komplex- bzw. synthetischem Medium mittels LC-ICPL .....	73
4.2.3.1 Abschätzung der Scoregrenzen .....	74
4.2.3.2 Die identifizierten Proteine.....	75
4.2.3.3 Berechnung und Umrechnung der Regulation .....	77
4.2.3.4 Stark regulierte Proteine - der dynamische Bereich.....	77
4.2.3.5 Die regulierten Proteine.....	78
4.2.3.6 Zur Biologie der regulierten Proteine.....	82
4.2.4 Vergleich von aerobem mit phototrophem Wachstum mittels LC-ICPL.....	85
4.2.4.1 Die identifizierten Proteine.....	86
4.2.4.2 Die regulierten Proteine.....	87
4.2.4.3 Zur Biologie der regulierten Proteine.....	89
4.2.5 Gesamtbetrachtung der vergleichenden Analysen.....	91
4.2.5.1 Die identifizierten Proteine.....	91
4.2.5.2 Gesamtbetrachtung und Interpretation der biologischen Regulationen.....	94
4.2.6 Vergleich mit 2-DE ICPL Daten .....	101
4.2.6.1 Vergleich mit den regulierten Proteinen .....	102
4.2.7 Analysen mittels LC ESI-MS/MS .....	105

### **4.3 Posttranslationale Modifikationen ..... 106**

4.3.1 Immunochemischer Nachweis von Phosphoproteinen .....	106
4.3.2 Radioaktive <i>in-vivo</i> Markierung der cytosolischen Proteine .....	107
4.3.3 Analyse der Phosphorylierungsstellen .....	109
4.3.3.1 Vorfractionierung mittels 1-DE.....	110
4.3.3.2 Vorfractionierung mittels SCX-Chromatographie .....	112
4.3.3.3 Zur Biologie der identifizierten Phosphoproteine.....	116
4.3.4 Identifizierung N-terminaler Acetylierungen.....	121

<b>5. Diskussion .....</b>	<b>123</b>
<b>5.1 Das zweidimensionale Referenzgel .....</b>	<b>123</b>
5.1.1 Das cytosolische Proteininventar.....	123
5.1.2 Die automatisierte MALDI-TOF PMF Analyse.....	126
<b>5.2 Quantitative Proteomanalysen.....</b>	<b>128</b>
5.2.1 Quantitative Proteomik mittels 2-DE.....	128
5.2.2 Quantitative Proteomik mittels ICPL.....	129
5.2.2.1 <i>Die identifizierten Proteine</i> .....	131
5.2.2.2 <i>Die regulierten Proteine</i> .....	132
<b>5.3 Das identifizierte Proteininventar .....</b>	<b>135</b>
5.3.1 Vergleich der unterschiedlichen MS-Technologie .....	135
5.3.2 Abschätzung des Expressionsniveaus .....	137
<b>5.4 Analyse posttranslationaler Modifikationen .....</b>	<b>140</b>
5.4.1 Analyse phosphorylierter Proteine.....	140
5.4.2 Identifizierung von Phosphorylierungsstellen .....	141
5.4.3 Identifizierung N-terminaler Acetylierungen .....	143
<b>6. Ausblick .....</b>	<b>145</b>
<b>7. Zusammenfassung.....</b>	<b>146</b>
<b>8. Referenzen .....</b>	<b>148</b>
<b>9. Anhang .....</b>	<b>163</b>
<b>Danksagung.....</b>	<b>165</b>
<b>Veröffentlichungen.....</b>	<b>167</b>
<b>Lebenslauf .....</b>	<b>168</b>

# 1 Einleitung

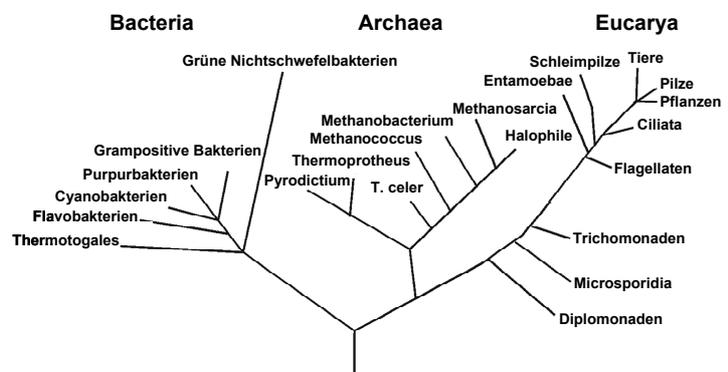
## 1.1 Halophile Archaea

Archaea bilden neben den Bacteria und Eukarya eine eigene monophyletische Gruppe der drei Domänen des Lebens [Woese, 1987; Woese *et al.*, 1990] (siehe Abb. 1-1). Der Name verdeutlicht, dass ihre Lebensbedingungen in zurückgezogenen ökologischen Nischen denen der frühgeschichtlichen Erde ähneln. Man findet sie ubiquitär unter extremen

Lebensbedingungen, wie bei Temperaturen über 100°C, pH-Werten unter 2 und über 11 sowie bei gesättigten Salzkonzentrationen, aber auch unter gemäßigten und normalen Lebensbedingungen. Trotz ihrer Erscheinungsform als einzellige Prokaryoten unterscheiden sie sich jedoch in vielerlei Hinsicht

von den Bakterien. Sie besitzen einen den Eukaryoten ähnlichen Transkriptions- und Translationsapparat, über Etherbindungen und nicht wie gewöhnlich über Esterbindungen verknüpfte Lipide sowie eine aus Glykoproteinen aufgebaute Zellwand (*S-Layer*) [Grabowski und Kelman, 2003; Allers und Mevarech, 2005].

Die halophilen und methanogenen Archaea bilden zusammen einen eigenen Ast des Urreiches der Archaea. Wie im Namen bereits angedeutet (halophil = salzliebend) benötigen halophile Archaea hohe Salzkonzentrationen für ein optimales Wachstum (>1.5 M NaCl). An ihnen lässt sich beispielhaft untersuchen, wie Leben unter diesen extremen Bedingungen möglich ist. Dafür bedarf es unterschiedlicher Strategien der Anpassung an den Lebensraum: Einerseits gleichen sie osmotisch die hohen Salzkonzentrationen der Umgebung durch hohe Kaliumchlorid-Konzentrationen im Zellinneren aus [Kushner, 1988; Kushner und Kamekura, 1988]. Zum anderen spiegelt sich die Anpassung an ihren Lebensraum in der Aminosäurekomposition ihrer Proteine wider. Diese enthalten im Gegensatz zu mesophilen Proteinen einen hohen Anteil an sauren Aminosäuren (Aspartate und Glutamate), die vornehmlich auf der



**Abb. 1-1:** Dreiteilung biologischer Organismen in die Domänen der Bacteria, Archaea und Eukarya (nach Woese *et al.*, 1990).

Proteinoberfläche die Löslichkeit und damit die für die biologische Funktion notwendige Struktur auch bei hoher Ionenstärke des umgebenen Mediums vermitteln [Britton *et al.*, 1998, Lanyi, 1974, Madern *et al.*, 2000]. Mit dieser Anpassung ist eine Erhöhung der Acidität halophiler Proteine verbunden [Lanyi, 1974; Madern *et al.*, 1995; Madern and Zaccai, 2000].

Halophile Archaea lassen sich in 10 Gattungen einteilen, wobei *Halobacterium*, *Haloferax*, *Halococcus*, *Halobaculum*, *Halorubrum* und *Haloarcula* bei neutralem pH wachsen, *Natronobacterium*, *Natronomonas*, *Natrialba* und *Natronococcus* hingegen ein alkalisches Milieu bevorzugen [Grant und Larsen, 1989; Kamekura, 1998, Falb *et al.*, 2005].

## 1.2 Zur Biologie von *Halobacterium salinarum*

*Halobacterium salinarum* (*H. salinarum*), welches zur Familie der *Halobacteria* gehört, besitzt ein optimales Wachstum bei 4 M NaCl. Unterhalb von ca. 2 M tritt die osmotische Lyse der Zellen ein. Die natürlichen Lebensräume von *H. salinarum* sind Salinen und hypersaline Seen wie das Tote Meer. Neben der Halophilie machen unter anderem seine einmalige Bioenergetik sowie seine Signaltransduktionsnetzwerke *H. salinarum* zu einem interessanten Objekt der biochemischen Forschung.

Neben dem Wachstum unter Standardbedingungen bei ausreichendem Angebot an Sauerstoff und organischen Nährstoffen (oxidative Phosphorylierung) besitzt *H. salinarum* unter anaeroben Wachstumsbedingungen, die in seiner natürlichen Umgebung überwiegend herrschen, drei alternative Wege zur Energiegewinnung: Zum einen können Dimethylsulfoxid oder Trimethylammoniumoxid als terminale Elektronenakzeptoren in der Atmungskette dienen [Oren und Truper, 1990; Oren, 1994]. Zum anderen kann Arginin zu Ornithin unter Bildung von Carbamoylphosphat und schließlich ATP fermentiert werden [Hartmann *et al.* 1980; Hochstein und Lang 1991; Ruepp *et al.*, 1995]. Die dritte Möglichkeit zur anaeroben Energiegewinnung ist die zweite Art der in der Natur vorkommenden Photosynthese. Dieser Prozess basiert nicht auf Chlorophyll-enthaltenden Reaktionszentren wie in photosynthetisch aktiven Bakterien oder Pflanzen, sondern auf Retinal-enthaltenden Proteinen als Energiewandler [Oesterhelt, 1998]. Diese Umwandlung beruht auf der thermoreversiblen Photoisomerisierung des Chromophors von der all-*trans* in die 13-*cis* Form.

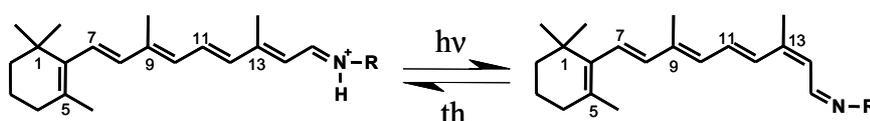


Abb. 1-2: Photoisomerisierung von all-*trans* zu 13-*cis* Retinal.



siehe Abb. 1-3). Neben den beiden Photorezeptoren existieren 16 weitere Rezeptoren, welche unterschiedliche Umwelteinflüsse wahrnehmen können [Hou *et al.*, 2000; Storch *et al.*, 1999; Kokoeva und Oesterhelt, 2000; Kokoeva *et al.*, 2002; Koch und Oesterhelt, 2005] und es dem Organismus so erlauben, die optimalen Wachstumsbedingungen zu suchen.

Zusätzlich zu dem eben beschriebenen Bewegungssystem kann eine Anpassung an äußere Lebensbedingungen durch 13 Sensor Kinasen und drei Response Regulatoren, welche die Genexpression regulieren, erfolgen.

### 1.3 Posttranslationale Modifikationen (in *H. salinarum*)

Posttranslationale Modifikationen (PTM's) spielen eine entscheidende Rolle in der Regulation von Proteinaktivitäten sowie in der Signalweiterleitung in Signaltransduktionsnetzwerken. Eine der wichtigsten Modifikationen von Proteinen ist dabei die reversible Phosphorylierung bestimmter Aminosäuren [Hunter, 2000]. Hierbei überträgt eine Proteinkinase eine Phosphatgruppe auf vornehmlich Serin-, Threonin- oder Tyrosin-Reste. Als Phosphatgruppen-Donor dient ATP. Wie später beschrieben können noch andere Aminosäuren als Phosphatgruppenakzeptoren auftreten. Die Entfernung der Phosphatgruppe und somit die Überführung des Proteins in seine Ausgangsform geschieht durch Proteinphosphatasen.

Entdeckt wurde die regulative Proteinphosphorylierung Mitte der 1950er Jahre von Edwin G. Krebs und Edmond H. Fischer am Beispiel der Phosphorylase, einem Enzym des Glycogenabbaus. Ausgehend von dieser fundamentalen Entdeckung, für die beide 1992 den Nobelpreis für Medizin erhielten, wurden in den darauf folgenden Jahren eine Vielzahl durch Proteinphosphorylierung regulierter Prozesse entdeckt. Im Zuge der fortschreitenden Genomsequenzierungen stieß man auf weitere Proteinkinasen (im humanen Genom 518) und Phosphatasen, was deren zentrale Rolle als Signalvermittler weiter unterstreicht [Johnson und Hunter, 2005; Pawson und Scott, 2005]. Man geht heute davon aus, dass jedes dritte humane Protein durch Phosphorylierung modifiziert und so reguliert werden kann. Beispiele dafür sind die Signaltransduktions-Netzwerke höherer Organismen zur Regulation der Genexpression wie der MAP-Kinase [Blumer und Johnson, 1994] Weg oder die Kontrolle des Zellzyklus [Morgan, 1995].

Man ging lange davon aus, dass die reversible Phosphorylierung durch Proteinkinasen und Proteinphosphatasen eine späte evolutionäre Entwicklung sei, die sich auf eukaryotische Lebensformen beschränke. Dieses Paradigma wurde jedoch widerlegt als gezeigt werden konnte, dass die katalytische Aktivität der Isocitrat Dehydrogenase aus *E. coli* durch Phosphorylierung eines Serin Restes reguliert wird

[Garnak und Reeves, 1979]. Durch die stetig wachsende Anzahl an genomischen Sequenzinformationen konnten in Prokaryoten und Archaea Homologe zu den in den Eukaryoten vorkommenden Proteinkinasen und Phosphatasen identifiziert und für einige deren Aktivität nachgewiesen werden [Masumoto *et al.*, 1994; Kennelly, 2002].

Ein weiteres Beispiel der Signalweiterleitung in Prokaryoten durch Phosphorylgruppen transfer sind die so genannten Zwei-Komponenten Systeme [Parkinson, 1993]. Diese werden hier am Beispiel der bereits in Abschnitt 1.2 eingeführten halobakteriellen Phototaxis beschrieben (siehe Abb. 1-3): Durch Aktivierung der Sensorrhodopsine (SR I und II) und deren nachgeschalteten Signalkoppler (Htr I und II) kommt es zur Aktivierung einer Proteinkinase (CheA). Hierbei handelt es sich um eine Histidinkinase, welche sich zunächst an einem Histidinrest autophosphoryliert und anschließend das Phosphat auf einen Aspartat-Rest des Antwort-Regulators (CheY) überträgt. Dieser bindet im aktivierten, phosphorylierten Zustand an den Flagellarmotor und steuert dort die Rotationsrichtung der Flagellen. In der eben beschriebenen Signalkaskade tritt noch eine weitere wichtige Form der posttranslationalen Modifizierung auf. Nach Methylierung des Signalkopplers durch eine aktivierte Methyltransferase (CheB) erfolgt das Abschalten des Sensorrhodopsin-Signalkoppler-Komplexes und somit die Adaptation an einen zeitlich konstant bleibenden Reiz [Alam *et al.*, 1989; Marwan *et al.* 1995].

Die Strategie der regulativen Proteinphosphorylierung ist also bereits auf der Entwicklungsstufe der Prokaryoten, wenn auch in geringerem Ausmaß in Vergleich zu höheren Organismen, anzutreffen. Inwieweit diese, neben der Funktion in Zwei-Komponenten Systemen, weitere regulative Funktionen in *H. salinarum* übernimmt, ist bisher noch unklar.

## 1.4 Das Genom von *Halobacterium salinarum*

Um einen besseren Überblick über die Gesamtphysiologie von *H. salinarum* zu erhalten, wurde sein Genom (R1, DSM 671) vollständig sequenziert [www.halolex.mpg.de]. Es besteht aus einem 2 Mbp großen Chromosom (siehe Abb. 1-4) und den 4 Megaplasmiden pHS1 (148 kbp), pHS2 (195 kbp), pHS3 (284 kbp) und pHS4 (41 kbp) (siehe dazu auch Abb. 4-14) wobei pHS3 eher als zweites Chromosom angesehen werden sollte (siehe 4.1.7). Das Genom gleicht besonders

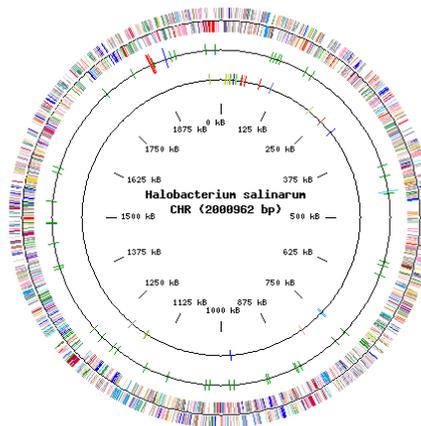


Abb. 1-4: Genom Karte des Chromosoms von *H. salinarum*.

in seinen chromosomalen Bereichen (Sequenzidentität nahezu 100%) dem des Laborstamms NRC-1 [Ng *et al.*, 2000]. Es besitzt einen hohen GC-Gehalt von 68% und einen starken Dinukleotid-Bias: CG>GC; AT>TA. Dies hat eine geringe Anzahl an Stoppcodons sowie eine sich daraus ergebende starke Diskrepanz von offenen Leseramen (ORFs mit über 100 Codons) und annotierten Genen zur Folge. Derzeit wird für 2821 ORFs, welche bereits 90% des Genoms ausmachen, angenommen, dass sie für reelle Proteine codieren (Stand Ende 2005; www.halolex.mpg.de).

Es bedarf also einer zusätzlichen Validierung, um eine genaue Annotation des Genoms von *H. salinarum* vorzunehmen. Einen Ansatz hierbei liefert die gezielte Identifizierung aller Genprodukte auf der Ebene der Proteine, um so bestimmte ORFs als Gene zu bestätigen oder auszuschließen.

## 1.5 Das Proteom - Proteomik

Der Begriff Proteom wurde erstmals von Marc Wilkins auf dem Siena-Meeting 1994 eingeführt. Er steht in Analogie zu dem Begriff Genom (Gesamtheit aller Gene eines Organismus) und bezeichnet die Gesamtheit aller in einem Organismus zu einem bestimmten Zeitpunkt vorkommenden Proteine. Im Gegensatz zum statischen Genom – ein Organismus hat immer die gleiche Anzahl an Genen, unabhängig vom Zell- oder Gewebetyp – ist das Proteom ein flexibler Zustand, der sich abhängig von Zellzustand oder Gewebetyp in höheren Organismen stark unterscheiden kann. Die Kenntnis über die Basensequenzen in einem Genom gibt nur einen ersten Überblick über das funktionelle Potential eines Organismus. Um diesen auf molekularer Ebene zu verstehen, muss man sich auf die Ebene der Proteine, d.h. der Funktionsträger einer Zelle, begeben und diese in ihrer Gesamtheit studieren. So sind z.B. die für die Funktion oder Regulation eines Proteins notwendigen posttranslationalen Modifikationen wie Phosphorylierungen (siehe 1.9) nicht unbedingt aus der Genomsequenz ableitbar. Zudem sind derartige Modifikationen oft von exogenen Parametern (Temperatur, Stress etc.) abhängig und treten daher nur temporär auf.

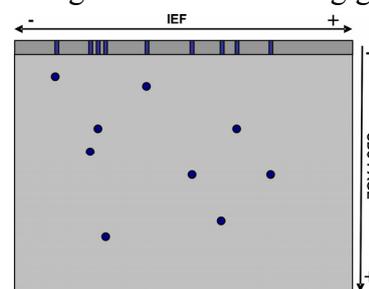
Die Proteomik (Proteomics) beschäftigt sich nun mit der Auftrennung komplexer Proteingemische sowie deren anschließender Analyse. Im Folgenden werden die methodischen Grundlagen der in dieser Arbeit verwendeten Techniken vorgestellt.

## 1.6 Trennmethode in der Proteomik

Man kann generell zwischen zwei Ansätzen der Probentrennung unterscheiden: Zum einen kann die Trennung des komplexen Proteingemisches (z.B. Zellextrakt) auf der Ebene der Proteine geschehen. Hierbei werden vor allem elektrophoretische Verfahren wie die zweidimensionale Gelelektrophorese eingesetzt [Görg *et al.*, 2004; Lottspeich, 1999]. Die aufgetrennten Proteine werden so der nachfolgenden Analytik zugänglich. Ein anderer Ansatz beruht auf der Trennung der Probe nach proteolytischem Verdau des Proteingemisches, also auf Peptidebene (*shotgun proteomics*) [McDonald and Yates 3rd, 2003; Swanson and Washburn, 2005]. Dieser Ansatz macht sich die Tatsache zu Nutze, dass die anschließende massenspektrometrische Analytik, d.h. die Identifizierung einzelner Proteine, in der Regel über seine proteolytischen Peptidfragmente geschieht. Dieser gehen chromatographische Trennungen des Peptidgemisches voraus.

### 1.6.1 Elektrophoretische Trennmethode

Proteine weisen zwitterionischen Charakter auf und können so elektrophoretisch, d.h. in einem elektrischen Feld, aufgetrennt werden. Ihr Ladungszustand ist abhängig vom pH-Wert des umgebenden Mediums. In der zweidimensionalen Gelelektrophorese (2-DE) werden zwei elektrophoretische Trennmethode miteinander kombiniert (siehe Abb. 1-5): In der ersten Dimension, der isoelektrischen Fokussierung (IEF), werden Proteine in einem in einer Gelmatrix ausgebildeten pH-Gradienten aufgetrennt. Dabei wandern die einzelnen Proteine einer Mischung in die Zone, deren pH-Wert dem isoelektrischen Punkt ( $pI$ ) des jeweiligen Proteins entspricht. Hier besitzen die Proteine ebenso viele positive wie negative Ladungen, d.h. ihre Nettoladung ist Null. Somit verlieren sie ihre elektrophoretische Mobilität – sie sind *fokussiert*. Die IEF ist eine konzentrierende Endpunktmethode, bei der sehr scharf begrenzte, hochkonzentrierte Proteinbanden entstehen. In der zweiten Dimension werden die fokussierten Proteine einer SDS-Polyacrylamid Gelelektrophorese (SDS-PAGE) unterzogen. Dabei werden die in der IEF-Matrix fokussierten Proteine mit dem anionischen Detergenz Natrium-Dodecylsulfat (SDS) beladen, auf ein Polyacrylamidgel gelegt und die Proteine in anodische Richtung elektrophoretisch nach ihrer Größe (d.h. Molekularmasse) getrennt. Zur nachfolgenden Visualisierung der aufgetrennten Proteine stehen unterschiedliche Färbemethode wie



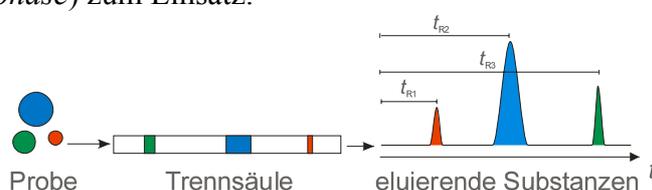
**Abb. 1-5:** Schematischer Aufbau der 2-DE.

die Coomassie- oder Silberfärbung sowie auf Fluoreszenz basierende Farbstoffe zur Verfügung.

Die 2-DE basiert auf den grundlegenden Arbeiten von Klose und O'Farrel [Klose, 1975; O'Farrel, 1975] und wurde im Laufe der Jahre stetig verbessert. Besonders die Einführung immobilisierter pH-Gradienten (IPG-Technologie) [Bjellqvist, 1982] führte zu einer erheblichen Verbesserung der Reproduzierbarkeit und machte sie so der breiten Anwendung zugänglich [Görg *et al.*, 2004]. Der Notwendigkeit eines hohen Probendurchsatzes zur parallelen Analyse unterschiedlicher Proben wurde durch die Entwicklung der entsprechenden Robotik Rechnung getragen [Nordhoff *et al.*, 2001]. Die 2-DE ist bis heute die Methode mit dem größten Auflösungsvermögen komplexer Proteingemische und wird als Standardmethode in der Proteomforschung eingesetzt. Über das Erstellen von *Proteom Referenz Karten* mittels 2-DE können Proteommuster analysiert und ausgewertet werden, wofür jedoch eine hohe Reproduzierbarkeit der Gele absolute Voraussetzung ist [Buttner *et al.*, 2001; Hecker und Völker, 2004; Giavalisco *et al.*, 2005].

### 1.6.2 Chromatographische Trennmethode

Der 2-DE stehen chromatographische Trennmethode gegenüber [Shi *et al.*, 2004]. Bei diesem Ansatz geht man jedoch meist nicht von dem *intakten* Proteingemisch aus, sondern spaltet dieses proteolytisch (z.B. mit Trypsin) in seine Peptidfragmente. Dies hat zum einen den Vorteil, dass Peptide aufgrund ihrer einheitlicheren physikochemischen Eigenschaften im Gegensatz zu Proteinen leichter zu handhaben sind. Zum anderen erhöht sich jedoch die Komplexität des Analyten immens. Ein Gemisch aus *nur* 1,000 Proteinen würde nach tryptischer Spaltung, bei durchschnittlich 30 proteolytischen Fragmenten pro Protein, eine Mischung aus 30,000 unterschiedlichen Peptidspezies ergeben. Um eine derart komplexe Mischung für die nachfolgende Analytik in ihre einzelnen Bestandteile aufzutrennen, bedient man sich chromatographischer Trennmethode – der Flüssigchromatographie (LC = *liquid chromatography*). Unter Chromatographie versteht man die Auftrennung von Analyten gelöst in einer mobilen (flüssigen) Phase durch ihre Wechselwirkung mit einer stationären Phase (spezifische Oberfläche). Bei der Auftrennung von Peptidgemischen kommen hauptsächlich die starke Kationenaustausch-Chromatographie (SCX-Chromatographie = *strong cation exchange*) sowie die Umkehrphasen-Chromatographie (RP-Chromatographie = *reversed phase*) zum Einsatz.



**Abb. 1-6:** Schematische Darstellung einer chromatographischen Auftrennung.  $t_{R1}$  = Elutionszeit Komponente 1,  $t_{R2}$  = Elutionszeit Komponente 2,  $t_{R3}$  = Elutionszeit Komponente 3.

Die Ionenaustausch-Chromatographie beruht auf der kompetitiven Wechselwirkung der geladenen Analyten (Peptide) mit der ebenfalls geladenen stationären Phase entgegen gesetzter Polarität. Im Sauren weisen Peptide einen kationischen Charakter auf und binden so an die geladenen funktionellen Gruppen der stationäre Phase. Je stärker das Peptid geladen ist, desto stärker bindet es. Die Fraktionierung erfolgt durch die Elution der gebundenen Analyten mit einem stetigen Anstieg der Salzkonzentration in der mobilen Phase. Die Peptide werden bei entsprechender Konzentration eines konkurrierenden Gegenions von ihrer Bindungsstelle verdrängt und werden so wieder mobil.

Die Umkehrphasen-Chromatographie beruht hingegen auf hydrophoben Wechselwirkungen zwischen den Analyten mit der unpolaren stationären Phase. Unter wässrigen Bedingungen binden die Peptide und können durch einen steigenden Gehalt an organischem Lösungsmittel der mobilen Phase mit geringer Polarität, wie Methanol oder Acetonitril, eluiert werden. Die RP-Chromatographie wird oft als letzte Auftrennung direkt vor der MS eingesetzt, da hier keine Salze in den Eluenten verwendet werden müssen, die die massenspektrometrische Analyse erheblich beeinträchtigen würden.

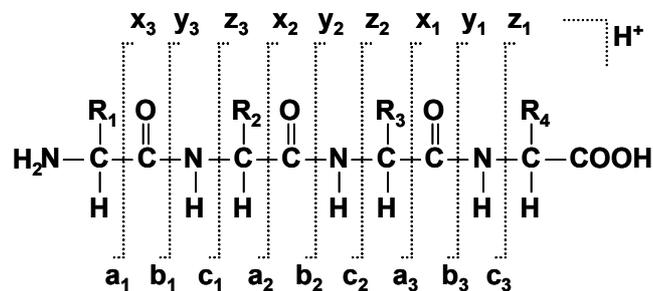
Chromatographische Trennmethode sind in ihrer Auflösung jedoch begrenzt. Um hochkomplexe Peptidgemische zu analysieren bedarf es daher entweder einer Kombination der chromatographischen Techniken wie der multidimensionalen Chromatographie [Link *et al.*, 1999; Peng *et al.*, 2003] oder einer Vorfraktionierung der anfänglichen Proteinprobe z.B. durch elektrophoretische Verfahren. Ein Vorteil der chromatographischen Methoden sind eine leichte Automatisierbarkeit und damit verbundene kürzere Analysezeiten gegenüber der 2-DE. Des Weiteren ist eine verbesserte Anbindung an die nachfolgende massenspektrometrische Analyse der aufgetrennten Peptide möglich.

## 1.7 Massenspektrometrie in der Proteomik

Ein zentraler Punkt der Proteomforschung ist die Proteinidentifizierung mittels Massenspektrometrie (MS), welche sich in den letzten Jahren als Analysemethode etabliert und die herkömmlichen Methoden wie N-terminale Sequenzierung oder immunologische Nachweisverfahren weitgehend ersetzt hat [Mann und Aebersold, 2003]. Die rasanten Fortschritte auf diesem Gebiet haben die Etablierung der Proteomik entscheidend mitgeprägt, wenn nicht gar erst ermöglicht. Generell versteht man unter Massenspektrometrie die Bestimmung der Molekülmasse im Hochvakuum. In der Proteomik wird in der Regel ein Protein jedoch nicht direkt massenspektrometrisch analysiert, sondern seine durch proteolytische Spaltung mit einer Protease (meist

Trypsin) entstandenen Peptide. Der Verdau kann nach der elektrophoretischen Auftrennung (siehe 1.6.1) direkt im Gel erfolgen und die entstandenen Peptide aus dem Gel eluiert werden [Shevshenko *et al.*, 1996]. Werden einzelne Proteinspezies analysiert, wie z.B. nach der Auftrennung mittels 2-DE, kann die Identifizierung des Proteins mit der *peptide mass fingerprint* (PMF) Methode erfolgen [Thiede *et al.*, 2005]. Hierbei wird das aufgenommene Massenspektrum der proteolytischen Peptide eines Proteins (der so genannte Peptid-Fingerabdruck) mit den theoretischen Massenspektren aller vorhergesagten Proteine eines Genoms verglichen. Bei genügender Übereinstimmung von experimentellem und theoretischem Massenspektrum kann eine eindeutige Zuordnung des Proteins zu einem Gen erfolgen [Pappin *et al.*, 1993; Yates *et al.*, 1993; Mann *et al.*, 1993; Gaevvert und Vandercherkhove, 2000].

Eine in der Proteomik immer mehr an Bedeutung gewinnende Methode ist die Kombination einer chromatographischen Auftrennung eines komplexen Peptidgemisches mit der Tandem Massenspektrometrie (MS/MS). Hierbei wird nach der Aufnahme des Massenspektrums (MS) das zu analysierende Peptid (in einer mit Kollisionsgas, meist Ar oder N<sub>2</sub>, gefüllten Stoßzelle) zusätzlich fragmentiert (CID = *collision induced fragmentation*). Die Fragmentierung geschieht vornehmlich entlang der Peptidbindung (siehe Abb. 1-7) in definierte Fragmentationen [Roepstorff und Fohlman, 1984; Johnson *et al.*, 1987; McLuckey, 1992]. Aus dem entstandenen Fragmentspektrum (MS/MS) lässt sich Sequenzinformation über das Peptid gewinnen, die als zusätzliches Identifizierungskriterium beim Abgleich mit der Datenbank herangezogen werden kann [Mann und Wilm, 1994; Steen und Mann, 2004]. Bei ausreichender Qualität des Fragmentspektrums ist hier im Gegensatz zur PMF-Methode die Identifizierung des Proteins anhand eines einzelnen Peptids innerhalb eines komplexen Peptidgemisches möglich (*shotgun proteomics*).

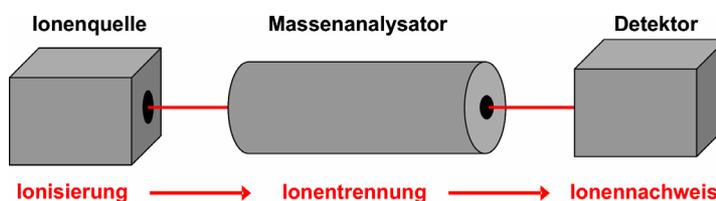


**Abb. 1-7:** Nomenklatur der wahrscheinlichsten Fragmentationen, die bei der kollisionsinduzierten Peptidfragmentierung auftreten.

Zusätzlich kann aus dem Tandem Massenspektrum durch die Analyse der entstehenden Fragmentserien (x, y, z bzw. a, b, c-Serie) z.B. die genaue Lokalisation von Peptidmodifikationen abgeleitet werden.

Ein Massenspektrometer besteht generell aus drei unterschiedlichen, voneinander unabhängigen Komponenten:

Der Ionenquelle, dem Massenanalysator und dem Detektor (siehe Abb. 1-8).



Für alle drei Komponenten gibt es

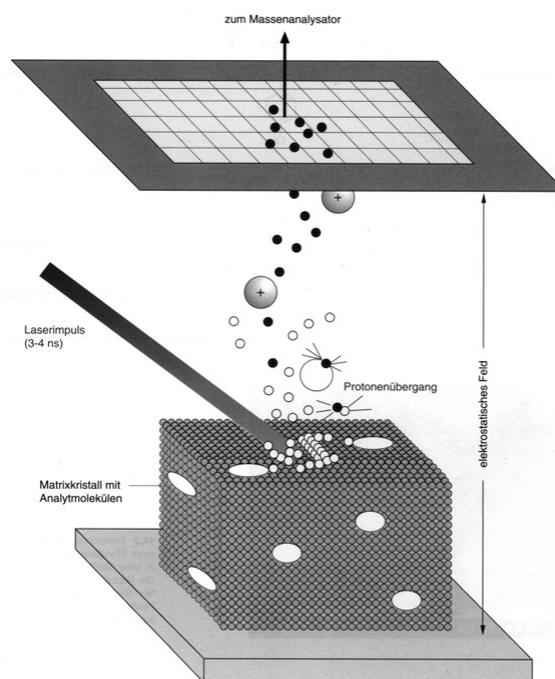
**Abb. 1-8:** Schematischer Aufbau eines Massenspektrometers.

unterschiedliche technologische Ansätze, die nahezu beliebig miteinander kombiniert werden können. Die Komponenten der in dieser Arbeit verwendeten Massenspektrometer werden im Folgenden kurz beschrieben.

### 1.7.1 Die Ionisation

Voraussetzung für die MS-Analytik ist die Ionisierung des Analyten. Hierbei haben sich zwei generell unterschiedliche so genannte „weiche“ Ionisationsmethoden etabliert, welche die Analyse von Biopolymeren wie Proteinen oder Peptiden erst ermöglichten. Für die im Folgenden vorgestellten Verfahren, welche Mitte der 1980er Jahre entwickelt wurden, erhielten John B. Fenn und Koichi Tanaka 2002 den Nobelpreis für Chemie.

Bei der *Matrix-assisted Laser Desorption/Ionization*-Methode (MALDI) werden die Analytpeptide mit einer meist aromatischen Substanz (Matrix) auf einem Probenhalter im Sauren kokristallisiert. Durch Laserbeschuss erfolgt die Einkopplung der für die Ionisierung erforderlichen Energie [Karas *et al.*, 2000; Karas und Krüger, 2003]. Die Beschleunigung der ionisierten Analyten

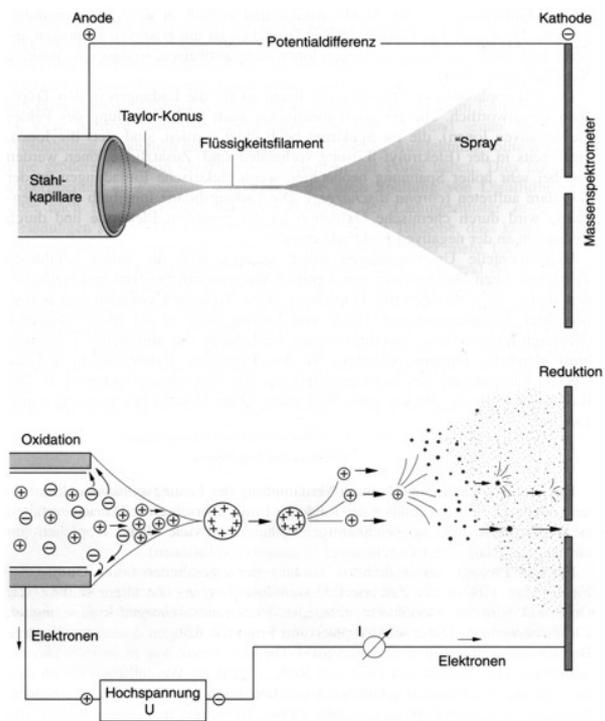


**Abb. 1-9:** Prinzip des MALDI-Prozesses (aus BIOANALYTIK; Lottspeich, Zorbas).

erfolgt durch das Anlegen eines elektrischen Feldes zwischen Probenhalter und einer der Probe gegenüberliegenden Elektrode (siehe Abb. 1-9). Die so ionisierten und

beschleunigten Ionen können nun in den Massenanalysator des Massenspektrometers eintreten.

Die zweite Ionisationsmethode beruht auf der Dispersion einer Flüssigkeit in sehr viele kleine geladene Tröpfchen mit Hilfe eines elektrostatischen Feldes (siehe Abb. 1-10). Bei dieser so genannten *Elektrospray-Ionisations-Methode* (ESI) werden die in Lösung befindlichen und ionisierten Analyten direkt in das Massenspektrometer gesprüht [Fenn *et al.*, 1989; Fenn, 2003]. Der genaue Ionisationsprozess ist bis heute noch nicht bis in alle Einzelheiten verstanden. Man nimmt an, dass die durch Verdampfen des Lösungsmittels immer kleiner werdenden Tropfen bei Erreichen einer bestimmten Ladungsdichte in viele kleine Tröpfchen von wenigen Nanometern Durchmesser zerfallen (Coulomb-Explosion). Dieser Prozess führt letztlich zu freien Gasphasen-Ionen, die in das Hochvakuum des Massenspektrometers eingesogen werden.



**Abb. 1-10:** Schematische Darstellung des makroskopischen (oben) und mikroskopischen (unten) ESI-Prozesses (aus BIOANALYTIK; Lottspeich, Zorbias).

### 1.7.2 Die Massenanalytoren

Bei der Massenanalyse haben sich folgende Technologien bewährt. Bei dem für die MALDI-MS überwiegend eingesetzten Massenanalysator handelt es sich um einen Flugzeitmassenanalysator (TOF = *time of flight*). Hier erfolgt die Massenbestimmung über die Messung der Zeit, die die Ionen zum Zurücklegen einer definierten Strecke (Driftstrecke) benötigen. Da die Ionen (mit einer Masse  $m$ , der Geschwindigkeit  $v$ , einer Ladungszahl  $z$  und der Elementarladung  $e$ ) in einem konstanten elektrischen Feld ( $U$ ) beschleunigt werden, besitzen sie alle die gleiche kinetische Energie:

$$E_{kin} = \frac{1}{2} \cdot m \cdot v^2 = z \cdot e \cdot U \quad \text{Gl. 1-1}$$

Bei bekannter Beschleunigungsspannung  $U$  und Flugstrecke  $L$  lässt sich aus der Messung der Flugzeit  $t$  (mit  $v = L/t$ ) das Verhältnis aus Masse und Ladung ( $m/z$ ) bestimmen:

$$\frac{m}{z} = \frac{2 \cdot e \cdot U}{L^2} \cdot t^2 \quad \text{Gl. 1-2}$$

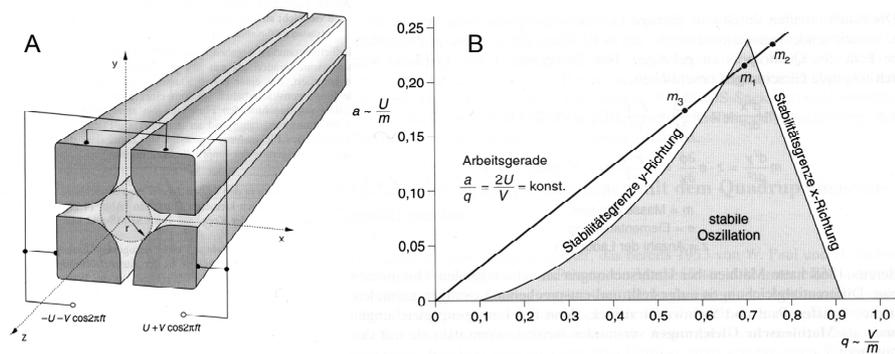
Die durch den Elektrospray-Prozess erzeugten Ionen wurden bis vor kurzem hauptsächlich durch zwei technologische Ansätze analysiert und detektiert. Ein gebräuchlicher Massenanalysator ist der sogenannte Quadrupol, welcher als Massenfilter eingesetzt wird. Er besteht aus vier hyperbolisch geformten, stabförmigen Elektroden, die auf einem Kreis mit dem Radius  $r$  um die  $z$ -Achse angeordnet sind (siehe Abb. 1-11 A). An den Stäben liegt eine Gleichspannung  $U$  und eine Wechselspannung ( $V \cdot \cos 2\pi vt$ ) mit der Frequenz  $v$  an, wobei gegenüberliegende Stäbe die gleiche Polarität der Gleichspannung und die gleiche Phase der Wechselspannung aufweisen. In der Nähe der  $z$ -Achse entsteht somit ein elektrisches Potential  $\Phi$ :

$$\Phi(x, y, t) = (U + V \cdot \cos 2\pi vt) \cdot \frac{x^2 - y^2}{r^2} \quad \text{Gl. 1-3}$$

Die Bewegung auf stabilen Bahnen der beschleunigten Ionen durch einen Quadrupol wird durch die Lösung der Mathieuschen Gleichungen beschrieben:

$$\frac{a}{q} = \frac{2zeU}{m(\pi vt)^2} \frac{m(\pi vt)^2}{zeV} = \frac{2U}{V} \quad \text{Gl. 1-4}$$

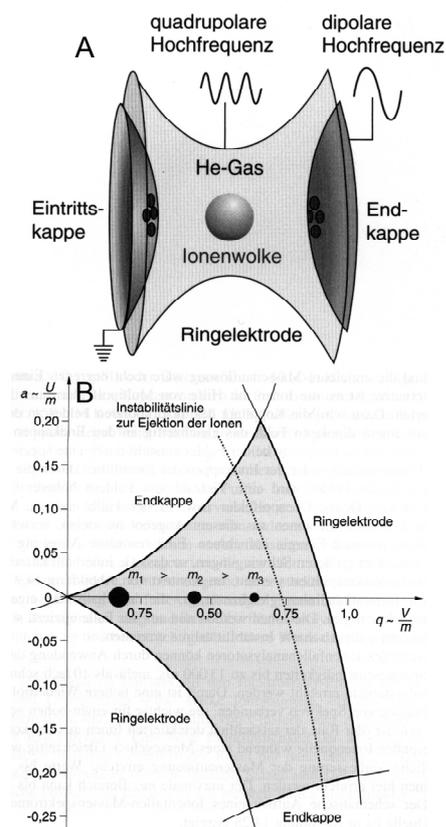
Mit den Parametern  $a$  und  $q$  wird die Beziehung zwischen einem zu transferierenden Ion der Masse  $m$  mit  $z$  Elementarladungen  $e$  und den Eigenschaften des Quadrupols festgelegt. Damit liegen bei gleichem  $z$  alle Massen auf derselben Geraden  $a/q = \text{konstant}$ , der sogenannten Arbeitsgeraden (siehe Abb. 1-11 B). Durch das Verändern der Gleichspannung  $U$  und der Amplitude des Wechselfeldes können Massenbereiche gescannt werden, da so Ionen unterschiedlicher Massen nacheinander in die stabilen Bereiche des Quadrupols gebracht werden können.



**Abb. 1-11:** (A) Elektrodenform und Elektrodenanordnung in einem Quadrupol Massenfilter. (B) Stabilitätsdiagramm der Mathieschen Gleichungen für das zweidimensionale Quadrupol (aus BIOANALYTIK; Lottspeich, Zorbas).

Durch das Hintereinanderschalten mehrerer Quadrupole, bei denen der erste als Massenfilter, der zweite als mit einem Stoßgas gefüllte Kollisionszelle und der dritte Quadrupol als Massenanalysator fungiert, können so nacheinander MS sowie MS/MS Spektren ausgewählter Vorläuferionen aufgenommen werden.

Alternativ zu den Quadrupolanalysatoren kann zur Massenanalyse eine sogenannte Ionenfalle verwendet werden, welche als dreidimensionaler Quadrupol angesehen werden kann. Für diese Entwicklung erhielt Wolfgang Paul zusammen mit Hans G. Dehmelt und Norman F. Ramsey 1989 den Nobelpreis für Physik. Die Ionenfalle besteht aus einer Ringelektrode und zwei Endkappen, an die Wechselspannungen angelegt werden (siehe Abb. 1.12 A). Die Funktion der Ionenfalle basiert im Wesentlichen auf dem gleichen Prinzip wie der Quadrupol. Auch hier bestimmt die Lösung der Mathieschen Differentialgleichungen die Wertebereiche von angelegter Gleich- und Wechselspannung, in denen Ionen stabile Bahnen beschreiben (siehe Abb. 1.12 B). Die Ionen können für variable Zeiten (bis Sekunden) auf diesen stabilen Bahnen gehalten und analysiert werden.



**Abb. 1-12:** Prinzipieller Aufbau (A) und Stabilitätsdiagramm der Mathieschen Gleichungen (B) einer Ionenfalle.

Beim Messvorgang werden zunächst Ionen eines breiten Massenbereichs in der Falle akkumuliert und zur Detektion mit Hilfe eines Multipolfeldes mit ansteigendem Molekulargewicht aus der Falle ejiziert. Für die MS/MS-Analyse können bis auf das gewünschte Vorläuferion alle anderen Ionen aus der Falle *geworfen* werden. Das Verbliebene wird durch Schwingungsanregung fragmentiert und die entstehenden Fragmente zur Detektion wiederum aus der Falle ejiziert.

Für die Detektion der Ionen der unterschiedlichen Massenspektrometer stehen Sekundärelektronenvervielfacher zur Verfügung, welche die auftreffenden Teilchen in einen messbaren Strom verstärken.

Eine weitere Methode ist die Fourier-Transform-Ionenzyklotron-Resonanz-Analyse (FT-ICR) [Amster, 1996; Marshall *et al.* 1998]. Hier werden die Ionen in einem starken konstanten Magnetfeld auf stabilen Kreisbahnen eingefangen. Dabei bewegen sich Ionen unterschiedlicher Masse mit unterschiedlicher Kreisfrequenz im Magnetfeld. Bei Einstrahlung eines elektromagnetischen Wechselfelds mit einer Frequenz, die der Kreisfrequenz der rotierenden Ionen entspricht, nehmen diese Energie auf. Diese Energieaufnahme wird in Abhängigkeit von der Frequenz gemessen und daraus der  $m/z$ -Wert der Ionen bestimmt.

### 1.7.3 Vergleich der Komponenten und Kombinationsmöglichkeiten

Die MALDI Ionisation wird überwiegend mit TOF Analysatoren kombiniert (MALDI-TOF MS) und so zur standardmäßigen Identifizierung von mittels 2-DE aufgetrennten Proteinen mit der PMF-Methode verwendet, da diese robuste und einfache instrumentelle Kombination eine weitgehende Automation ermöglicht. Die neueste Generation der MALDI TOF Massenspektrometer ermöglicht zusätzlich das Generieren von Fragmentspektren (MS/MS) in hohem Probendurchsatz. Dies erlaubt die Analyse komplexer Peptidgemische durch Kopplung mit chromatographischen Trennverfahren (*off-line* Methode) [Ericson *et al.*, 2003; Chen *et al.*, 2005; Migorodskaya *et al.*, 2005]. Dabei werden die chromatographischen Fraktionen mit der Matrix gemischt, auf das Probenziel aufgetragen und können so anschließend einzeln analysiert werden. Ein Vorteil der HPLC-Kopplung mit MALDI-TOF Massenspektrometern ist, dass die Probe nicht kontinuierlich in das Massenspektrometer eluiert und so die Aufnahme von Massen- und Fragmentspektren im Gegensatz zur LC-ESI-MS/MS zeitunabhängig ist.

Die ESI MS erlaubt hingegen die direkte Kopplung von Chromatographie und Massenspektrometrie (LC-ESI-MS). Hier können die von der Säule eluierenden Peptide direkt in das Massenspektrometer gesprüht werden [Ducret *et al.*, 1998; Washburn *et al.*, 2001; Ishihama, 2005]. Im Gegensatz zur Kopplung der LC mit der MALDI-TOF MS handelt es sich bei der LC-ESI MS um ein kontinuierliches (*on-line*) Verfahren, d.h. die

von der Säule in das Massenspektrometer eluierenden Peptide stehen nur eine bestimmte Zeit für die Aufnahme des Massenspektrums sowie des Fragmentspektrums zur Verfügung. In einem ersten Scan (*survey scan*) werden die Massen der zu dieser Zeit eluierenden Peptide aufgenommen (MS). Nach bestimmten Auswahlkriterien werden einzelne Peptide in der Kollisionszelle nacheinander fragmentiert und deren Fragmentspektren aufgenommen (MS/MS). Nach Beendigung der MS/MS-Messungen erfolgt wiederum ein *survey scan*, gefolgt von einem erneuten MS/MS Zyklus. Auf diese Weise werden die eluierenden Peptide vom Massenspektrometer *abgearbeitet*.

Bei den in dieser Arbeit verwendeten ESI-Massenspektrometern handelt es sich um sogenannte Hybridgeräte, die unterschiedliche Massenanalytoren miteinander kombinieren. In einem Fall erfolgte die Massenanalyse analog zur MALDI-MS mit einem Flugzeitanalysator (TOF, s.o.), dem zur Fokussierung und Auswahl der zu fragmentierenden Ionen (Massenfilter) sowie als Kollisionszelle Quadrupol-Elektroden vorgeschaltet sind. Aus den für dieses Massenspektrometer verwendeten Komponenten ergibt sich die gebräuchliche Abkürzung ESI-Q(uadrupol) TOF [Cernushevich *et al.*, 2001]. Ein weiteres Massenspektrometer kombiniert die Fähigkeit der schnellen Analysezyklen einer Ionenfalle mit der hohen Massengenauigkeit der FT-ICR Analyse [Syka *et al.*, 2004; Olsen *et al.*, 2004]. Bei der Ionenfalle handelt es sich abweichend von der zuvor vorgestellten dreidimensionalen Ionenfalle um eine lineare Ionenfalle, auf dessen Besonderheiten hier allerdings nicht eingegangen werden soll [Schwartz *et al.*, 2002; Mayya *et al.*, 2005].

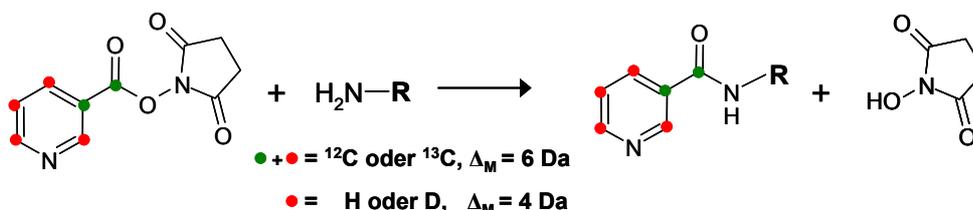
## 1.8 Quantifizierung in der Proteomik

Die Identifizierung der Proteine eines Organismus kann, wie am Beispiel von *H. salinarum* erläutert wurde, die korrekte Genomannotation entscheidend unterstützen. Ein solches Proteininventar ermöglicht allerdings nur begrenzte physiologische Aussagen. Einen wesentlichen Beitrag zum Verständnis biologischer Funktionen liefert die so genannte quantitative Proteomik. Hier werden die Proteome unterschiedlicher Zellzustände miteinander verglichen. In den Proteinmustern dieser verschiedenen Zustände (z.B. nach unterschiedlichen Wachstumsbedingungen, Stressfaktoren etc.) können Veränderungen einzelner Proteine beobachtet und daraus Rückschlüsse auf die beteiligten funktionellen Netzwerke (Stoffwechselwege, Regulationskaskaden etc.) gezogen werden. Durch eine Korrelation mit dem Phänotyp kann so eine funktionelle Aussage erreicht werden. Es ist jedoch anzumerken, dass keine absoluten Proteinkonzentrationen bestimmt werden, sondern nur das Mengenverhältnis eines bestimmten Proteins zwischen den gewählten Bedingungen. Es handelt sich daher eher

um relative Quantifizierung. Für diese *vergleichende* Proteomik existieren unterschiedliche Ansätze.

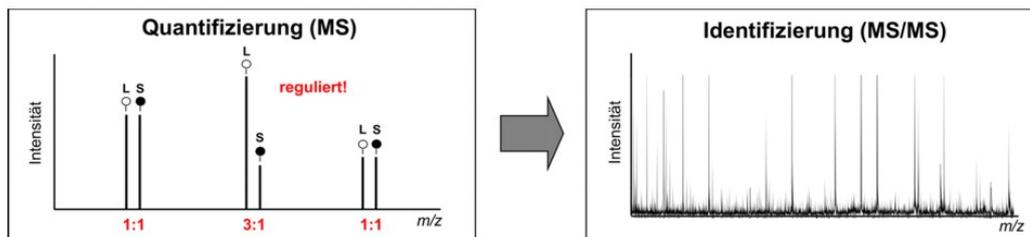
Die *traditionelle* Methode ist der visuelle Vergleich der in zweidimensionalen Gelen aufgetrennten Proteinmuster der unterschiedlichen Zustände. Die Entwicklung von computerunterstützten Bildverarbeitungsprogrammen hat diesen Ansatz erheblich vereinfacht [Raman *et al.* 2002; Luhn *et al.*, 2003; Rosengern *et al.* 2003; Dowesey *et al.*, 2003]. Ein Nachteil sind jedoch die bei den Färbemethoden von Gelen auftretenden Sättigungseffekte. Die Analyse von sehr großen und kleinen sowie basischen und hydrophoben Proteinen erweist sich mittels 2-DE ebenfalls als äußerst problematisch.

Ein weiterer Ansatz zum Vergleich unterschiedlicher Proteommuster basiert auf der kovalenten Modifikation von Proteinen mit isotope markierten Sonden (*isotope coded tags*). Dabei werden die Proteine eines Zellzustandes mit einer *leichten* Sonde, die des anderen Zustandes mit einer *schweren* Sonde (eine bestimmte Anzahl von Protonen oder  $^{12}\text{C}$  gegen Deuterium oder  $^{13}\text{C}$  substituiert) markiert. Ausgehend von der von Gygi und Aebersold vorgestellten ICAT-Methode (*isotope coded affinity tag*) [Gygi *et al.*, 1999] sind mittlerweile eine Vielzahl von unterschiedlichen isotope markierten Sonden entwickelt worden, welche die unterschiedlichen funktionellen Gruppen von Proteinen, bzw. Peptiden (Thiofunktionen etc.) kovalent modifizieren [Julka und Regnier, 2003]. In dieser Arbeit wurde zur Isotopenmarkierung von Proteinen die ICPL-Sonde eingesetzt (*isotope coded protein label*) [Schmidt *et al.*, 2005]. Diese markiert die Aminofunktionen der Lysine sowie die freien N-Termini aller Proteine:



**Abb. 1-13:** Struktur und Chemie der ICPL Sonde zur Markierung freier Aminofunktionen.

Nach Auftrennung der vereinten Proben beider Zustände (Proteine bzw. nach proteolytischen Verdau entstandene Peptide) geschieht die Quantifizierung anhand der im Massenspektrum erhaltenen Signalintensitäten. Dabei treten im Massenspektrum Paare des gleichen Peptids auf, deren Massenunterschied dem der eingesetzten Sonden entspricht. Die Identifizierung des zugehörigen Peptids erfolgt durch die MS/MS-Analyse der jeweiligen Peptide:



**Abb. 1-14:** Prinzip der Quantifizierung und Identifizierung mittels isotopenmarkierter Sonden. L = Peptid markiert mit leichter Sonde, S = Peptid markiert mit schwerer Sonde. Die Intensitätsverhältnisse sind in rot angegeben.

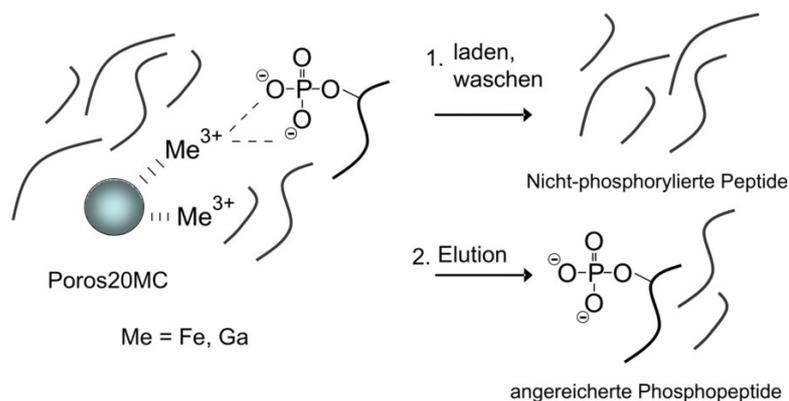
Die erhaltenen Verhältnisse der Signalintensitäten der Peptide entsprechen den Mengenverhältnissen der zugehörigen Proteine in den unterschiedlichen Zellzuständen. Die Quantifizierung mittels isotopenmarkierter Sonden erlaubt einen sehr genauen Vergleich der Proteinmengen. Da nach der Markierung der Proteine die Proben der unterschiedlichen Zustände vereinigt werden, sind Fehler aufgrund unterschiedlicher Probenbehandlung weitgehend auszuschließen. Eine Probenaufarbeitung entweder auf Protein- oder Peptidebene mit den unterschiedlichen elektrophoretischen oder chromatographischen Methoden ist möglich, da die finale Proteinidentifizierung mittels Tandem-MS geschieht.

## 1.9 Phosphoproteom Analyse

Ein weiterer wichtiger Bereich in der Proteomik ist die Analyse von posttranslationalen Modifikationen, die einen erheblichen Einfluss auf die Funktionen und Eigenschaften eines Proteins haben können (siehe 1.3) [Mann und Jensen, 2003; Jensen, 2004]. Die Analyse von Phosphorylierungen hat dabei aufgrund ihrer zentralen Funktion einen besonderen Stellenwert [Sickmann und Meyer, 2001; Loyet *et al.*, 2005]. Der Begriff Phosphoproteom beschreibt hierbei die Gesamtheit aller möglichen Phosphoproteine.

Neben den herkömmlichen Methoden zur Identifizierung von Phosphoproteinen wie der metabolen Markierung mit radioaktivem Phosphat wurden in den letzten Jahren im Bereich der Proteomanalytik neue methodische Ansätze entwickelt. Die massenspektrometrische Analyse gestaltet sich jedoch in vielerlei Hinsicht als problematisch: Zum einen liegen die Proteine oft nur in einem geringen Ausmaß und zudem abhängig von äußeren Faktoren in ihrem phosphorylierten Zustand vor. Zum anderen trägt die negativ geladene Phosphatgruppe zur Verschlechterung der Signalintensität im Massenspektrum bei (Signalsuppression). Diese Umstände machen eine Anreicherung von Phosphopeptiden vor der massenspektrometrischen Analyse unumgänglich. Dabei hat sich die auf immobilisierte Metallionen basierende

Affinitätschromatographie (IMAC = *immobilized metal affinity chromatography*) als eine wichtige Methode etabliert [Posewitz und Tempst, 1999; Ficarro *et al.*, 2002]. Sie beruht auf der Ausbildung von Salzbrücken zwischen der negativ geladenen Phosphatgruppe des Phosphopeptids und dem an die stationäre Phase chelatisierten positiv geladenen Metallion wie z.B. Gallium(III) oder Eisen(III) (siehe Abb.1-15).



**Abb. 1-15:** Generelles Prinzip der immobilisierten Metallionen Affinitätschromatographie (IMAC).

Ein wesentlicher Beitrag der Massenspektrometrie für die Analyse phosphorylierter Proteine liegt in der Identifizierung der genauen Phosphorylierungsstelle. Die Position der Phosphatgruppe in der Primärsequenz des Peptids bzw. Proteins kann aus dem MS/MS Spektrum des Phosphopeptids abgeleitet werden. Durch die Kombination der Phosphopeptid-Anreicherung mit massenspektrometrischer Hochdurchsatz-Analytik lassen sich eine Vielzahl von Phosphoproteinen und deren Phosphorylierungsstellen identifizieren [Beausoleil *et al.*, 2004; Ballif *et al.*, 2004; Gruhler *et al.*, 2005].

## 2 Ziel der Promotion

Im Rahmen dieser Arbeit wurde das cytosolische Proteom von *Halobacterium salinarum* unter folgenden Aspekten analysiert.

Zunächst sollte das Proteininventar mittels zweidimensionaler Gelelektrophorese (2-DE) auf standardisierten Gelen ermittelt und kartiert werden. Dies sollte zum einen das systematische Identifizieren von Proteinen ermöglichen, um unnötig wiederholtes Identifizieren bereits bekannter Proteine zu vermeiden. Das zweite Ziel der systematischen Proteomkartierung mittels 2-DE war die Validierung der bis dahin ausschließlich zur Proteinidentifizierung verwendeten MALDI-TOF-MS *peptide mass fingerprinting* Methode im entwickelten Hochdurchsatzverfahren (siehe Promotion Dr. Christian Klein), dessen Anwendbarkeit im Vergleich zu weiteren Methoden so abgeschätzt werden konnte.

Die erstellten Referenzgele sollten anschließend zur differentiellen Proteomanalytik herangezogen werden, um regulierte Proteine unter unterschiedlichen Wachstumsbedingungen durch Vergleich der Proteinmuster zu identifizieren und zu quantifizieren. Hierbei wurde zum einen das Wachstum unter aeroben Bedingungen ohne Licht mit phototrophen anaerobem Wachstum, zum anderen das Proteininventar von Zellen, die in Komplexmedium oder synthetischem Minimalmedium wuchsen, verglichen. Neben der 2-DE kam hierbei die auf Isotopenmarkierung von Proteinen beruhende ICPL-Technologie zum Einsatz, bei der die Quantifizierung der Proteine direkt aus den massenspektrometrischen Daten hervorgeht.

Des Weiteren wurde das Ausmaß der Proteinphosphorylierung in *Halobacterium salinarum* untersucht. Hierbei waren die Identifizierung von Phosphoproteinen sowie deren Phosphorylierungsstellen von vorrangigem Interesse. Neben den bekannten Proteinphosphorylierungen in den Signalkaskaden der 2-Komponenten Systeme, welche sich auf Histidin und Aspartat Phosphorylierungen beschränken, ist bisher wenig über Ser-, Thr- oder Tyr-Phosphorylierung und dessen Einfluss auf regulative Prozesse bekannt.

Die im Rahmen aller Arbeitsansätze identifizierten Proteine wurden fortlaufend zur Validierung der Genomdaten und zum Ausbau des Datenbanksystems HALOLEX herangezogen und vervollständigten so die Inventarisierung der Proteine von *H. salinarum*.

# 3 Materialien und Methoden

## 3.1 Material und Geräte

### 3.1.1 Materialien / Chemikalien

$\alpha$ -Cyano-4-Hydroxy-Zimtsäure	Bruker
Aceton	Riedel de Haen
Acetonitril	Merck
Acrylamid Proto Gel	National Diagnostics
Agarose	Gibco
Ameisensäure 98-100%	Merck
Ammoniumhydrogencarbonat	Fluka
Ammoniumperoxisulfat	BioRad
BCA-Kit	Pierce
Bradford-Kit	BioRad
$\beta$ -Casein	Sigma
Bromphenolblau	Sigma
C <sub>18</sub> -Extraction Disks	3M Empore
C <sub>18</sub> -Poros Material	Applied Biosystems
CHAPS	Sigma
Chloroform	Merck
Coomassie Brilliant Blau G-250	Serva
6-Deoxy-L-Galaktose (Fukose)	Sigma
Drystrip Cover Fluid (plus one)	Amersham Biosciences
Dithiothreitol	Merck
DNase I	Fluka
ECL-Kit	Pierce
Eisenchlorid	Merck
Elektrospray Nadel ( <i>on-line/off-line</i> )	New Objective / Proxeon Biosystems
Essigsäure 100%	Merck
Ethanol absolut	Riedel de Haen
Formaldehyd 37%	Merck
GeLoader-Tips	Eppendorf
Glycerin	Roth
Glycin	Merck
Harnstoff	Merck
Iodacetamid	Merck
Isopropanol	Merck
Kaliumchlorid	Merck
Kaliumhexacyanoferrat(III)	Fluka
Kerosin	Fluka
Magnesiumsulfat-heptahydrat	Merck
Methanol	Merck
N,N'-Methylenbisacrylamid	BioRad
Natriumacetat	Merck
Natriumcarbonat	Merck
Natriumfluorid	Sigma
Natriumhydrogenphosphat	Merck
Natriumhydroxid	Merck

Natriumthiosulfat-Pentahydrat	Merck
Pepton	Oxoid
Pharmalyte	Amersham Biosciences
Phosphorsäure	Merck
Poros 20 MC	Applied Biosystems
PVDF Membran	Millipore
Reaktionsgefäße 0.5, 1.5, 2.0 ml	Eppendorf
Salzsäure	Riedel de Haen
SDS, 2 x kristallisiert	Serva
Siedespeisesalz	Saline Bad Friedrichshall – Bad Reichenhall
Silbernitrat	Sigma
Steritops	Millipore
TECEP	Sigma
TEMED	Sigma
Thioharnstoff (plus one)	Amersham Biosciences
Trifluoressigsäure	Perkin Elmer
tri-Natriumcitrat-dihydrat	Merck
Trypsin, Modified Sequencing Grade	Promega
Antikörper	Biomol
Taurodexycolat	Sigma
Mikrotiterplattenschüttler	Heidolph
Orthovanadate	Sigma
Omnifix 1 <sup>®</sup> ml Einmalspritze	B.Braun
Whatmanpapier	VWR
Röntgenfilm	Agfa
PVP	Sigma
32-Phosphat	Hartmann Analytic
33-Phosphat	Hartmann Analytic
PEG-3500	Sigma
Tween-20	Merck

### 3.1.2 Geräte

MALDI-Massenspektrometer Reflex III	Bruker Daltonics
Proteiner SP	Bruker Daltonics
MAP II-Pipettierroboter	Bruker Daltonics
SMART HPLC System	Amersham Biosciences
Eppendorf Tischzentrifuge	Eppendorf
Tisch-Ultrazentrifuge (TL-100)	Beckmann
Q-TOF Ultimate Massenspektrometer	Waters
Multiphor II + IEF Zubehör	Amersham Biosciences
IPG-Phor	Amersham Biosciences
Film Entwickler	Agfa
Pipetten (2, 10, 20, 100, 200, 1000, 5000 µl)	Gilson
Pipetten (2, 10, 1000 µl)	Eppendorf
8-Kanal Pipetten	SL-Pette
Power-Supply EPS 1000 XL	Amersham Biosciences
Power-Supply EPS 3500 XL	Amersham Biosciences
IPG-Streifen Rehydratisierkassette	Amersham Biosciences
Semi-Dry Blotapparatur	Werkstatt MPI
Protean II Elektrophoresetank	BioRad
Reinstwasseranlage	Millipore
Ultraschall Bad (Sonorex RK510)	Bandelin
Ultraschall Bad (Branson Sonifier450)	Branson
Vortex Genie 2	Bender & Hobein AG
Szintillationszähler	Packard
Typhoon <sup>™</sup> 9400 Scanner	Amersham Biosciences

### 3.1.3 Programme

Biotoools 2.2	Bruker
Data Explorer 4.3 A	Applied Biosystems
Exel Office2003	Microsoft
GPMAW 6.2	Lighthouse Data
HALOLEX	Abteilung Oesterhelt
Mascot	Matrix Science
MassLynx 4.0	Waters
MSQuant (Version 1.4.0a12)	Center for Experimental BioInformatics, University of Southern Denmark
Origin 7.0 SR0	OriginLab Corporation
Proteomweaver	Definiens
Spotfire DecisionSite 8.0	Spotfire, Inc
Word Office2003	Microsoft

### 3.1.4 Medien und Lösungen

Vollmedium: (Mengen auf 1 Liter)	250g NaCl, 20g MgSO <sub>4</sub> ·7 H <sub>2</sub> O, 2 g KCl, 3 g Tri-Na-Citrat, 10 g Oxoid-Pepton, pH 7.0
synthetisches Medium: (Mengen auf 1 Liter)	235 g NaCl, 10 g MgSO <sub>4</sub> ·7 H <sub>2</sub> O, 2 g KCl, 0.1 g KNO <sub>3</sub> , 0.5 g Tri-Na-Citrat, 0.3g KH <sub>2</sub> PO <sub>4</sub> , 0.3 g K <sub>2</sub> HPO <sub>4</sub> , 0.05 mg CuSO <sub>4</sub> ·5 H <sub>2</sub> O, 2.3 mg FeCl <sub>2</sub> ·4 H <sub>2</sub> O, 0.3 mg MnSO <sub>4</sub> ·H <sub>2</sub> O, 0.44 mg ZnSO <sub>4</sub> ·7H <sub>2</sub> O, 0.02 mg Ascorbinsäure, 0.4 g L-Arginin, 0.44 g L-Isoleucin, 0.8 g L-Leucin, 0.224 g L-Lysin·H <sub>2</sub> O, 0.2 g L-Methionin, 0.5 g L-Threonin, 0.25 g L-Valin, 0.61 g L-Serin, 0.1 g L-Phenylalanin, 0.2 g L-Tyrosin, 5 mg Thiamin, 5 mg Folsäure, 0.5 mg Biotin, 0.5% (v/v) Glycerol
SDS-PAGE Auftragspuffer: (fünffach konzentriert)	50 mM Tris/HCl pH 6.8, 2 g DTT, 2% (w/v) SDS, 250 mg Bromphenolblau, 28.5 mL Glycerol (87% Lsg.)
Peptid-Standartt Mix: (Konzentration jeweils 100 pm/μl)	Bradykinin Fragment 1-7 (757.3998 g/mol), Angiotensin 2 Acetat Human (1046.542 g/mol), Angiotensin 1 Acetat Human (1296.6853 g/mol), Substance P (1347.7361 g/mol) Bombesin (1619.823 g/mol), ACTH (1-17) (2093.0868 g/mol), ACTH (18-39) (2465.199 g/mol), Somatostatin 28 (3147.4714 g/mol), Insulin Chain B Oxidized f. Bovine (3494.651 g/mol)
Elektrophoresepuffer: (Mengen auf 1 Liter)	1 g SDS, 12 g Tris, 57.6 g Glycin

### 3.1.5 Abkürzungen

1-D	eindimensional
2-D	zweidimensional
2-DE	zweidimensionale Gelelektrophorese
Abb.	Abbildung
ACN	Acetonitril
AK	Antikörper
APS	Ammoniumperoxodisulfat
BCA	Bichinoninsäure
BPB	Bromphenolblau
BS	Basalsalz
BSA	<i>bovine serum albumine</i>
CAI	<i>codon adaptation index</i>
CHAPS	3[(3-Cholamidopropyl)dimethylammonio]-1-propan-sulfonat
CID	Collision induced dissociation
DTT	1,4- Dithiothreitol

---

DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
EDTA	Ethylendiamintetraacetat
FA	Ameisensäure
FT	Rourrier Transform
h	Stunde(n)
HCCA	$\alpha$ -Cyano-4-Hydroxy-Zimtsäure
HOAc	Essigsäure
IAA	Iodacetamid
ICPL	<i>isotope coded protein label</i>
i.d.	Innendurchmesser
IEF	isoelektrische Fokussierung
IMAC	<i>immobilized metal affinity chromatography</i>
IPG	Immobilisierter pH Gradient
kbp	<i>kilo base pairs</i>
LC	Flüssigchromatographie ( <i>liquid chromatography</i> )
MALDI	Matrix assisted Laser Desorption/Ionisation
Mbp	<i>mega base pairs</i>
MeOH	Methanol
MilliQ	Millipore Wasser
min	Minuten
MS	Massenspektrometrie
MS/MS	Tanden- Massenspektrometrie
$M_r$	Molekulargewicht
PAA	Polyacrylamid
PAGE	Polyacrylamidgelelektrophorese
<i>pI</i>	Isoelektrischer Punkt
rpm	<i>rounds per minute</i>
RT	Raumtemperatur
SCX	<i>strong cation exchange</i>
sec.	Sekunden
SDS	Natriumdodecylsulfat
SDS-PAGE	SDS-Polyacrylamidgelelektrophorese
TCA	Trichloressigsäure
TCEP	Tris-(2-carboxyethyl)phosphine, hydrochloride
TEMED	N, N, N', N'-Tetramethylethylendiamin
TFA	Trifluoressigsäure
TM	Transmembran
TOF	<i>time-of-flight</i>
Tris	Tris-(Hydroxymethyl)-aminomethan
PBS	<i>phosphate buffered saline</i>
PEG	Polyethylene Glycol
PMF	<i>peptide mass fingerprint</i>
ppm	<i>parts per million</i>
PVP	Polyvenylpyrrolidon
U	<i>units</i>

## 3.2 Methoden

### 3.2.1 Anzucht von *Halobacterium salinarum*

#### 3.2.1.1 Anzucht der *Halobacterium*-Standkultur

Zum Generieren einer Standkultur die zum Animpfen aller nachfolgenden Kulturen dienen sollte, wurde ein 2-Liter-Erlenmeyerkolben mit 1 Liter Vollmedium (siehe 3.1.4) [Oesterhelt und Stoeckenius, 1974; Oesterhelt und Krippahl, 1983] mit 35 ml einer Stammkultur des *Halobacterium salinarum* Wildtypstamms R1 (DSM 671) angeimpft. Die Kultur wurde bei 37°C und 100 rpm im Dunkeln bis zur stationären Phase geschüttelt und danach bei Raumtemperatur gelagert.

#### 3.2.1.2 Anzucht der *Halobacterium* Gebrauchskultur

Ausgehend von der in 3.2.1.1 beschriebenen Standkultur wurden nun alle *Halobacterium* Kulturen angeimpft. Die Anzucht der Gebrauchskulturen erfolgte in 100 ml Erlenmeyer mit angebrachtem Reagenzglasfortsatz (Klettkolben). 35 ml wurden mit 1 ml der Standkultur angeimpft und bei 37°C im Dunkeln bis zu einer optischen Dichte (Trübung) von 30-40 Klett bei 100 rpm geschüttelt. Um sicherzustellen, dass alle Zellen die späte logarithmische Wachstumsphase erreicht hatten, wurde eine neue Kultur mit 1 ml der eben erstellten inokuliert, die wiederum bis 30-40 Klett unter gleichen Bedingungen geschüttelt wurde. Dieser Schritt wurde noch einmal wiederholt. Diese dritte Kultur wurde nun geerntet und wie in 3.2.2 beschrieben weiterverarbeitet.

Die Anzucht der Zellen in synthetischem Medium erfolgte analog wie oben beschrieben (für die Zusammensetzung des synthetischen Mediums siehe 3.1.4). Für die Anzucht der Zellen, die unter anaeroben, phototrophen Bedingungen wachsen sollten, wurden Klettkolben mit Schliff verwendet. Nach dem Inokulieren wurde dieser mit einem Septum Luft dicht verschlossen und die Kultur bei 37°C unter Belichtung mit 50 - 100 mW/cm<sup>2</sup> bei 100 rpm geschüttelt. Das Anziehen durch die drei aufeinander folgenden Kulturen erfolgte wie oben beschrieben.

### 3.2.2 Zellaufschlüsse

#### 3.2.2.1 Zellaufschluss und Proteinpräparation mittels Taurodeoxycholat

Taurodeoxycholat ist ein Gallensäurederivat, welches in geringen Konzentrationen zur Lyse von *Halobacterium salinarum* führt.

Zellen aus 2 ml Kultur wurden in einem 2 ml Eppendorf-Reaktionsgefäß bei 5000 g, 4°C für 5 min sedimentiert, und der Überstand verworfen. Dieser Schritt wurde wiederholt, so dass ein Reaktionsgefäß Zellen aus 4 ml Kultur enthielt. Die Zellen wurden einmal mit Basalsalz (4.3 M NaCl, 81 mM MgSO<sub>4</sub>, 27 mM KCl) gewaschen und erneut sedimentiert. Die Lyse der Zellen erfolgte wie beschrieben [Kamekura *et al.*, 1988; Kupper *et al.*, 1994] nach Zugabe von 200 µl Lysepuffer (0.0075% Taurodeoxycholat, Protease Inhibitor Complete ohne EDTA, 100 µg/mL DNase I in 50% BS; Phosphatase Inhibitor Mix (Roche)) für 45 min bei 4°C unter stetigem Schütteln. Bei Proben, welche zur Analyse des Phosphoproteoms dienen sollten, wurden zusätzlich 1 mM Orthovanadate und 1 mM NaF als Phosphataseinhibitoren zugesetzt. Die Suspension wurde anschließend bei 213,000 g, 4°C für 30 min zentrifugiert, um Zelltrümmer zu sedimentieren. Der Überstand wurde abgenommen und die in ihm befindlichen cytosolischen Proteine durch Zugabe des 10fachen Überschusses mittels Aceton gefällt. Die gefällten Proteine wurden mittels Zentrifugation bei 14,000 rpm, 4°C sedimentiert, mit eiskaltem 50%em Aceton gewaschen, um Salzkontaminationen zu entfernen, und erneut sedimentiert. Dieser Waschschrift wurde noch einmal wiederholt und das Proteinsediment im Vakuum getrocknet.

#### **3.2.2.2 Zellaufschluss mittels Sonifizierung in Basalsalz**

Zellen wurden wie in 3.2.2.1 beschrieben sedimentiert und in je 200 µl Lysepuffer (wie in 3.2.2.1 ohne Taurodeoxycholat) aufgenommen. Die Lyse erfolgte auf Eis durch gepulste Sonifizierung bei 50% Duty Cycle (Gerät). Die Proteine wurden wie in 3.2.2.1 aufgearbeitet.

#### **3.2.2.3 Zellaufschluss mittels Sonifizierung in Harnstofflösung**

Die Zellen einer 35 ml Kultur (30-40 Klett) wurden in einem 50 ml Falcon-Gefäß bei 5000 g für 15 min sedimentiert und in 1 ml Lysepuffer (8 M Harnstoff, 100 mM Tris/HCl pH 8, 1 mM Orthovanadate, 1 mM NaF) aufgenommen. Der Aufschluss der Zellen sowie die Zerkleinerung der frei werdenden DNA erfolgte durch längeres Sonifizieren im Ultraschallbad (Branson Sonifier 450). Die Proben wurden für in-Lösungs Verdauung (siehe 3.2.8.3) eingesetzt.

### 3.2.3 Proteinbestimmungen

#### 3.2.3.1 Proteinbestimmung nach Bradford

Die Proteinbestimmung nach Bradford [Bradford, 1976] beruht auf der Verschiebung des Absorptionsmaximums des Farbstoffs Coomassie Brilliantblau G250 von 465 auf 595 nm in Gegenwart von Proteinen. Grund dafür ist die Stabilisierung des Farbstoffs in seiner anionischen Form durch Bindung an Arginyl-Reste von Proteinen. Durch Messung der Absorption bei 595 nm kann man die Menge des gebundenen Farbstoffes quantifizieren. Die Proteinprobe wurde zu 1 ml Bradford Reagenz (800  $\mu$ l H<sub>2</sub>O + 200 ml Bradford-Dye, BioRad) gegeben und photometrisch bei 595 nm vermessen. Die Proteinkonzentration wurde durch Vergleich mit einer Eichkurve (BioRad Protein-Assay-Standard, 0 - 25  $\mu$ g/ml) ermittelt.

#### 3.2.3.2 Proteinbestimmung mittels BCA

Eine weitere Möglichkeit zur quantitativen Proteinbestimmung ist die Farbreaktion mit Bicinchoninsäure (BCA) [Smith *et al.*, 1985]. Proteine bilden mit Cu<sup>2+</sup>-Ionen in alkalischer Lösung einen Komplex (Biuret-Reaktion), wobei die Cu<sup>2+</sup>-Ionen zu Cu<sup>+</sup>-Ionen reduziert werden. Diese bilden dann mit der Bicinchoninsäure einen violetten Farbkomplex, der bei 562nm sein Absorptionsmaximum aufweist.

50 Volumenanteile von Lösung A (1 % BCA-Na<sub>2</sub>, 2 % Na<sub>2</sub>CO<sub>3</sub>·H<sub>2</sub>O, 0.16% Na<sub>2</sub>-tartrat, 0.4 % NaOH, 0.95 % NaHCO<sub>3</sub>, eingestellt auf pH 11.25) wurden mit 1 Volumenanteil von Lösung B (4% CuSO<sub>4</sub>·5H<sub>2</sub>O) vermischt und zu 1 ml dieses Gemisches 50  $\mu$ l der zu analysierenden Probe gegeben. Die Probe wurde für 30 min bei 37°C inkubiert, anschließend die Absorption bei 562 nm gemessen und die Proteinkonzentration aus dem Vergleich mit einer BSA-Eichgeraden (0.1-1.0 mg/ml) ermittelt. Der Vorteil dieser Methode liegt in einer geringeren Empfindlichkeit gegenüber Salzen und Detergentien. Die Temperatur muss jedoch sehr genau eingehalten werden.

### 3.2.4 *In-vivo* Markierung mit 33/32-P

Um einen Überblick über das Ausmaß der Proteinphosphorylierung in *Halobacterium salinarum* zu gewinnen, sollten Proteine *metabolisch* mit radioaktivem Phosphat markiert werden, um sie nach Auftrennung über SDS-Gelelektrophorese (1-D sowie 2-D) zu identifizieren. Dazu wurden Zellen bei ca. 30 Klett sedimentiert, dreifach konzentriert in BS aufgenommen und bei 37°C in einer temperierbaren Glasküvette gerührt. Die Markierung erfolgte durch Zugabe von bis zu 500  $\mu$ Ci 33-P oder bis zu 30  $\mu$ Ci 32-P entweder im Dunkeln oder durch Bestrahlung bei 50 - 100 mW/cm<sup>2</sup>.

Die Aufarbeitung der Proteine für die nachfolgende 2-DE erfolgte wie in [3.2.6.2](#) beschrieben. Dabei wurden von allen Fraktionen Proben genommen und deren spezifische Radioaktivität im Szintillationszähler (Hersteller) bestimmt.

Um die Phosphataufnahme der Zellen zu verfolgen, wurden die Zellen bei 5000 g sedimentiert, einmal mit BS gewaschen, erneut in BS aufgenommen und in Szintillationsflüssigkeit gegeben.

### 3.2.5 ICPL Markierung von Proteinen

Zum Einbringen der ICPL-Sonden in cytosolische Proteine wurden isolierte Proteinpellets aus zwei unterschiedlichen Zellzuständen in ICPL-Markierungspuffer (6 M Guanidiniumchlorid, 0.1 M HEPES, pH 8.5) aufgenommen und 10 min im Ultraschallbad sonifiziert. Die Proteinkonzentration der beiden Lösungen wurde nach Bradford (siehe [3.2.3.1](#)) oder mittels BCA (siehe [3.2.3.2](#)) bestimmt und durch Verdünnen mit ICPL-Markierungspuffer auf 5 mg/ml eingestellt. Folgende Mengen Reaktionslösung beziehen sich auf 40 µl Ausgangslösung bzw. 200 µg Protein.

Die Proteine wurden durch Zugabe von 1 µl 0.2 M TCEP-Lsg. (in 0.1 M HEPES, pH 8.5; Endkonzentration 5 mM) bei 60°C für 30 min reduziert und anschließend durch Zugabe von 1 µl 0.4 M IAA-Lsg. (in 0.1 M HEPES, pH 8.5; Endkonzentration 10 mM) für 30 min bei RT im Dunkeln alkyliert. Durch Zugabe von 1 µl 0.5 M N-Acetyl-Cystein-Lsg. (in 0.1 M HEPES, pH 8.5; Endkonzentration 12.5 mM) wurde restliches IAA abgefangen. Vor der Nicotinoylierung der Proteine wurde der pH-Wert der Lösungen überprüft und ggf. wieder auf 8.5 eingestellt. Die Markierung erfolgte durch die Zugabe eines zehnfachen molaren Überschusses von ICPL-Reagenz gegenüber allen freien Aminogruppen (je 6 µl einer 0.15 M N-Nicotinoyl-NHS-Lsg.). Dabei wurden eine Probe mit dem leichten und eine mit dem schweren Reagenz (entweder 4 x D oder 6 x <sup>13</sup>C) markiert. Die Proben wurden für 2 h bei RT inkubiert und die Reaktion durch Zugabe von 4 µl 1.5 M Hydroxylaminlösung (in H<sub>2</sub>O) abgebrochen. Anschließend wurden beide Proteinlösungen vereinigt. Um die bei der Nicotinoylierungsreaktion ebenfalls entstandenen O-Acylester mit Serinen, Threoninen oder Tyrosinen zu spalten, wurde die Lösung durch Zugabe eines entsprechenden Volumens (ca. 4 µl) 2 M NaOH für 20 min auf pH 12 verschoben und anschließend durch die Zugabe des gleichen Volumens 2 M HCl wieder neutralisiert.

Für die Auftrennung der markierten Proteine mittels 2-DE wurden die Proteinlösungen gegen eine Lösung aus 8 M Harnstoff, 2 M Thioharnstoff und 20 mM Tris über Nacht dialysiert. Die Probe wurde dann mit Rehydratisierungspuffer (8 M Harnstoff, 2 M Thioharnstoff, 20 mM Tris, 4% CHAPS, 2% Pharmalyte, etwas BPB) verdünnt und in die IPG-Streifen über Nacht eingequollen. Die Elektrophorese erfolgte wie weiter unten

beschrieben (siehe 3.2.6.2). Auf die Reduzierung und Alkylierung nach der Fokussierung konnte jedoch verzichtet werden, da diese bereits vor der ICPL-Markierung durchgeführt worden war.

Für die 1-D SDS PAGE wurden die Proteine aus der Guanidiniumchlorid Lösung gefällt. Dazu wurde Lösung mit dem gleichen Volumen H<sub>2</sub>O verdünnt und die Proteine durch Zugabe eines fünffachen Überschusses an Aceton über Nacht bei -20°C präzipitiert. Die ausgefallenen Proteine wurden bei 14,000 rpm, 4°C sedimentiert, noch einmal mit 80% Aceton gewaschen und anschließend im Vakuum getrocknet. Die markierten Proteine standen nun zur weiteren Analyse zur Verfügung.

### 3.2.6 Gelelektrophorese von Proteinen

#### 3.2.6.1 SDS-PAGE im Mini- und Medi-Format

Zur eindimensionalen SDS-Gelelektrophorese (1-DE) wurde das System von Laemmli [Laemmli, 1970] verwendet. Analytische Gele wurden im Mini-Gel Format (8 cm x 7 cm 0.075 cm; Breite, Höhe, Dicke), präparative Gele im Medi-Gel Format (13 x 13 x 0.15 cm, LKB) angefertigt.

Trenngel (12.5%):

	<b>Mengen für 2 Mini-Gele</b>	<b>Mengen für 1 Medi-Gel</b>
H <sub>2</sub> O	3.84 ml	11.53 ml
Tris (2M pH 8.8)	1.89 ml	5.67 ml
SDS 10%	100 µl	300 µl
Acrylamid (Protogel)	4.17 ml	12.5 ml
APS (10%)	150 µl	450 µl
TEMED	5 µl	15 µl

Sammelgel (3%):

	<b>Mengen für 2 Mini-Gele</b>	<b>Mengen für 1 Medi-Gel</b>
H <sub>2</sub> O	3.78 ml	11.34 ml
Tris (2M pH 6.8)	0.63 ml	1.89 ml
SDS 10%	100 µl	300 µl
Acrylamid (Protogel)	0.5 ml	1.5 ml
APS (10%)	25 µl	75 µl
TEMED	10 µl	30 µl

Proteinpellets wurden in 1 x SDS-PAGE Auftragspuffer gelöst, bei 99°C für 5 min inkubiert und ggf. sonifiziert. Analytische Gele wurden mit ca. 30 µg Protein beladen und die Elektrophorese bei maximal 130 V und 20 mA/Gel durchgeführt. Die präparativen Medi-Gele wurden mit bis zu 150 µg Protein beladen, die Elektrophorese erfolgte hier bei maximal 350 V.

### 3.2.6.2 Zweidimensionale Gelelektrophorese (2-DE)

Die zweidimensionale Gelelektrophorese, basierend auf den Arbeiten von Klose [Klose, 1975] und O'Farrell [O'Farrell, 1975], setzt sich aus zwei unterschiedlichen Trennmethoden zusammen: Die erste Trennung ist die isoelektrische Fokussierung (IEF). Hier werden die Proteine in einem pH-Gradienten gelelektrophoretisch nach ihrer Ladung getrennt. Die Proteine wandern nach Anlegen einer elektrischen Spannung in dem pH-Gradienten zu dem pH-Wert, der ihrem isoelektrischen Punkt entspricht. Dort besitzen die Proteine keine Nettoladung und können demnach nicht weiter im elektrischen Feld wandern; sie sind *fokussiert*. Die zweite Trennung ist eine Polyacrylamid Gelelektrophorese (PAGE), bei der die Proteine nach ihrer Größe getrennt werden. Die 2-DE ist bis heute die Trennmethode mit der höchsten Auflösung von komplexen Proteingemischen.

Für die IEF kamen Fokussierstreifen mit immobilisierten pH-Gradienten (IPG-Streifen) der Firmen Amersham Biosciences und BioRad zum Einsatz (pH 3.5-4.5, pH 4-5, pH 4.5-5.5; Amersham Biosciences, 18 cm und pH 5-8; Bio-Rad, 17 cm). Die IPG-Streifen wurden zunächst über Nacht in einer Rehydratisierungs-Kassette in je 300 µl (17 cm) oder 350 µl (18 cm) Rehydratisierungspuffer (7 M Harnstoff, 2 M Thioharnstoff, 4% CHAPS, 0.6% Pharmalytes 3-10, 20 mM DTT) rehydratisiert. Dabei wurde das Austrocknen der Streifen durch Überschichten mit Cover Fluid verhindert. Nach vollständiger Rehydratisierung wurden die Streifen aus der Kassette entnommen, mit MilliQ abgespritzt und unter leichtem Druck zwischen zwei befeuchtete Whatmanpapiere gelegt, um überschüssige Rehydratisierungslösung und eventuell entstandene Harnstoffkristalle zu entfernen. Die IPG-Streifen wurden nun in eine Multiphor II Fokussierapparatur eingebaut.

Die getrockneten Proteinpellets wurden in IEF-Probenpuffer (7 M Harnstoff, 2 M Thioharnstoff, 4% CHAPS, 0.6% Pharmalyte 3-10, 65 mM DTT) gelöst und für 30 min geschüttelt, um eine möglichst vollständige Solubilisierung der Proteine zu erhalten. Nicht gelöste Bestandteile wurden durch Zentrifugation bei 213.000g für 30 min bei 18°C sedimentiert. Für die IEF wurden ca. 300 µg Protein gelöst in 20 - 40 µl IEF-Probenpuffer über *Application Pieces* auf die IPG-Streifen am kathodischen Ende (IPG-Streifen pH 3.5-4.5 und 4-5) oder am anodischen Ende (IPG-Streifen pH 4.5-5.5 und 5-8) aufgetragen. Größere Proteinmengen bis 800 µg wurden während des Rehydratisierungsvorgangs durch passive In-Gel Rehydratisierung in die IGP-Streifen eingebracht. Die Bestimmung der für die IEF eingesetzten Proteinmenge erfolgte durch Abschätzung der eingesetzten Zellmasse [Michel und Oesterhelt, 1976].

Die mit den Proben beschickten IPG-Streifen wurden mit Cover Fluid überschichtet, die Fokussierung erfolgte bei 20°C mit folgendem Strom-Spannungsgradienten:

Phase1	in 1 min auf 150 V	Phase5	in 1 h auf 500 V
Phase2	1 h bei 150 V	Phase6	4 h bei 500 V
Phase3	in 1 h auf 300 V	Phase7	in 5 h auf 3500 V
Phase4	1 h bei 300 V	Phase8	11 h bei 3500 V

Die Fokussierzeit betrug so 51,500 Vh. Während der ersten fünf Phasen wurden die Elektrodenpapiere zwischen den Elektroden und den IPG-Streifen regelmäßig gewechselt. Nach Beendigung der IEF wurden die IPG-Streifen entweder in Cellofanfolie eingewickelt und bei -80°C gelagert oder direkt für die zweite Dimension eingesetzt. Dafür wurden die IPG-Streifen zweimal für je 20 min in Äquilibriumspuffer (30% Glycerol, 6 M Harnstoff, 50 mM Tris/HCl, pH 8.8, 2% SDS) inkubiert, um die Proteine mit SDS zu beladen und sie so für eine erneute Elektrophorese zugänglich zu machen [Görg *et al.*, 1995]. Im ersten Inkubationsschritt enthielt der Äquilibriumspuffer zusätzlich 65 mM DDT, um Disulfidbrücken zu reduzieren. Im zweiten Schritt war DTT durch 260 mM Iodacetamid (IAA) ersetzt worden, um die erneute Oxidation der reduzierten Cysteine zu Disulfidbrücken durch irreversible Carbamidomethylierung zu verhindern. Zusätzlich wurde der Lösung noch etwas Bromphenolblau zugesetzt, um bei der anschließenden PAGE eine Lauffront erkennen und so das Ende der Elektrophorese abschätzen zu können.

Von den äquilibrierten IPG-Streifen wurden anschließend ca. 0.5 – 1 cm des anodischen und kathodischen Endes abgeschnitten. Bei radioaktiven Streifen erfolgte das Abschneiden direkt nach der Fokussierung, da besonders das anodische Ende verstärkt Radioaktivität aufwies. Die IPG-Streifen wurden nun auf die Polyacrylamidgele aufgelegt und mit 50°C warmer 0.5%iger Agaroselösung in Elektrophoresespuffer (siehe 3.1.4) fixiert. Für die Auftrennung der cytosolischen Proteine im Massenbereich zwischen 200 und 25 kDa wurden 9%ige, für Proteine kleiner 25 kDa 14%ige Gele verwendet. Es wurden ebenfalls Gradientengele zwischen 9-16% und 9-14% eingesetzt. Die Elektrophorese erfolgte in einer Protean II Elektrophoresekammer (BioRad) bei 6°C bei 25 mA pro Gel.

### 3.2.6.3 IEF mittels IPG-Phor

Die IPG-Phor (Amersham Biosciences) ist neben der Multi-Phor (BioRad) die am häufigsten verwendete Apparatur zur isoelektrischen Fokussierung. Sie wurde hauptsächlich zur Fokussierung von radioaktiven Proben verwendet, da jeder einzelne

IPG-Streifen in einem Keramik-Schiffchen fokussiert werden kann. Dies hatte nur die Kontamination der einzelnen verwendeten Schiffchen und zugehörigen Elektroden, nicht aber der gesamten Apparatur zur Folge. Zusätzlich kann die IPG-Phor bei höheren Spannungen (bis 7000 V) betrieben werden, was eine schnellere Fokussierung erlaubt. Die Fokussierung erfolgte bei 20°C mit folgendem Strom-Spannungsgradienten:

Phase1	in 1 min auf 150 V	Phase5	in 1 h auf 500 V
Phase2	1 h bei 150 V	Phase6	4 h bei 500 V
Phase3	in 1 h auf 300 V	Phase7	in 5 h auf 3500 V
Phase4	1 h bei 300 V	Phase8	11 h bei 3500 V

Die Streifen wurden entweder direkt weiterverarbeitet oder bei -80°C gelagert.

#### 3.2.6.4 Gießen von Gelen für die 2-DE

Für die zweite Dimension (SDS-PAGE) wurden entweder uniprozentuale Gele oder Gradientengele (17 x 18 x 0.75 cm) verwendet:

	<b>9%</b>	<b>14%</b>	<b>16%</b>
Acrylamidlg. (Protogel)	75 ml	116.6	133.3
1.5M Tris/HCL pH8.8	63 ml	63ml	63ml
MilliQ	112 ml	70.4	53.7
10%-ige APS-Lsg	750 µl	750 µl	750 µl
TEMED	75 µl	75 µl	75 µl

Um die Bildung von SDS-Acrylamid Micellen zu vermeiden, die zu einer Modifikation der Proteine durch Acrylamid führen kann, wurde den 2-D Gelen kein SDS zugesetzt. Das für die Elektrophorese erforderliche SDS befand sich lediglich im Elektrophoresepuffer. Die Lösungen wurden vor der Zugabe von APS und TEMED über Steritops (Millipore) filtriert und entgast. Das Gießen erfolgte in einer Gießkassette (BioRad), die das gleichzeitige Anfertigen von sieben Sandwich-Gelen ermöglichte. Für Gradientengele wurde ein von der Zentralwerkstatt angefertigter Gradientenmischer verwendet. Die fertig gegossenen Gele wurden mit wassergesättigtem 2-Butanol überschichtet und bis zur Polymerisation (ca. 1 h) bei RT belassen. Danach wurden die Gele mit Wasser überschichtet und bei 4°C über Nacht aber maximal für zwei Wochen gelagert.

### 3.2.7 Proteinvisualisierung in SDS-Gelen

#### 3.2.7.1 Coomassie Färbung

Die Gele wurden nach der Elektrophorese für mindestens 1 h, oder über Nacht in Färbelösung unter stetigem Schütteln inkubiert (0.1% CBB-R250 in H<sub>2</sub>O/MeOH/HOAc

(45:45:10)). Anschließend wurde unter mehrmaligem Wechsel der Entfärbelösung (*i*-Prop/HOAc/H<sub>2</sub>O (12.5:10:77.5)) ungebundener Farbstoff entfernt. Bei der Entfärbung mittels *i*-Propanol bleibt das Coomassie an den Proteinen gebunden.

### 3.2.7.2 Silberfärbung nach Vorum

Die Silberfärbung ist neben den auf Fluoreszenzfarbstoffen basierenden Färbungen (Sypro Ruby) eine der sensitivsten Färbemethoden von gelelektrophoretisch aufgetrennten Proteinen. Die Nachweisgrenze liegt bei ca. 10 ng. Im Gegensatz zur herkömmlichen Silberfärbung nach Heukeshoven muss bei Silberfärbungen, die eine anschließende massenspektrometrische Analyse der gefärbten Proteine erlauben sollen, auf den Quervernetzer Glutaraldehyd verzichtet werden. Dies setzt zwar geringfügig die Sensitivität herab, welche aber immer noch deutlich besser ist als die einer herkömmlichen Coomassie-Färbung.

Zur Silberfärbung wurde ein leicht verändertes Protokoll nach Vorum [Mortz *et al.*, 2001] eingesetzt:

Fixieren I:	30 min oder über Nacht: 50% (v/v) Methanol 12% (v/v) Essigsäure
Fixieren II:	30 min: 50% (v/v) Methanol 12% (v/v) Essigsäure
Waschen:	3 x 20 min: 50% (v/v) Ethanol (absolut)
Inkubation:	1 min in 200 mg/ml Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub>
Waschen:	2 x 1 min in MilliQ
Sensibilisierung:	20 min in: 0.2% (w/v) AgNO <sub>3</sub> 0.0375% (v/v) Formaldehyd
Waschen:	2 x 10 sec in MilliQ
Entwickeln:	2-5 min in: 6% (w/w) Na <sub>2</sub> CO <sub>3</sub> 0.025% (v/v) Formaldehyd
Stoppen:	5 min in 5% HOAc

### 3.2.7.3 Immunoblotten von SDS-Gelen

Zur Immunodetektion von in SDS-Gelen aufgetrennten Phosphoproteinen ist seit einiger Zeit eine Vielzahl von unterschiedlichen Antikörpern erhältlich. Sie sind jedoch meist auf die Erkennung von phosphorylierten Epitopen höherer Eukaryoten optimiert. Um eine möglichst „unspezifische“ und dadurch breite Erkennung der in *H. salinarum* phosphorylierten Proteine zu erreichen, wurde ein Gemisch unterschiedlicher Antikörper verwendet. Dieser *Mix* (Biomol) enthielt fünf Anti-Phosphoserin (Sampler

Kit, Prod. No. 54694) und 3 Anti-Phosphothreonin (Sampler Kit, Prod. No. 54695) mit unterschiedlichen Erkennungsepitopen. Die unterschiedlichen Klone waren entweder IgG oder IgM Antikörper.

Cytosolische Proteinpellets, erhalten nach Taurodeoxycholat-Lyse (siehe 3.2.2.1), wurden mittels analytischer SDS-PAGE (siehe 3.2.6.1) aufgetrennt und das Gel für 15 min in Kathodenpuffer (192 mM Glycine, 25 mM /HCl, pH 8.5, 10% Methanol, 1% SDS) äquilibriert. Die PVDF-Membran (Millipore) wurde kurz in Methanol aktiviert und anschließend für 2 min in MilliQ und 5 min in Anodenpuffer (Kathodenpuffer ohne SDS) gewaschen. Der Proteintransfer auf die Membran erfolgte mit der *semi-dry* Methode in einer von der Werkstatt hergestellten Apparatur. Auf die Anode wurden 2 in Anodenpuffer getränkte Whatman-Papiere gelegt, auf die die äquilibrierte PVDF-Membran platziert wurde. Das Gel wurde luftblasenfrei aufgelegt und mit 2 in Kathodenpuffer getränkten Whatman-Papieren bedeckt. Die Apparatur wurde mit dem Deckel, welcher die Kathode darstellt, geschlossen. Der Proteintransfer erfolgte für 1 1/2 h bei 0.8 mA/cm<sup>2</sup> (46 mA). Die Membran wurde kurz in Blotpuffer (2 x PBS (274 mM NaCl, 5.4 mM KCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 3.6 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.2), 1% PEG 3500, 0.2% Tween 20, 1% PVP) gewaschen und anschließend für 2 h in Blockierlösung (Blotpuffer + 1% BSA) inkubiert.

Die Membran wurde dann für 1 h bei RT mit den 8 Antikörpern in Blockierlösung (Konzentration je Antikörper 0.1 µg/ml) inkubiert und nachfolgend für 3 x 10 min mit Blotpuffer gewaschen. Um die gebundenen Antikörper zu detektieren, folgte die Inkubation mit den Zweit-Antikörpern-HRP Konjugaten in Blotpuffer (Anti-IgG und Anti-IgM, beide 1:20,000). Die Chemilumineszenz-Detektion erfolgte mit dem ICL-Kit (Pierce) nach der Gebrauchsanweisung des Herstellers. Die Membran wurde in Saranfolie eingewickelt auf einen Photofilm (Agfa) gelegt und für 5 sec, 30 sec und 5 min entwickelt.

### 3.2.8 Proteinverdau

#### 3.2.8.1 Tryptischer In-Gel Verdau von Proteinen aus 2-D Gelen

Zum automatischen Ausschneiden der im Gel aufgetrennten und mittels Silberfärbung gefärbten Proteine kam der Bruker Proteineer SP „Spot-Picker“ zum Einsatz. Dazu wurde das Gel auf den Durchlichtscanner aufgelegt, fixiert und gescannt. Die auszuschneidenden Proteinspots wurden manuell ausgewählt und automatisch in 96ger Mikrotiter-Filterplatten (Multiscreen Durapor MAHV N45, Millipore, Bedford, MA, USA) überführt (Spot-Durchmesser 1.5 mm). Durch die Verwendung von Mikrotiter-Filterplatten konnten alle Wasch- und Inkubationslösungen durch

Zentrifugation aus den Platten entfernt werden, was ein schnelles Arbeiten mit mehreren Platten gleichzeitig ermöglichte.

Zunächst wurden die silbergefärbten Proteinspots durch Zugabe von je 30  $\mu\text{l}$  eines 1:1 Gemisches aus 30 mM Kaliumhexacyanoferrat (III) und 100 mM Natriumthiosulfat entfärbt [Gharahdaghi *et al.*, 1999]. Die Entfärbung beruht auf der Oxidation des elementaren Silbers ( $\text{Ag}^0$ ) zu Silber-Ionen ( $\text{Ag}^+$ ), die aus dem Gel heraus gewaschen werden können. Nach der Entfärbung wurden die Gel-Spots noch 2mal mit je 100  $\mu\text{l}$  MilliQ gewaschen. Es folgte eine abwechselnde Inkubation der Gelspots mit 50% ACN und 50 mM Ammoniumbicarbonat für je 10 min, wobei die beiden Inkubationsschritte noch einmal wiederholt wurden. Die Spaltung der Proteine erfolgte durch Zugabe von 25  $\mu\text{l}$  Trypsinlösung (20  $\mu\text{g}$  Trypsin gelöst in 25  $\mu\text{l}$  *Storage Buffer*, auf 5.2 ml 50 mM Ammoniumbicarbonat aufgefüllt) über Nacht bei 37°C. Der Überstand wurde am nächsten Tag in eine neue Mikrotiterplatte zentrifugiert und die Peptide in drei weiteren Schritten mittels MilliQ, 50% ACN und 50% ACN, 0.1% TFA eluiert (jeweils 50  $\mu\text{l}$  für 20 min unter Schütteln). Die vereinigten Eluate wurden in flüssigem Stickstoff eingefroren und in einer Vakuumzentrifuge getrocknet. Anschließend wurden die Peptide bis zur Salzfreiheit in jeweils 100  $\mu\text{l}$  Wasser aufgenommen, eingefroren und erneut getrocknet (in der Regel 2mal). Alle Inkubationsschritte erfolgten auf einem temperierbaren Schütteltisch, zum Pipettieren wurden 8-Kanal Pipetten verwendet.

### 3.2.8.2 Tryptischer Verdau aus eindimensionalen SDS-Gelen

Die ausgewählte Bande eines Coomassie-gefärbten präparativen 1-D Gels wurde mit einer Rasierklinge in Banden geschnitten (19-20 Stück), die einzelnen Banden in ca. 1 x 1 x 1.5 mm<sup>3</sup> Stückchen zerkleinert und vollständig in 50% MeOH und 50 mM Ammoniumbicarbonat entfärbt. Die Reduzierung und Caramidomethylierung der Cysteine erfolgte durch Inkubation der Gelstücke in 10 mM DTT, 50 mM Ammoniumbicarbonat bei 56°C und 55 mM IAA, 50 mM Ammoniumbicarbonat (je 30 min). Bei ICPL-makierten Proteinproben konnten diese beiden Schritte entfallen, da die Proteine im Rahmen der Markierung bereits Reduzierung und Caramidomethylierung erfahren hatten. Die Gelstücke wurden nun abwechselnd je 3mal in 50 mM Ammoniumbicarbonat und 50% ACN inkubiert und nach dem letzten 50% ACN-Schritt im Vakuum getrocknet.

Zu den getrockneten Gelstücken wurde soviel Trypsin-Lsg. (10  $\mu\text{g}/\text{ml}$  in 50 mM Ammoniumbicarbonat) hinzu gegeben, bis sie wieder vollständig gequollen waren. Um ein Austrocknen zu verhindern, wurden sie noch zusätzlich mit etwas Trypsin-Lsg. überdeckt. Der Verdau erfolgte bei 37°C über Nacht.

### 3.2.8.3 *In-Lösungsverdau von H. salinarum* Proteinen

Sedimentierte Zellen (5000 g, 15 min, 4°C) oder Proteinpellets wurden in Lysepuffer (8 M Harnstoff oder 5 M Guanidiniumchlorid, 100 mM Tris/HCl, pH 8, 1 mM NaF und je 5mM Orthovanadate bzw.  $\beta$ -Glycerophosphat) aufgenommen (je 1 ml pro Zellsediment aus 35 ml Kultur bei 40 Klett) und im Ultraschallbad lysiert. Die Proteine wurden zunächst durch Zugabe von DTT-Lsg. (Endkonzentration: 2 mM) für 30 min reduziert und mit IAA (Endkonzentration: 12 mM) für ebenfalls 30 min carbamidomethyliert. Entweder wurde die Lösung vor dem Verdau durch Zugabe von 100 mM Tris/HCl, pH 8 auf je 1 M Harnstoff bzw. Guanidiniumchlorid gebracht oder zweimal für 3 h gegen 1.5 M Harnstoff in Slide-A-Lyzer<sup>®</sup> 10K Dialysekassetten (Pierce, Rockford, IL, USA) dialysiert. Die proteolytische Spaltung der Proteine erfolgte durch Zugabe von Trypsin (Enzym:Protein, 1:100) über Nacht. Am nächsten Morgen wurde die Probe durch Zugabe von TFA angesäuert (Endkonzentration: 0.4%, pH ca. 2) und für 30 min bei 213.000g, 18°C zentrifugiert. Der Überstand wurde nach Bedarf mehrmals über C<sub>18</sub>-Kartuschen entsalzt (siehe 3.2.9.2). Bei der Verwendung von Harnstoff als chaotrope Substanz wurden alle Schritte bei Raumtemperatur durchgeführt, um die bei höheren Temperaturen auftretende Proteincarbamidomethylierung in harnstoffhaltigen Lösungen [McCarthy *et al.*, 2003] zu verhindern. Bei der Verwendung von Guanidiniumchlorid wurde die Reduktion bei 55°C und der Verdau bei 37°C durchgeführt.

### 3.2.8.4 *In-Lösungsverdau von BSA und $\beta$ -Casein*

BSA und  $\beta$ -Casein wurden wie in 3.2.8.3 beschrieben mit leichten Modifikationen in Guanidiniumchlorid verdaut. Dazu wurden 10 mg BSA in 424  $\mu$ l Lysepuffer (Proteinkonzentration: 23.6 mg/ml) und 20 mg  $\beta$ -Casein in 300  $\mu$ l Lysepuffer (Proteinkonzentration: 66.6 mg/ml) gelöst. Reduktion und Carbamidomethylierung erfolgte bei 10 mM DTT und 55 mM IAA. 45  $\mu$ l der BSA Lsg. und 16  $\mu$ l der  $\beta$ -Casein Lsg., was Proteinmengen von je 1 mg entsprach, wurden auf 1 M Guanidiniumchlorid verdünnt (Endvolumen: BSA-Lsg. 225  $\mu$ l,  $\beta$ -Casein-Lsg. 80  $\mu$ l) und über Nacht nach Zugabe von 20  $\mu$ g Trypsin (1:50) bei 37°C verdaut. Die Spaltung wurde durch Ansäuern mit TFA gestoppt (End-Peptidkonzentrationen: 70.5 pmol/ $\mu$ l BSA und 564 pmol/ $\mu$ l  $\beta$ -Casein). Zur Etablierung wurde ein Teil der Peptidlösungen wie in 3.2.9.1 beschrieben entsalzt. Definierte Mengen dienten nun zur Etablierung und Überprüfung der IMAC-Methode.

### 3.2.9 Entsalzung von Peptiden mittels *Reversed Phase Material*

Die in Proben enthaltenen Salze können die spätere massenspektrometrische Analyse erheblich beeinträchtigen, wobei Sensitivität und Auflösung der Messung stark herabgesetzt wird. Das Prinzip bei der in dieser Arbeit durchgeführten Entsalzung von Peptidgemischen beruht auf der Bindung von Peptiden an hydrophobe so genannte *reversed phase* (Umkehrphasen) Trägermaterialien in wässriger Lösung (hydrophobe Wechselwirkungen). Bei dem Trägermaterialien handelt es sich um, durch z.B. Oktadecyl-Gruppen ( $C_{18}$ ) hydrophob modifizierte Silicate. Unter wässrigen Bedingungen bleiben die Salze gelöst und können so entfernt werden. Die gebundenen Peptide werden anschließend mit organischem Lösungsmittel eluiert. Folgende Entsalzungsmethoden wurden im Rahmen dieser Arbeit durchgeführt:

#### 3.2.9.1 Entsalzung in GeLoader-Tips mittels $C_{18}$ -Material

Die Entsalzung kleinerer Peptidmengen z.B. aus In-Gel Verdauen kann bequem und schnell in GeLoader-Tips durchgeführt werden [Rappsilber *et al.*, 2003]. Durch Anlegen von Druck mittels einer 1 ml Einmalspritze können die Lösungen durch den Tip gepresst werden.

Dazu wurde ein ca. 1 mm großes Stück aus in Teflon eingebettetem  $C_{18}$ -Material (3M Empore<sup>TM</sup> Extraction Disks) herausgestochen und in den GeLoader eingeführt. Das Äquilibrieren des  $C_{18}$ -Materials erfolgte mit 10  $\mu$ l Isopropanol, gefolgt von 10  $\mu$ l 5% Ameisensäure (FA). Die Peptide wurden in 5% FA gelöst und über die Säule gegeben. Salze wurden durch 2maliges Waschen der Säule mit 5% FA entfernt und die entsalzten Peptide mit 80% MeOH, 5% FA eluiert. Das Trocknen der Peptide erfolgte im Vakuum.

#### 3.2.9.2 Entsalzung mit $C_{18}$ Kartuschen

Zum Entsalzen größerer Peptidmengen z.B. von In-Lösungs Verdauen (siehe 3.2.8.3) wurden 3 ml HD  $C_{18}$ -Kartuschen (Empore<sup>TM</sup> Extraction Cartridges) verwendet. Äquilibrieren, Waschen und Elution erfolgten analog zu 3.2.9.1, allerdings mit größeren Volumina (500  $\mu$ l Isopropanol, 1 ml 0.5% FA, Probe, 2 x 1 ml 0.5% FA, 300  $\mu$ l 80% MeOH, 5% FA).

#### 3.2.9.3 Entsalzung in GeLoader Tips mittels Oligo R3-Material

Die oben beschriebene Entsalzung von Peptiden mittels  $C_{18}$ -Material basiert wie erwähnt auf der hydrophoben Wechselwirkung zwischen Peptiden und Trägermaterial. Sind die Peptide allerdings sehr hydrophil, ist eine Bindung unter wässrigen Bedingungen nicht mehr gegeben. Das Oligo R3-Material, welches zur Aufreinigung von Oligonukleotiden entwickelt wurde, bindet auch hydrophile Peptide wie z.B.

Phosphopeptide. Bei Proben, die möglicherweise Phosphopeptide enthielten, wurde der Probendurchlauf nach der C<sub>18</sub>-Entsalzung noch einmal mittels Oligo R3 entsalzt und das Eluat mit dem ersten vereinigt.

Das Oligo R3-Material wurde in Isopropanol aufgeschlemmt und in einen mit 50 µl 5% FA beladenen GeLoader-Tip pipettiert. Dieser war an seiner Spitze mit einer Pinzette durch Drücken und leichtes Ziehen verengt worden. Die ca. 2 cm lange Säule wurde wie in 1.4.3.1 äquilibriert, gewaschen und eluiert.

### 3.2.10 Anreicherung von Phosphopeptiden

Die Analyse von Phosphopeptiden mittels MS gestaltet sich in mehrfacher Hinsicht als problematisch: Zum einen lassen sich Phosphopeptide aufgrund der zusätzlichen negativen Ladungen der Phosphatgruppe schlechter durch Protonierung ionisieren. Dies ist jedoch eine absolute Voraussetzung gängiger MS-Analytik. Zum anderen ist in einem Peptidgemisch nur ein Bruchteil der Peptide phosphoryliert. Diese Umstände machen eine Anreicherung von Phosphopeptiden vor der massenspektrometrischen Analyse unumgänglich.

#### 3.2.10.1 Anreicherung mittels SCX-Chromatographie

Die *strong cation exchange* (SCX) Chromatographie wurde kürzlich von Gygi und Mitarbeitern zur Anreicherung von Phosphopeptiden vorgestellt [Beausoleil *et al.*, 2004; Ballif *et al.*, 2004]. Die Methode basiert auf der früheren Elution von schwächer positiv geladenen tryptischen Peptiden. Diese besitzen im sauren (pH 2.7 – 3) durchschnittlich zwei positive Ladungen, die eine am C-terminalen Lysin oder Arginin (die Schnittstelle von Trypsin), die zweite an der N-terminalen Aminogruppe (Nettoladung: +2). Phosphopeptide bringen hingegen durch die stark saure Phosphatgruppe eine zusätzliche negative Ladung in das Peptid (Nettoladung: +1) ein. Daher sollten Phosphopeptide bereits bei niedrigen Salzkonzentrationen vom Kationenaustauscher eluieren, bei denen die restlichen Peptide noch gebunden bleiben.

Die Chromatographie wurde mit einer Resource S Säule (Amersham Biosciences) auf einem Äkta Explorer LC System (Amersham Biosciences) durchgeführt. Nach Äquilibrieren der Säule mit 100% Puffer A (5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 2.7, 30% ACN) wurden die Peptide nach In-Lösungsverdau in 8 M Harnstoff (siehe 3.2.8.3) direkt über einen 10 ml Loop auf die Säule gegeben. Die Flussgeschwindigkeiten des nachfolgenden Gradienten betragen 1 ml/min. Nach Beladen der Säule wurde diese für 15 min mit 100% Puffer A gewaschen. Die Elution der schwach bindenden Peptide erfolgte durch einen linearen Gradienten über 30 min auf 15% Puffer B (5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 2.7, 30%

ACN, 350 mM KCl). Nach einem 2 minütigen Plateau bei 15% B wurden die restlichen Peptide mit 100% Puffer B von der Säule gewaschen.

Es wurden 2 ml Fraktionen bis zum Anstieg auf 100% Puffer B gesammelt. Die eluierten Peptide der einzelnen Fraktionen wurden, wie im Ergebnisteil gezeigt, zusammengefasst und direkt für die TiO<sub>2</sub> Anreicherung verwendet.

### 3.2.10.2 Anreicherung mittels Fe<sup>3+</sup> IMAC in GeLoader Tips

Mittels IMAC können Phosphopeptide aus komplexen Peptidgemischen angereichert werden. Die Methode basiert auf der Interaktion zwischen negativ geladenen Phosphatgruppen und dem an das Trägermaterial koordiniertem Metall-Ion. Als Trägermaterial diente das POROS 20MC Harz (Applied Biosystems), welches mit Fe(III)-Ionen aktiviert wurde.

Zunächst wurde das POROS Material in 0.1M HOAc aufgeschlemmt und in 1 ml Mobicol-Säulchen (Mobictec) pipettiert. Das Waschen und Äquilibrieren des Säulenmaterials zur Aktivierung der Säule erfolgte mit einer aufgeschraubten 60 ml Spritze unter stetigem Druck:

- 30 ml H<sub>2</sub>O
- 30 ml 50 mM EDTA in 1M NaCl
- 30 ml H<sub>2</sub>O
- 30 mL 0.1 M HOAc
- 50 ml 0.1 M FeCl<sub>3</sub> in 0.1 M HOAc

Das aktivierte POROS Material wurde in einem 1.5 ml Eppendorfreaktionsgefäß in 0.1 M FeCl<sub>3</sub> in 0.1 M HOAc bei 4°C gelagert. Die IMAC Anreicherung erfolgte in GeLoader-Säulchen. Dazu wurde die Spitze eines GeLoader-Tips mit einer Pinzette zusammengedrückt und etwas in die Länge gezogen, was die Spitze stark verengte. Das aktivierte POROS Material wurde 1:1 mit 0.1 M HOAc gemischt und in einen GeLoader Tip gegeben, der mit 20 µl 0.1 M HOAc gefüllt war. Das Packen der Säule erfolgte unter Druck mit einer 1 ml Einmalspritze, so dass die Spitze des GeLoader mit POROS Material gefüllt war (ca. 3 - 4 cm). Die Säule wurde nun mit folgenden Lösungen unter Druck (s.o.) gewaschen und äquilibriert:

- 50 µl H<sub>2</sub>O
- 50 µl 250 mM NaCl
- 50 µl 0.1 M HOAc
- 50 µl 0.1 M HOAc : ACN (3:1 v/v)

- 50  $\mu$ l 0.1 M HOAc

Die vorher entsalzte Probe (siehe 3.2.9) wurde in 5 - 30  $\mu$ l 0.1 M HOAc gelöst und auf die äquilibrierte IMAC-Säule gegeben. Das Beladen und anschließende Waschen der Säule erfolgte unter leichtem Druck (50  $\mu$ l der 1 ml Spritze). Der Probendurchlauf sowie die Waschfraktionen wurden gesammelt. Das Waschen der Säule erfolgte mit 30 – 50  $\mu$ l der folgenden Lösungen:

- 0.1 M HOAc
- 0.1 M HOAc : ACN (3:1 v/v)
- 0.1 M HOAc

Die gebundenen Peptide wurden zweimal mit je 5  $\mu$ l 0.2 M Na<sub>2</sub>HPO<sub>4</sub> eluiert. Das basische Eluat wurde durch Zugabe von 20  $\mu$ l 2% TFA angesäuert (ca. pH 2) und wie in 3.2.9 beschrieben entsalzt.

Um die IMAC-Methode zu etablieren und ihre Güte sowie Reproduzierbarkeit zu überprüfen, wurden in regelmäßigen Abständen Testmessungen mit einem Gemisch aus tryptisch verdaulichem BSA und  $\beta$ -Casein (siehe 3.2.8.4) durchgeführt.  $\beta$ -Casein besitzt zwei Phosphopeptide, die als Testpeptide zur Evaluation der IMAC-Methode benutzt werden können. Es sollte also eine Anreicherung der Phosphopeptide des  $\beta$ -Caseins gegenüber der Abreicherung der nicht phosphorylierten BSA Peptide beobachtet werden. Dazu wurden je 1 pmol der entsalzten tryptischen Peptide von BSA und  $\beta$ -Casein gemischt, mit 0.1 M HOAc auf 30  $\mu$ l verdünnt und der IMAC-Prozedur unterzogen. Der Probendurchlauf, die Waschfraktionen und IMAC-Eluate wurden mittels MALDI-TOF nach Probenpräparation über Dünnschicht auf HCCA (siehe 3.2.11.3) analysiert.

### 3.2.10.3 Anreicherung mittels TiO<sub>2</sub> Material

Mit der Anreicherung von Phosphopeptiden über TiO<sub>2</sub> Material [Pinkse *et al.*, 2004] können Phosphopeptidfraktionen erhalten werden, die weniger mit störenden nicht phosphorylierten Peptiden verunreinigt sind [Larsen *et al.*, 2005]. Die TiO<sub>2</sub>-Anreicherung wurde ebenfalls mit den frühen SCX-Fraktionen durchgeführt. Dazu wurde in eine 200  $\mu$ l Pipettenspitze das in Methanol aufgeschlemmte TiO<sub>2</sub> Material eingebracht (ca. 2 mm). Das TiO<sub>2</sub> Material war zuvor einer TiO<sub>2</sub>-Kartusche (GL Sciences Inc. Tokyo, Japan) entnommen worden. Als Fritte diente ein Stück Mikro TFA-Filter (Applied Biosystems). Die Pipettenspitze wurde durch die durchstochene Kappe eines 2 ml Eppendorfreaktionsgefäß gesteckt. Die Lösungen zum Waschen,

Eluieren etc. konnten so durch Zentrifugation bei 5000 g über die Säule geleitet werden und wurden im Eppendorfreaktionsgefäß aufgefangen.

Die Säule wurde zunächst mit 30 µl Elutionspuffer 2 (NH<sub>3</sub>OH in 80% ACN, pH ca. 10.5) äquilibriert und danach mit 50 µl Waschpuffer 2 (40% ACN, 0.1% TFA) gewaschen. Hiernach erfolgte die Absättigung der Säule mit 30 µl 6 x DHB Puffer (30 mg/ml DHB in 80% ACN, 0.1% TFA) und mit 1 x DHB Puffer (verdünnt in Waschpuffer 2) gewaschen. Die Proben wurden 5:1 mit 6 x DHB Puffer (30 mg/ml DHB in 80% ACN, 0.1% TFA) verdünnt und auf die Säule gegeben. Danach wurde diese mit 50-100 µl Waschpuffer 1 (10% ACN, 0.1% TFA) und 50-100 µl Waschpuffer 2 gewaschen. Die Elution der gebundenen Peptide erfolgte mit 25 µl Elutionspuffer 1 (NH<sub>3</sub>OH in 10% ACN, pH ca. 10.5) und 25 µl Elutionspuffer 2 in ein frisches Eppendorfreaktionsgefäß. Die Proben wurden anschließend im Vakuum getrocknet.

### 3.2.11 Massenspektrometrische Methoden

#### 3.2.11.1 Automatische Dried-Droplet Probenpräparation

Große Probenmengen im Mikrotiterplatten Format (siehe 3.2.8.1) wurden automatisch für die MALDI-TOF Messung präpariert. Die in den Mikrotiterplatten getrockneten tryptischen Peptide wurden in je 10 µl 33% ACN, 0.1% TFA gelöst und mit dem MAP II Pipettierroboter auf ein 384er MALDI-Target (*stainless steel*) aufgetragen. Dabei wurden je 0.5 µl der Probe mit 0.5 µl Matrixlösung (gesättigte  $\alpha$ -Cyano-4-Hydroxy-Zimtsäure in 40% ACN, 0.1% TFA) gemischt (*Dried-Droplet Methode*). Der für die Kalibrierung der Massenspektren nötige Peptidstandard wurde nach jeder achten Probe erneut aufgetragen, so dass insgesamt die Proben von 3 Mikrotiterplatten auf ein MALDI-Target aufgebracht werden konnten (288 Proben + 36 Standardproben = 324 besetzte Positionen).

#### 3.2.11.2 Automatische Probenpräparation mittels GyroLab<sup>TM</sup>

Die Probenpräparation mittels Gyrolab<sup>TM</sup> (Gyros AB, Uppsala, Schweden) erfolgt auf runden Präparationsscheiben (CD-Format). Die Proben werden bei Rotation durch Mikrostrukturen geleitet und dabei zunächst über C<sub>18</sub> Material entsalzt und anschließend auf die MALDI-Analysebereiche eluiert. Die Präparation erfolgte nach den Angaben des Herstellers. Dazu wurden die Mikrostrukturen und das darin enthaltene C<sub>18</sub> Material mit 50% ACN, 0.1% TFA aktiviert und die in 0.1% TFA gelösten Proben auf dem C<sub>18</sub> Material aufkonzentriert. Nach Waschen der Proben mit 10% Ethanol, 0.1% TFA wurden die Peptide mit Matrixlösung (1 mg/ml  $\alpha$ -HCCA, 1 mg/ml Fukose, 0.4 mg/ml n-Octyl  $\beta$ -D-Glykoparanoside in 50% ACN, 0.1% TFA)

eluiert. Die MALDI-TOF Analyse erfolgte mit den üblichen Parametern (siehe 3.2.11.4).

### 3.2.11.3 Dünnschichtpräparation auf Anker-Targets

Im Gegensatz zu großen Probenmengen, welche automatisiert generiert, vermessen und prozessiert werden müssen (siehe 3.2.11.1), kann bei überschaubaren Probenmengen eine manuelle, komplexere und damit zeitaufwändigere Probenpräparation angewendet werden, die die Sensitivität der Messung allerdings verbessern kann. Die Dünnschichtpräparation auf Anker-Targets (Bruker Daltonics) bietet mehrere Vorteile: Zum einen wird auf den kleinen hydrophilen Ankern die Probe stark aufkonzentriert. Durch nachträgliches Waschen der Probe können Salzkontaminationen vor der Messung entfernt werden. Mit dieser Kombination wird eine Erhöhung der Sensitivität im Gegensatz zur herkömmlichen *Dried-Droplet* Methode auf Standart-Targets erreicht.

In Aceton gelöste HCCA-Matrix wurde durch Kapillarkraft in einen GeLoader aufgesogen und durch kurzes Berühren der Anker auf das Target aufgebracht. Für 600 µm Anker wurde eine Matrixkonzentration zwischen 15 und 18 mg/ml eingesetzt. Nach 1-2 sec. hatte sich eine dünne Schicht von Matrixkristallen auf dem Anker ausgebildet. Die Proben wurden in 0.2% TFA gelöst und 0.5 - 1 µl auf den Anker aufgetragen. Nach 3 – 5 min wurde der Probentropfen vom Anker mit einem leichten Luftstrom vom Anker entfernt und die Probe noch einmal kurz mit 0.5 µl 0.2% TFA gewaschen.

### 3.2.11.4 MALDI-TOF PMF-Messung

Bei der PMF (*peptide mass fingerprint*) Methode wird ein Protein proteolytisch in Peptide gespalten und die entstandenen Fragmente im Massenspektrometer analysiert. Besteht Kenntnis über die Sequenz des zu analysierenden Proteins (z.B. aus Genomsequenzen) kann dieses anhand seines Peptid Fingerabdrucks identifiziert werden, falls genügend Fragmente eine eindeutige Zuordnung erlauben. Beim Anfallen großer Probenmengen wie z.B. nach dem automatischen Ausschneiden und Probenverdau aus 2-D Gelen im Mikrotiterplatten Format (siehe 3.2.8.1) konnte auf einen automatisierten Messvorgang nicht verzichtet werden. Die MALDI-TOF Massenspektren wurden automatisch mit einem REFLEX III Massenspektrometer (Bruker Daltonics) durch Einschalten der *Fuzzy Control* Automatik aufgenommen. Wichtige Geräteeinstellungen waren hierbei die anfängliche Abschwächung des Lasers von 85% und die minimale Abschwächung von 60%. Für ein Einzelspektrum wurden 200 „Schüsse“ summiert bei einer maximalen Schuss-Summe von 20 auf einer Position. Bei kleineren Probenmengen konnte auf die Automatisierung verzichtet werden.

Die Beschleunigungsspannung des Spektrometers betrug 20 kV, die Detektorspannung lag bei einem neuen Detektor bei 1600 V und wurde nach 1.5 Jahren auf 1800 V erhöht. Die Kalibrierung der Massenspektren erfolgte zunächst extern über den auf benachbarten Positionen aufgetragenen Peptid-Standard-Mix. Dieser Standard-Mix enthielt acht Peptide mit Massen zwischen 1046.54 und 3494.65 Da. Für die automatische Annotation der aufgenommen Massenspektren wurde das Xmass-Programm (Version 5.1.7, Bruker Daltonics) verwendet. Anschließend erfolgte eine zweite Kalibrierung der Spektren [Zubarev *et al.*, 1996]. Diese wurde in enger Zusammenarbeit mit Bruker-Daltonics durchgeführt. Es konnte so eine durchschnittliche Massengenauigkeit von 25 ppm erhalten werden.

#### 3.2.11.5 Reinigung der MALDI-Targets

Standard Bruker Stahl-Targets wurden zunächst mit MeOH gewaschen und danach für 15 min in H<sub>2</sub>O (MilliQ) im Ultraschallbad sonifiziert. Das Target wurde nun an der Luft getrocknet, mit Isopropanol gewaschen und anschließend darin für 15 min sonifiziert. Nach erneutem Spülen mit Isopropanol wurde das Target im Ofen bei 100°C für mindestens 2 h getrocknet.

Bruker Anchor-Targets wurden mit Aceton unter vorsichtigem Reiben (Kimwipe) gewaschen, um die hydrophobe Beschichtung nicht zu beschädigen. Danach wurde das Target für 10 min in 50% MeOH sonifiziert, anschließend noch einmal mit MeOH gespült und an der Luft getrocknet.

Die Applied Biosystems *Blank*-Targets zur Messung der ICPL-Proben mittels LC-MALDI-TOF/TOF (siehe 3.2.11.9) wurden mit 50% ACN, 0.1% TFA benetzt und mit einem staubfreiem Tuch (Kimwipe) abgewischt. Es folgte die Ultraschallbehandlung in H<sub>2</sub>O (MilliQ) für 5 min. Das getrocknete Target wurde dann mit Pol-Paste poliert, für 5 min in Isopropanol sonifiziert und mindestens für 2 h im Ofen bei 100°C getrocknet.

#### 3.2.11.6 Packen von Nano-LC Säulen

Eine auf 15 cm zurechtgeschnittene *Fused Silica* Kapillare (75 µm i.d.) wurde zunächst mit einer Fritte versehen, um das Packen mit Säulenmaterial zu ermöglichen. Dazu wurden 80 µl Kasil (Kaliumsilicat) mit 16 µl Formamid gemischt, ein Ende der Kapillare kurz in die Mischung eingetaucht und die Fritte über Nacht im Ofen bei 100°C ausgehärtet. Zum Packen der Säule wurde Säulenmaterial (Reprosil-Pur120 C<sub>18</sub>-Material, 3 µm *particle size*) in MeOH aufgeschlemmt (ca. 50 mg/ml) und unter ständigem Rühren mit einer Druckbombe bei 50 bar in die Säule gepresst. Für das Anfertigen einer Vorsäule wurde eine 2.5 cm lange Kapillare (100 µm i.d.) verwendet (Säulenmaterial: Reprosil-Pur120 C<sub>18</sub> Material, 5 µm *particle size*).

### 3.2.11.7 Nano-LC ESI-MS/MS

Die Proben wurden in 5% FA aufgenommen, die chromatographische Auftrennung der Peptide erfolgte mit dem CapLC-System (Waters). Es wurden Probenvolumina von 5  $\mu\text{l}$  über einen *Autosampler* injiziert. Die Probe wurde zunächst auf die Vorsäule mit einem Fluss von 2  $\mu\text{l}/\text{min}$  geladen. Die Trennung erfolgte über eine selbstgepackte analytische Säule (siehe 3.2.11.6) bei einem Fluss von 200  $\text{nl}/\text{min}$  mit folgendem Gradienten:

Nach 8minütiger Spülphase bei 0% B (100% A) in 4 min auf 10% B, in 5 min auf 15% B, in 10 min auf 20% B, in 20 min auf 25% B, in 20 min auf 35% B, in 15 min auf 50% B, in 5 min auf 100% B (Eluent A: 0.5% FA, 2% ACN, Eluent B: 0.5% FA, 80% ACN).

Die Analyse der eluierenden Peptide erfolgte mit dem Q-TOF Ultimate Massenspektrometer (Waters) im *positive ion mode*. Wichtige Geräteeinstellungen hierbei waren eine Kapillarspannung von 2.3 kV. Das Massenfenster der zu analysierenden Ionen betrug 300-1500 Da für den MS- und 50-2000 Da für den MS/MS Modus. Nach der Aufnahme des MS Spektrums der eluierenden Peptide für 1.5 sec. erfolgten die Auswahl und Fragmentierung der 4 intensivsten Signale für je 1.5 sec. Die Kollisionsspannung wurde je nach Peptidmasse und Ladungszustand variiert. Sie betrug zwischen 16 und 40 kV. Helium diente als Kollisionsgas. Die weitere Prozessierung der Daten wurde mit der MassLynx Software (Waters, Version 4.0) durchgeführt, welche die Peaklisten zur anschließenden Mascot Suche (siehe 3.2.12) generierte.

Die Analysen auf dem LTQ FT-ICR Massenspektrometer wurden in der Arbeitsgruppe von Prof. Mann durchgeführt. Die Messungen der ICPL markierten Proben sowie der mittels SCX und  $\text{TiO}_2$  angereicherten Phosphopeptide erfolgten nach etablierten Protokollen [Olsen und Mann, 2004; Gruhler *et al.*, 2005].

### 3.2.11.8 Off-line ESI-MS/MS

Zur *off-line* Elektrospray Messung wurde die Probe 1-2  $\mu\text{l}$  in 50% MeOH, 2% FA aufgenommen und 0.5  $\mu\text{l}$  in eine auf die richtige Länge zurechtgeschnittene Elektrospray-Nadel (Protana) pipettiert. Durch Zentrifugation wurde der Probentropfen in das spitze Ende der Nadel gepresst. Die Nadel wurde in den entsprechenden Halter am Massenspektrometer eingebaut. Die Messung erfolgte bei einer Kapillarspannung von 1.3 kV. Nach manueller Auswahl des Peptids wurde dieses bei einer Kollisionsspannung zwischen 20-38 kV fragmentiert.

### 3.2.11.9 Nano-LC MALDI MS/MS der ICPL Proben

Die chromatographische Auftrennung der Peptide erfolgte mit dem Ultimate HPLC-System (LC Packings). Es wurden Probenvolumina von 50 µl injiziert, wobei die Peptide zunächst auf einer *reversed phase* C<sub>18</sub> Vorsäule gesammelt wurden. Die Trennung erfolgte über eine analytische *reversed phase* C<sub>18</sub> Säule (75 µm i.d., 15 cm), die direkt mit dem Probot Pipettiersystem (LC Packings) verbunden war. Die mobilen Phasen bestanden aus (A) 0.05% TFA und (B) 0.04% TFA in 80% (v/v) ACN. Die Elution der Peptide erfolgte über einen 65 min. linearen Gradienten von 10% bis 45% B, gefolgt von einer 20 min. Waschphase bei 100%. Die Flussrate betrug stets 200 nl/min. Das Eluat wurde in einem MicroTee Verbindungsstück (Upchurch), in dem die analytische Säule mündete, mit der Matrixlösung (5 mg/ml HCCA, 5 mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub>, 50% (v/v) ACN, 0.1% TFA) gemischt (Flussrate der Matrixlösung: 1.3 µl/min) und auf ein MALDI Probensteller (*blank, barcoded*) aufgetragen. Während der Elution über einen Zeitraum von 66.66 min wurden Fraktionen a 10 sec pipettiert, was eine Anzahl von 400 Fraktionen pro LC-Lauf auf einem MALDI-Probensteller ergab.

Die Aufnahme der Massenspektren erfolgte mit dem 4700 Proteomics Analyzer (Applied Biosystems), ausgestattet mit einem 200 Hz Nd-Yag Laser (355 nm) bei 2500 Laserschüssen im *positive-reflector* Modus. Mit Hilfe der Peakpicker-Software (Applied Biosystems) wurden nicht markierte Peptide sowie die isotopenmarkierten Peptide (Massenabstand der Peaks: 6, 12, 18, ....Da) ausgewählt und quantifiziert. Die MALDI-TOF/TOF (MS/MS) Analyse der ausgewählten Peptide erfolgte in einem zweiten Durchlauf bei 1500 Laserschüssen pro Peptid und einer Kollisionsenergie von 1 keV mit Stickstoff als Kollisionsgas.

### 3.2.12 Datenprozessierung

#### 3.2.12.1 Mascot Recherche der PMF-Daten

Die automatische Recherche der Daten gegen die *Halobacterium salinarum* Datenbank, welche alle 9,312 ORFs enthielt, erfolgte mit dem Biotools Programm, das die lokal installierte Mascot (Matrix Science) Suchmaschine integrierte. Die Suchparameter waren: 0 erlaubte *missed cleavage sites*, +/-100 ppm Massentoleranz, Carbamidomethylierung von Cysteinen als feste und keine erlaubten variablen Modifikationen.

Mascot Suchergebnisse basieren auf dem MOWSE Score Algorithmus (für Details siehe [www.matrixscience.com](http://www.matrixscience.com) [Perkins *et al.*, 1999], welcher für die *Halobacterium salinarum* Datenbank eine Signifikanzgrenze von 5% bei einem Score von 52 ansetzt. Dies bedeutet, dass Ergebnisse mit einem solchen Score oder höher

höchstens 5% falsch-positive Ergebnisse enthalten. Genaue bioinformatische Analysen der erhobenen Datensätze zeigten jedoch eine weitaus höhere Fehlerrate, was zur Erhöhung der Scoregrenzen für identifizierte Proteine führte. Nach Vereinbarung gelten Proteine ab einem Score von 72 als identifiziert ( $52 \leq \text{Score} < 92 = \textit{normal}$ ;  $\text{Score} > 92 = \textit{reliable}$ ).

Suchergebnisse, die trotz eines zu geringen Mascot Scores als „richtig identifiziert“ angesehen wurden (falsch-negative), konnten in einem nachträglichen manuellen Verfahren als identifizierte Proteine ausgesucht und annotiert werden (*m = manual*). Als Auswahlkriterien galten hier (a) der Score, (b) die Anzahl der identifizierten Peptide bezogen auf die Gesamtzahl der theoretisch möglichen Peptide, (c) eventuelle Überlagerung von zugewiesenen Peptiden mit Peptidmassen tryptischer Fragmente oder regelmäßig auftretenden Massenpeaks, (d) das Auftreten von *missed cleavage sites*, welche in der automatischen Datenanalyse nicht berücksichtigt wurden, und (f) die Korrelation zwischen dem theoretischen  $pI/M_r$  und der Gelposition des identifizierten Proteins.

### 3.2.12.2 Mascot Recherche der ICPL Daten und manuelle Nachanalyse

Die Recherche der ICPL Daten (LC-MALDI TOF/TOF) erfolgte über den GPS Explorer, der die Mascot Suchmaschine integrierte. Die Suchparameter waren hierbei: 1 erlaubte *missed cleavage site*, +/-150 ppm Massentoleranz für die MS und +/-0.4 Da für die MS/MS Messungen und Carbamidomethylierung von Cysteinen als feste Modifikation. Als variable Modifikation wurde die Oxidation von Methioninen erlaubt. Zusätzlich mussten die chemischen Modifikationen, die über die ICPL Tags in das Protein eingefügt wurden, berücksichtigt werden. Dies waren die Nicotinoylierung aller Lysine sowie der N-terminalen Aminogruppe des Proteins mit dem leichten (0 x D) und dem schweren Reagenz (4 x D oder 6 x  $^{13}\text{C}$ ). Als Spaltenzym musste Arg-C (Spaltung C-terminal nach Argininen) gewählt werden, da Trypsin nach der chemischen Markierung die Lysin-Schnittstelle nicht mehr erkennt. Für die Recherche wurde hier eine neuere Version der *H. salinarum* Datenbank verwendet, die nach Abgleich mit dem annotierten Genom von *Natronomonas pharaonis* korrigiert worden war. Ebenfalls enthielt die Datenbank nur die 2,821 ORFs, welche bis zu diesem Zeitpunkt als reale Proteine angesehen wurden.

Die vom GPS Explorer als Excel-Tabellen ausgegebenen Mascot Ergebnisse wurden nun nachträglich mit der Spotfire Software bearbeitet. Die festgelegten Scoregrenzen für die identifizierten Peptide wurden durch die Recherche eines gesamten Datensatzes gegen eine *nonsense* Datenbank ermittelt. In dieser Datenbank waren alle Aminosäuresequenzen der Proteine umgedreht, was zur Folge hatte, dass die

meisten tryptischen Fragmente zwar die gleiche Masse (ausgenommen sind die N- und C-terminalen Peptide), nicht aber das gleiche MS/MS Fragmentmuster aufwiesen. Nach Einstellung dieser Scoregrenzen (total ion score: 25, ion score 20 für Singletts und total ion score: 20, ion score 15 für pairs) wurden zunächst alle falsch positiv identifizierten Peptide gelöscht. Dieses waren Peptide, bei denen die Anzahl der Modifikationen nicht mit der Peptidsequenz übereinstimmte (Anzahl der Lysine + evtl. N-Terminus). Als zusätzliches Kriterium zur Überprüfung der Mascot Ergebnisse für die markierten Peptide diente die Anzahl der Isotopenlabel verglichen mit dem Abstand der Peakpärchen im Massenspektrum. Aufgrund dieser zusätzlichen Überprüfungsmöglichkeit ergab sich für markierte Peptide eine geringe Anzahl an falsch positiven Identifizierungen. Es konnten so niedrigere Scoregrenzen für Isotopenpaare angesetzt werden. Peptide, deren Regulation (Ratio) stark von denen anderer Peptiden des gleichen Proteins abwich, sowie alle Regulationen, die nur auf dem Ergebnis von einem einzigen Peptid beruhten, wurden nachträglich manuell überprüft.

### 3.2.12.3 Mascot Recherche der LC ESI-MS/MS Daten

Die von der MassLynx Software (Waters, Version 4.0) generierten Peaklisten konnten in die Mascot Suchmaschine eingeladen und recherchiert werden. Hierfür wurde die bis dato aktuelle Version der *H. salinarum* Protein Datenbank verwendet. Die Suchparameter waren hierbei: 1 erlaubte *missed cleavage site*, +/-0.15 Da Massentoleranz für die MS und +/-0.15 Da für die MS/MS Messungen sowie Carbamidomethylierung von Cysteinen als feste Modifikation. Die erlaubten variablen Modifikationen wurden je nach Art des Experiments ausgewählt (z. B. Phosphorylierung von Serinen, Threoninen und Tyrosinen oder N-terminale Acetylierung).

### 3.2.12.4 Mascot Recherche der LTQ FT-ICR Daten

Die Peaklisten wurden von dem in der Abteilung Mann geschriebenen Programm DTA Supercharge (Version 0.88) generiert. Die Mascot Recherche erfolgte ebenfalls gegen die aktuelle *H. salinarum* Protein Datenbank. Suchparameter waren hierbei: 1 erlaubte *missed cleavage site*, +/-10 ppm Massentoleranz für die MS und +/-0.5 Da für die MS/MS Messungen sowie Carbamidomethylierung von Cysteinen als feste Modifikation. Bei der Recherche von ICPL-modifizierten Peptiden wurden zusätzlich die entsprechenden chemischen Modifikationen als variable erlaubt. Bei der Recherche von Phosphopeptiden wurde zusätzlich die Phosphorylierung von Serinen, Threoninen und Tyrosinen sowie der *Neutral Loss* von Phosphoserinen und Phosphothreoninen

erlaubt. Die Addition der MS<sup>3</sup>-Scores erfolgte mit dem MSQuant Programm (Version 1.4.0a12; <http://msquant.sourceforge.net>), da dies mit der aktuellen Mascot Version noch nicht möglich war.

### 3.2.12.5 Korrelationsanalyse der 2-DE Identifizierungen

Die Korrelation zwischen dem theoretischen und dem sich aus der Gelposition ergebenden *pI* bzw. *M<sub>r</sub>*-Wert wurde wie folgt ermittelt:

Der theoretische *pI*-Wert wurde mit einem Programm des ExPasy Servers berechnet [Gasteiger *et al.*, 2003]. Das Molekulargewicht ergab sich aus der Primärsequenz der Proteine. Die Positionen der Proteinspots im 2-D Gel wurden vom Spotpicker Bruker SP aufgezeichnet. Für die weiteren Berechnungen wurden nur Gele mit mindestens 40 Identifizierungen berücksichtigt. Für die *pI*-Korrelation wurde eine lineare Ausgleichsgerade mit dem Programm R berechnet [Ihaka und Gentleman, 1996] und als Kalibrierung zur Ermittlung des experimentellen *pH*-Werts an der jeweiligen Position im Gel (horizontale Richtung) benutzt. Die Abweichungen der einzelnen Datenpunkte zur errechneten Ausgleichsgeraden zeigten eine Gauß-Verteilung. Es wurden für jedes Gel die Standardabweichungen berechnet. In einem iterativen Prozess wurden die Proteine, die mehr als drei Standardabweichungen aufwiesen, ausgeschlossen und die Berechnung wurde wiederholt, bis alle Proteine innerhalb der drei Standardabweichungen lagen. Es mussten jeweils maximal drei Datenpunkte (Proteine) ausgeschlossen werden. Anschließend wurden die Standardabweichungen zur finalen Ausgleichsgeraden für jeden Datenpunkt berechnet. Für die *M<sub>r</sub>*-Korrelation wurde die Ausgleichsgerade zwischen  $\log M_r$  und Gelposition (vertikale Richtung) sowie die sich daraus ergebenden Standardabweichungen in gleicher Weise wie für die *pI*-Daten beschrieben (s.o.) erhoben.

### 3.2.12.6 Berechnung des Codon Adaptation Index

Die Berechnung der *Codon Adaptation Indices* erfolgte wie bereits beschrieben [Carbone *et al.*, 2003; Grote *et al.*, 2005] von R.W.W. Brouwer in der Bioinformatikgruppe der Abteilung für Membranbiochemie unter der Leitung von Dr. Friedhelm Pfeiffer.

### 3.2.12.7 Integration der Proteomik-Daten in das HALOLEX System

Die in der Proteomik anfallenden Datenmengen (MS-, MS/MS-Spektren, die daraus resultierenden Proteinidentifikationen, Gelabbildungen etc.) bedürfen einer sorgfältigen Archivierung, anwenderorientierten Darstellung sowie einer ständigen Nachbearbeitung oder Reinterpretation (verbesserte Protein Datenbanken, veränderte

Suchparameter etc.). Dies ist nur durch die Zusammenarbeit mit der Bioinformatik und dem dadurch verbundenen Aufbau eines leistungsfähigen Datenbanksystems möglich.

Alle erhobenen Proteom-Daten werden in dem von der durch Dr. Friedhelm Pfeiffer geleiteten Bioinformatik Gruppe der Abteilung Membranbiochemie (Prof. Dieter Oesterhelt) entwickelten HALOLEX System ([www.halolex.mpg.de](http://www.halolex.mpg.de)) gespeichert. HALOLEX bildet hierbei die zentrale Plattform, in welcher neben den Proteom-Daten auch Genom- und Transkriptom-Daten zusammenfließen.

#### 3.2.12.8 2-D Gel Analyse mittels Proteomweaver

Die silbergefärbten 2-D Gele wurden als TIF-files mit 300 dpi über das BrukerSP System eingescannt und im Proteomweaver Programm (Definiens, München Deutschland) geöffnet. Für das automatische *matchen* der Gelbilder wurden die default Einstellungen gewählt unter Aktivierung der Kraterspoterkennung. Diese ist für die Bearbeitung silbergefärbter Gele vorzuziehen.

#### 3.2.12.9 Proteinstruktur Modellierung

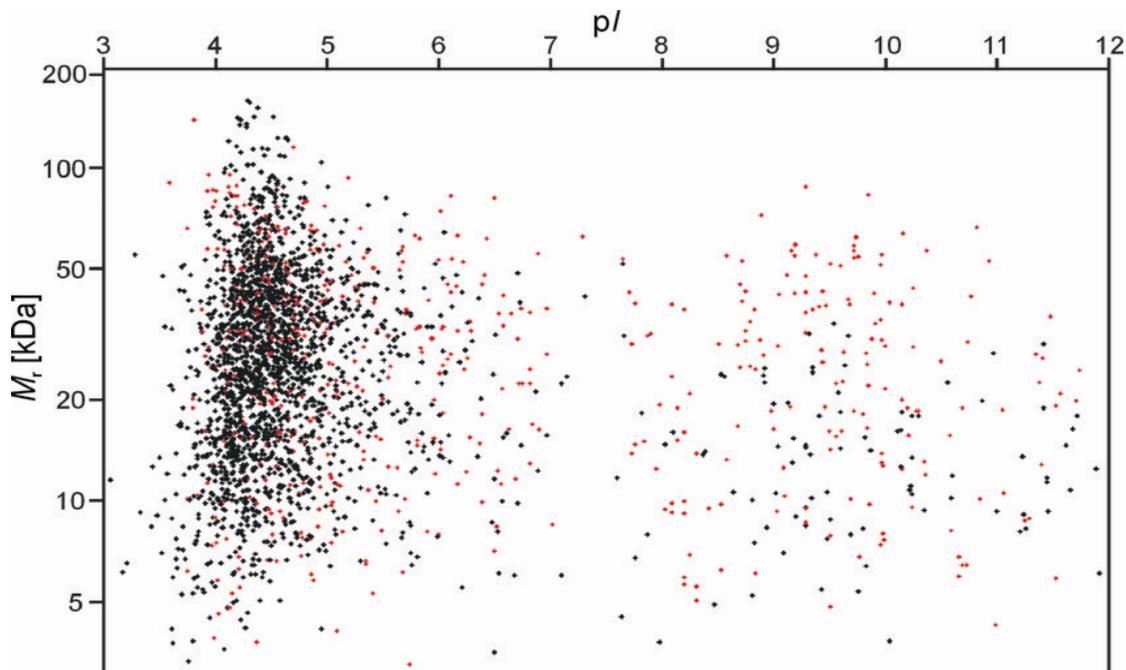
Die Protein Strukturmodelle wurden mit SWISSMODEL ([www.swissmodel.expasy.org](http://www.swissmodel.expasy.org)) im *first approach* Modus mit den jeweiligen Proteinsequenzen (OE2318R, OE4094F und OE4190F) als Template generiert. OE2318R und OE4094F konnten dabei erfolgreich in die Raumkoordinaten (pdb files) der bekannten Strukturen aus *Pseudomonas aeruginosa* (1pcj) und *Pyrococcus horikoshii* (1waq) modelliert werden. Die Darstellung und Bearbeitung der Strukturen erfolgte im DINO-Programm ([www.dino3D.org](http://www.dino3D.org)).

## 4 Ergebnisse

### 4.1 Das zweidimensionale Referenzgel

Die Proteinauftrennung mittels zweidimensionaler Gelelektrophorese (2-DE) erlaubt die Kartierung ganzer Proteome. Derartige Referenzgele, auf denen eine Vielzahl der aufgetrennten Proteine identifiziert ist, können als Grundlage unterschiedlicher experimenteller Ansätze dienen.

Durch die Kenntnis aller in einem Proteom vorkommenden Proteine lässt sich ein theoretisches zweidimensionales Gel erstellen, um die Lage der Proteine auf realen Gelen abzuschätzen. Dabei werden die aus der Primärsequenz der Proteine ermittelten theoretischen  $pI$ -Werte gegen die jeweiligen Molekulargewichte aufgetragen. Für die 2,821 Proteine aus *Halobacterium salinarum* ergibt sich dabei ein spezielles Bild (siehe Abb. 4-1):



**Abb. 4-1:** Theoretisches 2-D Gel von *H. salinarum*. Der theoretische  $pI$  (x-Achse) ist gegen das theoretische  $M_r$  (y-Achse) aller potentiellen Proteine von *H. salinarum* aufgetragen. Proteine, die vorhergesagte Transmembranbereiche (TMHMM) aufweisen, sind in rot dargestellt.

Halophile Proteine besitzen aufgrund ihres ungewöhnlich hohen Anteils an sauren Aminosäuren, welche deren Löslichkeit in konzentrierten Salzlösungen vermitteln, eine erhöhte Acidität (durchschnittlicher  $pI$ -Wert von 5.1 in *H. salinarum*). Dabei haben 82% aller Proteine einen  $pI$ -Wert zwischen 3.5 und 5.5. Dieses Phänomen macht sich noch einmal verstärkt bei den *cytosolischen* Proteinen (Proteine ohne vorhergesagte Transmembranbereiche, TMHMM, [Krogh *et al.*, 2001]) bemerkbar. Von ihnen weisen 90% einen  $pI$ -Wert zwischen 3.5 und 5.5 auf, wobei rund 80% innerhalb eines pH-Werts zwischen 4 und 5 liegen. Dies entspricht einer sehr engen Verteilung im sauren Bereich des theoretischen Gels. Mesophile Organismen haben im Vergleich dazu eine breitere  $pI$ -Verteilung ihrer Proteine mit einem sauer bis neutralen (pH 4.5 bis 7) und einem basischen Bereich (pH 9 bis 11), in welchem sich die Proteine verstärkt anhäufen [VanBogelen *et al.*, 1997; Link *et al.*, 1997; Buttner *et al.*, 2001]. Nur knapp 5% aller cytosolischen Proteine von *H. salinarum* weisen einen  $pI$  größer als 7 auf.

#### 4.1.1 Die cytosolische Standardprobe

In der Proteomik bedarf es häufig einer Vorfraktionierung der Gesamtprobe in so genannte Subproteome. Dies hat den Vorteil, dass die Komplexität der Probe verringert wird, mit der Identifizierung der analysierten Proteine eine zelluläre Lokalisation einhergehen kann und ferner bestimmte, in der Gesamtprobe stark verdünnte Proteine angereichert werden können [Righetti *et al.*, 2003]. Bei höheren Organismen kann eine derartige Vorfraktionierung anhand der vorgegebenen zellulären Kompartimentierung erfolgen und so die Subproteome der unterschiedlichen Organellen analysiert werden [Burnet *et al.*, 2003]. Eine einfache Art der Subfraktionierung ist die Einteilung in lösliche und nicht-lösliche Proteine. Nach Aufschluss der Zellen können diese beiden Fraktionen durch Zentrifugation voneinander getrennt werden. Dabei wird angenommen, dass sich in der löslichen Fraktion hauptsächlich Proteine des Cytoplasmas und in der nicht-löslichen Fraktion überwiegend Membranproteine befinden. Abhängig vom Zellaufschluss können diese Subfraktionen jedoch unterschiedlich stark mit Proteinen der anderen Fraktion kontaminiert sein.

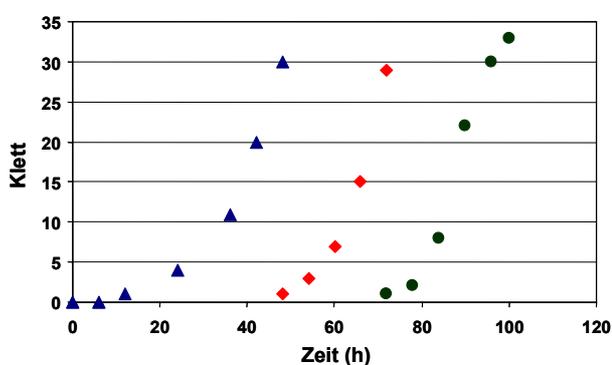


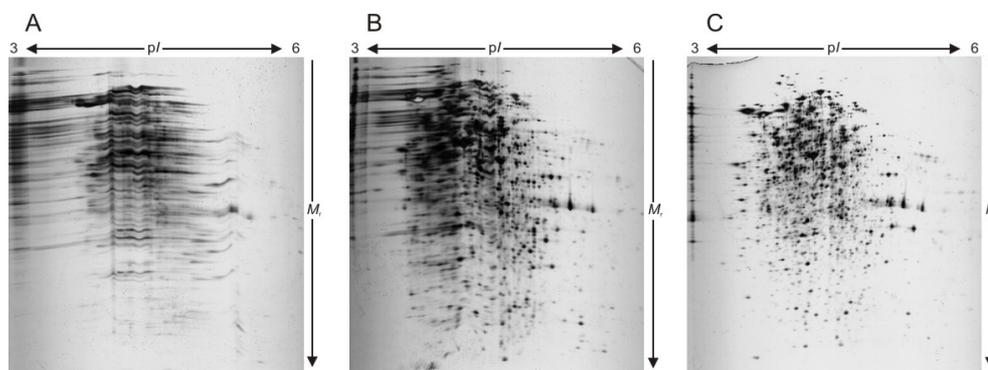
Abb. 4-2: Wachstumskurven zur Anzucht von *H. salinarum*.

Das Erstellen der Referenzgele für *Halobacterium salinarum* (R1, DSM 671) sollte sich auf die cytosolischen Proteine beschränken da die zweidimensionale gelelektrophoretische Auftrennung von Membranproteinen mit den herkömmlichen Verfahren nicht zu erreichen ist [Klein *et al.*, 2005]. Die Zellen wurden unter aeroben Wachstumsbedingungen ohne Licht zu einer Zelldichte von ca. 40 Klett ( $4 \times 10^8$  Zellen/ml) durch drei aufeinander folgende Überimpfungen in Vollmedium angezogen (siehe dazu Wachstumskurve, Abb. 4-2).

Um eine Kontamination der cytosolischen Probe mit Membranproteinen oder peripheren Membranproteinen zu vermeiden, wurde eine Zell-Lyse mit dem Detergenz Taurodeoxycholat [Kamekura *et al.*, 1988; Kupper *et al.*, 1994] unter Hochsalzbedingungen durchgeführt. Durch Einlagerung des Detergenz in die Zellmembran wird diese *löchrig* und die cytosolischen Proteine werden so freigesetzt. Die Zellhüllen, welche integrale als auch periphere Membranproteine enthalten, bleiben unter Hochsalzbedingungen erhalten und können durch Zentrifugation sedimentiert und so von den cytosolischen Proteinen abgetrennt werden. Dies ermöglichte eine signifikante Abreicherung des hochabundanten Glykoproteins, welches die S-Layer zur Strukturgebung der Zellen ausbildet.

#### 4.1.2 Zweidimensionale Auftrennung des cytosolischen Proteoms

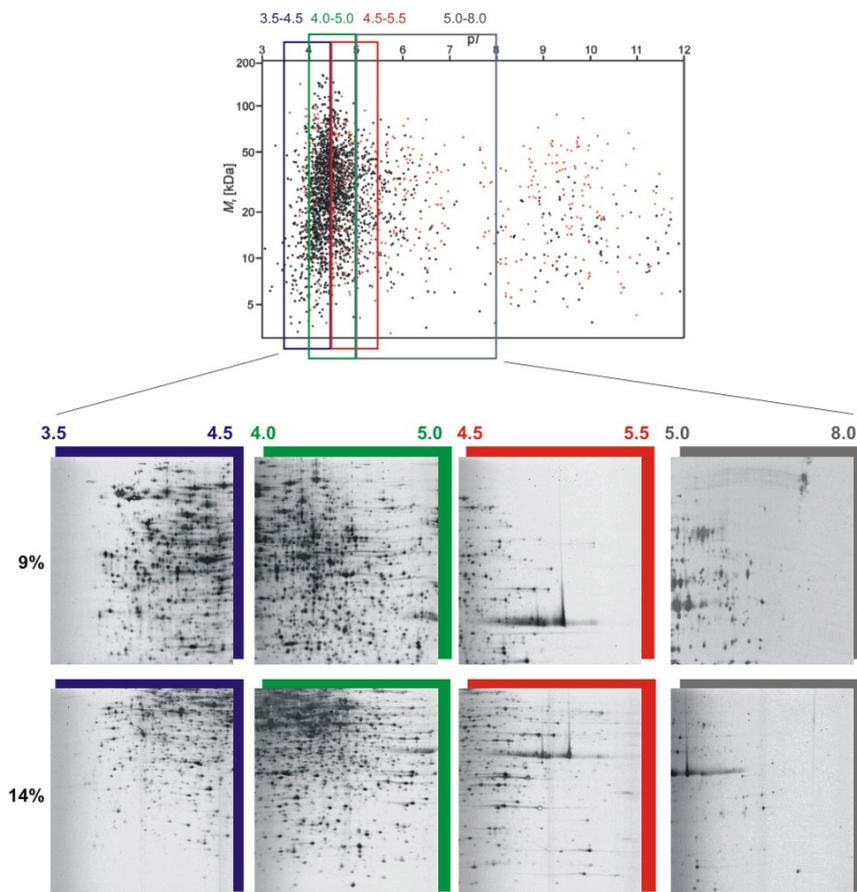
Die Auftrennung der cytosolischen Proteine erfolgte nun mittels 2-DE. Dabei wurden bei der ersten Dimension, der isoelektrischen Fokussierung (IEF), zunächst IPG-Streifen mit einem pH-Bereich von pH 3 bis pH 6 gewählt. Diese sollten den Großteil (>92%) aller cytosolischen Proteine enthalten. Um einen möglichst großen Massenbereich abzudecken, wurden Gradientengele von 9% bis 16% PAA für die zweite Dimension verwendet. Eine Voraussetzung für das Erstellen von Referenzgelen ist das Anfertigen reproduzierbarer 2-D Gele. Dabei stellt die zweidimensionale Gelelektrophorese halophiler Proteine eine besondere Herausforderung dar [Karadzic und Maupin-Furrow, 2005]. Abbildung 4-3 zeigt 2-D Gele unterschiedlicher Qualität:



**Abb. 4-3:** (A-C) Cytosolische 2-D Gele von *H. salinarum* im pI-Bereich zwischen 3.0 und 6.0.

Um reproduzierbare 2-D Gele hoher Qualität zu erhalten, mussten folgende wichtige Parameter optimiert werden: Hohe Salzkonzentrationen in der Proteinprobe waren zu vermeiden, da diese die IEF stark beeinträchtigen und durch so genanntes *streaking* im Gelbild die Auftrennung in einzelne Proteinspots nicht ermöglichen (siehe Abb. 4-3 A). Die Salze, die jedoch in großen Mengen (2M NaCl) aufgrund des gewählten Zellaufschlusses unter Hochsalzbedingungen in der Probe vorhanden waren, mussten durch sorgfältiges Waschen mit 50% Aceton entfernt werden. Des Weiteren stellte sich heraus, dass durch verlängerte Solubilisierungszeiten der Proteinprobe in IEF-Probenpuffer die Qualität des Gelbilds stark verbessert werden konnte. Es wurden so hochauflösende Gele weitestgehend ohne *streaking* erhalten, in denen die Proteinspots klar voneinander getrennt waren und so eine massenspektrometrische Analyse der Einzelproteine möglich wurde (siehe Abb. 4-3 C). Es bleibt jedoch anzumerken, dass trotz optimierter Bedingungen eine 100%ige Reproduzierbarkeit der Gelqualität nicht erreicht werden konnte. Identische Proben lieferten trotz paralleler Anfertigung bei IEF und SDS-PAGE unterschiedliche Gelqualitäten (siehe Abb. 4-3 B und C).

In den angefertigten 2-D Gelen ist die Anhäufung der Proteine in dem *pI*-Bereich zwischen 3.5 und 4.5 klar zu erkennen. Zur Verbesserung der Auflösung wurden nun für die erste Dimension so genannte *Zoom* Gele verwendet, welche nur eine *pI*-Einheit bei der IEF abdecken. Um den *pI*-Bereich, in dem die meisten *H. salinarum* Proteine liegen, zu analysieren, wurden mehrere *Zoom* Gele unterschiedlicher *pI*-Bereiche verwendet, die sich jeweils in einer halben *pH*-Einheit überschneiden. Zusätzlich wurden für jeden *pI*-Bereich 2 Gele mit unterschiedlichen PAA-Konzentrationen angefertigt. Dies sollte das Auflösungsvermögen der Gele als auch die Reproduzierbarkeit weiter verbessern und damit den Vergleich der Spotmuster unterschiedlicher Gele erleichtern. Abbildung 4-4 zeigt den abgedeckten *pI*- und *M<sub>r</sub>*-Bereich der zweidimensionalen Referenzgele anhand des theoretischen 2-D Gels sowie repräsentative Gele:



**Abb. 4-4:** Abgedeckte pI und Mr-Bereiche der gewählten 2-D Gele im Vergleich zum theoretischen 2-D Gel. Alle Gele sind online erhältlich unter [www.halolex.mpg.de](http://www.halolex.mpg.de).

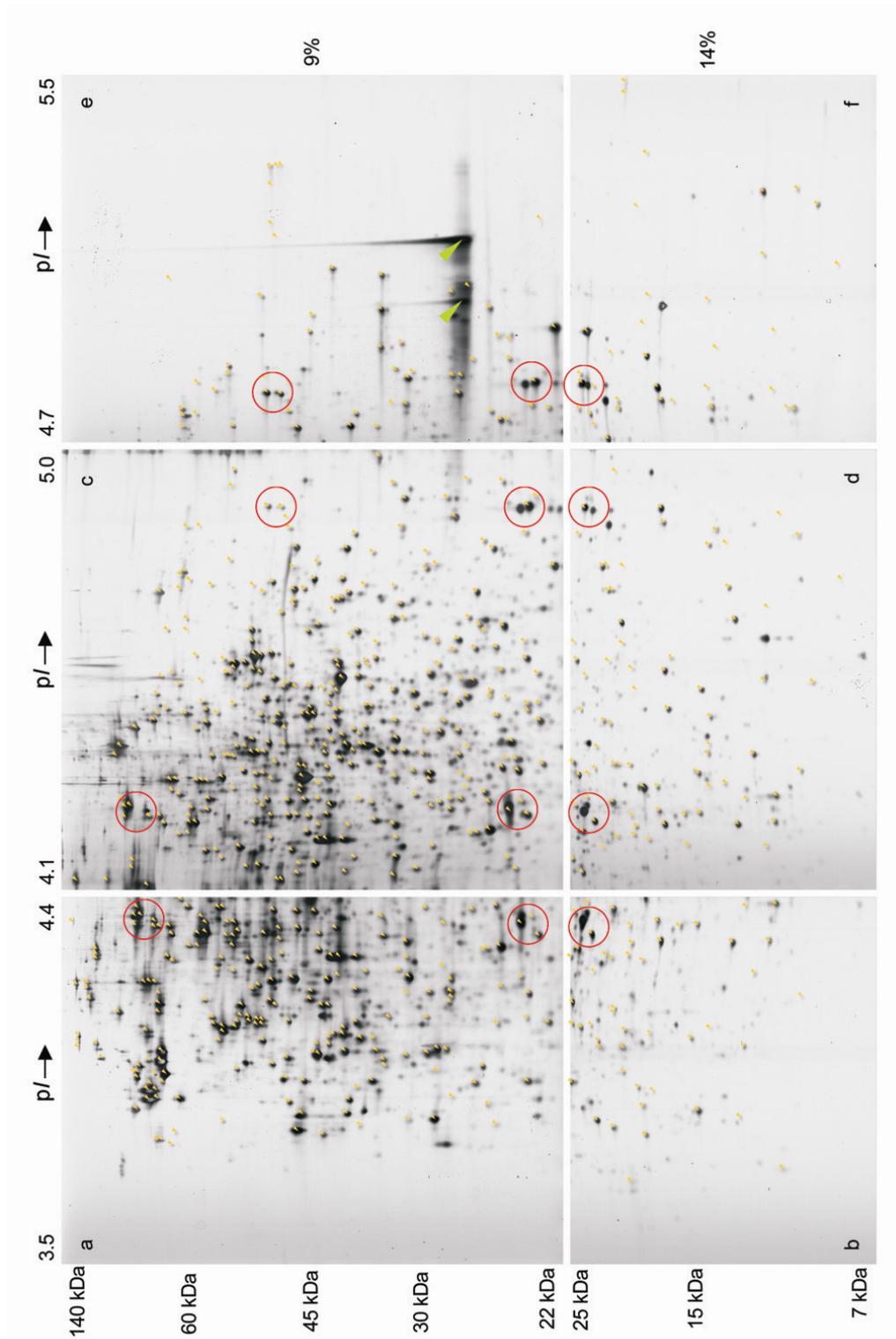
Die Visualisierung der Proteinspots geschah ausschließlich mit einer massenspektrometrisch-kompatiblen Silberfärbung (siehe 3.2.7.2). Mit dieser konnten insgesamt ca. 1,800 unterschiedliche Proteinspots auf den 2-D Gelen detektiert werden. Die reproduzierbare Anfertigung der Gele machte nun eine massenspektrometrische Analyse der 2-D Gele möglich.

#### 4.1.3 Massenspektrometrische Analyse der 2-D Referenzgele

Die massenspektrometrische Analyse der 2-D Gele erfolgte im semiautomatisierten Hochdurchsatzverfahren mittels MALDI-TOF *peptide mass fingerprinting* (siehe 3.2.11.4), welches in der Arbeitsgruppe Massenspektrometrie unter der Leitung von Dr. Frank Siedler etabliert worden war (siehe Dissertation Dr. Christian Klein). Dieses Verfahren beinhaltet (a) das automatische Ausstechen der Proteinspots nach Scannen der Gele und manueller Spotauswahl, (b) in-Gel Verdau in 96er Mikrotiter Filterplatten, (c) automatische MALDI-Target Präparation, (d)

automatische Aufnahme der PMF Massenspektren ebenso wie (e) die automatische Annotation der Spektren und (f) automatische Recherche der MS-Daten gegen die *H. salinarum* Protein Datenbank mit dem Mascot Programm. Die erhobenen Daten wurden im HALOLEX-System ([www.halolex.mpg.de](http://www.halolex.mpg.de)), welches von der Bioinformatik Gruppe unter der Leitung von Dr. Friedhelm Pfeiffer entwickelt und ständig ausgebaut wird, gebündelt.

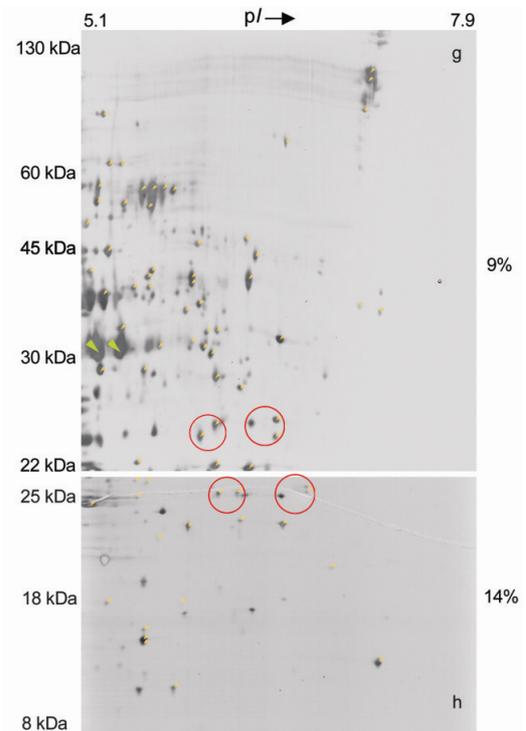
Für die erschöpfende Analyse der cytosolischen 2-D Gele wurde jeder Proteinspot mindestens dreimal aus unterschiedlichen, aber korrespondierenden Gelen ausgeschnitten. Die für die 2-DE verwendeten Proteinmengen betragen hierbei ca. 300 oder 800 µg. Für die anschließende MALDI-TOF PMF Analyse wurden die proteolytischen Peptide der ca. 7500 ausgestochenen Proteinspots zweimal unabhängig mit der *Dried-Droplet* Methode (siehe 3.2.11.1) automatisch präpariert und unabhängig voneinander vermessen (über 15,000 PMF-Massenspektren). Abbildungen 4-5 und 4-6 zeigen die Referenzgele im pI-Bereich zwischen 3.5 und 5.5 sowie zwischen 5.0 bis 8.0. Die identifizierten Proteinspots sind mit gelben Pfeilen markiert, identische Spotmuster in den überlappenden Gelregionen sind durch rote Kreise markiert. Die unterschiedlichen Gele sind als Sektoren (a-h) gekennzeichnet. Alle Gele mit den jeweiligen Proteinidentifizierungen sind im HALOLEX-System abgelegt.



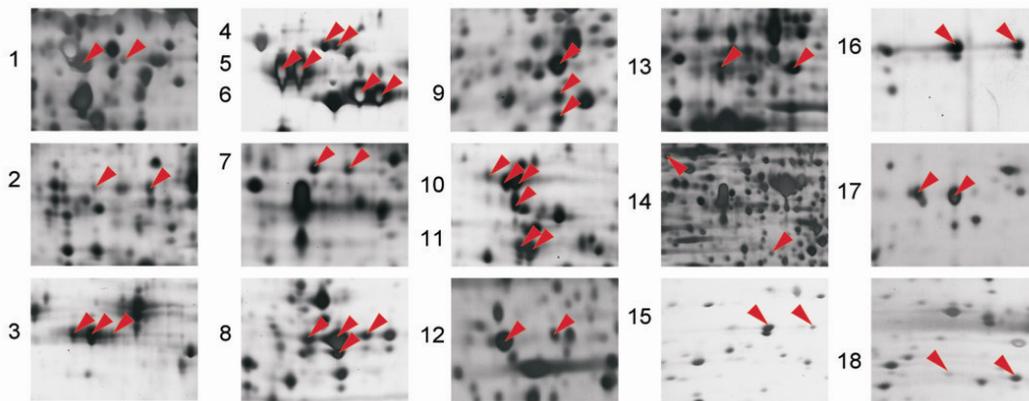
**Abb. 4-5:** 2-D Referenzgele der cytosolischen Proteine von *H. salinarum* im pI-Bereich zwischen 3.5 und 5.5. Jeder pI-Bereich wurde mit unterschiedlichen PAA-Konzentrationen (9% und 14%) abgedeckt. Die unterschiedlichen Sektoren sind markiert (a-f). Identifizierte Proteine sind mit gelben Dreiecken markiert. Die grünen Dreiecke zeigen die zugegebene DNase I. Identische Spotmuster auf den überlappenden Bereichen der Zoomgele sind mit roten Kreisen hervorgehoben.

Von den 1,959 cytosolischen Proteinen (ausgeschlossen sind Proteine die mindestens eine vorhergesagte TM-Region besitzen, die durch Insertionselemente unterbrochen sind sowie durch Genomduplikation auftretende zusätzliche Kopien desselben Proteins; Stand Ende 2004), die auf den Referenzgelen ( $pI$  3.5 - 8) theoretisch zu finden sein sollten, konnten 661 Proteine identifiziert werden, 227 von ihnen nur auf Gelen mit hoher Proteinbeladung (ca. 800  $\mu\text{g}$ ). Zusätzlich wurden 141 Proteine auf 2-D Gelen identifiziert, die nicht den Spezifikationen der Referenzgele entsprachen ( $pI$ -Bereich, Wachstumsphase der Zellen, Probenvorbereitung). Nur eines der identifizierten Proteine (OE2015R) enthielt eine vorhergesagte Transmembranregion. Sechs der identifizierten Proteine wurden von mehr als einem Gen codiert, was sich auf Genomduplikationen in den Plasmiden von *H. salinarum* zurückführen lässt. Überraschenderweise konnte ein Protein identifiziert werden, das von einem Insertionselement (ISH 2)

unterbrochen war, welches 5 Codons hinter dem Insertionspunkt ein Stoppcodon erzeugt. Dennoch wird dieses Gen weiterhin exprimiert und translatiert. Es wurden jedoch nur Peptide aus dem N-terminalen Bereich vor der Insertion aber keine aus dem C-terminalen Bereich nach der Insertion gefunden. Unter den identifizierten Proteinen wurden 94 in mehr als einem Proteinspot gefunden. Das Auftreten von Proteinen in unterschiedlichen Spots kann zum einen durch chemische Modifikation des Proteins wie Deamidierung oder Carbamoylierung und damit einhergehender Veränderung des  $pI$ -Wertes verbunden sein [McCarthy *et al.*, 2003]. Zum anderen können ein partieller proteolytischer Abbau eines Proteins oder biologisch relevante posttranslationale Modifikationen wie Phosphorylierungen den  $M_r$  bzw.  $pI$ -Wert des Proteins verändern. Von den in mehreren Spots auftretenden Proteinen wurden 45 mit unterschiedlichen  $pI$ -Werten (horizontale Verschiebung), 17 mit unterschiedlichen  $M_r$ -Werten (vertikale Verschiebung) und 35 mit Verschiebungen in beide Richtungen gefunden. Abbildung 4-7 zeigt eine repräsentative Auswahl an Proteinen die in unterschiedlichen Spots auf den 2-D Gelen gefunden wurden:



**Abb. 4-6:** 2-D Referenzgele der cytosolischen Proteine von *H. salinarum* im  $pI$ -Bereich zwischen 5.1 und 7.9

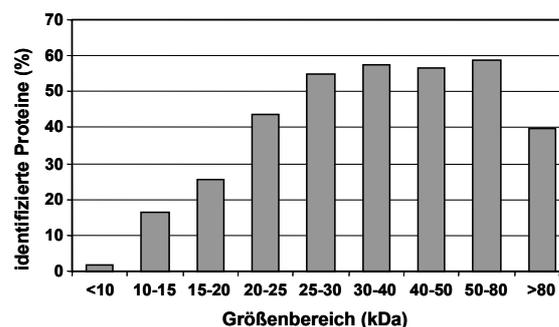


**Abb. 4.7:** Repräsentative Auswahl von Proteinen, die in mehreren Gelspots identifiziert wurden. 1: Glutamat Dehydrogenase (OE1270F), 2: Glutamat Amminoligase (OE3922R), 3: Acetat CoA Ligase (OE2450R), 4: Taxis Sensor Histidinkinase CheA (OE2415R), 5: Thermosom,  $\alpha$ -Kette (OE4122R), 6: Thermosom,  $\beta$ -Kette (OE3925R), 7: mutmaßliche Phosphat Acetyltransferase (OE2373F), 8: Pyridoxalphosphat abhängige Aminotransferase (OE3168R), 9: Aspartat Carbonyltransferase, katalytische Kette (OE5201F), 10: Isocitrat Dehydrogenase (OE3634F), 11:  $H^+$ -transportierende zwei-Sektoren ATPase, Kette B (OE3984R), 12: DNA Bindepotein dpsA (OE4427R), 13: Phosphopyruvat Hydratase (OE2640F), 14: Protein-Glutamat Methyltransferase cheB (OE2416R), 15: ABC-Transportsystem ATP Bindepotein (OE4301R), 16: Pyruvat-Ferredoxin Oxidoreduktase,  $\beta$ -Kette (OE2622R), 17: Pyridoxin Biosynthese Proteine (OE2622R), 18: Transkriptionsregulator arcR (OE5209R).

Auf den Referenzgelen konnten ca. 1,800 unterschiedliche Proteinspots aufgelöst werden. Nach Abziehen derer, die das gleiche Protein enthalten, kann man von 1,600 unterschiedlichen Proteinen auf den Referenzgelen ausgehen. Dies bedeutet, dass ca. 80% aller Gene in einem Ausmaß exprimiert werden, um die entsprechenden Proteine mittels Silberfärbung im Gel zu visualisieren. Dieser überraschend hohe Anteil an exprimierten Genen unter Standardbedingungen deckt sich mit Transkriptom-Analysen (siehe Dissertation Jens Twellmeyer).

Bei näherer Betrachtung der identifizierten Proteine ergaben sich zwei Faktoren, welche die Identifizierung mittels PMF stark beeinflussen: Der erste Faktor betraf die Größe der Proteine. Ungefähr 50% aller Proteine größer als 25 kDa, aber weniger als 20% kleiner als 20 kDa konnten identifiziert werden. Aus dem theoretischen 2-D Gel (Abb. 4-1) lässt

sich ableiten, dass 1/3 aller Proteine von *H. salinarum* ein Molekulargewicht kleiner als 20 kDa besitzen. Die entsprechenden Bereiche auf den 14% PAA Gelen (siehe Abb. 4-5 und 4-6, Sektoren b, d, f, g) zeigten jedoch, dass das „kleine Proteom“ nur zu einem



**Abb. 4-8:** Anteil (%) der identifizierten cytosolischen Proteine, die auf den Referenzgelen zu finden sein sollten, in Abhängigkeit von ihrer Größe.

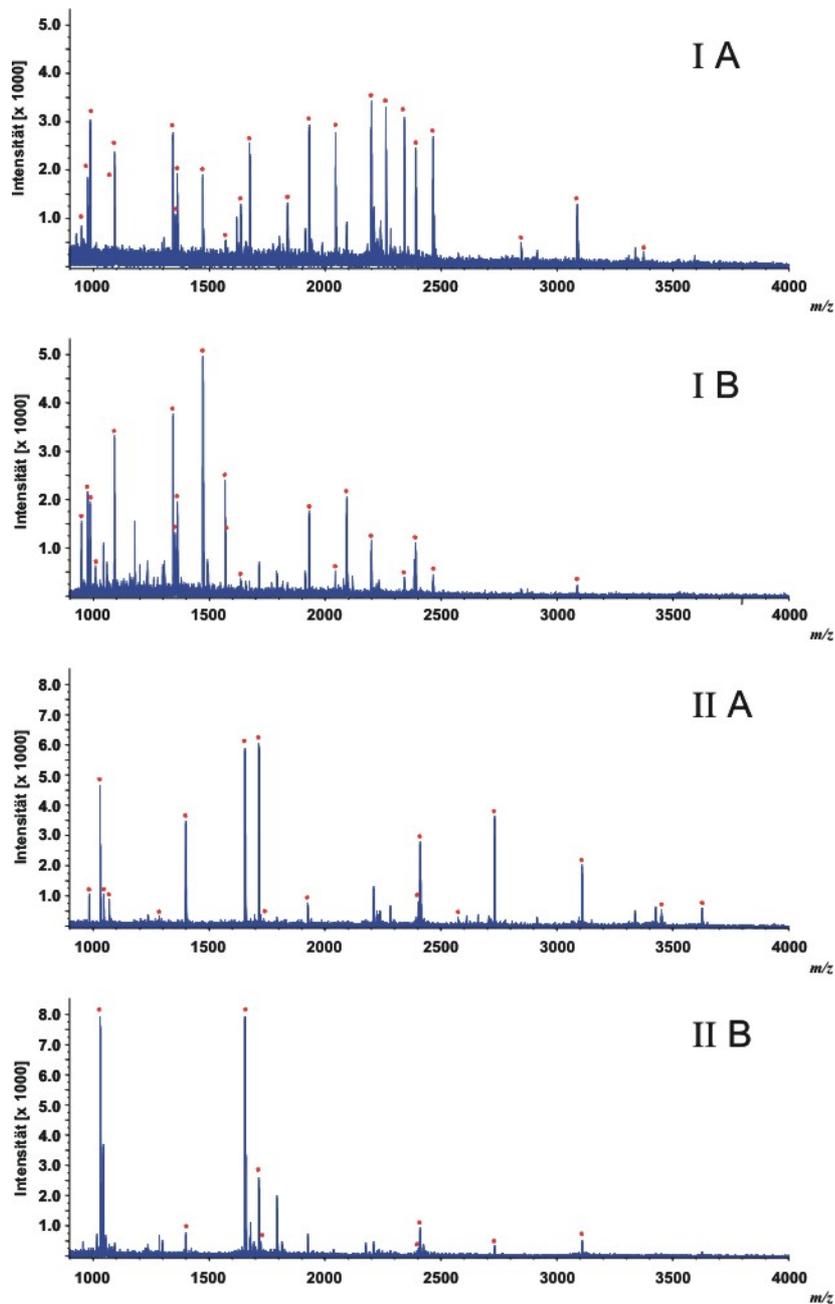
geringen Ausmaß auf den Gelen vertreten war. Zusätzlich dazu sind kleine Proteine mittels PMF schwieriger zu identifizieren, da sie weniger proteolytische Fragmente ergeben, die jedoch als Identifizierungskriterium bei dem Datenbankvergleich dienen. Abbildung 4-8 zeigt die Identifizierungsrate der Proteine in Abhängigkeit von ihren Molekulargewichten.

Weiter scheint die Lokalisation eines Gens im Genom die Identifizierungswahrscheinlichkeit des zugehörigen Proteins zu beeinflussen. Es konnten 47% aller chromosomalen Proteine (63% aller chromosomalen Proteine größer als 25 kDa), aber nur 14% aller der auf Plasmiden codierten Proteine identifiziert werden. Dies ergab eine Abdeckung von über 40 % aller cytosolischen Proteine.

#### 4.1.4 GyroLab™ als alternative Probenvorbereitung

Neben dem entwickelten Verfahren wurde die neue Technologie GyroLab™ zur Präparation von Peptidproben für die MALDI-TOF PMF Analytik getestet und mit der herkömmlichen *Dried-Droplet* Methode verglichen. Bei der GyroLab™ Methode wird die Probe durch Mikrostrukturen auf der Oberfläche der rotierenden runden Präparationsscheibe geleitet. Dabei werden die Proben über mit *reversed phase* Material versehene Kanäle entsalzt und direkt mit Matrix auf die MALDI-Analyse Positionen eluiert [Gustafsson *et al.*, 2004].

Aufgrund der Miniaturisierung und integrierten Entsalzung mittels *reversed phase* versprach das System eine höhere Empfindlichkeit gegenüber der herkömmlichen Methode. Es wurden zunächst Proben, die bereits präpariert und vermessen worden waren, mittels GyroLab™ für die MALDI-Messung vorbereitet und analysiert. Dabei stellte sich heraus, dass die gelagerten Proben mit der GyroLab™ Methode keine besseren Identifizierungsergebnisse lieferten. Auch ein Vergleich der beiden Methoden anhand von parallel angefertigten Proben aus 2-D Gelen ergab ebenfalls keine erhebliche Verbesserung der Ergebnisse mittels GyroLab™. In 75% der analysierten Proben wurden bessere Identifizierungsergebnisse mit der herkömmlichen *Dried-Droplet* Methode erreicht. Dies spricht für die Sensitivität dieser robusten Probenaufarbeitungs-Methode. Abbildung 4-9 zeigt repräsentative Massenspektren äquivalenter Proben nach Präparation mit beiden Methoden.

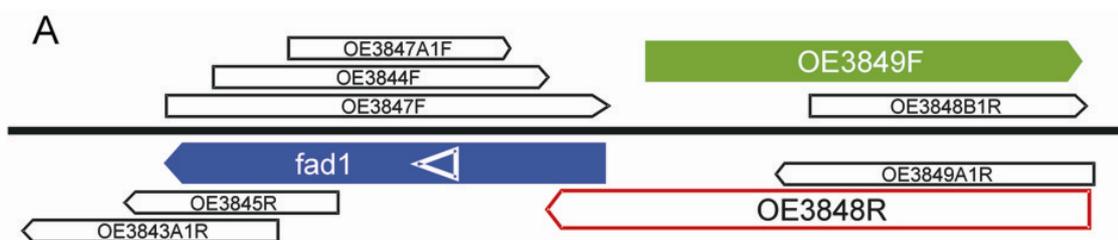


**Abb. 4-9:** MALDI-TOF PMF Massenspektren von äquivalenten Peptidproben. I: Spektren eines conserved proteins. (A) Herkömmliche *Dried-Droplet* Methode; MASCOT-Score 294, 24 *matching peptides* (23 gezeigt, rote Punkte). (B) Gyrolab™ Methode; MASCOT-Score 152, 21 *matching peptides* (20 gezeigt). II: Spektren der Pyruvat-Ferredoxin Oxidoreduktase, alpha Kette. (A) *dried droplet* Methode; MASCOT-Score 201, 18 *matching peptides* (17 gezeigt). (B) Gyrolab™ Methode; Mascot-Score 119, 9 *matching peptides*.

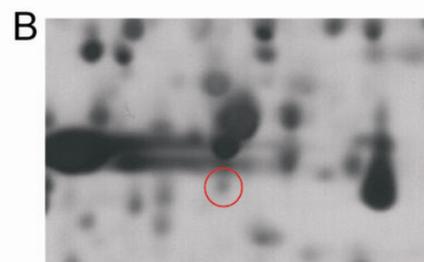
#### 4.1.5 Genomannotation anhand von MS-Daten

Probleme, die bei Genomannotationen anfallen können, wie die Vorhersage von zu vielen oder zu wenigen Genen oder die fehlerhafte Zuweisung des Startcodons, traten im Fall von *Halobacterium salinarum* verstärkt auf. Dies machte die Anwendung der Proteomik zu einem unersetzlichen Hilfsmittel der Genomik. Zum einen können durch die bloße Identifizierung eines Proteins überlappende oder auf dem gegenüberliegenden Strang befindliche ORFs als zusätzlich falsch vorhergesagte Gene entlarvt und gelöscht werden. Zum anderen kann durch die massenspektrometrische Identifizierung bestimmter Proteinfragmente eine fehlerhafte Vorhersage des Translationsstarts verbessert werden. Anhand der folgenden Beispiele wird gezeigt, wie die Proteomik zur Verbesserung und Korrektur der Genomannotation beitragen kann.

Das *fad1* Gen, welches für eine mutmaßliche Enoyl-3-Hydroxybuturyl-CoA Dehydratase codiert, liegt in der Genomregion um die Position 1,490,000 kbp. Das ORPHEUS Programm [Frishman *et al.*, 1998] sagte Met-1, das Glimmer Programm [Salzberg *et al.*, 1998; Delcher *et al.*, 1999] Met-81 als potentiellen Translationsstart voraus. Die Sequenz der 80 N-terminalen Aminosäuren wies jedoch eindeutige Homologien zu Proteinen anderer Organismen auf, woraufhin die Glimmer-Vorhersage verbessert wurde (siehe Abb. 4-10 A). Auf dem Gegenstrang in 5'-Richtung vom *fad1* Gen liegt ein weiteres Gen (OE3849F), das durch Massenspektrometrie eindeutig identifiziert werden konnte (siehe Abb. 4-10 B). Von Glimmer wurde wiederum das auf dem Gegenstrang befindliche Gen OE3848R vorhergesagt. Dieses überlappt mit der längeren und richtigen Version des *fad1* Gens. Zusätzlich befinden sich noch 7 weitere falsch vorhergesagte ORFs in dieser Region, welche lediglich 2 reelle Gene enthält.

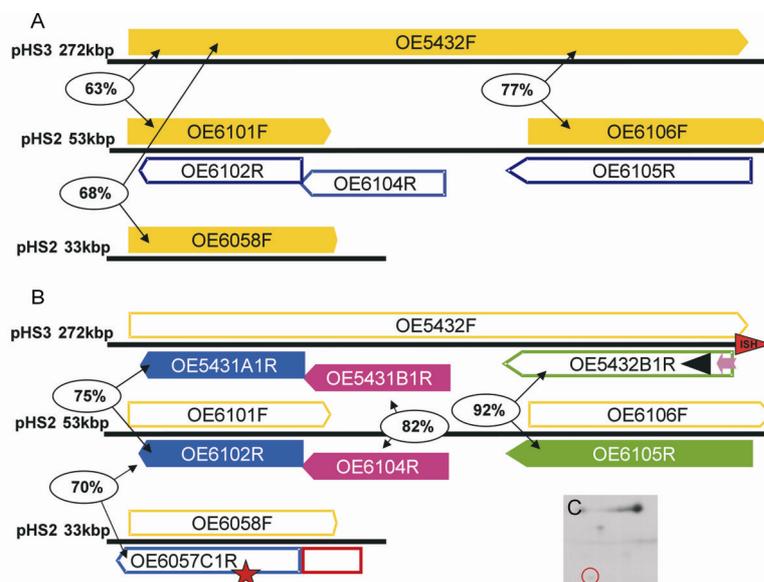


**Abb. 4-10:** (A) Genomannotation in der Region des *fad1* Gens. Das *fad1*-Gen ist in blau dargestellt mit dem von Glimmer falsch vorhergesagtem Startcodon (weißes Dreieck). Das eindeutig identifizierte Protein OE3849F aus einem schwachen silbergefärbten Spot (B) ist in grün dargestellt. Der von Glimmer falsch vorhergesagte ORF OE3848R ist in rot gezeigt. Die 7 zusätzlich vorhergesagten ORFs sind ebenfalls dargestellt.



Die Identifizierung eines Proteins aus einem schwach silbergefärbten Spot erlaubte die Lösung einer Reihe von Annotationsfehlern in drei unterschiedlichen Genombereichen auf zwei Megaplasmiden. Die beiden verwendeten ORF Vorhersageprogramme, Glimmer und ORPHEUS, sagten OE5432F als einzigen reellen ORF in der Region um die Genomposition 272 kbp auf dem Plasmid pHS3 voraus (siehe Abb. 4-11 A). Nach Definition lag dieser auf dem Leitstrang (F = *forward strand*).

In der Region um die Position 52 kbp auf dem Plasmid pHS2 sagten ORPHEUS und Glimmer drei Gene auf dem gegenüberliegenden Strang voraus (OE6102R, OE6104R und OE6105R; R = *reversed strand*), wobei ORPHEUS noch zusätzlich zwei ORFs auf dem Leitstrang berechnete (OE6101F und OE6106F). OE6106F wies 77% Proteinsequenzhomologie zum C-terminalen Bereich von OE5432F auf, wohingegen OE6101F 63% Sequenzhomologie zu dessen N-terminalen Ende besaß. Die drei auf dem Gegenstrang befindlichen Proteine besaßen keinerlei Homologien zu Proteinen aus *H. salinarum* oder einer anderen Species. In der Region um die Position 33 kbp auf demselben Plasmid (pHS2) wurde OE6058F von beiden Programmen vorhergesagt. Dieser ORF besaß 68% Sequenzhomologie zur N-terminalen Region von OE5432F auf



pHS3. Sowohl das hohe Ausmaß der Sequenzidentität zwischen OE6106F und OE5432F (77%) als auch das Fehlen eines Homologen für OE6105R legte die Vermutung nahe, dass der Leitstrang auf pHS2 an der Position um 53 kbp die reellen Gene enthält.

**Abb. 4-11:** Genomannotation vor (A) und nach (B) der Identifizierung des Proteins OE6105R. (C) Silbergefärbter Proteinspot von OE6105R in drei unterschiedlichen Genomregionen (Plasmid und Position links angegeben). ORFs (Pfeile), die ausgefüllt dargestellt sind, werden als kodierend angesehen. In (A) zeigen die gelb ausgefüllten ORFs 63-77 % Proteinsequenzhomologie untereinander. (B) zeigt die Reannotation der Genomregionen nach Identifizierung des Proteins OE6105R, Six-Frame Translation und weiterer manueller Bearbeitung. Grün, magenta und blau ausgefüllte ORFs zeigen 75-92% Proteinsequenzhomologie untereinander. Der Anfang des Insertionselements ISH3 ist als rotes Dreieck auf pHS3 dargestellt. Das schwarze Dreieck in OE5432B1R auf pHS3 markiert das nach Six-Frame Translation zugewiesene GTG Startcodon für Valin-21. Der rosa Pfeil in OE5432B1R markiert die Translation, die direkt dem Insertionselement ohne Startcodon folgen würde. Die N-terminale Region auf der anderen Seite des Insertionselements, die für 9 Aminosäuren codieren würde, ist nicht dargestellt. Der rote Stern in OE6057C1R auf pHS2 (33 kbp) ein *in-frame* Stoppcoden. Der rote Teil zeigt den N-terminalen Bereich, der *in-frame* an diesen ORFs fusioniert ist und Sequenzhomologie zu einem nicht verwandten Gen aufweist.

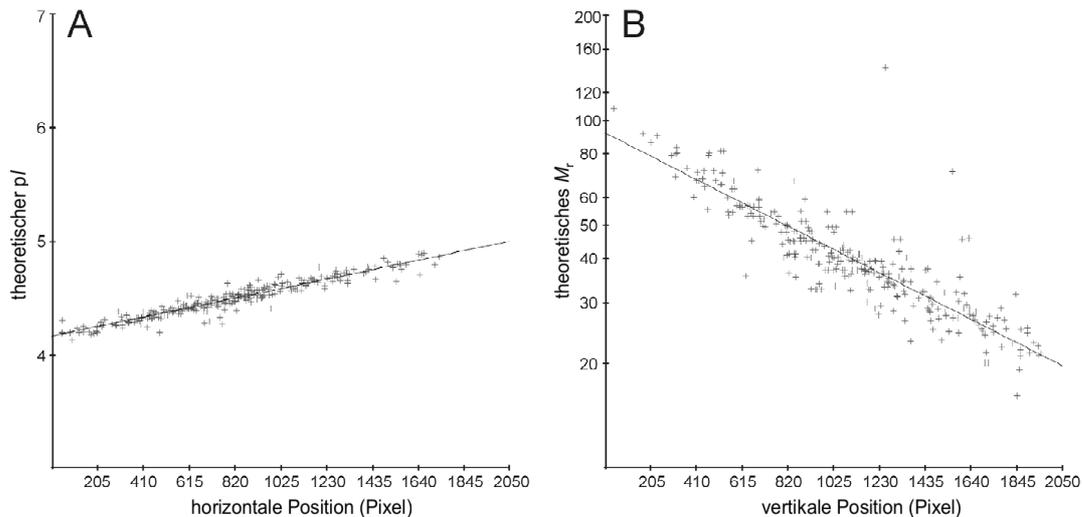
Diese Annahme konnte jedoch durch die eindeutige massenspektrometrische Identifizierung des Proteins OE6105R klar widerlegt werden, welches vom zugehörigen Gen auf dem Gegenstrang codiert wird. Daraus ergab sich, dass der ORF OE5432F nicht für ein reales Protein codiert, was daraufhin ebenfalls für die ORFs mit Sequenzhomologien gelten sollte (OE6101F und OE6058F). Die Region um die Position 272 kbp auf pHS3 enthielt entsprechend keine potentiellen Gene mehr.

Durch Sixframe-Translation mit einer minimalen Längenbegrenzung von 100 Codons wurden die zusätzlichen ORFs OE5432B1R mit 92% Sequenzhomologie zu Protein OE6105R und ORF OE5431A1R mit 75% Sequenzhomologie zu ORF6102R gefunden. Der ORF OE5431B1R, der nur aus 96 Aminosäuren besteht und 82% Sequenzhomologie zu ORF6104R aufweist, wurde nachträglich manuell erzeugt (siehe Abb. 4-11 B). Zusätzlich wurde das neu vorhergesagte Protein OE5432B1R durch das Insertionselement ISH3 nach dem Codon für Lys-9 unterbrochen (rotes Dreieck). Das durch Sixframe-Translation erzeugte Protein startet mit einem GTG-Kodon für Val-21 (schwarzes Dreieck). Die N-terminalen Aminosäuren auf der anderen Seite des Insertionselements sind nicht dargestellt. Lys-9 befindet sich auf beiden Seiten des Insertionselements aufgrund einer 5-Basenpaar Duplikation am Insertionspunkt.

Nach Inspektion der Genomregion um die Position 33 kbp auf pHS2 konnte ein Pseudogen (OE6057C1R) mit einem internen Stoppcodon (roter Stern) detektiert werden. Der C-terminale Teil nach dem Stoppcodon besaß 85% Sequenzhomologie zu dem ORF OE5431A1R. Die dem Stoppcodon vorausgehende Sequenz stellt den Bruch in der Plasmidsequenz dar. Der C-terminale Teil ist zu 77% homolog zu OE5431A1R, wohingegen der N-terminale Teil (rot gekennzeichnet) Sequenzhomologie zu einem anderen Protein aufweist. Auf Grundlage dieser Daten, die von der Identifizierung eines einzelnen Proteins ausgehen, wird angenommen, dass die ausgefüllten Pfeile in 4-11 B die realen Gene in den entsprechenden Genomregionen darstellen.

#### **4.1.6 Genomannotation anhand von MS-Daten und Korrelationsanalysen**

Werden die theoretischen  $pI$ - bzw.  $M_r$ -Werte der identifizierten Proteine gegen ihre jeweiligen Gelpositionen aufgetragen (siehe „Material und Methoden“ [3.2.12.5](#)), wird eine lineare Korrelation erhalten (siehe Abb. 4-12).



**Abb. 4-12:** Korrelationsanalyse der identifizierten Proteine aus einem 2-D Gel. Die identifizierten Proteine sind als Kreuze dargestellt, die Linie entspricht der berechneten Ausgleichsgeraden. (A) Korrelationsanalyse zwischen theoretischem  $pI$  und horizontaler Gelposition. (B) Korrelation zwischen theoretischem  $M_r$  und vertikaler Gelposition.

Anhand der Ausgleichsgeraden kann der experimentelle  $pI$ - bzw.  $M_r$ -Wert des identifizierten Proteins an der entsprechenden Gelposition definiert und so die theoretischen mit den experimentell ermittelten  $pI$ - bzw.  $M_r$ -Werten verglichen werden. Die große Anzahl an identifizierten Proteinen erlaubte eine detaillierte statistische Analyse.

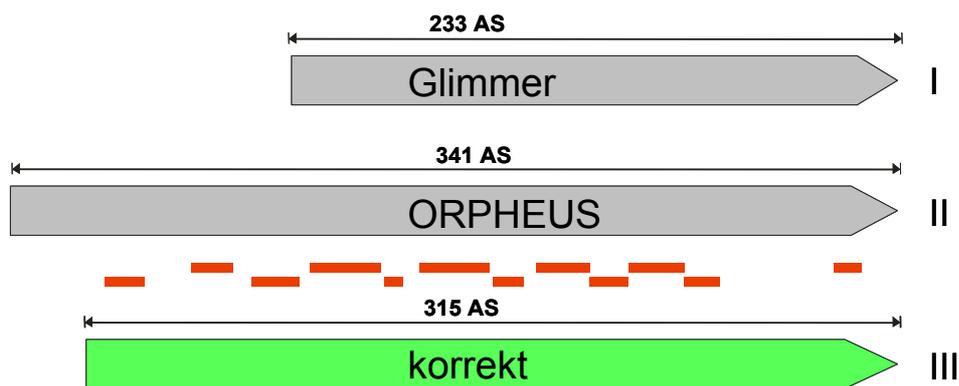
Für die  $pI$ -Werte wurde eine sehr gute Korrelation erhalten. Abbildung 4-12 A zeigt die Analyse der Daten eines 2-D Gels mit 242 identifizierten Proteinen (MOWSE Score  $\geq 92$ ). Die Standardabweichung betrug lediglich 0.034 pH-Einheiten, wobei mehr als 50% aller Proteine um weniger als 0.025 pH-Einheiten abwichen. Nur sehr wenige Proteine zeigten mehr als 3 Standardabweichungen von der Ausgleichsgeraden. Die anderen Gele, für die Korrelationsanalysen angefertigt wurden (insgesamt 25), zeigten ähnliche Ergebnisse, mit Standardabweichungen zwischen 0.03 und 0.06 pH-Einheiten. Die Korrelation zwischen den theoretischen und experimentellen  $M_r$ -Werten in Abbildung 4-12 B ergab eine Ausgleichsgerade mit Standardabweichungen von 0.177. Bei einer logarithmischen Skala würden 3 Standardabweichungen 0.351 entsprechen, was zur Folge hätte, dass ein Protein mit einer Masse von 100 kDa mit 3 Standardabweichungen zwischen 225 und 44 kDa zu finden wäre. Dies verdeutlicht die Ungenauigkeit der Abschätzung von Molekularmassen mittels SDS-PAGE.

Aufgrund der hohen Genauigkeit der  $pI$ -Korrelation konnten hohe Abweichungen von der Ausgleichsgeraden als Indikator für eine fehlerhafte Zuordnung des Proteins zu dem experimentell ermittelten *Peptid Fingerabdruck* im Massenspektrum angesehen werden (falsch-positive Identifizierungen). Weitere Gründe für Abweichungen können

entweder eine aus der Genomsequenz fehlerhaft abgeleitete Proteinsequenz durch z.B. falsche Startcodon Vorhersage oder das Auftreten posttranslationaler Modifikationen, welche den  $pI$ -Wert des Proteins verändern, sein.

Die Korrelationsanalyse wurde nun als Werkzeug für die Analyse der beiden erst genannten Phänomene herangezogen. Es konnten 5 falsch-positive Identifizierungen mit hohen Standardabweichungen von der Ausgleichsgeraden ermittelt und so eliminiert werden. Für zwei Identifizierungen lagen die Abweichungen bei 183 und 26 Standardabweichungen mit  $pI$ -Werten der Proteine bei 11.2 bzw. 6 auf 2-D Gelen mit einem  $pH$ -Bereich zwischen 3 bis 6 bzw. 3.5-5.5. Die restlichen drei Proteine besaßen alle  $pI$ -Werte über 9 und wurden auf sauren Gelen identifiziert. Des Weiteren konnten auf der Grundlage der  $pI$ -Korrelation mehrere falsch annotierte Startcodons erkannt werden. Nach manueller Auswahl des nächst möglichen Translationsstarts waren keine gravierenden Abweichungen von den Ausgleichsgeraden mehr zu erkennen. Kriterien für die Auswahl des neuen Startcodons waren (a) der  $pI$ -Wert der Region zwischen den alternativen Startcodons, (b) die Übereinstimmung der Sequenz zu homologen Proteinen und (c) der Abstand zum vorangehenden ORF.

Die problematische Startcodon-Vorhersage kann am Beispiel des Proteins OE3065R, einer mutmaßlichen Dehydrogenase, veranschaulicht werden. Das Programm ORPHEUS sagte ein Protein voraus, welches N-terminal 108 Aminosäuren länger war als das vom Programm Glimmer berechnete (siehe Abb. 4-13 I und II). Die Identifizierung von 3 Peptiden in der N-terminalen Verlängerung zeigte klar, dass das von Glimmer vorhergesagte Startcodon nicht zutreffen konnte.



**Abb. 4-13:** Unterschiedliche Genomannotation des Proteins OE3065R von zwei ORF Vorhersage Programmen. (I und II), sowie korrekte manuelle Annotation auf der Grundlage von MS-Daten und  $pI$ -Korrelationsanalyse (III). Die roten Balken repräsentieren die mittels MS gefundenen Peptide.

Die von ORPHEUS berechnete Sequenz zeigte jedoch immer noch große Abweichungen (zwischen 2.4 und 5.1 Standardabweichungen) in der *pI*-Korrelation auf mehreren 2-D Gelen. Nach Zuordnung eines alternativen Startcodons, das die von ORPHEUS vorgeschlagene Proteinsequenz um 36 Aminosäuren verkürzte, wurde eine gute *pI*-Korrelation erhalten (siehe Abb. 4-13 III).

Dieses Beispiel zeigt, dass selbst die aus massenspektrometrischen Daten hervorgehenden Verifizierungen von Proteinsequenzen nicht ausreichend sein können, um fehlerhafte Genomannotationen zu erkennen und zu korrigieren. Erst die Kombination der Proteom-Daten mit der *pI*-Korrelationsanalyse erlaubte in diesem Fall die richtige Zuordnung des ORFs, welche von zwei Vorhersageprogrammen falsch berechnet wurde.

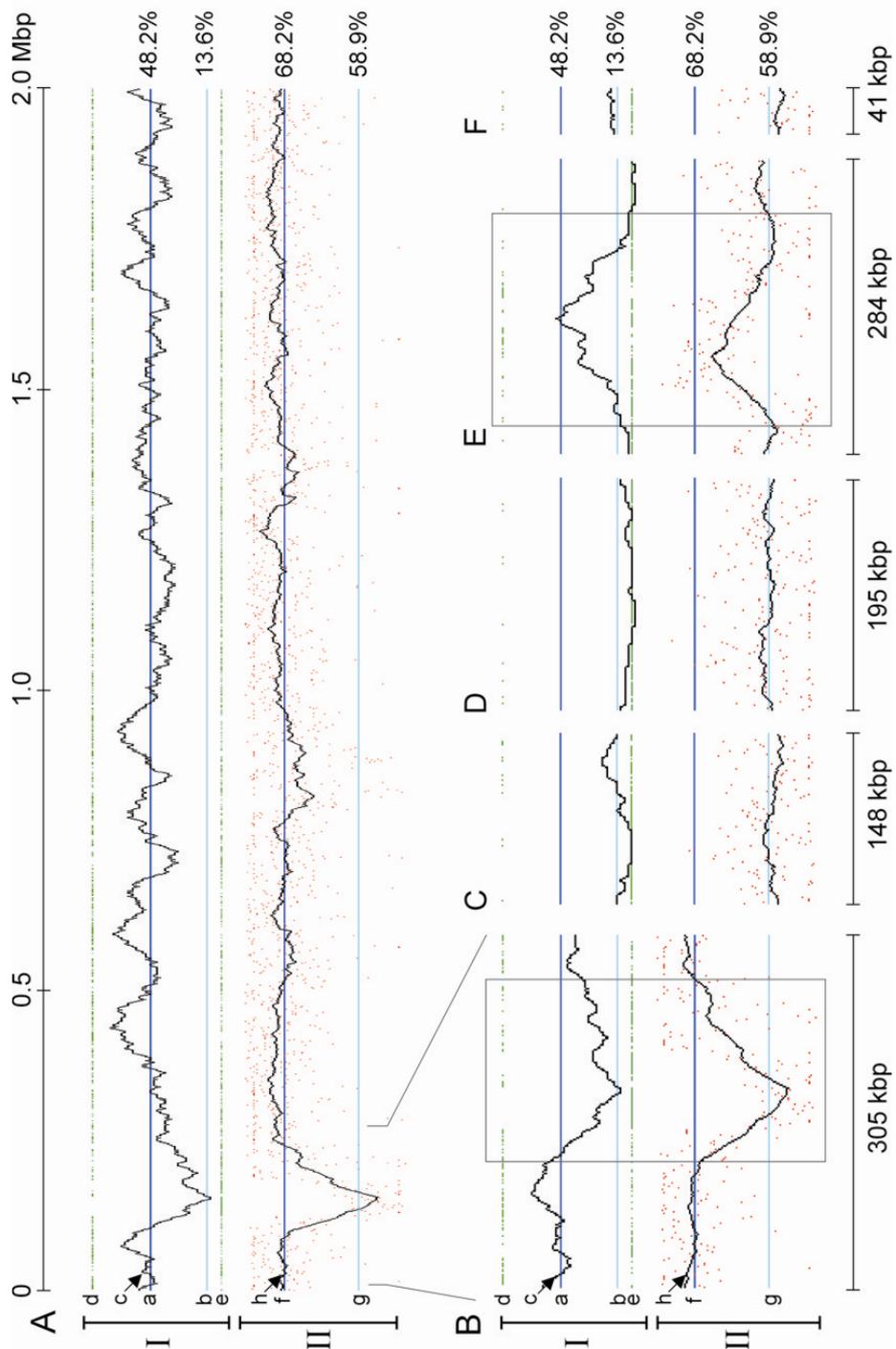
#### 4.1.7 Genominterpretation anhand von Proteom- und Genom-Daten

Das Genom von *H. salinarum*, Stamm R1 (DSM 671, [www.halolex.mpg.de](http://www.halolex.mpg.de)), besteht aus einem 2 Mbp großen Chromosom und 4 Megaplasmiden mit einer Größe zwischen 40 und 284 kbp. Das Chromosom besitzt einen sehr hohen GC-Gehalt von 68%, wohingegen die Plasmide einen moderaten GC-Gehalt von 58% aufweisen. In Bezug auf die Proteomanalyse konnte ebenfalls ein signifikanter Unterschied zwischen chromosomalen und auf Plasmiden kodierten Proteinen festgestellt werden. Es wurden 47% aller Proteine, die auf dem Chromosom, aber nur 14% der Proteine, die auf den Plasmiden codiert werden, identifiziert.

Auf der Grundlage dieser Genom- und Proteom-Daten konnte gezeigt werden, dass (a) eine 60 kbp episomale Region eines ehemaligen Plasmids in das Chromosom integriert ist, und (b) dass das Plasmid pHS3 ehemalig chromosomale DNA einer Länge von ca. 70 kbp enthält, auf der wichtige und essentielle Genen kodiert sind. Abbildung 4-14 A-F zeigt den GC-Gehalt des Chromosoms und der Plasmide sowie den Identifizierungsgrad der auf ihnen kodierten Proteine. Zur Berechnung der Daten dienten nur die cytosolischen Proteine. GC-Gehalt sowie Identifizierungsgrad an der jeweiligen Stelle des Genoms ergeben sich aus dem durchschnittlichen Wert eines Fensters von 45 ORFs um die betreffende Stelle. Die jeweiligen GC-Gehalte der einzelnen ORFs sind als rote Punkte dargestellt. GC-Gehalt sowie Identifizierungsgrad unterscheiden sich signifikant zwischen dem Chromosom und den Plasmiden. Die chromosomale Region zwischen Position 11,500 und 71,500 zeigt einen ungewöhnlich starken Abfall des GC-Gehalts sowie der Identifikationsrate auf Werte, die typisch für die Plasmide sind (siehe Kasten Abb. 4-14 B). Dies kann als Hinweis für eine

episomale Insertion eines Plasmids in das Chromosom angesehen werden [Pfeifer und Betlach, 1985; Kennedy *et al.*, 2001].

Die umgekehrte Situation wird auf dem Plasmid pHS3 in der Region zwischen 54,000 und 126,000 vorgefunden. Hier steigt sowohl der GC-Gehalt als auch die Identifizierungsrate auf Werte, die im Chromosom gefunden werden (siehe Abb. 4-14 E). Diese Region wird folglich als ein ehemaliges chromosomales DNA Fragment interpretiert. Diese Annahme wird dadurch gestützt, dass auf dieser Region eine Reihe wichtiger und sogar essentieller Proteine, wie z.B die einzige Catalase, die Arginin-tRNA Ligase oder die beiden Ketten der Aspartate Carbamoyltransferase (pyrB and pyrI), liegen. Schlussfolgernd sollte Plasmid pHS3 eher als zweites Chromosom angesehen werden.

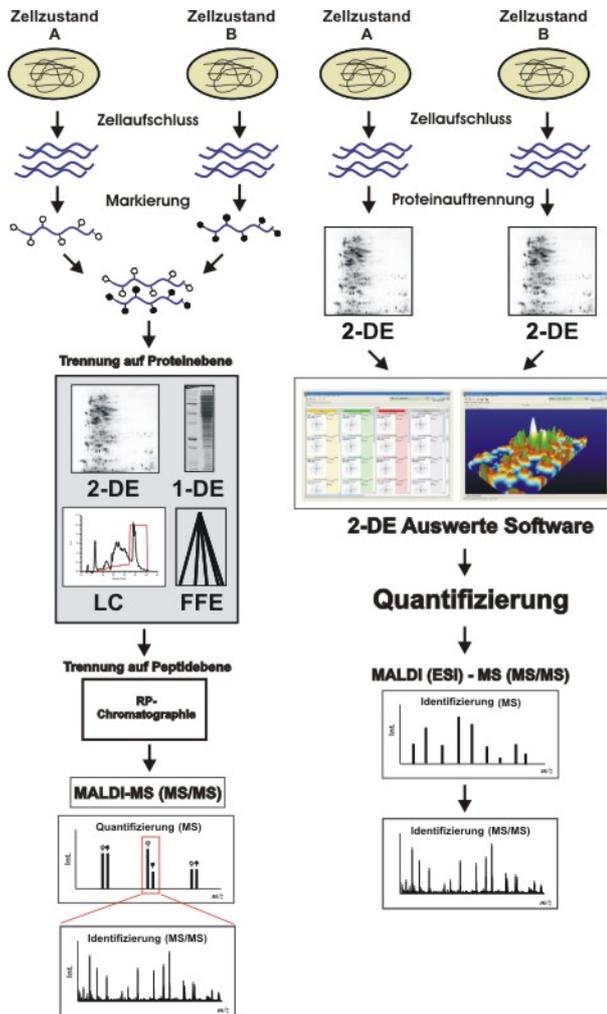


**Abb. 4.14:** Identifizierungsrate und GC-Gehalt des Chromosoms (A+B) und der Plasmide (C-F). Die Darstellung bezieht sich auf die cytosolischen Proteine.

I: (a) Durchschnittliche Identifizierungsrate aller chromosomalen Proteine (dunkelblaue Linie, 48.2%), (b) Durchschnittliche Identifizierungsrate aller Plasmid-Proteine (hellblaue Linie, 13.6%) und (c) durchschnittliche Identifizierungsrate für ein Fenster von 45 ORFs (schwarze Kurve). Identifizierte bzw. nicht-identifizierte Proteine sind als grüne Punkte oberhalb bzw. unterhalb der Linie dargestellt.

II: (f) Durchschnittlicher GC-Gehalt des Chromosoms (dunkelblaue Linie, 68.2%), (g) Durchschnittlicher GC-Gehalt der Plasmide (hellblaue Linie, 58.9%); der GC-Gehalt eines jeden ORFs ist als roter Punkt dargestellt. (h) Durchschnittlicher GC-Gehalt für ein Fenster von 45 ORFs. Die episodale Insertion eines Plasmids in das Chromosom ist vergrößert in B (Kasten) dargestellt. Chromosomale Region, die in das Plasmid pHS3 integriert ist, ist in E hervorgehoben (Kasten).

## 4.2 Quantitative Proteomik



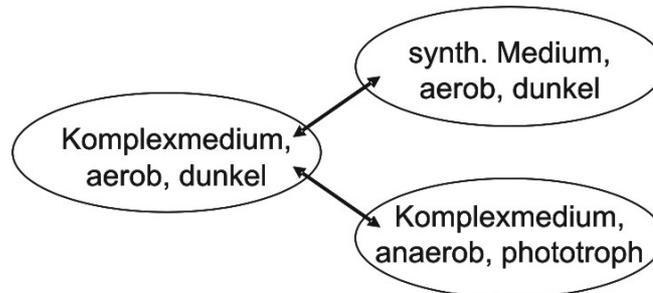
**Abb. 4-15:** Schematischer Ablauf der quantitativen Proteomik mittels 2-DE oder isotopenmarkierter Sonden.

In einem quantitativen Vergleich der Proteinexpression wurden nun die unterschiedlichen Proteome von *H. salinarum* bei unterschiedlichen Wachstumsbedingungen untersucht. Aus der Kenntnis der regulierten (unterschiedlich expremierten) Proteine können die physiologischen Veränderungen und Anpassungen des Organismus auf Proteomebene studiert werden. Die Analyse erfolgte entweder durch den visuellen oder computerunterstützten Vergleich der Spotmuster unterschiedlicher 2-D Gele oder durch den Einsatz isotopenmarkierter Sonden zur Quantifizierung regulierter Proteine.

Der Einsatz isotopenmarkierter Sonden bietet mehrere Vorteile gegenüber dem Vergleich der Proteommuster mittels 2-DE. Die Quantifizierung beruht hier auf den Intensitäten der Peptidsignale im Massenspektrum und ist somit

unabhängig von Färbemethoden. Die ICPL-Methode liefert zusätzlich den großen Vorteil, dass direkt Proteine markiert werden, was eine Kombination der unterschiedlichen Trennmethoden auf Protein- und Peptidebene erlaubt (siehe Abb. 4-15). Ein Nachteil der ICPL-Methode in Bezug auf ihren Einsatz für *H. salinarum* ist jedoch die Acidizität halophiler Proteine, die mit einer verringerten Anzahl an Lysinen einhergeht (Lysinanteil: knapp 2% bei *H. salinarum* gegenüber 6% bei mesophilen Proteinen). Dies bedeutet, dass ebenfalls die Anzahl markierbarer Peptide pro Protein und so die Wahrscheinlichkeit, Regulationsinformationen ableiten zu können, verringert ist. Bei den 308 Proteinen in *H. salinarum*, die überhaupt kein Lysin besitzen, ist die Aminofunktion des N-Terminus die einzige funktionelle Gruppe, die eine Modifikation dieser Proteine mit der ICPL Sonde erlaubt, um Regulationsinformationen zu erhalten.

Bei den angestellten Vergleichen handelte es sich um stark unterschiedliche Lebensbedingungen. Es wurde die Standard-Anzuchtbedingung, Wachstum in komplexem Medium unter aeroben Bedingungen, zum einen mit aerobem Wachstum in synthetischem Medium und zum anderen mit anaerobem, phototrophen Wachstum in Komplexmedium verglichen (siehe Abb. 4-16).



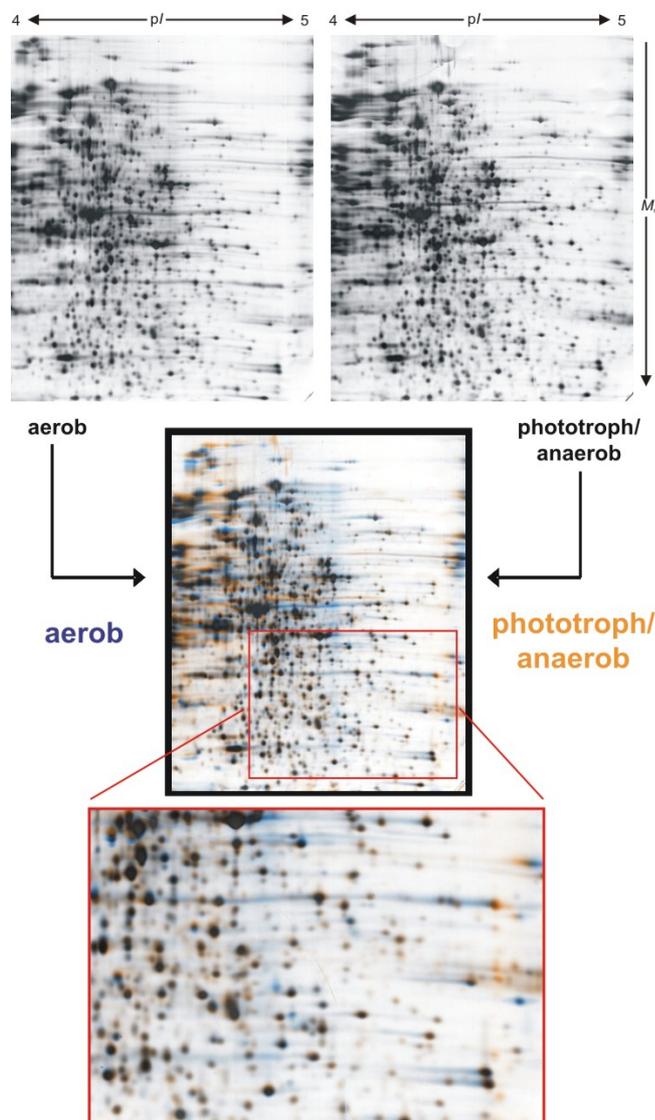
**Abb. 4.16:** Schematische Darstellung der angestellten Proteomvergleiche zwischen unterschiedlichen Wachstumsbedingungen.

Es wurde so das Wachstum unter idealen bioenergetischen Bedingungen mit genügend Nährstoffen mit zwei Mangelbedingungen verglichen, bei denen einerseits die Versorgung mit genügend Nährstoffen (synthetisches Medium) und andererseits die Variabilität des Redoxzustands (anaerob, phototroph) limitierend ist.

Für die vergleichenden Ansätze wurden Zellen wie in „Material und Methoden“ beschrieben angezogen und die cytosolische Proteinprobe unter Standardbedingungen gewonnen. Der Vergleich der cytosolischen Proteome erfolgte nun entweder mittels 2-DE oder der ICPL-Methode.

#### 4.2.1 Vergleich von aerobem mit phototrophem Wachstum mittels 2-DE

Die 2-D Gele wurden wie in „Material und Methoden“ beschrieben unter standardisierten Bedingungen angefertigt (pI-Bereich 4-5) und die aufgetrennten Proteine mittels Silberfärbung visualisiert. Die visuelle Inspektion der Gele zeigte keine nennenswerten Veränderungen im Proteommuster bei dem Vergleich von aerobem mit phototrophem Wachstum. Der Einsatz der 2-D Gel Auswertesoftware Proteomweaver™ ergab ebenfalls keine eindeutige Regulation von Proteinen. Abbildung. 4-17 zeigt 2-D Gele beider Zustände, das mittels Proteomweaver erzeugte Überlagerungsbild beider Gele sowie einen Ausschnitt dieses Vergleichs.

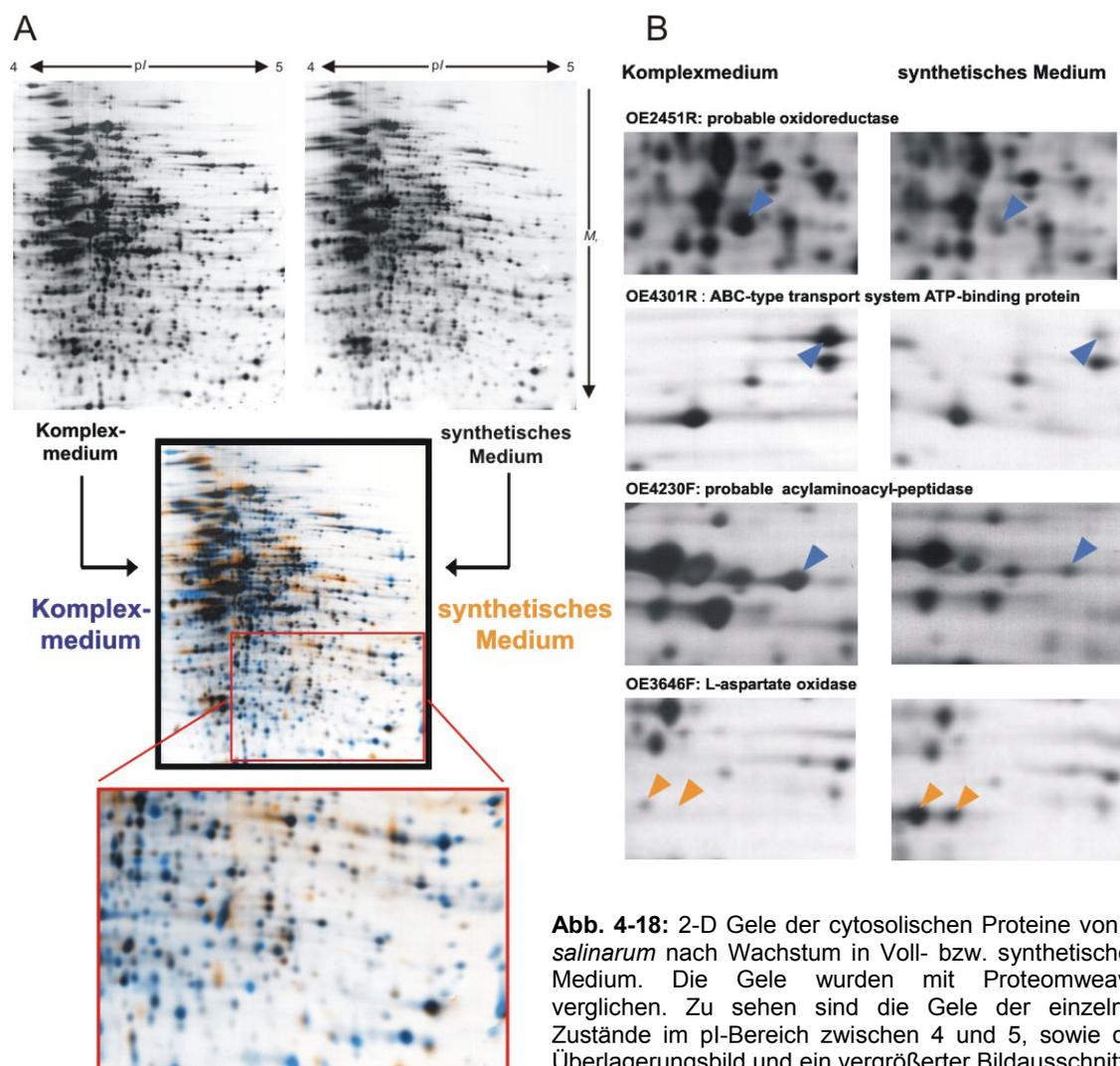


**Abb. 4-17:** 2-D Gele der cytosolischen Proteine von *H. salinarum* nach Wachstum unter aeroben bzw. anaeroben phototrophen Bedingungen. Die Gele wurden mit Proteomweaver verglichen. Zu sehen sind die Gele der einzelnen Zustände im pI-Bereich zwischen 4 und 5 sowie das Überlagerungsbild und ein vergrößerter Bildausschnitt.

Einige Proteinspots, die nach der automatischen Überlagerung der Gele mittels Proteomweaver™ reguliert zu sein schienen stellten sich jedoch bei näherer manueller Betrachtung als nicht signifikant reguliert heraus (Regulation <50%; für eine Nähere Beschreibung biologisch signifikanter Regulationen siehe 4.2.3.5).

#### 4.2.2 Vergleich des Wachstums in Komplex- bzw. synthetischem Medium mittels 2-DE

Der nächste Vergleich mittels 2-DE bezog sich auf die Proteomänderungen von Zellen, die entweder in komplexem oder in synthetischem Medium wuchsen. Die Gele wurden unter Standardbedingungen angefertigt. Der Vergleich der silbergefärbten Gele wurde jedoch durch deren geringere Qualität eingeschränkt. Zusätzlich schien das Gel der Proteine aus dem Vollmedium mit einer höheren Proteinmenge beladen zu sein, was eine vergleichende Analyse zusätzlich erschwerte und eine sichere Bestimmung regulierter Proteine nicht zuließ.



**Abb. 4-18:** 2-D Gele der cytosolischen Proteine von *H. salinarum* nach Wachstum in Voll- bzw. synthetischem Medium. Die Gele wurden mit Proteomweaver verglichen. Zu sehen sind die Gele der einzelnen Zustände im pI-Bereich zwischen 4 und 5, sowie das Überlagerungsbild und ein vergrößerter Bildausschnitt.

Abbildung 4-18 A zeigt 2-D Gele beider Zustände, das mittels Proteomweaver erzeugte Überlagerungsbild beider Gele sowie einen Ausschnitt dieses Vergleichs. Die überwiegende Blaufärbung vieler Proteinspots verdeutlicht die Überladung des Gels mit den Proteinen der in Komplexmedium gewachsenen Zellen.

Es konnten lediglich einige wenige Proteine als reguliert erkannt werden, deren Färbintensität sich stark voneinander unterschied. Abbildung 4-18 B zeigt vier Proteine für die eine Regulation unter den gewählten Bedingungen aus den 2-D Gelen abgeleitet werden konnte. Die Proteine OE2451R (probable oxidoreductase), OE4301R (ABC-transport system ATP-binding protein) und OE4230F (probable acylaminoacyl-peptidase) schienen im Komplexmedium hochreguliert zu sein, wohingegen das Protein OE3636F (L-aspartate oxidase) auf den Gelen der Proteine aus synthetischem Medium intensivere Proteinspots lieferte. Für weitere Proteine war der Nachweis einer Regulation aus bereits genannten Gründen leider nicht möglich.

Anhand der geringen Anzahl an identifizierten regulierten Proteinen konnten auch hier keine Aussagen über die Anpassung des Organismus an die jeweiligen Wachstumbedingungen getroffen werden.

### 4.2.3 Vergleich des Wachstums in Komplex- bzw. synthetischem Medium mittels LC-ICPL

Für den anstehenden ICPL-Versuch wurde folgende methodische Kombination gewählt (siehe Abb. 4-19): Die Proteinprobe wurde nach Markierung und Vereinigung

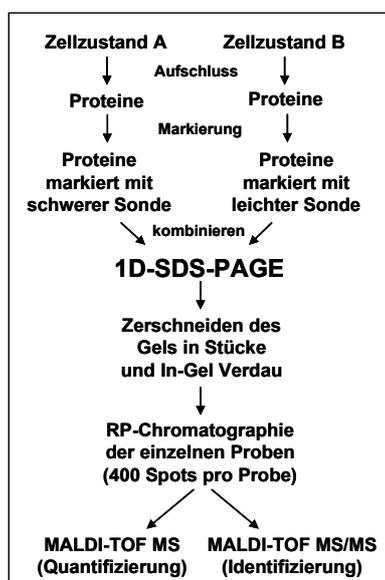


Abb. 4-19: Methodischer Ablauf des ICPL-Versuchs zur Analyse differenziell regulierter Proteine.

mittels 1D-SDS PAGE nach Größe getrennt, das Gel nach Coomassiefärbung in 19 einzelne Stücke (*Slices*) geschnitten und die in ihnen befindlichen Proteine mit Trypsin gespalten. Das komplexe Peptidgemisch einer jeden Gelprobe, welches die tryptischen Fragmente mehrerer hundert Proteine enthalten kann, wurde anschließend mittels RP-Chromatographie aufgetrennt und automatisch zusammen mit Matrixlösung in 400 Fraktionen auf ein Proben target aufgebracht. Die Regulationsinformation der Proteine wird dem aufgenommenen Massenspektrum durch den Vergleich der jeweiligen Isotopenpaare entnommen, wobei die Identifizierung über die Fragmentspektren (MS/MS) der einzelnen Peptide möglich ist.

### 4.2.3.1 Abschätzung der Scoregrenzen

Um die Anzahl an falsch-positiven Identifizierungen zu minimieren, wurden zunächst die Mascot Scoregrenzen für die Identifizierung der Proteine über die MS/MS-Spektren der Peptide überprüft. Dazu wurden alle 45,319 aufgenommenen MS/MS-Spektren gegen die umgedrehte *H. salinarum* Protein Datenbank recherchiert. In dieser Datenbank sind alle Proteinsequenzen invertiert. Dies hat zur Folge, dass die internen tryptischen Peptide (Trypsin spaltet nach ICPL-Markierung nur noch nach Arginin, aber nicht mehr nach dem modifizierten Lysin) zwar die gleiche Masse besitzen, sich die Fragmentserien, die zur MS/MS-Identifizierung herangezogen werden, jedoch unterscheiden. Die Recherche gegen eine solche Protein Datenbank eignet sich demnach gut, um die Anzahl an falsch positiven Identifizierungen unter den gewählten Scoregrenzen abzuschätzen [Moore *et al.*, 2002; Peng *et al.*, 2003].

**Tab. 4-1:** Ergebnisse nach Recherche des Datensatzes gegen die richtige und umgedrehte Datenbank. Gezeigt sind Anzahl an identifizierten Peptiden (2), Singlets (4) sowie Isotopenpaaren (6) bei der Recherche gegen die richtige *H. salinarum* Datenbank bei unterschiedlichen Mascot Score cut-offs (1) sowie Ergebnisse der Recherche gegen die umgedrehte Datenbank zur Abschätzung des Anteils an falsch positiven Identifizierungen für alle Peptide (3), Singletts (5) sowie Isotopenpaare (7). In Klammern ist jeweils der prozentuale Anteil der Ergebnisse bei Suche gegen die umgedrehte Datenbank an den Suchergebnissen gegen die richtige Datenbank angegeben.

MASCOT SCORE <sup>1</sup>	identifizierte Peptide		identifizierte Singletts		identifizierte Isotopenpaare	
	normale DB <sup>2</sup>	umgedrehte DB <sup>3</sup>	normale DB <sup>4</sup>	umgedrehte DB <sup>5</sup>	normale DB <sup>6</sup>	umgedrehte DB <sup>7</sup>
70%	22161	3211 (14.5%)	9908	2434 (24.5%)	12253	777 (6%)
75%	21668	3208 (14.8%)	9659	2432 (25.2%)	12009	776 (6.1%)
80%	20969	2657 (12.7%)	9324	2004 (21.5%)	11645	653 (5.3%)
85%	19973	1885 (9.4%)	8870	1415 (16.1%)	11103	470 (4.1%)
90%	18684	1154 (6.1%)	8251	865 (10.5%)	10433	289 (2.7%)
93%	17639	748 (4.2%)	7770	558 (7.2%)	9869	190 (1.9%)
95%	16684	495 (3%)	7344	362 (4.9%)	9340	133 (1.4%)
98%	14383	185 (1.3%)	6361	123 (1.9%)	8022	62 (0.8%)

Es wurden unterschiedliche Scoregrenzen für die Recherche des Datensatzes gegen die reelle und die umgedrehte Datenbank eingestellt und die Anzahl der identifizierten Peptide beider Recherchen miteinander verglichen. Dabei wurden bei der Auswertung zusätzlich zwischen ICPL-modifizierten Peptiden (Isotopenpaare) und nicht-modifizierten Peptiden (Singletts) unterschieden. Tabelle 4-1 zeigt die Ergebnisse der Recherchen für alle Peptide, für ICPL-modifizierte Peptide (Isotopenpaare) sowie für nicht-modifizierte Peptide (Singletts) bei unterschiedlichen Scoregrenzen.

Es ist ersichtlich, dass für den gesamten Datensatz eine Mascot-Sicherheit von etwa 98% zu wählen ist, um einen Anteil an Falschpositiven von unter 2% zu erhalten (hier 1.3 % bei 98%). Dies gilt ebenfalls für die Singletts. Auch hier wurde eine Rate an falsch positiven Suchergebnissen von unter 2% gegen die umgedrehte Datenbank erst bei einer Mascot-Sicherheit von 98% erreicht (hier 1.9%). Es zeigte sich, dass die von Mascot vorgeschlagenen Sicherheiten in einem guten Maß mit den tatsächlichen

prozentualen Anteilen an Falschpositiven übereinstimmten. Für die Isotopenpaare hingegen entsteht ein anderes Bild. Es konnte bereits bei einer Mascot Grenze von 93% ein Anteil an Falschpositiven von unter 2% erreicht werden (hier 1.9%). Hätte man für die Recherche der Isotopenpaare eine Mascot Sicherheit von 98% gewählt, die nur noch 0.8% falsch-positive Ergebnisse liefert, wären 1318 Peptide als falsch-negative Identifizierungen verloren gegangen, was 14% aller identifizierten Peptide entspricht.

Bei den Isotopenpaaren kann der Abstand zwischen den beiden Isotopensignalen als zusätzliches Korrekturkriterium für die von Mascot vorgeschlagene Peptidsequenz herangezogen werden (siehe „Material und Methoden“ 3.2.12.2). Dadurch kann ein Teil der falschen Suchergebnisse direkt ausgeschlossen werden, was es erlaubt, für die Mascot Recherche geringere Scoregrenzen zu verwenden. Ausgehend von diesen Ergebnissen wurden nun die Scoregrenzen für die Recherchen der ICPL-Datensätze gegen die reelle *H. salinarum* Proteindatenbank festgelegt.

**Isotopenpaare: total ion-score: 20 (95%), ion-score: 15**

**Singletts: total ion-score: 25 (98%), ion-score: 20**

Mit diesen festgelegten Scoregrenzen wurden die ICPL-Datensätze gegen beide Datenbanken recherchiert (die umgedrehte sowie die reelle). Es bleibt anzumerken, dass die Scoregrenze für die Isotopenpaare mit 95% bei einer Quote an Falschpositiven von 1.3% und somit unter der 2% Grenze lag.

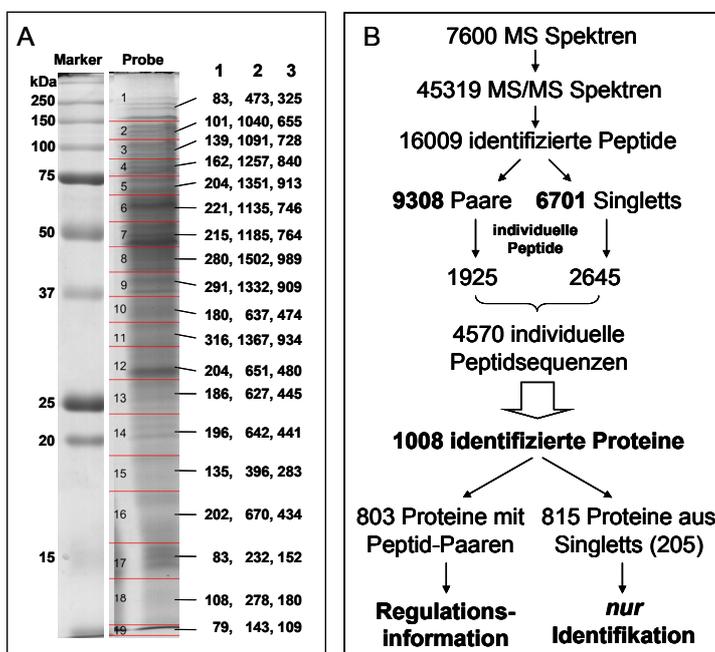
Bei näherer Betrachtung der falsch identifizierten Peptide zeigte sich, dass es sich bei ihnen im Vergleich zu den real identifizierten um überwiegend kurze Peptide handelte. Gut 60% aller falsch-positiven Identifizierungen besaßen eine kleinere molekulare Masse als 1000 Da, im Gegensatz zu nur 8% bei den reell identifizierten Peptiden. Kurze Sequenzen neigen eher zu falsch-positiven Identifizierungen, da sie im MS/MS-Spektrum weniger Fragmente liefern, was eine zufällige Übereinstimmung der Signale mit den theoretischen Fragmentationen erleichtert. Zusätzlich wiesen die Falschpositiven eine höhere Anzahl an *missed cleavage sites* auf (40% gegenüber 4%).

#### 4.2.3.2 Die identifizierten Proteine

Mit den in 4.2.3.1 ermittelten Mascot Scoregrenzen wurden für die Recherche des Datensatzes gegen die reelle *H. salinarum* Proteindatenbank folgende Ergebnisse erhalten. Aus den 19 Gelstücken ergaben sich 7,600 Massenspektren, da die Peptide jeder Gelprobe in 400 *Analyse-Spots* mittels RP-Chromatographie fraktioniert wurden. Jede dieser Fraktionen wurde mittels MALDI-TOF MS vermessen und lieferte so ein

Massenspektrum, welches automatisch nach Peptidpaaren mit einem Massenabstand entsprechend dem Massenunterschied der ICPL-Sonden (ein Vielfaches von 6 Da: 6, 12, 18,...36 Da), sowie nach einzelnen Signalen (Singletts) für die nachfolgende MS/MS-Messung abgesucht wurde. Bei den Peptid-Paaren erfolgte die MS/MS-Messung nur von dem intensiveren der beiden Signale. Aus den 7,600 MS-Spektren konnten so insgesamt 45,319 MS/MS-Spektren aufgenommen werden. Daraus ergaben sich 16,009 identifizierte Peptide, wobei es sich um 9,308 Peptidpaare (1926 individuelle Peptidsequenzen) und 6,701 Singletts (2,645 individuelle Peptidsequenzen) handelte. Aus diesen Peptididentifizierungen ergaben sich 1,008 identifizierte Proteine, 803 von ihnen mit isotopenmarkierten Peptidpaaren, was die Ableitung einer Regulationsinformation ermöglichte.

Abbildung 4-20 A zeigt das mit Coomassie gefärbte 1-D Gel der ICPL-markierten cytosolischen Proteine. Weiter sind die Proteinidentifizierungen sowie die Anzahl der gefundenen Peptide und der individuellen Peptidsequenzen der einzelnen Gelstücke (rote Linien) angegeben. In B sind für den gesamten Versuch die aus allen MS- und MS/MS-Spektren gewonnenen Peptididentifizierungen, ihre Aufteilung in Isotopenpaare und Singletts sowie die sich daraus ergebende Anzahl an Proteinidentifizierungen und Regulationsinformationen illustriert.



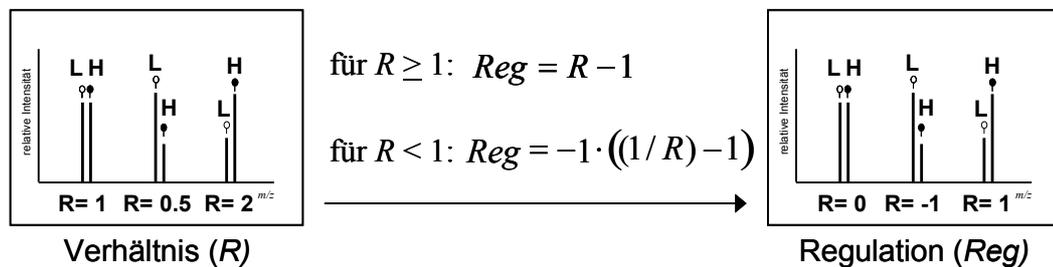
**Abb. 4-20:** (A) Übersicht über die aus den 1-D Gel-slices erhaltenen Proteinidentifikationen (1), die Anzahl der identifizierten Peptide (2) sowie den eindeutigen Peptidsequenzen (3). In (B) sind die aus den MS bzw. MS/MS Spektren hervorgehenden Identifizierungen zusammengefasst.

Die Summe aller Protein-Identifizierungen beträgt 3,385. Bei einer Gesamtzahl von 1,008 gefundenen unterschiedlichen Proteinen bedeutet dies, dass ein Protein durchschnittlich in 3 Banden identifiziert wurde. Dieses eher überraschende Ergebnis kann zum Teil durch proteolytische Abbaufragmente erklärt werden, die aufgrund ihrer geringeren Masse ein anderes Laufverhalten im Gel aufweisen als das intakte Protein. Dem enorm hohen Ausmaß der in unterschiedlichen Banden gefundenen Proteine

müsste demnach jedoch eine starke proteolytische Aktivität bereits in der intakten Zelle oder während der Probenvorbereitung vorliegen.

#### 4.2.3.3 Berechnung und Umrechnung der Regulation

Die Regulationsinformation ergibt sich aus dem Verhältnis (Ratio) des Signals des schweren zum leichten Peptidsignal der Isotopenpaare im Massenspektrum ( $R = \text{Signal}_{\text{schwer}} / \text{Signal}_{\text{leicht}}$ ). Die graphische Darstellung aller regulierten Proteine anhand dieses Verhältnisses ist jedoch problematisch, da alle Regulationen mit einem Verhältnis kleiner eins zwischen null und eins abgebildet werden. Der Ratio wurde deshalb in positive und negative Regulation (Reg) wie in Abbildung 4-21 gezeigt umgerechnet.

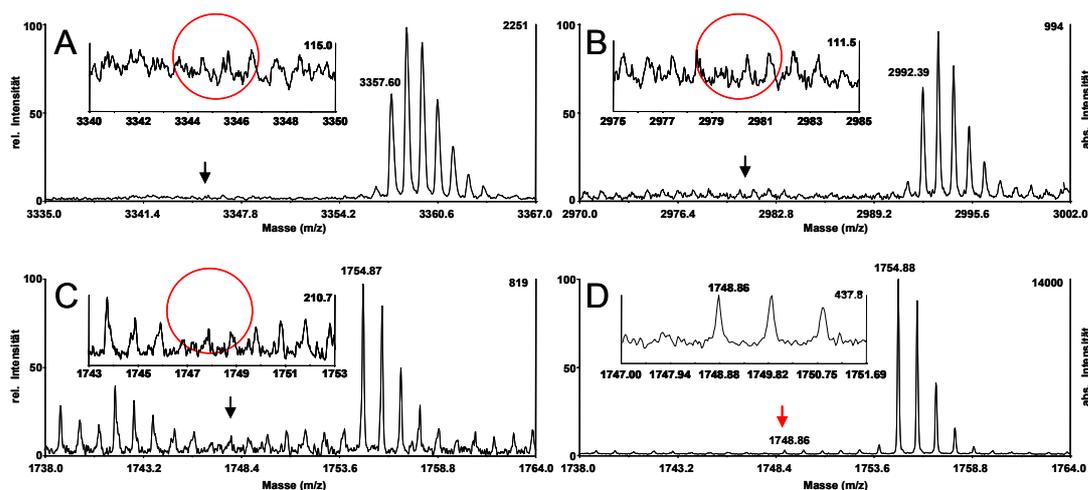


**Abb. 4-21:** Berechnung des Verhältnisses  $R$  der Isotopenpaare sowie Umrechnung in die entsprechenden Regulationen  $R_{\text{neu}}$ .

Diese Regulation (multipliziert mit 100) entspricht dem prozentualen Verhältnis der Isotopensignale (ein Verhältnis von 2 entspricht einer Regulation von 100%). Um die Regulationen aller Proteine zu normieren und so experimentelle Fehler auszugleichen, wurde zunächst der Durchschnittswert aller Regulationen zwischen -1 und +1 berechnet und von allen erhaltenen Werten abgezogen.

#### 4.2.3.4 Stark regulierte Proteine - der dynamische Bereich

Als stärkste Regulation wurde ein Wert von +/-25 vergeben. Derartige Regulationen wurden Proteinen zugeteilt, für die entweder eine Regulation >25 oder <-25 aus den Signalunterschieden der Isotopenpaare berechnet wurde oder von denen nur ein Isotopensignal im Massenspektrum gefunden werden konnte, also nur das Peptid eines Zustands zu detektieren war. Zwischen diesen beiden Fällen kann jedoch nicht endgültig mit der hier verwendeten Methode unterschieden werden. Abbildung 4-22 zeigt Ausschnitte von MALDI-Massenspektren, in denen ICPL-modifizierte Peptide eines Proteins identifiziert wurden.



**Abb. 4-22:** Ausschnitte von Massenspektren mit ICPL-modifizierten Peptiden eines Proteins. (A-C) Massenspektren, in denen nur ein Peptid des jeweiligen Isotopenpaars detektiert wurde. Der Massenbereich in dem der zweite Isotopenpartner zu erwarten wäre (schwarzer Pfeil), ist zusätzlich vergrößert dargestellt und mit einem roten Kreis markiert. (D) Für dasselbe Peptid aus (C) konnte in einem weiteren Massenspektrum jedoch das zugehörige isotopenmarkierte Peptid detektiert werden (roter Pfeil). Der Massenbereich ist ebenfalls vergrößert dargestellt.

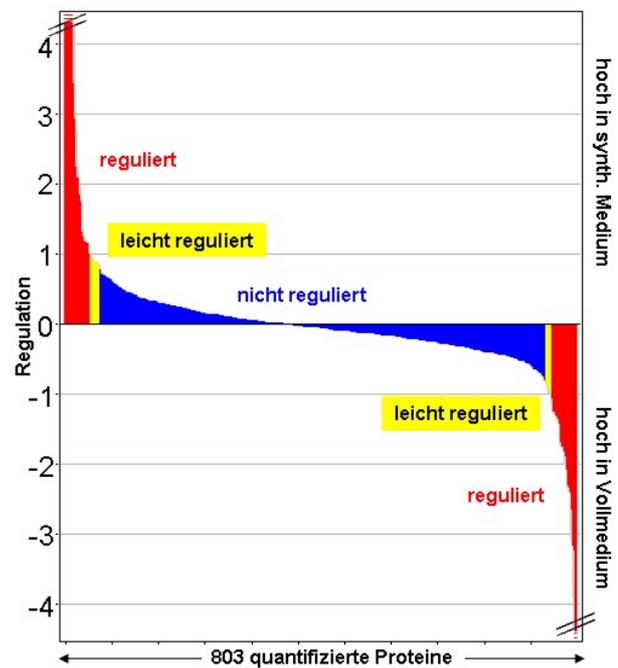
In drei Massenspektren (A-C) wurde nur ein Peptid des Isotopenpaars detektiert, was dafür sprechen würde, dass das Protein nur in einem Zustand in der Zelle vorliegt. In einem weiteren Massenspektrum (D) konnte jedoch der zugehörige Isotopenpartner eines ICPL-Peptids wenn auch mit geringer Intensität eindeutig gefunden werden. Dabei handelte es sich um dasselbe Peptid, für das zuvor in einem anderen Spektrum (C) der zugehörige Isotopenpartner nicht detektiert werden konnte. Es ist darauf hinzuweisen, dass in dem Massenspektrum, in dem das Isotopenpaar gefunden wurde (D), das Signal des schweren ICPL-Peptids mit einer weitaus stärkeren Intensität detektiert wurde als in den anderen drei Spektren (A-C). Dieses Beispiel zeigt, dass es bei stark regulierten Proteinen, d.h. bei Isotopenpaaren, bei denen sich die Konzentrationen der Isotopenpartner stark unterscheiden, sehr von der Sensitivität der Messung bzw. von den in der Probe vorhandenen Absolutmengen der Isotopenpartner abhängt, ob beide detektiert werden. Der dynamische Bereich, der auf Isotopenmarkierungen beruhenden Quantifizierungsmethoden ist demnach ebenfalls limitiert. Er reicht jedoch aus um stark regulierte Proteine zu identifizieren, die bis zu 25fach reguliert sind.

#### 4.2.3.5 Die regulierten Proteine

Für die Betrachtung der unter den unterschiedlichen Wachstumsbedingungen regulierten Proteine wurde ein Verhältnis der Peptidsignale von größer als 2 oder kleiner als 0.5 (Regulation  $\geq 1$  oder  $\leq -1$ ; entsprechend  $>100\%$ ) als biologisch relevante

Regulation angesehen. Dies bedeutet auf Proteinebene, dass ein Protein in einem Zustand in einer mindestens doppelt so hohen Konzentration vorliegen muss. Proteine, die eine Regulation von  $\geq 10.81$ ,  $< 111$  zeigten, wurden als leicht reguliert angesehen. Regulationen zwischen  $0.8$  und  $-0.8$  wurden hingegen als nicht biologisch relevant betrachtet. Die Proteine des Wachstums in synthetischem Medium wurden mit der schweren Isotopensonde, die Proteine aus den Zellen, die im Vollmedium wuchsen, mit der leichten Sonde markiert. Dies ergab positive Regulation für hoch regulierte Proteine aus dem synthetischen Medium und negative Regulationen für Proteine, die im Vollmedium hoch reguliert waren.

Abbildung 4-23 zeigt eine graphische Darstellung aller 803 Proteine, von denen Regulationsinformationen über zugehörige Isotopenpaare verfügbar waren. Es konnten insgesamt 78 regulierte und 20 leicht regulierte Proteine identifiziert werden. Dabei waren 47 beim Wachstum in Vollmedium (6 davon leicht) und 51 bei Wachstum in synthetischem Medium hoch reguliert (14 davon leicht). Die Proteinennamen aller regulierten Proteine, sowie die Anzahl der gefundenen Isotopenpaare, die Regulation und ihre jeweiligen Standardabweichungen sind in den nachfolgenden Tabellen 4-2 A und 4-2 B zu finden.



**Abb. 4-23:** Verteilung der regulierten Proteine in Abhängigkeit ihrer Regulationen. In rot gezeigt sind Proteine mit relevanten Regulationen ( $R \geq 111$ ), gelb bedeutet leichte Regulation ( $10.81 \leq R < 111$ ). Die blauen Proteine wurden als nicht reguliert eingestuft.

**Tabelle 4-2:** Proteine die als hoch reguliert in Komplexmedium (A) und in synthetischem Medium (B) identifiziert wurden. Gegeben sind Gen-Code und Proteinname (1), die Anzahl der identifizierten Isotopenpaare die zur Berechnung der Regulation beigetragen haben sowie in Klammern die Anzahl an gefundenen unterschiedlichen Isotopenpaaren pro Protein (2), die berechnete Regulation (3) sowie die Standardabweichung (4).

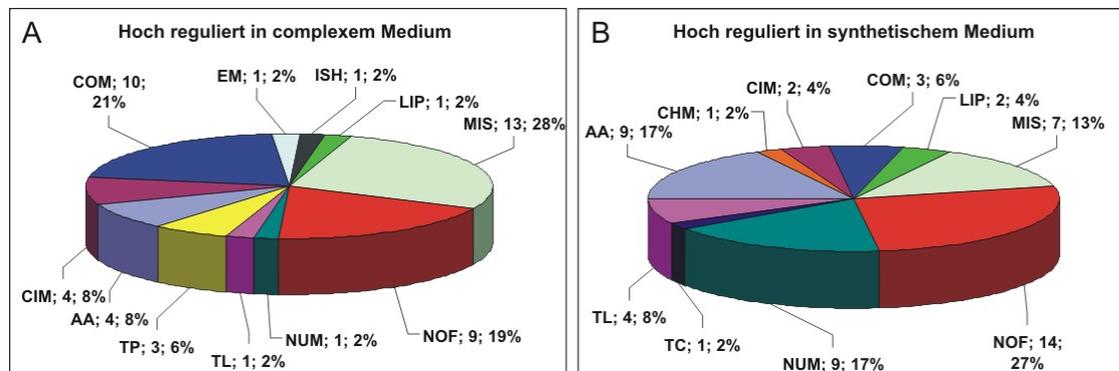
### A: hoch reguliert im Komplexmedium

<b>(Gen-Code) Proteinname<sup>1</sup></b>	<b>Pep<sup>2</sup></b>	<b>Reg.<sup>3</sup></b>	<b>SD<sup>4</sup></b>
(OE4651F) thiamine biosynthesis protein (thiazole biosynthetic enzyme)	9 (1)	<-25.00	0.00
(OE4230F) probable acylaminoacyl-peptidase (EC 3.4.19.1)	2 (2)	-6.44	0.49
(OE2057F) thiamin biosynthesis protein thiC	5 (3)	-6.30	1.28
(OE1613R) probable acylaminoacyl-peptidase (EC 3.4.19.1)	5 (3)	-5.33	1.04
(OE4714F) protein OE4714F	1 (1)	-3.93	-
(OE2451R) probable oxidoreductase	5 (2)	-3.40	0.50
(OE3045F) conserved protein	1 (1)	-3.22	-
(OE1500R) pyruvate water dikinase (EC 2.7.9.2) (phosphoenolpyruvate synthase)	4 (2)	-3.20	1.13
(OE2614F) conserved protein	1 (2)	-3.10	-
(OE4301R) ABC-type transport system ATP-binding protein	7 (4)	-2.63	0.92
(OE3237F) precorrin-8X methylmutase (EC 5.4.1.2)	3 (2)	-2.44	0.23
(OE3221F) conserved cobalamin operon protein	5 (1)	-2.43	0.20
(OE4438F) leucyl aminopeptidase (EC 3.4.11.1)	2 (1)	-2.38	0.27
(OE4331R) tryptophanase (EC 4.1.99.1)	1 (1)	-2.37	-
(OE2186R) conserved protein	2 (1)	-2.33	0.15
(OE4157F) probable chlorohydrolase	1 (1)	-2.04	-
(OE3209F) probable precorrin-2 C20-methyltransferase (EC 2.1.1.130)	4 (3)	-1.95	0.20
(OE3229R) probable chelatase (EC 4.99.1.-) (cobalamin cluster)	3 (2)	-1.94	0.54
(OE3216F) precorrin-3B C17-methyltransferase (EC 2.1.1.131) 2	5 (2)	-1.92	0.31
(OE4673F) carboxypeptidase (EC 3.4.-.-)	24 (4)	-1.85	0.34
(OE3212F) precorrin-4 C11-methyltransferase (EC 2.1.1.133)	3 (1)	-1.85	0.36
(OE2014F) farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (squalene synthase)	1 (1)	-1.80	-
(OE3230F) cobN protein (probable cobalt chelatase) (EC 4.99.1.-)	24 (6)	-1.79	0.32
(OE1405R) conserved protein	2 (1)	-1.78	0.38
(OE3000R) conserved protein	3 (1)	-1.76	0.14
(OE5208R) arginine deiminase (EC 3.5.3.6)	38 (8)	-1.71	0.45
(OE1154F) glyceraldehyde -3-phosphate dehydrogenase (NAD(P))	15 (3)	-1.50	0.16
(OE3214F) precorrin-3B C17-methyltransferase (EC 2.1.1.131) 1	14 (3)	-1.47	0.45
(OE1903F) glycine-rich protein	3 (2)	-1.43	0.36
(OE3213F) cbtG protein (unassigned function)	5 (3)	-1.43	0.23
(OE1711R) oxoglutarate-ferredoxin oxidoreductase (EC 1.2.7.3) alpha chain	11 (5)	-1.40	0.39
(OE2190R) aldehyde dehydrogenase (EC 1.2.1.-)	4 (3)	-1.34	0.18
(OE4302R) ABC-type oligopeptide transport system ATP-binding protein	2 (1)	-1.33	0.32
(OE1465F) endopeptidase La (EC 3.4.21.53)	2 (2)	-1.32	0.67
(OE4018F) conserved protein	1 (1)	-1.30	-
(OE3488R) probable creatininase (EC 3.5.2.10)	2 (1)	-1.29	0.27
(OE3207F) probable precorrin-8W decarboxylase (EC 1.-.-.-)	5 (3)	-1.28	0.21
(OE3224F) conserved cobalamin operon protein	3 (2)	-1.21	0.75
(OE3425F) adenylate kinase (EC 2.7.4.3)	3 (1)	-1.16	0.15
(OE4727R) probable IS1341-type transposase	3 (3)	-1.09	0.01
(OE4571R) probable leucyl aminopeptidase (EC 3.4.11.1)	10 (2)	-1.02	0.22
(OE1710R) oxoglutarate-ferredoxin oxidoreductase (EC 1.2.7.3) beta chain	6 (1)	-0.99	0.29
(OE1954F) NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain B	3 (2)	-0.93	0.18
(OE3566F) conserved protein	4 (3)	-0.87	0.13
(OE2247R) conserved protein	1 (1)	-0.83	-
(OE2532R) conserved protein	1 (1)	-0.81	-
(OE3807R) oligoendopeptidase	17 (5)	-0.81	0.21

**B: hoch reguliert im synthetischen Medium**

<b>(Gen- Code) Proteinname<sup>1</sup></b>	<b>Pep<sup>2</sup></b>	<b>Reg.<sup>3</sup></b>	<b>SD<sup>4</sup></b>
(OE1470F) tryptophan synthase (EC 4.2.1.20) beta chain	6 (1)	>25.00	0.41
(OE4132R) tryptophan-tRNA ligase (EC 6.1.1.2)	6 (4)	>25.00	0.82
(OE2367F) aldehyde dehydrogenase	6 (2)	>25.00	1.63
(OE1471F) tryptophan synthase (EC 4.2.1.20) alpha chain	4 (4)	>25.00	3.50
(OE4399F) probable oxidoreductase (EC 1.1.1.-)	1 (1)	18.90	-
(OE3554F) carbamoyl-phosphate synthase (EC 6.3.-.-) large chain	14 (6)	12.53	1.76
(OE3556R) carbamoyl-phosphate synthase (EC 6.3.-.-) small chain	3 (1)	11.05	2.76
(OE3334R) anthranilate phosphoribosyltransferase (EC 2.4.2.18)	2 (2)	9.40	0.71
(OE3646F) L-aspartate oxidase (EC 1.4.3.16)	2 (2)	5.83	0.14
(OE1951F) phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21)	24 (4)	5.41	1.67
(OE4187R) probable DNA-binding protein	1 (1)	5.21	-
(OE2274R) phosphoribosylformylglycinamide synthase (EC 6.3.5.3) component II	2 (2)	3.65	0.35
(OE3038F) methylenetetrahydrofolate dehydrogenase (NAD(P)+) (EC 1.5.1.5)	5 (2)	3.50	0.64
(OE4329F) conserved protein	3 (3)	3.39	0.59
(OE2864F) phosphoribosylamine-glycine ligase (EC 6.3.4.13)	18 (3)	2.89	0.48
(OE1821R) ribosomal protein L36a.eR (HLA)	2 (1)	2.71	0.08
(OE2813R) conserved protein	1 (1)	2.20	-
(OE3732R) conserved protein	2 (1)	2.10	0.29
(OE1495R) pyruvate kinase (EC 2.7.1.40)	17 (6)	2.05	0.35
(OE1620R) phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)	36 (5)	2.05	0.97
(OE7027R) gas-vesicle operon protein gvpl1	4 (4)	2.01	0.33
(OE3591F) glutamyl-tRNA(Gln) amidotransferase (EC 6.3.5.-) chain D	1 (1)	1.90	-
(OE5162D1F) conserved protein	2 (1)	1.80	0.16
(OE3644F) quinolinate synthetase A (EC 4.1.99.-)	1 (1)	1.70	-
(OE1623F) adenylosuccinate lyase (EC 4.3.2.2)	14 (5)	1.43	0.40
(OE2122F) conserved protein	1 (1)	1.31	-
(OE4412R) threonine synthase (EC 4.2.99.2)	1 (1)	1.27	-
(OE5318F) protein OE5318F	6 (2)	1.21	0.34
(OE5202F) aspartate carbamoyltransferase (EC 2.1.3.2) regulatory chain	11 (3)	1.20	0.14
(OE2084R) transcription initiation factor TFB	2 (2)	1.14	0.33
(OE3724F) phosphoribosylaminoimidazolesuccinocarboxamide synthase (EC 6.3.2.6)	13 (4)	1.13	0.35
(OE3901R) probable N-acetyltransferase (EC 2.3.1.-)	2 (1)	1.13	0.04
(OE3439F) conserved protein	1 (1)	1.13	-
(OE2912F) probable fatty-acid-CoA ligase (EC 6.2.1.-)	56 (8)	1.12	0.33
(OE4420R) argininosuccinate synthase (EC 6.3.4.5)	4 (2)	1.12	0.27
(OE1602F) glycerol-1 phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261)	1 (1)	1.10	-
(OE5201F) aspartate carbamoyltransferase (EC 2.1.3.2) catalytic chain	19 (4)	1.00	0.18
(OE1146R) probable molybdenum cofactor biosynthesis protein moeA 1	1 (1)	0.96	-
(OE4260R) probable N-acetyltransferase (EC 2.3.1.-)	2 (2)	0.94	0.08
(OE3953R) orotate phosphoribosyltransferase (EC 2.4.2.10)	5 (1)	0.91	0.05
(OE5106F) trkA domain protein	2 (1)	0.90	0.28
(OE8047F) conserved protein	4 (2)	0.89	0.27
(OE2602R) ribosomal protein L1	1 (1)	0.86	-
(OE4408F) phosphoglycerate dehydrogenase (EC 1.1.1.95)	10 (1)	0.86	0.26
(OE1016R) glucose-1-phosphate thymidyltransferase (EC 2.7.7.24)	2 (1)	0.85	0.08
(OE5285R) hypothetical protein	2 (2)	0.85	0.07
(OE1614F) hexosyltransferase homolog	1 (1)	0.84	-
(OE1729R) protein N-acetyltransferase homolog	1 (1)	0.84	-
(OE3922R) glutamate-ammonia ligase (EC 6.3.1.2)	11 (5)	0.84	0.13
(OE2138F) probable acyl/butyryl-CoA dehydrogenase (EC 1.3.99.-)	13 (4)	0.82	0.35
(OE5211F) conserved protein	5 (4)	0.81	0.16

Die Einteilung der regulierten Proteine in ihre jeweiligen Funktionsklassen ließ bereits eine grobe Betrachtung der unter den verschiedenen Wachstumsbedingungen regulierten Proteine zu:



**Abb. 4-24:** Aufteilung der induzierten Proteine nach Funktionsklassen. (A) Wachstum in Komplexmedium und (B) Wachstum in synthetischem Medium. Abkürzungen der Funktionsklassen: COM: Cofaktor Metabolismus, CIM: Intermediärmetabolismus, AA: Aminosäuren Metabolismus, TP: Transport kleiner Moleküle, TL: Translation, TC: Transkription, NUM: Nukleotid Metabolismus, NOF: No Function, MIS: Miscellaneous, ISH: ISH-kodierende Transposasen, EM: Energie Metabolismus, LIP: Lipid Metabolismus

Aus Abbildung 4-24 ist deutlich zu erkennen, dass im komplexen Medium verstärkt Proteine des Cofaktor-Metabolismus (COM) gegenüber dem Wachstum in synthetischem Medium exprimiert waren. Des Weiteren waren Transportproteine (TP) nur im komplexen Medium hoch reguliert. Dagegen werden Proteine des Nukleotid (NUM) sowie Aminosäure Metabolismus (AA) im synthetischen Medium hoch reguliert.

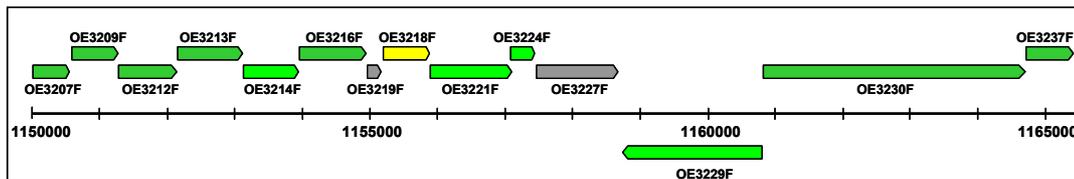
Für eine genauere Analyse der physiologischen Antwort von *H. salinarum* auf die hier verglichenen unterschiedlichen Lebensbedingungen bedurfte es einer genaueren Analyse der biologischen Funktionen der regulierten Proteine. Diese konnten nach detaillierter Betrachtung verschiedenen biologischen und metabolen Reaktionswegen zugeordnet werden.

#### 4.2.3.6 Zur Biologie der regulierten Proteine

##### (A) Hoch reguliert im Komplexmedium (negative Regulation):

Die Biosynthesewege von zwei Cofaktoren, von Thiamin (Vitamin B<sub>1</sub>) und Cobalamin (Vitamin B<sub>12</sub>), sind im Komplexmedium hoch reguliert. Das Thiamin Biosyntheseprotein OE4651F ist dabei eines der am stärksten regulierten Proteine (Regulation: <-25). Thiamin scheint im Komplexmedium nicht in ausreichenden Mengen vorzuliegen, während es in das synthetische Medium in einer Konzentration von 19 mM zugegeben wird.

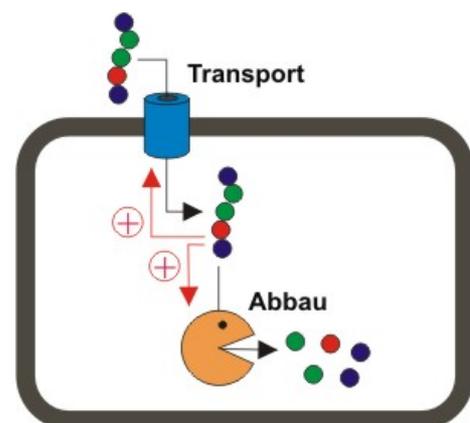
Insgesamt 11 Proteine, die an der Biosynthese des Porphyrins Cobalamin (Vitamin B<sub>12</sub>) beteiligt sind, werden im Komplexmedium hoch reguliert. Einige dieser Proteine werden in einem Operon transkribiert (siehe Abb. 4-25), und zeigen erwartungsgemäß Regulationen in annähernd gleichem Maß (Regulationen zwischen -1.21 und -2.43). Proteine, die an der Biosynthese anderer Porphyrine wie z.B. Häm beteiligt sind, konnten durch ICPL-Daten als nicht reguliert klassifiziert werden. Es ist also die Cobalamin Biosynthese im Reaktionsweg der Porphyrin Biosynthese selektiv reguliert.



**Abb. 4-25:** Genomregion der 11 hoch regulierten Proteine (grüne Pfeile) der Cobalamin Biosynthese. Das gelb markierte Protein konnte lediglich ohne Regulationsinformation identifiziert werden, wohingegen das grau markierte Protein nicht identifiziert wurde. Einige dieser Gene werden in einem Operon expremiert.

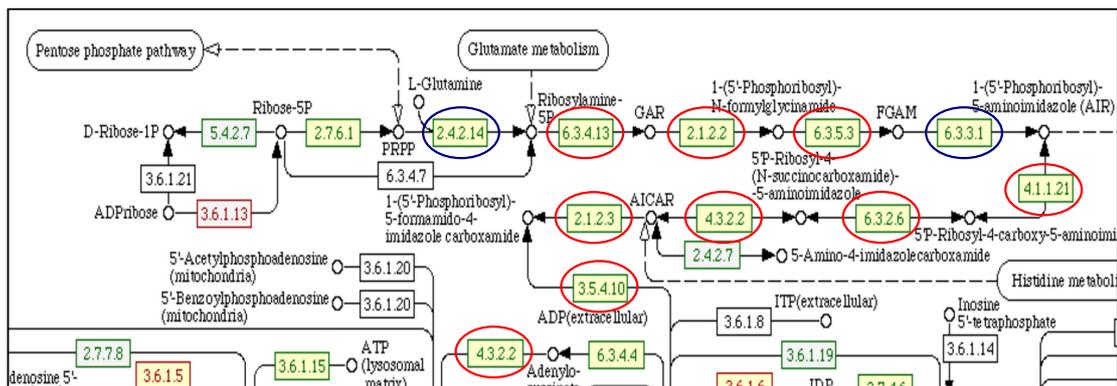
Mehrere Enzyme mit bekannten Funktionen werden im Komplexmedium hoch reguliert. Dazu gehören unter anderem die ‘pyruvate water dikinase’ (OE1500R) sowie die beiden Untereinheiten der ‘oxoglutarate-ferridoxine oxidoreductase’ (OE1710R und OE1711R). Beide Untereinheiten dieses Enzyms zeigten ähnliche Regulationen (Regulation: -0,99 und -1.40). Beide Enzyme katalysieren Schlüsselreaktionen der reduktiven CO<sub>2</sub>-Fixierung und des Beginns der Gluconeogenese.

Eine weitere Gruppe der im Komplexmedium hoch regulierten Proteine besitzt proteolytische Aktivität. Hierzu gehörten zwei Acylaminoacyl-Peptidasen, welche acetylierte N-terminale Aminosäuren abspalten, zwei Leucyl-Aminopeptidasen, zwei Endopeptidasen, sowie eine Carboxypeptidase. Das Komplexmedium enthält neben Aminosäuren eine beträchtlich hohe Anzahl an Peptiden. Diese scheint *H. salinarum* als Nahrungsquelle zu erkennen und gezielt Proteasen für deren Abbau bereitzustellen. Interessanterweise konnten zusätzlich zwei Proteine als reguliert identifiziert werden, welche Homologien zu Peptid-Transportsystemen aufweisen. Dies komplettiert das Bild der Metabolisierung von



**Abb. 4-26:** Induktion von Proteinen des Peptidtransports und Peptidabbaus durch im Medium vorliegende Peptide.





**Abb. 4-28:** Ausschnitt der Purin Biosynthese nach KEGG (Rekonstruktionsdaten aus HALOLEX). Für die grün gefärbten Enzyme konnten *H. salinarum* Proteine zugeordnet werden, für die rot gefärbten nicht. Die induzierten Proteine sind durch rote Kreise hervorgehoben. Die blau umkreisten Enzyme wurden lediglich identifiziert, konnten aber nicht quantifiziert werden.

Weiter zeigen sich im synthetischen Medium mehrere Enzyme der Biosynthese von Aminosäuren als hoch reguliert. So waren z.B. beide Untereinheiten der Tryptophan-Synthase stark reguliert (beide Regulationen: >25). Tryptophan wird im synthetischen Medium nicht zugegeben und muss daher von der Zelle selbst synthetisiert werden. Weitere Aminosäure-Biosynthesewege, aus denen Enzyme als hoch reguliert identifiziert wurden, waren der Serin/Threonin-Metabolismus, der Glutamat/Glutamin-Metabolismus sowie die Biosynthese von Arginin. Lediglich Glutamat und Glutamin sind nicht im synthetischen Medium enthalten.

Im Gegensatz zu den Acylaminoacyl Peptidasen, die N-terminal acetylierte Aminosäuren abspalten und im Komplexmedium verstärkt exprimiert wurden, konnten drei N-Acetyltransferasen im synthetischen Medium in erhöhten Konzentrationen identifiziert werden (Regulationen zwischen 0.84 und 1.13). Proteinacetylierung kann als Schutzmechanismus vor Proteinabbau gesehen werden. Im synthetischen Medium unter Mangelbedingungen scheint *H. salinarum* demnach seine Proteine vor proteolytischem Abbau zu schützen.

Interessanterweise wurde die Pyruvate Kinase (OE1495R) als reguliert identifiziert (Regulation: 2.05). Dieses Enzym katalysiert den letzten irreversiblen Schritt der Glykolyse vom Phosphoenolpyruvate zum Pyruvate und stellt somit ein Schlüsselenzym der Glykolyse dar.

#### 4.2.4 Vergleich von aerobem mit phototrophen Wachstum mittels LC-ICPL

Für den Vergleich der Proteome von *H. salinarum* unter aeroben und anaeroben phototrophen Wachstumsbedingungen wurde die gleiche methodische Strategie wie

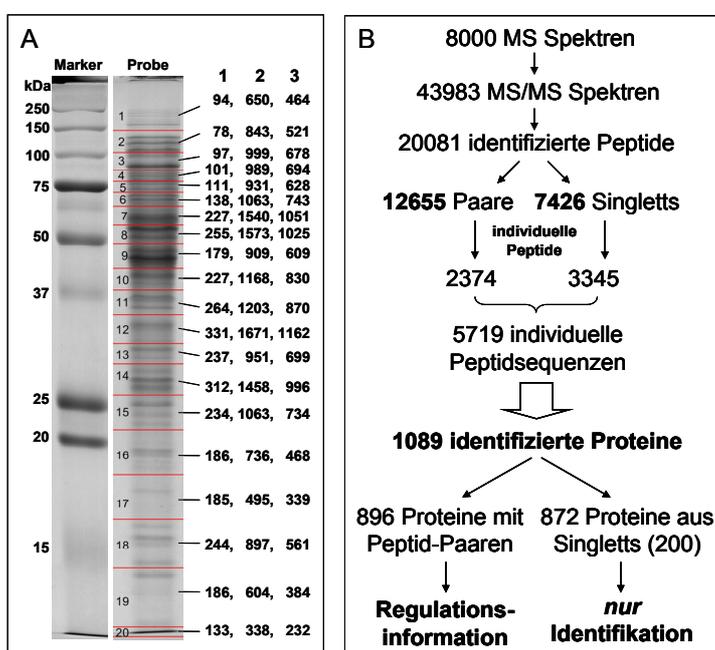
bereits zuvor gewählt. Die Proteine der beiden Zustände wurden nach Markierung und Vereinigung mittels 1-D SDS-PAGE nach Größe getrennt, das Gel in Stücke geschnitten, die Proteine im Gel proteolytisch mit Trypsin gespalten und die entstandenen Peptidfraktionen mittels RP-Chromatographie aufgetrennt. Quantifizierung und Identifizierung der Peptide erfolgte mittels MALDI-TOF MS bzw. MS/MS (siehe dazu Abb. 4-19).

#### 4.2.4.1 Die identifizierten Proteine

Das Gel wurde diesmal nicht in 19 Stücke wie im vorherigen ICPL-Versuch, sondern in 20 Stücke unterteilt. Daraus ergaben sich 8000 MS-Spektren bei 400 Analyse-Spots pro Peptidprobe. Die Quantifizierung der Isotopenpaare sowie die

Auswahl der MS/MS Analysen erfolgten wie beschrieben. Es wurden 43,983 MS/MS Spektren aufgenommen. Dies waren 1,336 weniger als im ersten Versuch. Für die Recherche gegen die *H. salinarum* Datenbank wurden die in 4.2.3.1 ermittelten Mascot Scoregrenzen gewählt. Die Anzahl der identifizierten Peptide stieg um 4,072 auf 20,081 an. Darunter waren 12,655 Isotopenpaare (3,347 mehr) und 7,226 Singletts (725 mehr). Die Anzahl der individuellen Peptidsequenzen

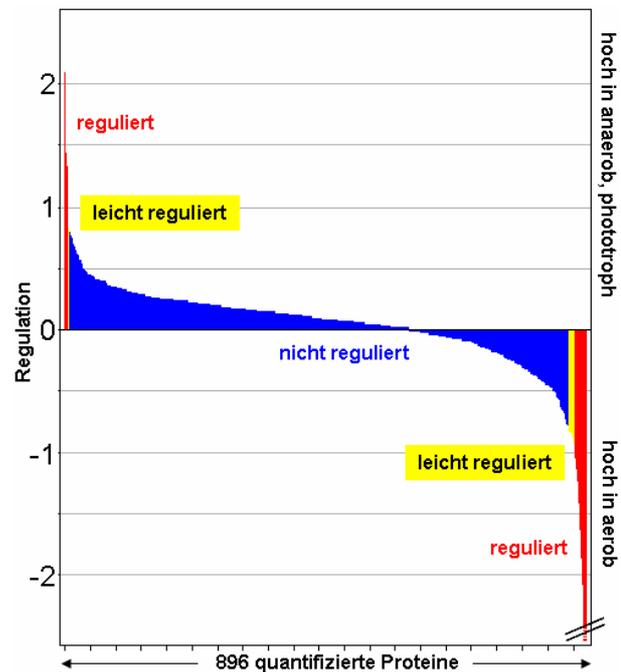
erhöhte sich ebenfalls um 1,159 auf 5,719 individuelle Peptide (2,374 Isotopenpaare, 3,355 Singletts). Anhand der identifizierten Peptide konnten 1,089 Proteine identifiziert werden, also knapp 90 Proteine mehr als im ersten Versuch. Von 896 Proteinen (93 mehr) konnten Regulations-Informationen über identifizierte Isotopenpaare gewonnen werden, wohingegen 200 Proteine nur über Singletts identifiziert werden konnten. Abbildung 4-29 fasst die Identifizierungen aus diesem Versuch zusammen.



**Abb. 4-29:** (A) Übersicht über die aus den 1-D Gel-Slices erhaltenen Proteinidentifikationen (1), die Anzahl der identifizierten Peptide (2) sowie die eindeutigen Peptidsequenzen (3). In (B) sind die aus den MS bzw. MS/MS Spektren hervorgehenden Identifizierungen zusammengefasst.

#### 4.2.4.2 Die regulierten Proteine

Für die Betrachtung der regulierten Proteine wurden die in 4.2.3.5 beschriebenen Kriterien zur Ermittlung der Regulationen herangezogen. Die Proteine der Zellen, die unter anaeroben phototrophen Bedingungen wuchsen, wurden mit der schweren Isotopensonde, die Proteine der Zellen die aerob wuchsen, mit der leichten Sonde markiert. Dies ergab positive Regulationen für hoch regulierte Proteine unter anaeroben, phototrophen Bedingungen und negative Regulationen für Proteine, die im aeroben Wachstum hoch reguliert waren. Abbildung 4-30 zeigt eine graphische Darstellung aller 896 Proteine, von denen Regulationsinformationen über zugehörige Isotopenpaare verfügbar waren.



**Abb. 4-30:** Verteilung der regulierten Proteine in Abhängigkeit ihrer Regulationen. In rot gezeigt sind Proteine mit relevanten Regulationen ( $R \geq 111$ ), gelb bedeutet leichte Regulation ( $10.8 \leq R < 111$ ). Die blauen Proteine wurden als nicht reguliert eingestuft.

Es fiel auf, dass im Vergleich zum vorangegangenen ICPL-Versuch (Vollmedium vs. synthetisches Medium) trotz der größeren Zahl identifizierter Proteine nur knapp die Hälfte an regulierten Proteinen gefunden werden konnte. Besonders unter anaeroben phototrophen Bedingungen waren nur wenige Proteine hoch reguliert. Insgesamt konnten 29 regulierte und 12 leicht regulierte Proteine identifiziert werden. Dabei waren 31 Proteine beim Wachstum unter aeroben Bedingungen (11 davon leicht) und nur 10 Proteine unter anaeroben phototrophen Bedingungen hoch reguliert (1 davon leicht). Die Proteinamen aller regulierten Proteine sowie die Anzahl aller identifizierten Peptide, die Anzahl der gefundenen Isotopenpaare und die Regulation sind in den nachfolgenden Tabellen zu finden.

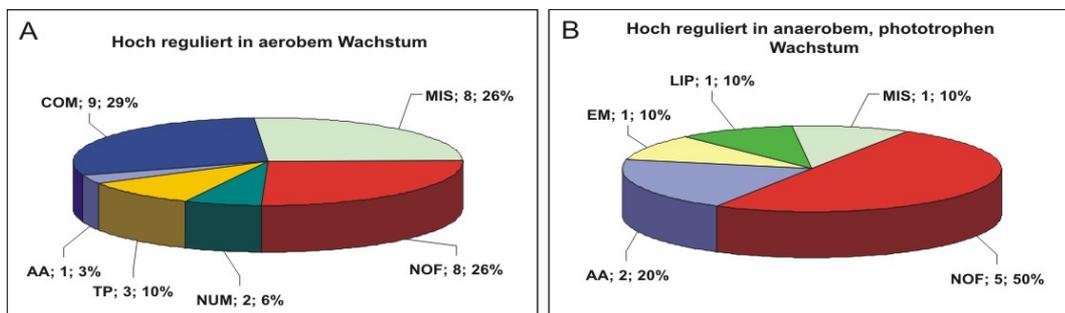
**Tab. 4-3:** Proteine, die als hoch reguliert in aerobem Wachstum (A) und in anaerobem phototrophen Wachstum (B) identifiziert wurden. Gegeben sind Gen-Code und Proteinname (1), die Anzahl der identifizierten Isotopenpaare, die zur Berechnung der Regulation beigetragen haben, sowie in Klammern die Anzahl an gefundenen unterschiedlichen Isotopenpaaren pro Protein (2), die berechnete Regulation (3) sowie die Standardabweichung (4).

<b>A (Gene Code) Proteinname<sup>1</sup> (hoch in aerobem Wachstum)</b>	<b>Pep<sup>2</sup></b>	<b>Reg.<sup>3</sup></b>	<b>SD<sup>4</sup></b>
(OE4346R) ribonucleoside-diphosphate reductase (EC 1.17.4.1) large chain	2 (2)	-9.50	0.79
(OE4345R) probable ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain	1 (1)	-8.94	-
(OE2057F) thiamin biosynthesis protein thiC	10 (4)	-5.75	1.30
(OE3221F) conserved cobalamin operon protein	1 (1)	-4.50	-
(OE4651F) thiamine biosynthesis protein (thiazole biosynthetic enzyme)	12 (1)	-2.90	0.73
(OE3213F) CbiG protein (unassigned function)	6 (4)	-2.18	0.57
(OE4654F) phosphomethylpyrimidine kinase (EC 2.7.4.7)	2 (2)	-2.11	0.30
(OE3209F) probable precorrin-2 C20-methyltransferase (EC 2.1.1.130)	5 (4)	-1.99	1.00
(OE3224F) conserved cobalamin operon protein	2 (1)	-1.89	0.54
(OE4311F) ABC-type transport system periplasmic substrate-binding protein	2 (2)	-1.66	0.08
(OE3207F) probable precorrin-8W decarboxylase (EC 1.-.-.-)	5 (4)	-1.53	0.58
(OE3216F) precorrin-3B C17-methyltransferase (EC 2.1.1.131) 2	4 (3)	-1.44	0.62
(OE3315R) ABC-type cobalt transport system ATP-binding protein	1 (1)	-1.42	-
(OE1932R) conserved protein	1 (1)	-1.33	-
(OE5326R) hypothetical protein	1 (1)	-1.27	-
(OE3214F) precorrin-3B C17-methyltransferase (EC 2.1.1.131) 1	19 (4)	-1.23	0.49
(OE4398F) O-acetylhomoserine (thiol)-lyase (EC 4.2.99.10)	1 (1)	-1.16	-
(OE3212F) precorrin-4 C11-methyltransferase (EC 2.1.1.133)	2 (1)	-1.16	0.50
(OE4551F) ABC-type transport system periplasmic substrate-binding protein	1 (1)	-1.07	-
(OE3237F) precorrin-8X methylmutase (EC 5.4.1.2)	5 (3)	-1.01	0.32
(OE1613R) probable acylaminoacyl-peptidase (EC 3.4.19.1)	30 (5)	-0.96	0.33
(OE4673F) carboxypeptidase (EC 3.4.-.-)	29 (4)	-0.95	0.28
(OE2671R) HemK protein homolog (protoporphyrinogen oxidase homolog)	1 (1)	-0.94	-
(OE5032R) hypothetical protein	1 (1)	-0.90	-
(OE1626F) conserved protein	2 (1)	-0.89	0.07
(OE3036F) glycine hydroxymethyltransferase (EC 2.1.2.1)	8 (1)	-0.89	0.38
(OE2703F) probable copper-containing oxidoreductase	5 (3)	-0.86	0.46
(OE3710R) trkA domain protein	9 (2)	-0.85	0.31
(OE3045F) conserved protein	2 (2)	-0.81	0.08
(OE4230F) probable acylaminoacyl-peptidase (EC 3.4.19.1)	11 (2)	-0.80	0.36
(OE5248F) conserved protein	1 (1)	-0.80	-

<b>B (Gene Code) Proteinname<sup>1</sup> (hoch in phototrophen Wachstum)</b>	<b>Pep<sup>2</sup></b>	<b>Reg.<sup>3</sup></b>	<b>SD<sup>4</sup></b>
(OE5083R) heat shock protein homolog	2 (2)	2.26	0.48
(OE3815R) conserved hypothetical protein	6 (5)	2.06	0.63
(OE2991F) transcription factor TFB homolog	1 (1)	1.56	-
(OE3556R) carbamoyl-phosphate synthase (EC 6.3.-.-) small chain	1 (1)	1.41	-
(OE3101R) bacterioopsin activator	4 (1)	1.40	0.42
(OE1651F) hypothetical protein	2 (2)	1.29	0.30
(OE1876R) probable menaquinol-cytochrome-c reductase (EC 1.10.2.-) Rieske Protein	4 (1)	1.09	0.41
(OE3554F) carbamoyl-phosphate synthase (EC 6.3.-.-) large chain	13 (5)	1.06	0.39
(OE6046F) conserved hypothetical protein	1 (1)	1.06	-
(OE3468R) phytoene dehydrogenase (EC 1.14.99.-) (phytoene desaturase) 2	3 (2)	0.80	0.34

Abbildung 4-31 zeigt die regulierten Proteine entsprechend ihrer biologischen Funktionsklassen.



**Abb. 4-31:** Aufteilung der induzierten Proteine nach Funktionsklassen. (A) Wachstum unter aeroben und (B) Wachstum unter anaeroben phototrophen Bedingungen. Abkürzungen der Funktionsklassen: COM: Cofaktor Metabolismus, AA: Aminosäuren Metabolismus, TP: Transport kleiner Moleküle, NUM: Nukleotid Metabolismus, NOF: No Function, MIS: Miscellaneous, EM: Energie Metabolismus, LIP: Lipid Metabolismus

Im aeroben Wachstum waren wie bereits beim vorangegangenen Vergleich in der korrespondierenden Bedingung (Komplexmedium) Proteine des Cofaktor-Metabolismus (COM) sowie Transportproteine (TP) verstärkt expremiert. Der Vergleich der regulierten Proteine beider Bedingungen war hier jedoch schwieriger, weil unter anaeroben phototrophen Bedingungen insgesamt nur 10 Proteine als reguliert gefunden wurden und davon 5 ohne bekannte biologische Funktion (NOF). Eine genauere Betrachtung der regulierten Reaktionswege ergab folgendes Bild.

#### 4.2.4.3 Zur Biologie der regulierten Proteine

##### (A) Hoch reguliert unter aeroben Wachstumsbedingungen:

Bei der Betrachtung der biologischen Funktionen der Proteine, die beim Übergang von anaerobem phototrophen zu aeroben Wachstum in Komplexmedium hoch reguliert sind, fällt auf, dass die meisten regulierten Proteine bzw. Funktionsklassen auch beim Übergang vom aeroben Wachstum in synthetischem Medium zu Wachstum in Komplexmedium (siehe 4.2.3.6) hoch reguliert sind.

So waren z.B. wiederum Proteine der Thiamin- und Cobalamin-Biosynthese hoch reguliert. Bei letzterer wurden 9 Proteine, darunter etliche eines Operons als reguliert gefunden. Auch 2 Proteine des Peptidtransports (OE4551F und OE4311F) sind reguliert. Interessanterweise handelte es sich jedoch nicht um die im vorherigen Versuch als reguliert gefundenen Proteinen. Zusätzlich ist ein Protein, das am Cobalt-Transport beteiligt ist, hoch reguliert (OE3315R). Dies deckt sich mit der verstärkten Expression der Cobalamin Biosynthese Proteine. Des Weiteren wurden 2 Proteine mit proteolytischer Aktivität gefunden, die bereits beim Übergang vom aeroben Wachstum

in synthetischem Medium zu Wachstum hoch reguliert waren (OE4673F und OE1613F). Die restlichen Proteasen, die ebenfalls bei diesem Übergang reguliert waren, wiesen in diesem Vergleich jedoch keine Regulationen auf.

Die am stärksten regulierten Proteine waren die beiden Untereinheiten der in die Nukleotid-Biosynthese involvierten 'ribonucleoside-diphosphate reductase' (OE4345R und OE4346R) mit Regulationen von 9.5 und 8.94.

### **(B) Hochreguliert unter anaeroben phototrophen Bedingungen**

Wie bereits angemerkt konnten nur 10 Proteine als reguliert identifiziert werden. Davon konnten drei Proteinen keine biologischen Funktionen zugewiesen. Wie bereits im synthetischen Medium zeigten sich die beiden Untereinheiten der Carbamoylphosphat Synthase (OE3556R und OE3554F; Regulationen: 1.41, und 1.06) als hoch reguliert.

Es ist anzunehmen, dass hier die induzierte Aktivität dieses Enzyms, das aus Ammoniak und CO<sub>2</sub> unter dem Verbrauch von 2 ATP Carbamoylphosphat generiert, nicht für die Pyrimidine-Biosynthese gebraucht wird, sondern der Zelle Arginin bereit stellt. Da unter phototrophen Wachstum durch die der Protonenpumpe Bakteriorhodopsin genügend ATP vorliegt, könnte die CO<sub>2</sub>-Fixierung zu Arginin als langfristiger Energiespeicher dienen, da dieses unter anaeroben Bedingungen im Dunkeln unter ATP Bildung fermentiert werden kann.

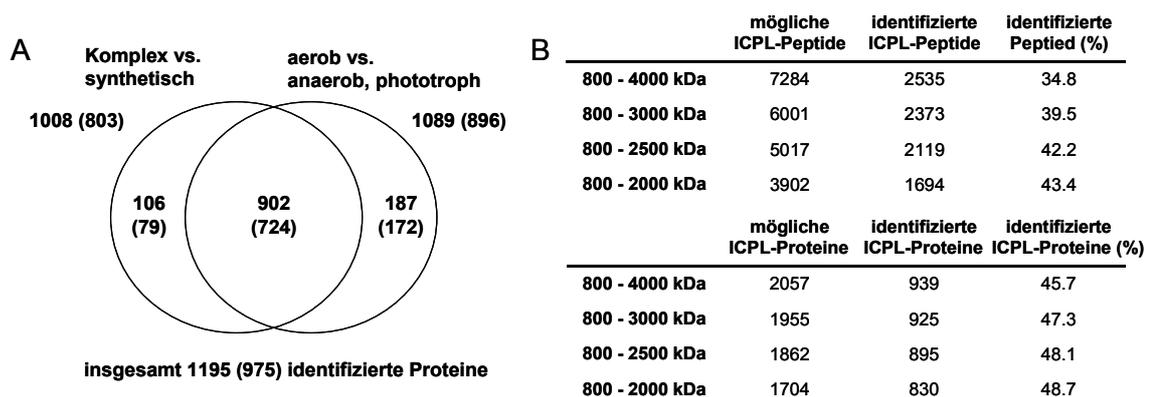
Der hoch regulierte Bacterio-opsin Regulator BAT sowie das TFB Homolog (OE3101R und OE2991F; Regulationen: 1.40 und 1.56) könnten an der Transkriptions Regulation des Bakterio-opsins beteiligt sein, dessen Induktion unter anaeroben phototrophen Bedingungen unabdingbar ist und auch spektroskopisch nachgewiesen werden kann. Da sich die Analyse in dieser Arbeit auf die cytosolischen Proteine beschränkte, konnte Bakteriorhodopsin nicht als reguliert identifiziert werden. Weiter wurde die Phytoene Dehydrogenase als hoch reguliert gefunden, die ein Schlüsselenzym in der Retinal-Biosynthese darstellt. Dieser Chromophor, der dem Retinalprotein Bakteriorhodopsin erst seine biologische Funktion als lichtgetriebene Protonenpumpe ermöglicht, muss erwartungsgemäß ebenfalls in erhöhten Mengen hergestellt werden. Eine plausible Erklärung für die Regulation des Rieske Proteins (OE1876R, Regulation 1.09) sowie für das Heat Shock Protein Homolog (OE5083R, Regulation: 2.26) kann nicht gegeben werden.

## 4.2.5 Gesamtbetrachtung der vergleichenden Analysen

Beide vergleichenden Analysen wurden mit demselben experimentellen Ansatz durchgeführt, was einen detaillierten Vergleich der Ergebnisse beider Experimente untereinander, aber auch die Betrachtung der Ergebnisse als Gesamtbild erlaubt.

### 4.2.5.1 Die identifizierten Proteine

Durch die beiden ICPL-Versuche (1. Komplex vs. synthetisches Medium und 2. aerob vs. anaerob, phototroph) wurden zusammen insgesamt 1,195 Proteine identifiziert, 975 von ihnen mit zugehörigen Isotopenpaaren. Es konnten also von 82% aller identifizierten Proteine Regulationsinformationen erhalten werden. 902 Proteine (724 mit Isotopenpaaren) wurden in beiden Experimenten wiedergefunden, was 75% (74%) aller identifizierten Proteine entspricht. 106 Proteine wurden hingegen nur im ersten (79 mit Isotopenpaaren) und 187 nur im zweiten Versuch (172 mit Isotopenpaaren) gefunden (siehe Abb. 4-32 A). Dies verdeutlicht eine recht gute Reproduzierbarkeit des gewählten methodischen ICPL-Ansatzes.



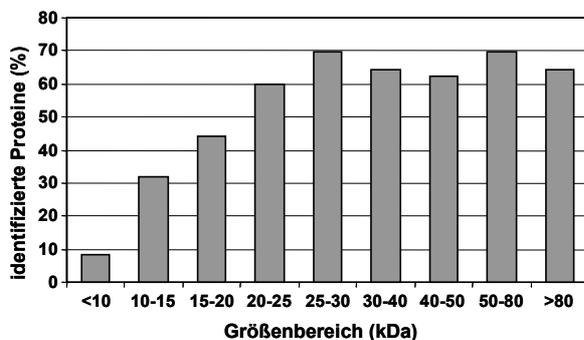
**Abb. 4-32:** Übersicht über identifizierte Proteine und Isotopenpaare aus beiden Versuchen. (A) Vergleich der identifizierten Proteine die nur in einem oder in beiden ICPL-Versuchen gefunden wurden. Die Zahlen in Klammern stehen für die Proteine mit Regulationsinformationen. (B) Vergleich der identifizierten cytosolischen Isotopenpaare (Peptide) und Proteine mit Regulationsinformation mit den theoretisch möglichen in Abhängigkeit von unterschiedlichen Massenfenstern.

Die in den beiden Versuchen identifizierten Proteine entsprachen 53% des cytosolischen Proteoms (ohne vorhergesagte TM) von *H. salinarum*. Bezieht man sich nur auf die chromosomalen Proteine, wurden sogar 65% derer identifiziert, was wiederum die Vermutung untermauert, dass die Plasmidproteine in einem geringeren Ausmaß exprimiert werden. Von gut 42% aller cytosolischen Proteine konnten Regulationsinformationen über zugehörige Isotopenpaare erhalten werden. Dieser Anteil erhöht sich auf rund 55% bei Betrachtung der chromosomalen Proteine. Obwohl sich die ICPL-Analysen auf die cytosolischen Proteine beschränkten, wurden 54

Proteine mit vorhergesagten Transmembranregionen identifiziert. Obwohl der Großteil von ihnen (32) nur 1 bis 2 TM-Regionen besaß, konnten ebenfalls Proteine mit bis zu 16 Transmembranhelices gefunden werden.

Wie bereits erwähnt haben halobakterielle Proteine aufgrund ihres aciden Charakters weniger Lysine, was die Anzahl der lysinhaltigen Peptide und somit auch die der möglichen Isotopenpaare herabsetzt. Bezieht man sich für die cytosolischen Proteine nur auf den Massenbereich, der in den Analysen mittels MALDI-TOF abgedeckt wurde (800 – 4,000 Da), können maximal 7,284 ICPL-Isotopenpaare und daraus resultierend 2,057 der möglichen 2,239 Proteine quantifiziert werden (siehe Abb. 4-32 B). Davon wurden tatsächlich 2,535 Peptide (34.8%) bzw. 939 Proteine (45.7%) gefunden. Der prozentuale Anteil sowohl an identifizierbaren ICPL-Peptiden als auch an quantifizierbaren Proteinen steigt, wenn man die obere Grenze des abgedeckten Massenbereichs herabsetzt. Das zeigt, dass größere Peptide mittels MALDI-TOF MS schwerer detektiert werden können.

Auch in dem hier auf Tandemmassenspektrometrie basierenden methodischen Ansatz zeigte sich eine deutliche Präferenz für Proteine mit höherer molekularer Masse. Die Beobachtung war bereits bei der massenspektrometrischen Analyse der 2-D Gele mittels MALDI-TOF-PMF gemacht worden (siehe 4.1.3). Aus Abbildung 4-33 wird deutlich, dass ca. 60 bis 70% der Proteine größer als 20 kDa identifiziert werden konnten, der Großteil der kleinen Proteine jedoch nicht gefunden wurde. In Bezug auf die chromosomalen Proteine wurden sogar 80% derer >20 kDa identifiziert (nicht



**Abb. 4-33:** Anteil (%) der mittels LC-TOF/TOF identifizierten cytosolischen Proteine in Abhängigkeit von unterschiedlichen Molekulargrößenbereichen.

gezeigt in Abb. 4-33). Nur 139 der 1,195 identifizierten Proteine (11%) lagen auf den halobakteriellen Plasmiden, die jedoch 27% aller vorhergesagten cytosolischen Proteine beherbergen.

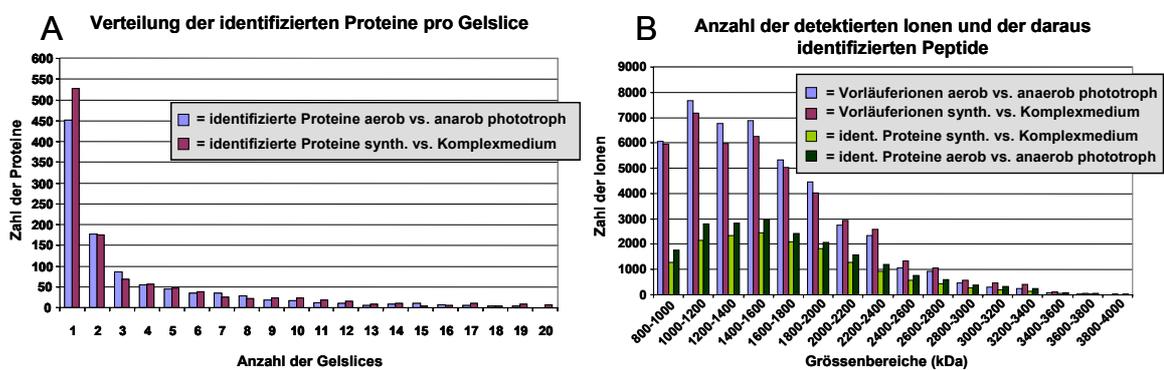
Bei der Betrachtung der identifizierten Proteine pro Gelslice fiel bereits auf, dass ein Großteil in mehreren Gelstücken gefunden wurde.

Abbildung 4-34 A verdeutlicht, dass die Anzahl der Proteine, die in mehreren Fraktionen identifiziert wurden, mit steigender Zahl exponentiell abnimmt und die meisten Proteine in einem oder zwei bis drei Gelslices gefunden wurden. Einige Proteine wurden jedoch in sehr vielen bis allen Gelstücken identifiziert. Bei diesen „schmierenden“ Proteinen stellte sich bei näherer Betrachtung heraus, dass es sich in den beiden Versuchen dabei zum größten Teil um die gleichen handelte. Ungefähr 95%

aller Proteine, die in 10 oder mehr Gelstücken in dem ersten ICPL-Versuch gefunden wurden (Komplexmedium vs. synthetisches Medium), fanden sich auch im zweiten Versuch (aerob vs. anaerob, phototroph) in 5 oder mehr Fraktionen wieder. Es wurden sowohl Proteine in Gelstücken identifiziert, die zu große Massen für den jeweiligen Größenbereich des Gelslices aufwiesen als auch Proteine mit zu kleinen molekularen Massen.

Der erste Fall könnte teilweise durch proteolytische Abbauprodukte hervorgerufen werden, würde allerdings nicht erklären warum ein Protein in allen Gelslices identifiziert wird. Auf den analysierten 2-D Gelen wurden ebenfalls Proteine in mehreren Spots aus unterschiedlichen  $M_r$ -Bereichen des Gels identifiziert (siehe 4.1.3), die sich allerdings auf bestimmte Größenbereiche beschränkten (diskrete Spots). Der zweite Fall, das Auftreten von Proteinen mit einer zu kleinen theoretischen Masse für den jeweiligen Größenbereich, wäre durch eine eingeschränkte Löslichkeit des Proteines oder durch Proteinaggregate zu erklären, die selbst unter den starken denaturierenden Bedingungen des SDS-Auftragpuffers stabil bleiben.

Generell kann man sagen, dass in der SDS-PAGE die Aufgetrennten Proteine nur begrenzt in scharfe Molekularmassenbereiche *fokussiert* werden.



**Abb. 4-34:** Statistiken zu den identifizierten Proteinen und Peptiden. In (A) ist gezeigt, in wie vielen unterschiedlichen Gelstücken die identifizierten Proteine beider Experimente gefunden wurden (blau: synthetisches vs. Komplexmedium; lila: aerob vs. anaerob phototroph). In (B) wird die Anzahl der detektierten Vorläuferionen (blau: synthetisches vs. Komplexmedium; lila: aerob vs. anaerob, phototroph) sowie die sich daraus ergebenden identifizierten Peptide (hellgrün: synthetisches vs. Komplexmedium; dunkelgrün: aerob vs. anaerob phototroph) dargestellt.

Eine weitere interessante Beobachtung war bei der Betrachtung der für die MS/MS Analyse ausgewählten Vorläuferionen in Abhängigkeit von ihrer Molekülmasse und den sich daraus ergebenden Identifizierungen zu machen (siehe Abb. 4-34 B). Es zeigte sich, dass die meisten Vorläufer eine Masse zwischen 1000 und 1600 Da aufwiesen und sich mit zunehmender Masse deren Anzahl stetig verringerte. Ein ähnliches Verhalten war bei den identifizierten Peptiden auszumachen. Dennoch nahm

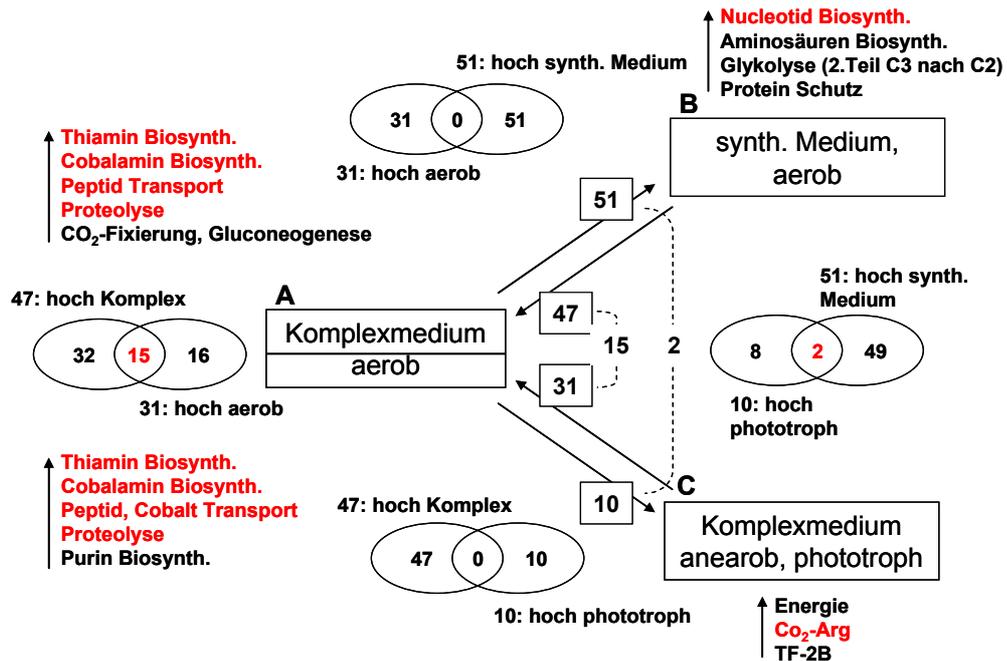
der prozentuale Anteil der aus den Vorläuferionen identifizierten Peptide mit steigender Masse zu. Ebenfalls zu erkennen war, dass in dem 2. ICPL-Versuch (aerob vs. phototroph) jeweils weniger Vorläufer in den unteren Massenbereichen (bis 2000 Da), jedoch mehr in den höheren Massenbereichen für die nachfolgende MS/MS Analyse ausgewählt wurden. Der prozentuale Anteil der identifizierten Peptide war dabei im Vergleich zum ersten Versuch (Komplexmedium vs. synthetisches Medium) bis auf den größten Massenbereich (3800 - 4000 Da) stets um ca. 10% höher, was durch die Einstellung einer höheren Laserintensität in diesem Versuch zu erklären wäre.

Generell kann aus diesen Daten abgeleitet werden, dass ein Peptid mittels MALDI-TOF/TOF umso leichter identifiziert werden kann, desto größer seine molekulare Masse ist.

#### ***4.2.5.2 Gesamtbetrachtung und Interpretation der biologischen Regulationen***

Für die Analysen wurden zwei Mangelbedingungen bei denen zum einen die Nährstoffverfügbarkeit (B: synthetisches Medium, aerob) und zum anderen Variabilität des Redoxzustands limitierend waren (C: Komplexmedium, anaerob, phototroph), mit der Standardwachstumsbedingung (A: Komplexmedium, aerob) verglichen. Da sich beide Versuche auf die gleiche Standardbedingung (A) beziehen (A vs. B bzw. A vs. C), kann auch ein indirekter Vergleich der beiden limitierenden Bedingungen untereinander vorgenommen werden.

Für die Regulation von Proteinen bei den untersuchten unterschiedlichen Wachstumsbedingungen sind generell folgende Erwartungen festzuhalten: Wachstum unter phototrophen Bedingungen sollte Proteine, die an der Energieumwandlungen beteiligt sind, induzieren, wohingegen unter aerobem Wachstum in Redoxvorgänge involvierte Proteine als hoch reguliert gefunden werden sollten. In synthetischem Medium, in dem nur eine begrenzte Anzahl an Nährstoffen vorliegt, sollten dagegen Enzyme der unterschiedlichen Biosynthesewege induziert sein. In Komplexmedium, der Bedingung für optimales Wachstum, wäre zu erwarten, dass die Proteine, welche für eine maximale Wachstumsgeschwindigkeit unabdingbar sind, als hoch reguliert identifiziert werden. Diese Annahmen können nun auf der Grundlage der erhobenen Daten verifiziert werden. Abbildung 4-35 fasst die erhaltenen Regulationen zusammen.



**Abb. 4-35:** Zusammenfassung der Regulationsergebnisse des paarweisen Vergleichs zwischen aerobem Wachstum in Komplexmedium (A) mit aerobem Wachstum in synthetischem Medium (B) und Wachstum unter anaeroben phototrophen Bedingungen (C). Verglichen sind ebenfalls die regulierten Proteine der unabhängigen Experimente (Schnittmengen). Nur die unter Standardbedingungen induzierten Proteine (A) zeigen mehrere gemeinsame (15) beim paarweisen Vergleich mit den Mangelbedingungen (B und C). Die in diesen beiden Mangelbedingungen regulierten Proteine haben nur die beiden Untereinheiten eines Enzyms gemeinsam. Die induzierten Proteine der limitierenden Bedingungen beim Vergleich mit der Standardbedingung zeigten erwartungsgemäß keine Überschneidungen.

Die geringste Anzahl von 10 hoch regulierten Proteinen wurde beim Übergang zu anaerobem phototrophen Wachstum gefunden (A→C). Es handelt sich dabei um Proteine der Energieversorgung, der Transkriptionkontrolle sowie der CO<sub>2</sub>-Fixierung. Beim umgekehrten Übergang (C→A) handelt es sich um insgesamt 31 induzierte Proteine, die hauptsächlich in der Cofaktor Biosynthese, dem Peptidtransport und Abbau sowie der Purin Biosynthese involviert sind. Die 47 hochregulierten Proteine beim Wechsel von synthetischem Medium zu Komplexmedium (B→A) können zu einem großen Teil den eben genannten Klassen zugeordnet werden. Zusätzlich wurden induzierte Proteine der CO<sub>2</sub>-Fixierung und der Gluconeogenese identifiziert. Die 51 regulierten Proteine beim entgegengesetzten Übergang (A→B) hin zu Wachstum in synthetischem Medium haben ihre biologischen Funktionen hauptsächlich in der Nucleotid- und Aminosäurebiosynthese sowie der Proteinprotektion.

Beim Vergleich der drei Zustände untereinander wird nur eine geringe Überschneidung der regulierten Proteine festgestellt. Beim Übergang von phototrophen zu aerobem Wachstum in Komplexmedium und anschließendem Wechsel zu Wachstum unter aeroben Bedingungen in synthetischem Medium (C→A→B) werden insgesamt 82

Proteine hoch reguliert die sich nicht überlappen. Der umgekehrte Wechsel ( $B \rightarrow A \rightarrow C$ ) beinhaltet 57 regulierte ebenfalls nicht redundante Proteine. Vergleicht man die Regulationen die ausgehend von der Standardwachstumsbedingung zu den beiden limitierenden Zuständen ( $A \rightarrow B$  und  $A \rightarrow C$ ) auftreten, findet man lediglich 2 Proteine die bei beiden Übergängen hoch reguliert sind. Dabei handelt es sich um die beiden Untereinheiten der Carbamoylphosphat Synthase. Insgesamt 15 gemeinsam hoch regulierte Proteine finden sich beim umgekehrten Übergang ausgehend vom Wachstum in synthetischem Medium oder unter phototrophen Bedingungen hin zur Standardbedingung ( $B \rightarrow A$  und  $C \rightarrow A$ ). Dieses größte Set an überlappenden Proteinen beinhaltet Enzyme der Thiamin- und Cobalamin-Biosynthese. Diese beiden Cofaktoren erfüllen ihre Aufgaben in Dekarboxylierungsreaktionen (Thiamin) und radiakalinduzierten C-C Bindungsumlagerungen (Cobalamin) wie im Aminosäurestoffwechsel. Weitere Mitglieder der Gruppe, die in beiden Übergängen (hin zur Standardbedingung) gefunden werden, sind Proteine des Peptidtransports und Peptidabbaus. Spezifisch für den Übergang von anaeroben zu aeroben Bedingungen ( $C \rightarrow A$ ) ist die Induktion von Proteinen der Purin Biosynthese, wohingegen beim Wechsel von synthetischem zu Komplexmedium ( $B \rightarrow A$ ) Proteine der  $\text{CO}_2$ -Fixierung und Gluconeogenese spezifisch induziert werden.

Auf eine nähere biologische Interpretation der erhaltenen Ergebnisse soll nun im Folgenden eingegangen werden. Bei dem Übergang zu phototrophen Wachstum, der nur eine überraschenderweise geringe Anzahl an hoch regulierten Proteinen beinhaltet, wurden die beiden Ketten der Carbamoylphosphat Synthase, der Bakterio-opsin Aktivator, BAT, die Phytoen Dehydrogenase sowie ein Transkriptionsfaktor Homolog als induziert gefunden. Des Weiteren wurde ein Heat Shock Protein und ein Rieske Protein Homolog des bc1 Komplexes neben drei weiteren Proteinen, denen keine Homologien zugewiesen waren, als hoch reguliert klassifiziert.

Der hoch regulierte Bakterio-opsin Regulator BAT sowie das TFB Homolog könnten an der Transkriptions Regulation des Bakterio-opsins beteiligt sein, dessen Induktion unter anaeroben, phototrophen Bedingungen unabdingbar für die Energieversorgung der Zelle ist und auch spektroskopisch nachgewiesen werden kann. Da sich die Analyse in dieser Arbeit auf die cytosolischen Proteine beschränkte, konnte Bakteriorhodopsin nicht als reguliert identifiziert werden. Weiter wurde die Phytoene Dehydrogenase als hoch reguliert gefunden, die ein Schlüsselenzym in der Retinal-Biosynthese darstellt. Dieser Chromophor, der dem Retinalprotein Bakteriorhodopsin erst seine biologische Funktion als lichtgetriebene Protonenpumpe ermöglicht, muss erwartungsgemäß ebenfalls in erhöhten Mengen hergestellt werden.

Eine Erklärung für die induzierte Aktivität der Carbamoylphosphat Synthase, die aus Ammoniak (oder Glutamin) und  $\text{CO}_2$  unter dem Verbrauch von 2 ATP Carbamoylphosphat generiert, wäre, dass unter diesen Wachstumsbedingungen die Biosynthese von Arginin verstärkt wird. Da unter phototrophem Wachstum durch die stimulierte Aktivität der Protonenpumpe Bakteriorhodopsin genügend ATP vorliegt, könnte die  $\text{CO}_2$ -Fixierung zu Arginin als langfristiger Energiespeicher dienen, da dieses unter anaeroben Bedingungen im Dunkeln zur Energiegewinnung fermentiert werden kann. Eine plausible Erklärung für die Regulation des Rieske Proteins sowie für das Heat Shock Protein Homolog kann leider nicht gegeben werden.

Beim Wechsel zu Wachstum in synthetischem Medium (A→B) wurde die größte Zahl an regulierten Proteinen gefunden. Diese lassen sich erwartungsgemäß in biosynthetische Reaktionswege einordnen, da nur eine begrenzte Anzahl an Nährstoffen vorliegt und Aminosäuren oder Nukleotide von der Zelle selbst hergestellt werden müssen. Wie beim vorherigen Übergang bereits erwähnt wurden auch hier die beiden Untereinheiten der Carbamoylphosphat Synthase als hoch reguliert gefunden. Dieses Enzym ist jedoch unter Nährstofflimitation eher der Pyrimidin Biosynthese zuzuordnen. Diese Annahme wird dadurch weiter unterstützt, dass zwei weitere Enzyme aus diesem Reaktionsweg als induziert gefunden wurden. In *H. salinarum* kann nicht zwischen den zwei Formen der Carbamoylphosphat Synthase, die entweder Glutamin (Pyrimidine Biosynthese) oder Ammoniak (Arginin Biosynthese) als N-Substrat verwenden, unterschieden werden.

Eine Reihe wichtiger Enzyme der Purin Biosynthese, darunter die Schrittmacherenzyme, wurden als induziert identifiziert, was eine Regulation des Reaktionsweges dieses zweiten Nukleotids klar belegt. Auf die Pyruvate Kinase, einem Schlüsselenzym der Glykolyse, sowie die Proteinprotektion wird im folgenden näher eingegangen.

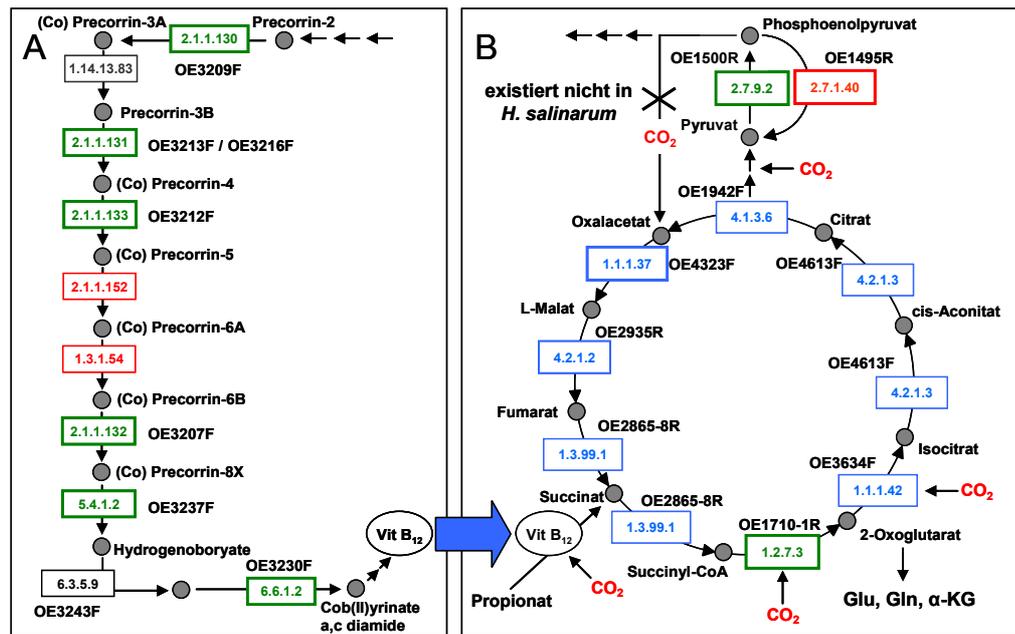
Bei den Übergängen von beiden limitierenden Bedingungen hin zur Standardbedingung (B→A und C→A) wurden 47 bzw. 31 Proteine reguliert, von denen jeweils 15 bei beiden gefunden wurden. Die Übergänge stellen einen Wechsel zu einem vielseitigeren Metabolismus dar, entweder durch die Verfügbarkeit der Energiegewinnung durch die oxidative Phosphorylierung (C→A) oder durch eine nährstoffreiche Umgebung (B→A). Dies entspricht in beiden Fällen besseren Wachstumsbedingungen, die eine schnellere Vermehrung erlauben sollten. Die Proteine, die bei beiden Übergängen gefunden wurden, gehören der Thiamin und Cobalamin Biosynthese an. Zu dem gemeinsamen Set gehören ebenfalls Proteine des Peptidtransports und -Abbaus. In synthetischem Medium (B) wurden interessanterweise Proteine, die bei dem Proteinschutz eine Rolle spielen, gefunden. In den beiden

unterschiedlichen Bedingungen wurden demnach Proteine mit entgegengesetzter Funktion als induziert gefunden.

Spezifisch für den Vergleich gegen synthetisches Medium war die Induktion von Proteinen der CO<sub>2</sub>-Fixierung sowie der Gluconeogenese. Diese Regulationen können wie in Abbildung 4-36 dargestellt mit dem ebenfalls induzierten Cofaktormetabolismus und mit der entgegengesetzten Regulation eines Enzyms der Glykolyse in synthetischem Medium in Verbindung gebracht werden. Die Induktion von 11 Proteinen der Cobalamin Biosynthese bestätigt eindeutig die Regulation dieses Reaktionswegs. Das Endprodukt Cobalamin (Vitamin B<sub>12</sub>) ist u.a. der Cofaktor der Methylmalonyl CoA Mutase, dem Schlüsselenzym der Umwandlung von Propionat in Succinat durch CO<sub>2</sub> Einbau (der Abbau von ungradzahligen Fettsäuren zu Propionat ist in *H. salinarum* möglich). Dieses Produkt kann im Reaktionsweg der reduktiven CO<sub>2</sub>-Fixierung, in dem ebenfalls zwei Enzyme induziert sind ('oxogluterate oxidoreduktase' und die 'puruvate water dikinase'), über Oxalacetat und Puruvat in Phosphoenolpyruvete umgewandelt werden und so als Substrat für die Gluconeogenese dienen. Bei ausreichender Negativität der Redoxpotentiale wäre hier die Möglichkeit eines funktionell umgekehrt laufenden Citratzyklus gegeben.

Das Genom von *H. salinarum* codiert für 6 Ferredoxine. Das fdx-Gen, welches für das Coenzym (fer2, OE4217R) der Pyruvat und Oxogluterat Oxidoreduktase des Citratzyklus codiert, wird konstitutiv expremiert [Kerscher und Oesterhelt, 1976]. Es besitzt ein Redoxpotential von -345 mV [Kerscher *et al.*, 1976] und kann so nur die Hinreaktionen zum Acetyl- und Succinyl-CoA co-katalysieren, da die Eisen-Schwefel-Cluster der zugehörigen Enzyme Redoxpotentialile von etwa -520 mV aufweisen. Diese ferredoxinabhängigen Pyruvat bzw.  $\alpha$ -Ketoglutarat umsetzenden Enzyme wurden in unterschiedlichen Archaea gefunden (Kerscher *et al.*, 1982; Wakabayashi *et al.*, 1983). Im anaerob wachsenden *Thermoplasma acidophilum* wird ein weiteres Ferredoxin expremiert, das einen zusätzlichen 3Fe-4S Komplex aufweist, der ein Redoxpotential bei -520 mV besitzt. Dies ermöglicht einen reduktiven Citratzyklus unter anaeroben Bedingungen. Interessanterweise konnte ein Homolog zu diesem Ferredoxin auch in *H. salinarum* gefunden werden (fer4, OE4496R), was die Vermutung eines funktionierenden reduzierenden Citratzyklus rechtfertigt.

Die von der Pyruvat Kinase katalysierte Rückreaktion vom Phosphoenolpyruvat zum Pyruvat scheint hingegen bei Wachstum in synthetischem Medium bevorzugt zu sein, da hier dieses Enzym als hoch reguliert identifiziert wurde.

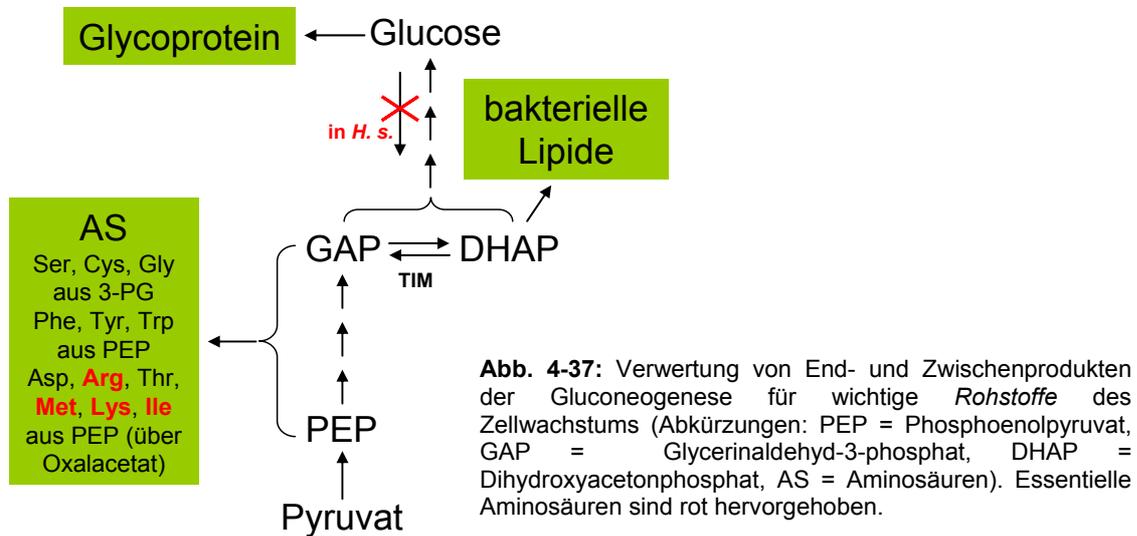


**Abb. 4-36:** Verbindung zwischen Cobalamin Biosynthese und CO<sub>2</sub>-Fixierung. In (A) ist ein Teil der Cobalamin Biosynthese dargestellt. Die Enzyme sind mit ihren E.C. Nummern angegeben. Grüne Enzyme wurden in Komplexmedium als induziert identifiziert, graue Proteine wurden nicht identifiziert und rot markierte Enzyme konnten in *H. salinarum* bisher nicht zugeordnet werden. Für die regulierten Proteine ist zusätzlich der OE-Code angegeben. In (B) ist der CO<sub>2</sub>-Fixierungsweg gezeigt. Für die beteiligten Enzyme sind E.C. Nummern sowie der OE-Code angegeben. Grün markierte Enzyme wurden in Komplexmedium als induziert identifiziert, blaue Enzyme zeigten keine signifikante Regulation. Das rot markierte Enzym wurde in synthetischem Medium als hoch reguliert gefunden. Die Verbindung zwischen Cobalamin Biosynthese und CO<sub>2</sub>-Fixierung ist durch den blauen Pfeil hervorgehoben.

Das bevorzugte Ablaufen der reduktiven CO<sub>2</sub>-Fixierung und die sich daran anschließende Gluconeogenese im Komplexmedium (optimales Wachstum) scheinen auf ersten Blick ebenso widersprüchlich zu sein wie eine Verstärkung der Glycolyse im synthetischen Medium (Mangelbedingung). Für diese Beobachtungen bzw. Interpretationen, die wie eben beschrieben aus den Regulationsdaten der ICPL-Versuche abgeleitet wurden, soll nun eine mögliche Erklärung gegeben werden.

Abbildung 4-37 verdeutlicht, dass aus Zwischen- sowie Endprodukten der Gluconeogenese wichtige Moleküle gebildet werden, die für das Wachstum der Zelle notwendig sind. Im Komplexmedium können die Zellen schnell wachsen und benötigen demnach große Mengen an den Produkten, die die Wachstumsgeschwindigkeit begrenzen würden. So kann das Endprodukt der Gluconeogenese zwar nicht als Energiespeicher in *H. salinarum* dienen, da weder die Biosynthese noch die Elemente der Verstoffwechslung von Glucose existieren, dennoch ist es ein wichtiger Bestandteil des Glycoproteins, das der Organismus zur Aufrechterhaltung seiner Zellstruktur benötigt. Aus Zwischenprodukten wie dem Dihydroxyacetonphosphat (DHAP) können bakterielle Lipide hergestellt werden, die ebenfalls für das schnelle Wachstum der Zelle

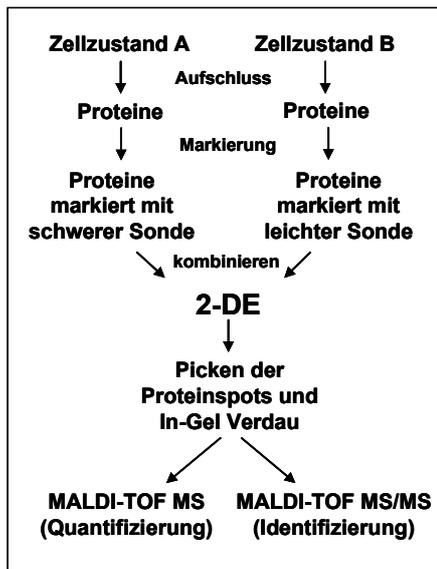
in großen Mengen benötigt werden. Aus weiteren früheren Zwischenprodukten können die unterschiedlichsten Aminosäuren synthetisiert werden.



Es ist schwierig die Hochregulation von glykolytischen Enzymen zu erklären. Die Intermediate der Glykolyse/Gluconeogenese auf der Ebene von C-3 und C-6 bilden das Kohlenstoff-Grundgerüst für viele wichtige Grundbausteine der Zelle. Eine mögliche Erklärung wäre der Bedarf an erhöhten Kapazitäten zum Umbau des Kohlenstoff Gerüsts auf der Ebene von C-3 bis C-6. Die Hochregulation von Enzymen, die am Proteinschutz beteiligt sind könnte andeuten, dass die Zelle unter diesen Lebensbedingungen einen ökonomischen Stoffwechsel benötigt, und mit der eingesparten Energie biosynthetische Reaktionen antreiben kann.

#### 4.2.6 Vergleich mit 2-DE ICPL Daten

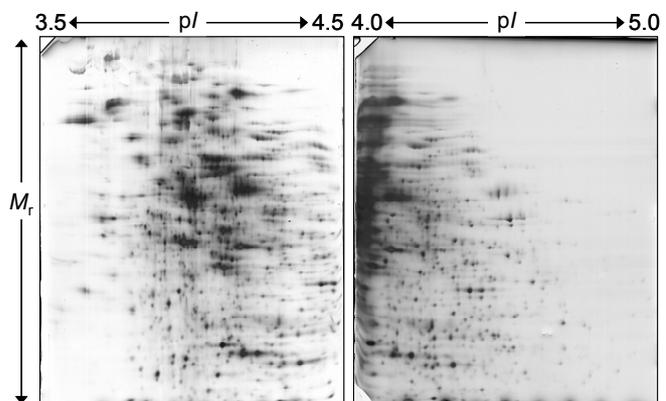
Wie in Abbildung 4-15 dargestellt geschieht die kovalente ICPL-Modifikation an intakten Proteinen, was den Einsatz unterschiedlicher Trennmethode der Proteinprobe vor der massenspektrometrischen Analyse zur Identifizierung differentiell regulierter Proteine erlaubt. Parallel zu den oben beschriebenen ICPL-Analysen mittels 1-D SDS PAGE LC-TOF/TOF wurde in dieser Arbeit ein weiterer methodischer ICPL-Ansatz für beide Vergleichsexperimente (1. Komplexmedium vs. synthetisches Medium, 2. aerob vs. anaerob phototroph) wie in Abbildung 4-38 dargestellt angewendet.



**Abb. 4-38:** Schematische Darstellung des auf 2-DE beruhenden Ansatzes zur ICPL Analyse.

Hierbei werden die markierten Proteine mittels 2-DE aufgetrennt. Die mit der schweren bzw. leichten Sonde markierten Proteine sollten hierbei identisches Laufverhalten aufweisen und so in demselben Proteinspot vorliegen. Diese Proteine wurden anschließend einzeln nach tryptischem In-Gel Verdau analysiert und die Regulationsinformation aus den Isotopenpaaren des PMF-Spektrums gewonnen. Zusätzlich zur PMF-Analyse erfolgte die MS/MS Messung der erhaltenen Isotopenpaare sowie weiterer intensiver Peptidsignale. Abbildung 4-39 zeigt die 2-D Gele der aufgetrennten Proteine im  $pI$ -Bereich zwischen 3.5 und 4.5 sowie zwischen 4.0 und 5.0. Es ist deutlich zu erkennen, dass im Vergleich zur unmodifizierten *H. salinarum* Probe (siehe Abb. 4-5) die Proteine noch weiter in den sauren  $pI$ -Bereich verschoben sind. Fast alle Proteine sind zwischen  $pI$  3.5 und 4.5 zu finden. Dies ist dadurch zu erklären, dass durch die ICPL-Modifikation der Lysine vom Amin zum Amid die Proteine basische Gruppen verlieren. Das Fehlen von Proteinspots auf dem basischeren Gel ( $pI$  4.0-5.0) zeigt jedoch, dass die Modifikation der Proteine vollständig abgelaufen ist.

**Abb. 4-39:** 2-D Gele ICPL-markierter Proteine von *H. salinarum* im  $pI$  Bereich zwischen 3.5 und 4.5 sowie zwischen 4.0 und 5.0.



#### 4.2.6.1 Vergleich mit den regulierten Proteinen

Die Präparation der in den 2-D Gelen aufgetrennten Proteine erfolgte wie bereits beschrieben im Hochdurchsatz-Verfahren. Dazu wurden ca. 800 Proteinspots aus den Gelen ausgeschnitten und wie in „Material und Methoden“ beschrieben prozessiert. Für diese Analysen wurden aus frisch angezogenen *H. salinarum* Kulturen präparierte Proteinproben verwendet. Der Vergleich der mit diesem Ansatz erhaltenen regulierten Proteine mit den Ergebnissen der vorangegangenen Versuche (1-D SDS PAGE LC-MALDI TOF/TOF, kurz LC-ICPL) sollte demnach Rückschlüsse auf die biologische Reproduzierbarkeit der quantitativen Proteomstudien an *H. salinarum* erlauben. Hierfür sollen im Folgenden lediglich die identifizierten regulierten Proteine betrachtet werden.

Die Messung der Proben aus dem Vergleich von Wachstum in Komplexmedium mit dem Wachstum in synthetischem Medium ergab 136 Proteine, von denen Regulationsinformationen erhalten werden konnten. Von ihnen waren lediglich 26 Proteine nach den in 4.2.3 beschriebenen Kriterien als differentiell reguliert anzusehen, wobei 12 von ihnen im Komplexmedium und die restlichen 14 unter Wachstum in synthetischem induziert waren. Beim Vergleich von aerobem mit anaerobem phototrophem Wachstum wurden Regulationsinformationen von 170 Proteinen erhalten. Von ihnen waren lediglich 12 Proteine in aerobem Wachstum induziert. Proteine, die im anaeroben phototrophem Wachstum hoch reguliert waren, konnten nicht gefunden werden.

Die in diesem Versuch identifizierten regulierten Proteine sind in den Tabellen 4-4 A, B und C zusammengefasst. Ebenfalls dort zu finden ist ein Abgleich mit den in den vorangegangenen ICPL-Versuchen erhaltenen Regulationen sowie dort erhaltenen Regulationsfaktoren.

**Tab. 4-4:** Proteine, die als hoch reguliert in Komplexmedium (A), in synthetischem Medium (B) und unter aerobem Wachstum (C) identifiziert wurden. Angegeben sind Gen-Code und Proteinname (1), die Anzahl der identifizierten Isotopenpaare (2), die sich aus ihnen berechnete Regulation (3) sowie die Standardabweichung (4). Weiter zu finden ist ein Vergleich, ob das Protein bereits im vorangegangenen ICPL Versuch als reguliert gefunden wurde (5) sowie die dort erhaltene Regulation (6). Ein gelbes „Nein“ in (5) bedeutet, dass das Protein dennoch einen erhöhten Regulationsfaktor aufwies, der jedoch nicht als reguliert gewertet wurde. Rot-markierte Regulationfaktoren in (6) deuten an, dass das Protein im LC-ICPL Versuch eine höhere Regulation aufwies.

#	A (Gen-Code) Proteinname <sup>1</sup> (hoch in Komplex)	# Paare <sup>2</sup>	Reg 2DE <sup>3</sup>	Reg LC <sup>4</sup>	Vergleich 2DE-LC <sup>5</sup>
1	(OE3214F) precorrin-3B C17-methyltransferase	2	-5.43	-1.47	OK
2	(OE4673F) carboxypeptidase	4	-4.53	-1.85	OK
3	(OE1154F) glyceraldehyde-3-phosphate dehydrogenase	3	-3.73	-1.5	OK
4	(OE4301R) ABC-type transport system ATP-binding protein	1	-3.50	-2.63	OK
5	(OE3296F) hydroxymethylglutaryl-CoA synthase	1	-2.51	-0.68	Nein
6	(OE1500R) pyruvate, water dikinase	2	-2.28	-3.2	OK
7	(OE5208R) arginine deiminase	2	-1.33	-1.71	OK
8	(OE5206R) carbamate kinase	5	-1.10	-0.25	Nein
9	(OE2871F) probable 3-hydroxyacyl-CoA dehydrogenase	7	-1.03	-0.37	Nein
10	(OE4735R) ribosomal protein S7	3	-1.02	-0.33	Nein
11	(OE4703R) UDP-N-acetylglucosamine 2-epimerase	2	-0.99	-0.01	Nein
12	(OE3565F) molybdenum cofactor biosynthesis protein B	1	-0.87	0.07	Nein

#	B (Gen-Code) Proteinname <sup>1</sup> (hoch in synthetisch)	# Paare <sup>2</sup>	Reg 2DE <sup>3</sup>	Reg LC <sup>4</sup>	Vergleich 2DE-LC <sup>5</sup>
1	(OE3554F) carbamoyl-phosphate synthase	1	7.80	12.53	OK
2	(OE1951F) phosphoribosylaminoimidazole carboxylase	2	5.09	5.41	OK
3	(OE3038F) methylenetetrahydrofolate dehydrogenase	1	3.96	3.5	OK
4	(OE1620R) phosphoribosylglycinamide formyltransferase	3	3.42	2.05	OK
5	(OE4408F) phosphoglycerate dehydrogenase	3	2.76	0.86	OK
6	(OE3724F) phosphoribosylaminoimidazolesuccinocarboxamide	1	2.04	1.13	OK
7	(OE4405R) phosphoserine phosphatase	3	1.61	0.32	Nein
8	(OE1495R) pyruvate kinase	2	1.48	2.05	OK
9	(OE4159F) adenosylhomocysteinase	9	1.35	0.3	Nein
10	(OE4391F) probable serine--pyruvate aminotransferase	2	1.13	0.67	Nein
11	(OE3580R) threonine--tRNA ligase	7	0.85	0.03	Nein
12	(OE2864F) phosphoribosylamine--glycine ligase	1	0.85	2.89	OK
13	(OE1584R) conserved hypothetical protein	6	0.84	-0.23	Nein
14	(OE2015R) 3-hydroxybutyryl-CoA dehydrogenase	1	0.83	0.21	Nein

#	C (Gen-Code) Proteinname <sup>1</sup> (hoch in aerob)	# Paare <sup>2</sup>	Reg 2DE <sup>3</sup>	Reg LC <sup>4</sup>	Vergleich 2DE-LC <sup>5</sup>
1	(OE4651F) thiamine biosynthesis protein	1	-3.50	-2.9	OK
2	(OE3207F) probable precorrin-8W decarboxylase	3	-2.73	-1.5	OK
3	(OE4654F) phosphomethylpyrimidine kinase	1	-2.48	-2.1	OK
4	(OE3216F) precorrin-3B C17-methyltransferase	1	-2.43	-1.44	OK
5	(OE3212F) precorrin-4 C11-methyltransferase	2	-2.32	-1.16	OK
6	(OE3209F) probable precorrin-2 C20-methyltransferase	7	-2.29	-1.99	OK
7	(OE3213F) CbiG protein (unassigned function)	3	-1.94	-2.18	OK
8	(OE3214F) precorrin-3B C17-methyltransferase	2	-1.66	-1.23	OK
9	(OE1613R) probable acylaminoacyl-peptidase	1	-1.50	-0.96	OK
10	(OE1584R) conserved hypothetical protein	15	-1.32	-0.46	Nein
11	(OE4673F) carboxypeptidase	1	-1.05	-0.95	OK
12	(OE2013R) probable acyl/butyryl-CoA dehydrogenase	2	-0.95	-0.7	Nein

Trotz der eher geringen Anzahl an identifizierten Proteinen konnte für einen großen Teil der Proteine bzw. Reaktionswege, auf dessen Regulationsdaten die Beschreibung der Übergänge zwischen den untersuchten Lebensbedingungen im vorherigen LC-ICPL Versuch basiert, auch hier die entsprechenden Regulationen gefunden werden.

So waren beispielsweise im Komplexmedium (Tabelle 4-4 A, Vergleich zu synthetischem Medium) und im aeroben Wachstum (Tabelle 4-4 C, Vergleich zu anaerobem, phototrophem Wachstum) die bereits beschriebenen Proteine der Thiamin und Cobalamin Biosynthese sowie die des Peptidtransports und -Abbaus als induziert gefunden. Im Komplexmedium wurde wiederum die 'pyruvat water dikinase' als induziert identifiziert. Zusätzlich wurden hier noch 5 Proteine gefunden, die im vorangegangenen Versuch als nicht reguliert klassifiziert worden waren. Diese Proteine ergänzten weder die zuvor getroffene Interpretation der Regulationsdaten, noch wurden sie in Frage gestellt. Lediglich die Carbamat Kinase (OE5206R) wird im 2-DE Experiment reguliert gefunden, und zwar gegenläufig zur Carbamoylphosphat Synthase. Dies war auch zu erwarten, da die beiden Enzyme gegenläufige Reaktionen katalysieren, die einen *futile cycle* bilden könnte.

Im synthetischen Medium waren erneut Proteine der Nukleotid sowie Aminosäure Biosynthese hoch reguliert. Bei letzterer wurden zusätzlich Proteine gefunden, die zuvor als nicht reguliert klassifiziert waren. Ebenfalls wurde die Pyruvat Kinase als induziert gefunden, was sich wiederum mit den bereits erhobenen Ergebnissen deckte.

Grundsätzlich konnte mit dem Vergleich der Daten aus zwei unabhängigen Experimenten gezeigt werden, dass sich die Tendenzen der erhaltenen Regulationen für den überwiegenden Teil der Proteine entsprachen. In einigen Fällen unterschieden sich die Werte der Regulationen stärker (siehe z.B. OE3214F und OE4673F in Tabelle 4-4 A [Regulationen 2-DE: -5.43 bzw. -4.53, Regulationen LC-ICPL: -1.47, -1.85] oder OE3554F in Tabelle 4-4 B [Regulationen 2-DE: 7.8, Regulationen LC-ICPL: 12.53]). Es wurden jedoch keine Proteine gefunden, die in beiden experimentellen Ansätzen (2-DE-ICPL, 1-DE LC-ICPL) entgegengesetzt reguliert waren.

Es zeigte sich aber auch, dass in den 2-DE Versuchen Proteine hoch reguliert waren, die in den vorangegangenen Analysen eindeutig als nicht reguliert klassifiziert werden konnten. Diese beobachteten Regulationen können eventuell durch Probenpräparationsartefakte oder durch nicht identische *Sekundärstimuli* in beiden Versuchsdurchführungen erklärt werden. Derartige Probenpräparationsartefakte oder nicht identische *Sekundärstimuli* könnten ebenfalls unterschiedlich stark ausfallende Regulationen eines Proteins in beiden Versuchsdurchführungen erklären.

Die Tatsache unterschiedlicher Adaptationszustände an die gewählten Wachstumsbedingungen sollte auszuschließen sein, da identische Anzuchtbedingungen, Wachstumszyklen und Erntezeiten in beiden Versuchsreihen vorlagen.

#### 4.2.7 Analyse mittels LC-ESI-MS/MS

In einer Kooperation mit der Abteilung von Prof. Matthias Mann wurden äquivalente ICPL markierte Proben des Vergleichs von Wachstum in Komplexmedium mit Wachstum in synthetischem Medium (siehe 4.2.3) mit einem LTQ FT-ICR Massenspektrometer vermessen. Bei den Proben handelte es sich um tryptische Peptide aus 19 Gelslices eines 1-D SDS Gels. Die so erhobenen Daten werden im Rahmen dieser Arbeit nicht mehr in Bezug auf die Quantifizierung mittels ICPL beschrieben, sondern dienen lediglich der Vervollständigung des cytosolischen Proteininventars von *H. salinarum*.

Im Vergleich mit den bisherigen verwendeten Methoden können Proteine mit dem hier verwendeten Massenspektrometer anhand der aufgenommenen Tandem Massenspektren zugehöriger proteolytischer Peptide mit einer höheren Sicherheit identifiziert werden. Zum einen werden die Peptidmassen mit einer sehr hohen Genauigkeit bestimmt, was die Anzahl der möglichen Peptide aus der verwendeten Datenbank, die für die Recherche in Frage kommen herabsetzt. Zum anderen wird aus dem MS/MS-Spektrum zusätzlich das intensivste Signal isoliert und erneut fragmentiert. Die aus diesem MS<sup>3</sup> Spektrum erhaltenen Fragmentmassen gehen zusätzlich in die Datenbank Recherche ein und verifizieren so das Ergebnis.

Für die nachfolgende Betrachtung wurden Proteine mit einem Mascot Score  $\geq 25$  als identifizierte Proteine angesehen. Mit dieser Scoregrenze konnten 1414 Proteine (68 mit vorhergesagten Transmembranregionen) von *H. salinarum* identifiziert werden, davon 876 mit Isotopenpaaren. Im Vergleich dazu wurden mit dem LC-MALDI TOF/TOF Verfahren 1008 Proteine aus einer äquivalenten Probe identifiziert, davon 803 mit Isotopenpaaren (siehe 4.2.3.2 und 4.2.3.5). Bei dem Vergleich der Proteine mit Regulationsinformationen mit den LC-MALDI TOF/TOF Analysen (*off-line* Verfahren) ist zusätzlich zu berücksichtigen, dass bei der LC-ESI Messung mit dem LTQ FT-ICR Massenspektrometer keine gesonderte Auswahl von Isotopenpaaren für die nachfolgende Identifizierung mittels MS/MS getroffen wurde (*on-line* Verfahren). Die Anzahl an identifizierten Proteinen mit Regulationsinformationen sollte demnach noch zu steigern sein, wenn eine derartige datenabhängige Auswahl der Vorläuferionen wie bei den LC-MALDI TOF/TOF Analysen genützt wird.

### 4.3 Posttranslationale Modifikationen

Bei der Analyse posttranslationaler Modifikationen stand die Frage im Mittelpunkt, inwieweit Proteinphosphorylierungen in *Halobacterium salinarum* neben den bekannten, in der Signalweiterleitung der Zwei-Komponenten Systeme auftretenden Modifikationen eine regulative Rolle spielen. Neben diesen bereits erwähnten Histidin- und Aspartyl-Phosphorylierungen ist für *H. salinarum* nicht bekannt, ob und in welchem Ausmaß Serin-, Threonin- oder sogar Tyrosin-Phosphorylierungen als regulative posttranslationale Modifikationen auftreten. In dem Genom von *H. salinarum* sind 6 Proteine mit Homologien zu Protein Kinasen codiert (siehe Tabelle 4-5).

**Tab. 4-5:** Im Genom von *H. salinarum* annotierte Protein Kinasen.

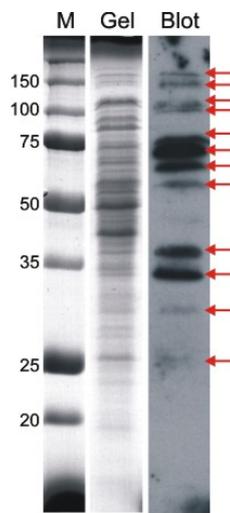
<b>(Gen-Code) Proteinname</b>	<b>MS-Status</b>
(OE2103R) protein kinase prkA homolog	identifiziert
(OE2104R) probable protein kinase prkA (serine kinase)	identifiziert
(OE2354R) protein kinase distant homolog	identifiziert
(OE5071F) protein kinase weak homolog	nicht identifiziert
(OE5188F) protein kinase weak homolog	identifiziert
(OE6278R) protein kinase afsK homolog	identifiziert

Dies war als erster Hinweis zu werten, dass Ser-, Thr- oder Tyrosin-Phosphorylierungen in *H. salinarum* auftreten könnten. Aufgrund der fundamentalen Bedeutung der Proteinphosphorylierung steht heute eine Vielzahl an unterschiedlichen Analyseverfahren von phosphorylierten Proteinen zur Verfügung. Die in dieser Arbeit angewendeten Methoden und die daraus resultierenden Ergebnisse werden im Folgenden vorgestellt.

#### 4.3.1 Immunochemischer Nachweis von Phosphoproteinen

Für den immunochemischen Nachweis phosphorylierter Proteine stehen unterschiedliche Antikörper zur Verfügung, welche spezifisch die phosphorylierten Aminosäuren erkennen. Bei phosphorylierten Threoninen und insbesondere Serinen sind jedoch die zum phosphorylierten Thr/Ser benachbarten Aminosäuren ebenfalls für die Antigen-Erkennung durch den Antikörper von Bedeutung. Da im Falle von *H. salinarum* nichts über eventuelle Phosphorylierungsepitope bekannt ist, wie z.B. bei höheren Organismen, wurde für die immunochemische Analyse der cytosolischen Proteine eine Mischung aus verschiedenen Anti-Phospho-Ser und Anti-Phospho-Thr Antikörpern verwendet, um möglichst viele unterschiedliche Phosphorylierungsstellen detektieren zu können.

Abbildung 4-40 zeigt die mittels 1-D SDS PAGE aufgetrennten cytosolischen Proteine sowie den Blot eines äquivalenten Gels. Es sind deutlich mehrere diskrete

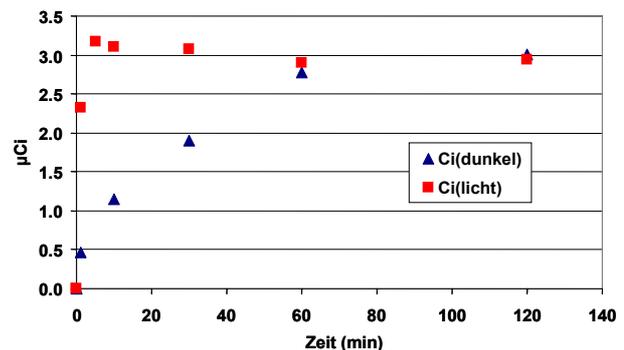


Signale bei ca. 150 kDa, 100 kDa, 55-75 kDa, 30 kDa und schwache Signale bei ca. 20 und 25 kDa zu erkennen. Die Antikörper scheinen also phosphorylierte Proteine erkannt zu haben, was ein weiterer Hinweis für das Vorkommen von Ser/Thr Phosphorylierungen in *H. salinarum* ist. Das Auflösungsvermögen der 1-D SDS-PAGE reicht jedoch nicht aus, den Signalen eine Proteinbande im Gel zuzuweisen, um diese anschließend zu identifizieren. Dennoch konnte gezeigt werden, dass die Phospho-Ser/Thr spezifischen Antikörper Proteine von *H. salinarum* erkennen und spezifisch binden.

**Abb. 4-40:** 1-D Gel sowie Anti-phospho Ser/Thr Blot cytosolischer Proteine.

### 4.3.2 Radioaktive *in-vivo* Markierung der cytosolischen Proteine

Die metabolische Markierung der Proteine von *H. salinarum* und deren Auftrennung mittels 2-DE sollte nun eine bessere Übersicht über das Ausmaß der Proteinphosphorylierung liefern. Zunächst wurde die Phosphataufnahme von *H. salinarum* untersucht, da das Einbringen des radioaktiven Phosphats in die Zellen die Voraussetzung für die metabolische

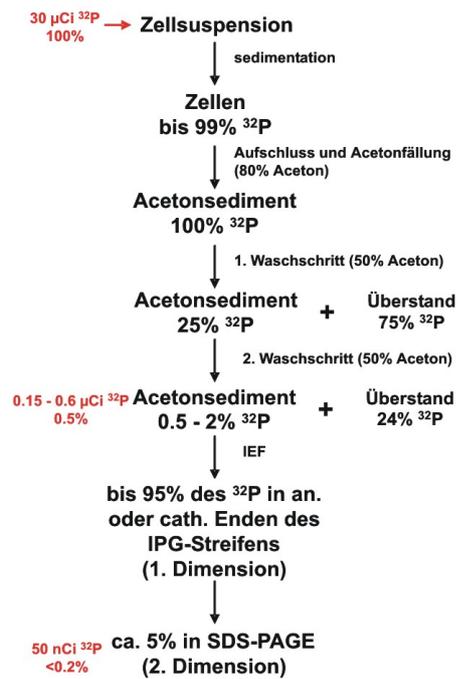


**Abb. 4-41:** Zeitabhängige Phosphataufnahme im Dunkeln und unter Belichtung der Zellen.

Markierung darstellt. Dazu wurden Zellen bei ca. 40 Klett 5fach in BS aufkonzentriert und zu der stetig gerührten und bei 37°C temperierten Suspension 3 µCi <sup>32</sup>Phosphat zugegeben. Nach regelmäßigen Abständen wurden Proben entnommen, die Zellen sedimentiert und die aufgenommene Radioaktivität im Szintillationszähler bestimmt. Der Versuch wurde im Dunkeln und unter Belichtung durchgeführt.

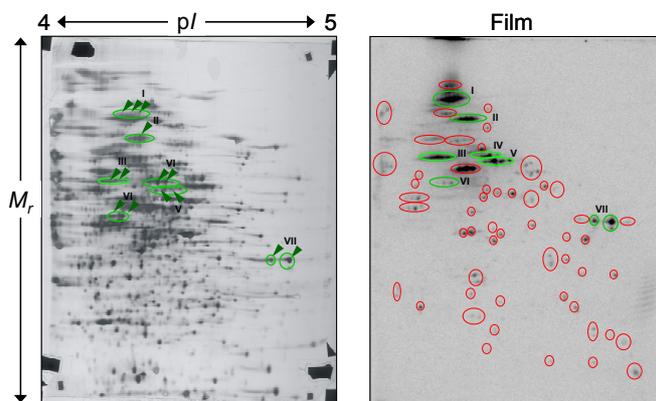
Im Dunkeln hatten die Zellen nach ca. einer Stunde das gesamte zugegebene radioaktive Phosphat aufgenommen, wohingegen es dafür im Licht nur weniger Minuten bedurfte (siehe Abb. 4-41). Die Aufnahme des gesamten zugegebenen radioaktiven Phosphats stellte so eine optimale Voraussetzung für die metabolische Markierung der Proteine dar.

Im nächsten Schritt wurden nun Zellen mit  $30 \mu\text{Ci } ^{32}\text{P}$  pro ml Zellsuspension für 5 min im Dunkeln inkubiert, die Proteine nach Standardbedingungen präpariert und mittels 2-DE aufgetrennt. Abbildung 4-42 zeigt den Verbleib der eingesetzten Radioaktivität während der Aufarbeitung der Proteine. Wie bereits gezeigt hatten die Zellen das gesamte radioaktive Phosphat aufgenommen. Nach Aufschluss der Zellen, Sedimentation von Membranbestandteilen und Acetonfällung der Proteine waren immer noch 100% der aufgenommenen Radioaktivität im Sediment zu finden. Gut  $\frac{3}{4}$  der Radioaktivität ging jedoch beim ersten Waschschriff mit 50% Aceton verloren. Nach dem zweiten Waschschriff waren nur noch 0.5 - 2% der eingesetzten Radioaktivität im Proteinpellet zu finden, welches nun für die 2-DE eingesetzt wurde. Hierbei stellte sich heraus, dass von der in den IPG-Streifen eingebrachten Radioaktivität der größte Teil in dem kathodischen und vor allem in dem anodischen Ende nach Bedingung der Fokussierung zu finden war. In die 2. Dimension, der SDS PAGE, gingen so nur 50 nCi ein, was weniger als 0.2% der eingesetzten Radioaktivität entsprach.



**Abb. 4-42:** Verbleib der eingesetzten Radioaktivität während der Aufarbeitung und Auftrennung der Proteine.

Die aufgetrennten Proteine wurden mittels Ag-Färbung visualisiert, die Gele zwischen Celofanfolie getrocknet und für mehrere Tage bei  $-80\text{C}$  auf einen Röntgenfilm aufgelegt. Abbildung 4-43 zeigt ein *radioaktives* Gel im *pI*-Bereich zwischen 4 und 5



**Abb. 4-43:** Zweidimensionales Gel radioaktiv markierter Proteine. Die radioaktiven Signale auf dem Film (rechts), die auf dem Gel (links) Proteinspots zugewiesen und identifiziert werden konnten, sind grün umkreist und nummeriert (I-VII). Signale, die nicht eindeutig zugewiesen werden konnten, sind auf dem Film rot umkreist.

(links) und einen entsprechenden entwickelten Film (rechts). Auf dem Film sind deutlich mehrere Signale unterschiedlicher Intensität zu erkennen (umkreiste Flecken). Durch Anbringen von fluoreszierenden Markern auf dem getrockneten Gel, welche nach Entwicklung des Films entsprechende Signale hinterließen, konnte der Film mit dem getrockneten Gel zur

Deckung gebracht und die radioaktiven Signale mit dem Proteinmuster auf dem 2-D Gel abgeglichen werden. Die *radioaktiven* Proteine wurden anschließend aus korrespondierenden, nicht-radioaktiven Gelen identifiziert oder konnten durch Vergleich der Spotmuster aus den angefertigten Referenzgelen entnommen werden. Es konnten insgesamt 7 Proteine (grüne Kreise) als radioaktives Phosphat tragend identifiziert werden (siehe Tabelle 4-6).

**Tab. 4-6:** Als phosphattragend identifizierte Proteine.

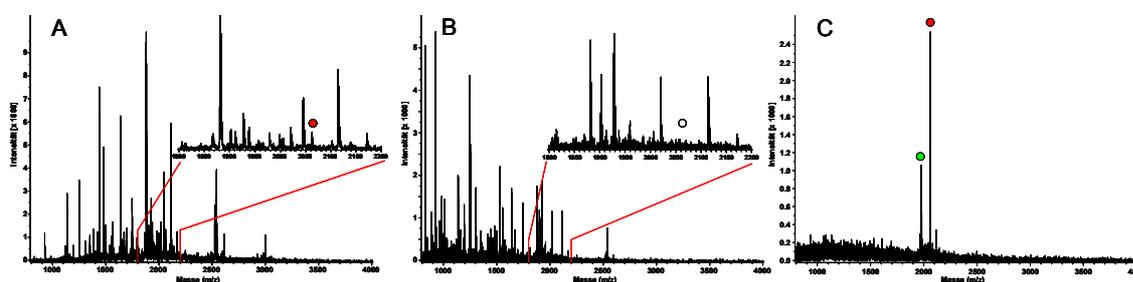
Nr.	Code	Protein Name
I	OE1500R	pyruvate water dikinase
II	OE2623R	pyruvate synthase (EC 1.2.7.1) alpha chain
III	OE3762R	glycerol kinase
IV	OE4169F	conserved protein
V	OE2318R	phosphomannomutase
VI	OE2373F	propable phosphate acyltransferase
VII	OE3524F	pyridoxine biosynthesis protein

Bei den identifizierten Phosphoproteinen handelt es sich um 2 Kinasen sowie weitere phosphatgruppenübertragende Enzyme. In keinem der Massenspektren konnte jedoch das phosphattragende Peptid identifiziert werden. Es war also nicht möglich, die Phosphorylierungsstelle, bzw. die Art der Phosphorylierung zu bestimmen. Neben der beschriebenen Übertragung von Phosphat als Signalgeber kann dieses auch als metabolisches Intermediat oder in Form anderer Kofaktoren an Proteine gebunden sein. Um genauere Aussagen über die Art der nachgewiesenen Phosphorylierungen treffen zu können, bedurfte es einer genauen Charakterisierung der Phosphorylierungsstelle.

### 4.3.3 Analyse der Phosphorylierungsstellen

Um die Phosphorylierungsstelle eines modifizierten Proteins mittels MS bzw. MS/MS zu bestimmen, bedarf es einer Anreicherung der phosphorylierten Peptide. Hierfür wurde die IMAC-Methode (*immobilized metal affinity chromatography*) gewählt, mit der phosphorylierte Peptide angereichert werden können. Diese Methode wurde zunächst mit tryptischen Peptiden von  $\beta$ -Casein etabliert, da dieses ein einfach phosphoryliertes Peptid mit der Masse 2061.83 Da enthält, welches gut mit gängigen massenenspektrometrischen Verfahren wie der MALDI-TOF-MS detektiert werden kann. Zu dem  $\beta$ -Casein Verdau wurden zusätzlich tryptische BSA Peptide im äquimolaren Verhältnis hinzugemischt, um den Aufreinigungseffekt des Phosphopeptids mittels IMAC aus einer komplexen Peptidmischung zu simulieren. Abbildung 4-44 A zeigt das aufgenommene MALDI-TOF Massenspektrum des Gemisches aus je 200 fmol tryptischer BSA- und  $\beta$ -Casein Peptide sowie einen

vergrößerten Ausschnitt, in dem das phosphorylierte  $\beta$ -Casein Peptid (roter Kreis) zu sehen ist (roter Kreis). In Abbildung 4-44 B ist das Massenspektrum der nicht gebundenen Peptide (Durchfluß und Waschfraktion) zu sehen. In dem vergrößerten Bereich ist kein Signal, welches dem Phosphopeptid entsprechen würde (offener Kreis), zu erkennen. Dies zeigt, dass das  $\beta$ -Casein Phosphopeptid komplett von der IMAC-Säule gebunden wurde. Abbildung 4-44 C zeigt das Massenspektrum des eluierten Phosphopeptids (roter Kreis). Es sind neben dem Phosphopeptid keine weiteren Signale zu erkennen, was zeigt, dass es zu keinen unspezifischen Bindungen an das IMAC Material gekommen ist. Der Peak bei 1978.35 Da (grüner Kreis) entspricht dem durch den Ionisationsprozeß dephosphorylierten Phosphopeptid.



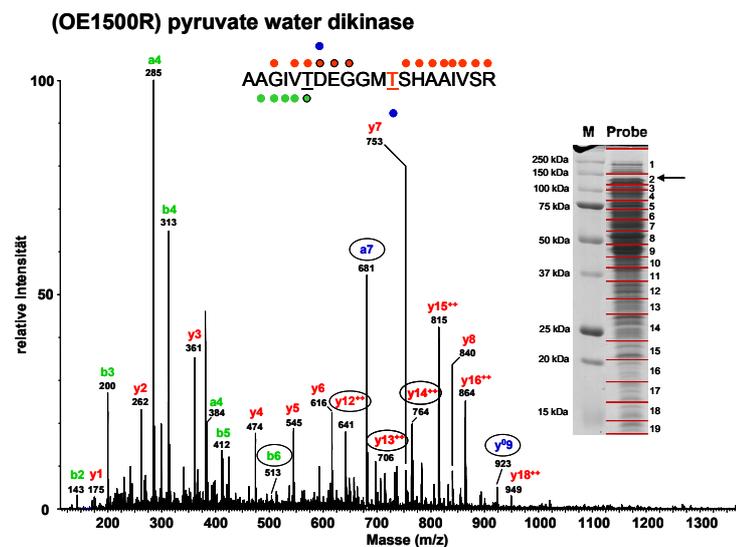
**Abb. 4-44:** MALDI-Spektren unterschiedlicher Stufen der Anreicherung eines Phosphopeptids mittels IMAC. (A) Gemisch aus tryptischen BSA und  $\beta$ -Casein Peptiden. Der Massenbereich, in dem das phosphorylierte  $\beta$ -Casein Peptid (roter Punkt) liegt, ist vergrößert dargestellt. (B) Massenspektrum der Peptide, die nicht an die IMAC Säule gebunden haben. In dem vergrößerten Massenbereich ist kein Signal dem phosphorylierten  $\beta$ -Casein Peptid (weißer Punkt) zuzuordnen. (C) IMAC Eluat. Als einzige Signale sind das phosphorylierte  $\beta$ -Casein Peptid (roter Punkt) sowie das Peptid, welches während des MALDI-Ionisationsprozesses die Phosphatgruppe verloren hat (grüner Punkt), zu sehen.

Die IMAC-Methode ist demnach in der Lage, spezifisch Phosphopeptide aus Peptidmischungen zu isolieren. Es ist jedoch bekannt, dass es bei steigender Komplexität der Mischung zu unspezifischen Bindungen kommen kann. Gerade saure Peptide, die vermehrt negative Ladungen tragen, können eine verstärkte Affinität zum IMAC-Material zeigen und so mit den Phosphopeptiden koeluiieren. Dieses Problem sollte demnach vermehrt bei halobakteriellen Proteinen auftreten, da diese einen überdurchschnittlich hohen Anteil an sauren Aminosäuren besitzen. Um die möglichen Kontamination zu minimieren und die Komplexität der für die IMAC Aufreinigung eingesetzten Peptidprobe möglichst klein zu halten, wurden unterschiedliche Vorfraktionierungen auf Protein oder Peptidebene eingesetzt.

#### 4.3.3.1 Vorfraktionierung mittels 1-DE

Zunächst wurden die cytosolischen Proteine über 1-D SDS PAGE aufgetrennt, das Gel in 19 Slices geschnitten und die Proteine im Gel proteolytisch mit Trypsin gespalten. Die daraus entstandenen Peptidfraktionen wurden der IMAC unterzogen und

mittels LC-ESI-MS/MS (Q-TOF, siehe „Material und Methoden“ 3.2.11.7) analysiert. Für die anschließende Mascot-Recherche wurde als variable Modifikation die Phosphorylierung von Serin und Threonin Resten erlaubt. In Slice 2 identifizierte Mascot ein phosphoryliertes Peptid, welches der ‘pyruvat water dikinase’ (OE1500R) zugeordnet werden konnte. Dieses Protein war bereits nach radioaktiver metaboler Markierung als phosphattragend auf 2-D Gelen identifiziert worden (siehe 4.3.2). Abbildung 4-45 zeigt das aufgenommene Fragmentspektrum des für die MS/MS Messung ausgewählten Peptidions sowie die Sequenzannotation des phosphorylierten Peptids.



**Abb. 4-45:** Tandemspektrum eines Phosphopeptides der ‚pyruvate water dikinase‘. Die zugeordneten y-, b- und a-Fragmentionen (rot bzw. grün) sind in den jeweiligen Spektren als auch über (y-Ionen, rot) bzw. unter (a- und b-Ionen, grün) der zugehörigen Peptidsequenz angegeben. Das phosphorylierte Threonin-12 ist rot hervorgehoben und unterstrichen. Die Ionen, die Threonin-12 als phosphorylierte Aminosäure belegen, sind in Spektrum umkreist bzw. an der Sequenz schwarz umrandet. Die Ionen, die für Threonin-6 (unterstrichen) als phosphorylierte Aminosäure sprechen sind im Spektrum blau und an der Sequenz als blaue Punkte dargestellt. Ebenfalls gezeigt ist das 1-D Gel. Das Gelslice, aus dem das aufgenommene Spektrum stammt, ist markiert (Pfeil).

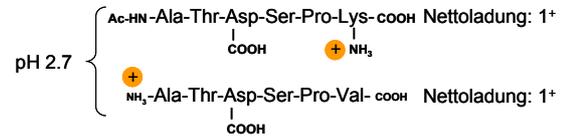
Aus der Sequenzannotation konnte Threonin 12 als phosphorylierte Aminosäure abgeleitet werden. Die Zuweisung des phosphorylierten Threonins beruht jedoch auf nur vier Fragmentionen, von denen einige nur eine geringe Intensität aufweisen (umkreist in Abb. 4-45). Ebenfalls können zwei im Spektrum auftretende Signale (blau in Abb. 4-45) dem Peptid, welches die Phosphatgruppe am Threonin-6 tragen würde, zugeordnet werden. Eine eindeutige Zuweisung der Phosphorylierungsstelle ist letztendlich nicht möglich, obwohl mehr Signale für die Position 12 sprechen.

In den massenspektrometrischen Analysen der restlichen Gelslices (siehe Abb. 4-45, Einsatz) konnten keine weiteren Phosphorylierungsstellen identifiziert werden. Aus dem immunodetektierten 1-D Gel (siehe 4.3.1) sowie aus den radioaktiven 2-D



Das starke Signal in Fraktion 1 (Durchlauf) ist auf den in der Probe befindlichen Harnstoff zurückzuführen. Beim langsamen Anstieg des Hochsalzeluenten bis 15 % B sind deutliche Signale zu erkennen. In diesen Fraktionen (2-6) sollten sich die möglichen Phosphopeptide befinden.

Es zeigte sich jedoch, dass diese frühen Fraktionen immer noch viele nicht phosphorylierte Peptide enthielten. Dabei handelte es sich überwiegend um zwei Arten von Peptiden: N-terminal acetylierte



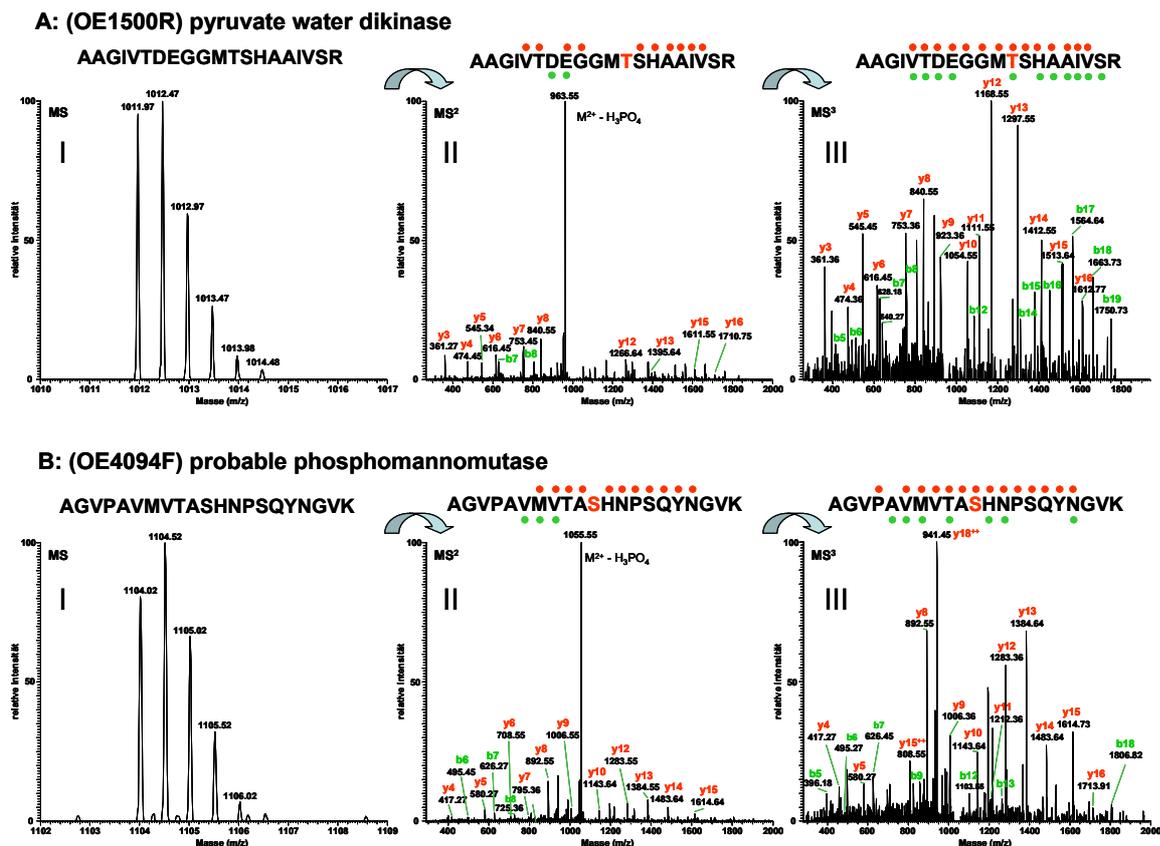
**Abb. 4-48:** Nettoladung tryptischer N-terminal acetylierter und C-terminaler Peptide bei pH 2.7.

sowie C-terminale Peptide, die ebenfalls wie Phosphopeptide weniger positive Ladungen tragen (siehe Abb. 4-48). Für die C-terminalen Peptide trifft dies allerdings nur zu, wenn das Protein nicht mit einem Arginin oder Lysin endet. Um aus diesen Fraktionen die in ihnen möglicherweise enthaltenen Phosphopeptide weiter anzureichern, wurde die kürzlich vorgestellte TiO<sub>2</sub>-Methode [Pinkse *et al.*, 2004] angewendet. Diese beruht ebenfalls wie die IMAC-Methode auf der affinitätschromatographischen Anreicherung, soll jedoch im Vergleich selektiver für Phosphopeptide sein, wenn die TiO<sub>2</sub>-Säule vor der Elution der gebundenen Peptide mit DHB-Lösung gewaschen wird [Larsen *et al.*, 2005]. Die massenspektrometrische Analyse der angereicherten Peptide erfolgte diesmal in Kooperation mit der Abteilung von Prof. Matthias Mann mit einem LTQ FT-ICR Massenspektrometer. Dieses Gerät besteht aus zwei Massenanalysatoren, einer linearen Ionenfalle (LTQ) sowie dem Fourier-Transform Ionenzyklotron-Resonanz-Analysator (FT-ICR), die unabhängig voneinander operieren können.

In vielen Tandem Massenspektren (MS/MS) von Phosphoserin- sowie Phosphothreonin Peptiden ist zu beobachten, dass ein Großteil des phosphattragenden Vorläuferions einer  $\beta$ -Eliminierung des neutralen Phosphats (H<sub>3</sub>PO<sub>4</sub>) unterliegt (*Neutral Loss*) [Huddleston *et al.*, 1993; Schlosser *et al.*, 2001], da die Spaltung des Phosphoesters energetisch günstiger ist als die einer Peptidbindung. Das sich daraus ergebene Fragmentspektrum ist deshalb häufig schwierig zu interpretieren, da es als einziges dominantes Signal oftmals nur das vom Phosphat befreite Peptidsignal enthält (MH<sup>n+</sup> - H<sub>3</sub>PO<sub>4</sub>). Mit geringeren Intensitäten tritt häufig ein Gemisch aus Fragmentionen des Phosphopeptids sowie des *Neutral Loss* Peptids auf, was eine eindeutige Identifizierung und insbesondere die Zuweisung der Phosphorylierungsstelle erschwert. Ebenfalls werden durch die verwendeten Anreicherungsstrategien Phosphoproteine oftmals nur durch ein einziges Peptid, das Phosphopeptid, identifiziert, was die Sicherheit der Identifizierung zusätzlich herabsetzt.

Die Verwendung des LTQ FT-ICR Massenspektrometers ermöglicht in zweierlei Weise eine zuverlässigere Identifizierung von Phosphopeptiden: Zum einen trägt die durch den FT-ICR Massenanalysator erreichte hohe Massengenauigkeit zu einer sichereren Identifizierung bei, so dass bei der Datenbankrecherche weniger theoretische Fragmentspektren zum Vergleich herangezogen werden müssen. Zum anderen ermöglicht die LTQ-Ionenfalle des Instrument aufeinanderfolgende Stufen der Fragmentierung: In der Ionenfalle kann das im ersten Fragmentierungsschritt (MS/MS) entstandene *Neutral Loss* Ion ( $MH^{n+} - H_3PO_4$ ) isoliert und erneut fragmentiert werden (MS/MS/MS oder  $MS^3$ ) [Beausoleil *et al.*, 2004]. Dem durch den *Neutral Loss* entstandenen Dehydroalanin (ehemalig Phosphoserin) und der 2-Aminodehydrobutansäure (ehemalig Phosphothreonin) können dann im  $MS^3$ -Spektrum die Phosphorylierungsstelle zugewiesen werden.

Abbildung 4-49 A und B zeigt die Massenspektren (I) und die sich daraus ergebenden Tandemspektren (II und III) identifizierter Phosphopeptide aus Fraktion 3 der SCX-Chromatographie nach  $TiO_2$  Anreicherung.

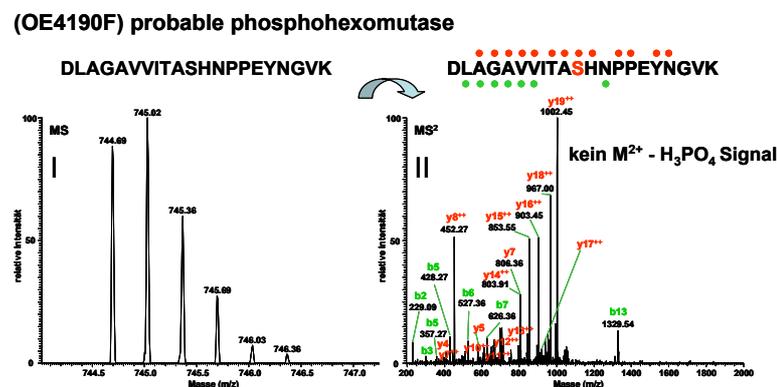


**Abb. 4-49:** Gezeigt sind die MS (I), MS/MS (II) sowie MS/MS/MS (III) Spektren identifizierter Phosphopeptide (A und B). Die zugeordneten y- und b-Fragmentionen (rot bzw. grün) sind in den jeweiligen Spektren als auch über (y-Ionen, rot) bzw. unter (b-Ionen, grün) der zugehörigen Peptidsequenz angegeben. Die phosphorylierte Aminosäure ist jeweils rot hervorgehoben. In beiden  $MS^2$ -Spektren ist das dominante *Neutral Loss* Signal zu erkennen.

Für beide Peptide ist im MS/MS Spektrum (II) klar das dominante *Neutral Loss* Signal zu erkennen. Es können dennoch bereits hier einige entstandene Fragmentationen der Peptidsequenz zugeordnet werden, was die Annotation der phosphorylierten Aminosäure zulässt. In den jeweiligen MS<sup>3</sup> Spektren (III) wurde eine große Sequenzabdeckung durch die entstandenen Fragmentensignale erhalten, die die vorhergesagte Phosphorylierungsstelle aus den MS<sup>2</sup> Spektren klar bestätigt.

Somit konnte die für die 'pyruvat water dikinase' mittels LTQ FT-ICR MS Threonin-12 des Peptids (Rest 392 des Proteins) als eindeutige Phosphorylierungsstelle identifiziert werden. Dies war im vorangegangenen Versuch mittels Q-TOF MS von IMAC-Proben aus 1D-Gel Slices nicht möglich (siehe 4.3.3.1). Als zweite Phosphorylierungsstelle konnte Serin-11 der Peptids (Rest 94 des Proteins) der 'probable phosphomannomutase' mittels MS<sup>3</sup> eindeutig identifiziert werden.

Das Auftreten des charakteristischen *Neutral Loss* Signals im MS<sup>2</sup> Spektrum ist jedoch nicht immer gegeben. Abbildung 4-50 zeigt Massen- sowie Tandemspektrum des Phosphopeptids der 'probable phosphohexomutase' (OE4190F) aus Fraktion 3 der SCX-Chromatographie, für das ohne *Neutral Loss* Signal Ser-10 (Rest 103 des Proteins) als Phosphorylierungsstelle identifiziert wurde.



**Abb. 4-50:** Gezeigt sind die MS (I) und MS/MS (II) Spektren eines identifizierten Phosphopeptids. In diesem Fall ist kein *Neutral Loss* Signal erkennbar (kein MS<sup>3</sup>). Die zugeordneten y- und b-Fragmentationen (rot bzw. grün) sind in den jeweiligen Spektren als auch über (y-Ionen, rot) bzw. unter (b-Ionen, grün) der zugehörigen Peptidsequenz angegeben. Die phosphorylierte Aminosäure ist jeweils rot hervorgehoben.

Im MS<sup>2</sup> Spektrum ist kein Signal zu finden, welches der Masse des *Neutral Loss* Fragments entsprechen würde. Demnach wurde auch kein MS<sup>3</sup> durchgeführt, da die automatische Einstellung des Massenspektrometers dies nur für *Neutral Loss* Signale vorsieht. Dennoch konnte aus dem MS<sup>2</sup> Spektrum die Phosphorylierungsstelle identifiziert werden. Neben den hier gezeigten wurden keine weiteren Phosphopeptide in den analysierten frühen SCX-Fractionen gefunden.

### 4.3.3.3 Zur Biologie der identifizierten Phosphoproteine

Nach der Identifizierung von halobakteriellen Proteinphosphorylierungen war nun die Frage zu beantworten von welcher funktionellen Bedeutung diese Modifikationen sind. Tabelle 4-7 fasst die drei Phosphorylierungsstellen der identifizierten Phosphoproteine zusammen.

**Tab. 4-7:** Identifizierte Phosphorylierungsstellen halobakterieller Proteine.

<b>(Gen-Code) Proteinname</b>	<b>phosphorylierte Aminosäure</b>
1. (OE1500R) pyruvate water dikinase	Threonin-392
2. (OE4190F) probable phosphohexomutase	Serin-103
3. (OE4094F) probable phosphomannomutase	Serin-94

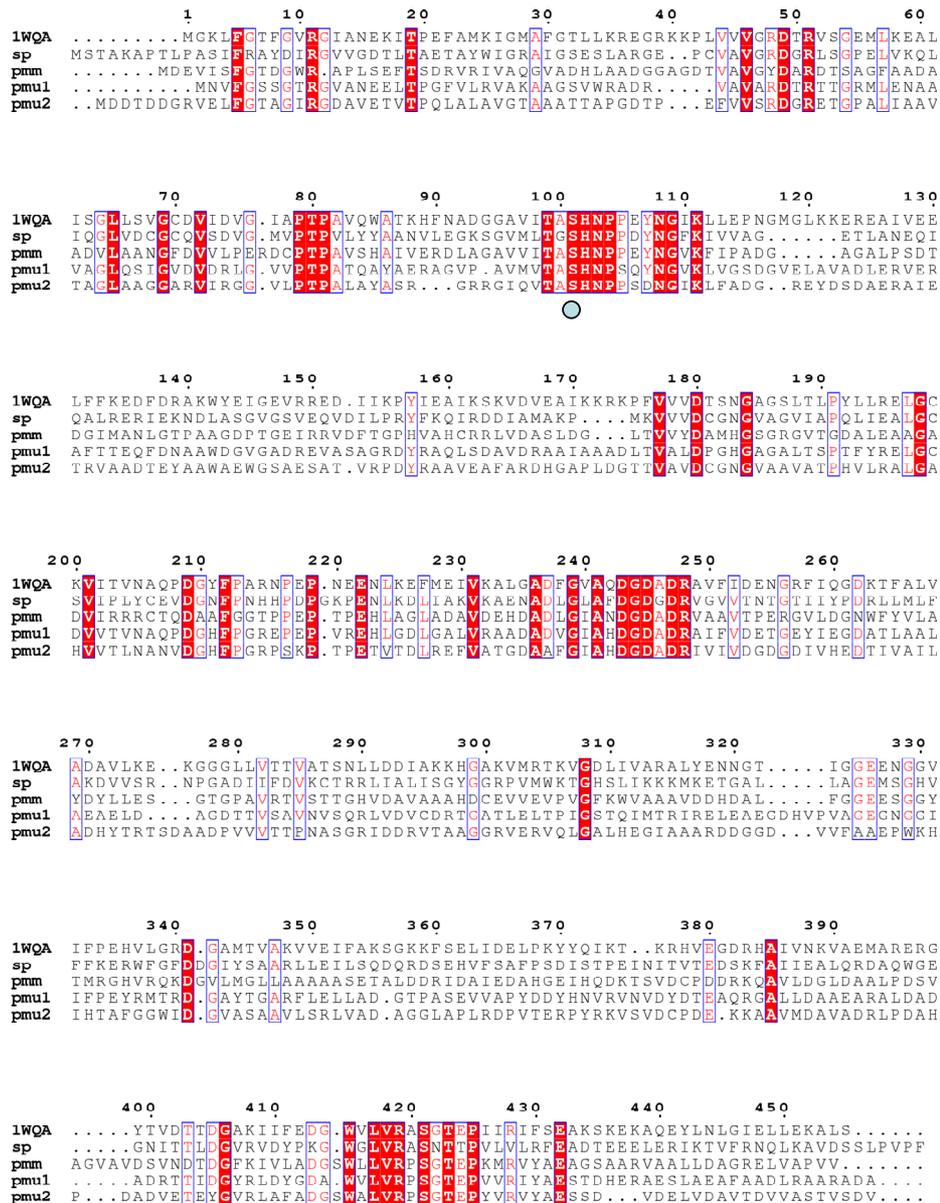
Die ‘pyruvate water dikinase’ (OE1500R) katalysiert den ersten Schritt der Gluconeogenese, indem sie Pyruvat und ATP in Phosphoenolpyruvat, AMP und Phosphat umwandelt. Es handelt sich also um ein phosphatgruppenübertragendes Enzym (Kinase). Aus *E. coli* wurde eine ebenfalls phosphorylierte ‘pyruvate water dikinase’ isoliert, für die gezeigt wurde, dass es sich hierbei um ein Intermediat des Phosphatgruppen-Übertragungsmechanismus handelt [Berman und Cohn, 1970; Hutchins *et al.*, 2001]. Dies lässt vermuten, dass die identifizierte Proteinphosphorylierung der halobakteriellen ‘pyruvate water dikinase’ ebenfalls ein Zwischenprodukt des Reaktionszyklus darstellt.

Die beiden anderen Proteine, für die Phosphorylierungsstellen identifiziert werden konnten, sind funktionell miteinander verwandt. Es handelt sich bei ihnen um Phosphohexomutasen, die die gegenseitige Umwandlung von 1-Phosphat Hexosen in 6-Phosphat Hexosen katalysieren. Bei Betrachtung der Sequenzen der identifizierten Phosphopeptide ist eine Übereinstimmung der Phosphorylierungsepitope zu erkennen, was bei der funktionellen Verwandtschaft und der hohen Sequenzhomologie zu erwarten war. Die identifizierten Phosphorylierungsepitope erlauben eine indirekte Zuweisung der Phosphorylierungsstelle eines weiteren Phosphoproteins OE2318R, das mittels *in vivo* Markierung als phosphattragend identifiziert wurde. Es handelt sich hierbei ebenfalls um eine Phosphohexomutase (‘phosphomannomutase’), weshalb anzunehmen ist, dass auch hier das homologe Serin-102 die Phosphorylierungsstelle darstellt. Tabelle 4-8 zeigt die Sequenzen der identifizierten Phosphopeptide sowie die tryptischen Phosphopeptide der ‘phosphohexomutase’ aus *Pyrococcus horikoshii* und *Pseudomonas aeruginos*, die bereits als phosphoryliert beschrieben wurden.

**Tab. 4-8:** Sequenzen der identifizierten Phosphopeptide (1-3) sowie Sequenz eines möglichen Phosphopeptids eines phosphattragenden Proteins. Ebenfalls abgebildet sind die Sequenzen des Phosphopeptids einer phosphomannomutase aus *Pyrococcus horikoshii* und *Pseudomonas aeruginosa*.

(Gen-Code) Proteinname	Sequenz
1. (OE1500R) pyruvate water dikinase	AAGIVTDEGGMTSHAAIVSR (ident.)
2. (OE4190F) probable phosphohexomutase, pmu3	DLGAVVITASHNPEYNGVK (ident.)
3. (OE4094F) probable phosphomannomutase, pmu1	AGVPAVMVTASHNPSQYNGVK (ident.)
-- (OE2318R) phosphomannomutase, pmu2	GIQVTASHNPPSDNGIK (aus 32-G Gelen)
-- Phosphohexomutase, <i>Pyrococcus horikoshii</i> , WQA1	HFNADGGAVITASHNPEYNGIK
-- Phosphomannomutase, <i>Pseudomonas aeruginosa</i> , SP	SGVMLTGSHNPPDYNGFK

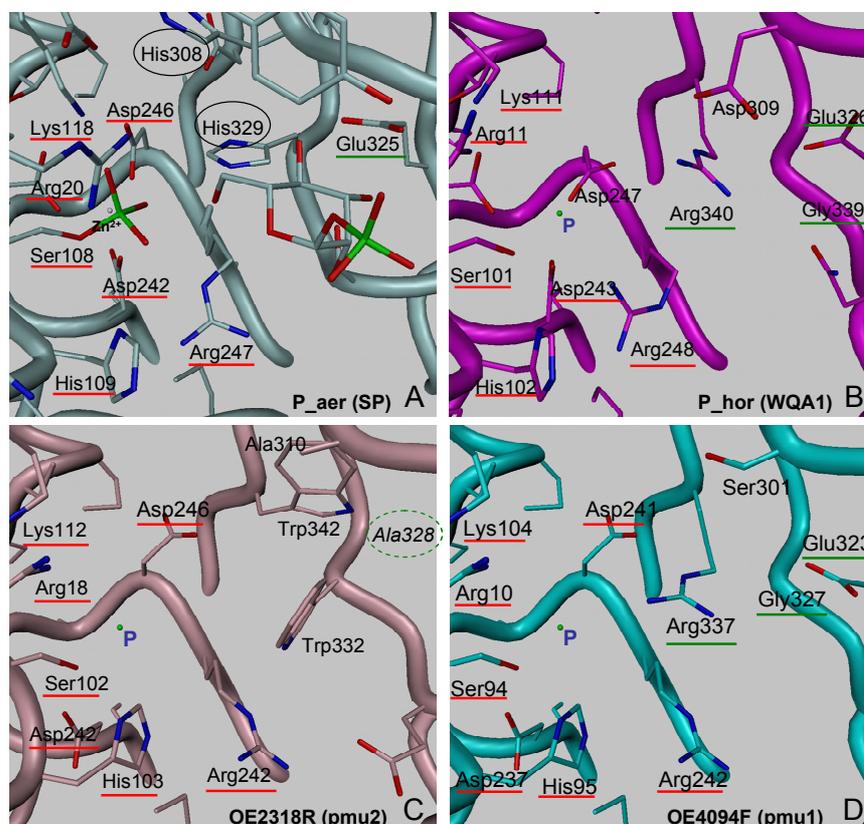
Für den Reaktionsmechanismus der ‘phosphohexomutase’ aus *Pseudomonas aeruginosa* konnte gezeigt werden, dass das Phosphat vom Serin-108 auf die 6- bzw. 1-Position der 1- bzw. 6-Phosphat Hexose und das andere Phosphat von der 1- bzw. 6-Position zurück auf das Protein übertragen wird [Naught und Tipton, 2001; Naught *et al.*, 2003]. Dieses Enzym und ebenfalls das Phosphorylierungsepitop weisen hohe Sequenzhomologie zu den Phosphohexomutasen aus *H. salinarum* auf. Es ist deshalb anzunehmen, dass es sich auch bei den hier identifizierten Phosphorylierungen wiederum um Zwischenprodukte des Reaktionszyklus handelt. Abbildung 4-51 zeigt den kompletten Sequenzabgleich zwischen den drei halobakteriellen Phosphohexomutasen und dem Homolog aus *Pyrococcus horikoshii* und *Pseudomonas aeruginosa*.



**Abb. 4-51:** Sequenzabgleich der drei halobakteriellen Phosphohexomutasen pmm (OE4190F), pmu1 (OE4094F) und pmu2 (OE2318R) mit dem Homolog aus *Pyrococcus horikoshii* (1WQA) und *Pseudomonas aeruginosa* (sp). Zu 100% konservierte Reste sind in weiß mit rotem Hintergrund dargestellt. Reste, die in drei der vier Proteinen konserviert sind, sind rot gedruckt und durch einen blauen Kasten hervorgehoben. Das phosphorylierte Serin ist mit einem hellblauen Punkt unter der Sequenz gekennzeichnet.

Für die Phosphohexomutase aus *Pyrococcus horikoshii* [noch nicht veröffentlicht] und *Pseudomonas aeruginosa* [Regni *et al.*, 2003] konnten die Kristallstrukturen gelöst werden. Dies erlaubte die Rekonstruktion der Strukturen für zwei der halobakteriellen Phosphohexomutase. Abbildung 4-52 zeigt die Kristallstrukturen der aktiven Zentren der Homologen aus *Pyrococcus horikoshii* und

*Pseudomonas aeruginosa* (A+B) sowie die Modellierungen für OE2318R (pmu2) und OE4094F (pmu1) aus *H. salinarum*.



**Abb. 4-52:** Kristallstrukturen der aktiven Zentren der Phosphohexomutasen aus *Pseudomonas aeruginosa* (A) und *Pyrococcus horikoshii* (B) sowie die anhand dieser Strukturen erhaltene Modellierungen der beiden halobakteriellen Enzyme (C+D). In B bis D ist die Position des Phosphats lediglich anhand der Koordinaten aus A angegeben. Reste, die in allen Proteinen konserviert sind, sind in rot unterstrichen, Reste die lediglich in A, B und D übereinstimmen, sind grün unterstrichen. Reste, die in A als wichtig für die Substratkoordination angesehen werden, jedoch in den anderen Strukturen fehlen, sind in A schwarz umrandet.

Die Struktur aus *Pseudomonas aeruginosa* enthält die phosphorylierte Aminosäure (Ser108) sowie Glucose-1-Phosphat als Substrat (A). Die Struktur des Enzyms aus *Pyrococcus horikoshii* (B) und die gemodelten Strukturen der Homologen aus *H. salinarum* (C+D) beinhalten weder das Substrat noch das phosphorylierte Serin. Die Position des Phosphats ist lediglich anhand der Raumkoordinaten aus A angegeben (blauer Punkt mit P markiert). Es zeigt sich, dass eine Vielzahl der Reste des aktiven Zentrums konserviert sind (rot unterstrichen). So sind das dem phosphorylierten Serin folgende Histidin sowie weitere umgebende Aminosäuren, die an der Koordination des Phosphats beteiligt sind (Arg, Lys, Asp), in allen Strukturen konserviert. Reste, die das Substrat koordinieren, sind jedoch weniger stark konserviert. So sind z.B. die beiden Histidine (308 und 329) in dem Enzym aus *Pseudomonas aeruginosa* (A) in dem

halobakteriellen Homolog pmu2 (C) durch zwei Tryptophane funktionell ersetzt. In den anderen beiden Enzymen aus *Pyrococcus horikoshii* (B) und *H. salinarum* (pmu1, D) scheint ein Arginin die Koordination zu übernehmen. Das Glutamat, welches in (A) die Koordination des Phosphats des gebundenen Zuckers übernimmt, ist strukturell lediglich in (B) und (D) wiederzufinden.

Dieses eher uneinheitliche Bild der Reste die für die Substratkoordination verantwortlich sind, lässt vermuten, dass die halobakteriellen Enzyme eventuell eine andere Substratspezifität aufweisen.

### 4.3.4 Identifizierung N-terminaler Acetylierungen

Bei der Analyse der frühen Fraktionen der SCX-Chromatographie zur Anreicherung phosphorylierter Peptide, aber auch bei den LC-TOF/TOF Analysen isotopenmarkierter Proteine wurden eine Vielzahl an N-terminal acetylierten Peptiden identifiziert. Tabelle 4-9 listet die N-terminalen Sequenzen der acetylierten Proteine auf.

Tab. 4-9: Peptidsequenzen (1) der Proteine (2), die als N-terminal acetyliert identifiziert wurden.

Peptidsequenz <sup>1</sup>	(Gene Code) Proteinname <sup>2</sup>
AANTHADDTR	(OE4303R) ABC -type transport system permease protein
AATTPVIAAAYR	(OE3884F) acetyl -CoA C-ac(et)yltransferase
AQQMGNQPLIVLSEDSQR	(OE4122R) thermosome alpha chain
AQTPPPEAPSTDESEPETALETAR	(OE1270F) glutamate dehydrogenase
AQTQLPHDAKAGPTKPEVR	(OE3207F) probable precorrin -8W decarboxylase
ATETDLEELRR	(OE3524F) pyridoxine biosynthesis protein
FGGGGMNPR	(OE1633F) conserved hypothetical protein
GQQQSVILSFEDGAR	(OE4568R) conserved protein
MNGDEVRR	(OE4344F) excinuclease ABC chain C
MQGNLPPEAQEK	(OE1371R) prefoldin beta chain
MQGQNNQQAYDR	(OE1275F) proteasome alpha chain
MTAASTR	(OE2478R) conserved hypothetical protein
SAEATEADLSEDER	(OE2554R) conserved protein
SAESEEAGTAEVIEVVGK	(OE2664F) ribosomal protein S28.eR
SAEPIATFESSLDEGVVELVR	(OE6062R) conserved hypothetical protein
SAESQDGSLEER	(OE4292F) conserved hypothetical protein
SAPDTVSPDLR	(OE3843F) NAD <sup>+</sup> synthase
SAPEEEFYSEER	(OE3289F) conserved hypothetical protein
SASGPANIEYVVAR	(OE3987R) H <sup>+</sup> -transporting two -sector ATPase
SATDDLVSALR	(OE3953R) orotate phosphoribosyltransferase
SDADTAPDPSDIGETDAPPIEDAPYK	(OE3696F) hypothetical protein
SDANEAADYGTLYVVGIGPLPGGMTQR	(OE3214F) precorrin -3B C17-methyltransferase
SDAQPAEADPDALVGLR	(OE4099R) hypothetical protein
SDAQPAEDETAEAAESEDTEVAANAK	(OE4735R) ribosomal protein S7
SDSEPAQGSIEDQGPVEISPELAR	(OE4718F) ribonuclease R homolog
SEAQPDAR	(OE1405R) conserved protein
SESEYKIADGSGKFLQAVKDGR	(OE2402F) conserved protein
SETDSASETVEEPTTWPDLAVGLYDR	(OE2872F) hypothetical protein
SETTDASGLSAGLPSR	(OE2548F) probable transmembrane oligosaccharyl transferase
SGATSTGDVGYDELAVLK	(OE3964R) conserved protein
SGFGDVPDQYDPAGVEER	(OE4068F) isoleucine --tRNA ligase
SGLELDATQLDR	(OE1140R) molybdenum cofactor biosynthesis protein moeB
SGTPDTLFAAAGATR	(OE3242F) conserved cobalamin operon protein
SKAFSAIDEDR	(OE2622R) pyruvate --ferredoxin oxidoreductase
SKHSNAAHAPGVGR	(OE3963R) 3,4 -dihydroxy -2-butanone 4 -phosphate synthase
SLAEFDADVVDAR	(OE2633F) ribosomal protein L13
SNDTDTPELAR	(OE2450F) acetate --CoA ligase
SQQLPDVQATEPDVSVGLSEVGTGVEK	(OE3673F) conserved protein
SQSGGSDMTLAFELSALK	(OE3547F) conserved protein
SQSSLDDDELFGAAEEMR	(OE1205R) hypothetical protein
SSDAPDDQPPADAGVR	(OE2696F) hypothetical protein
SSEDIQDTEAR	(OE1782F) conserved protein
SSEDIQDTEAR	(OE3973F) peptide chain release factor eRF-1
SSEDL SADDL TLPVK	(OE3328R) ribonucleotide reductase
SSEEQEDVDEDIQYFVR	(OE2627F) ribosomal protein S13
SSESPDVPDDVADR	(OE2304F) DNA topoisomerase (ATP -hydrolyzing)
SSESPESVDDSTKVQATAEWDPLQAVR	(OE5208R) arginine deiminase
SSGQNSGGLMSAGLVR	(OE4646R) hypothetical protein
SSIELTASQK	(OE2306F) conserved protein
SSKGSNYGDIHR	(OE1878R) protein OE1878R
SSLAGETTADGPPVATSCR	(OE2566R) isochorismate synthase
SSPNLAHPPTDYTGIVLFP	(OE7033R) gas -vesicle operon protein gvpD1
SSPPTSSAPLGDIEPAGSR	(OE3227F) nicotinate -nucleotide -phosphoribosyltransferase homolog
STKPNEGLLAR	(OE3907R) ABC -type transport system permease protein
STKPSQSETETETDTEPATTHGQQR	(OE2866R) succinate dehydrogenase
STQKNAR	(OE4427R) DNA -binding protein dpsA
SVELPFAPVDSLIR	(OE1220F) archaeal histone
SVREEFDEWAASGR	(OE4360R) probable methyltransferase
SYNDTWQPKTR	(OE3415F) ribosomal protein S5
TFDSFADANAEVTR	(OE4651F) thiamine biosynthesis protein

Es fiel auf, dass bei den meisten, aber nicht bei allen der acetylierten Proteine das Initiationsmethionin abgespalten war und so die zweite Aminosäure die Acetylgruppe trug. Lediglich 4 der 60 identifizierten modifizierten Proteine trugen die Acetylierung am Methionin-1. Bei Betrachtung der zweiten Aminosäure wurde deutlich, dass neben Alanin, Phenylalanin, Glycin und Threonin bevorzugt Serin acetyliert wird (47 mal). Die dem Serin folgende Aminosäure scheint jedoch keinen starken Einfluss zu haben, da hier sowohl hydrophobe (Alanin, Glycin, Leucin, Valin) als auch hydrophile (Serin, Threonin, Tyrosin, Glutamin) sowie saure (Aspartat, Glutamat) und basische (Lysin) Aminosäuren folgen.

Der N-terminalen Proteinacetylierung kann eine Schutzfunktion vor proteolytischem Abbau zugeschrieben werden. Durch die Kenntnis der N-acetylierten Proteine ist so deren Lebensdauer abzuschätzen. Wie bereits in 4.1 angesprochen werden durch die stetige Erweiterung der Proteininventarisierung mittels MS die Genomannotation fortlaufend validiert und verbessert. Der Identifizierung N-terminaler Peptide ist dabei eine wichtige Stellung zuzuweisen, da so die Startcodonvorhersage weiter überprüft werden kann. Dies ist durch die bloße Identifizierung der Proteine durch interne Peptide, wie in Abschnitt 4.1.6 gezeigt wurde, nicht unbedingt möglich.

## 5 Diskussion

In dieser Arbeit wurden unterschiedliche Trenn- und massenspektrometrische Methoden zur Analyse cytosolischer Proteine von *H. salinarum* angewendet. Diese sollen im Folgenden für sich, aber auch im Vergleich zueinander diskutiert werden.

### 5.1 Das zweidimensionale Referenzgel

Die Analyse der cytosolischen Proteine von *H. salinarum* mittels zweidimensionaler Gelelektrophorese und anschließender Identifizierung über MALDI-TOF *peptide mass fingerprinting* verfolgte im Wesentlichen zwei übergeordnete Ziele:

Zum einen sollte das Proteininventar auf standardisierten 2-D Gelen erfasst werden. Die daraus erhaltenen Proteinidentifikationen wurden zur ständigen Validierung und Verbesserung der Genomannotation herangezogen. Des Weiteren sollte die Effizienz der automatisierten massenspektrometrischen Analyse der in 2-D Gelen aufgetrennten Proteine auf der Grundlage der erhaltenen Identifizierungen abgeschätzt werden.

#### 5.1.1 Das cytosolische Proteininventar mittels 2-DE

Für das Erstellen des 2-DE Referenzproteoms wurden in Komplexmedium wachsende Zellen in der späten logarithmischen Wachstumsphase gewählt. In diesem Zellzustand sollten alle Proteine, die für den Standardstoffwechsel und das Zellwachstum benötigt werden, vorhanden sein, Proteine, die durch das Altern der Zellen oder Hungerzustände induziert werden, jedoch fehlen. Das cytosolische Proteom wurde durch den in „Material und Methoden“ beschriebenen „milden“ Zellaufschluss unter Hochsalzbedingungen definiert. Diese Vorgehensweise wurde gewählt, um eine Verunreinigung der Probe durch periphere Membranproteine, Zellwandbestandteile oder Flagellen zu vermeiden, welche unter Niedrigsalzbedingungen löslich werden. Die Präparation von Proteinextrakten aus stark salzhaltigen Lösungen stellt jedoch eine besondere Herausforderung dar, da das Salz für die nachfolgenden Analyseschritte wie die isoelektrische Fokussierung nahezu vollständig entfernt werden muss. Dies konnte effektiv durch Acetonpräzipitation und nachfolgende Waschschrte erreicht werden und stellt eine schnelle Alternative zur Trichloressigsäure Präzipitation und insbesondere zur Entsalzung mittels Dialyse oder Filtration durch *cut-off* (Ausschluss-) Membranen dar

[Cho *et al.*, 2003; Karadzic and Maupin-Furlow 2005]. Die Verwendung von hochauflösenden *narrow-range* IPG-Streifen zur IEF [Tonella *et al.*, 2001; O'Neill *et al.*, 2002] war aufgrund der engen *pI*-Verteilung der halobakteriellen Proteine eine unabdingbare Voraussetzung für das Erstellen reproduzierbarer 2-D Gele.

Mit dem im folgenden Abschnitt diskutierten Hochdurchsatzverfahren konnten 661 Proteine aus den Standard Referenzgelen sowie 141 weitere auf nicht standardisierten 2-D Gelen identifiziert werden. Dies entspricht ca. 40% des gesamten cytosolischen Proteininventars von *H. salinarum* und repräsentiert die höchste Proteomabdeckung, die mit der Kombination aus 2-DE und anschließender Proteinidentifizierung über MALDI-TOF PMF bisher erreicht wurde. Dabei ist zu erwähnen, dass bei der Betrachtung des cytosolischen Proteoms alle Proteine ohne vorhergesagte Transmembranregionen (TMHMM) berücksichtigt wurden. Dies schließt also membranlokalisierte periphere Membranproteine mit ein, und deshalb sollte die tatsächliche Anzahl der in der cytosolischen Probe befindlichen Proteine noch geringer ausfallen und so die Proteomabdeckung bezogen auf die analysierte cytosolische Probe noch größer sein. Nur eines der 802 identifizierten Proteine enthielt eine vorhergesagte Transmembranregion, was darauf hindeutet, dass die Proteinprobe nur gering mit eventuellen Transmembran- oder peripheren Membranproteinen verunreinigt war. Es ist jedoch anzumerken, dass in der Probe befindliche Membranproteine mittels 2-DE und MALDI-PMF zum größten Teil nicht aufgetrennt und identifiziert werden können [Klein *et al.*, 2005] und deshalb eine Abschätzung der *Reinheit* der Probe mit diesen Daten nur schwer möglich ist. Für die Proteinidentifizierungen wurden zur Verringerung der möglichen falschpositiven Ergebnisse die Mascot Score Anforderungen zusätzlich verschärft, um einen möglichst zuverlässigen Datensatz zu generieren. Die von Mascot vorgeschlagene Scoregrenze zur Erlangung einer Falsch Posiven Rate von 5% stellte sich für die MALDI-TOF PMF Analyse als zu niedrig heraus.

Die identifizierten Proteine können zu einem großen Teil den Reaktionswegen des zentralen Stoffwechsels zugeordnet werden. Bis zu 90% der Proteine der Gluconeogenese, des Citrat Zyklus, des Purin und Pyrimidine Metabolismus sowie des Cofaktor und der Aminosäure Biosynthese konnten identifiziert werden. Ausnahme ist hier die Atmungskette zu nennen, die aufgrund ihrer membranständigen Komponenten nicht im cytosolischen Proteom zu finden war, wohl aber im Membranproteom [Klein *et al.*, 2005]. Ebenfalls waren ribosomale Proteine unterrepräsentiert, was jedoch an der häufig geringen molekularen Masse der ribosomalen Untereinheiten liegen könnte, da kleine Proteine in den vorliegenden Analysen dieser Arbeit unterrepräsentiert sind. Für ca. 50% der identifizierten Proteine konnten keine der üblichen funktionellen Klassen

zugewiesen werden, und wurden deshalb in der Klasse *No Function* (NOF) zusammengefasst.

Bei der Betrachtung der Genomlokalisierung der identifizierten Proteine fielen starke Unterschiede zwischen den auf dem Chromosom und den auf Plasmiden codierten Proteinen auf. Fast 50% der chromosomalen Proteine, jedoch nur 14% der Plasmidproteine wurden identifiziert. Daraus kann geschlossen werden, dass ein Großteil der auf den Plasmiden codierten Proteine, über deren biologische Funktionen in vielen Fällen keine Aussage getroffen werden kann, nicht exprimiert werden, da sie vermutlich nur unter speziellen Lebensbedingungen gebraucht werden. Auf die Expressionsunterschiede zwischen Plasmid- und chromosomalen Proteinen wird noch einmal genauer eingegangen. Die Unterschiede werden wie bereits erwähnt auch in der Nukleotidzusammensetzung mit 68% GC-Gehalt des Chromosoms und einem durchschnittlichen GC-Gehalt von 59% der Plasmide deutlich. Diese Differenzen gehen einher mit unterschiedlichen CAI-Werten der chromosomalen und den auf Plasmiden lokalisierten ORFs. Die Berechnung des CAI-Werts kann zur Abschätzung der Expressionsstärke eines Proteins herangezogen werden und zeigte im Fall von *H. salinarum* klar, dass die chromosomalen Proteine stärker exprimiert werden sollten als die der Plasmide (in 5.3.1 wird noch einmal näher darauf eingegangen). Dies ist eine weitere Erklärung der unterschiedlichen Identifizierungsraten von chromosomalen- und Plasmidproteinen.

Dennoch gibt es Abweichungen von den unterschiedlichen GC-Gehalten der beiden genetischen Elemente sowie den Identifizierungsraten der auf ihnen codierten Proteine. Das Chromosom weist eine ca. 60 kbp große Region mit einem eher moderaten GC-Gehalt auf, die bereits als AT-reiche Insel [Pfeifer und Betlach, 1985] beschrieben wurde. In diesem Bereich des Chromosoms ist die Identifizierungsrate der codierten Proteine stark verringert und entspricht eher den für Plasmiden typischen Werten. Dies lässt vermuten, dass diese Region einem ehemaligen Plasmid oder Teil eines Plasmids entspricht, der in das Chromosom integriert ist. Überraschend war jedoch die Identifizierung einer Region auf dem Plasmid pHS3 mit hohem, für das Chromosom typischen GC-Gehalt und ebenfalls hoher Identifizierungsrate der in dieser Region codierten Proteine. Daraus kann geschlossen werden, dass diese 70 kbp große Region ehemals chromosomaler DNA entspricht, die in das Plasmid integriert ist und dort ihren hohen GC-Gehalt und die entsprechende Expressionsrate der auf ihr codierten Gene bewahrt hat. In dieser Region und in ihrer direkten Nachbarschaft werden wichtige und essentielle Gene wie die Catalase, die beiden Ketten der Carbamoylphosphattransferase und die einzige Arginin-tRNA Ligase codiert, weswegen der Plasmid pHS3 eher als zweites Chromosom angesehen werden sollte.

Die eben beschriebene Untermauerung des DNA Austausches zwischen Chromosom und Plasmiden sowie die Bestätigung eines Zusammenhangs zwischen GC-Gehalt und Expressionsrate sind nicht die einzigen Beispiele, in denen die Proteomik essentiell zum genomischen Verständnis von *Halobacterium salinarum* beigetragen hat. Wie im Ergebnisteil gezeigt konnten eine Reihe von Annotationsfehlern erst mit Hilfe massenspektrometrischer Proteom-Daten oder 2-DE Korrelationsanalyse korrigiert bzw. erst eine sinnvolle Genomannotation erzielt werden. Es ist anzumerken, dass Organismen mit hohen GC-Gehalten und den daraus resultierenden Schwierigkeiten für die Genomannotation einen Spezialfall darstellen (siehe Dissertation Michaela Falb). Die Tatsache, dass für *H. salinarum* neben den ca. 2,800 für Proteine codierenden Genen 6,000 zusätzliche ORF vorhergesagt wurden, was einer durchschnittlichen Anzahl von 2.8 offenen Leseramen pro Genomregion entspricht, verdeutlicht die für diesen Organismus spezielle Situation. Dennoch ist abzusehen, dass auch für andere Organismen, deren Genome sequenziert wurden oder noch werden, die Proteomik einen entscheidenden Beitrag zur Verbesserung der Genomannotation leisten kann [Desiere *et al.*, 2005]. Dies gilt insbesondere für die Überprüfung der Vorhersage von unterschiedlichen Spleißvarianten in eukaryotischen Genomen. In welchem Ausmaß jedoch die Genomdatenbanken einer substantiellen Verbesserung bedürfen, bleibt abzuwarten.

### 5.1.2 Die automatisierte MALDI-TOF PMF Analyse

Die Hochdurchsatzanalyse von Proteinidentifizierungen aus 2-D Gelen bedarf einer erheblichen Automation, um die anfallenden Probenmengen zu bearbeiten. Dazu wurden in den letzten Jahren unterschiedliche Ansätze vorgestellt [Nordhoff *et al.*, 2001], die mittlerweile von verschiedenen Herstellern auch kommerziell vertrieben werden. Das in der Arbeitsgruppe Massenspektrometrie unter der Leitung von Dr. Frank Siedler entwickelte Verfahren (siehe „Material und Methoden“ sowie Dissertation Christian Klein) erlaubte die Bearbeitung von 7,500 Gelspots sowie das Aufnehmen von 15,000 Massenspektren in einem passablen Zeitrahmen von ca. 6 Wochen. Mit diesem Verfahren konnten 40% des cytosolischen Proteoms identifiziert werden. Dies entspricht der höchsten Proteomabdeckung, die bisher mit der Kombination aus 2-DE und MALDI-TOF PMF erzielt werden konnte. Die 2-D Gele wurden ausschließlich mit Silbernitrat gefärbt. Diese einfache und sensitive Färbemethode ermöglicht zwar die massenspektrometrische Identifizierung der aufgetrennten Proteine [Shevshenko *et al.*, 1996], weist jedoch im Vergleich zu anderen Färbungen mit z.B. Coomassie [Neuhoff *et al.*, 1988] oder Fluoreszenzfarbstoffen [Breggren *et al.*, 2000; Rabilloud *et al.*, 2001; Berggren *et al.*, 2002] Nachteile auf, wie eine verringerte Sequenzabdeckung der

proteolytischen Peptide nach In-Gel Verdau im Massenspektrum [Richert *et al.*, 2004] oder eine verstärkte Oxidation der Peptide. Bei der Verwendung solcher Alternativen wäre vielleicht ein noch besseres Ergebnis zu erwarten gewesen.

Die Verwendung einer alternativen Probenpräparation mit der kürzlich entwickelten Gyrolab<sup>TM</sup> Technologie, welche eine höhere Sensitivität und somit mehr Identifizierungen verspricht [Gustansson *et al.*, 2004], brachte in keiner Weise eine Verbesserung. Es wurden überwiegend mit dem herkömmlichen Verfahren bessere Ergebnisse erzielt. Dies demonstriert die Qualität dieses robusten und einfachen Hochdurchsatzverfahrens. Dennoch könnten weitere technologische Neuerungen wie die Verwendung von Anchor Targets [Nordhoff *et al.*, 2003] oder der Einsatz von Verdaurobottern den Ablauf weiter beschleunigen und verbessern.

Bisher konnten von keinem Proteom alle zugehörigen Proteine identifiziert werden. Trotz der hier, im Vergleich zu anderen Arbeiten, erreichten hohen Proteomabdeckung muss die Frage nach den Proteinen gestellt werden, die nicht mit dem gewählten Ansatz identifiziert werden konnten. Neben den allgemeinen bei der 2-DE Probenaufarbeitung anfallenden Probenverlusten [Zhou *et al.*, 2005] können folgende Gründe angeführt werden: (a) die Proteine wurden unter den gewählten Wachstumsbedingungen nicht expremiert oder ihr Expressionsniveau war zu gering, um detektiert und identifiziert zu werden, (b) nicht alle Proteine waren Teil des analysierten Subproteoms (in diesem Fall des cytosolischen Subproteoms), da sie im Zuge der Fraktionierung und Aufarbeitung verloren gingen, oder (c) die Proteine konnten mit der gewählten massenspektrometrischen Analyseverfahren (hier das *peptide mass fingerprinting*) trotz ausreichender Mengen nicht identifiziert werden.

Wie bereits erwähnt konnte aus den unterschiedlichen Identifizierungsraten chromosomaler und auf Plasmiden codierter Proteine zusammen mit den unterschiedlichen GC-Gehalten beider genomischen Elemente abgeleitet werden, dass chromosomale Proteine zu einem größeren Ausmaß expremiert werden. Daraus folgt, dass ein Großteil der Plasmidproteine gar nicht identifiziert werden können. Der Verlust von Proteinen während der Probenaufarbeitungen ist nur schwer zu vermeiden, kann aber abgeschätzt werden. Dennoch konnte eine Abreicherung von kleinen Proteinen (<25 kDa) sowohl auf den Gelbildern als auch daraus resultierend in ihrer prozentualen Identifizierungsrate beobachtet werden. Es ist deshalb anzunehmen, dass spezifisch kleine Proteine bei der gewählten Probenaufarbeitung verloren gehen. Als Grund hierfür ist anzunehmen, dass kleine Proteine aufgrund ihrer physikochemischen Eigenschaften, die mit abnehmender Größe eher Peptidcharakter aufweisen und so nur schlecht oder gar nicht mittels organischen Lösungsmitteln gefällt werden können, in Lösung bleiben [Chertov *et al.*, 2004]. Vor kurzem konnte gezeigt werden, dass beim In-Gel Verdaus

ein Teil der Proteine bereits vor der proteolytischen Spaltung während der Äquilibrierschritte aus dem Gel ausgewaschen wird und so verloren geht (Klein *et al.*, in Vorbereitung). Dies betrifft in verstärktem Maße kleine Proteine. Zusätzlich zu diesen Problemen in der Probenbereitung ist die massenspektrometrische Identifizierung kleiner Proteine mittels PMF problematisch, da sie oftmals nicht genügend tryptische Fragmente im analysierten Massenfenster liefern und so, auch wenn in der Probe in ausreichender Menge vorhanden ist, der Identifizierung mittels PMF entgehen.

Dieses Problem sollte durch die Anwendung von auf Tandemmassenspektrometrie (MS/MS) beruhenden Methoden umgangen werden können, da hier die Identifizierung eines Proteins durch die Analyse weniger oder sogar einzelner Peptide möglich ist. Die Anwendung von MS/MS Techniken für die Analyse des cytosolischen Proteoms von *Halobacterium salinarum* wird im Abschnitt 5.3.1 diskutiert.

## 5.2 Quantitative Proteomanalysen

Die bloße Identifizierung von Proteinen ist zwar die Voraussetzung für Proteomanalysen und im Falle von *Halobacterium salinarum* aufgrund der mehrfach genannten Gründe von entscheidender Bedeutung. Dennoch liegt eine der Hauptaufgaben der Proteomforschung in der vergleichenden Betrachtung von Proteomen unter unterschiedlichen Bedingungen. Aus der Kenntnis der regulierten Proteine kann so ein besseres biologisches Verständnis des analysierten Systems abgeleitet werden. Es wurden halobakterielle Zellen unter maximal unterschiedlichen Bedingungen angezogen und die Proteome mit unterschiedlichen Methoden analysiert. Die in dieser Arbeit angewendeten Methoden werden nun in den folgenden Abschnitten diskutiert.

### 5.2.1 Quantitative Proteomik mittels 2-DE

Die 2-DE als Standardmethode in der Proteomforschung kann zur relativen Quantifizierung von Proteinen herangezogen werden [Hecker und Völker, 2004; Völker und Hecker, 2005]. Nach der Etablierung reproduzierbarer zweidimensionaler Referenzgele für das cytosolische Proteom von *H. salinarum* wurde die 2-DE für die Analyse differentiell exprimierter Proteine unter verschiedenen Wachstumsbedingungen eingesetzt. Es konnten jedoch keine Regulationen im ersten Vergleichsexperiment (aerobes gegen anaerobes phototrophes Wachstum) identifiziert

werden. Auch im zweiten Versuch waren nur wenige Regulationen auf den Gelen auszumachen. Hier erschwerte zusätzlich eine geringere Gelqualität die Analyse der unterschiedlichen Wachstumsbedingungen (Anzucht in Komplexmedium gegen Anzucht in synthetischem Medium). Die gewählte Silbernitrat Färbemethode bietet nicht nur wie bereits erwähnt bei der nachfolgenden massenspektrometrischen Analyse Nachteile, sondern besitzt auch erhebliche Defizite beim Einsatz in der vergleichenden Gelanalyse. Gründe dafür sind der limitierte dynamische Bereich der Färbung sowie das Auftreten von sogenannten Kraterspots. Die Silberfärbung ist im Vergleich zu alternativen Färbemethoden keine Endpunktmethode, d.h. die Spotintensitäten sind abhängig von der Inkubationszeit und können sich so stark in Abhängigkeit vom Färbeprozess voneinander unterscheiden.

Es ist offensichtlich, dass die verwendete Quantifizierungsstrategie mittels 2-DE mit den verfügbaren instrumentellen Voraussetzungen eine möglichst erschöpfende Analyse der regulierten Proteine nicht erlaubte. Neuere auf Fluoreszenzfarbstoffen basierende differentielle Analysetechniken mittels 2-DE wie die DIGE-Methode [Unlu *et al.*, 1997], bei der die Proteine der unterschiedlichen Zustände in einem Gel aufgetrennt und analysiert werden, konnten zur Verbesserung der differentiellen Analyse mittels 2-DE beitragen [Tonge *et al.*, 2001; Friedman *et al.*, 2004]. Dazu bedarf es jedoch spezieller Ausstattung wie Fluoreszenzscannern und zugehörige Auswertesoftware, die in dieser Arbeit nicht zur Verfügung standen.

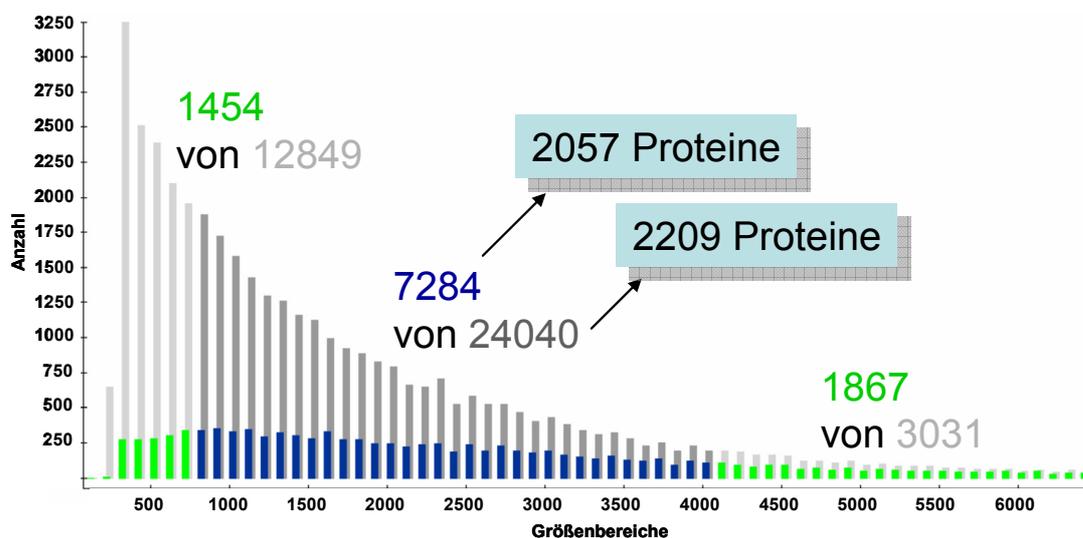
## 5.2.2 Quantitative Proteomik mittels ICPL

Die Verwendung von isotoopenmarkierten Sonden zur Quantifizierung differentiell regulierter Proteine bietet im Vergleich zur Anwendung der 2-DE mehrere Vorteile. Wie bereits beschrieben ist die Quantifizierung hier unabhängig von Färbemethoden und basiert auf dem Vergleich von Peptidsignalen in den aufgenommenen Massenspektren. Diese Methodik kann mit gängiger LC-MS/MS Analytik kombiniert werden und so die zeitintensive Analyse mittels 2-DE, bei der jeder einzelne Proteinspot gesondert bearbeitet werden muss, umgehen.

Ausgehend von der ICAT-Methode, die 1999 von Aebersold und Mitarbeitern vorgestellt wurde [Gygi *et al.*, 1999], sind in letzter Zeit eine Reihe unterschiedlicher Markierungsstrategien entwickelt worden [Julka und Regnier, 2003]. Diese basieren auf unterschiedlichen Chemien zur kovalenten Modifikation der verschiedenen Aminosäuren in Proteinen oder Peptiden. Dabei werden überwiegend die Thiofunktionen von Cysteinen, die Carboxylgruppen der Aspartate und Glutamate sowie der N-Terminus als reaktive Zentren ausgenutzt. Die in dieser Arbeit verwendete ICPL-Methode modifiziert hingegen die in einem Protein auftretenden

Aminofunktionen der Lysine sowie des N-terminus [Schmidt *et al.*, 2005]. Dies hat den Vorteil gegenüber Cysteinmarkierungen, dass eine höhere Sequenzabdeckung an modifizierten Peptiden eines Proteins erreicht werden kann, da Lysine statistisch häufiger auftreten als Cysteine. Dies sollte eine genauere Quantifizierung der Regulation erlauben sowie die Analyse auf z.B. auftretende posttranslationale Modifikationen erweitern.

Ein Nachteil der ICPL-Methode in Bezug auf ihre Anwendung auf *H. salinarum* ist jedoch, dass halobakterielle Proteine einen geringeren Anteil an Lysinen aufweisen als mesophile Proteine. Dieser ist mit 1.8% immer noch höher als der Anteil an Cysteinen (0.75%), liegt aber deutlich unter dem anderer Prokaryoten wie *E. coli* (4.4%) oder *Bacillus subtilis* (7%). Dadurch werden nur verhältnismäßig wenig markierte Peptide nach proteolytischer Spaltung der Proteine erhalten. Dieser Zustand wird in Abbildung 5-1 dargestellt.



**Abb. 5-1:** Größenverteilung aller sowie mittels ICPL zugänglichen Peptide nach Arg-C Spaltung von cytosolischen Proteinen aus *H. salinarum*. Die Anzahl der Peptide, die mittels ICPL markiert werden können (grün bzw. blau), sowie aller Peptide (grau bzw. dunkelgrau) in den unterschiedlichen Größenbereichen (jeweils 100 Da) sind angegeben. Der Massenbereich, der mittels MALDI-TOF abgedeckt wird, ist blau bzw. dunkelgrau hervorgehoben. Ebenfalls gezeigt ist die Anzahl der aus diesem Massenbereich zugänglichen Proteine (blau hinterlegt).

Es wird deutlich, dass nur knapp ein Drittel aller mittels MALDI-TOF MS theoretisch nachweisbaren Peptide (Massenbereich 800 – 4000 Da) markiert werden können. Dies hat zur Folge, dass die Bestimmung der Regulation eines Proteins oftmals nur über wenige Isotopenpaare vorgenommen werden kann. Von 152 cytosolischen Proteinen sind mittels ICPL-Markierung keine Regulationinformationen zu erhalten, da

sie weder ein Lysin enthalten noch das N-terminale Peptid im zugänglichen Massenfenster liegt.

Ein Vorteil der ICPL-Methode ist hingegen, dass die Markierung des intakten Proteins vor der proteolytischen Spaltung erfolgt. Dies erlaubt die Kombination der unterschiedlichen Trennmethode auf Proteinebene, um die Komplexität der Probe bereits hier zu verringern. Andere Methoden gehen von der Markierung aller Peptide nach proteolytischer Spaltung der Proteinprobe aus [Goodlett *et al.*, 2001; Chakraborty und Regnier, 2002]. Die mögliche Sequenzabdeckung liegt hier zwar bei nahezu 100%, da alle Peptide markiert werden können (z.B. über den N-Terminus), jedoch bedarf es für die anschließende Analyse einer effizienten Auftrennung des hochkomplexen Peptidgemisches. Obwohl durch die Kopplung multidimensionaler Chromatographie mit Tandemmassenspektrometern eine beeindruckende Anzahl an Peptiden sowie Proteinen identifiziert werden kann [Washburn *et al.*, 2001; Peng *et al.*, 2003], ist es dennoch zu bezweifeln, dass hochkomplexe Peptidgemische wie die aus ganzen Zellysaten mit mehreren tausend Proteinen mit diesen Technologien befriedigend aufgetrennt und massenspektrometrisch analysiert werden können.

Eine weitere elegante Methode zur Analyse von Zellkulturen basiert auf der metabolen Inkorporation von isotoopenmarkierten Aminosäuren [Ong *et al.*, 2002; Zhu *et al.*, 2002; Ong *et al.*, 2003]. Es wird also eine nachträgliche kovalente Modifikation der Protein- bzw. Peptidprobe umgangen. Dies sollte den experimentellen Fehler der durch die notwendige getrennte Behandlung der zu vergleichenden Proben bis zum Einbringen der Isotopensonden auftreten kann, weiter verringern. Durch eine entsprechende Wahl der isotoopenmarkierten Aminosäure wird auch eine bedeutend bessere Ausbeute an isotoopenmarkierten Peptiden im Vergleich zur ICPL-Methode erzielt.

Diese Methode zur Analyse differentiell regulierter Proteine befindet sich derzeit in der Abteilung Membranbiochemie im Aufbau (Dissertationen Kostas Konstantinidis und Matthias Schlesner). Ein Nachteil dieser Methode ist jedoch, dass sie nicht auf jede Proteinprobe wie z.B. Gewebe angewendet werden kann, sondern auf wachsende Zellen in synthetischem Medium beschränkt ist. Ebenfalls ist darauf zu achten, dass die eingesetzte isotoopenmarkierte Aminosäure für den jeweiligen Organismus essentiell ist, nicht verstoffwechselt werden kann und zu 100% in die Proteine eingebaut wird.

### 5.2.2.1 Die identifizierten Proteine

Die ICPL-Analysen beschränkten sich in dieser Arbeit ebenfalls ausschließlich auf das cytosolische Proteom von *H. salinarum* (für die Analyse halobakterieller Membranproteine siehe Disertation Birgit Bisle). Es wurde ebenfalls die beschriebene

Proteinaufarbeitung unter Hochsalzbedingungen angewendet. Nach Markierung der Proteine der unterschiedlichen Zustände und deren Vereinigung erfolgte eine Fraktionierung der Proteinprobe mittels 1-D SDS PAGE nach Molekulargewicht. Die aus den unterschiedlichen Größenbereichen durch tryptischen in-Gel Verdau generierten Peptidfraktionen wurden mittels LC MALDI-TOF/TOF analysiert. Diese Methode beruht auf der Identifizierung der Proteine durch Tandem Massenspektrometrie der mittels *reversed phase* Chromatographie aufgetrennten Peptide.

In beiden Versuchen (1. Komplexmedium vs. synthetisches Medium (siehe 4.2.3) und 2. aerob vs. anaerob phototroph (siehe 4.2.4)) wurden 1,195 Proteine von *H. salinarum* identifiziert. Dieser Satz an neu identifizierten Proteinen erweiterte um 464 auf 1,265 Identifizierungen. Ein tiefergehender Vergleich der identifizierten Proteine findet sich in Abschnitt 5.3.1, in dem zusätzlich die Daten, die mit einer dritten ebenfalls auf LC-MS/MS basierenden Methode (LC FT-ICR) erhoben wurden, mit einfließen.

#### 5.2.2.2 Die regulierten Proteine

Für die Betrachtung differentiell regulierter Proteine wurde eine relevante Regulation von größer als 100% wie in 4.2.3.3 beschrieben festgelegt. Für die Wahl dieses Grenzwertes können zwei Hauptgründe genannt werden:

Die ICPL-Methode ist zwar ab Werten von 30% zuverlässig [Schmidt *et al.*, 2005], jedoch bedarf es zum Erreichen dieser Genauigkeit einer manuellen Auswertung der jeweiligen Massenspektren. Dies war jedoch im Zuge der durchgeführten umfassenden Analysen des cytosolischen Proteininventars von *H. salinarum* nicht möglich. Die Berechnung der Regulationen erfolgte wie in Material und Methoden beschrieben automatisch. Hierbei kann es jedoch zu höheren Abweichungen und Fehlzusweisungen kommen. Die größten Fehler und Ausreißer wurden zwar nachträglich manuell überprüft und korrigiert, auf eine umfassende manuelle Nachanalyse musste jedoch aus Zeitgründen verzichtet werden.

Die Frage, welche Konzentrationsänderung eines Proteins eine biologische Antwort auslöst, ist sicherlich nicht eindeutig zu beantworten. So sollten bereits kleine Änderungen im Expressionsniveau von Signalproteinen eine große Wirkung entfalten können, wohingegen sich die vergleichbare Regulation eines Enzyms im beteiligten Stoffwechselweg eher gering auswirken sollte. Durch die Wahl des oben angeführten *Cut-Offs* von 100% sollten demnach differentiell expremierte Proteine detektiert werden, die nach Abschätzung des methodischen Fehlers eine biologisch eindeutige Regulation aufweisen.

Die Analyse der beiden Vergleichsexperimente zeigt, dass mit der ICPL-Methode von einem Großteil der identifizierten cytosolischen Proteine Regulationsinformationen erhalten werden können. Im ersten Vergleichsexperiment (Wachstum in Komplexmedium vs. Wachstum in synthetischem Medium) wurden 12% der Proteine von denen Regulationsinformationen gewonnen wurden, als reguliert klassifiziert. Die Anzahl der regulierten Proteine war in beiden Bedingungen nahezu gleich. Im zweiten Vergleichsexperiment (Wachstum in Komplexmedium vs. Wachstum in synthetischem Medium) wurden nur 4.6% der Proteine reguliert. Es zeigte sich, dass besonders im anaeroben phototrophen Wachstum nur 10 Proteine als reguliert identifiziert wurden.

Weiter fiel in beiden Analysen auf, dass nur wenige Proteine hohe Regulationen aufwiesen. Nur 4 Proteine waren im Komplexmedium und 11 Proteine im synthetischen Medium um mehr als 500% reguliert (Regulationen  $> 5$  oder  $< -5$ ). Die geringe Regulation fiel noch stärker beim zweiten Vergleich (aerob vs. phototroph) auf. Hier waren gerade einmal 3 Proteine unter aeroben Wachstumsbedingungen mehr als 500% reguliert. Beim anaeroben phototrophen Wachstum zeigte keines der identifizierten Proteine mehr als 240% Regulation.

Diese Ergebnisse waren einerseits nicht zu erwarten, deckten sich jedoch andererseits mit Transkriptomanalysen der identischen Wachstumsbedingungen (siehe Dissertation Jens Twellmeyer). Insbesondere die geringe Regulation beim Vergleich von aerobem und anaerobem phototrophen Wachstum war überraschend, da es sich bei diesen Wachstumsbedingungen um stark unterschiedliche Lebensbedingungen mit sehr unterschiedlichen bioenergetischen Anforderungen handelt. Dennoch scheint *H. salinarum* auch auf derart extreme Änderungen lediglich mit einem verhältnismäßig geringen Umbau des cytosolischen Proteininventars zu antworten. Dieses Verhalten wird von der Beobachtung weiter gestützt, dass unter der Standardwachstumsbedingung bereits ein Großteil der cytosolischen Proteine in der Zelle expremiert wird (siehe 5.3.1).

Anhand des eher dürftigen Regulationsverhaltens und der hohen Abdeckung an identifizierten Proteinen kann geschlossen werden, dass *H. salinarum* stets alle oder zumindest einen Großteil der Proteine seines Inventars jederzeit herstellt, selbst wenn es diese unter den gegebenen Bedingungen nicht unbedingt benötigt. Dies würde es dem Organismus erlauben, schnell auf Änderungen der Umgebung zu reagieren, ohne zusätzliche Energiemengen aufbringen zu müssen um sein Proteinrepertoire anzupassen.

Es soll an dieser Stelle noch einmal darauf hingewiesen werden, dass, wie im Ergebnisteil dargestellt, mit der ICPL-Methode nicht eindeutig gezeigt werden kann, ob ein Protein exklusiv nur in einer Bedingung der gewählten Zustände vorkommt, d.h. wenn nur ein Isotopenpartner im Massenspektrum detektiert wird. Eine derartige

Situation kann neben der gezeigten Begrenzung des dynamischen Bereichs der ICPL-Methode zusätzlich durch das Auftreten von differentiell regulierter proteolytischer Aktivität entstehen. Derartige Artefakte sollten insbesondere bei einer Vorfraktionierung der Proteinprobe nach Molekulargröße (z.B. 1D SDS-PAGE) anzutreffen sein. Ein Hinweis darauf wäre eine identifizierte Regulation in einer Fraktion, die nicht der theoretischen Masse des regulierten Proteins entspricht, oder entgegengesetzte Regulationen eines Proteins in unterschiedlichen Größenfraktionen.

Die in den ICPL-Analysen als differentiell reguliert klassifizierten Proteine ließen sich zum überwiegenden Teil in den biologischen Kontext einordnen. Dabei wurden zum einen Proteine des Intermediär-, Aminosäure- und Nukleotid-Stoffwechsels gefunden, aber auch unerwartete zelluläre Antworten auf die untersuchten Wachstumsbedingungen aufgedeckt (siehe Peptidimport und Abbau bei aerobem Wachstum in Komplexmedium). Die Cofaktoren Thiamin und Cobalamin scheinen unter optimaler Bedingung, welche schnelles Zellwachstum ermöglicht, nicht in ausreichenden Mengen vorzuliegen, da ihre Biosyntheseproteine hoch reguliert sind. Eine Erklärung bzw. Interpretation der verstärkten reduktiven CO<sub>2</sub>-Fixierung und anschließender Gluconeogenese im Komplexmedium und dem bevorzugten Ablaufen der Glykolyse in synthetischem Medium konnte unter Betrachtung des bioenergetischen Potentials des Organismus (siehe Ferredoxine mit unterschiedlichen Redox-Potentialen) ebenfalls gegeben werden.

Die Annahme einer gegenläufig regulierten Glykolyse und Gluconeogenese beruht auf der gegenläufigen Regulation der Pyruvat Kinase (Glykolyse, synthetisches Medium) bzw. 'pyruvat water dikinase' (Gluconeogenese, Komplexmedium) die die Hin- und Rückreaktion des letzten Schrittes der Glykolyse bzw. Anfangsreaktion der Gluconeogenese katalysieren (siehe Ergebnisteil 4.2.5.2). Für Bakterien aus Viehmägen konnte gezeigt werden, dass bei Wachstum unter hohem Kohlenstoffangebot beide Enzyme einen sogenannten *Futile Cycle* (sinnlose Zyklen) bilden können, um überschüssiges ATP zu verbrauchen, dass bei zu hoher zellulärer Konzentration für die Zelle toxisch werden kann [Russell und Cook, 1995; Russell, 1998]. Die Existenz eines derartigen *Energy Spilling* Zyklus durch die gegenseitige Umwandlung von Pyruvat in Phosphoenolpyruvat, katalysiert durch die eben genannten Enzyme, wurde bereits auch für Archaea diskutiert [Hutchins *et al.*, 2001].

Eine derartige Situation ist für *H. salinarum* weniger wahrscheinlich, da bei phototrophem Wachstum, durch die von der lichtgetriebenen Protonenpumpe Bakteriorhodopsin bewirkte Photosynthese, eine maximale zelluläre ATP-Konzentration vorliegt. Es kommt lediglich zur Lichthemmung der Atmung unter

aeroben phototrophen Wachstumsbedingungen. Demnach ist anzunehmen, dass der Organismus auch mit einem hohen ATP-Level zurechtkommt.

Die Reproduzierbarkeit und Zuverlässigkeit der in der Proteomforschung generierten und publizierten Daten wird in letzter Zeit verstärkt diskutiert [Carr *et al.*, 2004]. Die Wiederholung der durchgeführten Experimente sowie die Abschätzung der Zuverlässigkeit der massenspektrometrisch erhaltenen Proteom-Daten werden zunehmend gefordert. Die Einführung einer standardisierten Datenevaluation wird diesen Ansprüchen Rechnung tragen. In dieser Arbeit wurden die Score-Grenzen zur massenspektrometrischen Identifizierung von Proteinen durch die Verwendung einer *Nonsense* Datenbank festgelegt. Dies sollte das Vorkommen von falsch positiv als auch falsch negativ identifizierten Peptiden und Proteinen in den vorgestellten Datensätzen minimieren. Eine komplette Wiederholung der in dieser Arbeit vorgestellten Versuchsreihen war jedoch aus Zeitgründen nicht möglich. Hierbei ist anzumerken, dass eine derartige Wiederholung nicht den methodischen Fehler, sondern auch die biologische Varianz des untersuchten biologischen Systems berücksichtigen muss. Wie im Ergebnisteil gezeigt konnten für einen Teil der Proteine die Regulationen, auf die sich die biologische Interpretation der erhaltenen Ergebnisse stützt, in unabhängigen Proben reproduzierbar nachgewiesen werden. Auf weitere, nicht wieder identifizierte regulierte Proteine konnte jedoch indirekt über funktionell verwandte identifizierte Regulationen geschlossen werden.

In einer anderen Arbeit haben Baliga und Mitarbeiter mit dem ICAT-Ansatz unterschiedliche Mutanten des *H. salinarum* Labostamms NRC-1 untersucht [Baliga *et al.*, 2002]. Hierbei wurde eine defiziente mit einer überexprimierenden Mutanten des Transkriptionsfaktor *bat*, der an der Regulation von Genen die Proteine der halobakteriellen Purpurchromatophorensäule exprimieren, verglichen. In dieser Arbeit werden die Nachteile des ICAT-Ansatzes deutlich. Es konnten lediglich Regulationsinformationen von 272 Proteinen gewonnen werden. Die Bestimmung der Regulation erfolgte in den meisten Fällen anhand eines einzelnen Peptids.

## 5.3 Das identifizierte Proteininventar

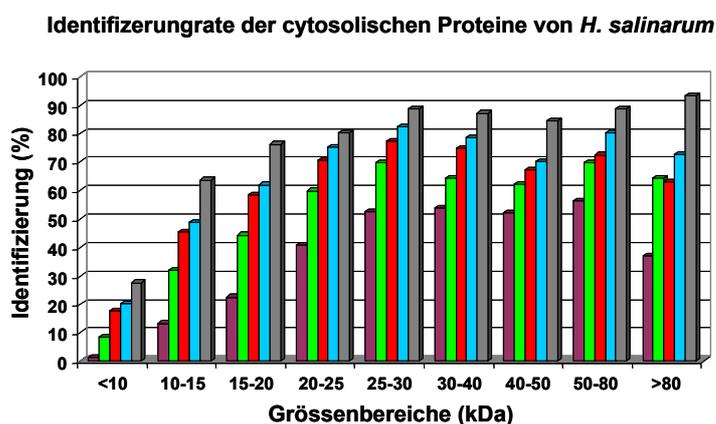
### 5.3.1 Vergleich der unterschiedlichen MS-Technologien

In dieser Arbeit wurden unterschiedliche methodische Ansätze zur Analyse des cytosolischen Proteoms von *H. salinarum* angewendet. Diese sollen im Folgenden bezüglich ihrer Effizienz hinsichtlich der Identifizierung cytosolischer Proteine

diskutiert werden. Ein Vergleich der Ergebnisse ist möglich, da sich die Präparationen der cytosolischen Proteinproben entsprachen.

Im klassischen Proteomikansatz verfolgt wurden die Proteine mittels 2-DE aufgetrennt und durch MALDI-TOF PMF identifiziert. Die Analyse der ICPL-markierten Proteine diente vorrangig der Identifizierung differentiell regulierter Proteine. Die so erhaltenen auf Tandemmassenspektrometrie beruhenden Proteinidentifizierungen können dennoch mit den Ergebnissen der vorherigen Methode verglichen werden. Als weiterer methodischer Ansatz wurden die Peptidfraktionen einer ICPL-Probe mittels LC-ESI MS analysiert. Hierbei kam ein LTQ FT-ICR Massenspektrometer zum Einsatz, welches sich durch hohe Massengenauigkeit und Auflösungsvermögen sowie durch die Möglichkeit, aufeinanderfolgende Fragmentierungen vorzunehmen, auszeichnet.

Wie in 4.1.3 beschrieben wurden bei dem 2-DE MALDI TOF Ansatz, mit dem 801 cytosolische Proteine identifiziert werden konnten, insbesondere Proteine mit geringer molekularer Masse zu einem geringeren Ausmaß identifiziert. Gründe sind zum einen, dass die PMF Methode für kleine Proteine, die nur wenige proteolytische Peptide liefern, schlecht geeignet ist. Der zweite Grund ist, dass das *kleine Proteom* nur zu einem begrenzten Teil in der analysierten Proteinprobe vorliegt, also kleine Proteine spezifisch bei der Probenaufarbeitung verloren gehen. In Abbildung 5-2 sind die mit den unterschiedlichen Methoden erreichten Proteomabdeckungen in Abhängigkeit vom Molekulargewicht dargestellt.



**Abb. 5-2:** Proteomabdeckung in Abhängigkeit unterschiedlicher Molekulargewichtsbereiche. Die prozentualen Identifizierungen, die mittels 2-DE MALDI TOF (lila Balken), LC-MALDI TOF/TOF (grüner Balken), LC-ESI FT-ICR (roter Balken) sowie mit allen drei angewendeten methodischen Ansätze zusammen erreicht werden konnten (türkiser Balken) werden dargestellt. Ebenfalls gezeigt ist die prozentuale Abdeckung aller cytosolischen chromosomalen Proteine (grauer Balken).

Es ist deutlich zu sehen, dass mit auf MS/MS beruhenden Methoden mehr Proteine identifiziert werden können [Wang *et al.* 2004; Chong and Wright, 2005]. Mit der LC-MALDI TOF/TOF Methode wurden insgesamt 1141 cytosolische Proteine identifiziert, auf 2-D Gelen insgesamt 801. Der Zugewinn wirkte sich wie erwartet bei kleinen Proteinen (<20 kDa) am stärksten aus. Die besten Ergebnisse wurden mit der LC-ESI FT-ICR Methode mit 1,346 identifizierten cytosolischen Proteinen erhalten. Hier konnten bis auf den Größenbereich > 80 kDa durchweg die meisten Proteine identifiziert werden.

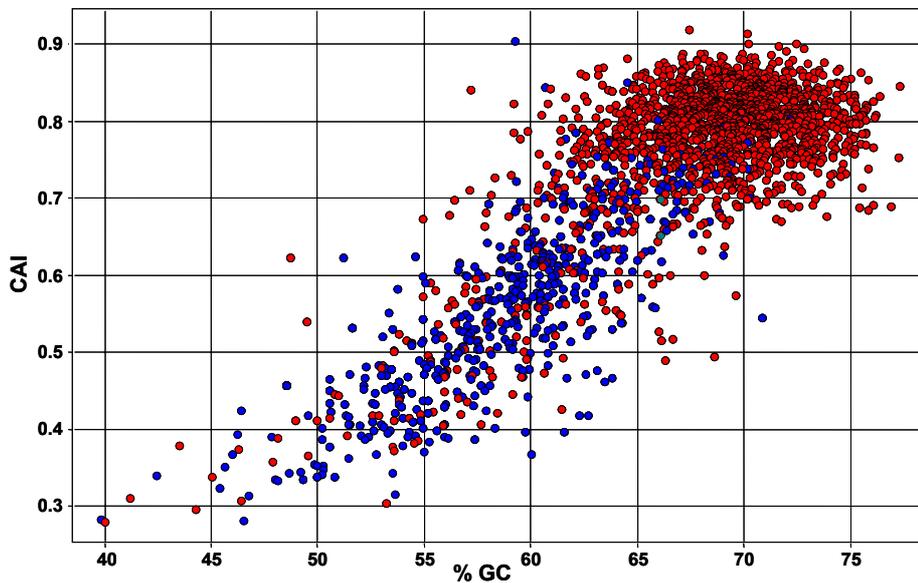
Bei allen drei Methoden sind die kleinen Proteine <15 kDa und insbesondere <10 kDa im Gesamtinventar stetig unterrepräsentiert. Dies untermauert die bereits an den 2-D Gelen für kleine Proteine (siehe Abb. 4-5 und Abb. 4-6 Sektoren b, d, e und g, 14% Gele) gemachte Vermutung, dass ein beträchtlicher Teil des kleinen Proteoms nicht in der analysierten cytosolischen Proteinprobe enthalten ist. Allerdings ist grundsätzlich die Frage zu stellen, inwieweit diese kleinen ORFs überhaupt expremiert werden. Die Tatsache, dass für die meisten dieser kleinen Proteine keine Homologen gefunden werden konnten, stellt dies zusätzlich in Frage. Weitere Arbeiten zur spezifischen Anreicherung und Identifizierung des kleinen Proteoms die derzeit vorgenommen werden, sollen diese Frage näher beantworten.

### 5.3.2 Abschätzung des Expressionsniveaus

Aus Abbildung 5-2 geht noch einmal hervor, dass die Identifizierungsrate für cytosolische chromosomale Proteine zusätzlich erhöht ist (grauer Balken). Insgesamt konnten in dieser Arbeit 76% aller cytosolischen chromosomalen Proteine, allerdings nur 32% aller auf Plasmiden codierten Proteine identifiziert werden. Es bestätigte sich demnach das Bild, das bereits bei der Analyse der cytosolischen Proteine mittels 2-DE und MALDI-TOF PMF erhalten wurde (siehe 4.1). Diese Diskrepanz kann dadurch erklärt werden, dass die Plasmidproteine nur in einem geringeren Ausmaß expremiert werden und so entweder in geringen Mengen oder gar nicht unter den analysierten Wachstumsbedingungen in der Zelle vorliegen.

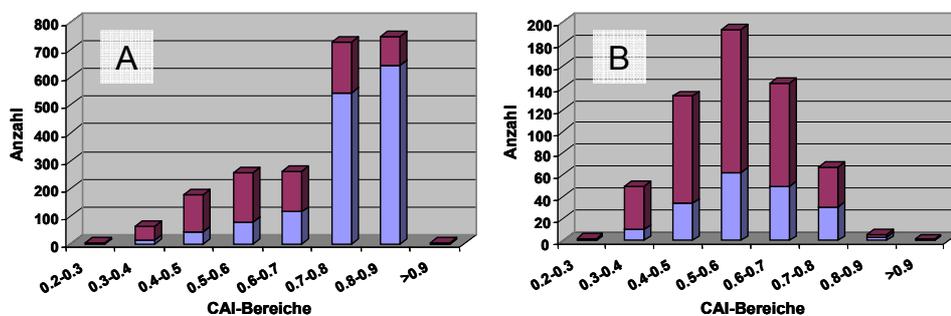
Eine Abschätzung der Expressionsrate kann durch den *Codon Adaptation Index* (CAI) vorgenommen werden. Dieser setzt die Auswahl der für die Aminosäuren eines Proteins benutzten Codons in das Verhältnis zu den Codons, die in dem gesamten Organismus am häufigsten gebraucht werden [Karlin *et al.*, 2005]. Dabei wird angenommen, dass ein Gen umso stärker expremiert wird, je mehr häufig verwendete Codons für die jeweiligen Aminosäuren benutzt werden. Ein CAI von 1 entspricht dabei einem Gen, welches ausschließlich die häufigsten Codons verwendet. Abbildung 5-3

zeigt die berechneten CAI-Werte für alle Gene der cytosolischen Proteine von *H. salinarum* in Abhängigkeit von ihren jeweiligen GC-Gehalten.



**Abb. 5-3:** CAI in Abhängigkeit von den zugehörigen GC-Gehalten aller für cytosolische Proteine codierenden Gene. Chromosomale Gene sind in rot, Gene auf Plasmiden in blau dargestellt.

Es wird deutlich, dass mit einem abnehmenden GC-Gehalt ab ungefähr 65% auch die CAI-Werte abnehmen und ein direkter Zusammenhang zwischen GC-Gehalt und CAI-Werten besteht. Die Mehrheit der chromosomalen Gene besitzen CAI-Werte zwischen 0.7 und 0.9, was einer mittleren bis starken Expression entspricht. Der Großteil der Plasmidgene hingegen weist CAI-Werte zwischen 0.28 und 0.7 auf, was nach der CAI-Theorie überwiegend geringer bis mittlerer Expression entsprechen würde. Dies wird durch die vorliegenden Proteomstudien unterstützt. Abbildung 5-4 A und B zeigt die Identifizierungsraten cytosolischer chromosomaler sowie Plasmidproteine in unterschiedlichen CAI-Bereichen.



**Abb. 5-4:** Identifizierungsrate chromosomaler (A) und Plasmidproteine (B) in Abhängigkeit unterschiedlicher CAI-Bereiche. Die Balken repräsentieren die Anzahl aller cytosolischen Proteine in dem jeweiligen CAI-Bereich. Der blaue Anteil entspricht identifizierten, der lila Anteil den nicht identifizierten Proteinen.

Aus den Tabellen wird deutlich, dass der Anteil an identifizierten Proteinen mit dem CAI-Wert steigt. Dies lässt den Schluss zu, dass auch in *H. salinarum* der CAI-Werte mit der Expressionrate der Gene zu korreliert.

Durch die Verwendung der unterschiedlichen analytischen Methoden bei den Proteomanalysen an *H. salinarum* konnten insgesamt 1441 der 2239 cytosolischen Proteine identifiziert werden (64.4%), was einer der bisher größten erzielten Proteomabdeckungen im Vergleich zu anderen Arbeiten [Peng *et al.* 2003; Zhu *et al.*, 2004; Wang *et al.*, 2004; Chong und Wright, 2005] entspricht. Tabelle 5-1 fasst die erzielten Proteinidentifizierungen mit den unterschiedlichen Methoden zusammen.

**Tab. 5-1:** Mit unterschiedlichen technologischen Ansätzen identifizierte Proteine. Gezeigt ist die Anzahl der gesamten (obere Reihe) sowie der cytosolischen Proteine, die mittels 2-DE (1), LC MALDI-TOF/TOF (2), LC FT-ICR (3) und auf Tandem MS beruhenden Methoden ((2) +(3); (4)) identifiziert werden konnten. In Klammern ist jeweils die Anzahl an Proteinen angegeben, die mit der jeweiligen Methode und den 2-DE Daten (1) identifiziert werden konnte.

	2-DE <sup>1</sup>	LC MALDI-TOF/TOF <sup>2</sup>	LC FT-ICR <sup>3</sup>	LC MS/MS <sup>4</sup>
<b>identifizierte Proteine</b>	<b>803</b>	<b>1195 (1265)</b>	<b>1414 (1453)</b>	<b>1495 (1522)</b>
<b>identifizierte cyt. Proteine</b>	<b>802</b>	<b>1141 (1210)</b>	<b>1346 (1385)</b>	<b>1414 (1441)</b>

Wenn man berücksichtigt, dass eine beträchtliche Anzahl der kleinen sowie der auf Plasmiden codierten oder membranassoziierten Proteine nicht in der cytosolischen Standardprobe enthalten sein sollte, kann davon ausgegangen werden, dass der überwiegende Teil der in der cytosolischen Standardprobe vorhandenen Proteine identifiziert werden konnte. Überraschenderweise wurden trotz der Abreicherung von Membranbestandteilen 81 Membranproteine identifiziert. Dies kann durch Membrankontaminationen erklärt werden, was die Reinheit der Cytosolfraktion in Frage stellen würde. Bei diesen Membranproteinen kann es sich aber auch um natürliche Abbau- oder auch Aufbauprodukte handeln, die im Cytosol der Zelle zu finden sind. Nicht überraschend hingegen war die Tatsache, dass alleine 80 der Membranproteine nur mit LC-MS/MS Techniken identifiziert wurden.

Es bleibt abzuwarten, inwieweit das cytosolische Proteininventar durch die Verwendung neuerer, in Sensitivität und Durchsatz verbesserter MS Technologien, die es erlauben, noch tiefer in komplexe Protein bzw. Peptidgemische einzudringen, komplettiert werden kann. Die Analyse weiterer Subproteome wie Membranfraktionen oder das kleine Proteom wird ebenfalls zur weiteren Inventarisierung des Gesamtproteoms von *H. salinarum* beitragen und es der Komplettierung nahe bringen.

Die Daten der unterschiedlichen Analysen werden derzeit im HALOLEX System zugänglich gemacht.

## 5.4 Analyse posttranslationaler Modifikationen

Neben den vergleichenden Analysen ist eine weitere wichtige Aufgabe der Proteomforschung die Analyse posttranslationaler Modifikationen. Hierbei lag der Schwerpunkt dieser Arbeit in der Identifizierung von Proteinphosphorylierungen. Die Analyse konzentrierte sich auch hier vornehmlich auf die Untersuchung cytosolischer Proteine der späten logarithmischen Wachstumsphase.

### 5.4.1 Analyse phosphorylierter Proteine

Die Regulation komplexer Signaltransduktionskaskaden oder die Steuerung von Proteinaktivitäten durch reversible Phosphorylierung wurde lange Zeit ausschließlich höheren Lebensformen zugeschrieben. Man ging davon aus, dass Proteinaktivitäten in Prokaryoten vornehmlich durch allosterische Effekte reguliert werden. Trotz der Entdeckung von Serin, Threonin oder gar Tyrosin Phosphorylierungen, denen teilweise auch regulative Funktionen zugesprochen werden können, ist neben den Proteinphosphorylierungen, die im Zuge der Signalübertragung über Zwei-Komponentensysteme zur Steuerung der Taxis in bakteriellen oder archaealen Lebensformen auftreten (siehe 1.3), bis dato nur wenig bekannt [Kennelly, 2002]. Kürzlich wurde für *Bacillus subtilis* gezeigt, dass eine in Prokaryoten auftretende Hyperphosphorylierung von Proteinen als Degradationssignal eine Rolle spielen könnte [Rosen *et al.*, 2004].

Bei den in dieser Arbeit durchgeführten Analysen handelt es sich um die erste proteomweite Untersuchung zur Identifizierung von Serin, Threonin oder Tyrosin Phosphorylierungen in *H. salinarum* und Archaeobakterien im Allgemeinen.

Die im Genom annotierten Proteinkinasen waren zwar als erster Hinweis zu verstehen, dass Phosphorylierungen halobakterieller Proteine auftreten können, dennoch waren die Homologien dieser Gene zu bekannten Kinasen nicht immer eindeutig. Es konnten auch keine Proteinphosphatasen gefunden werden, welche die charakteristische Reversibilität dieses Regulationsvorgangs gewährleisten. Das Auftreten von posttranslationalen Modifikationen in *H. salinarum* wurde jedoch bereits durch das Identifizieren von Proteinen in unterschiedlichen Proteinspots auf 2-D Gelen untermauert.

Die immunochemische Detektion des gelelektrophoretisch aufgetrennten cytosolischen Proteoms mit Phosphoserin- bzw. Phosphothreonin spezifischen Antikörpern zeigte jedoch klar, dass Serrin- oder Threonin Phosphorylierungen in *H. salinarum* auftreten. Durch die metabolische Markierung mit radioaktivem Phosphat und anschließender zweidimensionaler Auftrennung der markierten Proteine konnten einige phosphattragende Proteine identifiziert werden. Das Ausmaß an Proteinphosphorylierungen war dennoch eher gering (Durchschnittlich 30 Spots). Lediglich 7 Proteine wurden identifiziert, die ein radioaktives Signal im Gel hinterließen. In einer vergleichbaren Arbeit zur Analyse des Phosphoproteoms des Gram-positiven Bodenbakteriums *Corynebacterium glutanicum* wurden mittels radioaktiver Markierung und immunochemischen Nachweis bis zu 120 Proteinspots auf 2-D Gelen gefunden, die Phosphoproteinen entsprachen. Von ihnen konnten 41 Proteine durch MALDI-TOF PMF Analyse identifiziert [Bendt *et al.*, 2003]. Auch im Genom von *Corynebacterium glutanicum* sind nur wenige Proteinkinasen codiert. Wie in *H. salinarum* codieren lediglich 0.2% aller Gene für mögliche Proteinkinasen im Gegensatz zu ca. 5% in Eukaryoten [Blom *et al.*, 1998; Johnson und Hunter, 2005]. Dennoch war der Phosphorylierungsgrad in dem Gram-positiven Bodenbakterium deutlich ausgeprägter als in *H. salinarum*. Daraus könnte geschlossen werden, dass es sich bei einigen der annotierten Proteinkinasen nicht um solche handelt.

Aus den in dieser Arbeit mittels metabolischer Markierung und 2-DE erhobenen Daten konnte nicht abgeleitet werden, um welche Art der Phosphorylierung (Ser- Thr oder Tyr-Phosphorylierung) es sich handelte.

#### 5.4.2 Identifizierung von Phosphorylierungsstellen

Für die Analyse von Phosphorylierungsstellen wurde zum einen die cytosolische Proteinprobe mittels 1-D SDS PAGE vorfraktioniert und aus den entstandenen Peptidfraktionen die in ihnen enthaltenen Phosphopeptide affinitätschromatographisch angereichert. Zum anderen wurde eine gelfreie Methode zur Phosphopeptidanreicherung eingesetzt.

Mit beiden Methoden konnten lediglich drei Phosphorylierungsstellen halobakterieller Proteine identifiziert werden. Dabei handelte es sich ausschließlich um metabolische Intermediate, aber nicht um regulative Proteinphosphorylierungen wie man sie im Rahmen von Signaltransduktionskaskaden höherer Organismen findet. Durch diese Analysen konnte nur für ein Protein, das zuvor durch radioaktive Markierung als phosphattragend auf 2-D Gelen identifiziert worden war, die Phosphorylierungsstelle identifiziert werden. Lediglich für ein weiteres Protein auf den *radioaktiven* Gelen konnten die Phosphorylierungsstellen durch einen indirekten Vergleich mit den beiden

anderen identifizierten Phosphoproteinen aufgrund der hohen Sequenzhomologien und gleicher vorhergesagter biologischer Funktion zugewiesen werden.

Die Anzahl phosphorylierter Proteine konnte demnach nicht entscheidend erweitert werden. Weiter ist die Frage zu stellen, weshalb nicht weitere Phosphorylierungsstellen der radioaktiv markierten Proteine identifiziert werden konnten. Hierfür gibt es mehrere Gründe: Zum einen könnten die tryptischen Phosphopeptide weiterer Phosphoproteine entweder zu groß, aber auch zu klein für die gewählten massenspektrometrischen Methoden sein und sich so der Detektion entziehen. Um dies auszuschließen wäre der Gebrauch von alternativen Proteasen denkbar, wodurch Proteine unterschiedlich zum verwendeten Trypsin gespalten werden. Dann wäre allerdings die Anreicherung von Phosphopeptiden über SCX nicht mehr anzuwenden, da diese auf tryptische Phosphopeptide angewiesen ist, die C-terminal eine weitere positive Ladung tragen (Trypsin spaltet nach Arg und Lys). Eine weitere Möglichkeit wäre, dass das radioaktive Phosphat nicht direkt an das Protein sondern in Form eines anderen phosphattragenden Kofaktors gebunden ist, und so ein Signal auf dem 2-D Gel ergeben würde. Bei den beschriebenen Analysen wurde jedoch ausschließlich nach phosphorylierten Serinen, Threoninen oder Tyrosinen (Einstellungen in Mascot) gesucht. Als dritte Möglichkeit ist in Betracht zu ziehen, dass die metabolische Markierung auch geringste Mengen an phosphorylierten Proteinen in den 2-D Gelen sichtbar machen kann. Die massenspektrometrische Analyse ist jedoch immer einem Sensitivitätslimit unterworfen, welches nicht ausreicht, derartige Spuren zu detektieren und so nachzuweisen.

Mit den in dieser Arbeit angewendeten Methoden zur spezifischen Anreicherung von Phosphopeptiden mittels SCX-Chromatographie und anschließender Isolierung über TiO<sub>2</sub> Säulen (siehe „Material und Methoden“) konnten in vergleichbaren Arbeiten 700 Phosphorylierungsstellen aus Hefe identifiziert werden [Gruhler *et al.*, 2005]. Diese Methode stellt somit bis dato den Ansatz in der Literatur dar, mit dem die meisten Phosphorylierungsstellen aus verschiedenen Organismen identifiziert werden konnten. In den letzten Jahren wurden unterschiedliche Methoden zur spezifischen Anreicherung und anschließender massenspektrometrischen Analyse von Phosphopeptiden entwickelt. Durch chemische [Zhou *et al.*, 2001; Oda *et al.*, 2001; Brittain *et al.*, 2005] oder enzymatische [Gaevert *et al.*, 2005] Modifizierung der Phosphorylierungsstelle und anschließender affinitätschromatographischer bzw. auf differentiellen Chromatographietechniken beruhender Isolierung konnten ebenfalls gute Erfolge erzielt werden.

Es ist jedoch zu bezweifeln, dass mit diesen eben genannten Methoden eine weitaus größere Zahl an Phosphorylierungsstellen von *H. salinarum* identifiziert werden könnten.

### 5.4.3 Identifizierung N-terminaler Acetylierungen

Die N-terminale Acetylierung von Proteinen tritt in allen drei Domänen des Lebens auf. Während der überwiegende Teil cytoplasmatischer Proteine (80 - 90%) bei höheren Organismen (z.B. Säugern) in acetylierter Form vorliegt, gibt es über das Auftreten dieser Proteinmodifikation in Bakterien und Archaea nur wenig experimentelle Informationen [Polevoda und Sherman, 2003]. Die in dieser Arbeit erhobenen Daten stellen somit die erste globale Analyse N-terminal acetylierter Proteine in Prokaryoten dar. Es konnte gezeigt werden, dass die N-terminale Acetylierung in *H. salinarum* bevorzugt, aber nicht ausschließlich nach Abspaltung des Initiator-methionins erfolgt. Es zeigte sich eine starke Präferenz für Serin gefolgt von Alanin als dem Initiator-methionin folgende Aminosäure. Dies entspricht der Spezifität der N-Acetyltransferase A (NatA) der Hefe, die noch zwei weitere Formen dieses Enzyms besitzt (NatB und NatC), welche auch das Initiator-methionin acetylieren. Derartige Acetylierungen konnten auch in *H. salinarum* gefunden werden. Es wurden demnach vergleichbare Spezifitäten der Acetylierung wie in Eukaryoten beschrieben gefunden. Es zeigte sich, dass die Komplexität der Proteinacetylierung in *H. salinarum* zwar geringer zu sein scheint als in Prokaryoten, sie aber dennoch stärker anzutreffen ist als bisher für Prokaryoten angenommen [Polevoda und Sherman, 2003]. Alleine 544 Proteine in *H. salinarum* besitzen als zweite Aminosäure (dem Initiator-methionin folgend) ein Serin. Man kann also davon ausgehen, dass neben den identifizierten noch weitere Proteine in acetylierter Form vorliegen.

Die biologische Funktion der N-terminalen Acetylierung ist dennoch nicht verstanden. Neben Hinweisen, dass diese Modifikation eukaryotische Proteine vor Degradation schützt (N-end rule), gibt es nur wenige Beispiele, dass Proteine die Acetylierung für ihre Aktivität (Actin in *Drosophila*) oder Stabilität (Glucose Dehydrogenase in *Neurospora crassa*) benötigen [Perrier *et al.*, 2005; Polevoda und Sherman, 2003]. Es ist demnach schwierig, etwas über die Funktion der N-terminalen Acetylierung in *H. salinarum* auszusagen.

Die Kenntnis der identifizierten N-terminalen Peptide konnte jedoch zur Überprüfung der vorhergesagten Startcodons genutzt werden. Diese ist in mikrobiellen und speziell in archaealen Genomen oftmals problematisch, da hier die Bindungsstellen des Ribosoms (Shine-Dalgarno Sequenzen) am Genstart nur gering konserviert sind [Sartorius-Neef und Pfeiffer, 2004; Torarinsson *et al.*, 2005; siehe Dissertation

Michaela Falb]. Die identifizierten N-terminalen Acetylierungen erlaubten die Verifizierung der vorhergesagten Startcodons und ergänzen so weitere Arbeiten zur spezifischen Identifizierung N-terminaler Peptide mittels COFRADIC [Gevaert *et al.*, 2003; Gevaert *et al.*, 2005] zur Analyse der Startcodon Vorhersage. Auf diese Methoden soll jedoch in dieser Arbeit nicht näher eingegangen werden.

## 6 Ausblick

Die Proteomforschung ist stark von den methodischen Fortschritten der letzten Jahre geprägt. Insbesondere die rapide fortschreitende Entwicklung von immer sensitiveren und schnelleren Massenspektrometern trägt dazu bei, nicht nur die Spitze des Eisberges, die hochabundanten Proteine eines Proteoms, zu analysieren, sondern immer tiefer in dieses vorzudringen. Gesamte Proteome als Ganzes zu analysieren nimmt immer mehr Gestalt an.

Die zweidimensionale Gelelektrophorese als traditionelle Methode der Proteomanalytik hat sich in den letzten Jahren steigender Beliebtheit erfreut. Bis heute ist sie die Methode mit dem höchsten Auflösungsvermögen von komplexen Proteingemischen. Durch die Einführung der immobilisierten pH Gradienten (IPGs) sowie von *Ultra-Zoom* Gelen konnten Reproduzierbarkeit und Auflösungsvermögen weiter gesteigert werden. Die Entwicklung automatisierter Analyseverfahren macht die Bearbeitung der 2-D Gele im Hochdurchsatzverfahren möglich. In dieser Arbeit konnte jedoch gezeigt werden, dass modernere methodische Ansätze, die lediglich auf einer groben Vorfraktionierung der Proteinprobe und anschließender Analyse komplexer Peptidgemische beruhen, in kürzerer Zeit mehr Ergebnisse liefern. Die Kopplung chromatographischer Methoden mit modernen Tandem Massenspektrometern ermöglicht die schnellste Identifizierung von Proteinen.

Durch die Einführung von Isotopenmarkierungen können vergleichende Proteomanalysen schnell und präzise durchgeführt werden. Eine Erweiterung dieser Strategie durch die metabolische Markierung mit isotopenmarkierten Aminosäuren sollte die Abdeckung und Zuverlässigkeit der Identifizierung differentiell regulierter Proteine zusätzlich verbessern. Ein weiterer wichtiger Schritt der Proteomanalyse wird es sein, nicht nur das Proteininventar und differentiell regulierte Proteine, sondern insbesondere Interaktionen von Proteinkomplexen zu erforschen.

Insbesondere cytosolische Proteome von mikrobiellen Organismen lassen sich mit den beschriebenen Methoden in ihrer Gesamtheit darstellen und analysieren. Eine Erweiterung auf schwerer zu bearbeitende Proteine wie die der Zellmembran ist unerlässlich für das Verständnis wichtiger zellulärer Prozesse und wird derzeit in anderen Arbeiten angegangen.

Die Analyse posttranslationaler Modifikationen wird ebenfalls weiter zum Verständnis des Organismus beitragen. Inwieweit jedoch Proteinphosphorylierungen dabei eine entscheidende Rolle spielen werden, muß abgewartet werden.

## 7 Zusammenfassung

Im Rahmen dieser Arbeit wurden unterschiedliche Proteomstudien an *Halobacterium salinarum* durchgeführt, die sich generell in drei unabhängige Themenblöcke unterteilen lassen.

Zunächst erfolgte die Erfassung des cytosolischen Proteoms auf standardisierten 2-D Gelen. Hierfür wurde die Auftrennung auf sogenannten *Zoom-Gelen* aufgrund der engen *pI*-Verteilung halobakterieller Proteine etabliert. Die erschöpfende Analyse der so aufgetrennten Proteine wurde mit der *peptide mass fingerprinting* Methode im Hochdurchsatzverfahren durchgeführt. So konnten so 40% des zugänglichen cytosolischen Proteininventars identifiziert und auf Referenz 2-D Gelen kartiert werden. Des Weiteren konnten die massenspektrometrischen Proteinidentifizierungen zur intensiven Verbesserung der in vielen Fällen nicht eindeutigen Genomannotation herangezogen werden. Durch die Korrelationsanalyse von experimentellen und theoretischen *pI*-Werten konnten weitere Annotationsfehler behoben werden. Das Zusammenspiel von Genomik und Proteomik erlaubte die Identifizierung eines Bereichs ehemaliger Plasmid DNA, die in das Chromosom eingewandert ist. Weiter konnte für ein Plasmid gezeigt werden, dass dieser ehemals chromosomale DNA und somit eine Reihe wichtiger und essentieller Gene enthält und deshalb eher als zweites Chromosom angesehen werden sollte.

Nach der Inventarisierung folgten Analysen zur differentiellen Expression cytosolischer Proteine in Abhängigkeit unterschiedlicher Wachstumszustände. Hierbei wurde das aerobe Wachstum unter Standardbedingungen mit Wachstum unter Mangelbedingungen wie in synthetischem Medium oder unter anaeroben phototrophen Bedingungen verglichen. Bei der Analyse dieser Zustände mittels 2-DE konnten nur wenig regulierte Proteine identifiziert werden. Generell stellte sich die 2-DE aufgrund schwankender Reproduzierbarkeit und des geringen dynamischen Bereichs der gewählten Färbemethode für diese vergleichenden Analysen als wenig geeignet heraus. Die Verwendung isotope markierter Sonden erlaubte hingegen einen detaillierten Einblick in die Regulationen auf Proteomebene unter besagten Bedingungen. Mit dem methodischen Ansatz der 1D-SDS PAGE LC-MALDI-TOF/TOF Analyse konnten von einem beträchtlichen Teil (ca. 40%) des cytosolischen Proteoms Regulationsinformationen erhalten werden. Auch hier zeigte sich, dass unter anaeroben phototrophen Bedingungen überraschend wenige Proteine reguliert werden müssen, damit sich der Organismus auf diese Lebensbedingung einstellen kann. Wachstum unter

Nährstoffmangel in synthetischem Medium induziert überwiegend Proteine der Aminosäure- und Nukleotid-Biosynthese sowie Proteine, die am Proteinschutz und der Glykolyse beteiligt sind. Unter Standardwachstum wurden Proteine als induziert identifiziert, die bei nicht ausreichendem Vorhandensein wachstumslimitierend wirken würden. Dies waren hauptsächlich Enzyme der Thiamin und Cobalamin Biosynthese sowie des Peptid-Transportes und -Abbaues. Auch die CO<sub>2</sub>-Fixierung und die nachfolgende Gluconeogenese sind induziert. Durch die Verwendung unterschiedlicher Proteomik Ansätze konnte am Ende der Großteil der cytosolischen Proteine von *H. salinarum* identifiziert werden. Es scheinen nahezu alle chromosomalen Proteine auch unter Standard Wachstumsbedingungen expremiert zu werden, was starke Regulationen der Proteinkonzentrationen in Abhängigkeit unterschiedlicher Lebensbedingungen nicht notwendig macht. Der Organismus kann sich so schnell ohne großen energetischen Aufwand an wechselnde Lebensbedingungen anpassen.

Im letzten Teil dieser Arbeit wurde das Ausmaß der Proteinphosphorylierung in *H. salinarum* untersucht. Hierbei konnten zunächst über radioaktive *in vivo* Markierung phosphattragende Proteine identifiziert werden. Ebenfalls gelang es, die ersten halobakteriellen Serin und Threonin Phosphorylierungsstellen zu identifizieren. Generell zeigte sich jedoch, dass das Ausmaß der Proteinphosphorylierung in *H. salinarum* eher gering ist. Bei den identifizierten Phosphorylierungen handelte es sich ausschließlich um katalytische Zwischenprodukte. Für die Existenz von regulativen Ser/Thr oder Tyr Proteinphosphorylierungen waren keine eindeutigen Hinweise zu finden. Weiter gelang es, eine Vielzahl von N-terminal acetylierten Proteinen als eine weitere Form der posttranslationalen Modifikationen zu identifizieren.

## 8 Referenzen

1. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. Nature, 2003. **422**(6928): p. 198-207.
2. Alam, M., *et al.*, *Methyl-accepting taxis proteins in Halobacterium halobium*. Embo J, 1989. **8**(2): p. 631-9.
3. Allers, T. and M. Mevarech, *Archaeal genetics - the third way*. Nat Rev Genet, 2005. **6**(1): p. 58-73.
4. Altschul, S.F., *et al.*, *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.
5. Amster, I.J., *Fourier transform mass spectrometry*. Journal of Mass Spectrometry, 1996. **31**(12): p. 1325-1337.
6. Baliga, N.S., *et al.*, *Coordinate regulation of energy transduction modules in Halobacterium sp analyzed by a global systems approach*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(23): p. 14913-14918.
7. Ballif, B.A., *et al.*, *Phosphoproteomic analysis of the developing mouse brain*. Mol Cell Proteomics, 2004. **3**(11): p. 1093-101.
8. Beausoleil, S.A., *et al.*, *Large-scale characterization of HeLa cell nuclear phosphoproteins*. Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12130-5.
9. Bendt, A.K., *et al.*, *Towards a phosphoproteome map of Corynebacterium glutamicum*. Proteomics, 2003. **3**(8): p. 1637-1646.
10. Berggren, K., *et al.*, *Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex*. Electrophoresis, 2000. **21**(12): p. 2509-2521.
11. Berggren, K.N., *et al.*, *An improved formulation of SYPRO Ruby protein gel stain: Comparison with the original formulation and with a ruthenium II tris (bathophenanthroline disulfonate) formulation*. Proteomics, 2002. **2**(5): p. 486-498.
12. Berman, K.M. and M. Cohn, *Phosphoenolpyruvate synthetase. Partial reactions studied with adenosine triphosphate analogues and the inorganic phosphate-H<sub>2</sub>O exchange reaction*. J Biol Chem, 1970. **245**(20): p. 5319-25.

13. Bjellqvist, B., *et al.*, *Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications*. J Biochem Biophys Methods, 1982. **6**(4): p. 317-39.
14. Blom, N., A. Kreegipuu, and S. Brunak, *PhosphoBase: a database of phosphorylation sites*. Nucleic Acids Research, 1998. **26**(1): p. 382-386.
15. Blumer, K.J. and G.L. Johnson, *Diversity in function and regulation of MAP kinase pathways*. Trends Biochem Sci, 1994. **19**(6): p. 236-40.
16. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-54.
17. Brittain, S.M., *et al.*, *Enrichment and analysis of peptide subsets using fluoruous affinity tags and mass spectrometry*. Nature Biotechnology, 2005. **23**(4): p. 463-468.
18. Britton, K.L., *et al.*, *Insights into the molecular basis of salt tolerance from the study of glutamate dehydrogenase from Halobacterium salinarum*. Journal of Biological Chemistry, 1998. **273**(15): p. 9023-9030.
19. Brunet, S., *et al.*, *Organelle proteomics: looking at less to see more*. Trends Cell Biol, 2003. **13**(12): p. 629-38.
20. Buttner, K., *et al.*, *A comprehensive two-dimensional map of cytosolic proteins of Bacillus subtilis*. Electrophoresis, 2001. **22**(14): p. 2908-2935.
21. Carbone, A., A. Zinovyev, and F. Kepes, *Codon adaptation index as a measure of dominating codon bias*. Bioinformatics, 2003. **19**(16): p. 2005-2015.
22. Carr, S., *et al.*, *The need for guidelines in publication of peptide and protein identification data - Working group on publication guidelines for peptide and protein identification data*. Molecular & Cellular Proteomics, 2004. **3**(6): p. 531-533.
23. Chakraborty, A. and F.E. Regnier, *Global internal standard technology for comparative proteomics*. J Chromatogr A, 2002. **949**(1-2): p. 173-84.
24. Chen, H.S., *et al.*, *High-speed, high-resolution monolithic capillary LC-MALDI MS using an off-line continuous deposition interface for proteomic analysis*. Anal Chem, 2005. **77**(8): p. 2323-31.
25. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole-time-of-flight mass spectrometry*. J Mass Spectrom, 2001. **36**(8): p. 849-65.
26. Chertov, O., *et al.*, *Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry*. Proteomics, 2004. **4**(4): p. 1195-203.

27. Cho, C.W., *et al.*, *Improvement of the two-dimensional gel electrophoresis analysis for the proteome study of Halobacterium salinarum*. *Proteomics*, 2003. **3**(12): p. 2325-2329.
28. Chong, P.K. and P.C. Wright, *Identification and Characterization of the Sulfolobus solfataricus P2 Proteome*. *J Proteome Res*, 2005. **4**(5): p. 1789-98.
29. Delcher, A.L., *et al.*, *Improved microbial gene identification with GLIMMER*. *Nucleic Acids Research*, 1999. **27**(23): p. 4636-4641.
30. Desiere, F., *et al.*, *Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry*. *Genome Biology*, 2005. **6**(1): p. -.
31. Dowsey, A.W., M.J. Dunn, and G.Z. Yang, *The role of bioinformatics in two-dimensional gel electrophoresis*. *Proteomics*, 2003. **3**(8): p. 1567-96.
32. Ducret, A., *et al.*, *High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry*. *Protein Sci*, 1998. **7**(3): p. 706-19.
33. Ericson, C., *et al.*, *An automated noncontact deposition interface for liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry*. *Anal Chem*, 2003. **75**(10): p. 2309-15.
34. Falb, M., *et al.*, *Living with two extremes: Conclusions from the genome sequence of Natronomonas pharaonis*. *Genome Res*, 2005. **15**(10): p. 1336-43.
35. Fenn, J.B., *Electrospray wings for molecular elephants (Nobel lecture)*. *Angew Chem Int Ed Engl*, 2003. **42**(33): p. 3871-94.
36. Fenn, J.B., *et al.*, *Electrospray ionization for mass spectrometry of large biomolecules*. *Science*, 1989. **246**(4926): p. 64-71.
37. Ficarro, S.B., *et al.*, *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae*. *Nat Biotechnol*, 2002. **20**(3): p. 301-5.
38. Friedman, D.B., *et al.*, *Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry*. *Proteomics*, 2004. **4**(3): p. 793-811.
39. Frishman, D., *et al.*, *Combining diverse evidence for gene recognition in completely sequenced bacterial genomes*. *Nucleic Acids Research*, 1998. **26**(12): p. 2941-2947.
40. Garcia-Rizo, C., *et al.*, *Improvement of proteomic data analysis and its application to Halobacterium salinarum*. submitted, 2003.
41. Garnak, M. and H.C. Reeves, *Phosphorylation of Isocitrate Dehydrogenase of Escherichia-Coli*. *Science*, 1979. **203**(4385): p. 1111-1112.

42. Gasteiger, E., *et al.*, *ExPASy: the proteomics server for in-depth protein knowledge and analysis*. *Nucleic Acids Research*, 2003. **31**(13): p. 3784-3788.
43. Gevaert, K., *et al.*, *Diagonal reverse-phase chromatography applications in peptide-centric proteomics: Ahead of catalogue-omics?* *Anal Biochem*, 2005. **345**(1): p. 18-29.
44. Gevaert, K., *et al.*, *Global phosphoproteome analysis on human HepG2 hepatocytes using reversed-phase diagonal LC*. *Proteomics*, 2005. **5**(14): p. 3589-99.
45. Gevaert, K., *et al.*, *Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis: identification of more than 800 Escherichia coli proteins*. *Mol Cell Proteomics*, 2002. **1**(11): p. 896-903.
46. Gevaert, K. and J. Vandekerckhove, *Protein identification methods in proteomics*. *Electrophoresis*, 2000. **21**(6): p. 1145-1154.
47. Gharahdaghi, F., *et al.*, *Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: A method for the removal of silver ions to enhance sensitivity*. *Electrophoresis*, 1999. **20**(3): p. 601-605.
48. Giavalisco, P., *et al.*, *Proteome analysis of Arabidopsis thaliana by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry*. *Proteomics*, 2005. **5**(7): p. 1902-13.
49. Gobom, J., *et al.*, *alpha-cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics*. *Analytical Chemistry*, 2001. **73**(3): p. 434-438.
50. Goo, Y.A., *et al.*, *Proteomic analysis of an extreme halophilic archaeon, Halobacterium sp. NRC-1*. *Molecular & Cellular Proteomics*, 2003. **2**(8): p. 506-524.
51. Goodlett, D.R., *et al.*, *Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation*. *Rapid Commun Mass Spectrom*, 2001. **15**(14): p. 1214-21.
52. Gorg, A., *et al.*, *2-Dimensional Polyacrylamide-Gel Electrophoresis with Immobilized Ph Gradients in the First Dimension (Ipg-Dalt) - the State-of-the-Art and the Controversy of Vertical Versus Horizontal Systems*. *Electrophoresis*, 1995. **16**(7): p. 1079-1086.
53. Gorg, A., *et al.*, *The current state of two-dimensional electrophoresis with immobilized pH gradients*. *Electrophoresis*, 2000. **21**(6): p. 1037-1053.
54. Grabowski, B. and Z. Kelman, *Archeal DNA replication: eukaryal proteins in a bacterial context*. *Annu Rev Microbiol*, 2003. **57**: p. 487-516.
55. Grant, W.D. and H. Larsen, *Extremely halophilic archaeobacteria*. *Williams and Wilkins*, 1989: p. 2216-2233.

56. Grote, A., *et al.*, *JCat: a novel tool to adapt codon usage of a target gene to its potential expression host*. *Nucleic Acids Research*, 2005. **33**: p. W526-W531.
57. Gruhler, A., *et al.*, *Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway*. *Mol Cell Proteomics*, 2005. **4**(3): p. 310-27.
58. Gustafsson, M., *et al.*, *Integrated Sample Preparation and MALDI Mass Spectrometry on a Microfluidic Compact Disk*. *Analytical Chemistry*, 2004. **76**(2): p. 345-350.
59. Gygi, S.P., *et al.*, *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. *Nat Biotechnol*, 1999. **17**(10): p. 994-9.
60. Hartmann, R., H.D. Sickinger, and D. Oesterhelt, *Quantitative aspects of energy conversion in halobacteria*. *FEBS Lett*, 1977. **82**(1): p. 1-6.
61. Hartmann, R., H.D. Sickinger, and D. Oesterhelt, *Anaerobic Growth of Halobacteria*. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, 1980. **77**(7): p. 3821-3825.
62. Haupts, U., J. Tittor, and D. Oesterhelt, *Closing in on bacteriorhodopsin: Progress in understanding the molecule*. *Annual Review of Biophysics and Biomolecular Structure*, 1999. **28**: p. 367-399.
63. Hecker, M. and U. Volker, *Towards a comprehensive understanding of Bacillus subtilis cell physiology by physiological proteomics*. *Proteomics*, 2004. **4**(12): p. 3727-3750.
64. Hochstein, L.I. and F. Lang, *Purification and Properties of a Dissimilatory Nitrate Reductase from Haloferax-Denitrificans*. *Archives of Biochemistry and Biophysics*, 1991. **288**(2): p. 380-385.
65. Hochstein, L.I. and F. Lang, *Purification and properties of a dissimilatory nitrate reductase from Haloferax denitrificans*. *Arch Biochem Biophys*, 1991. **288**(2): p. 380-5.
66. Hoff, W.D., K.H. Jung, and J.L. Spudich, *Molecular mechanism of photosignaling by archaeal sensory rhodopsins*. *Annu Rev Biophys Biomol Struct*, 1997. **26**: p. 223-58.
67. Hou, S., *et al.*, *Myoglobin-like aerotaxis transducers in Archaea and Bacteria*. *Nature*, 2000. **403**(6769): p. 540-4.
68. Huddleston, M.J., *et al.*, *Selective Detection of Phosphopeptides in Complex-Mixtures by Electrospray Liquid-Chromatography Mass-Spectrometry*. *Journal of the American Society for Mass Spectrometry*, 1993. **4**(9): p. 710-717.
69. Hunter, T., *Signaling--2000 and beyond*. *Cell*, 2000. **100**(1): p. 113-27.

70. Hutchins, A.M., J.F. Holden, and M.W.W. Adams, *Phosphoenolpyruvate synthetase from the hyperthermophilic archaeon Pyrococcus furiosus*. Journal of Bacteriology, 2001. **183**(2): p. 709-715.
71. Ihaka, R. and R. Gentleman, *R: A language for data analysis and graphics*. Journal of Computational and Graphical Statistics, 1996. **5**(3): p. 299-314.
72. Ishihama, Y., *Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry*. J Chromatogr A, 2005. **1067**(1-2): p. 73-83.
73. Jensen, O.N., *Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry*. Curr Opin Chem Biol, 2004. **8**(1): p. 33-41.
74. Johnson, R.S., et al., *Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine*. Anal Chem, 1987. **59**(21): p. 2621-5.
75. Johnson, S.A. and T. Hunter, *Kinomics: methods for deciphering the kinome*. Nat Methods, 2005. **2**(1): p. 17-25.
76. Julka, S. and F. Regnier, *Quantification in proteomics through stable isotope coding: a review*. J Proteome Res, 2004. **3**(3): p. 350-63.
77. Kamekura, M., *Diversity of extremely halophilic bacteria*. Extremophiles, 1998. **2**(3): p. 289-95.
78. Kamekura, M., et al., *Lysis of Halobacteria in Bacto-Peptone by Bile-Acids*. Applied and Environmental Microbiology, 1988. **54**(4): p. 990-995.
79. Karadzic, I.M. and J.A. Maupin-Furlow, *Improvement of two-dimensional gel electrophoresis proteome maps of the haloarchaeon Haloferax volcanii*. Proteomics, 2005. **5**(2): p. 354-9.
80. Karas, M., M. Gluckmann, and J. Schafer, *Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors*. J Mass Spectrom, 2000. **35**(1): p. 1-12.
81. Karas, M. and R. Kruger, *Ion formation in MALDI: the cluster ionization mechanism*. Chem Rev, 2003. **103**(2): p. 427-40.
82. Karlin, S., et al., *Genomic and proteomic comparisons between bacterial and archaeal genomes and related comparisons with the yeast and fly genomes*. Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7309-14.
83. Karlin, S., et al., *Predicted highly expressed genes in archaeal genomes*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(20): p. 7303-7308.

84. Kennedy, S.P., *et al.*, *Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence*. Genome Research, 2001. **11**(10): p. 1641-1650.
85. Kennelly, P.J., *Protein kinases and protein phosphatases in prokaryotes: a genomic perspective*. FEMS Microbiol Lett, 2002. **206**(1): p. 1-8.
86. Kerscher, L., S. Nowitzki, and D. Oesterhelt, *Thermoacidophilic Archaeobacteria Contain Bacterial-Type Ferredoxins Acting as Electron-Acceptors of 2-Oxoacid-Ferredoxin Oxidoreductases*. European Journal of Biochemistry, 1982. **128**(1): p. 223-230.
87. Kerscher, L. and D. Oesterhelt, *Ferredoxin from Halobacteria*. Febs Letters, 1976. **67**(3): p. 320-322.
88. Kerscher, L., *et al.*, *New Plant-Type Ferredoxin from Halobacteria*. European Journal of Biochemistry, 1976. **71**(1): p. 101-107.
89. Klein, C., *et al.*, *The membrane proteome of Halobacterium salinarum*. Proteomics, 2005. **5**(1): p. 180-97.
90. Klose, J., *Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals*. Humangenetik, 1975. **26**(3): p. 231-43.
91. Koch, M.K. and D. Oesterhelt, *MpcT is the transducer for membrane potential changes in Halobacterium salinarum*. Mol Microbiol, 2005. **55**(6): p. 1681-94.
92. Kokoeva, M.V. and D. Oesterhelt, *BasT, a membrane-bound transducer protein for amino acid detection in Halobacterium salinarum*. Mol Microbiol, 2000. **35**(3): p. 647-56.
93. Kokoeva, M.V., *et al.*, *A novel mode of sensory transduction in archaea: binding protein-mediated chemotaxis towards osmoprotectants and amino acids*. Embo J, 2002. **21**(10): p. 2312-22.
94. Krogh, A., *et al.*, *Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes*. Journal of Molecular Biology, 2001. **305**(3): p. 567-580.
95. Kupper, J., *et al.*, *The Flagellar Bundle of Halobacterium-Salinarium Is Inserted into a Distinct Polar-Cap Structure*. Journal of Bacteriology, 1994. **176**(16): p. 5184-5187.
96. Kushner, D.J., *What Is the True Internal Environment of Halophilic and Other Bacteria*. Canadian Journal of Microbiology, 1988. **34**(4): p. 482-486.
97. Laemmli, U.K., *Cleavage of Structural Proteins During Assembly of Head of Bacteriophage-T4*. Nature, 1970. **227**(5259): p. 680-&.

98. Lanyi, J.K., *Salt-Dependent Properties of Proteins from Extremely Halophilic Bacteria*. Bacteriological Reviews, 1974. **38**(3): p. 272-290.
99. Lanyi, J.K. and R.E. MacDonald, *Existence of electrogenic hydrogen ion/sodium ion antiport in Halobacterium halobium cell envelope vesicles*. Biochemistry, 1976. **15**(21): p. 4608-14.
100. Larsen, M.R., *et al.*, *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns*. Mol Cell Proteomics, 2005. **4**(7): p. 873-86.
101. Link, A.J., *et al.*, *Direct analysis of protein complexes using mass spectrometry*. Nat Biotechnol, 1999. **17**(7): p. 676-82.
102. Link, A.J., *et al.*, *Identifying the major proteome components of Haemophilus influenzae type-strain NCTC 8143*. Electrophoresis, 1997. **18**(8): p. 1314-1334.
103. Link, A.J., K. Robison, and G.M. Church, *Comparing the predicted and observed properties of proteins encoded in the genome of Escherichia coli K-12*. Electrophoresis, 1997. **18**(8): p. 1259-1313.
104. Lottspeich, F., *Bioanalytik*. Spektrum Akademischer Verlag (eds. F. Lottspeich, H. Zorbas), 1998.
105. Lottspeich, F., *Proteome Analysis: A Pathway to the Functional Analysis of Proteins*. Angew Chem Int Ed Engl, 1999. **38**(17): p. 2476-2492.
106. Loyet, K.M., J.T. Stults, and D. Arnott, *Mass spectrometric contributions to the practice of phosphorylation site mapping through 2003: a literature review*. Mol Cell Proteomics, 2005. **4**(3): p. 235-45.
107. Luhn, S., *et al.*, *Using standard positions and image fusion to create proteome maps from collections of two-dimensional gel electrophoresis images*. Proteomics, 2003. **3**(7): p. 1117-27.
108. MacDonald, R.E., R.V. Greene, and J.K. Lanyi, *Light-activated amino acid transport systems in Halobacterium halobium envelope vesicles: role of chemical and electrical gradients*. Biochemistry, 1977. **16**(14): p. 3227-35.
109. Madern, D., C. Ebel, and G. Zaccai, *Halophilic adaptation of enzymes*. Extremophiles, 2000. **4**(2): p. 91-98.
110. Madern, D., C. Pfister, and G. Zaccai, *Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from Haloarcula marismortui in physiological salts*. Eur J Biochem, 1995. **230**(3): p. 1088-95.
111. Mann, M., P. Hojrup, and P. Roepstorff, *Use of Mass-Spectrometric Molecular-Weight Information to Identify Proteins in Sequence Databases*. Biological Mass Spectrometry, 1993. **22**(6): p. 338-345.

112. Mann, M. and O.N. Jensen, *Proteomic analysis of post-translational modifications*. Nat Biotechnol, 2003. **21**(3): p. 255-61.
113. Mann, M. and M. Wilm, *Error-tolerant identification of peptides in sequence databases by peptide sequence tags*. Anal Chem, 1994. **66**(24): p. 4390-9.
114. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: a primer*. Mass Spectrom Rev, 1998. **17**(1): p. 1-35.
115. Marwan, W., et al., *Mechanism of Photosensory Adaptation in Halobacterium-Salinarium*. Journal of Molecular Biology, 1995. **246**(4): p. 493-499.
116. Marwan, W. and D. Oesterhelt, *Archaeal vision and bacterial smelling*. Asm News, 2000. **66**(2): p. 83-89.
117. Matsumoto, A., et al., *Phosphorylation of the AfsR protein involved in secondary metabolism in Streptomyces species by a eukaryotic-type protein kinase*. Gene, 1994. **146**(1): p. 47-56.
118. Mayya, V., et al., *Systematic comparison of a two-dimensional ion trap and a three-dimensional ion trap mass spectrometer in proteomics*. Mol Cell Proteomics, 2005. **4**(2): p. 214-23.
119. McCarthy, J., et al., *Carbamylation of proteins in 2-D electrophoresis - Myth or reality?* Journal of Proteome Research, 2003. **2**(3): p. 239-242.
120. McDonald, W.H. and J.R. Yates, 3rd, *Shotgun proteomics: integrating technologies to answer biological questions*. Curr Opin Mol Ther, 2003. **5**(3): p. 302-9.
121. McLuckey, S.A., *Principles of Collisional Activation in Analytical Mass-Spectrometry*. Journal of the American Society for Mass Spectrometry, 1992. **3**(6): p. 599-614.
122. Michel, H. and D. Oesterhelt, *Light-Induced-Changes of Ph Gradient and Membrane-Potential in H-Halobium*. Febs Letters, 1976. **65**(2): p. 175-178.
123. Mirgorodskaya, E., et al., *Nanoflow liquid chromatography coupled to matrix-assisted laser desorption/ionization mass spectrometry: sample preparation, data analysis, and application to the analysis of complex peptide mixtures*. Proteomics, 2005. **5**(2): p. 399-408.
124. Moore, R.E., M.K. Young, and T.D. Lee, *Qscore: an algorithm for evaluating SEQUEST database search results*. J Am Soc Mass Spectrom, 2002. **13**(4): p. 378-86.
125. Morgan, D.O., *Principles of CDK regulation*. Nature, 1995. **374**(6518): p. 131-4.

126. Moritz, B. and H.E. Meyer, *Approaches for the quantification of protein concentration ratios*. Proteomics, 2003. **3**(11): p. 2208-2220.
127. Mortz, E., *et al.*, *Improved silver staining protocols for high sensitivity protein identification using matrix-assisted laser desorption/ionization-time of flight analysis*. Proteomics, 2001. **1**(11): p. 1359-1363.
128. Nakamura, K., *et al.*, *[From genome to proteome--aim of human proteomics]*. Seikagaku, 2004. **76**(10): p. 1271-4.
129. Naught, L.E., *et al.*, *Roles of active site residues in Pseudomonas aeruginosa phosphomannomutase/phosphoglucomutase*. Biochemistry, 2003. **42**(33): p. 9946-9951.
130. Naught, L.E. and P.A. Tipton, *Kinetic mechanism and pH dependence of the kinetic parameters of Pseudomonas aeruginosa phosphomannomutase/phosphoglucomutase*. Archives of Biochemistry and Biophysics, 2001. **396**(1): p. 111-118.
131. Neuhoff, V., *et al.*, *Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250*. Electrophoresis, 1988. **9**(6): p. 255-62.
132. Ng, W.V., *et al.*, *Genome sequence of Halobacterium species NRC-1*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(22): p. 12176-12181.
133. Nordhoff, E., *et al.*, *Large-gel two-dimensional electrophoresis-matrix assisted laser desorption/ionization-time of flight-mass spectrometry: an analytical challenge for studying complex protein mixtures*. Electrophoresis, 2001. **22**(14): p. 2844-55.
134. Nordhoff, E., *et al.*, *Sample preparation protocols for MALDI-MS of peptides and oligonucleotides using prestructured sample supports*. International Journal of Mass Spectrometry, 2003. **226**(1): p. 163-180.
135. Oda, Y., T. Nagasu, and B.T. Chait, *Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome*. Nat Biotechnol, 2001. **19**(4): p. 379-82.
136. Oesterhelt, D., *Structure and function of halorhodopsin*. Israel Journal of Chemistry, 1995. **35**(3-4): p. 475-494.
137. Oesterhelt, D., *The structure and mechanism of the family of retinal proteins from halophilic archaea*. Current Opinion in Structural Biology, 1998. **8**(4): p. 489-500.

138. Oesterhelt, D. and G. Krippahl, *Phototropic Growth of Halobacteria and Its Use for Isolation of Photosynthetically-Deficient Mutants*. *Annales De Microbiologie*, 1983. **B134**(1): p. 137-150.
139. Oesterhelt, D. and W. Stoeckenius, *Isolation of the Cell Membrane of Halobacterium halobium and its Fractionation into Red and Purple Membrane*. *Methods Enzymol.*, 1974. **31**: p. 667-678.
140. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins*. *J Biol Chem*, 1975. **250**(10): p. 4007-21.
141. Olsen, J.V. and M. Mann, *Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation*. *Proc Natl Acad Sci U S A*, 2004. **101**(37): p. 13417-22.
142. Olsen, J.V., S.E. Ong, and M. Mann, *Trypsin cleaves exclusively C-terminal to arginine and lysine residues*. *Mol Cell Proteomics*, 2004. **3**(6): p. 608-14.
143. O'Neill, E.E., *et al.*, *Towards complete analysis of the platelet proteome*. *Proteomics*, 2002. **2**(3): p. 288-305.
144. Ong, S.E., *et al.*, *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics*. *Molecular & Cellular Proteomics*, 2002. **1**(5): p. 376-386.
145. Ong, S.E., I. Kratchmarova, and M. Mann, *Properties of C-13-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC)*. *Journal of Proteome Research*, 2003. **2**(2): p. 173-181.
146. Oren, A., *The Ecology of the Extremely Halophilic Archaea*. *Fems Microbiology Reviews*, 1994. **13**(4): p. 415-439.
147. Oren, A. and H.G. Truper, *Anaerobic Growth of Halophilic Archaeobacteria by Reduction of Dimethylsulfoxide and Trimethylamine N-Oxide*. *Fems Microbiology Letters*, 1990. **70**(1): p. 33-36.
148. Pappin, D.J.C., P. Hojrup, and A.J. Bleasby, *Rapid Identification of Proteins by Peptide-Mass Fingerprinting*. *Current Biology*, 1993. **3**(6): p. 327-332.
149. Parkinson, J.S., *Signal transduction schemes of bacteria*. *Cell*, 1993. **73**(5): p. 857-71.
150. Pawson, T. and J.D. Scott, *Protein phosphorylation in signaling--50 years and counting*. *Trends Biochem Sci*, 2005. **30**(6): p. 286-90.
151. Peng, J., *et al.*, *Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome*. *J Proteome Res*, 2003. **2**(1): p. 43-50.

152. Perkins, D.N., *et al.*, *Probability-based protein identification by searching sequence databases using mass spectrometry data*. *Electrophoresis*, 1999. **20**(18): p. 3551-3567.
153. Perrier, J., *et al.*, *Catabolism of intracellular N-terminal acetylated proteins: involvement of acylpeptide hydrolase and acylase*. *Biochimie*, 2005. **87**(8): p. 673-685.
154. Pfeifer, F. and M. Betlach, *Genome Organization in Halobacterium-Halobium - a 70 Kb Island of More (at) Rich DNA in the Chromosome*. *Molecular & General Genetics*, 1985. **198**(3): p. 449-455.
155. Pinkse, M.W., *et al.*, *Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns*. *Anal Chem*, 2004. **76**(14): p. 3935-43.
156. Polevoda, B. and F. Sherman, *N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins*. *Journal of Molecular Biology*, 2003. **325**(4): p. 595-622.
157. Posewitz, M.C. and P. Tempst, *Immobilized gallium(III) affinity chromatography of phosphopeptides*. *Anal Chem*, 1999. **71**(14): p. 2883-92.
158. Rabilloud, T., *et al.*, *A comparison between Sypro Ruby and ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels*. *Proteomics*, 2001. **1**(5): p. 699-704.
159. Raman, B., A. Cheung, and M.R. Marten, *Quantitative comparison and evaluation of two commercially available, two-dimensional electrophoresis image analysis software packages, Z3 and Melanie*. *Electrophoresis*, 2002. **23**(14): p. 2194-202.
160. Rappsilber, J., Y. Ishihama, and M. Mann, *Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics*. *Anal Chem*, 2003. **75**(3): p. 663-70.
161. Regni, C., P.A. Tipton, and L.J. Beamer, *Crystal structure of PMM/PGM: an enzyme in the biosynthetic pathway of P. aeruginosa virulence factors*. *Structure (Camb)*, 2002. **10**(2): p. 269-79.
162. Richert, S., *et al.*, *About the mechanism of interference of silver staining with peptide mass spectrometry*. *Proteomics*, 2004. **4**(4): p. 909-16.
163. Richmond, C.S., *et al.*, *Genome-wide expression profiling in Escherichia coli K-12*. *Nucleic Acids Research*, 1999. **27**(19): p. 3821-3835.
164. Righetti, P.G., *et al.*, *Prefractionation techniques in proteome analysis*. *Proteomics*, 2003. **3**(8): p. 1397-407.

165. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides*. Biomed Mass Spectrom, 1984. **11**(11): p. 601.
166. Rosen, R., *et al.*, *Highly phosphorylated bacterial proteins*. Proteomics, 2004. **4**(10): p. 3068-3077.
167. Rosengren, A.T., *et al.*, *Comparison of PDQuest and Progenesis software packages in the analysis of two-dimensional electrophoresis gels*. Proteomics, 2003. **3**(10): p. 1936-46.
168. Ruepp, A., *et al.*, *Catabolic Ornithine Transcarbamylase of Halobacterium-Halobium (Salinarium) - Purification, Characterization, Sequence Determination, and Evolution*. Journal of Bacteriology, 1995. **177**(5): p. 1129-1136.
169. Russell, J.B., *Strategies that ruminal bacteria use to handle excess carbohydrate*. J Anim Sci, 1998. **76**(7): p. 1955-63.
170. Russell, J.B. and G.M. Cook, *Energetics of bacterial growth: balance of anabolic and catabolic reactions*. Microbiol Rev, 1995. **59**(1): p. 48-62.
171. Salzberg, S.L., *et al.*, *Microbial gene identification using interpolated Markov models*. Nucleic Acids Research, 1998. **26**(2): p. 544-548.
172. Sartorius-Neef, S. and F. Pfeifer, *In vivo studies on putative Shine-Dalgarno sequences of the halophilic archaeon Halobacterium salinarum*. Mol Microbiol, 2004. **51**(2): p. 579-88.
173. Schlosser, A., *et al.*, *Analysis of protein phosphorylation by a combination of elastase digestion and neutral loss tandem mass spectrometry*. Analytical Chemistry, 2001. **73**(2): p. 170-176.
174. Schmidt, A., J. Kellermann, and F. Lottspeich, *A novel strategy for quantitative proteomics using isotope-coded protein labels*. Proteomics, 2005. **5**(1): p. 4-15.
175. Schuerenbeg, M., *et al.*, *Prestructured MALDI-MS sample supports*. Analytical Chemistry, 2000. **72**(15): p. 3436-3442.
176. Schwartz, J.C., M.W. Senko, and J.E. Syka, *A two-dimensional quadrupole ion trap mass spectrometer*. J Am Soc Mass Spectrom, 2002. **13**(6): p. 659-69.
177. Shevchenko, A., *et al.*, *Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels*. Analytical Chemistry, 1996. **68**(5): p. 850-858.
178. Shi, Y., *et al.*, *The role of liquid chromatography in proteomics*. J Chromatogr A, 2004. **1053**(1-2): p. 27-36.
179. Sickmann, A. and H.E. Meyer, *Phosphoamino acid analysis*. Proteomics, 2001. **1**(2): p. 200-6.

180. Smith, P.K., *et al.*, *Measurement of protein using bicinchoninic acid*. Anal Biochem, 1985. **150**(1): p. 76-85.
181. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing*. Nat Rev Mol Cell Biol, 2004. **5**(9): p. 699-711.
182. Storch, K.F., J. Rudolph, and D. Oesterhelt, *Car: a cytoplasmic sensor responsible for arginine chemotaxis in the archaeon Halobacterium salinarum*. Embo J, 1999. **18**(5): p. 1146-58.
183. Swanson, S.K. and M.P. Washburn, *The continuing evolution of shotgun proteomics*. Drug Discov Today, 2005. **10**(10): p. 719-25.
184. Syka, J.E., *et al.*, *Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications*. J Proteome Res, 2004. **3**(3): p. 621-6.
185. Tebbe, A., *et al.*, *Analysis of the cytosolic proteome of Halobacterium salinarum and its implication for genome annotation*. Proteomics, 2005. **5**(1): p. 168-79.
186. Thiede, B., *et al.*, *Peptide mass fingerprinting*. Methods, 2005. **35**(3): p. 237-47.
187. Tonella, L., *et al.*, *New perspectives in the Escherichia coli proteome investigation*. Proteomics, 2001. **1**(3): p. 409-423.
188. Tonge, R., *et al.*, *Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology*. Proteomics, 2001. **1**(3): p. 377-96.
189. Torarinsson, E., H.P. Klenk, and R.A. Garrett, *Divergent transcriptional and translational signals in Archaea*. Environ Microbiol, 2005. **7**(1): p. 47-54.
190. Unlu, M., M.E. Morgan, and J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts*. Electrophoresis, 1997. **18**(11): p. 2071-7.
191. VanBogelen, R.A., *et al.*, *Escherichia coli proteome analysis using the gene-protein database*. Electrophoresis, 1997. **18**(8): p. 1243-1251.
192. Volker, U. and M. Hecker, *From genomics via proteomics to cellular physiology of the Gram-positive model organism Bacillus subtilis*. Cellular Microbiology, 2005. **7**(8): p. 1077-1085.
193. Wakabayashi, S., *et al.*, *Amino-Acid-Sequence of a Ferredoxin from Thermoacidophilic Archaeobacteria, Thermoplasma-Acidophilum*. Febs Letters, 1983. **162**(1): p. 21-24.
194. Wang, J.Q., *et al.*, *An analysis of the proteomic profile for Thermoanaerobacter tengcongensis under optimal culture conditions*. Proteomics, 2004. **4**(1): p. 136-150.

195. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology*. Nat Biotechnol, 2001. **19**(3): p. 242-7.
196. Wasinger, V.C., B. Bjellqvist, and I. HumpherySmith, *Proteomic 'contigs' of Ochrobactrum anthropi, application of extensive pH gradients*. Electrophoresis, 1997. **18**(8): p. 1373-1383.
197. Whetstone, P.A., *et al.*, *Element-coded affinity tags for peptides and proteins*. Bioconjug Chem, 2004. **15**(1): p. 3-6.
198. Woese, C.R., *Bacterial Evolution*. Microbiological Reviews, 1987. **51**(2): p. 221-271.
199. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya*. Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.
200. Yates, J.R., *et al.*, *Peptide Mass Maps - a Highly Informative Approach to Protein Identification*. Analytical Biochemistry, 1993. **214**(2): p. 397-408.
201. Zhou, H., J.D. Watts, and R. Aebersold, *A systematic approach to the analysis of protein phosphorylation*. Nat Biotechnol, 2001. **19**(4): p. 375-8.
202. Zhou, S., *et al.*, *A quantitative investigation into the losses of proteins at different stages of a two-dimensional gel electrophoresis procedure*. Proteomics, 2005. **5**(11): p. 2739-47.
203. Zhu, H.N., *et al.*, *Amino acid residue specific stable isotope labeling for quantitative proteomics*. Rapid Communications in Mass Spectrometry, 2002. **16**(22): p. 2115-2123.
204. Zhu, W., *et al.*, *Shotgun proteomics of Methanococcus jannaschii and insights into methanogenesis*. J Proteome Res, 2004. **3**(3): p. 538-48.
205. Zubarev, R.A., P. Hakansson, and B. Sundqvist, *Accuracy requirements for peptide characterization by monoisotopic molecular mass measurements*. Analytical Chemistry, 1996. **68**(22): p. 4060-4063.

## 9 Anhang

Auf der beiliegenden CD sind die Tabellen der identifizierten Proteine der in dieser Arbeit durchgeführten Proteomstudien zu finden.

**Tabelle 1:** Liste der identifizierten Proteine aus 2-D Referenzgelen mittels MALDI-TOF PMF (siehe Abschnitt [4.1.3](#)).

Funktionsklasse, Gen-Code, Genname, Proteinname, theoretischer  $pI$  und theoretisches  $M_r$  sind aufgeführt. „Sector“ steht für den Bereich der 2D-Gele, in dem das Protein identifiziert wurde (a, c, e, g = 9% Gele; pH 3.5-4.5, 4.0-5.0, 4.5-5.5 und 5.0-8.0; b, d, f, h = 14% Gele, pH 3.5-4.5, 4.0-5.0, 4.5-5.5, and 5.0-8.0. Wenn kein Sektor-Index angegeben ist wurde das Protein nur auf Gelen identifiziert die nicht den Referenz Gelen entsprechen. Wenn das Protein nur auf Gelen mit hoher Proteinbeladung identifiziert wurde ist der Sektor-Index mit einem Stern markiert. Die Spalte „Multiple“ zeigt an, ob ein Protein in mehreren Spots auf dem gleichen Gel identifiziert wurde. Dabei steht  $pI$  für Verschiebungen in  $pI$ -Richtung und  $M_r$  für Verschiebungen in  $M_r$ -Richtung. „Identifizierungslevel“ klassifiziert die Proteine wie in „Material und Methoden“ [3.2.12.1](#) beschrieben (r = reliable, n = normal, m = manual). Abkürzungen für die Funktionsklassen: AA = *aminoacid metabolism*, CE = *cell envelope*, CHM = *carbohydrate metabolism*, CHP = *chaperones*, CIM = *central intermediary metabolism*, COM = *coenzyme metabolism*, CP = *cellular processes*, EM = *energy metabolism*, LIP = *lipid metabolism*, MIS = *miscellaneous*, NOF = *no function*, NUM = *nucleotide metabolism*, REG = *gene regulation*, RMT = *RNA maturation*, RRR = *replication, repair, recombination*, SEC = *protein secretion*, SIG = *signal transduction*, TC = *transcription*, TL = *translation*, TP = *small molecule transport*

**Tabelle 2:** Identifizierte Proteine der ICPL-Studie (LC-MALDI TOF/TOF) des Vergleichs des Wachstums in Komplexmedium mit dem Wachstum in synthetischem Medium (siehe Abschnitt [4.2.3](#)).

Angegeben sind Gen-Code und Proteinname. Weiter gezeigt sind: Die Anzahl aller identifizierten Peptide des jeweiligen Proteins über alle analysierten Fraktionen (**All Pept.**), die Anzahl der identifizierten einzigartigen Peptidsequenzen (**Unique Pept.**), die Anzahl aller identifizierten Singletts des jeweiligen Proteins über alle analysierten Fraktionen (**All Singl.**), die Anzahl der identifizierten einzigartigen Singlettsequenzen (**Unique Singl.**), die Anzahl aller identifizierten Isotopenpaare des jeweiligen Proteins über alle analysierten Fraktionen (**All Pairs**), die Anzahl der identifizierten einzigartigen Isotopenpaarsequenzen (**Unique Pairs**), die Berechnete Regulation (Reg.) sowie die zugehörigen Standardabweichung (**StdDev**).

**Tabelle 3:** Identifizierte Proteine der ICPL-Studie (LC-MALDI TOF/TOF) des Vergleichs des Wachstums unter aeroben Bedingungen mit dem Wachstum unter anaeroben phototrophen Bedingungen (siehe Abschnitt 4.2.4).

Angegeben sind Gen-Code und Proteinname. Weiter gezeigt sind: Die Anzahl aller identifizierten Peptide des jeweiligen Proteins über alle analysierten Fraktionen (**All Pept.**), die Anzahl der identifizierten einzigartigen Peptidsequenzen (**Unique Pept.**), die Anzahl aller identifizierten Singletts des jeweiligen Proteins über alle analysierten Fraktionen (**All Singl.**), die Anzahl der identifizierten einzigartigen Singlettsequenzen (**Unique Singl.**), die Anzahl aller identifizierten Isotopenpaare des jeweiligen Proteins über alle analysierten Fraktionen (**All Pairs**), die Anzahl der identifizierten einzigartigen Isotopenpaarsequenzen (**Unique Pairs**), die Berechnete Regulation (Reg.) sowie die zugehörigen Standardabweichung (**StdDev**).

**Tabelle 4:** Identifizierte Proteine mit Regulationsinformationen der ICPL-Studie (2-DE MALDI TOF/TOF) des Vergleichs des Wachstums in Komplexmedium mit dem Wachstum (siehe Abschnitt 4.2.6).

Angegeben sind Gen-Code und Proteinname sowie die Anzahl der identifizierten einzigartigen Isotopenpaarsequenzen des jeweiligen Proteins (**Pairs**), die sich daraus berechnete Regulation (**Reg**) und die Standardabweichung (**StdDev**).

**Tabelle 5:** Identifizierte Proteine mit Regulationsinformationen der ICPL-Studie (2-DE MALDI TOF/TOF) des Vergleichs des Wachstums unter aeroben Bedingungen mit dem Wachstum unter anaeroben phototrophen Bedingungen (siehe Abschnitt 4.2.6).

Angegeben sind Gen-Code und Proteinname sowie die Anzahl der identifizierten einzigartigen Isotopenpaarsequenzen des jeweiligen Proteins (**Pairs**), die sich daraus berechnete Regulation (**Reg**) und die Standardabweichung (**StdDev**).

**Tabelle 6:** Identifizierte Proteine der ICPL-Proben (LC-ESI FT-ICR MS) des Vergleichs des Wachstums unter aeroben Bedingungen mit dem Wachstum unter anaeroben phototrophen Bedingungen (siehe Abschnitt 4.2.7).

Funktionsklasse, Gen-Code, Gennamen, Proteinname, theoretischer  $pI$ , theoretisches  $M_r$  und die Anzahl der vorhergesagten Transmembranregionen sind aufgeführt. Für die Abkürzungen der Funktionsklassen siehe Tabelle 1.

## Danksagung

Zu dem Gelingen dieser Arbeit haben viele beigetragen, bei denen ich mich an dieser Stelle bedanken möchte.

Mein erster Dank gebührt meinem Doktorvater Prof. Dr. Dieter Oesterheld, der mich mit einer spannenden Aufgabenstellung betraute, für alle Mittel und Freiheiten zur Realisierung dieser Arbeit und zur Umsetzung meiner Ideen sorgte, sowie das Fortschreiten des Projektes stets unterstützend begleitete. Die vielen Diskussionen mit ihm wirkten nicht nur bewusstseinsweiternd sondern auch enorm motivierend – ich habe viel von ihm gelernt.

Ebenso möchte ich mich bei P.D. Dr. Christoph Eckerskorn für die Bereitschaft bedanken, das Zweitgutachten für die vorliegende Arbeit zu übernehmen.

Während der vergangenen Jahre hatte ich das große Glück in einem außerordentlich motivierenden Umfeld zu arbeiten. Hierbei möchte ich mich ganz herzlich bei der Arbeitsgruppe „Massenspektrometrie“ unter der Leitung von Dr. Frank Siedler bedanken, von dem ich nicht nur das kleine und große Einmaleins der Proteomforschung erlernen durfte, sondern auch von ihm in Sachen Projektplanung und Dokumentation profitieren konnte. Dr. Christian Klein danke ich für die Einführung in das Arbeitsgebiet. Das Arbeiten mit ihm hat immer großen Spaß gemacht. Nicht nur in Punkto des regionalen Zugehörigkeitsgefühls und daraus resultierenden fußballerischen Feindbildern fühlte ich mich ihm stets verbunden.

Beatrix Scheffer und Birgit Bisle danke ich für die nette Zusammenarbeit speziell in Bezug auf ihre Hilfestellungen in Sachen nano-LS-MS und dem „Grossen Versuch“. Weiter Danke ich Sigrid Bauer und Barbara Fischer für ihre stete Hilfsbereitschaft und den netten Umgang. Ebenfalls danke ich sehr herzlich Dr. Michalis Aivaliotis sowie dem Kostas Konstantinidis für eine nette und gute Zusammenarbeit. Letzterem danke ich besonders für die Übernahme der Organisation von Kongressbesuchen, die mich oftmals heillos überfordert hätte.

Weiter möchte ich mich bei der Bioinformatikgruppe bedanken, die uns „Proteomiker“ stets unermüdlich bei der Auswertung der Daten unterstützt hat. Mein besonderer dank gilt hierbei Dr. Friedhelm Pfeiffer, für etliche kleine Programme aber auch die zahlreichen Diskussionen und Erklärungen zum biologischen und biochemischen Verständnis des eigenen Handelns. Er hat substantiell zum Fortschreiten dieser Arbeit beigetragen.

Ein weiteres Dankeschön gilt der Arbeitsgruppe von Dr. Friedrich Lottspeich. Mein besonderer Dank gilt hierbei Alexander Schmidt für die tolle Zusammenarbeit auf

dem ICPL-Gebiet sowie die vielen und lehrreichen Diskussionen zur *Proteomics*. Ebenfalls danke ich Monica Zobawa für die zahlreichen Messungen meiner Proben.

Prof. Dr. Matthias Mann sowie Dr. Boris Macek und Jesper V. Olsen möchte ich für die Einführung in neuste massenspektrometrische Technologien danken

Dr. Jörg Tittor danke ich sehr herzlich für Käsespäzle, Ratschläge und die Unterstützung beim Umgang mit Radioaktivität. Frau Haack möchte ich meinen Dank aussprechen für die vielen Hilfen in unterschiedlichsten organisatorischen Angelegenheiten. Silvia Haslbeck danke ich für das Bereitstellen von sauberem Labormaterial und angesetzten Medien.

Bei allen nicht genannten Mitgliedern der Abteilung „Oesterhelt“, die mich während der letzten Jahre begleitet haben bedanke ich mich für die tolle und freundliche Arbeitsatmosphäre. Es hat wirklich großen Spaß gemacht.

Kaya, Mimi und Gabi danke ich für das Korrekturlesen dieser Arbeit.

Zum Schluss aber dafür um so herzlicher danke ich meinen Eltern und meiner ganzen Familie, die mich immer mit aller Kraft unterstützt haben.

## Veröffentlichungen

Teile dieser Dissertation finden sich in folgenden Publikationen wieder:

- Tebbe A., Klein C., Bisle B., Siedler F., Scheffer B., Garzia-Rizo C., Wolferzt J., Hickmann V., Pfeiffer F. and Oesterhelt D., Analysis of the cytosolic proteome of *Halobacterium salinarum* and its implication for genome annotation, *Proteomics*, 5(1) 168 - 179 (2005)
- Tebbe A., Schmidt A., Konstantinidis K., Siedler F., Pfeiffer F., Kellermann J., Lottspeich F., and Oesterhelt D., Life-Style changes of a halophilic archaeon analyzed by quantitative proteomics, in Vorbereitung
- Aivaliotis M., Falb M., Tebbe A., Siedler F., Pfeiffer F., Oesterhelt D., Large scale identification of N-terminal peptides in *Halobacterium salinarum* and *Natronomonas pharaonis*, in Vorbereitung

Posterbeiträge auf wissenschaftlichen Konferenzen zum Thema dieser Arbeit:

- Tebbe A., Klein C., Bisle B., Siedler F., Scheffer B., Garcia-Rizo C., Wolfertz C., Hickmann V., Pfeiffer F. and Oesterhelt D. Analysis of the cytosolic proteome of *Halobacterium salinarum* and its implication for genome annotation, Poster Präsentation, *From Genome To Proteome: Biomarker Discovery and Imaging Proteomics*; 6<sup>th</sup> Siena Meeting, Siena, Italien, 30 August- 2 September 2004
- Bisle B., Klein C., Tebbe A., Koch M., Scheffer B., Pfeiffer F., Siedler F. and Oesterhelt D., Proteomics and Systems Biology – The Model Organism *Halobacterium salinarum*, Poster Präsentation, 2. *BMBF Status Seminar*, Potsdam, Deutschland, Juni 2005
- Tebbe A., Klein C., Schmidt A., Bisle B., Konstantinidis K., Scheffer B., Lottspeich F., Siedler F., Pfeiffer F. and Oesterhelt D., Analysis of the Cytosolic Proteome of *Halobacterium salinarum* - implications for genome annotation and differential expression, Poster Präsentation, 53<sup>rd</sup> *ASMS Conference on Mass Spectrometry*, San Antonio Texas, USA, 5.-9. Juni 2005
- Tebbe A., Klein C., Schmidt A., Bisle B., Konstantinidis K., Scheffer B., Lottspeich F., Siedler F., Pfeiffer F. and Oesterhelt D., Analysis of the Cytosolic Proteome of *Halobacterium salinarum* - Genome annotation and differential expression, Poster Präsentation, *HUPO 4<sup>th</sup> Annual World Congress*, München, Deutschland, 28. August - 1. September 2005 (ausgezeichnet mit dem Poster Award aus über 1000 Beiträgen)
- Tebbe A., Aivaliotis M., Siedler F., Pfeiffer F. and Oesterhelt D., Analysis of protein phosphorylation in *Halobacterium salinarum*, Poster Präsentation, *HUPO 4<sup>th</sup> Annual World Congress*, München, Deutschland, 28. August - 1. September 2005

## Lebenslauf

### Schulbildung

08/1982 – 07/1986

08/1986 – 06/1995

Andreas Tebbe

Geboren am 02.05.1976 in Dortmund

Olpketal Grundschule, Dortmund

Max-Planck-Gymnasium, Dortmund

### Zivildienst

08/1995 – 08/1996

Zivildienst im Altenheim Weiße Taube in Dortmund

### Hochschulstudium

10/1996 – 07/2001

10/2000 – 07/2001

Studium der Biochemie an der Ruhr-Universität Bochum

Diplomarbeit

am Max-Planck-Institut für molekulare Physiologie, Dortmund  
in den Abteilungen Strukturelle Biologie und Chemische  
Biologie

Diplomarbeitsthema:

*Zum modularen Einsatz von Lipopeptiden mit Sequenzmotiven  
von Ras Proteinen: Synthese und Charakterisierung von Proben  
für Lokalisationsstudien und Affinitätsmarkierung*

07/2001

Abschluss Diplom

### Berufspraktikum

10/2001 – 12/2001

Praktikum bei der Morphochem AG – München, Tätigkeitsfeld:  
*Etablierung eines Assays zur Messung der Interaktion zwischen  
kleinen Molekülen und dem Eliminationstransporter P-  
Glycoprotein (ABCI/MDR1)*

### Promotion

02/2002 – 11/2005

Dissertation

am Max-Planck-Institut für Biochemie, Martinried in der  
Abteilung für Membranbiochemie unter der Leitung von Prof.  
Dr. Dieter Oesterhelt

Dissertationsthema:

*Das Proteom eines halophilen Archaeons un seine Antwort auf  
Änderungen der Lebensbedingungen – Inventarisierung,  
Quantifizierung und posttranslationale Modifikationen*

### Post-Doc

ab 12/2005

Post-Doc Aufenthalt bei der Hoffmann-La Roche AG  
am *Medical Genomics Research Center*, Basel, Schweiz