

**Charakterisierung von mikrobiellen Gemeinschaften
aus Biogasanlagen mittels Metaproteomanalyse
und Korrelationen der Ergebnisse
zu den Prozessparametern**

Dissertation
zur Erlangung des akademischen Grades
Doktoringenieur
(Dr.-Ing.)

von Dipl.-Ing. Robert Steven Heyer

geb. am 18.02.1986 in Magdeburg

genehmigt durch die Fakultät für Verfahrens- und Systemtechnik
der Otto-von-Guericke-Universität Magdeburg

Gutachter:

Prof. Dr. Udo Reichl

Prof. Dr. Katharina Riedel

Prof. Dr. Reinhard Pätz

Promotionskolloquium am 01.06.2017

*Wir stehen immer noch vor der Tür, hinter
der die großen Antworten warten.*

Arthur Miller

Wege entstehen dadurch, dass man sie geht.

Franz Kafka

Danksagung

Die vorliegende Dissertation wurde am Lehrstuhl Bioprozesstechnik der Otto-von-Guericke-Universität und am Max-Planck-Institut für Dynamik komplexer technischer Systeme in Magdeburg unter der Leitung von Herrn Prof. Dr.-Ing. U. Reichl angefertigt. In dieser Zeit haben mich viele Personen unterstützt, denen ich an dieser Stelle meinen Dank aussprechen möchte.

Herrn Prof. Dr.-Ing. U. Reichl möchte ich dafür danken, dass ich in seiner Arbeitsgruppe promovieren durfte, für die vielen fruchtbaren Diskussionen und für die vielen hilfreichen Anmerkungen bei diversen Publikationen, Anträgen und Abstracts.

Der **Deutschen Bundesstiftung Umwelt** und ihren Mitarbeitern danke ich nicht nur für die finanzielle Unterstützung, sondern auch dafür, dass sie an mich und mein Promotionsvorhaben geglaubt haben. Die vielen Seminare, Events und Workshops waren eine schöne Bereicherung meiner Doktorandenzeit.

Ganz herzlich möchte ich meinem Betreuer **Herrn Dr. rer. nat. Dirk Benndorf** danken. Zum Einen hat er sich immer Zeit genommen, um sämtliche fachliche und methodische Fragen zu erörtern. Zum Anderen war er eine große Hilfe beim Schreiben des Stipendienantrages, der Publikationen und letztendlich dieser Dissertation. Ferner danke ich ihm dafür, dass er immer bemüht war mir den Rücken freizuhalten und mir geholfen hat mich auf das Wesentliche dieser Arbeit zu konzentrieren.

Dann gilt mein Dank der gesamten **Bioprozesstechnik Arbeitsgruppe**, die für gute Arbeitsbedingungen und ein kollegiales Miteinander gesorgt hat. Insbesondere möchte ich mich bei meinen Mitdoktoranden **Frau Dipl.-Biol. Sabine Kluge, Herrn Dipl.-Ing Fabian Kohrs, Herrn M.Sc. Sebastian Püttker** und **Herrn Dipl.-Ing. Marcus Hoffmann** für den fachlichen Austausch und die Zusammenarbeit sowie für das freundschaftliche Miteinander bedanken. Weiterer Dank gebührt **Frau Corina Siewert** für ihre Unterstützung im Labor und für die gute Stimmung, die sie verbreitet hat.

Weiterhin danke ich meinen Studenten **Frau Jennifer Rüdiger, Frau Rebekka Düsedau, Herrn Alexander Behne** und **Herrn To Dinh Nguyen Ngu**, deren Abschluss-

arbeiten ich betreuen durfte. Ihre Arbeiten waren sehr hilfreich für meine Dissertation. Auch hat es mir sehr viel Freude bereitet mit Ihnen zusammenzuarbeiten.

Herzlich möchte ich auch unseren Kooperationspartnern danken, mit denen wir im regen wissenschaftlichen Austausch standen und die uns mit Proben von Biogasanlagen versorgt haben. Insbesondere danke ich **Herrn Dr. rer. nat. Michael Klocke**, **Frau Dr. rer. nat. Angelika Hanreich** und **Frau Dr. rer. nat. Susanne Theuerl** vom Leibniz Institut für Agrartechnik Potsdam-Bornim, sowie **Herr Dr. Alexander Sczyrba**, **Herr Dr. Andreas Schlüter** und **Frau M.Sc. Irena Maus** vom CeBiTec Bielefeld.

Herrn Prof. Dr.-Ing. Richard Hanke-Rauschenbach und **Frau Dipl.-Ing. Astrid Bensmann** aus der Systemverfahrenstechnikarbeitsgruppe der Otto-von-Guericke-Universität Magdeburg danke ich für die außerordentlich angenehme Zusammenarbeit und die Einblicke in die Modellierung des Biogasprozesses.

Als sehr angenehm empfand ich den Austausch nach Gent (Belgien) in die Arbeitsgruppe „Laboratory of Microbial Ecology and Technology“. Ich möchte daher dieser Arbeitsgruppe, insbesondere **Herrn Prof. Dr. ir. Nico Boon** und **Herrn ir. Jo De Vrieze**, und dem Deutschen Akademischen Austauschdienst für die finanzielle Unterstützung während des Austausches meinen Danken aussprechen.

Mein besondere Dank gilt meinen **Eltern**, meiner **Schwiegermutter** und meiner **Frau Anne Heyer**, die mich unterstützt haben und immer für mich da waren.

Zusammenfassung

Der Abbau von biologischen Abfällen zu Methan in Biogasanlagen durch eine komplexe mikrobielle Gemeinschaft stellt eine der wichtigsten erneuerbaren Energiequellen dar. Voraussetzung für die Erhöhung der Methanproduktion ist ein detailliertes Verständnis über die Funktionsweise der beteiligten mikrobiellen Gemeinschaften. Daher sollte im Rahmen dieser Promotionsarbeit die taxonomische und funktionelle Zusammensetzung der mikrobiellen Gemeinschaften in möglichst vielen industriellen Biogasanlagen basierend auf den Proteinen (Metaproteom) charakterisiert werden. Zusätzlich sollte der Einfluss einzelner Prozessparameter auf die taxonomische und funktionelle Zusammensetzung der mikrobiellen Gemeinschaften untersucht und mögliche Markerproteine für die entsprechenden Betriebsweisen identifiziert werden.

Zu diesem Zweck wurde ein zuverlässiger und hoch auflösender Metaproteomworkflow etabliert und weiterentwickelt. Dieser umfasste die Extraktion der Proteine mit Phenol in einer Kugelmühle, das Einlaufen der Proteine in eine SDS-PAGE, den tryptische Verdau und die Messung mittels Elite Hybrid Ion Trap-Orbitrap MS/MS. Für die anschließende Proteinidentifikation wurde mit Mascot gegen UniProtKB/Swiss-Prot und mehrere Metagenome gesucht, bevor die Proteine nach UniRef50 Clustern gruppiert wurden und die taxonomische und funktionelle Auswertung mit der neu entwickelten MetaProteomeAnalyzer Software erfolgte. Mit diesem Workflow konnten 40 Biogasanlagenproben analysiert und pro Biogasanlage bis zu 859 Metaproteine identifiziert werden.

Die mikrobiellen Gemeinschaften von robust arbeitenden Biogasanlagen änderten sich wenig und waren funktionell sehr ähnlich. Schlüsseltaxa waren *Bakterien* der Ordnungen *Bacillales*, *Enterobacteriales*, *Bacteriales* und *Clostridiales* sowie die Archaeenordnungen *Methanobacteriales*, *Methanosarcinales* und *Methanococcales*. Wesentliche Stoffwechselfade von Biogasprozessen ausgehend von der Hydrolyse bis zur Methanogenese konnten durch die identifizierten Metaproteine belegt werden. Trotz der großen funktionellen Ähnlichkeit der mikrobiellen Gemeinschaften in den verschiedenen Biogasanlagen zeigte die Cluster- und Hauptkomponentenanalyse die Existenz von vier Gruppen von mikrobiellen Gemeinschaften. Spezifisch für die einzelnen Gruppen waren entweder thermophile oder mesophile Prozesstemperaturen sowie „upflow anaerobic sludge blanket“-Reaktoren oder Klärschlamm als Substrat. Weiterhin wurde die taxonomische und funktionelle Zusammensetzung der mikrobiellen Gemeinschaften durch die Raumbelastung, die Schlammverweilzeit sowie die Gesamtstickstoffmenge in den Biogasanlagen beeinflusst. Passend da-

zu wurden die Proteine 5,10-Methyltetrahydromethanopterinreduktase (Q8TXY4 *Euryarchaeota*) und Methyl Coenzyme M Reduktase (P07962, *Methanosarcina barkeri str. Fusaro*) als Marker für die Gesamtstickstoffkonzentration bzw. Raumbelastung in Biogasanlagen identifiziert. Das Vorhandensein der Ordnung *Thermotogales* zeigte den Betrieb der Biogasanlagen unter thermophilen Bedingungen an.

Metaproteomuntersuchungen erlauben somit nicht nur die taxonomische und funktionelle Beschreibung der am Biogasprozess beteiligten mikrobiellen Gemeinschaften, sondern ermöglichen auch die Identifikation von geänderten Prozessbedingungen oder Prozessstörungen.

Abstract

Degradation of biogenic waste to methane by complex microbial communities in a biogas plant represents a major renewable energy source. Detailed knowledge about the functioning of the microbial communities is a precondition for the improvement of the methane production. Thus, this PhD thesis characterized the taxonomic and functional structure of microbial communities from several industrial biogas plants based on the proteins (metaproteome). Of particular interest was the impact of process parameters on the taxonomic and functional composition of microbial communities and the identification of marker proteins for certain process parameters.

Therefore, a reliable and high resolution metaproteome workflow was established and improved. The workflow comprised protein extraction with phenol in a ball mill, entering of proteins into a SDS-PAGE, tryptic digestion, and measurement with an Elite Hybrid Ion Trap-Orbitrap MS/MS. For protein identification a Mascot search against UniProtKB/Swiss-Prot database and several metagenomes was carried out. Afterwards proteins were grouped according to the Uniref50 clusters, followed by taxonomic and functional evaluation of the results using the recently established software tool MetaProteomeAnalyzer. Forty biogas plant samples were analyzed with this workflow, and up to 859 metaproteins were identified for each biogas plant.

Microbial communities of robust operating biogas plants were stable over time and showed similar biological functions. Key taxa were *Bacteria* belonging to the orders *Bacillales*, *Enterobacteriales*, *Bacteriales*, and *Clostridiales* as well as *Archaea* belonging to the orders *Methanobacteriales*, *Methanosarcinales*, and *Methanococcales*. Major metabolic pathways of the biogas processes were identified by corresponding metaproteins starting from hydrolysis to methanogenesis. Despite a high functional similarity within the microbial communities in biogas plants, cluster and principal component analyses showed the existence of four groups of microbial communities. Specific for the discrete groups were thermophilic or mesophilic process regimes as well as upflow anaerobic sludge blanket reactors or sewage sludge as substrate. Furthermore, organic loading rate, sludge retention time as well as total ammonia nitrogen amount affected the taxonomic and functional composition of microbial communities. In line the proteins 5,10-methylenetetrahydromethanopterin reductase (Q8TXY4 *Euryarchaeota*) and methyl-coenzyme M reductase (P07962, *Methanosarcina barkeri str. Fusaro*) were identified as markers for total ammonia concentration

resp. organic loading rate in biogas plants. Moreover, the presents of the order *Thermotogales* indicated operation of the biogas plants under thermophilic conditions.

In conclusion, metaproteome analysis enabled not only the taxonomic and functional description of the microbial communities of biogas plants, but also the monitoring of changes in process conditions or process disturbances.

Inhaltsverzeichnis

| | |
|--|-------------|
| Danksagung | V |
| Zusammenfassung/Abstract | VIII |
| Inhaltsverzeichnis | XII |
| Abkürzungsverzeichnis | XIII |
| Einheitenverzeichnis | XVI |
| 1 Einleitung | 1 |
| 2 Theorie | 3 |
| 2.1 Biogasanlagen | 3 |
| 2.1.1 Biogasprozesse | 3 |
| 2.1.2 Aufbau von Biogasanlagen | 4 |
| 2.1.3 Aufbau von Biogasreaktoren und wichtige Prozessparameter | 5 |
| 2.1.3.1 Reaktordesign | 6 |
| 2.1.3.2 pH Wert | 7 |
| 2.1.3.3 Prozesstemperatur | 8 |
| 2.1.3.4 Inoculum | 8 |
| 2.1.3.5 Substrat | 9 |
| 2.2 Mikrobieller Abbau von Biomasse | 11 |
| 2.2.1 Hydrolyse | 11 |
| 2.2.2 Gärung | 11 |
| 2.2.3 Acetogenese | 14 |
| 2.2.4 Methanogenese | 16 |
| 2.3 Mikrobielle Gemeinschaften | 17 |
| 2.3.1 Methoden zur Charakterisierung von mikrobiellen Gemeinschaften . | 17 |
| 2.3.2 Mikrobielles Ressourcen Management | 19 |

| | | |
|----------|--|-----------|
| 2.4 | Methoden der Proteinanalytik | 21 |
| 2.4.1 | Proteine | 22 |
| 2.4.2 | Proteinextraktion | 24 |
| 2.4.3 | Bestimmung der Proteinkonzentration | 25 |
| 2.4.4 | Protein- und Peptidseparation | 26 |
| 2.4.5 | Proteinverdau | 27 |
| 2.4.6 | Massenspektrometrie | 27 |
| 2.4.7 | Massenspektrometrie-basierte Proteinidentifizierung | 30 |
| 2.4.8 | Bestimmung der Proteinfunktion und Taxonomie | 33 |
| 2.4.9 | Quantifizierung einzelner Proteinen | 34 |
| 2.5 | Statistische Methoden zur Datenanalyse | 36 |
| 3 | Material und Methoden | 38 |
| 3.1 | Beprobung von Biogasanlagen | 38 |
| 3.2 | Probenvorbereitung | 38 |
| 3.2.1 | Phenolextraktion | 38 |
| 3.2.2 | Proteinquantifizierung mittels Amidoschwarz | 39 |
| 3.2.3 | Acetonfällung | 39 |
| 3.3 | Proteinseparation | 39 |
| 3.3.1 | SDS-PAGE | 40 |
| 3.3.2 | IEF mit dem 3100 OFFGEL-Fraktionator | 40 |
| 3.4 | Visualisierung und Quantifizierung einzelner Proteine | 40 |
| 3.4.1 | Kolloidale Coomassiefärbung | 41 |
| 3.5 | Tryptischer Verdau und LC-MS/MS Messung | 41 |
| 3.5.1 | Tryptischer Verdau | 41 |
| 3.5.2 | Massenspektrometrie | 42 |
| 3.6 | Bioinformatische und statistische Datenauswertung | 42 |
| 3.6.1 | Proteinidentifikation mittels Datenbanksuche | 42 |
| 3.6.2 | Taxonomische und funktionale Datenauswertung mit dem MetaPro- teomeAnalyzer | 43 |
| 3.6.3 | Statistische Auswertung mit Matlab | 43 |
| 3.6.4 | Erstellung und Analyse von Graphnetzwerken mit Gephi | 44 |
| 3.6.5 | Multiples Sequenzalignment mit Jalview | 44 |
| 4 | Methodenoptimierung | 45 |
| 4.1 | Optimierung der Proteinseparation | 45 |
| 4.1.1 | Optimierung der Beladungsmenge für die SDS-PAGE | 46 |

| | | |
|----------|---|-----------|
| 4.1.2 | Robustheit, Reproduzierbarkeit und Trennqualität der SDS-PAGE . | 47 |
| 4.1.3 | Einfluss der Trenndimensionen auf die Anzahl an identifizierten Peptiden und Proteinen | 50 |
| 4.1.4 | Diskussion | 53 |
| 4.2 | Optimierung der LC-MS/MS Messung | 55 |
| 4.2.1 | Überprüfung des tryptischen Verdaus | 55 |
| 4.2.2 | Vergleich verfügbarer Massenspektrometer | 55 |
| 4.2.3 | Optimierung von LC-MS/MS Messungen | 56 |
| 4.2.4 | Reproduzierbarkeit von Orbitrap Elite Messungen | 58 |
| 4.2.5 | Diskussion | 61 |
| 4.3 | Optimierung der Proteinidentifikation | 63 |
| 4.3.1 | Auswahl von Proteindatenbanken für die Proteinidentifikation . . . | 63 |
| 4.3.2 | Qualität der Spektrenidentifikationen | 64 |
| 4.3.3 | Multiple Suchalgorithmen | 66 |
| 4.3.4 | Posttranslationale Modifikationen | 66 |
| 4.3.5 | <i>De novo</i> Sequenzierung und BLAST-Suche | 68 |
| 4.3.6 | Diskussion | 70 |
| 4.4 | Bioinformatische Datenauswertung mit dem MetaProteomeAnalyzer | 73 |
| 4.4.1 | Workflow MetaProteomeAnalyzer | 73 |
| 4.4.2 | Metaproteinkonzept | 74 |
| 4.4.3 | Diskussion | 81 |
| 5 | Korrelationen von mikrobiellen Gemeinschaften mit Prozessparametern | 83 |
| 5.1 | Langzeitstabilität von mikrobiellen Gemeinschaften in Biogasanlagen . . . | 83 |
| 5.1.1 | SDS-PAGE Profile von sechs landwirtschaftlichen Biogasanlagen . . | 84 |
| 5.1.2 | Charakterisierung der mikrobiellen Taxonomien und Funktionen . . | 87 |
| 5.1.3 | Diskussion | 89 |
| 5.2 | Vergleich einer mesophilen und einer thermophilen Biogasanlage | 91 |
| 5.2.1 | Taxonomie- und Funktionsprofile | 91 |
| 5.2.2 | Zuordnung der identifizierten Proteine zum Biogasprozess | 92 |
| 5.2.3 | Diskussion | 95 |
| 5.3 | Charakterisierung von mikrobiellen Gemeinschaften von 40 Biogasanlagen . | 97 |
| 5.3.1 | Beurteilung der Proteinextrakte und der Proteinidentifikationen . . | 98 |
| 5.3.2 | Clusteranalyse und Hauptkomponentenanalyse von verschiedenen mikrobiellen Gemeinschaften | 101 |
| 5.3.3 | Untersuchung der ökologischen Indizes Richness und Evenness . . . | 104 |
| 5.3.4 | Taxonomische Zusammensetzung von mikrobiellen Gemeinschaften | 106 |

| | |
|---|---------------|
| 5.3.5 Funktionale Zusammensetzung von mikrobiellen Gemeinschaften . . | 109 |
| 5.3.6 Zusammenhang zwischen mikrobiellen Gemeinschaften und Prozessparametern | 114 |
| 5.3.7 Diskussion | 118 |
| 6 Fazit und Ausblick | 123 |
| Literaturverzeichnis | 159 |
| Tabellenverzeichnis | 160 |
| Abbildungsverzeichnis | 162 |
| Publikationsübersicht | 163 |
| 7 Anhang | i |
| 7.1 Hemmstoffe | i |
| 7.2 Generationszeiten methanogener Archaeen | ii |
| 7.3 Biochemie von Biogasprozessen | iii |
| 7.4 Chemikalienliste | vi |
| 7.5 Geräteliste | xi |
| 7.6 Liste der beprobten Biogasanlagen | xii |
| 7.7 Komponenten und Parameter für das Elite Hybrid Ion Trap-Orbitrap MS/MS | xvii |
| 7.8 Komponenten und Parameter für das AmaZon ETD ESI-Ion Trap MS/MS | xx |
| 7.9 Multiples Sequenzalignment | xxiii |
| 7.10 Methanogenesestoffwechsel für eine mesophile und eine thermophile Biogasanlage | xxv |
| 7.11 Abhängigkeit der Richness und Evenness vom Schwellwert | xxvi |
| 7.12 Zusammenhänge zwischen Prozessparametern visualisiert als GPLOM . . . | xxvii |
| 7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern | xxviii |
| 7.14 CD | xxxvi |
| Lebenslauf | xxxvii |

Abkürzungsverzeichnis

| Abkürzung | Bedeutung |
|---|--|
| ADP | Adenosindiphosphat |
| AK | Antikörper |
| ATP | Adenosintriphosphat |
| BLAST | <i>basic local alignment search tool</i> |
| BGP | Biogasanlagenproben |
| C | Kohlenstoff |
| CH ₄ | Methan |
| CID | <i>collision-induced dissociation</i> |
| CO ₂ | Kohlenstoffdioxide |
| CO ₃ ²⁻ | Carbonation |
| Co | Coenzym |
| CoM | Coenzym M |
| CSTR | Rührkesselreaktoren |
| DBU | Deutsche Bundesstiftung Umwelt |
| DGGE | <i>denaturing gradient gel electrophoresis</i> |
| DNA | Desoxyribonukleinsäure |
| DTT | 1,4-Dithiothreitol |
| EC | <i>enzyme commission number</i> |
| EFTU | Elongationsfaktor TU |
| ESI | <i>electrospray ionization</i> |
| ETD | <i>electron-transfer dissociation</i> |
| GAPD | Glyceraldehyd-3-phosphatdehydrogenase |
| GPLOM | <i>generalized plot matrix</i> |
| H ⁺ | Hydroniumion bzw. Protonen |
| H ₂ | Wasserstoff |
| H ₂ CO ₃ bzw. HCO ₃ ⁻ | Kohlensäure bzw. Hydrogencarbonat |
| H ₂ O | Wasser |
| H4MPT | Tetrahydromethanopterin |

| Abkürzung | Bedeutung |
|-----------------------------------|--|
| HCD | <i>higher-energy collisional dissociation</i> |
| HF | Hauptfermenter |
| HPLC | <i>high performance liquid chromatography</i> |
| IAA | Ioadacetamid |
| KEGG | <i>Kyoto encyclopedia of genes and genomes</i> |
| KO | <i>KEGG orthology</i> |
| LC | <i>liquid chromatography</i> |
| m | milli |
| m/z-Verhältnis | Masse-zu-Ladungs-Verhältnis |
| MALDI | Matrix-unterstützte Laser-Desorption/Ionisation |
| MCR | Methyl Coenzym M Reduktase |
| MPA | MetaProteomeAnalyzer |
| MRM | Mikrobielles Ressourcen Management |
| mRNA | <i>messenger</i> -RNA |
| MS | Massenspektrometer/ Massenspektrometrie |
| MS/MS | Tandem-Massenspektrometer/ -Massenspektrometrie |
| N | Stickstoff |
| NADH | Nicotinamidadenindinukleotid |
| NADPH | Nicotinamidadenindinukleotidphosphat |
| NCBI | <i>national center for biotechnology information</i> |
| NH₃ | Ammoniak |
| NH₄⁺ | Ammonium |
| OH⁻ | Hydroxidion |
| O | Sauerstoff |
| P | Phosphor |
| PCR | Polymerasekettenreaktion |
| PTMs | posttranslationale Modifikationen |
| qPCR | <i>quantitative realtime</i> PCR |
| RNA | Ribonukleinsäure |
| rRNA | ribosomale RNA |
| S | Schwefel |
| SDS | <i>sodium dodecyl sulfate</i> |
| SDS-PAGE | SDS-Polyacrylamidgelelektrophorese |

| Abkürzung | Bedeutung |
|---|--|
| TRFLP | <i>terminal restriction fragment length polymorphism</i> |
| tRNA | Transfer-RNA |
| UASB | <i>upflow anaerobic sludge blanket</i> |
| UPLC | <i>ultra performance liquid chromatography</i> |
| 2D-PAGE | zweidimensionale Polyacrylamidgelelektrophorese |
| 5,10-Methylen- H4MPT Reduktase | 5,10-Methylen-tetrahydromethanopterinreduktase |

Symbolverzeichnis

| Symbol | Bedeutung | Einheit |
|------------------------|---|---|
| c | Stoffmengenkonzentration | [M] Molar, [mol/m ³] Mol je Kubikmeter |
| E | Energie | [J] Joule |
| emPAI | <i>exponentially modified protein abundance index</i> | [-] |
| f | Frequenz | [Hz] Hertz |
| FDR | <i>false discovery rate</i> | [-] |
| FM | Frischmasse | [g] Gramm |
| FOS/TAC | Verhältnis von flüchtigen organischen Säuren (FOS) zur Carbonat-Pufferkapazität (TAC) | [-] |
| f_{rot} | Drehzahl | [min ⁻¹] Umdrehung pro Minute |
| g | Gravitation | [m/s ²] Meter je Quadratsekunde |
| Hac | Essigsäureäquivalent | [mg/L] Milligramm je Liter |
| HRT | Hydraulische Verweilzeit | [d] Tagen |
| I | Stromstärke | [A] Ampere |
| m | Masse | [g] Gramm, [kg] Kilogramm, [Da] Dalton |
| NSAF | <i>normalized spectral abundance factor</i> | [-] |
| OLR | Raumbelastung | [kg _{oTS})/(m ³ d] Kilogramm organische Trockensubstanz je Kubikmeter Reaktorvolumen und Tag |

| Symbol | Bedeutung | Einheit |
|--------------|----------------------------|---|
| oTS | organische Trockensubstanz | [g] Gramm |
| p | Druck | [atm] Atmosphären, [Pa] Pascal, [bar] Bar |
| pI | isolektrische Punkt | [-] |
| pm | Partialdruck | [Pa] Pascal, [bar] Bar |
| n | Stoffmenge | [mol] Mol, [ppm] <i>parts per million</i> |
| RT | Raumtemperatur | [°C] Grad Celsius |
| r | Korrelationskoeffizient | [-] |
| t | Zeit | [s] Sekunde, [min] Minute, [h] Stunden, [d] Tag, [a] Jahr |
| T | Temperatur | [°C] Grad Celsius, [K] Kelvin |
| TAN | Gesamtammoniumstickstoffk | [mg/L] Milligramm je Liter |
| TS | Trockengewicht | [g] Gramm |
| V | Volumen | [m ³] Kubikmetern |
| (v/v) | Volumenprozente | [% vol] |

Einleitung

Biogasanlagen haben sich seit dem Erneuerbare-Energien-Gesetz im Jahr 2000 zu einer der wichtigsten Möglichkeiten zur Bereitstellung von Energie aus erneuerbaren Energiequellen entwickelt. In diesen bauen komplexe mikrobielle Gemeinschaften Energiepflanzen, Gülle, Mist und Pflanzenreste zu Biogas ab, welches aus Methan (CH_4) und Kohlenstoffdioxid besteht (CO_2). Das erzeugte Biogas wiederum wird in Blockheizkraftwerken zur Produktion von Strom und Wärme genutzt. Vorreiter beim Einsatz dieser erneuerbaren Energiequelle war die Bundesrepublik Deutschland. Im Jahr 2014 wurden von knapp 8.000 Biogasanlagen circa 4,7% der Gesamtstrommenge produziert [Agentur für erneuerbare Energien (2014), Bundesministerium für Wirtschaft und Energie (2015)]. Damit die Bundesrepublik Deutschland den kompletten Ausstieg aus fossilen Energiequellen schaffen kann, wird allerdings ein weiterer Zuwachs erneuerbarer Energiequellen benötigt. Auf Grund des steigenden Bedarfs an Ackerflächen für die Nahrungsmittelproduktion ist eine weitere Steigerung der Energiepflanzenproduktion für Biogasanlagen sehr schwierig. Folglich kann die Stromproduktion durch Biogasanlagen nur durch die Erhöhung der Produktivität und Ausbeute der Biogasanlagen oder die Erschließung von neuen Substraten gesteigert werden, wie z. B. durch die Nutzung von Grünschnitten oder anderen Pflanzenresten [Angelidaki et al. (2011)].

Eine wesentliche Voraussetzung für die weitere Optimierung von Biogasanlagen ist es, die genaue Funktionsweise der mikrobiellen Gemeinschaften in Biogasanlagen besser zu verstehen. Insbesondere ist es wichtig zu analysieren, welche Mikroorganismen die Biomasse abbauen, welche Stoffwechselwege sie dafür benutzen und mit welchen anderen Mikroorganismen sie dabei zusammenarbeiten [Verstraete et al. (2007)]. Ferner gilt es zu klären, welchen Einfluss die einzelnen Prozessparameter auf die mikrobiellen Gemeinschaften haben.

Während in bisherigen Untersuchungen von Biogasprozessen meist nur die chemischen und technischen Parameter von Biogasanlagen [Weiland (2010), Ward et al. (2008), Holm-Nielsen et al. (2009), Appels et al. (2008), Angelidaki et al. (2011)] sowie die theoretischen Abbauewege der Biomasse [Schink (1997)] untersucht wurden, ist die Charakterisierung

von mikrobiellen Gemeinschaften ein relativ junges Forschungsgebiet. Sie wird ermöglicht durch die jüngere Entwicklung neuer molekularbiologischer [Hofman-Bang et al. (2003)] und Omiks-Methoden wie Metagenomik [Schlüter et al. (2008)], Metatranskriptomik [Zakrzewski et al. (2012)] und Metaproteomik [Wilmes and Bond (2006)].

Molekularbiologische Untersuchungen der taxonomischen Zusammensetzungen von Biogasanlagenproben anhand des in allen *Archaeen* und *Bakterien* vorkommenden 16S-rRNA Gens zeigten das Vorhandensein von bestimmten Schlüsseltaxonomien in Biogasanlagen [Hofman-Bang et al. (2003), Klocke et al. (2008)]. Aber erst durch die Entwicklung von hochdurchsatzfähigen Gensequenzierungsverfahren wie 454-Pyrosequenzierung [Margulies et al. (2005)] und Illumina-Sequenzierung [Bentley et al. (2008)] war es möglich, die komplette taxonomische und funktionelle Zusammensetzung von mikrobiellen Gemeinschaften zu charakterisieren [Schlüter et al. (2008), Wirth et al. (2012)]. Nachteilig bei diesen Methoden ist, dass das Vorhandensein von Genen nicht die Aktivität der einzelnen Spezies oder metabolischer Wege nachweist. Eine bessere Korrelation zu den tatsächlichen Aktivitäten kann durch die Betrachtung der transkribierten Gene (Transkriptom) [Zakrzewski et al. (2012)] oder der translatierten Proteine (Proteom) [Wasinger et al. (1995)] erreicht werden. Zwar sind Untersuchungen auf Proteinebene dank der hohen Proteinstabilität robuster als Untersuchungen auf der Transkriptomebene, wurden aber für die Charakterisierung von komplexen mikrobiellen Gemeinschaften (Metaproteomanalyse) nur selten eingesetzt. Wesentliche Gründe dafür waren die schwierige Extraktion der Proteine aus der Probenmatrix, der Mangel an geeigneten Proteindatenbanken für die Proteinidentifikation und die große Probenkomplexität [Becher et al. (2013), Benndorf and Reichl (2014), Hettich et al. (2013), Muth et al. (2013a), Otto et al. (2014), Seifert et al. (2013), Siggins et al. (2012), Wöhlbrand et al. (2013)]. Diese Probleme zeigten sich auch in den Ergebnissen bisheriger Metaproteomstudien über Biogasprozesse [Abram et al. (2011), Abram et al. (2009), Hanreich et al. (2012)], welche für Laborbiogasanlagen immerhin die Machbarkeit von Metaproteomstudien von Biogasanlagenproben belegten. Allerdings wurden diese Laborbiogasanlagen teilweise nur mit synthetischem Medium betrieben. Praxisrelevanter, aber wegen den großen Mengen an Probenverunreinigungen schwieriger als die Untersuchung von Laborbiogasanlagen, ist die Charakterisierung der taxonomischen und funktionellen Zusammensetzung der mikrobiellen Gemeinschaften in industriellen Biogasanlagen. Dafür sollte im Rahmen dieser Promotion ein zuverlässiger und hoch auflösender Metaproteomworkflow etabliert und zur Untersuchung von möglichst vielen industriellen Biogasanlagen genutzt werden. Ziel dieser Untersuchungen sollte neben der Beschreibung der mikrobiellen Gemeinschaften die Untersuchung des Einflusses einzelner Parameter sein, wie z. B. der Prozesstemperatur. Ausgehend von diesen Ergebnissen sollte zusätzlich nach Markerproteinen für die einzelnen Parameter gesucht werden.

Theorie

2.1 Biogasanlagen

2.1.1 Biogasprozesse

Biogasanlagen sind neben Wind, Solar- und Wasserkraftanlagen eine der wichtigsten Möglichkeiten für die Bereitstellung von Energie aus erneuerbaren Energieformen. So gibt es allein in der Bundesrepublik Deutschland 8000 Biogasanlagen, welche ca. 4,7% der jährlichen Stromproduktion und fast 1% des Wärmebedarfs bereitstellen [Agentur für erneuerbare Energien (2014), Bundesministerium für Wirtschaft und Energie (2015)].

In Biogasanlagen werden unter Luftabschluss biologische Abfälle wie Pflanzensilage, Gülle, Mist, sowie Lebensmittelreste aus Haushalt und Lebensmittelindustrie zu Biogas abgebaut. Biogas wiederum ist ein hervorragender Energieträger für die Produktion von Strom und Wärme in Blockheizkraftwerken und kann nach entsprechender Aufbereitung sogar im Erdgasnetz gespeichert werden. Ein weiterer Vorteil gegenüber der reinen Verbrennung von Biomasse ist, dass für Biogasprozesse auch Substrate mit einem Wasseranteil von über 50% genutzt werden können. Aus ökologischer Sicht entspricht der Biogasprozess der Nutzung der freiwerdenden Energie des natürlichen Biomasseabbaus und passt sich somit in die natürlichen Stoffkreisläufe ein. Im Gegensatz zur anaeroben Zersetzung von Biomasse zu CH_4 und CO_2 in Sedimenten, Gewässern, Tiermägen und Reisfeldern wird das klimaschädliche CH_4 in Biogasanlagen aufgefangen und zu CO_2 verbrannt. Zusätzlich enthält der verbleibende Gärrest noch sämtlichen Stickstoff, Phosphor sowie Mineralstoffe und Spurenelemente. Er ist daher ein hervorragender Dünger [Ward et al. (2008)], womit sich der Stoffkreislauf schließt. Im Vergleich zu Gülle als Dünger hat der Gärrest aus Biogasanlagen den Vorteil, dass geruchsbelästigende Stoffe abgebaut werden [Smet et al. (1999)] und der Anteil an pathogenen Mikroorganismen deutlich reduziert wird [Ward et al. (2008), Sahlström (2003), Eikmeyer et al. (2013)].

Ein Hauptproblem von Biogasprozessen ist, dass faserige Materialien wie Cellulose und Lignin sowie aromatische Verbindungen (z. B. einige Aminosäuren) nur schlecht abgebaut

werden können [Schieder et al. (2013), Seite 24]. Insbesondere Gras und Stroh enthalten jedoch bis zu 80% Lignin [Schieder et al. (2013), Seite 26]. Aber auch Gülle kann nur zu 40-65% abgebaut werden [Angelidaki et al. (2011)]. Im Umkehrschluss folgt daraus, dass bei nahezu kompletter Verwertung der Cellulose und des Lignins die Ausbeuten von Biogasprozessen verdoppelt werden könnten. Mögliche Ansätze liefern dabei Termiten, Kühe und Schafe, welche in der Lage sind, Lignin und Cellulose mit den mikrobiellen Gemeinschaften in ihrem Verdauungstrakt effektiv zu verstoffwechseln [Toyoda et al. (2009), Burnum et al. (2010)]. Wenn es möglich wäre, diese Funktionsweise auf Biogasanlagen zu übertragen, könnte Biogasprozesse auch mit Pflanzenresten wirtschaftlich betrieben werden. Dadurch könnte eine Konkurrenz von landwirtschaftlichen Flächen für die Produktion von Nahrungsmitteln und Substraten für Biogasanlagen vermieden werden.

In Zukunft könnte die Bedeutung von Biogasprozessen noch weiter ansteigen. Sollte es z. B. bei der Abwasserbehandlung zum Paradigmenwechsel kommen und Fäkalien anstatt verdünnt, gesammelt werden, könnten diese als neue Substrate für den Biogasprozess dienen. Gegenüber der normalen, aeroben Abwasserbehandlung entsteht dabei auch weniger Klärschlamm, welcher anschließend aufwendig entsorgt werden muss. [Angelidaki et al. (2011)]. Eine andere potentielle Anwendung des Biogasprozesses ist die Speicherung von Energie [Luo et al. (2012), Bensmann et al. (2014)]. Bei einem Überschuss an Strom ist es möglich durch Elektrolyse Wasser zu Wasserstoff (H_2) und Sauerstoff (O) zu spalten. Der H_2 kann dann in Biogasanlagen eingespeist und zusammen mit überschüssigem CO_2 in CH_4 umgewandelt werden.

2.1.2 Aufbau von Biogasanlagen

Biogasanlagen bestehen im wesentlichen aus einem luftdichtem Fermenter dem meist (semi)-kontinuierlich frische Biomasse zugesetzt und Gärrest entnommen wird [Abbildung: 2.1]. Im Fermenter selbst sorgen Horizontal- oder Tauchmotorrührwerke für eine gleichmäßige Durchmischung. Das entstehende Biogas wird im Fermenterkopf gesammelt und nach biologischer oder chemischer Entschwefelung in einem Blockheizkraftwerk verbrannt. Alternativ kann mittels Gaswäsche CO_2 und Schwefelwasserstoff aus dem Biogas entfernt und das verbleibende CH_4 in das Erdgasnetz eingespeist werden [Appels et al. (2008), Fachagentur Nachwachsende Rohstoffe (2010), Seite 112]. Das genaue Design von Biogasanlagen variiert je nach Anlagenbetreiber, Substratmenge- und art. Neben Rührkeselfermentern werden auch Membran-, Aufstromrohr- (UASB) und Strömungsrohrfermenter eingesetzt. Der anfallende Gärrest wird meist in Nachgärfermentern bis zur weiteren Verwendung gelagert, wobei diese noch zu 5-15% der Gasproduktion beitragen können. Bauliche Besonderheit einiger Biogasanlagen sind Anmischgruben oder vorgeschaltete

Hydrolysestufen. Eine Anmischgrube ist ein Behälter, in der das Substrat vor der Zugabe in den Hauptfermenter für eine bessere Viskosität mit Gärresten gemischt wird. Mit dem Einbau einer Hydrolysestufe wird bezweckt, dass die Schritte des Biogasprozesses [Kapitel: 2.2] räumlich getrennt werden. Dabei sollen Hydrolyse und Teile der Fermentation in der Hydrolysestufe unter optimalen pH-Bedingungen ablaufen [Weiland (2010)].

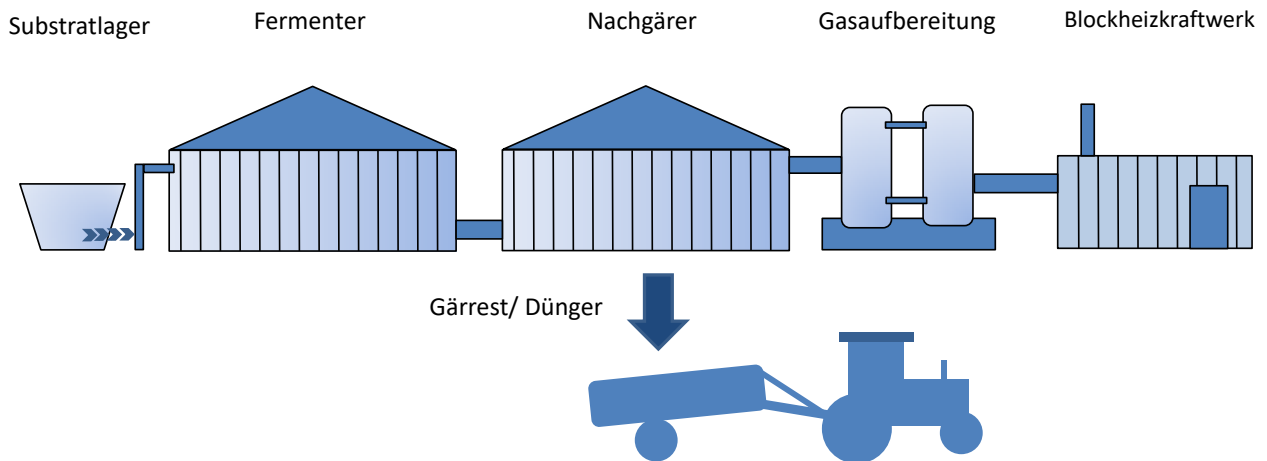


Abbildung 2.1: *Allgemeiner Aufbau einer Biogasanlage.*

2.1.3 Aufbau von Biogasreaktoren und wichtige Prozessparameter

Für den Betreiber einer Biogasanlage sind die Menge an produziertem Biogas und die Biogasqualität die wichtigsten Prozessgrößen. Neben der reinen Volumenmessung des produzierten Biogases wird dafür die Biogaszusammensetzung regelmäßig mit Hilfe von entsprechenden Gassonden oder Gaschromatographen überprüft. Hochwertiges Biogas zeichnet sich durch eine hohe Methankonzentration von über 50% und geringe Mengen an Schwefelwasserstoff ($< 500 \text{ mg/m}^3$) und Ammoniak ($< 0,7 \text{ mg/m}^3$) aus.

Um eine möglichst hohe Produktivität bei Biogasanlagen zu erzielen, ist es notwendig, dass die mikrobiellen Gemeinschaften möglichst gute Lebensbedingungen vorfinden. Wesentlichen Einfluss auf die mikrobiellen Gemeinschaften haben neben dem Reaktordesign [Kapitel: 2.1.3.1] der pH-Wert [Kapitel: 2.1.3.2], die Prozesstemperatur [Kapitel: 2.1.3.3], das Inoculum [Kapitel: 2.1.3.4] und das Substrat [Kapitel: 2.1.3.5]. Für die einzelnen Parameter gibt es allgemeine Empfehlungen [Tabelle: 2.1], welche in den einzelnen Abschnitten diskutiert werden. Allerdings müssen diese auf die verfügbaren Substrate, sowie die ökonomischen und baulichen Gegebenheiten angepasst werden.

2.1 Biogasanlagen

Tabelle 2.1: *Empfohlene Prozessparameter für den Biogasprozess*

basierend auf [Weiland (2010), Jørgensen (2009), Seite 5, Chen et al. (2008), Lossie and Pütz (2008)]. Abkürzungen: FOS/TAC Verhältnis der flüchtigen organischen Säuren zur Carbonat-Pufferkapazität, HRT hydraulische Verweilzeit, OLR Raumbelastung, TAN Gesamtammoniumstickstoff.

| Prozessparameter | Empfohlen | Problematisch |
|--------------------------|---|-----------------------------|
| pH-Wert | | |
| pH | 6,8-7,2 | < 6,6 |
| FOS/TAC | < 0,3 | > 0,6 |
| Prozesstemperatur | | |
| Temperaturbereich | psychrophil (<30 °C) mesophil (30-44 °C) thermophil (45–55 °C) | <0 °C |
| Substrat | | |
| HRT | 10-100 Tage | |
| OLR | 2-4 kg/(d*m ³) | < 10 kg/(d*m ³) |
| Substrat | Kohlenstoff:Stickstoff:Sulfat:Phosphor-Verhältnis: 600:15:5:1 | |
| TAN | 200 mg/L | 1.000-6.000 mg/L |
| Spurenelemente | Barium (Ba) 0,05 g/m ³ Calcium (Ca) 0,03 g/m ³ Cobalt (Co) 0,005 g/m ³ Eisen (Fe) 0,2 g/m ³ Magnesium (Mg) 0,02 g/m ³ Molybdenum (Mo) 0,005 g/m ³ Nickel (Ni) 0,01 g/m ³ | |

2.1.3.1 Reaktordesign

Wie schon vorab diskutiert [Kapitel: 2.1.2], kommen in Biogasanlagen verschiedene Reaktordesigns und Konzepte sowie deren Kombination zum Einsatz. Prinzipiell wird dabei in Nass- und Trockenfermentation unterschieden. Der Unterschied zwischen beiden Fermentationen ist, dass bei Trockenfermentationen gröbere Substrate mit höheren Trockengewichte (TS) eingesetzt werden, aber dafür nicht gerührt wird. Passend dazu gibt es drei grundsätzliche Reaktortypen für Biogasanlagen: Rührkesselreaktoren (CSTR), Propfenstromreaktoren und *upflow anaerobic sludge blanket* (UASB)-Reaktoren. CSTR werden meist in landwirtschaftlichen Biogasanlagen mit größeren Gülleanteilen und TS von bis zu 10% genutzt. Bei höheren TS und sehr faserigen Substraten eignen sich Propfenstromre-

aktoren besser. Im Gegensatz dazu werden UASB-Reaktoren bei sehr kleinen TS, wie bei Industrieabwässern und Klärschlamm, eingesetzt. In vielen Biogasanlagen werden auch mehrere Reaktoren kombiniert um für die verschiedenen Mikroorganismen und Prozessstufen optimale Bedingungen zu schaffen und somit die Biogasproduktion zu maximieren [Bensmann et al. (2013)].

Einen eher sekundären Einfluss auf die mikrobielle Gemeinschaft haben die Rührwerke. Diese sorgen für eine gleichmäßige Durchmischung des Substrates und verhindern die Ausbildung von Schwimmdecken [Weiland (2010), Lienen et al. (2013)]. Dadurch sorgen sie für eine bessere Verfügbarkeit der Nährstoffe für die Mikroorganismen und einen stabileren pH-Wert. Zwar kann durch zu starkes Rühren die Interaktionen zwischen den Mikroorganismen gestört werden, aber da aus ökonomischen Gründen Biogasreaktoren so wenig wie möglich gerührt werden, ist dies ein eher untergeordnetes Problem.

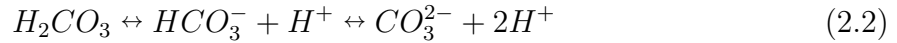
2.1.3.2 pH Wert

Der pH-Wert, welcher meist online über eine pH-Sonde gemessen wird, ist ein Schlüsselparameter von Biogasprozessen. Die methanproduzierenden *Archaeen* haben ein pH Optimum zwischen pH 6,8-7,2 und werden inhibiert bzw. abgetötet, sobald der pH-Wert unter pH 6,6 fällt [Ward et al. (2008)]. Bei einer zu schnellen Substratzufuhr (zu hohen Raumbelastung (OLR)) wird die Biomasse zu schnell zu kurzkettigen Fettsäuren abgebaut, welche akkumulieren. Als Konsequenz fällt der pH-Wert unter den für die methanproduzierenden *Archaeen* tolerierbaren Wert und die Biogasprozesse kommen zum Erliegen [Munk et al. (2010)]. In diesem Fall wird von einer Übersäuerung der Biogasanlage gesprochen.

Neben zu niedrigen pH-Werten sind aber auch zu hohe pH-Werte ungünstig, da Ammonium (NH_4) dann zunehmend zu Ammoniak (NH_3) dissoziiert [Gleichung: 2.1]. Während viele Mikroorganismen relative hohe Mengen von NH_3 tolerieren können, ist es schon in niedrigen Konzentrationen für die methanproduzierenden *Archaeen* toxisch [Anhang: 7.1]. Folglich wird in Biogasanlagen regelmäßig die Gesamtstickstoffkonzentration (TAN) untersucht und daraus die Menge an NH_4 und NH_3 berechnet. Kritisch wird dies besonders bei höheren Prozesstemperaturen, bei denen das Dissoziationsgleichgewicht noch weiter in Richtung des NH_3 verschoben ist. Um in solchen Fällen eine Ammoniakvergiftung zu verhindern, kann NH_3 durch Fällungen oder Strippen aufwendig entfernt [Kabdasli et al. (2000)] oder zumindest durch Verdünnen mit stickstoffarmen Substraten reduziert werden [Chen et al. (2008)].

Bei der Regelung des pH-Wertes einer Biogasanlagen sollte beachtet werden, dass der pH-Wert durch NH_4 in der Gärflüssigkeit [Gleichung: (2.1)] und CO_2 in Form von Kohlensäure (HCO_3^-) [Gleichung: (2.2)] gepuffert wird. Deshalb bleibt der pH-Wert lange stabil bis die Pufferkapazität erschöpft ist und ändert sich dann schlagartig. Daher wird, um die

Pufferbedingungen in einer Biogasanlage zu berücksichtigen, zusätzlich zum pH-Wert das Verhältnis der flüchtigen organischen Säuren zur Carbonat-Pufferkapazität (FOS/TAC Wert) bestimmt [Rieger and Weiland (2006)].



2.1.3.3 Prozesstemperatur

Für den Betrieb von Biogasanlagen werden zwei Temperaturbereiche unterschieden: mesophil (30-44 °C) und thermophil (45–55 °C) [Schieder et al. (2013)]. In diesen dominieren jeweils die Mikroorganismen, die speziell für diesen Temperaturbereich angepasst sind. Da sich die Geschwindigkeit von biochemischen Reaktionen bei einer Erhöhung der Temperatur um 10 K in etwa verdoppelt und dadurch die Hydraulische Verweilzeit (HRT) durch den schnelleren Abbau der Biomasse verringert werden kann, sind höhere Prozesstemperaturen von Vorteil. Die Keimbelastung im Gärrest wird zusätzlich durch die höhere Prozesstemperatur verringert. In der Praxis dominieren allerdings mesophile Biogasanlagen. Diese sind wirtschaftlicher, weil die Biogasfermenter weniger beheizt werden müssen und weisen eine höhere Prozessstabilität auf. So werden mesophile Biogasanlagen von ± 3 K Temperaturschwankungen nicht gestört [Weiland (2010)], während sich die Temperatur von thermophilen Biogasanlagen nur um maximal $\pm 0,5$ K [Ward et al. (2008)] ändern sollte.

Theoretisch könnten Biogasanlagen auch psychrophil (<30 °C) betrieben werden. Bis jetzt gibt es aber aufgrund der langsameren Geschwindigkeit der biochemischen Reaktionen bei niedrigeren Temperaturen nur Labor- und Pilotanlagen für die Behandlung von Abwasser [Lettinga et al. (2001), Bialek et al. (2013)].

Ein weiterer Effekt von Temperaturänderungen in Biogasanlagen ist, dass sich mit steigender Prozesstemperatur die Löslichkeit der Gase in der Gärflüssigkeit reduziert. Dies beeinflusst zwar die Biogasproduktion nicht sonderlich, kann aber erheblichen Einfluss auf die taxonomische und funktionelle Zusammensetzung der mikrobiellen Gemeinschaften haben. Zum Beispiel wird durch Temperaturerhöhung der Wasserstoffpartialdruck reduziert, wodurch alternative Pfade des Biogasprozesses [Kapitel: 2.2] für die Mikroorganismen energetisch günstiger werden.

2.1.3.4 Inoculum

Für das Anfahren von Biogasanlagen werden diese meist mit Gärrest aus anderen Biogasanlagen befüllt. Zwar kommen in der Umwelt nahezu alle für den Biogasprozess relevanten

Mikroorganismen vor [Blagodatskaya and Kuzyakov (2013)], aber diese benötigen Zeit sich an die entsprechenden Umweltbedingungen zu adaptieren. Dabei können nach entsprechender Adaptionszeit auch größere Mengen von Hemmstoffen, wie NH_4 und Fettsäuren, toleriert werden [Melbinger et al. (1971), Chen et al. (2008), Fachagentur Nachwachsende Rohstoffe (2010), Seite 15]. Um diese Anpassungszeit so kurz wie möglich zu halten, empfiehlt es sich jedoch Inocula zu nutzen, welche bereits an die späteren Prozessbedingungen angepasst sind.

2.1.3.5 Substrat

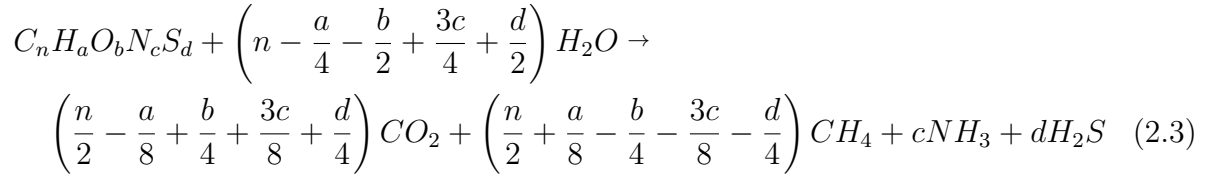
Standardsubstrate für Biogasanlagen sind Energiepflanzen, Gülle, Mist, Klärschlamm, Haushaltsabfälle und Reste aus der Lebensmittelindustrie. Nach ihrer Zerkleinerung werden sie vor der Zugabe mit Gärflüssigkeit gemischt, um die Fluidität zu erhöhen und den pH-Wert zu stabilisieren. Je feiner das Substrat zerkleinert wurde, desto besser können die Mikroorganismen es nutzen. Allerdings ist die Zerkleinerung der Biomasse auch sehr energieintensiv.

Um eine ganzjährige Verfügbarkeit von Energiepflanzen zu gewährleisten, werden diese siliert. Bei der Silierung werden die Pflanzen gehäckselt und unter Sauerstoffabschluss zersetzen. Milchsäurebakterien freilösliche Kohlenhydrate vornehmlich zu Milchsäure. Durch den damit verbundenen Abfall des pH-Wertes der Silage zu pH 3 oder 4 wird die Silage konserviert [Weiland (2010)].

Die wichtigsten Parameter bei der Substratzufuhr einer Biogasanlage sind die Raumbelastung (OLR) und die HRT. Die OLR gibt an wieviel Kilogramm Trockenmasse pro Tag und Kubikmeter Fermentervolumen in die Biogasanlage gegeben werden. Normalerweise beträgt sie zwischen $2\text{-}4 \text{ kg}/(\text{d} * \text{m}^3)$, kann aber für gewisse Biogasanlagen auch bei $10 \text{ kg}/(\text{d} * \text{m}^3)$ liegen [Weiland (2010)]. Die HRT gibt an wie lange die Biomasse und damit auch die einzelnen Mikroorganismen durchschnittlich in der Biogasanlage verbleiben. Folglich werden bei HRT die kleiner als die Wachstumsgeschwindigkeit bestimmter Mikroorganismen sind, diese ausgewaschen. Als Konsequenz können dann einzelne Substrate nicht mehr abgebaut werden oder die Biogasprozesse kommen komplett zum Erliegen. Typische HRT für Rührkesselreaktoren, welche insbesondere die Generationszeit der methanbildenden *Archaeen* berücksichtigen [Anhang: 7.2], liegen zwischen 10-100 Tagen [Ward et al. (2008), Appels et al. (2008), Fachagentur Nachwachsende Rohstoffe (2010), Seite 151] je nach Substrat und Prozesstemperatur.

Bei der Substratzufuhr sollte außerdem ein ausgewogenes Verhältnis von C, Stickstoff (N), Phosphor (P) und Schwefel (S), Mineralstoffen und Spurenelementen eingehalten werden [Tabelle: 2.1]. So empfehlen *Weiland et al. (2010)* [Weiland (2010)] ein Verhältnis von C zu N zu P zu S von 600:15:5:1. Ferner kann aus der Menge an C, H, O, N und S die

Biogasproduktion berechnet werden [Gleichung: (2.3)] [Buswell and Hatfield (1936), Boyle (1976)].



Da nicht jedes Substrat ein ausgewogenes Nährstoffverhältnis aufweist, werden diese Substrate mit anderen Substraten kofermentiert. Beispielsweise kann Klärschlamm gut mit landwirtschaftlichen Abfällen kombiniert oder bei Stickstoffmangel Maissilage mit Gülle bzw. Grassilage versetzt werden [Angelidaki and Ellegaard (2003), Hartmann and Ahring (2005)]. Bei spezifischen Prozessproblemen können auch Zusätze, wie z. B. Spurenelemente, zum Substrat beigemischt werden [Tabelle: 2.2].

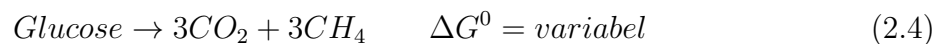
Tabelle 2.2: *Substratzusätze für Biogasanlagen bei bestimmten Prozessproblemen* [Suárez Quiñones et al. (2012), Appels et al. (2008), Schmack and Reuter (2010), Schmack and Reuter (2011)].

| Prozessproblem | Additive |
|-----------------------|--|
| Spurenelementmangel | Zugabe der fehlenden Spurenelemente |
| Schaumbildung | Zugabe von Pflanzenölen |
| Zu hohe Viskosität | Zugabe von Enzympräparaten (ZYmaXX) |
| Zu langsame Hydrolyse | Zugabe von speziellen Mikroorganismen (METHANOS) |

Ferner kann das Substrat Hemmstoffe, wie z.B. NH_3 , Antibiotika, Schwefelverbindungen und Schwermetalle enthalten [Anhang: 7.1]. Die Hemmung beruht meist auf der Inhibition/ Vergiftung der Mikroorganismen, kann aber auch wie im Fall von Sulfat durch die Nutzung als alternative Elektronenakzeptoren verursacht werden. Dabei werden organische Verbindungen verbraucht, welche dann nicht mehr für die Bildung des CH_4 zur Verfügung stehen [Chen et al. (2008), Appels et al. (2008)].

2.2 Mikrobieller Abbau von Biomasse

Der Biogasprozess ist der anaerobe Abbau von Biomasse zu CH_4 und CO_2 durch eine mikrobielle Gemeinschaft bestehend aus *Bakterien* und *Archaeen* [Abbildung: 2.2, Anhang: 7.3]. Aus Mangel an O als geeigneter Elektronenakzeptor unter anaeroben Bedingungen sind die Mikroorganismen nicht in der Lage, die organische Biomasse komplett bis zum CO_2 zu oxidieren. Daher muss ein Teil des Kohlenstoffes zu CH_4 reduziert werden [Gleichung: 2.4]. Insgesamt kann der Biogasprozess in die vier Schritte Hydrolyse [Kapitel: 2.2.1], Gärung [Kapitel: 2.2.2] Acetogenese [Kapitel: 2.2.3] und Methanogenese [Kapitel: 2.2.4] unterteilt werden [Madigan and Martinko (2009), Seite 755].



2.2.1 Hydrolyse

Damit Mikroorganismen unlösliche, komplexe Biopolymere aufnehmen und durch Gärung Energie gewinnen können, müssen sie diese vorher spalten. Dafür produzieren *Bakterien*, insbesondere Vertreter der Stämme *Bacteroidetes* und *Firmicutes*, diverse hydrolytische Enzyme und setzen sie frei. Dabei werden Fette durch Lipasen zu Fettsäuren und Glycerin (*enzyme commission number* (EC): 3.1.-.-.), komplexe Kohlenhydrate durch Amylasen und Cellulasen zu Einfachzuckern (EC: 3.2.-.-.), DNA durch Nukleasen zu Nukleotiden (EC: 3.1.-.-.) und Proteine durch Peptidasen (EC: 3.4.-.-.) zu Aminosäuren gespalten. Eine Besonderheit stellen faserige Moleküle bestehend aus Cellulose, Hemicellulose und Lignin da. Je nach Zusammensetzung können diese gar nicht oder nur langsam zu Hexosen und Pentosen aufgeschlossen werden. Immerhin sind einige Mikroorganismen wie *Clostridium thermocellum* oder *Ruminococcus flavefaciens* in der Lage, Cellulose effektiv zu verwerten. Dafür bilden sie einen extrazellulären Cellulosomkomplex [Shoham et al. (1999)]. Dieser bleibt an den Mikroorganismen haften und kann darüber hinaus auch die Bindung an Pflanzenfasern vermitteln. Neben diversen Struktur-, Binde- und Ankerproteinen, umfasst das Cellulosom diverse Exo- und Endoenzyme. Während Cellulasen (EC: 3.2.1.91) die Cellulose an der β -Glycosidbindung spalten, sind Xylanasen (EC: 3.2.1.8.), Xyloglucanasen (EC: 3.2.1.155) und Mannosen (EC: 3.2.1.24) in der Lage Hemicellulose zu hydrolysieren [Dykstra et al. (2014)].

2.2.2 Gärung

Für die Gewinnung von Energie unter anaeroben Bedingungen sind Mikroorganismen in der Lage, die verschiedensten energiereichen Substrate zu fermentieren. Die dabei freier-

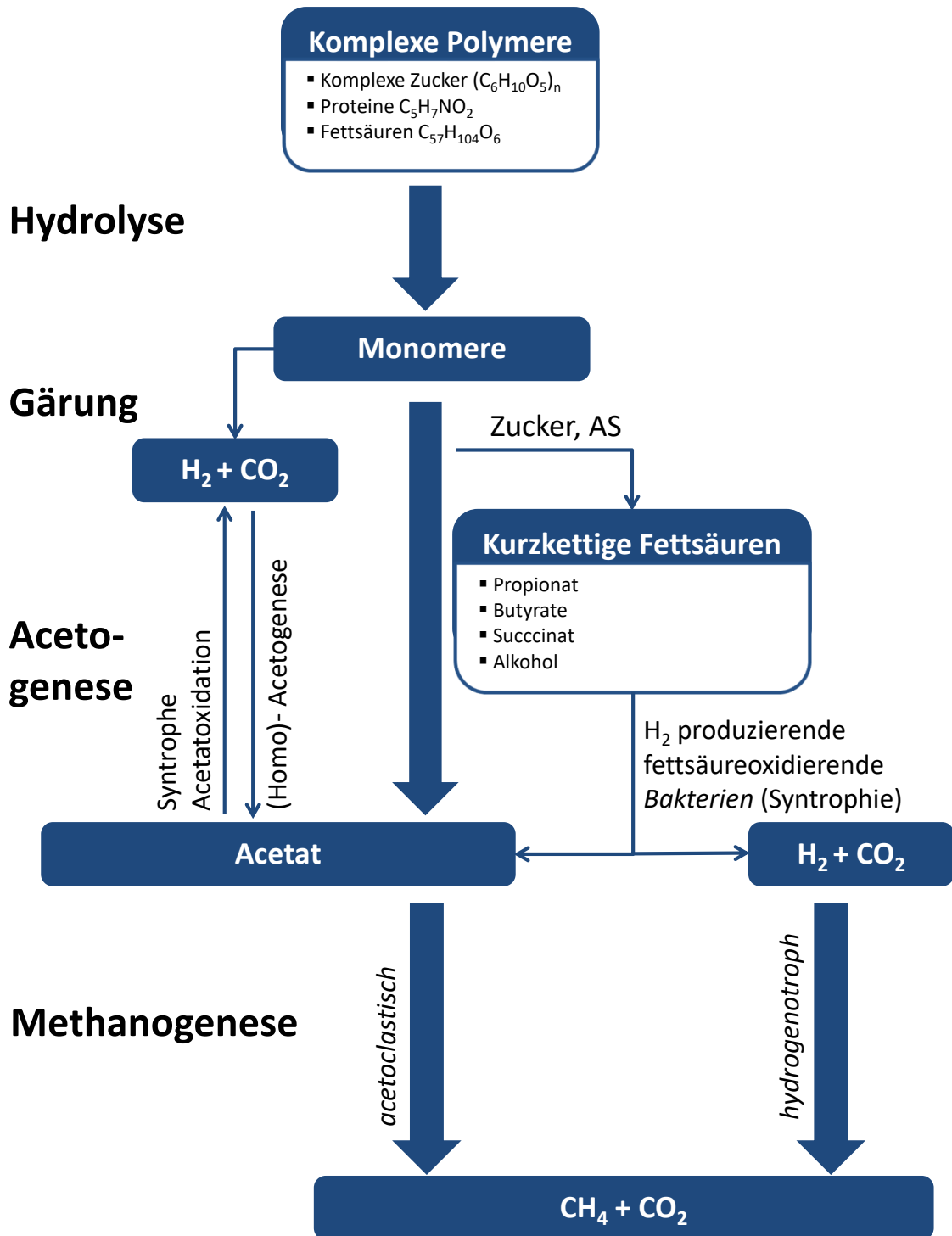


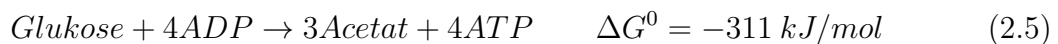
Abbildung 2.2: Vier Stufen des Biogasprozesses

Diese Abbildung zeigt die vier Stufen des Biogasprozesses [Madigan and Martinko (2009), Westerholm (2012)]: Hydrolyse, Gärung, Acetogenese und Methanogenese.

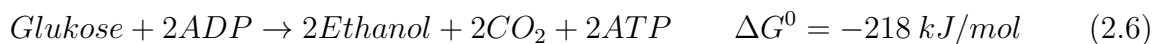
dende Energie wird durch Substratkettenphosphorylierung in ATP gespeichert. Da für diese energiebringenden Reaktionen das Substrat oft erst unter Zuhilfenahme des Redoxäquivalens NAD^+ oxidiert werden muss, erfolgt anschließend in einem zweiten Schritt die Regeneration des NAD^+ .

Hexosen und Pentosen, welche in Biogasanlagen den Hauptbestandteil des Substrates ausmachen, werden dabei zuerst durch die Glykolyse oder den Pentosephosphatweg zu Pyruvat bzw. Acetyl-Coenzym (Co)A abgebaut. Anschließend erfolgt die weitere Vergärung zu Ethanol (EC: 1.1.1.1/2; EC: 1.1.2.7/8) [Gleichung: 2.6], Acetat (-) [Gleichung: 2.5], Laktat (EC: 1.1.1.27.) [Gleichung: 2.7/ 2.8], Butyrat (-) [Gleichung: 2.9] oder Propionat (-) [Gleichung: 2.10]. Typische Vertreter für die einzelnen Gärungswege sind *Zymomonas mobilis* für die Ethanolgärung [Schlegel 2007, Seite 360], *Acetobacter woodii* für die Acetatgärung [Buschhorn et al. (1989)], *Lactobacillus amylovorus* für die Laktatgärung [Shin et al. (2010)], diverse *Butyrivibrio*-Spezies für die Butyratgärung [Weiss et al. (2008)] und *Propionibacterium freudenreichii* für die Propionatgärung [Riedel and Britz (1993)]. Auch kombinieren einige Mikroorganismen die verschiedenen Gärungsformen [Gleichung: 2.11] oder bauen Pentosen über die Glykolyse bzw. Hexosen über den Entner-Doudoroff-Weg und den Pentosephosphatweg ab.

Acetatgärung



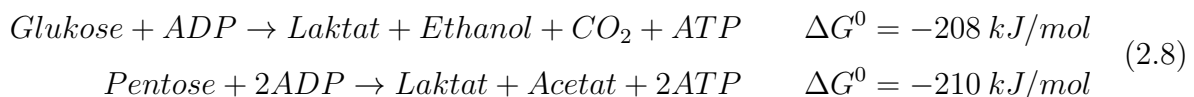
Ethanolgärung



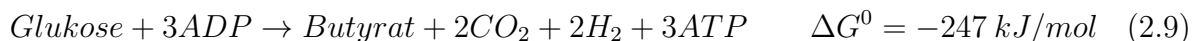
Homofermentative Laktatgärung



Heterofermentative Laktatgärung



Butyratgärung



Propionatgärung



Gemischte Säuregärung



Für den Abbau von Desoxyribonukleinsäure (DNA) und Ribonukleinsäure (RNA) werden diese Makromoleküle zuerst in ihre Nukleotide und dann in Phosphatreste, Pentosen und organische Basen gespalten. Während die Pentosen dann wie oben beschrieben über den Pentosephosphatweg abgebaut werden, ist der Abbau der organischen Basen sehr komplex und erfolgt über verschiedene Wege [Vogels and Van der Drift (1976)].

Die Degradation der Aminosäuren erfolgt unter anaeroben Bedingungen meistens paarweise zu Fettsäuren wie Acetat, Butyrat, Propionat oder Valeriat über die Sticklandreaktion durch Mikroorganismen wie *Clostridium sporogenes* [Nisman (1954), Angelidaki et al. (2011), Ramsay and Pullammanappallil (2001)]. Dabei wird eine Aminosäure oxidiert (EC: 4.1.1.X), die andere reduziert (EC: 1.21.4.2) und zeitgleich NAD^+ regeneriert.

Um Fettsäuren unter anaeroben Bedingungen zu verwerten, werden diese zyklisch durch β -Oxidation um je zwei Kohlenstoffatome verkleinert und Acetyl-CoA gebildet. Im Fall einer ungeraden Anzahl an Kohlenstoffatomen entsteht im letzten Schritt anstelle von Acetyl-CoA die Fettsäure Propionyl-CoA. Ein typischer Vertreter für diesen Abbauweg ist z. B. *Syntrophomonas wolfei* [Lorowitz et al. (1989)]. Um das beim Abbau der Fettsäuren benötigte NAD^+ zu regenerieren, wird parallel dazu H_2 gebildet. Voraussetzung dafür ist allerdings, dass das entstehende H_2 von anderen Mikroorganismen (syntrophe Mikroorganismen) sofort wieder verbraucht wird.

2.2.3 Acetogenese

Als Resultat der Gärung wurden neben Alkoholen und H_2 kurzkettige Fettsäuren gebildet. Voraussetzung für den weiteren Abbau der Fettsäuren und Alkohole sind niedrige Wasserstoffpartialdrücke. Folglich können diese nur weiter degradiert werden, wenn andere Mikroorganismen den entstehenden H_2 sofort wieder verbrauchen. Dabei kommt es wahrscheinlich zum Interspezies-Wasserstofftransfer zwischen fettsäureabbauenden und wasserstoffverwertenden Mikroorganismen [Stams and Plugge (2009)]. Allgemein wird dieses gemeinsame Verstoffwecheln von Substraten durch verschiedene Mikroorganismen

als Syntrophie bezeichnet. Eine für Biogasanlagen wichtige Form der Wasserstoffverwertung ist die Acetogenese. Mikroorganismen wie *Clostridium acetivum* [Poehlein et al. (2015)] nutzen H_2 zusammen mit CO_2 , um über den reduktiven Acetyl-CoA-Weg (Wood-Ljungdahl-Weg) Acetat zu produzieren [Gleichung: 2.12]. Zusätzlich können die Mikroorganismen bei diesem Prozess Energie in Form eines Natriumgradienten speichern. Auf der anderen Seite gibt es auch Mikroorganismen wie *Syntrophaceticus schinkii*, die diesen Weg rückwärts nutzen, um Acetat zu H_2 und CO_2 zu spalten. Dieser Prozess heißt dann syntrophe Acetatoxidation [Gleichung: 2.13] [Müller et al. (2013), Westerholm (2012)]. Welcher Weg während der Biogasprozesse tatsächlich genutzt werden, hängt im Wesentlichen von der freien Reaktionsenthalpie (Gibbs-Energie) dieser Reaktion ab, welche neben der Temperatur vom Wasserstoffpartialdruck und der Acetatkonzentration abhängt [Gleichung: 2.14]. Karakashev et al. (2006) [Karakashev et al. (2006)] haben z. B. beobachtet, dass hauptsächlich die syntrophe Acetat-Oxidation abläuft, solange keine Mikroorganismen der Familie *Methanosaetaceae* anwesend sind, die die Acetatkonzentration sehr niedrig halten. Nach Müller et al. (2013) [Müller et al. (2013)] kann ein wesentlicher Faktor auch der Anteil an Protein im Substrat sein. Beim Abbau der Proteine wird viel Ammonium freigesetzt, welches die acetoclastische Methanogenese hemmt und somit verhältnismäßig hohe Acetatkonzentrationen vorliegen.

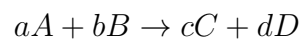
Homoacetogenese



Acetatoxidation



Freie Reaktionsenthalpie (Gibbs-Energie)



$$\Delta G = \Delta G^0 - RT * \ln \left(\frac{[C]^c * [D]^d}{[A]^a * [B]^b} \right) \quad (2.14)$$

- ΔG : Änderung der freien Reaktionsenthalpie
 ΔG^0 : Änderung der freien Reaktionsenthalpie unter Standardbedingungen
R: Universelle Gaskonstante
T: Absolute Temperatur

[A], [B]: Konzentration der Substrate
 [C], [D]: Konzentration der Produkte

2.2.4 Methanogenese

Die Bildung von CH₄ während des Biogasprozesses [Thauer et al. (1977), Thauer et al. (2008), Ferry (2011)] erfolgt durch methanogene *Archaeen*, welche hauptsächlich CO₂ und H₂ (hydrogenotroph) [Gleichung: 2.15] sowie Acetat (acetoclastisch) [Gleichung: 2.16] nutzen. Die bei den einzelnen Schritten dieses Weges freiwerdende Energie reicht nicht für die Bildung eines Moleküls ATP aus. Sie wird daher für den Aufbau eines Protonen- bzw. Natriumgradienten genutzt. Dieser wird in einem zweiten Schritt für die Bildung von ATP verwendet.

Bei der Bildung von ATP während der Methanogenese gibt es einen wesentlichen Unterschied zwischen der Ordnung *Methanosarcinales*, welche Cytochrome besitzen und den Methanogenen ohne Cytochrome. Die zusätzlichen Cytochrome erlauben statt einem Energiegewinn von bis zu 0,5 Mol ATP pro Mol gebildetem Methan den Gewinn von bis zu 1,5 Mol ATP für den hydrogenotrophen Weg und die Nutzung weiterer Substrate wie Acetat, Methanol, sowie Di- und Trimethylamide [Koch et al. (2016), Thauer et al. (2008)]. Die Nachteile sind aber, dass diese *Archaeen* höhere Wasserstoffpartialdrücke (>10 Pa) und Temperaturen unter 60 °C brauchen sowie Verdopplungszeiten von 10 h anstatt von 1 h aufweisen [Thauer et al. (2008)].

Die wichtigsten Vertreter für methanogene *Archaeen* mit Cytochromen sind die Familien *Methanosaetaceae* und *Methanosarcinaceae*. Aber auch zwischen diesen Familien gibt es einen Unterschied. Während sich Vertreter der Familie *Methanosarcinaceae* auf die Nutzung verschiedenster Substrate spezialisiert haben, nutzen Vertreter der Familie *Methanosaetaceae* nur Acetat. Dafür besitzen sie aber statt eines schwach Acetat-affinen Acetatkinase-phosphotransacetylase Systems, eine hoch-affine Adenosinmonophosphat-bildene Acetyl-CoA Synthetase (EC: 6.2.1.1), welche ein Wachstum [Liu and Whitman (2008)] bei Acetatkonzentration von 5-20 μMol anstatt von 1 mMol erlaubt.

Hydrogenotrophe Methanogenese



Acetoclastische Methanogenese



2.3 Mikrobielle Gemeinschaften

In nahezu allen Lebensräumen kommt nicht nur eine, sondern viele verschiedene mikrobielle Spezies vor. Diese mikrobiellen Gemeinschaften können neben *Bakterien*, *Archaeen*, auch Bakteriophagen und eukaryotische Einzeller umfassen. Beispielsweise haben *Schlüter et al. (2008)* gezeigt, dass in Biogasanlagen bis zu 1000 verschiedene *Bakterien* und *Archaeen* vorkommen. Darüber hinaus treten zwischen den einzelnen Arten sehr komplexe Interaktionen wie Konkurrenz um Nährstoffe, Räuber-Beute-Beziehungen und auch Formen von Zusammenarbeit wie Syntrophie auf.

2.3.1 Methoden zur Charakterisierung von mikrobiellen Gemeinschaften

Für die Optimierung von Biogasprozessen ist es notwendig, die mikrobielle Gemeinschaft in Biogasanlagen umfassend zu verstehen [Weiland (2010)]. Die Schlüsselfragen sind dabei: Welche Mikroorganismen sind in welcher Häufigkeit vorhanden und welche metabolischen Funktionen haben diese [Hofman-Bang et al. (2003)]? Da die Kultivierung der meisten Mikroorganismen als Reinkultur schwierig oder gar unmöglich ist [Amann et al. (1995)], müssen die angewendeten Methoden geeignet sein, die Zusammensetzung und Funktion der einzelnen Mikroorganismen auch in komplexen Proben zu charakterisieren.

Geeignete Methoden, um die Häufigkeit von Mikroorganismen zu untersuchen sind mikroskopische Methoden oder molekularbiologische Methoden. Für die Charakterisierung des Metabolismus hingegen eignen sich Metaproteom- und Metabolomuntersuchungen, sowie Enzymaktivitätsassays.

Die Mikroskopie erlaubt die Charakterisierung von einzelnen Mikroorganismen nach ihrer Morphologie, z. B. nach Stäbchen, Kokken oder Spirillen. Die Identifizierung einzelner Spezies [Grotenhuis et al. (1991)] ist allerdings sehr schwierig und setzt umfangreiches Expertenwissen voraus. Weiterhin kann durch Mikroskopie geklärt werden, ob die Mikroorganismen vereinzelt, als Agglomerate oder gebunden an das Substrat vorliegen. Eine weitere Anwendung für mikroskopische Untersuchungen, ist die Frage ob z. B. anaerobe Pilze [Trinci et al. (1994), Kazda et al. (2014)] oder eukaryotische Einzeller eine Rolle während des Biogasprozesses spielen. Für eine genauere taxonomische Charakterisierung können die Mikroorganismen, z. B. mittels Fluoreszenz-in-situ-Hybridisierung [Sekiguchi et al. (1999), Nettmann et al. (2010), Nettmann et al. (2013)] angefärbt werden. Die mikroskopische Quantifizierung von hydrogenotrophen *Archaeen* kann alternativ über die Eigenfluoreszenz von deren F420-Komplexen erfolgen [Nader et al. (1995)]. Allerdings kann die Erfassung von Fluoreszenzsignalen durch Huminstoffe in der Probe gestört werden [Hofman-Bang et al. (2003), Bastida et al. (2009), Schlegel (2013)].

Da die Quantifizierung der Mikroorganismen mittels Mikroskopie sehr zeitaufwendig ist,

setzte sich in den letzten Jahren die automatische Zählung von Zellen mittels Durchflusszytometrie kombiniert mit entsprechenden Färbungen der Mikroorganismen immer mehr durch [Nader et al. (1995), Müller et al. (2012), Rüger et al. (2014)]. Bei der Durchflusszytometrie ist jedoch zu beachten, dass die Mikroorganismen vorher vereinzelt werden müssen.

Molekularbiologische Methoden basieren im Wesentlichen auf der Charakterisierung und Quantifizierung einzelner Gene bzw. *messenger*-RNAs (mRNA). Für die Untersuchung einzelner Gene wird nach der Extraktion der DNA [Hofman-Bang et al. (2003)]/ RNA [Ibrahim and Ahring (1999)] ein spezifisches Primerpaar an das Gen gebunden und dieses mittels Polymerasekettenreaktion (PCR) amplifiziert. Dabei erfolgt die Amplifikation der mRNA nicht direkt, sondern sie wird vorher in DNA übersetzt. Da in mikrobiellen Gemeinschaften die einzelnen Mikroorganismen verschiedene Versionen eines Genes besitzen, erfolgt anschließend die Separation der Genvarianten mittels *denaturing gradient gel electrophoresis* (DGGE) oder *terminal restriction fragment length polymorphism* (TRFLP). Anschließend können die Genvarianten mit einer Klonbibliothek und Sequenzierung identifiziert werden. Die Quantifizierung der Genvarianten erfolgt dann entweder aus der Intensität des DGGE- und TRFLP-Profiles oder mittels quantitativer PCR [Takai and Horikoshi (2000), Blume et al. (2010)]. Letztere setzt allerdings spezifische Primer voraus. Ein wichtiges Untersuchungsobjekt in der Molekularbiologie ist das 16S RNA-Gen, welches für die taxonomische Einordnung der Mikroorganismen genutzt wird [Amann et al. (1995)]. Dieses Gen hat den Vorteil, dass es in allen Spezies wenigstens einmal vorkommt (16S-RNA Häufigkeit siehe [Klappenbach et al. (2001)]) und sowohl hochkonservierte Bereiche für die Primer als auch variable Bereiche für die taxonomische Charakterisierung aufweist. Zusätzlich kann auch die 16S-RNA selbst untersucht werden, welche mit 10^3 - 10^5 Kopien sehr abundant ist [Hofman-Bang et al. (2003)]. Neben der taxonomischen Charakterisierung von mikrobiellen Gemeinschaften in Biogasanlagen [Klocke et al. (2008), Nelson et al. (2011), Ziganshin et al. (2013)], können auch einzelne funktionelle Gene dieser analysiert werden, wie die Methyl CoM Reduktase [Munk et al. (2012)].

Dank der Entwicklung neuer Sequenzierungstechniken wie der 454-Pyrosequenzierung [Margulies et al. (2005)] und Illuminat Sequenzierung [Bentley et al. (2008)] ist es mittlerweile auch möglich, zeitgleich alle mikrobiellen Gene/ mRNAs (Metagenom/ Metatranskriptom) in Proben aus Biogasanlagen zu charakterisieren [Schlüter et al. (2008), Wirth et al. (2012), Zakrzewski et al. (2012)]. Ermöglicht wird dadurch sowohl die Untersuchung der taxonomischen Zusammensetzung als auch der Funktion der mikrobiellen Gemeinschaften. Der wesentliche Nachteil dieser molekularbiologischen Methoden ist, dass aus dem Vorhandensein von Genen nicht auf ihre tatsächliche Aktivität geschlossen werden kann. Des Weiteren ist zu beachten, dass die mRNA nicht sehr stabil ist [Belasco and

Brawerman (1993)] und einer starken zeitlichen Dynamik unterliegt [Blagodatskaya and Kuzyakov (2013)].

Eine bessere Korrelation mit dem tatsächlichen Metabolismus (Metaproteom) [Kapitel: 2.4] erlaubt die Untersuchung der mikrobiellen Enzyme, welche die einzelnen Stoffwechselreaktionen katalysieren. Für die Identifizierung der Proteine ist es aber vorteilhaft, wenn die Komplexität der mikrobiellen Gemeinschaften bekannt ist oder bereits alle Gene der mikrobiellen Gemeinschaft sequenziert wurden.

Die Häufigkeit des einzelnen Proteins lässt Rückschlüsse auf die metabolische Aktivität zu, trotzdem muss beachtet werden, dass die reale Enzymaktivität noch von der Temperatur, dem pH- Wert, sowie der Substrat- und Produktkonzentrationen abhängt. Daher ist für die Untersuchung der metabolischen Flüsse auch die Untersuchung des Metaboloms nötig. Bei der Untersuchung des Biogasprozesses wird neben der Untersuchung des Substrates (Futtermittelanalytik) auch die Menge an kurzkettigen Fettsäuren und die Gaszusammensetzung mittels GC bestimmt. Durch die Untersuchung des Isotopenverhältnisses von Kohlenstoff und Wasserstoff der Komponenten im Gas kann zusätzlich indirekt auf die Nutzung bestimmter Stoffwechselreaktionen geschlossen werden [Nikolausz et al. (2013)]. Die für Reinkulturen oft angewandte Messung der Enzymaktivität kann zwar auch für Proben aus dem Biogasprozess genutzt werden, ist aufgrund der Probenverunreinigungen in den Biogasanlagenproben aber schwierig und wird selten angewandt. Beispiele sind die Bestimmung der Aktivität von hydrolytischen Enzymen [Gasch et al. (2013)] und von dem Methanogeneseenzym Heterodisulfatereduktase [Refai et al. (2014)].

2.3.2 Mikrobielles Ressourcen Management

Mikrobielles Ressourcen Management (MRM) ist eine Sammlung von Methoden zur Beschreibung von mikrobiellen Gemeinschaften mit dem Ziel, diese gezielt zu beeinflussen [Verstraete et al. (2007)]. Dabei ist es nach *Verstraete et al. (2007)* wichtig, drei Schlüsselfragen zu beantworten: „Wer ist da?“, „Wer macht was?“ und „Wer macht was mit wem?“. Darüber hinaus ist es aber auch wichtig genau zu definieren, welche Anforderungen an die mikrobiellen Gemeinschaften gestellt werden. So kann es für den Biogasprozess durchaus wünschenswert sein, neben einer hohen und stabilen Methanproduktion geringe Wachstumsraten der Mikroorganismen zu haben. Dies würde die Menge an zu entsorgendem Gärrest minimieren.

Für die praktische Umsetzung des MRM ist es notwendig, die mikrobiellen Gemeinschaften mit Hilfen von geeigneten Größen zu beschreiben. *Verstraete et al. (2007)* und seine Mitarbeiter [Wittebolle et al. (2009), Marzorati et al. (2008), Mertens et al. (2005), deRoy (2014)] schlagen die Untersuchung der Speziesanzahl (*Richness*), deren Gleichverteilung

(*Evenness/Community Organisation*) sowie die Dynamik in der Zusammensetzung der mikrobiellen Gemeinschaften vor. Während die Speziesanzahl, die Anzahl an verschiedenen Spezies oder anderer Taxa in einer Probe ist, lässt sich die Dynamik der Spezies durch Pearsonkorrelation zwischen den taxonomischen Profilen beschreiben und als Änderung der Ähnlichkeit ausdrücken [Gleichung: 2.17].

$$\%_{\text{Änderung}} = 100 - r(t_1, t_2) * 100 \quad (2.17)$$

$\%_{\text{Änderung}}$: Prozentuale Dynamik der mikrobiellen Gemeinschaft
 $r(t_1, t_2)$: Pearsonkorrelation zwischen Zeitpunkt 1 und Zeitpunkt 2

Für die Beschreibung der mikrobiellen Gleichverteilung (*Evenness/Community Organisation*) wird die kumulative Häufigkeit der Mikroorganismen in Abhängigkeit deren Speziesanzahl, sortiert nach absteigender Häufigkeit (Lorenzkurve), aufgetragen und daraus der Gini-Index berechnet [Abbildung: 2.3]. Dabei ist der Gini-Index das Verhältnis der Fläche zwischen der Lorenzkurve und der Linie der perfekten Gleichverteilung normalisiert auf die Fläche unter der perfekten Gleichverteilung.

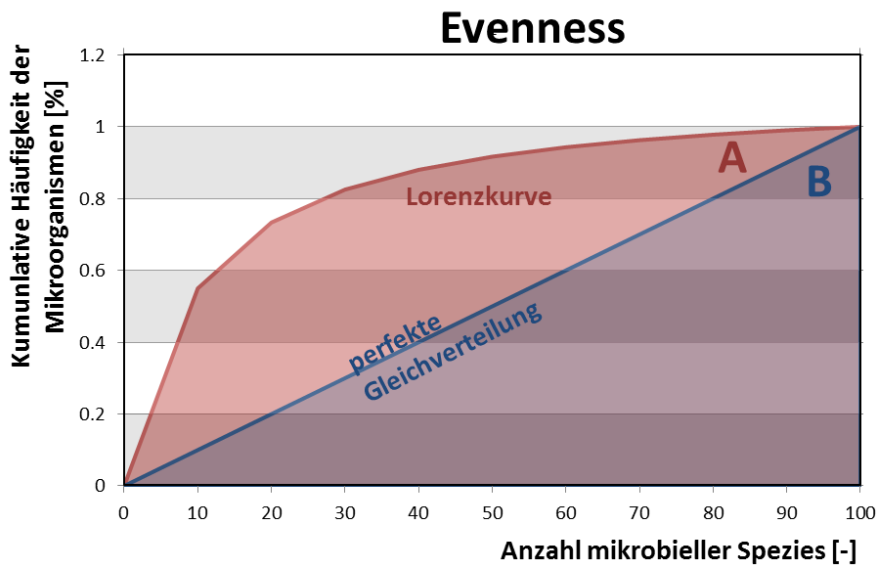


Abbildung 2.3: *Mikrobielle Gleichverteilung (Evenness/Community Organisation)*
 Diese Abbildung zeigt die mikrobielle Gleichverteilung (*Evenness/Community Organisation*). Dabei wird die mikrobielle Gleichverteilung durch den Gini-Index angegeben. Der Gini-Index ist das Verhältnis der Fläche (A) zwischen der Lorenzkurve und der Linie der perfekten Gleichverteilung normalisiert auf die Fläche (B) unter der perfekten Gleichverteilung.

Bei der Untersuchung von sehr komplexen mikrobiellen Gemeinschaften mit den Methoden des MRM stellt sich das Problem, dass wenig abundante Spezies durch Analysen mit geringer Auflösung nicht erfasst werden können [Verstraete et al. (2007)]. Als Konsequenz gelten die Ergebnisse dieser Untersuchungen dann nur unter der Annahme, dass die dominierenden Spezies auch die Schlüsselrollen in der mikrobiellen Gemeinschaft einnehmen. Ein anderer Ansatz für die Charakterisierung von mikrobiellen Gemeinschaft kommt von *Shade et al. (2012)* [Shade and Handelsman (2012)], die versuchen die Schnittmenge an mikrobiellen Spezies (*Core Community*) zu bestimmen, die für einen bestimmten Prozess notwendig ist. Dabei bleibt aber fraglich, ob nicht einzelne Funktionen von den verschiedensten Spezies eingenommen werden können. Erschwerend kommt hinzu, dass nicht genau klar ist, ob sich mikrobielle Gemeinschaften nur nach deterministischen Prinzipien zusammen finden oder inwieweit der Zufall eine Rolle spielt [Vanwonterghem et al. (2014)].

2.4 Methoden der Proteinanalytik

Viele lebensnotwendige Prozesse werden in Zellen von Proteinen/ Enzymen ausgeführt. Die Gesamtheit aller exprimierten Proteine einer Zelle/ Spezies wird als Proteom und deren Untersuchung als Proteomanalyse oder Proteomik bezeichnet [Wasinger et al. (1995), Wilkins et al. (2006)].

Bei komplexeren Proben, welche die Proteine von mehreren verschiedenen Spezies enthalten, wird der Begriff Proteom zu Metaproteom erweitert [Rodriguez-Valera (2004), Wilmes and Bond (2006)]. Bei der Untersuchung des Metaproteoms gibt es drei große Herausforderungen [Muth et al. (2013a)]. Die Erste ist die enorme Probenkomplexität [Gotelli et al. (2012), Muth et al. (2013a)]. Nach *Schlüter et al. (2008)* sind am Biogasprozess schätzungsweise 1000 verschiedene mikrobielle Spezies beteiligt. Unter der Annahme, dass jede Spezies 1000 Proteine exprimiert, enthält jede Probe ungefähr eine Million verschiedene Proteine. Die zweite Herausforderung sind redundante Proteinidentifikationen. Identifizierte Peptide gehören nicht nur zu einem Protein, sondern können mehreren, ähnlichen Proteinen von verschiedenen Spezies zugeordnet werden. Die Konsequenz ist eine Vielzahl von Proteinidentifikationen basierend auf nur einem Peptid. Das dritte Problem ist, dass Proteine nur identifiziert werden können, wenn deren Sequenzen in entsprechenden Datenbanken hinterlegt sind. Laut *Amann et al. (1995)* sind aber nur 0,001-15% der Bakterien kultivierbar, was dazu führt, dass nur von diesen das Genome komplett sequenziert werden konnte und somit auch nur von diesen Proteine identifiziert werden können.

Für die Untersuchung des Metaproteoms müssen die Proteine [Kapitel: 2.4.1] zuerst aus ihrer Matrix extrahiert [Kapitel: 2.4.2] und deren Konzentration bestimmt werden [Kapitel:

2.4.3]. Dann werden die Proteine separiert um die Probenkomplexität zu reduzieren [Kapitel: 2.4.4], bevor sie zu kleineren Peptiden verdaut [Kapitel: 2.4.5] und mittels Tandem-Massenspektrometrie (MS/MS) [Kapitel: 2.4.6] analysiert werden. Abschließend erfolgt die Identifizierung der Proteine [Kapitel: 2.4.7] und die taxonomische und funktionelle Klassifikation [Kapitel: 2.4.8]. Im Folgenden wird nach einer kurzen Einführung zu Proteinen auf die einzelnen Schritte des Metaproteomworkflows [Abbildung: 2.4] eingegangen und Strategien für die Untersuchung von mikrobiellen Gemeinschaften in Biogasanlagen vorgestellt.

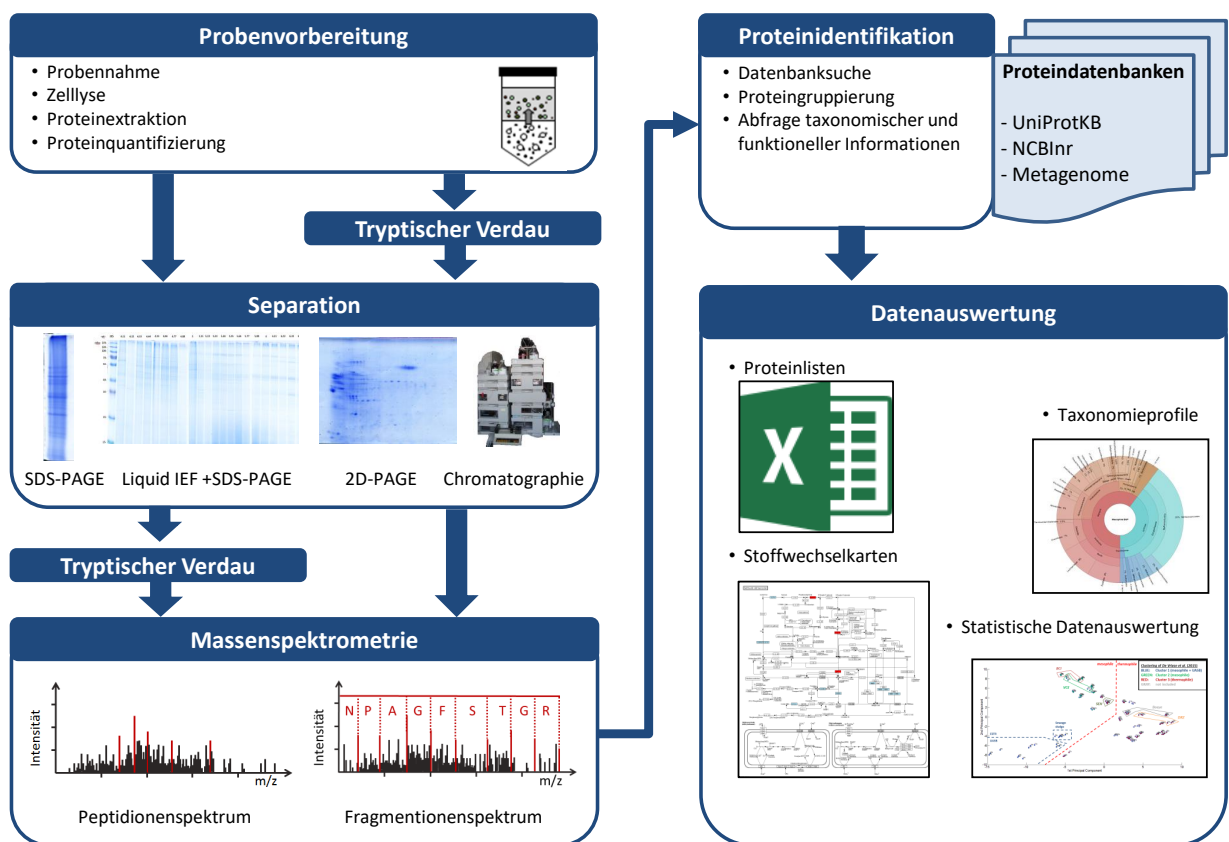


Abbildung 2.4: Allgemeiner Workflow für Metaproteomanalysen.

2.4.1 Proteine

Proteine sind Biomoleküle, welche von Zellen aus Aminosäuren synthetisiert werden. Sie katalysieren biochemische Reaktionen, geben den Zellen Struktur, transportieren Stoffe und leiten Signale weiter. Der Aufbau der Proteine ist in der DNA verschlüsselt. Vor der Synthese der Proteine wird der für das Protein relevante DNA-Abschnitt (Gen) in eine

mRNA übersetzt. Dieser Vorgang wird als Transkription bezeichnet. Anschließend wird die mRNA zu den Ribosomen transportiert und dient dort bei der Translation als Vorlage für die Synthese der Proteine aus den einzelnen Aminosäuren. Anschließend können die Aktivität, Stabilität und Funktion von Proteinen durch posttranslationale Modifikationen (PTMs), wie z. B. Glykosylierung, Phosphorylierung oder Acetylierung verändert werden [Dunn (2000), Seite 2].

Im Unterschied zu Genen ist die Häufigkeit von mRNAs und Proteinen in der Zelle veränderlich. In einer Bakterienzelle variiert die Häufigkeit von einzelnen mRNA-Kopien eines Gens zwischen 0,04-1,4 Kopien pro Zelle [Maier et al. (2011)] und die einzelner Proteine zwischen 10-100.000 Kopien pro Zelle [Wöhlbrand et al. (2013)]. Die unterschiedliche Häufigkeit der Proteine und der mRNA wird zum Einen durch die zeitlich verschobene Dynamik nach einem Stimulus verursacht [Blagodatskaya and Kuzyakov (2013)]. Zum anderen gibt es verschiedene Halbwertszeiten, welche für bakterielle mRNA durchschnittlich 3 min [Belasco and Brawerman (1993)] und für bakterielle Proteine durchschnittlich 23 h [Maier et al. (2011)] betragen. Als Konsequenz gibt es nur eine mäßige Korrelation zwischen den Kopienanzahlen von mRNA und Proteinen. So haben z. B. Maier et al. (2011) gezeigt, dass die Korrelation zwischen mRNA und Proteinen für *Mycoplasma pneumoniae* nur 0.52 beträgt. Das Vorhandensein von Proteinen ist daher besser geeignet als das Vorhandensein von mRNAs, die ablaufenden biochemischen Prozesse in einer Zelle zu analysieren. Allerdings ist auch die Präsenz eines Proteins kein Beweis für seine Aktivität. Aus Effizienzgründen kann aber davon ausgegangen werden, dass Zellen Proteine nur in der Menge bilden, welche sie auch benötigen. Veränderungen in der Proteinhäufigkeit sind dabei auf Änderungen von Umwelteinflüsse, wie neue Substrate [Schmidt et al. (2016)], Hitzeschock [Lüders et al. (2009)] und Schadstoffe [Santos et al. (2004)] zurückführbar. Wichtig bei der Untersuchung von Veränderungen in der Proteinexpression ist ihre zeitliche Dynamik. Einerseits kann schon 10 min nach einem Umweltstimulus die Bildung neuer Proteine nachgewiesen werden [Kramer et al. (2010)]. Andererseits können auch 1-2 Tage vergehen, bis die Aktivität von großen, neugebildeten Enzymkomplexen wie dem Cellulosom nachweisbar wird [Blagodatskaya and Kuzyakov (2013)]. Umgekehrt kann die Aktivität dieses Komplexes bis zu 4 Monate nach Substratzugabe erhalten bleiben [Blagodatskaya and Kuzyakov (2013)].

Die Stabilität, Aktivität und Funktion von Proteinen wird weiterhin von PTMs beeinflusst. Da PTMs von Bakterien allerdings seltener sind als bei Eukaryoten und nicht ohne vorherige Anreicherung mittels MS identifiziert werden können [Cain et al. (2014)], werden sie bei umfassenden Metaproteomanalysen derzeit noch vernachlässigt.

Proteine mit ähnlicher Funktion weisen meist eine ähnliche Aminosäuresequenz auf. Dies wird als Homologie bezeichnet. Die Verwandtschaft einzelner Proteine zeigt sich insbeson-

dere an Bindungsstellen und den aktiven Zentren von Enzymen, welche eine besonders konservierte Aminosäurestruktur aufweisen. Ein Vorteil dieser Eigenschaft ist, dass die taxonomische Verwandtschaft basierend auf einzelnen Proteinen wie RecA, Hsp60, Elongationsfaktor Tu, ATPase oder RNA-Polymerase bestimmt werden kann. [Ludwig and Klenk (2001)]. Auf der anderen Seite erschwert die Homologie, die genaue taxonomische Zuordnung einzelner Peptide.

2.4.2 Proteinextraktion

Die Analyse von Proteinen beginnt mit der Lyse der mikrobiellen Zellen und der Extraktion der Proteine aus der Probenmatrix. In Biogasanlagen besteht diese Probenmatrix neben extrazellulären, mikrobiellen Enzymen aus den Bestandteilen der Gülle und nicht abgebauten pflanzlichen Materialien. Die Zellen können dafür mechanisch, enzymatisch oder chemisch aufgeschlossen werden. Zu beachten ist dabei, dass es zu keiner Verfälschung des Proteoms, z. B. durch Proteinabbau oder -neusynthese durch Proteasen bzw. Proteinbiosynthese kommt. Der mechanische Zellaufschluss in einer Kugelmühle hat den Vorteil, dass alle Mikroorganismen, insbesondere auch grampositive *Bakterien* und *Archaeen* mit stabilen Zellwänden zuverlässig aufgeschlossen werden [Hofman-Bang et al. (2003)]. Dagegen funktionieren chemische und enzymatische Zellaufschlüsse meist nur für bestimmte Zelltypen. Zusätzlich kann die Probenmatrix die Effektivität von Enzymen und Chemikalien reduzieren. Um den Einfluss der Probenmatrix zu verringern, können die Mikroorganismen vor der Zellyse von der Probenmatrix separiert werden (indirekte Proteinextraktion) [Becher et al. (2013)]. Gängige Verfahren sind diverse Wasch-, Filtrations- oder Zentrifugationschritte [Taylor and Williams (2010)], welche aber die Gefahr bergen, bestimmte Mikroorganismen an- oder abzureichern [Becher et al. (2013)]. Bei Proben aus Biogasanlagen sind diese Methoden insbesondere problematisch, da die mikrobiellen Gemeinschaften zum Teil als einzelne Zellen, als Flocken oder als Biofilm vorliegen [Hofman-Bang et al. (2003)]. Alternativ können die Zellen zusammen mit der Probenmatrix lysiert werden. Die Proteine müssen dann aber aufwendig von Verunreinigungen wie z. B. Huminstoffe separiert werden. Huminstoffe sind Abbauprodukte von Biomasse, welche ähnliche biochemische Eigenschaften wie Proteine haben [Piccolo (2002), Senesi et al. (1989)]. Sie stören die Quantifizierung, Separation und Identifizierung der Proteine [Bastida et al. (2009)]. Daher ist es notwendig, die Proteine nach dem Aufschluss zu reinigen.

Vor Beginn dieser Promotion gab es bereits erste erfolgreiche Versuche, Proteine aus Biogasanlagen zu untersuchen. *Abram et al. (2009)* haben mittels Ultraschall und einem Zentrifugationsschritt Proteine aus einem Laborfermenter extrahiert und analysiert. Dieser

wurde jedoch nur mit synthetischem Medium gefüttert. *Hanreich et al. (2012)* charakterisierten dagegen einen Laborfermenter, der mit realem Substrat betrieben wurde. Dabei wurden die Zellen durch Ultraschall aufgeschlossen, die Proteine mit Phenol extrahiert und anschließend durch Fällung mit Ammoniumacetat in Methanol gereinigt und konzentriert. Der Vorteil der Phenolextraktion ist, dass DNA und Probenverunreinigungen wie Huminstoffe effektiv abgereichert werden. Zugleich werden durch die denaturierende Wirkung von Phenol Proteasen inaktiviert. Ursprünglich wurde dieses Protokoll für Proben aus Grundwasser und Boden [Benndorf et al. (2007)] entwickelt und auch für Klärschlammproben [Kuhn et al. (2011)] genutzt. In den Diplomarbeiten von *Heyer (2011)* und *Kohrs (2011)* wurden kleine Verbesserungen des ursprünglichen Protokolls von *Hanreich et al. (2012)* vorgeschlagen. Die Zellyse erfolgt nun durch Phenolextraktion in einer Kugelmühle. Außerdem werden die Proteinpellets nach der Extraktion in Harnstoffpuffer gelöst.

Entscheidend für die Qualität der Proteinextraktion sind die Ausbeute an extrahiertem Protein und die Reproduzierbarkeit der schlussendlich identifizierten Proteine. *Leary et al. (2014)* zeigten, dass die Ausbeuten von verschiedenen Extraktionsmethoden für Metaproteomuntersuchungen nur 1-5% der ursprünglichen Proteinmenge und die Reproduzierbarkeiten der Proteinidentifikationen weniger als 10% betragen. Im Gegensatz dazu bestimmten *Keiblinger et al. (2012)* die Ausbeute von Phenolextraktionen mit einem Protokoll ähnlich dem von *Benndorf et al (2007)* mit etwa 50%. Problematischer als schlechte Ausbeuten ist allerdings die an oder Abreicherung bestimmter Proteingruppen bei der Proteinextraktion. Es ist z. B. bekannt, dass es bei der Phenolextraktion zur Abreicherung von Glykoproteinen kommt und dass bei den anschließenden Fällungsschritten hydrophobe Proteine [Eckert and Kartenbeck (1997), Seite 50] verloren gehen, weil diese nicht mit ausfallen. Folglich können nur Ergebnisse von Metaproteomuntersuchungen mit gleicher Extraktionsmethode quantitativ verglichen werden.

2.4.3 Bestimmung der Proteinkonzentration

Für die weiterführende Untersuchung von Proben ist es notwendig, mit vergleichbaren Proteinmengen zu arbeiten. Dafür wird die Proteinkonzentration standardmäßig kolorimetrisch mittels Lowry- [Lowry et al. (1951)], Bradford- [Bradford (1976)] oder Bicinchoninsäure- [Smith et al. (1985)] Proteintests bestimmt. Während der Lowry- und der Bicinchoninsäure-Test auf der Färbung der Seitenketten von Tyrosin und Tryptophan basieren, färbt der Bradford-Test kationische und unpolare Aminosäureseitenketten an. Für reine Proteinproben ist es außerdem möglich, spektrophotometrisch die Proteinkonzentration durch die Messung der Absorption der Peptidbindung bei 205 nm oder der aromatischen

Seitenketten bei 280 nm [Goldfarb et al. (1951)] zu messen [Desjardins et al. (2009)]. Allerdings werden alle diese Tests durch Detergenzien und Probenverunreinigungen mehr oder weniger stark gestört. Besonders für Umweltproben, welche große Mengen an Huminstoffen enthalten, kann es zu signifikanten Störungen dieser Tests kommen [Criquet et al. (2002), Kuhn et al. (2011)]. In den Diplomarbeiten von *Heyer (2011)* und *F. Kohrs (2011)* konnte dagegen gezeigt werden, dass mittels Amidoschwarz-Tests [Popov et al. (1975), Schweikl et al. (1989)] die Proteinkonzentration auch für stark verunreinigte Umweltproben erfolgreich bestimmt werden kann. Der Amidoschwarz-Test basiert auf der Komplexbildung zwischen den basischen Aminogruppen der Proteine und den zwei Sulfonsäuregruppen des Azofarbstoffs Amidoschwarz, wobei dieser Komplex bei niedrigen pH-Werten ausfällt [Kohrs (2010), Schlegel (2013)].

2.4.4 Protein- und Peptidseparation

Aufgrund der großen Komplexität von Proben aus Biogasanlagen ist es sinnvoll, die Proteine vor der massenspektrometrischen Analyse zu trennen. Andernfalls würden sich die einzelnen Peptidionen im MS überlagern und ihre Identifizierung würde erschwert. Da die Trennleistung eines Separationsschrittes oft nicht ausreicht, werden in der Praxis meist mehrere Methoden kombiniert. Die erste Möglichkeit, um die Probenkomplexität zu reduzieren, ist die Präfraktionierung bestimmter Zellen/ Mikroorganismen bzw. bestimmter Zellorganellen/ Kompartimente [Gatto et al. (2010)] z. B. mittels Durchflusszytometrie [Jehlich et al. (2010)] oder Dichtegradientenzentrifugation [Taylor and Williams (2010), Haange et al. (2012)].

Für die Trennung der extrahierten Proteine wird zwischen elektrophoretischen (Gel-basierenden) und chromatographischen (Gel-freien) Methoden unterschieden. Typische Elektrophoreseverfahren sind die SDS-PAGE [Laemmli (1970)] oder die 2D-PAGE [Klose (1975), O'Farrell (1975)], bei denen die Auftrennung der Proteine nach ihrem Molekulargewicht bzw. isoelektrischem Punkt (pI) und Molekulargewicht erfolgt.

Für die chromatographische Auftrennung werden die Proteine zuerst zu Peptiden verdaut [Kapitel: 2.4.5]. Anschließend erfolgt die Trennung der Peptide basierend auf ihren unterschiedlich starken Interaktionen mit der Matrix in einer Chromatographiesäule. Routinemäßig werden die Peptide mittels Umkehrphasenchromatographie nach ihrer Hydrophobizität getrennt. Allerdings sind auch mehrdimensionale Chromatographieverfahren, insbesondere die Kombination aus Ionenaustauschchromatographie und Umkehrphasenchromatographie verbreitet [Link et al. (1999), Wolters et al. (2001), Washburn et al. (2001)]. Die Vorteile der chromatographischen gegenüber der Gel-basierenden Verfahren sind der größere dynamische Bereich von bis zu 1:10.000, sowie die Möglichkeit, Proteine mit pI

kleiner 3 oder größer 11 und einem Molekulargewicht kleiner als 10 kDa oder größer als 200 kDa zu untersuchen [Washburn et al. (2001)]. Darüber hinaus ist der Gel-freie Ansatz weniger arbeitsaufwendig und weniger empfindlich gegen Kontaminationen mit Keratin [Bell et al. (2009)]. Mittels alternativer Chromatographieverfahren wie immobilisierter Metalaffinitätschromatographie [Kalume et al. (2003)] können gezielt Peptide mit bestimmten PTMs angereichert werden [Aguilar (2004), Seite 77]. Nachteilig für chromatographische Verfahren ist die größere Anfälligkeit für Probenverunreinigungen, wie die in Biogasanlagenproben vorkommenden Huminstoffe.

Für bisherige Metaproteomanalysen [Abram et al. (2011), Hanreich et al. (2012)] wurden die Proteine mittels 2D-PAGE aufgetrennt und einzelne Proteinspots mittels LC-MS/MS vermessen. Allerdings zeigen mehrere, neuere Arbeiten, dass für komplexe Umweltproben die Kombination von SDS-PAGE und LC-MS/MS Messungen mit sehr langem Gradienten auch vielversprechende Ergebnisse liefert [Schneider et al. (2012), Becher et al. (2013)]. Insbesondere durch die SDS-PAGE werden Verunreinigungen, wie Huminstoffe, vor der Chromatographie entfernt.

2.4.5 Proteinverdau

Vor der massenspektrometrischen Analyse müssen die Proteine zu Peptiden verdaut werden. Eine reproduzierbare Methode ist der enzymatische Verdau der Proteine. Meist wird dafür Trypsin genutzt, welches Proteine C-terminal nach den Aminosäuren Arginin und Lysin zerschneidet, solange sich kein Prolin vor diesen befindet. Für die bessere Zugänglichkeit der Schnittstelle werden zuvor die Disulfidbrückenbindungen durch die Zugabe Dithiothreitol reduziert und deren Rückbildung durch Alkylierung der Schwefelreste von Cysteinen verhindert. In Abhängigkeit davon, ob die Proben Gel-basierend oder Gel-frei analysiert werden, ist es möglich, die Proteine direkt (in solution digest) [Link and LaBaer (2011)] oder aus dem Gel (in gel digest) [Rosenfeld et al. (1992), Shevchenko et al. (1996)] heraus zu verdauen. Da der direkte Proteinverdau durch Probenverunreinigungen und Proteinagglomeration gestört wird, ist für Umweltproben der Umweg über das Gel oder ein Filter [Wiśniewski et al. (2009)] zweckmäßiger.

2.4.6 Massenspektrometrie

Massenspektrometrie stellt eine Möglichkeit zur Identifizierung von Proteinen dar. Bei dieser Methode wird das Masse-zu-Ladungs (m/z)-Verhältnis von Peptidionen und ihrer Bruchstücke im Hochvakuum bestimmt. Anschließend werden die Proteine auf der Basis des gemessenen m/z Verhältnisses mittels Datenbankabgleich identifiziert. Die Hauptbestandteile eines Massenspektrometers sind die Ionenquelle, der Massenanalysator und der

Detektor [Abbildung: 2.5].

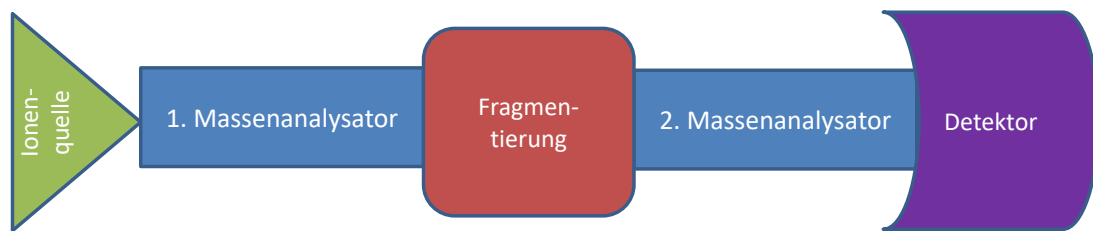


Abbildung 2.5: Schematischer Aufbau eines Massenspektrometers)

Ionenquelle

Für die Ionisation von Peptiden entwickelten *Tanaka et al. (1988)* die Matrix-unterstützte Laser-Desorption/Ionisation (MALDI) und *Fenn et al. (1989)* [Fenn et al. (1989)] die Elektronenspray-Ionisation (ESI). Bei MALDI werden die Peptide zusammen mit kleinen organischen Molekülen zu einer Matrix kokristallisiert. Durch Einwirken eines Lasers werden dann einzelne Moleküle aus der Matrix herausgelöst, wobei je ein Proton auf die Peptide übertragen wird. Im Gegensatz dazu werden bei ESI die Peptide zusammen mit einem sauren Lösungsmittel in ein elektrisches Feld injiziert. Dabei wird über eine Nadel das Gemisch vernebelt, das Lösungsmittel verdampft und die vorhandenen Protonen gehen auf die Peptide über. Auf diese Art entstehen bei ESI mehrfach geladene Ionen, während bei MALDI ausschließlich einfach geladene Ionen gebildet werden. Da beide Ionisationstechniken auf verschiedenen Mechanismen beruhen und die Ionisation verschiedener Peptidionen begünstigen, ist es mitunter sinnvoll, die Ergebnisse beider Techniken zu kombinieren [Yang et al. (2007)]. Da ESI-MS online an eine HPLC koppelbar sind, werden sie meistens für komplexe Proben eingesetzt. MALDI-MS erlaubt dagegen einen höheren Probendurchsatz und werden daher meist für die Identifikation von Proteinspots aus 2D-PAGE Gelen genutzt.

Massenanalysator

Für die Separation der Peptidionen können verschiedene Massenanalysatoren wie z. B. Flugzeitanalysatoren, Quadrupole oder Ionenfallen eingesetzt werden.

Im Flugzeitanalysator werden die Peptidionen durch ein elektrisches Feld beschleunigt. Da Peptidionen mit einem großen m/z Verhältnis weniger stark durch das elektrische Feld beschleunigt werden als Peptidionen mit kleinem m/z Verhältnis, sind diese langsamer. Abschließend kann aus der Zeit, welche die Peptidionen zum Detektor brauchen, das m/z Verhältnis berechnet werden.

Ein Quadrupol Massenanalysator trennt Peptidionen mit Hilfe von zwei orthogonalen elektrischen Feldern. Dabei funktioniert ein Feld als Tiefpass- und das andere als Hochpassfilter, wodurch nur Peptidionen mit einem bestimmten m/z -Verhältnis zum Detektor gelangen. Durch Veränderung der Stärken der elektrischen Felder erfolgt die Analyse des gesamten m/z -Bereichs.

Eine Weiterentwicklung der Quadrupole sind Ionenfallen, in denen die Peptidionen vor der Analyse in einem elektrischen Feld gesammelt werden, wodurch die Sensitivität steigt. Eine Besonderheit stellt die Orbitrap dar, da bei ihr Massenanalysator und Detektor kombiniert sind. In ihr werden die Peptide durch elektrische Felder in einer Kammer in einem radialen Orbit um eine Spindel gefangen. Da die Injektion der Peptide dezentral in die Orbitrapkammer erfolgt, schwingen die Peptide auf ihren Umlaufbahnen auch axial zur Spindel. Dabei induzieren sie einen Strom, aus dem durch Fouriertransformation das m/z -Verhältnis bestimmt werden kann.

Detektor

Während der Nachweis der Peptide bei der Orbitrap durch die Induktion eines Stromes während der Rotation um eine Spindel erfolgt, zählen andere Detektoren (meist Mikrokanalplatte) wie viele Peptide auf diesem Einschlagen. Dabei induziert der Einschlag der Peptide einen Stromfluss, welcher durch den Detektor weiter verstärkt und gemessen wird. Durch die Detektion der Peptidionen über den gesamten m/z -Bereich entsteht dann das typische MS-Spektrum. Entscheidend für die Qualität der Messung ist die Genauigkeit (Auflösung) und Sensitivität mit der das MS die Peptidionen analysiert. Von den aufgeführten MS weisen die Orbitrap-MS die beste Auflösung (bis zu 240.000 für Peptidionen mit m/z von 400), Sensitivität und Massengenauigkeit (etwa 10 ppm) auf.

Fragmentierung von Peptiden

Da Peptide aufgrund gleicher oder ähnlicher Aminosäurezusammensetzung das gleiche m/z -Verhältnis aufweisen können, ist eine genauere Charakterisierung notwendig. Dafür werden die wichtigsten Peptidionen (meist die 5-30 häufigsten Peptide im MS-Spektrum) selektiert und fragmentiert. Die entstehenden Bruchstücke werden als Fragmentionen bzw. deren Gesamtheit als Fragmentionenspektrum bezeichnet [Abbildung: 2.6]. Meist wird für die Fragmentierung von Peptiden die *collision-induced dissociation* (CID) Fragmentierung [Mitchell Wells and McLuckey (2005)] verwendet, bei der die Peptidionen durch Kollision mit einem Inertgas vorzugsweise zu b-, y- Ionen zerfallen. Eine Weiterentwicklung des CID speziell für Orbitrap-MS ist die *higher-energy collisional dissociation* (HCD) [Olsen et al. (2007)], bei der die Kollision in einer zusätzlichen HCD-Zelle erfolgt. Eine Alternative zu diesen Techniken ist die *electron-transfer dissociation* (ETD) [Syka et al. (2004)]. Bei

dieser fragmentieren die Peptidionen weniger stark als bei CID und HCD, wodurch sie sich insbesondere für die Detektion von PTMs eignet. Dabei werden die Peptide durch Elektronenübertragung mit einem radikalischen Anion vorzugsweise an der Peptidbindung zu c- und z-Ionen gespalten.

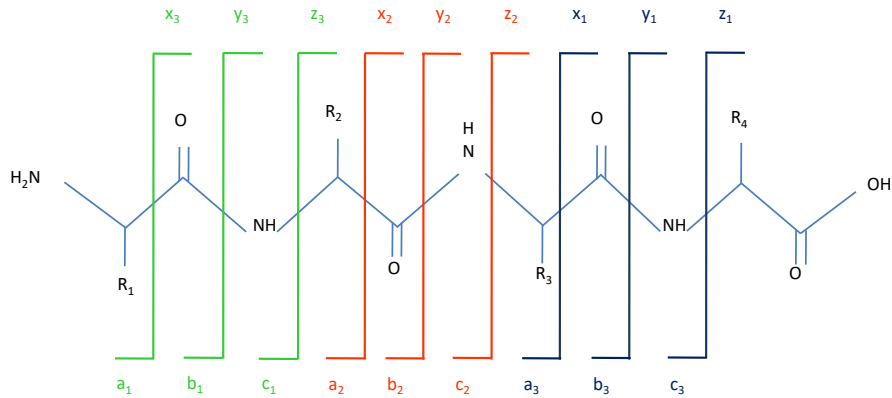


Abbildung 2.6: *Charakteristische Fragmentationen von Peptiden* nach *Biemann et al. (1988)* [Biemann (1988)]

Massenspektrometrie und Metaproteomanalysen von Biogasanlagenproben

Zu Beginn dieser Arbeit gab es zwei Metaproteomanalysen von Biogasprozessen [Abram et al. (2011), Hanreich et al. (2012)] in welchen die Identifizierung der Proteine durch Identifizierung von 2D-PAGE Proteinspots mittels Q-Star XL MS/MS (Triple Quadrupol) erfolgte. Allerdings zeigte die Arbeit von *Schneider et al. (2012)* [Schneider et al. (2012)] deutlich, dass für die Identifizierung von Proteinen aus komplexen Umweltproben ein Orbitrap-MS/MS aufgrund der höheren Massengenauigkeit und Sensitivität besser geeignet ist.

2.4.7 Massenspektrometrie-basierte Proteinidentifizierung

Die Protein- und Peptididentifizierung erfolgt durch Vergleich der MS/MS-Spektren mit den Spektren von theoretisch verdauten und fragmentierten Proteinen aus Proteindatenbanken mittels Datenbanksuchalgorithmen. Entsprechende Algorithmen sind u.a. Mascot [Perkins et al. (1999)], X!Tandem [Craig and Beavis (2004)], OMMSA [Geer et al. (2004)], InsPecT [Tanner et al. (2005)] und Crux [Park et al. (2008)]. Für die Aufarbeitung der massenspektrometrischen Rohdaten und die Visualisierung der Identifizierungsergebnisse sind diese Algorithmen meist in eine entsprechende bioinformatische Plattform wie OpenMS [Sturm et al. (2008)], den Proteome DiscovererTM oder ProteinScape [Thiele et al. (2010)]

eingebettet. Zum Teil besitzen diese Plattformen auch eine Schnittstelle, um die Rohdaten und Ergebnisdaten in entsprechenden öffentlichen MS-Ergebnisdatenbanken wie z. B. PRIDE [Vizcaíno et al. (2013)] zu speichern.

Proteindatenbanken

Die am häufigsten verwendeten Datenbanken für die Proteinidentifikation sind NCBIInr [Acland et al. (2013)], UniProtKB/SwissProt und UniProtKB/TrEMBL [UniProt Consortium (2012)]. Ein Unterschied zwischen den Datenbanken ist die Größe, wobei UniProtKB/SwissProt (551.193 Proteineinträgen) deutlich kleiner als NCBIInr (58.496.614 Proteineinträgen) und UniProtKB/TrEMBL (62.148.086 Proteineinträgen) ist (Stand 11 Mai 2016). Auf der anderen Seite sind die Einträge für UniProtKB/SwissProt manuell überprüft und mit Informationen über Taxonomie, Funktion und Struktur verknüpft.

Neben der Verwendung dieser umfassenden Datenbanken ist es auch möglich für die jeweilige Probe ein spezifisches Metagenom zu sequenzieren und als Datenbank zu verwenden [Tyson et al. (2004), Ram et al. (2005)].

Datenbanksuchalgorithmen

Die Arbeitsweise der Datenbanksuchalgorithmen basiert auf drei wichtigen Schritten. Im ersten Schritt werden alle Peptide aus der Datenbank ausgewählt, welche die gleiche Masse wie das gemessene Peptidion aufweisen. Dann wird für diese Peptide im zweiten Schritt ihr theoretisches MS/MS Spektrum ermittelt und berechnet wie gut sie das gemessene MS/MS Spektrum erklären. Im dritten Schritt wird dann entschieden, welche Peptidvorschläge valide Identifizierungen des Peptides sind. Dafür kann entweder die Wahrscheinlichkeiten der einzelnen Peptidvorschläge aus der Güteverteilung der Peptidvorschläge berechnet werden oder ein Schwellwert für glaubwürdige Identifikationen durch Berechnung der *false discovery rate* (FDR) [Gleichung: 2.18] [Elias et al. (2005)] ermittelt werden.

Die FDR ist das Verhältnis der Spektren, die in einer randomisierten Version der Originaldatenbank identifiziert wurden, geteilt durch die Zahl aller Spektren, die in der Original- und Zufallsdatenbank gefunden wurden. Aus diesem Wert kann abgelesen werden, wie viele identifizierte Peptide wahrscheinlich falsch positiv identifiziert wurden, aber nicht welche. Typische Verfahren für das Randomisieren der Originaldatenbank sind das Umdrehen oder Durchmischen der Proteineinträge, sowie das Erstellen einer kompletten Zufallsdatenbank [Elias and Gygi (2010)].

Wie schon die Arbeit von *Jagtap et al. (2013)* zeigt, ist die Größe der Datenbank für die FDR Berechnung entscheidend. Allgemein gilt, dass sich durch Verdopplung der Datenbankgröße auch die Wahrscheinlichkeit für ein Zufallstreffer und damit auch die FDR verdoppelt. Es ist daher wichtig, dass die Datenbank nicht nur alle für die Probe relevante

Proteineinträge enthält, sondern die Komplexität der Datenbank auch der Komplexität der Probe entspricht. Wenn dies nicht beachtet wird, die Datenbank also zu viele oder zu wenige Datenbankeinträge beinhaltet, wird die FDR über- bzw. unterbestimmt. Als Folge werden mehr korrekte Identifizierungen als nötig ignoriert oder eigentlich falsche Identifikationen als wahr gewertet. Basierend auf den Peptiden, welche in den Proteindatenbanken bzw. in den Metagenomen identifiziert wurden, erfolgt dann die Identifizierung der Proteine. Je nach Ansatz gilt ein Protein als identifiziert, wenn ein oder mehrere Peptide identifiziert wurden [Bradshaw et al. (2006)]. Dabei ist die Identifizierung eines Proteins durch zwei Peptide nicht zwangsläufig besser als durch ein Peptid. Es kommt vielmehr auf die Qualität der Identifizierung an [Gupta and Pevzner (2009)]. Auch hat sich in der Praxis gezeigt, dass für circa 30% der Proteine von komplexen Proben nur ein Peptid identifiziert werden kann. Die Glaubwürdigkeit von Peptididentifikationen kann durch Kombination verschiedener Datenbanksuchalgorithmen [Vaudel et al. (2011), Shteynberg et al. (2013), Muth et al. (2013b)] oder die Überprüfung der Proteinidentifikation durch komplexe stochastische Modelle gesteigert werden.

$$FDR = \frac{FP}{TP + FP} \quad (2.18)$$

FDR: *false discovery rate*
 FP: Falsch positive Identifikationen
 TP: Richtig positive Identifikationen

Alternative Methoden zur Proteinidentifikation

Um Proteine, deren Sequenz nicht in Datenbanken vorhanden ist, zu identifizieren, kann eine *de novo* Sequenzierung [Frank and Pevzner (2005)] durchgeführt werden. Dabei wird versucht aus den einzelnen Fragmentspektren die ursprüngliche Aminosäuresequenz zu bestimmen. Anschließend wird durch *MS-basic local alignment search tool* (BLAST)-Suche [Shevchenko et al. (2001)] nach homologen Proteinen mit ähnlicher Aminosäuresequenz gesucht. Der Nachteil der *de novo* Sequenzierung ist, dass jedes Ergebnis manuell überprüft werden muss.

Alternative ist die Suche gegen Spektralbibliotheken möglich. Hier werden neue Spektren mit denen aus alten Messungen verglichen und um Redundanzen zu entfernen zu Clustern gruppiert [Lam et al. (2007)]. Die Spektralbibliothek profitiert davon, dass die Bibliotheksspektren die Peptidionisation besser widerspiegeln als die theoretische Berechnung der Spektren. Zusätzlich können im Anschluss relevante Spektren durch *de novo* Sequenzierung identifiziert werden.

2.4.8 Bestimmung der Proteinfunktion und Taxonomie

Für die Charakterisierung von mikrobiellen Gemeinschaften ist es wichtig, die Proteinidentifikationen mit einer bestimmten Taxonomie und Funktion zu verknüpfen.

Die Arbeiten von *Abram et al. (2011)* und *Hanreich et al. (2012)* haben bereits das Problem aufgezeigt, dass die identifizierten Peptide zu mehreren, funktionell ähnlichen Proteinen gehören können. Die Konsequenz ist, dass redundante Proteinidentifikationen, welche vor der Zuordnung der Taxonomie und Funktion erst gruppiert werden müssen. *Lü et al. (2014)* schlagen dafür vor, Proteinidentifikationen mit einem gemeinsamen Peptid zusammenzufassen, während *Keiblinger et al. (2012)* und *Schneider et al. (2012)* dies nur bei Übereinstimmung des gesamten Peptidsets tun. Die Gruppierung nach dem gesamten Peptidset folgt dabei der Idee die einfachste Proteinerklärung für alle identifizierten Peptide zu finden (Prinzip von Parsimonie) [*Nesvizhskii et al. (2003)*]. Im Gegensatz zu Reinkulturproteomik, aus der dieser Ansatz stammt, ist aber die Peptidzuordnung für Metaproteomanalysen komplizierter, wie von *Nesvizhskii et al. (2005)* diskutiert. Eine Erweiterung dieser Regeln könnte sinnvoll oder sogar nötig sein. Identifizierte Peptide können nämlich nicht nur zu einem oder mehreren identifizierten Proteinen gehören, sondern auch zu einem weiteren nicht sequenzierten Protein [*Hettich et al. (2013)*]. Bisherige Möglichkeiten um dies zu berücksichtigen sind Gruppierungen der Proteinidentifikationen nach Sequenzähnlichkeit basierend auf den UniRef Clustern [*Lü et al. (2014)*] oder den Klassen der Genontologien [*Gotelli et al. (2012)*].

Für die anschließende Festlegung der Taxonomie der Proteingruppen kann entweder der gemeinsame phylogenetische Vorfahre aller identifizierten Proteine [*Huson et al. (2007)*] oder die Taxonomie, die die Peptididentifikationen am besten erklärt, genutzt werden [*Nesvizhskii and Aebersold (2005)*].

Schwieriger ist die Beschreibung der Funktion, für welche es verschiedene Quellen gibt. Ursprünglich wurde die Funktion und mögliche Bindungspartner der Proteine einzeln untersucht und in Datenbanken abgelegt. Für neue Proteine wird mittlerweile die Funktion durch Sequenzvergleich mit diesen Datenbanken bestimmt und nach funktional ähnlichen Proteinen gruppiert. Allerdings gibt es für diesen Sequenzvergleich verschiedene Ansätze wie z. B. *KEGG orthologies* (KO) [*Kanehisa et al. (2004)*], *cluster of orthologues genes* [*Tatusov et al. (2000)*] oder *evolutionary genealogy of genes: non-supervised orthologous groups* [*Powell et al. (2012)*]. Eine Weiterentwicklung dieser Datenbanken sind PFAM [*Finn et al. (2010)*], TIGRFAMs [*Haft et al. (2013)*] und Interpro [*Hunter et al. (2012)*], welche neben der reinen Sequenzähnlichkeit auch funktionelle Domänen berücksichtigt. Ein anderes System der Funktionsklassifikation ist die EC [*Bairoch (2000)*]. Dabei erfolgt die Gruppierung der Proteine nach den Enzymreaktionen, die sie katalysieren.

Alternative Klassifikationen, wie die UniProt *keywords* [UniProt Consortium (2012)] oder die *gene ontology* [Ashburner et al. (2000)], nutzen neben der spezifischen Funktion noch weitere Kriterien, wie die biologische Funktion, die zelluläre Lokalisation oder mögliche Bindungspartner (Liganden).

Für die Visualisierung vieler Proteine gibt es entsprechende Internetapplikationen, wie die KEGG *pathway* Datenbank [Kanehisa and Goto (2000)], MeTaCyc [Caspi et al. (2012)] und Reactome [Croft et al. (2011)]. Zusätzlich bieten diese Applikationen die Möglichkeit, die Expression und Repression der Proteine in Folge eines Stimulus zu visualisieren.

2.4.9 Quantifizierung einzelner Proteinen

Der Vergleich verschiedener Proben erfordert eine reproduzierbare Quantifizierung der Proteine über einen möglichst großen dynamischen Bereich. Die absolute Quantifizierung von Proteinen ist allerdings wegen der unterschiedlichen Ionisierungseffizienz der einzelnen Peptide schwierig. Folglich basieren die meisten Quantifizierungsverfahren auf einer relativen Quantifizierung zwischen zwei Proben.

Bei Gel-basierenden Proteomuntersuchungen erfolgt die Quantifizierung meist durch Vergleich der Intensitäten von Banden oder Spots, welche mit kolloidaler Coomassie [Neuhoff et al. (1985)] oder Silber [Blum et al. (1987)] gefärbten wurden. Die Silberfärbung ist mit einer Sensitivität von einem Nanogramm Protein zehnmal so empfindlich wie die kolloidale Coomassie Färbung, erfordert aber auch deutlich mehr Arbeitsaufwand. Schneller aber deutlich teurer ist die Arbeit mit Fluoreszenzfärbung, die nicht nur ähnlich sensitiv wie die Silberfärbung ist, sondern auch einen deutlich größeren dynamischen Bereich hat [Patton (2002)]. Zusätzlich können mit verschiedenen Fluoreszenzfarbstoffen markierte Proben zusammen in einem Gel analysiert und verglichen werden [Ünlü et al. (1997)], wodurch Gel zu Gel Variationen ausgeschlossen werden können.

Bei Gel-freien Proteomuntersuchungen kann die Quantifizierung durch Einbau von stabilen Isotopen in die Proteine, durch chemische Modifikationen oder Label-frei erreicht werden. Es ist sogar möglich, durch Zugabe einer definierten Menge an synthetischen Peptiden, welche schwere Isotope enthalten, absolut zu quantifizieren. Dabei kann aus der Menge und dem Intensitätsverhältnis zum unmarkierten Peptid die absolute Menge des Proteins berechnet werden [Gerber et al. (2003)]. Der Nachteil dieser Methode ist, dass für jedes zu quantifizierende Protein ein eigenes Peptid hergestellt werden muss.

Eine andere Möglichkeit mit stabilen Isotopen zu quantifizieren ist die Kultivierung von Zellen in speziellem Medium, welches z. B. in der Zellkultur $^{13}\text{C}/^{15}\text{N}$ markiertes Arginin oder Lysin enthält [Ong et al. (2002)]. Anschließend wird diese Probe mit einer nicht markierten Probe verglichen. Als weitere Möglichkeit können Mikroorganismen mit einem

Substrat gefüttert werden, welches mit stabilen Isotopen markiert wurde (Protein-SIP [Jehlich et al. (2008)]). Durch Einbau der Isotopen in einzelne Proteine kann dann geprüft werden, welche Spezies stoffwechselaktiv waren. Der Vorteil beider Methoden ist, dass durch die biochemische Äquivalenz von markierten und nicht markierten Proben, sowie der Markierung vor der Probenvorbereitung keine Verzerrung bei der Proteomuntersuchung entsteht.

Für die Quantifizierung können Peptide und Proteine auch mit synthetischen Molekülen modifiziert werden. Diese Moleküle gibt es mit verschiedenen Molekulargewichten, als leichte und als schwere Variante. Für die Quantifizierung der Proben mit MS werden die Peptidintensitäten beider, verglichen. Typische Bindestellen für die chemische Modifikation sind die freien Aminosäuregruppen von Lysine wie z. B. für ICPL [Lottspeich and Kellermann (2011)] oder iTRAQ [Ross et al. (2004)], oder die Thiolgruppe der Cysteine wie z. B. bei ICAT [Gygi et al. (1999)].

Während die meisten Quantifizierungsmethoden mit Rein- und synthetische Kulturen hervorragend funktionieren, stören bei Umweltproben häufig Huminstoffe alle Quantifizierungen, die auf einer Modifikation des Proteins/ Peptides basieren [Bastida et al. (2009)]. Eine preisgünstige und zeitsparende Alternative zu schweren Isotopen und chemischen Markierungen ist die Label-freie Quantifizierung. Proteine können durch die Intensität der Peptidionen oder nach Häufigkeit ihrer Fragmentionen [Podwojski et al. (2010)] quantifiziert werden. Zwei gängige Maße für die Häufigkeit der Fragmentionen sind die Anzahl der identifizierten Peptide und Spektren. Da diese Größen aber von der Aminosäuresequenzlänge der Proteine und der Komplexität der Probe abhängen, werden oft zwei normalisierte Formen dieser Größen verwendet. Die erste Form ist der *exponentially modified protein abundance index* (emPAI [Ishihama et al. (2005)], [Formel: 2.19, 2.20]). Dabei wird das Verhältnis aus identifizierten Peptiden zu theoretisch identifizierbaren Peptiden gebildet und exponentiell modifiziert.

$$PAI = \frac{N_{beobachtet}}{N_{beobachtbar}} \quad (2.19)$$

PAI: Proteinhäufigkeitsindex

$N_{beobachtet}$: Anzahl an Peptiden, welche für das Protein identifiziert wurden

$N_{beobachtbar}$: Anzahl an Peptiden, welche für das Protein theoretisch beobachtet werden können

$$emPAI = 10^{PAI} - 1 \quad (2.20)$$

emPAI: *exponentially modified protein abundance index*

Die zweite Form ist der *normalized spectral abundance factors* (NSAF) [Zybailov et al. (2007)], welcher von *Schneider et al. (2012)* erfolgreich für Umweltproben genutzt wurde. Für diesen wird die Spektrenanzahl aller Peptide eines Proteins mit der Aminosäuresequenzlänge des Proteins multipliziert und durch die Summe der Spektrenanzahl für alle Proteine multipliziert mit deren individuellen Längen geteilt [Gleichung: 2.21]. Der Vorteil ist, dass dabei sowohl die Proteinlänge als auch die Komplexität der Proben berücksichtigt werden. Laut *Liu et al. (2004)* ist es dabei möglich Proteine verlässlich über zwei Größenordnungen zu quantifizieren.

$$NSAF_{Protein} = \frac{SpeC_{Protein} * L_{Protein}}{\sum_{i=1}^n SpeC_{Protein,i} * L_{Protein,i}} \quad (2.21)$$

NSAF_{Protein}: *normalized spectral abundance factors*
 SpeC_{Protein}: Anzahl Spektren zum Protein
 L_{Protein}: Aminosäuresequenzlänge des Proteins
 n: Gesamtanzahl an identifizierten Proteinen

2.5 Statistische Methoden zur Datenanalyse

Für das bessere Verständnis von Biogasprozessen ist es erforderlich zu erkennen, wie einzelne Prozessparameter mit der Taxonomie und Funktion der mikrobiellen Gemeinschaften zusammenhängen.

Ausgangspunkt ist die Datenmatrix, in welche alle Messgrößen zusammengefasst werden. Nach entsprechender Normalisierung der Daten kann mittels Median, Standardabweichung und Varianz eine Übersicht über die Verteilung der einzelnen Messgrößen geschaffen werden, z.B. durch Visualisierung als Box Plots [Benjamini (1988)]. Alternativ könnte auch der Mittelwert betrachtet werden. Dieser ist allerdings nicht so robust gegenüber Ausreißern [Rousseeuw and Bassett Jr (1990)].

Korrelationsanalysen helfen Zusammenhänge zwischen einer Vielzahl von Parametern zu erkennen. Typische Berechnungen für den Korrelationskoeffizienten sind die Pearson-Korrelation [Gleichung: (2.22)] [Pearson (1895)] und die Spearmans Rangkorrelation [Gleichung: (2.23)] [Myers et al. (2010)]. Die Vorteile der Spearmans Rangkorrelation sind, dass

sie robuster ist und auch nichtlineare Zusammenhänge anzeigen kann. Da bei entsprechend großen Datensätzen die Wahrscheinlichkeit von zufälligen Korrelationen besteht, ist es wichtig die Korrelationen zusätzlich einem Signifikanztest zu unterziehen. Dafür kann für die Pearson-Korrelation der Student's t-Test bzw. für die Spearmans Rangkorrelation der Permutationstest verwendet werden.

$$r_{xy} = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)} \quad (2.22)$$

r_{xy} : Pearsons Rangkorrelationskoeffizient
 $Cov(X, Y)$: Kovarianz der Verteilung X und Y
 $\sigma(X), \sigma(Y)$: Standardabweichung der Verteilung X und Y

$$r_s = \frac{Cov(rg_x, rg_y)}{srg_x srg_y} \quad (2.23)$$

r_s : Spearmans Rangkorrelationskoeffizient
 $Cov(rg_x, rg_y)$: Kovarianz der Ränge der Verteilung von X und Y
 srg_x/srg_y : Standardabweichung der Verteilung X und Y

An Biogasprozessen sind eine Vielzahl an mikrobiellen Spezies und Funktionen beteiligt, welche wiederum von einer großen Anzahl an Prozessgrößen wie z. B. Temperatur, pH-Wert beeinflusst werden. Eine Möglichkeit diese komplexen Daten besser zu analysieren stellen Lernalgorithmen dar. Unüberwachte Lernalgorithmen, wie Clusteranalyse [Sokal (1958)] und Hauptkomponentenanalyse [Person (1901)] ordnen die Proben ohne Vorwissen. Überwachte Lernalgorithmen basieren auf einer vorherigen Klassifikation der Daten und versuchen aus diesen eine Erklärung der Klassifikation zu generieren. Ein Beispiel dafür ist das *decision tree learning* [Quinlan (1986)]. Bei diesem wird ein Model in Form eines Baumes erstellt, welches durch einfache Entscheidungen, ob ein Schwellenwert einer Messgröße überschritten wurde, versucht, die Klassifikation der Daten wiederzugeben. Um abschließend die Qualität des gelernten Models zu beurteilen, wird diese anschließend mit einem zweiten Testdatensatz überprüft.

Material und Methoden

Im folgenden Abschnitt werden die in dieser Arbeit verwendeten Methoden vorgestellt. Detaillierte Arbeitsanweisungen sind zusätzlich auf einer CD [Anhang: 7.14 CD] abgelegt. Die für die einzelnen Methoden verwendeten Geräte [Anhang: 7.4] und Chemikalien [Anhang: 7.3] sind ebenfalls im Anhang zu finden.

3.1 Beprobung von Biogasanlagen

Im Rahmen dieser Doktorarbeit wurde eine Vielzahl von Biogasanlagen sowie einige Laborreaktoren untersucht [Anhang: 7.5]. Die Beprobung und Aufnahme der Prozessparameter erfolgte dabei meist seitens des Biogasanlagenbetreibers [Arbeitsanweisung (SOP) 10.5 siehe CD]. Um für vergleichbare Untersuchungsbedingungen zu sorgen, wurde Wert darauf gelegt, dass die Biogasanlagen vor der Beprobung gemischt und das entsprechende Totvolumen des Abflusses verworfen wurde. Anschließend wurden die Proben bei $-20\text{ }^{\circ}\text{C}$ gelagert.

3.2 Probenvorbereitung

Die Probenvorbereitung umfasst die Extraktion der Proteine aus den Biogasproben [Kapitel: 3.2.1], die Quantifizierung der extrahierten Proteinmenge [Kapitel: 3.2.2] sowie die Fällung einer definierten Proteinmenge mit Aceton [Kapitel: 3.2.3] für die anschließende Proteinseparation [Kapitel: 3.3].

3.2.1 Phenolextraktion

Für die Extraktion der Proteine [SOP 2.3 siehe CD] wurden je 1 g Silikaperlen, 400 mg oder 400 μL Probe, 400 μL 2 M Saccharoselösung und 700 μL Phenollösung in 2 ml Reaktionsgefäße gegeben und diese 30 min bei 30 Hz in einer Kugelmühle geschüttelt. Um danach die Phenolphase von der wässrigen Phase zu trennen, wurden die Reaktionsgefäße zentrifugiert (10 min, 10.000 g, RT) und jeweils die oberen Phenolphasen in einem 15 mL Reaktionsgefäß gesammelt. Anschließend wurde die Phenolphase mit der gleichen Menge

1 M Saccharoselösung auf einem Thermomixer gewaschen (10 min, 300 min⁻¹, RT) und nach erneuter Zentrifugation (10 min, 10.000 g, RT) in ein neues 15 mL Reaktionsgefäß überführt.

Um die Proteine aus der Phenolphase auszufällen, wurden einmal die vierfache (4 v/v) und einmal die dreifache (3 v/v) Menge an eiskaltem 0.1 M Ammoniumacetat in Methanollösung zugegeben und je für 1 h bei -20 °C im Tiefkühlschrank inkubiert. Nach jeder Inkubation wurden die Proben zentrifugiert und der Überstand verworfen. Um restliche Verunreinigungen zu entfernen, wurde jeweils zweimal abwechselnd mit der dreifachen Menge (3 v/v) eiskaltem 80% Aceton und 70% Ethanol gewaschen und 15 min bei -20 °C im Tiefkühlschrank inkubiert. Dazwischen erfolgte jeweils ein Zentrifugationsschritt (10 min, 10.000 g, 4 °C) und das Verwerfen des Überstandes. Nach dem letzten Waschschrift wurde das Proteinpellet in 2 mL Harnstoffpuffer gelöst und bis zur weiteren Nutzung bei -20 °C gelagert.

3.2.2 Proteinquantifizierung mittels Amidoschwarz

Für die Quantifizierung der extrahierten Proteine wurde der Amidoschwarzassay [SOP 3.1, siehe CD] genutzt. Dafür wurden die Proteine aus 50 µL Probe mittels 300 µL Färbelösung gefällt und der Überstand nach Zentrifugation (16.400 g, 5 min, RT) verworfen. Nach zwei Wasch- und Zentrifugationsschritten (16.400 g, 5 min, RT) mit je 500 µL Waschlösung wurde das Pellet in 1 mL 0,1 M Natriumhydroxidlösung aufgenommen. Abschließend wurde die Proteinlösung photometrisch bei 615 nm analysiert und die Proteinmenge durch Vergleich mit einer Eichkurve aus einer definierten Menge Rinderserumalbumin im Bereich von 0 -1,0 µg/µg berechnet.

3.2.3 Acetonfällung

Um die Proteine aus dem Harnstoffpuffer mit Aceton auszufällen [SOP 2.4, siehe CD], musste die Proteinlösung zuerst mit der gleichen Menge VE Wasser verdünnt werden. Anschließend erfolgte die Proteinfällung durch Zugabe der fünffachen Menge (5 v/v) eiskalter 100% Acetonlösung. Nach Inkubation bei -20 °C über Nacht und Zentrifugation (30 min, 4 °C, 10.000 g/ 16.400 g je nach verfügbarer Zentrifuge/ Probenmenge) wurde der Überstand verworfen und das restliche Aceton unter einem Abzug ausgedampft.

3.3 Proteinseparation

Die Probenkomplexität kann durch die Separation der Proteine nach ihrem Molekulargewicht mittels SDS-PAGE [Kapitel: 3.3.1] oder nach isoelektrischem Punkt und Moleku-

largewicht mittels OFFGEL-Fraktionierung [Kapitel: 3.3.2] verringert werden.

3.3.1 SDS-PAGE

Nach der Reinigung der Glasplatten und dem Zusammenbau der Gelgießeinrichtung wurde als erstes das Trenngel für die SDS-PAGE [SOP 4.1, siehe CD] gegossen, mit wassergesättigtem iso-Butanol überschichtet und für 30 min polymerisiert. Im Anschluss wurde das Butanol wieder entfernt, das Sammelgel auf das Gel pipettiert und der Kamm luftblasenfrei eingesetzt. 30 Minuten später konnte der Kamm entfernt werden, das Gel in die Elektrophoresekammer eingebaut und diese mit SDS-Laufpuffer befüllt werden.

Über Nacht wurde die gewünschte Proteinmenge mit Aceton gefällt [Kapitel: 3.2.3], in 21 μL SDS-Probenpuffer auf einem Thermomixer inkubiert (5 min, 60 °C, 1400 min^{-1}) und unlösliche Partikel abzentrifugiert (10 min, 16.400 g, RT). Anschließend wurden die Taschen im SDS-Gel mit 20 μL Probe bzw. 2 μL Proteinstandard beladen und die Elektrophoresekammer an das Stromversorgungsgerät angeschlossen. Bis die Proben das Trenngel erreicht hatten, wurde ein Strom von 10 mA pro Gel und danach bis zum unteren Ende des Gels ein Strom von 20 mA pro Gel angelegt. Nach Abschalten des Stromversorgungsgerätes konnten die Gele entnommen und die Proteine visualisiert werden [Kapitel: 3.4].

3.3.2 IEF mit dem 3100 OFFGEL-Fraktionator

Der 3100 OFFGEL-Fraktionator wurde zur IEF von Proteinen in Lösung verwendet [SOP 4.4, siehe CD]. Nach dem Einlegen der IPG-Streifen und der 12er Kämme (13 cm IPG Streifen, pH 4–7) wurden die Streifen mit 40 μL Rehydratisierungslösung pro Kammer für 15 min rehydratisiert. Anschließend wurde je 1 mg Proteinprobe mit 1,44 mL OFFGEL Stock Lösung (1,25X) verdünnt (OFFGEL Protein Sample) und jeweils 150 μL in jede Tasche pipettiert. Vor der Fokussierung wurden die Kontaktstellen mit feuchten Filterpapieren (je 10 μL VE Wasser) sowie zusätzlich an der Anode mit 400 μL bzw. an der Kathode mit 600 μL Mineralöl bedeckt, bevor die Elektroden aufgesetzt wurden. Für die Separation der Proteine wurde das Programm OG12 für 200 kWh (4500 V, 50 μA) angewendet.

3.4 Visualisierung und Quantifizierung einzelner Proteine

Um den Erfolg der Probenvorbereitung und Proteinseparation zu überprüfen, wurden die Gele nach der SDS-PAGE mittels kolloidaler Coomassiefärbung gefärbt [Kapitel: 3.4.1].

3.4.1 Kolloidale Coomassiefärbung

Nach der Entnahme der Gele aus der Elektrophoresekammer begann die kolloidale Coomassiefärbung [SOP 5.1, siehe CD] mit der Inkubation der Gele in Fixierlösung für eine 1 h auf einem Schüttler (20 min^{-1} , RT). Nach dem Abgießen der Fixierlösung wurden diese zweimal mit VE Wasser gewaschen und über Nacht in Färbelösung geschüttelt (20 min^{-1} , RT). Am nächsten Tag konnte überschüssiger Farbstoff durch mehrfaches Waschen in VE Wasser entfernt und die Gele gescannt werden (Transmissionslicht, 48 Bit Farbe, Helligkeit: 0, Kontrast: 0, Gamma: 1,2).

Für eine längere Lagerung der Gele wurden diese zwischen Cellophanfolien getrocknet [SOP 5.5, siehe CD]. Dafür wurden die Cellophanfolie 3 min und die Gele 10 min in Glycerollösung eingeweicht und dann in einem Trocknungsrahmen über Nacht unter einem Abzug getrocknet.

3.5 Tryptischer Verdau und LC-MS/MS Messung

Für die MS-basierte Analyse der Proteine wurden die Gelabschnitte der SDS-PAGE [Kapitel: 3.3.1] mit einem Skalpell ausgeschnitten, in Würfel mit 1 mm Kantenlänge zerkleinert und in einem 1,5 mL Reaktionsgefäß gesammelt. Anschließend wurden die Proteine im Gel tryptisch zu Peptiden verdaut [Kapitel: 3.5.1] und die Peptide mittels MS analysiert [Kapitel: 3.5.2].

3.5.1 Tryptischer Verdau

Um den Coomassie-Farbstoff nach der Proteinvisualisierung zu entfernen, wurden als erster Schritt des tryptischen Verdaus [SOP 6.3, siehe CD] die Gelwürfel einmal über Nacht und einmal für 1 h auf einem Schüttler (500 min^{-1} , RT) mit $900 \mu\text{L}$ Waschlösung gewaschen. Anschließend erfolgte ein weiterer Waschschriff mit $900 \mu\text{L}$ Acetonitril (500 min^{-1} , RT), gefolgt vom kompletten Verdampfen des Acetonitrils durch Vakuumzentrifugation. Für die Reduktion der Proteine wurden die Gelwürfel 30 min in 45 mM DTT (500 min^{-1} , RT) inkubiert und danach mit $900 \mu\text{L}$ Acetonitril gewaschen (500 min^{-1} , RT). Die Alkylierung der Proteine erfolgte durch Zugabe von 100 mM IAA (500 min^{-1} , RT, im Dunkeln) für 30 min, gefolgt von Waschschriffen mit jeweils $900 \mu\text{L}$ Acetonitril, 25 mM ABC-Puffer und Acetonitril.

Nachdem das restliche Acetonitril durch Vakuumzentrifugation entfernt wurde, erfolgte der eigentliche tryptische Verdau durch die Zugabe von $100 \mu\text{L}$ Trypsinpuffer über Nacht ($37 \text{ }^\circ\text{C}$, Verhältnis Protein zu Trypsin: 1/100).

Für die Extraktion der tryptischen Peptide wurde einmal $100 \mu\text{L}$ Extraktionspuffer I und

einmal 100 μL Extraktionspuffer II (500 min^{-1} , RT) für 30 min zugegeben. Die Überstände wurden in einem neuen Reaktionsgefäß gesammelt und durch Vakuumzentrifugation eingetrocknet. Vor der MS-Analyse wurden die Peptide dann in 20-30 μL Probenpuffer aufgenommen, zentrifugiert (30 min, 13.000 g, 4 °C) und alle löslichen Peptide in ein HPLC-Vial überführt.

3.5.2 Massenspektrometrie

Für die massenspektrometrische Analyse standen ein AmaZon ETD ESI-Ion Trap MS/MS [SOP 6.5, siehe CD, Anhang: 7.8] und ein Elite Hybrid Ion Trap-Orbitrap MS/MS zur Verfügung [SOP 6.6, siehe CD, Anhang: 7.7]. Beide Geräte waren online an ein UltiMate 3000 RSLCnano splitless Chromatographiesystem gekoppelt, in welches je eine Umkehrphasensäule zum Sammeln der Peptide und eine zum Trennen eingebaut waren. Nach dem Binden der Peptide an die Sammelsäule erfolgte die Elution der Peptide über die Trennsäule mit einem 30 min bzw. 120 min Gradient von 4% Lösungsmittel B auf 55% Lösungsmittel B. Die entsprechenden einzelnen Geräte und Methodenparameter können dem Anhang entnommen werden [Anhang: 7.7, 7.8].

3.6 Bioinformatische und statistische Datenauswertung

Für die Auswertung der Metaproteomuntersuchungen war der Einsatz von entsprechender Software für die Proteinidentifikation [Kapitel: 7.8], die Interpretation der Ergebnisse [Kapitel: 3.6.2], die statistische Auswertung [Kapitel: 3.6.3] und die Datenvisualisierung [Kapitel: 3.6.4] essentiell. Weiterhin wurde Software genutzt, um die Ähnlichkeit von Proteinsequenzen zu berechnen [Kapitel: 3.6.5].

3.6.1 Proteinidentifikation mittels Datenbanksuche

Für die Proteinidentifikation wurden die MS-Ergebnisse als *mascot generic format*-Datei in das Programm Proteinscape (Version: 3.1.3.461) [Thiele et al. (2008)] geladen und mittels Mascot (Version 2.5) [Perkins et al. (1999)] identifiziert. Zusätzlich wurden die Suchmaschinen X!TANDEM [Craig and Beavis (2004)] und OMMSA [Geer et al. (2004)], welche in die MetaProteomeAnalyzer (MPA) [Muth et al. (2015a)] eingebettet sind, getestet. Die entsprechenden Suchparameter sind im Anhang 7.7 und 7.8 aufgeführt. Als Datenbank wurde NCBI Inr [Acland et al. (2013)] und UniProtKB/SwissProt [UniProt Consortium (2012)] genutzt, wobei die Versionen der jeweiligen Datenbanken den referenzierten Publikationen entnommen werden können. Darüber hinaus waren Metagenome von Biogasanlagenproben der Universität Bielefeld [Stolze et al. (2016), Hanreich et al.

(2013), Rademacher et al. (2012), Schlüter et al. (2008)] verfügbar. Um nicht-annotierte Sequenzeinträge der Metagenome mit Proteininformationen zu verknüpfen, wurden die identifizierten Proteine mittels BLAST-Suche (StandardEinstellung, *e-value*: 0,0001) [Altschul et al. (1990)] annotiert.

3.6.2 Taxonomische und funktionale Datenauswertung mit dem MetaProteomeAnalyzer

Die taxonomische und funktionelle Einordnung der identifizierten Proteine erfolgte mit der MPA Software (Version 1.03-1.09, <https://code.google.com/p/meta-proteome-analyzer/>) [SOP 12.2 siehe CD] [Muth et al. (2015a)]. Dafür wurden die Mascot Suchergebnisse als *.dat*-Datei mit einer FDR-Schwelle von 5% bzw. 1% und dem „MascotDatFile“-Parser [Helsens et al. (2007)] geladen. Anschließend wurden redundante Proteinidentifikationen mit überlappenden Peptiden oder bei Zugehörigkeit zum selben UniRef50 Cluster [Suzek et al. (2007)] gruppiert und die Taxonomie dieser auf den gemeinsamen taxonomischen Vorfahren im phylogenetischen Baum gesetzt. Für die funktionelle Zuordnung wurden die identifizierten Metaproteine basierend auf ihrer EC-Nummer [Bairoch (2000)] und ihrer KO [Kanehisa et al. (2004)] den entsprechenden KEGG-Stoffwechselwegen [Kanehisa and Goto (2000)] zugeordnet.

3.6.3 Statistische Auswertung mit Matlab

Für die statistische Auswertung der Ergebnisse wurde die Software Matlab (Version 2011a) verwendet. Die entsprechenden Matlab-Skripte können der beigelegten CD entnommen werden [Anhang CD: 7.14] und umfassen im Wesentlichen die vorgefertigten Funktionen für „Hierarchisches Clustern“, „Hauptkomponentenanalyse“, „Clustergram“, „Spearman's Rangkorrelationskoeffizient“ und „Decision Tree Learning“.

Für das „Decision Tree Learning“ wurden die Biogasanlagen vorab in Gruppen von Biogasanlagen mit hohen bzw. niedrigen Werten von jedem Prozessparameter geteilt. Anschließend versuchte der Lernalgorithmus für 60% der Biogasanlagen Metaproteine, taxonomische Ordnungen und biologische Prozesse vorzuschlagen, welche die vorherige Klassifikation erklärten. Abschließend wurde die Performanz der gelernten Marker an den restlichen 40% der Biogasanlagen getestet.

Darüber hinaus wurde die Software zur Analyse und zum Clustern von SDS-PAGEs genutzt. Dafür wurden die Gelbilder geladen, die Intensitätsprofile der Spuren in der SDS-PAGE ermittelt und Clustergramme (*unweighted pair group method with arithmetic mean*) basierend auf der Kreuzkorrelation zwischen den einzelnen SDS-PAGE Spuren erstellt. Da die Länge der einzelnen SDS-PAGE Spuren etwas variierte, wurde die Länge der Spuren

vorher auf den Beginn des Trenngels und die Lauffront normalisiert.

3.6.4 Erstellung und Analyse von Graphnetzwerken mit Gephi

Zur besseren Visualisierung der Zusammenhänge zwischen den einzelnen Prozessparametern und den mikrobiellen Gemeinschaften wurden die taxonomischen Ordnungen, biologischen Prozesse, Metaproteine und Prozessparameter in einer Matrix korreliert und als Graphnetzwerk mit Hilfe der Software Gephi (Version 0.8.2-beta) und dem „force atlas“-Algorithmus („autostab strength“: 2000, „repulsion strength“: 1000, „attraction strength“: 1, „gravity“ 100, „attraction distrib.“ „checked“) dargestellt [Bastian et al. (2009)]. Die Knoten waren dabei die einzelnen taxonomischen Ordnungen, biologischen Prozesse, Metaproteine oder Prozessparameter. Die Kanten entsprachen den einzelnen Korrelationen.

3.6.5 Multiples Sequenzalignment mit Jalview

Der Vergleich von verschiedenen Proteinsequenzen [SOP 12.1, siehe CD] wurde mit der Software Jalview (Version Jalview 2.8.0b1 (<http://www.jalview.org/>)) [Waterhouse et al. (2009)] und dem Clustal W Algorithmus [Larkin et al. (2007)] bei Standardeinstellungen ausgeführt. Als Eingabe wurden die FASTA Dateien der einzelnen Proteine genutzt.

Methodenoptimierung

Das Ziel der vorliegenden Arbeit war es, die taxonomische und funktionelle Zusammensetzung der mikrobiellen Gemeinschaften in Biogasanlagen zu untersuchen und deren Wechselwirkungen mit einzelnen Prozessparametern und Prozessproblemen aufzudecken. Für die Bearbeitung dieser Thematik wurden mikrobielle Enzyme/ Proteine (Metaproteom) untersucht, da diese die ablaufenden biochemischen Reaktionen widerspiegeln.

Essentiell für die umfassende Charakterisierung von mikrobiellen Gemeinschaften ist die Identifikation von möglichst vielen Proteinen. Nichtsdestotrotz muss die aufgewendete Arbeits- und Messzeit für die Untersuchungen adäquat sein. Dementsprechend sollte ein Workflow etabliert werden, welcher erlaubt die wesentlichen taxonomischen Gruppen und Funktionen der mikrobiellen Gemeinschaften zu erfassen (mindestens 300 Proteinidentifikationen), zugleich für Proben aus allen Biogasanlagen geeignet ist und einen hohen Probendurchsatz erlaubt. Das Ziel war die komplette Untersuchung, von Proben aus 30 Biogasanlagen in 3 Personenmonaten.

In diesem Zusammenhang wurde die Proteinseparation [Kapitel: 4.1], die LC-MS/MS Messung [Kapitel: 4.2] und die Proteinidentifikation [Kapitel: 4.3] optimiert sowie der MPA [Kapitel 4.4] für die Datenauswertung entwickelt.

4.1 Optimierung der Proteinseparation

Die Publikation von *Hanreich et al. (2012)* sowie die Diplomarbeiten von *Kohrs* und *Heyer* hatten gezeigt, dass Proteine aus Biogasanlagenproben bzw. Klärschlamm erfolgreich durch Zellaufschluss in einer Kugelmühle, Phenolextraktion sowie 2D-PAGE analysiert werden können. Allerdings variierte in diesen Arbeiten die Qualität der Proteinmuster bei der 2D-PAGE für die verschiedenen Biogasanlagen deutlich und konnte auch durch intensive Bemühungen nur geringfügig verbessert werden.

Eine zuverlässigere Trennmethode war die eindimensionale Separation der Proteine mittels SDS-PAGE. Für diese wurde die Beladungsmenge optimiert [Kapitel: 4.1.1] sowie deren Robustheit, Reproduzierbarkeit und Trennqualität überprüft [Kapitel: 4.1.2]. Da allerdings die Auflösung der SDS-PAGE kleiner als die der 2D-PAGE war, wurde zu-

sätzlich getestet, ob die Auflösung durch Kombination mit einer weiteren Trenntechnik gesteigert werden kann [Kapitel: 4.1.3].

4.1.1 Optimierung der Beladungsmenge für die SDS-PAGE

Standardmäßig wird für die SDS-PAGE eine Beladungsmenge von 20 μg Proteinextrakt pro Probe empfohlen [Bio-Rad, Laboratory Inc. (2015)]. Wird weniger geladen sind die Proteinbanden nicht zu erkennen, wird mehr geladen kann es zur Bildung von Präzipitationen und Agglomeraten kommen. Aus der Arbeit von *Hanreich et al. (2012)* war bekannt, dass die Proteinkonzentrationen in Proteinextrakten aus Biogasanlagenproben aufgrund von Probenverunreinigungen deutlich geringer waren als die Konzentrationen, welche durch entsprechende Quantifizierungsmethoden bestimmt wurden [Kapitel: 2.4.3]. Um die optimale Proteinmenge zu evaluieren, wurde eine SDS-PAGE mit 50-500 μg Phenol extrahierten Proteinen einer Biogasprobe aus Bielefeld [Anhang: 7.6] durchgeführt. Der Vergleich der entsprechenden Proteinprofile im SDS-PAGE Gelbild [Abbildung: 4.1] zeigt für alle Proteinmengen das gleiche Proteinmuster. Die Farbintensität der Proteinprofile nahm allerdings von 50 μg bis 200 μg Proteinextrakt deutlich zu. Für höhere Mengen an Proteinextrakt hingegen konnte keine weitere Intensitätszunahme beobachtet werden, sondern nur eine Zunahme eines bräunlichen Pellets beim letzten Zentrifugationsschritt vor dem Auftragen der Probe auf das SDS-Gel. Bei dem braunen Pellet handelte es sich vermutlich um verbleibende Huminstoffe.

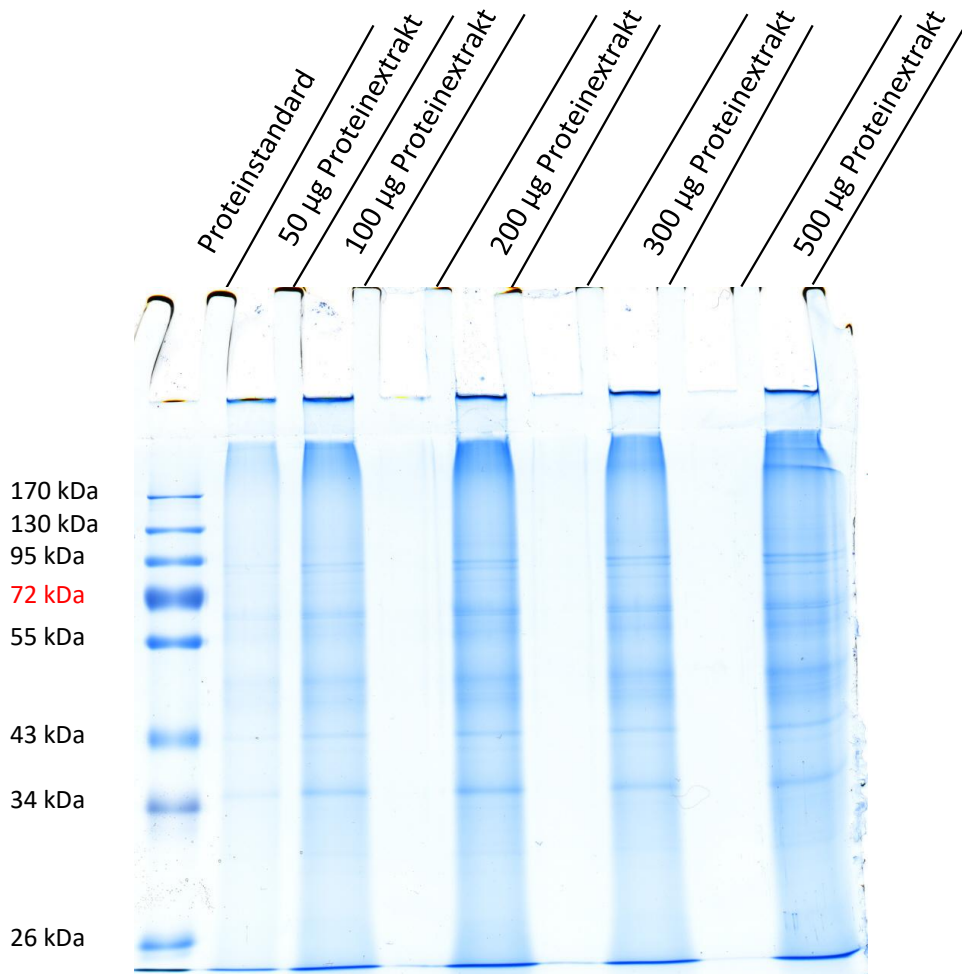


Abbildung 4.1: *SDS-PAGE* mit verschiedene Proteinmengen im Bereich von 50-500 μg . Für die Extraktion der Proteine wurden die Zellen in einer Kugelmühle lysiert und parallel mit Phenol extrahiert.

4.1.2 Robustheit, Reproduzierbarkeit und Trennqualität der SDS-PAGE

¹Um die Robustheit und Reproduzierbarkeit der SDS-PAGE basierten Proteinseparation von Biogasanlagenproben zu testen, wurden sechs verschiedene Biogasanlagen (BGP 3, BGP 4, BGP 5, BGP 7, BGP 8, BGP 9) über mehrere Zeitpunkte untersucht. Für alle Biogasanlagen [Anhang: 7.6] konnten über alle untersuchte Zeitpunkte erfolgreich Proteine extrahiert werden. Exemplarisch finden sich in Abbildung 4.2 die SDS-PAGE Profile und die Clustergramme aller sechs Biogasanlagen sowie elf Zeitpunkte der stabil operierenden Biogasanlagen BGP 3 und BGP 4. Die Menge an Proteinextrakt betrug 100 μg .

Insgesamt waren die SDS-PAGE Muster sämtlicher Proben sehr ähnlich. Sie wiesen einen intensiven blauen Hintergrund auf, welcher von einigen dominierenden Banden überlagert

¹Basiert teilweise auf der Publikation Heyer et al. (2013)

wurde. Eine Unterscheidung der verschiedenen Biogasanlagen aufgrund dieser Bandenmuster war möglich, was durch anschließende Clusteranalyse der SDS-PAGE Spuren weiter bestätigt wurde. Zusätzlich zeigten die Clusterbilder, dass neben den verschiedenen Zeitpunkten (BGP 3, BGP 4) die technischen Replikate (BGP 4_1a, BGP 4_1b) die größte Ähnlichkeit zueinander hatten.

Um zusätzlich eine Aussage über die Trennqualität zu treffen, wurde nochmals 500 μg der Probe BGP 3_5 mittels SDS-PAGE separiert, in zehn Fraktionen geteilt und alle Fraktionen mittels Elite Hybrid Ion Trap-Orbitrap MS/MS analysiert. Die entsprechende Identifizierung erfolgt dann mit Mascot-Suche gegen UniProtKB/SwissProt (Peptide Score 15).

Während der Großteil der identifizierten Proteine nur in maximal drei benachbarten Fraktionen vorkamen [Abbildung: 4.3] gab es einige hochabundante Proteine wie z. B. das Protein *Methyl-Coenzyme M Reduktase* (P07964, 30 Peptide) welche zum Teil in allen Fraktionen gefunden wurde.

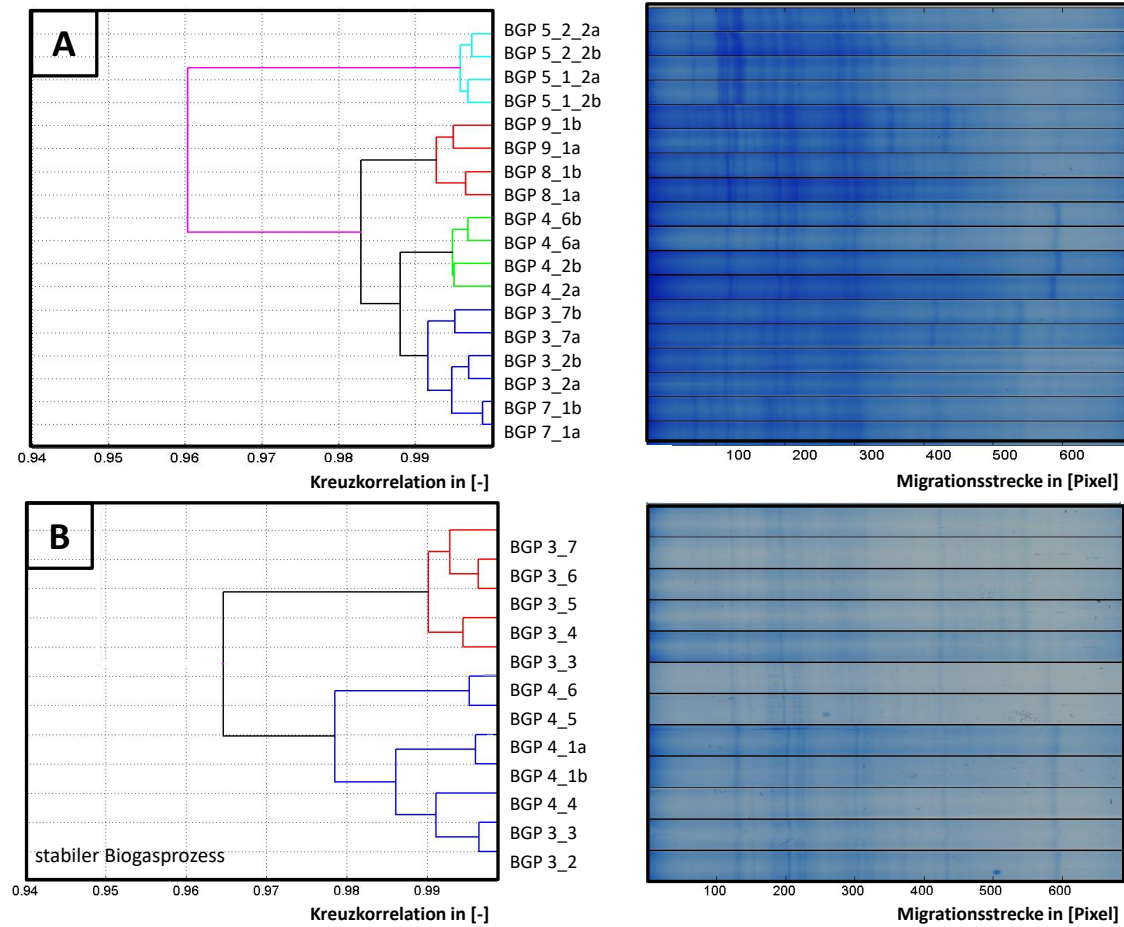


Abbildung 4.2: *Robustheit und Reproduzierbarkeit von SDS-PAGEs aus Biogasanlagenproben*, untersucht am Beispiel von sechs verschiedenen Biogasanlagen (A) sowie an verschiedenen Zeitpunkten zweier Biogasanlagen (B) mit jeweils $100 \mu\text{g}$ Proteinextrakt. Dafür sind rechts die SDS-PAGE Spuren horizontal abgebildet und links das Clustering basierend auf der Kreuzkorrelation der SDS-PAGE Profile [Kapitel: 3.6.3]. Für die bessere Visualisierung der SDS-PAGE Spuren wurde die Helligkeit um 40% reduziert und die Kontraste um 40% erhöht.

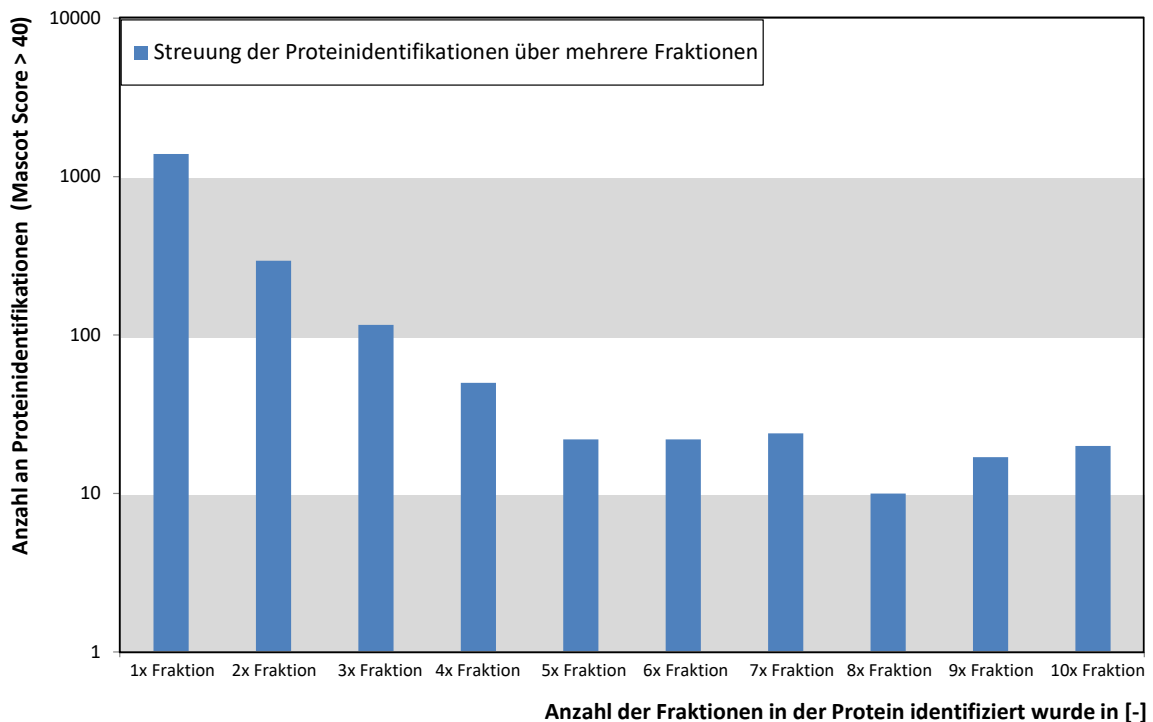


Abbildung 4.3: *Streuung der Proteinidentifikationen über die Anzahl an Fraktionen der SDS-PAGE*

4.1.3 Einfluss der Trenndimensionen auf die Anzahl an identifizierten Peptiden und Proteinen

²Die Kombination verschiedener Trenntechniken ermöglicht eine Separation der Proteine nach verschiedenen biochemischen Eigenschaften. Dadurch kann die Komplexität einer Probe effektiver als mit einer Trenntechnik reduziert und damit die Auflösung erhöht werden. Im Umkehrschluss multipliziert sich durch die Kombination der Trenntechniken aber auch die Gesamtzahl an zu untersuchenden Fraktionen. Daher muss die Separationsstrategie immer an die Probe, die Fragestellung und die verfügbare Messzeit angepasst werden.

Innerhalb der Arbeitsgruppe etablierte Separationsmethoden waren die SDS-PAGE und die flüssig-IEF für Proteine sowie die Umkehrphasenchromatographie für tryptische Peptide. Ziel dieses Experimentes war die Untersuchung des Nutzen verschiedener Trenndimensionen auf die Anzahl an identifizierten Peptiden und Proteinen, um die Planung neuer Experimente zu erleichtern.

²Beruhet auf der Publikation Kohrs et al. (2013).

Als Proben dienten die Proteinextrakte aus einer mesophilen und einer thermophilen Biogasanlage [Anhang: 7.6 BGP 5, BGP 4]. Für den eindimensionalen Ansatz (1D-Ansatz) wurden 100 μg Proteinextrakt in eine SDS-PAGE einlaufen gelassen und als eine Fraktion dreimal mittels Umkehrphasenchromatographie und 120 min Gradient analysiert. Das Einlaufen in die SDS-PAGE diente dazu, verbleibende Verunreinigungen über das Polyacrylamidgel zu entfernen und um eine bessere Vergleichbarkeit der verschiedenen Ansätze zu gewährleisten. Für den 2D-Ansatz wurden die 100 μg Proteinextrakt vor der Umkehrphasenchromatographie mittels SDS-PAGE in 10 Fraktionen vorsepariert. Beim 3D-Ansatz wurden 1000 μg Proteinextrakt zusätzlich vor der SDS-PAGE und Umkehrphasenchromatographie durch flüssig-IEF (pH 4-7) in 12 Fraktionen vorsepariert, wodurch letztendlich 120 Fraktionen entstanden sind. Die Identifizierung der Proteine erfolgte durch ein AmaZon ETD ESI-Ion Trap MS/MS und durch Datenbanksuche mit Mascot gegen UniProtKB/SwissProt und einer FDR von 5%.

Beim Vergleich der identifizierten Peptide und Proteine [Abbildung: 4.4 A + B] konnten für den 1D-Ansatz bis zu 74 Peptide bzw. 30 Proteine, für die Triplikatmessung bis zu 118 Peptide bzw. 48 Proteine, für den 2D-Ansatz bis zu 785 Peptide bzw. 383 Proteine und für den 3D-Ansatz bis zu 3452 Peptide bzw. 1663 Proteine identifiziert werden. Interessanterweise wurden von den 540 bzw. 785 die mit 2D-Ansatz gefunden wurden 172 bzw. 422 nicht mit dem höher auflösenden 3D-Ansatz gefunden [Abbildung: 4.4 C]. Ferner zeigten die Triplikate des 1D-Ansatzes relativ große Unterschiede bei den identifizierten Peptiden und Proteinen, welche im Kapitel [Kapitel: 4.2.4] ausführlicher betrachtet werden.

4.1 Optimierung der Proteinseparation

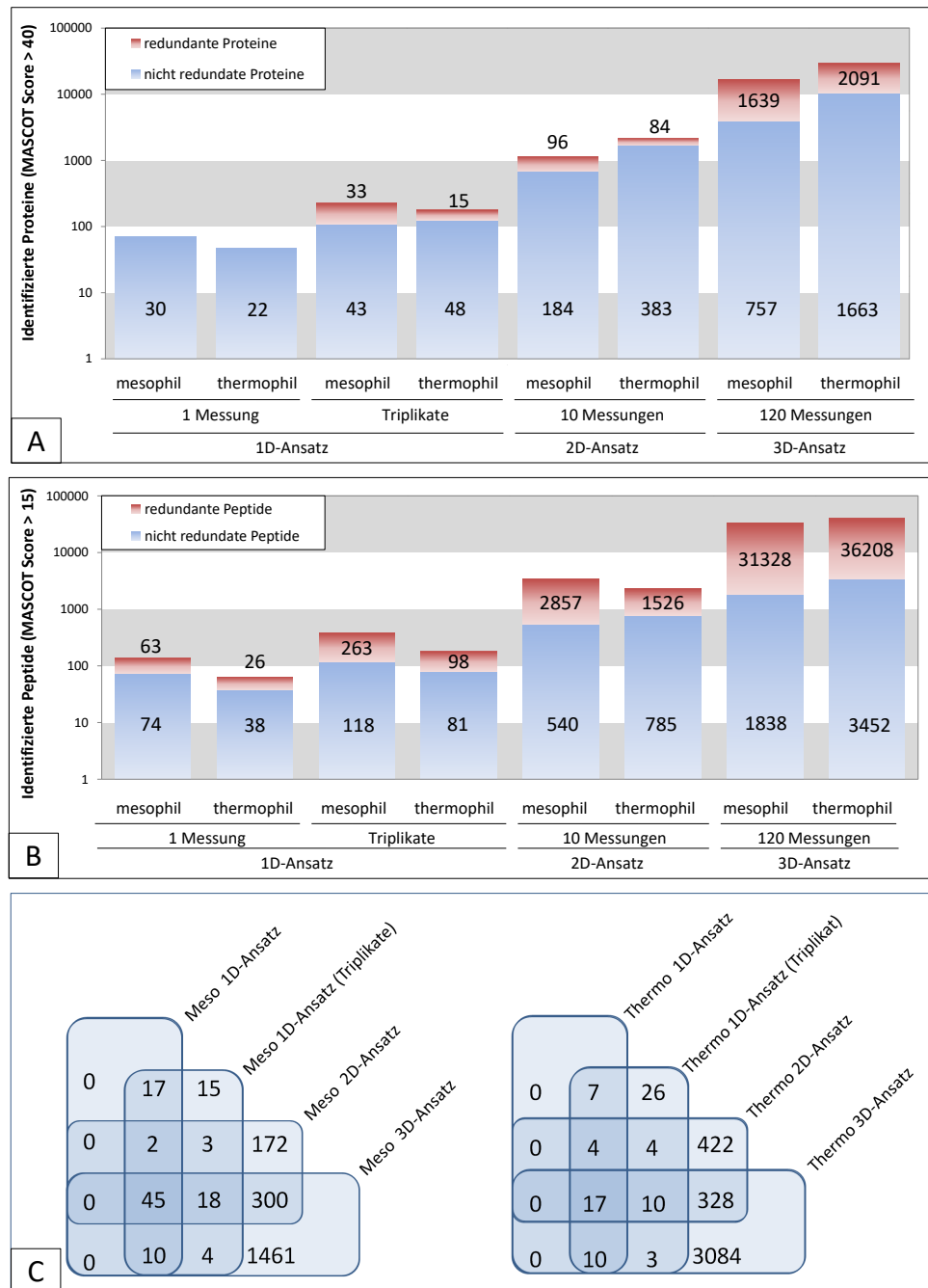


Abbildung 4.4: Identifizierte Proteine (A) und Peptide (B) einer mesophilen und einer thermophilen Biogasanlage (BGP 4, BGP 5) mit verschiedener Anzahl an Trennstufen.. C zeigt die Überschneidungen der Peptididentifikationen. 1D-Ansatz: 100 μ g Proteinextrakt in eine SDS-PAGE einlaufen gelassen und mittels LC-MS/MS-Messung (AmaZon ETD ESI-Ion Trap MS/MS) und 120 min Gradient analysiert. 2D-Ansatz 100 μ g Proteinextrakt vor der LC-MS/MS-Messung mittels SDS-PAGE in 10 Fraktionen vorsepariert. 3D-Ansatz: 1000 μ g Proteinextrakt zusätzlich vor der SDS-PAGE und LC-MS/MS-Messung durch flüssig-IEF (pH 4-7) fraktioniert. Die Proteinidentifizierung erfolgte mit Mascot gegen UniProtKB/SwissProt (FDR von 5%).

4.1.4 Diskussion

Proteinextrakte aus allen untersuchten Biogasanlagenproben konnten zuverlässig und reproduzierbar mittels SDS-PAGE charakterisiert werden [Ehlers and Cloete (1999), Schneider et al. (2012)]. Beim 2D-Ansatz konnte mit 383 identifizierten Proteinen im Vergleich zu etwa 300 Proteinspots bei der ursprünglichen 2D-PAGE [Hanreich et al. (2013)] eine ähnlich hohe Anzahl an Identifikationen erreicht werden. Der 2D-Ansatz benötigte aber deutlich weniger Arbeitsaufwand und MS-Messzeit als die 2D-PAGE. Darüber hinaus kann die Anzahl an Identifikationen beim 2D-Ansatzes über die Anzahl der Fraktionen und Kombination mittels flüssig-IEF (3D-Ansatz) noch weiter erhöht werden [Lü et al. (2014)]. Im Vergleich zum 1D-Ansatz konnte mit dem 2D-Ansatz durch die verzehnfacht der Fraktionszahl auch die Anzahl an Peptididentifikationen verzehnfacht werden. Im Gegensatz dazu wurde durch die zusätzlich Verzwölfachung der Fraktionsanzahl durch den 3D-Ansatz die Anzahl der Peptididentifikationen des 2D-Ansatzes nur noch verdreifacht. Darüberhinaus gingen beim 3D-Ansatz einige Proteine verloren, weil deren pI außerhalb des untersuchten pI-Bereichs lag.

Neben der Separation der Proteine bzw. Peptide in die einzelnen Fraktionen ist die Trennqualität entscheidend für die Anzahl der identifizierten Proteine. Wie die Bandenmuster in den SDS-PAGE Gelen [Abbildung: 4.1, 4.2] zeigten, konnten die Proteine getrennt werden. Allerdings konnten sehr abundante Proteine auch in mehrere SDS-PAGE Fraktionen gefunden werden [Abbildung: 4.3]. Passend dazu war der Hintergrund der SDS-PAGEs deutlich dunkler als bei SDS-PAGEs von Reinkulturen (Daten nicht gezeigt). Eine Erklärung für die Streuung der Proteine über mehrere Fraktionen und den dunkleren Hintergrund könnten verbliebene Huminstoffe in den Proteinextrakten sein [Bastida et al. (2009)]. Huminstoffe bilden in Umweltproben mit Proteinen feste Komplexe, welche mit gängigen chromatographischen oder proteinbiochemischen Methoden nicht trennbar sind und zum Verschmieren der Proteine in der SDS-PAGE führen [Stiehm (2011)]. Passend dazu benötigte die SDS-PAGE von Biogasanlagenproben eine zehnmals größere Menge an Proteinextrakt als bei Reinkulturproben und durch das Zentrifugieren der Proben nach dem Aufnehmen in SDS-PAGE Probenpuffer entstand ein bräunliches Pellet.

Um für zukünftige Studien die Anzahl an identifizierten Proteinen noch weiter zu erhöhen, könnten zusätzliche Trennmethoden, wie z. B. die Umkehrphasenchromatographie von Proteinen [Eschelbach and Jorgenson (2006)], ergänzt oder eine mehrdimensionale Peptidchromatographie [Washburn et al. (2001)] etabliert werden. Im Sinne einer Routineuntersuchung von Biogasanlagen wäre aber die Reduzierung der gesamten Analysenzeit von über 1 Woche auf unter 24 h dringlicher. Voraussetzung dafür wäre, die zeitaufwendige

Gel-basierte Proteinseparation und den In-Gel Verdau zu umgehen. Stattdessen könnten die Proteine direkt mit einem schnelleren FASP-Verdau [Wiśniewski et al. (2009)] zu Peptiden verdaut und mittels LC-MS/MS analysiert werden. Allerdings sollte dann die Gradientenlänge bei der Umkehrgradientenchromatographie von zwei auf vier oder mehr Stunden erhöht werden, um trotzdem die Probenkomplexität ausreichend zu reduzieren.

Zusammenfassung: Durch die Trennung mittels SDS-PAGE beim 2D-Ansatz konnten die Proteine aus Biogasanlagenproben erfolgreich und reproduzierbar separiert werden. Die Anzahl an identifizierten Proteinen beim 2D-Ansatz war adäquat zur 2D-PAGE basierten Untersuchung und konnte durch Kombination mittels flüssig-IEF beim 3D-Ansatz noch weiter gesteigert werden.

4.2 Optimierung der LC-MS/MS Messung

Für die Anzahl an identifizierten Peptiden und Proteinen ist neben der Proteinseparation, der erfolgreiche tryptische Verdau der Proteine zu Peptiden [Kapitel: 4.2.1], das zur Verfügung stehende LC-MS/MS System [Kapitel: 4.2.2], die optimale Beladungsmenge für die LC-MS/MS Messung [Kapitel: 4.2.3] sowie die Reproduzierbarkeit der LC-MS/MS Messung [Kapitel: 4.2.4] entscheidend.

4.2.1 Überprüfung des tryptischen Verdau

Der vorherige Verdau der Proteine zu Peptiden war essentiell für die Proteinidentifikation mittels LC-MS/MS [Kapitel: 2.4.6]. Standardmäßig wird für den Verdau von in-Gel separierten Proteinen das Enzym Trypsin und das Protokoll von *Shevchenko et al. (1996)* [Shevchenko et al. (1996)] genutzt. Wie in Kapitel 4.1.4 angesprochen enthalten Proteinextrakten aus Biogasanlagenproben noch Verunreinigungen wie Huminstoffe. Diese Verunreinigungen könnten das Trypsin inhibieren, wodurch vermehrt unvollständig verdaute Peptide (*missed cleavages*) auftreten könnten.

Für die Überprüfung dieser Problematik wurden 100 μg , 200 μg und 300 μg Proteinextrakt aus einer Biogasanlage [Anhang: 7.6 BGP 5] mit dem empfohlenen Protein-zu-Trypsinverhältnis von 1:50 und dem Verhältnis 1:250 verdaut. Anschließend wurde die Probe in 12 μL Lösungsmittel aufgenommen und je 1 μg als Triplikat mit einem Elite Hybrid Ion Trap-Orbitrap MS/MS und einem 120 min Gradienten vermessen.

Die Ergebnisse der Proteinidentifikation zeigten [Abbildung: 4.5], dass die Anzahl an identifizierten Spektren bei beiden Trypsinverhältnissen und allen drei Proteinmengen annähernd gleich war. Allerdings konnte für alle untersuchten Protein-zu-Trypsinverhältnis eine geringe Zunahme an identifizierten Spektren beobachtet werden, sobald ein *missed cleavage* zugelassen wurde.

4.2.2 Vergleich verfügbarer Massenspektrometer

Je sensitiver und hochauflösender ein Massenspektrometer ist, umso besser ist die Qualität der gemessenen Spektren und folglich auch die Anzahl der Peptid- und Proteinidentifikationen. Im Verlauf dieser Arbeit konnte neben dem AmaZon ETD ESI-Ion Trap MS/MS ein hochauflösendes Elite Hybrid Ion Trap-Orbitrap MS/MS genutzt werden. Um den Mehrwert dieses Gerätes zu testen, wurden 100 μg einer Biogasanlagenprobe mittels SDS-PAGE in 10 Fraktionen geteilt, jede Fraktion in 12 μL Lösungsmittel aufgenommen und je 5 μL in die AmaZon ETD ESI-Ion Trap MS/MS und je 1 μL in die Elite Hybrid Ion Trap-Orbitrap MS injiziert. Die LC-Separation erfolgte mit nahezu identischen

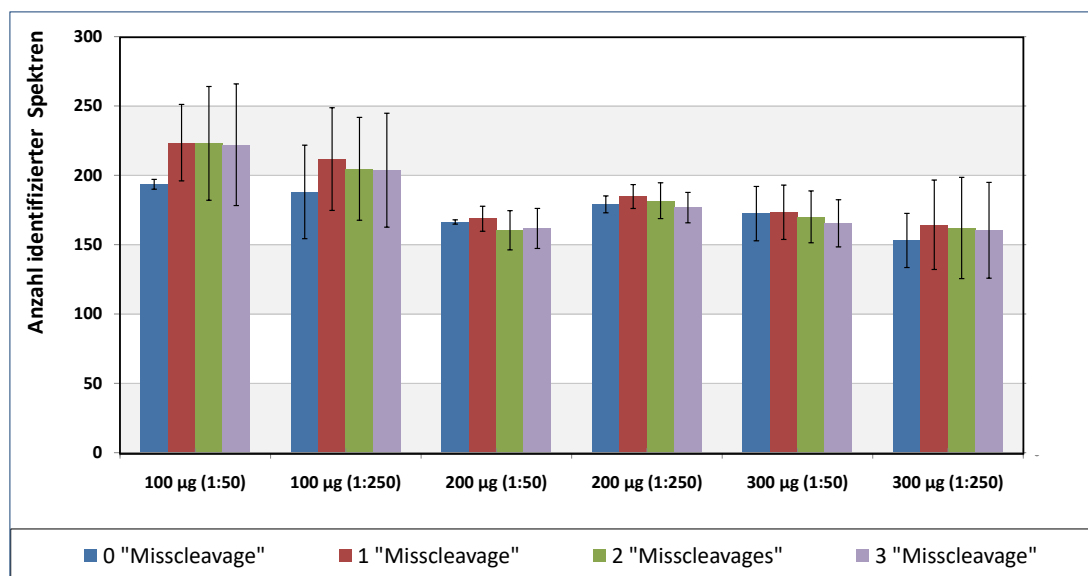


Abbildung 4.5: *Einfluss der Trypsinmenge auf die Anzahl an identifizierten Spektren und missed cleavages*

UltiMate 3000 RSLCnano LC-Systemen (eines „gesplitted“/ eines „ungesplitted“) und einem 120 min Gradienten für die AmaZon ETD ESI-Ion Trap MS/MS bzw. einem 30 min Gradienten für Elite Hybrid Ion Trap-Orbitrap MS/MS.

Während mit dem AmaZon ETD ESI-Ion Trap MS/MS eine Gesamtzahl von 255 nicht-redundanten Peptiden identifiziert wurde, waren es mit dem Elite Hybrid Ion Trap-Orbitrap MS/MS eine Gesamtzahl von 917 nicht-redundanten Peptiden. Trotz der geringeren Probenmenge und dem kürzeren Gradienten war das Elite Hybrid Ion Trap-Orbitrap MS/MS drei- bis viermal besser. Infolgedessen wurde dieses für alle folgenden Messungen genutzt und die weitere Optimierung der LC-MS/MS Methoden auch nur für dieses Gerät durchgeführt.

4.2.3 Optimierung von LC-MS/MS Messungen

Sowohl die Gradientenlänge als auch die Probenbeladung einer LC-MS/MS-Messung beeinflussen entscheidend die Anzahl an identifizierten Spektren, Peptiden und Proteinen. Die Beladung mit zu vielen Peptiden kann zur Überladung der Säulen führen, wodurch die chromatographische Auftrennung gestört würde oder auch Peptide verloren gehen. Bei zu geringer Probenmenge fallen niederabundante Peptide unter die Sensitivitätsgrenze des MS und werden nicht identifiziert. Zusätzlich sollte die Beladungsmenge jeweils an die Probenkomplexität und die Gradientenlänge angepasst werden.

Am Beispiel der Probe BGP 13 [Anhang: 7.5] wurden daher verschiedene Beladungsmen-

gen für drei gängige Separationsstrategien getestet. Zuerst wurden 26,67 μg , 53,33 μg und 80 μg Proteinextrakt in ein SDS-Gel (1D-Ansatz) aufgegeben und mittels 120 min Gradient analysiert. Zusätzlich wurden zwischen 1,6 μg und 133 μg Proteinextrakt mittels SDS-PAGE in zehn Fraktionen vorsepariert (2D-Ansatz) und zweitens mittels eines 30 min sowie drittens mittels eines 120 min Gradienten vermessen. Die MS-Messung erfolgte an der Elite Hybrid Ion Trap-Orbitrap MS/MS und die Proteinidentifikation wurde mit Mascot gegen UniProtKB/SwissProt vorgenommen (Peptide Score 15, Proteinscore 40). Während für die in das Gel eingelaufenen Proben bis zu 259 Spektren identifiziert wurden, konnte durch die vorherige SDS-PAGE-Trennung die Anzahl weiter auf 2745 Spektren (30 min Gradienten) bzw. auf bis zu 7971 Spektren (120 min Gradient) gesteigert werden [Abbildung: 4.6]. Eine Erhöhung der Proteinmenge führte für die eingelaufenen Proben zu keiner Erhöhung der Anzahl an identifizierten Spektren. Im Gegensatz dazu konnte für beide Ansätze mit SDS-PAGE Vortrennung eine Zunahme der identifizierten Spektren durch die Erhöhung der Proteinmenge von 1,6 μg auf bis zu 133 μg bis zum Fünffachen beobachtet werden. Erst ab Proteinmengen von über 50 μg setzte für diese eine Sättigung der Spektrenanzahl ein. Parallel dazu konnte für größere Probenmengen als 50 μg auch eine Verschleppung von Proteinen aus einer LC-MS/MS-Messung in die nächste LC-MS/MS-Messung beobachtet werden (Daten nicht gezeigt).

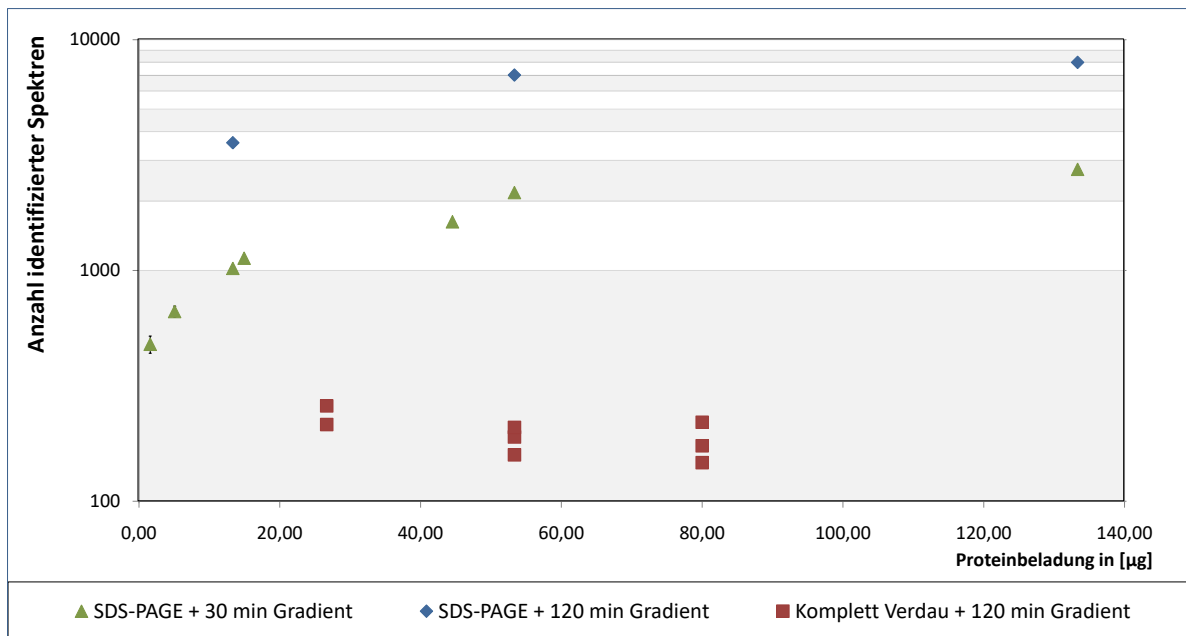


Abbildung 4.6: *Identifizierte Spektren bei verschiedener Probenbeladung und Gradientenlänge.* Die dargestellten Spektrenanzahlen und Proteinmengen beziehen sich für die SDS-PAGE Vortrennung auf die Gesamtanzahl über alle Fraktionen.

4.2.4 Reproduzierbarkeit von Orbitrap Elite Messungen

Eine weitere wesentliche Voraussetzung für die erfolgreiche Untersuchung von mikrobiellen Gemeinschaften aus Biogasanlagen ist die Reproduzierbarkeit der LC-MS/MS-Messung. Da die Separation von Biogasanlagenproben mittels SDS-PAGE in zehn Fraktionen und anschließender Messung aller zehn Fraktionen mittels 120 min Gradient und Elite Hybrid Ion Trap-Orbitrap MS/MS die höchste Anzahl an Peptididentifikationen ergab [Kapitel: 4.2.3], wurde die Reproduzierbarkeit der LC-MS/MS-Messungen für diesen Ansatz untersucht. Für diese Untersuchung wurden zwei Proben der stabil arbeitenden Biogasanlage BGP 9 [Anhang: 7.5] im Abstand von zwei Wochen genutzt. Die Proteine aus diesen Proben wurden je zweimal extrahiert und 13,3 μg Proteinextrakt je zweimal massenspektrometrisch vermessen.

Die Reproduzierbarkeit der identifizierten Peptide bzw. Proteine bezogen auf die Gesamtmenge an Identifizierungen betrug für die reine MS/MS Messung etwa 65,5% bzw. 62,0% (A + B), für die Extraktion 54,3% bzw. 53,6% (1 + 2) und für die Probennahmen 51,8% bzw. 51,8% (T07 + T09) [Tabelle: 4.1]. Um die genauen Ursachen für die geringe Reproduzierbarkeit von zwei LC-Läufen von derselben Probe zu ergründen, wurde die Identifikation des Proteins *Acetyl-CoA Decarboxylase/Synthase Komplex (Q49163)* intensiver für die erste Fraktion der SDS-PAGE analysiert [Abbildung: 4.7]. In der ersten Messung wurden 16 Peptide dieses Proteins identifiziert und in der zweiten 18. Von den insgesamt identifizierten 22 Peptiden waren 12 Peptide gleich. Allerdings konnten die Peptidionen für die fehlenden Peptididentifikationen in beiden LC-MS/MS Messungen beobachtet werden. Diese waren aber entweder nicht vom MS für die Fragmentierung ausgewählt oder die Identifikation waren nicht erfolgreich.

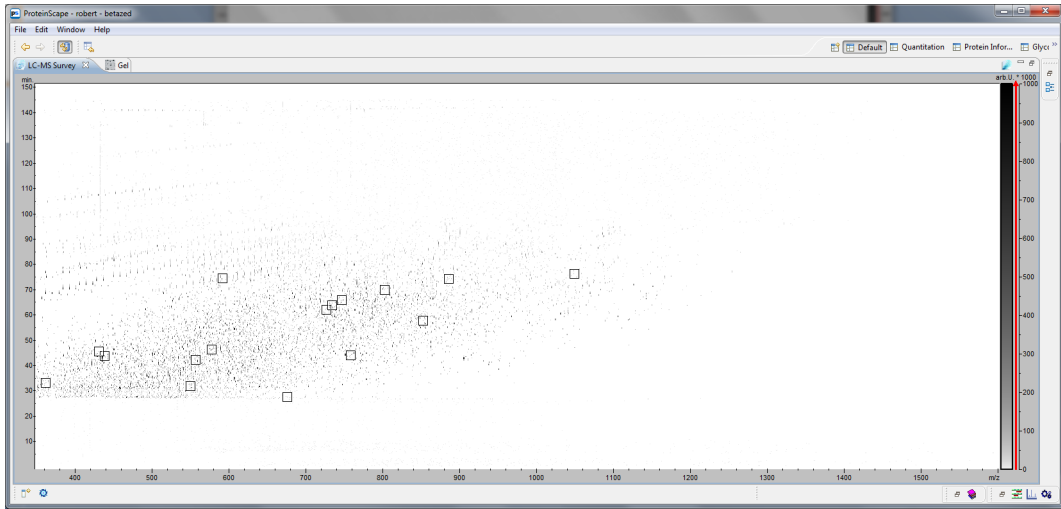
4.2 Optimierung der LC-MS/MS Messung

Probe

Identifizierte Peptidionen der LC-MS/MS Messung

T07_Ext1_A_Fraktion 1

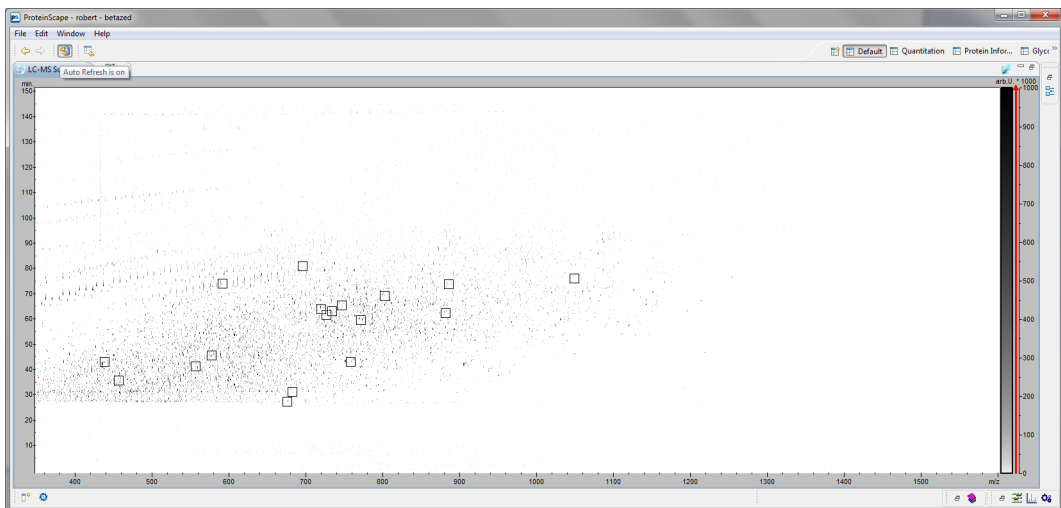
Retentionszeit in [min]



m/z in [Da]

T07_Ext1_B_Bande1

Retentionszeit in [min]



m/z in [Da]

Abbildung 4.7: *Reproduzierbarkeit von LC-MS/MS Messungen*, untersucht am Protein Acetyl-CoA Decarboxylase/ Synthase (Q49163). Die Darstellung zeigt alle gemessenen Peptidionen über den Verlauf der LC-MS/MS Messung sowie ihre Masse-Ladungsverhältnisse. Identifizierte Peptide des untersuchten Proteins sind durch eckige Kästchen hervorgehoben.

Tabelle 4.1: *Reproduzierbarkeit der SDS-PAGE LC-MS/MS Methode* bei zehn SDS-PAGE Fraktionen (120 min LC Gradienten, Elite Hybrid Ion Trap-Orbitrap MS/MS). Es wurden jeweils technische Replikate der LC-MS/MS Messung (A, B) und der gesamten Probenaufarbeitung (1,2) sowie zwei separate Zeitpunkte (T07, T09) als „biologische Replikate“ untersucht.

| Messung | Peptide/ Proteine exklusiv erste Messung | Peptide/ Proteine Schnittmenge | Peptide/ Proteine exklusiv zweite Messung | Peptide/ Proteine Reproduzierbarkeit in [%] |
|--|--|-----------------------------------|---|---|
| Technische Replikate MS/MS | | | | |
| T07_1A vs. 1B | 1207/ 651 | 2097/ 1054 | 1175/ 630 | 63,8/ 62,2 |
| T07_2A vs. 1B | 1821/ 765 | 4357/ 1241 | 1676/ 731 | 71,4/ 62,4 |
| T09_1A vs. 1B | 1237/ 672 | 2056/ 1048 | 1146/ 626 | 63,3/ 61,8 |
| T09_2A vs. B | 1426/ 779 | 2363/ 1196 | 1312/ 712 | 63,4/ 61,7 |
| Technische Replikate Extraktion | | | | |
| T09_1A vs. 2A | 1346/ 721 | 1947/ 1000 | 1842/ 976 | 55,3/ 54,4 |
| T09_1B vs. 2B | 1348/ 716 | 1854/ 958 | 1821/ 950 | 54,2/ 53,7 |
| T09_1A vs. 2B | 1439/ 762 | 1854/ 958 | 1821/ 950 | 53,4/ 53,0 |
| T09_1B vs. 2A | 1323/ 709 | 1879/ 965 | 1919/ 1010 | 54,1/ 53,3 |
| Biologische Replikate Probennahme | | | | |
| T07_1A vs. T09_1A | 1604/ 830 | 1700/ 875 | 1593/ 845 | 51,6/ 51,1 |
| T07_1B vs. T09_1B | 1563/ 808 | 1709/ 876 | 1493/ 798 | 52,8/ 52,2 |
| T07_1A+B vs. T09_1A+1B | 2210/ 1120 | 2269/ 1215 | 2170/1131 | 50,9/ 51,9 |

4.2.5 Diskussion

Wie die bis zu viermal höhere Anzahl an identifizierten Peptiden bei der Messung mit dem Elite Hybrid Ion Trap-Orbitrap MS/MS im Vergleich zum AmaZon ETD ESI-Ion Trap MS/MS zeigte [Kapitel: 4.2.2], ist die Auflösung des MS am wichtigsten für die Qualität der LC-MS/MS-Messung. Mit noch besseren MS, wie z. B. Fouriertransformations-MS/MS [Marshall et al. (1998), Scigelova et al. (2011)] könnte die Anzahl an Peptidentifikationen noch weiter gesteigert werden. Alternative könnten auch Ionen-Mobilitäts-MS/MS [Silveira et al. (2014)] genutzt werden. Letztere trennen die Peptidionen zusätzlich nach ihrer Mobilität auf, wodurch die Probenkomplexität reduziert und die Anzahl an Identifikationen gesteigert werden könnte.

Ungeachtet des rasanten technischen Fortschritts hat die LC-MS/MS basierte Messung eine Schwachstelle: die geringe Schnittmenge der identifizierten Peptide und Proteine zwischen zwei LC Läufen. Für das untersuchte Beispiel [Kapitel: 4.2.4] betrug sie analog zu entsprechenden anderen Studien [Tabb et al. (2010), Wilkins et al. (2006)] nur zwischen 50-6%. Infolgedessen können Proteine mit wenigen identifizierten Peptiden schlecht verglichen werden bzw. es wird eine große Anzahl an Wiederholungsmessungen benötigt. So müssten bei 50% Schnittmenge bei den identifizierten Peptiden zwei, zehn, bzw. fünfzehn Messungen durchgeführt werden, um 75%, 95% bzw. nahezu 100% aller messbaren Peptide zu identifizieren [Wilkins et al. (2006)]. Eine Ursache für die begrenzte Reproduzierbarkeit der LC-MS/MS-Messungen sind kleine Unterschiede bei der LC oder bei der Peptidionisation, welche zur Auswahl verschiedener Peptidionen für die Fragmentierung führen [Tabb et al. (2010)]. Eine andere Ursache ist das Scheitern der Peptidentifikation aufgrund von Qualitätsunterschieden bei der CID Fragmentierung [Berg et al. (2006)].

Trotz des Scheiterns der Peptidentifikation waren aber die Signale der Peptidionen vorhanden [Abbildung: 4.7]. Anstelle die Proteinhäufigkeiten aus der Anzahl an identifizierten Peptiden und Spektren zu ermitteln, ist daher die Quantifizierung aus den Flächen der Peptidionen genauer und reproduzierbarer [Neilson et al. (2011), Otto et al. (2014)]. Allerdings wird dafür spezielle Software wie z. B. Progenesis QI (NonLinear, Newcastle, England) benötigt, um die Peptidflächen mit den Peptidentifikationen zu verknüpfen. Auf der anderen Seite kann bei dieser Strategie auch die Dauer der MS-Messung durch Messmethoden wie *data-independent acquisition* [Egertson et al. (2013)] oder *Swath-MS* [Schubert et al. (2015)] reduziert werden.

Neben der Qualität des MS wird die Anzahl an Identifikationen durch die Probenkomplexität beeinflusst. Bei zu großer Probenkomplexität kann es zur Überlagerung der Fragmentationsspektren und dadurch zum Scheitern der Peptidentifikationen. Folglich ist

analog zur Separation der Proteine mittels SDS-PAGE [Kapitel: 3.3] auch die Trennung der Peptide mittels LC wichtig. So ermöglichte es die Verlängerung des LC Gradienten von 30 min auf 120 min, dreimal mehr Spektren zu identifizieren. Vermutlich kann die Anzahl an identifizierten Spektren durch den Einsatz von 50 cm anstatt 25 cm Säulen und Gradientenlängen von 240 min sogar noch weiter erhöht werden.

Parallel zur Anpassung der Separationsstrategie war es zusätzlich nötig, die Probenmenge zu optimieren [Kapitel: 4.2.3]. Überraschenderweise war eine Steigerung der Peptidmenge für die LC-MS/MS-Messung über die Empfehlung von 2-4 μg Peptid hinaus sinnvoll [Michalski et al. (2011), Nagaraj et al. (2012)]. Eine Erklärung könnte die Überschätzung der Proteinmenge bei der Proteinquantifizierung [Kapitel: 4.1.1] und weitere Verluste während der SDS-PAGE oder dem tryptischen Verdau gewesen sein. Daher sollte die Peptidmenge vor der LC-MS/MS Messung z. B. mit microBCA [Smith et al. (1985)], Aminosäureanalyse [Rutherford and Gilani (2009)], Tricin-SDS-PAGE [Schägger (2006)] oder NanoDrop [Desjardins et al. (2009)] überprüft werden.

Eine untergeordnete Rolle bei der Optimierung der LC-MS/MS Messungen spielte der tryptische Verdau. Dieser funktionierte für die untersuchten Proben einwandfrei [Kapitel: 4.2.1]. Nur vereinzelt traten nicht komplett verdaute Proteine (*missed cleavages*) auf. Für diese Fälle von nicht komplett verdauten Proteinen war jedoch bekannt, dass Trypsin durch die sterische Behinderung nicht nach C-terminalem Prolin schneiden kann [Siepen et al. (2007)]. Um in Zukunft den Arbeitsaufwand für den tryptischen Verdau zu reduzieren, wäre es noch möglich den In-Gel-Verdau durch einen schnelleren filterbasierten In-Lösungsverdau [Wiśniewski et al. (2009)] zu ersetzen.

Zusammenfassung: Die Anzahl an identifizierten Peptiden und Proteinen konnte durch die Nutzung eines hochauflösenden Elite Hybrid Ion Trap-Orbitrap MS/MS sowie die richtige Probenbeladung und eine möglichst umfassende Fraktionierung der Probe um ein Vielfaches im Vergleich zum ursprünglichen Workflow gesteigert werden.

4.3 Optimierung der Proteinidentifikation

Die Identifikation der Proteine erfolgte durch Suche der MS-Spektren gegen Proteindatenbanken mittels Datenbanksuchalgorithmen. Im Wesentlichen wurde die Anzahl und die Qualität der Proteinidentifikationen durch die Auswahl der Proteindatenbanken [Kapitel: 4.3.1] und der Datenbanksuchalgorithmen [Kapitel: 4.3.3] sowie die Definition von validen Schwellwerten für Proteinidentifikationen [Kapitel: 4.3.2] beeinflusst. Ferner mussten bei der Datenbanksuche mögliche PTMs beachtet werden [Kapitel: 4.3.4]. Eine andere Möglichkeit, die Anzahl der Proteinidentifikationen zu erhöhen, war die *de novo* Sequenzierung und anschließende BLAST Suche [Kapitel: 4.3.5].

4.3.1 Auswahl von Proteindatenbanken für die Proteinidentifikation

Die benutzten Datenbanksuchalgorithmen ermöglichen nur die Identifikation von Proteinen, welche in der entsprechenden Datenbank vorhanden sind. Prinzipiell ist die Suche gegen Datenbanken mit nahezu allen bekannten Proteinen möglich, aber dadurch werden wiederum die Suchzeit und die Chance für Zufallstreffer deutlich erhöht. Dementsprechend war die Auswahl einer zur Probe passenden Datenbank entscheidend für die Anzahl und Qualität der Proteinidentifikationen.

Im Folgenden wurden die Datenbanken UniProtKB/Swiss-Prot, UniProtKB/TrEMBL und NCBIInr, sowie eine Mischdatenbank aus mehreren Biogasmetagengenomen [Schlüter et al. (2008), Hanreich et al. (2013), Rademacher et al. (2012), Stolze et al. (2016)] und UniProtKB/Swiss-Prot [Tabelle: 4.2 auf ihrer Eignung für Biogasanlagenproben getestet. Als Probe diente die eingelaufene und mittels Elite Hybrid Ion Trap-Orbitrap MS/MS vermessene Probe Gent 24 (1D-Ansatz) [Anhang: 7.6]. Als Schwellwert für eine erfolgreiche Proteinidentifikation wurde ein Mascot Score >40 gewählt.

Die Suchen gegen die größeren Datenbanken NCBIInr und UniProtKB/TrEMBL erlaubten doppelt so viele Spektren zu identifizieren wie die Suche gegen UniProtKB/Swiss-Prot [Tabelle: 4.2]. Allerdings benötigten diese bis zu fünfzehnmal längere Suchzeiten. Im Gegensatz zu diesen großen Datenbanken dauerte die Suche gegen die Mischdatenbank nur doppelt so lange wie die gegen UniProtKB/Swiss-Prot, ermöglichte es aber, fünfmal so viele Spektren als diese zu identifizieren.

Tabelle 4.2: Anzahl an identifizierten Spektren für Datenbanken verschiedener Größe

| Datenbank | Identifizierte Spektren [Mascot Score >40] | Datenbankgröße in [Sequenzen] | Suchzeit in [h] |
|--------------------------------------|---|----------------------------------|--------------------|
| UniProtKB/Swiss-Prot | 821 | 548.872 | 0,5-0,75 |
| UniProtKB/TrEMBL | 1.487 | 27.122.814 | 8-8,5 |
| NCBI nr | 1.614 | 35.149.712 | 7-11 |
| Metagenome + UniProtKB/Swiss-Prot | 4.288 | 2.340.186 | 1-1,25 |

4.3.2 Qualität der Spektrenidentifikationen

Die Qualität von Spektrenidentifikationen konnte entweder durch die Berechnung der FDR oder entsprechende Scores der Datenbanksuchalgorithmen beurteilt werden. Während die Scores die Wahrscheinlichkeit mit der ein Peptid ein Spektrum erklärt widerspiegeln, zeigt die FDR an, wie viel Prozent der Spektrenidentifikationen wahrscheinlich falsch sind. Standard in der wissenschaftlichen Gemeinschaft ist eine FDR von maximal 1% [Barnouin (2011)]. Voraussetzung für die Berechnung der FDR ist allerdings, dass die Datenbank die taxonomische Zusammensetzung und Komplexität der Probe repräsentiert. Da dies für die verschiedenen Biogasanlagenproben nicht immer zutrifft, wurde die Nutzung des FDR-Konzeptes am Beispiel von 40 Biogasanlagenproben evaluiert.

Diese Proben wurden mittels Elite Hybrid Ion Trap-Orbitrap MS/MS (1D-Ansatz) vermessen und gegen UniProtKB/Swiss-Prot sowie eine Mischdatenbank aus mehreren Biogasmengenomen [Schlüter et al. (2008), Hanreich et al. (2013), Rademacher et al. (2012), Stolze et al. (2016)] UniProtKB/Swiss-Prot gesucht. Neben dem Mascot Score von 40 wurde für die Mischdatenbank eine FDR von 1% als Schwellenwert angewendet.

Wie schon in Kapitel 4.3.1 gezeigt, konnte durch Datenbanksuche gegen Metagenome anstelle der UniProtKB/Swiss-Prot-Datenbank die Anzahl an identifizierten Spektren deutlich gesteigert werden [Abbildung: 4.8]. Während die Anzahl der identifizierten Spektren mit FDR-Schwelle bzw. festem Mascot Score für die meisten Biogasanlagenproben ähnlich war, wurden für einige wenige Proben mit FDR-Schwelle bis zu 50% weniger Spektren identifiziert als mit festem Mascot Score. Auffälligerweise war der Nutzen der Metagenome bei diesen Proben sehr gering. Bei der Probe Gent 16 wurden bei Nutzung der FDR-Schwelle sogar weniger Spektren bei der Datenbanksuche gegen die Metagenome identifiziert als gegen die UniProtKB/Swiss-Prot-Datenbank.

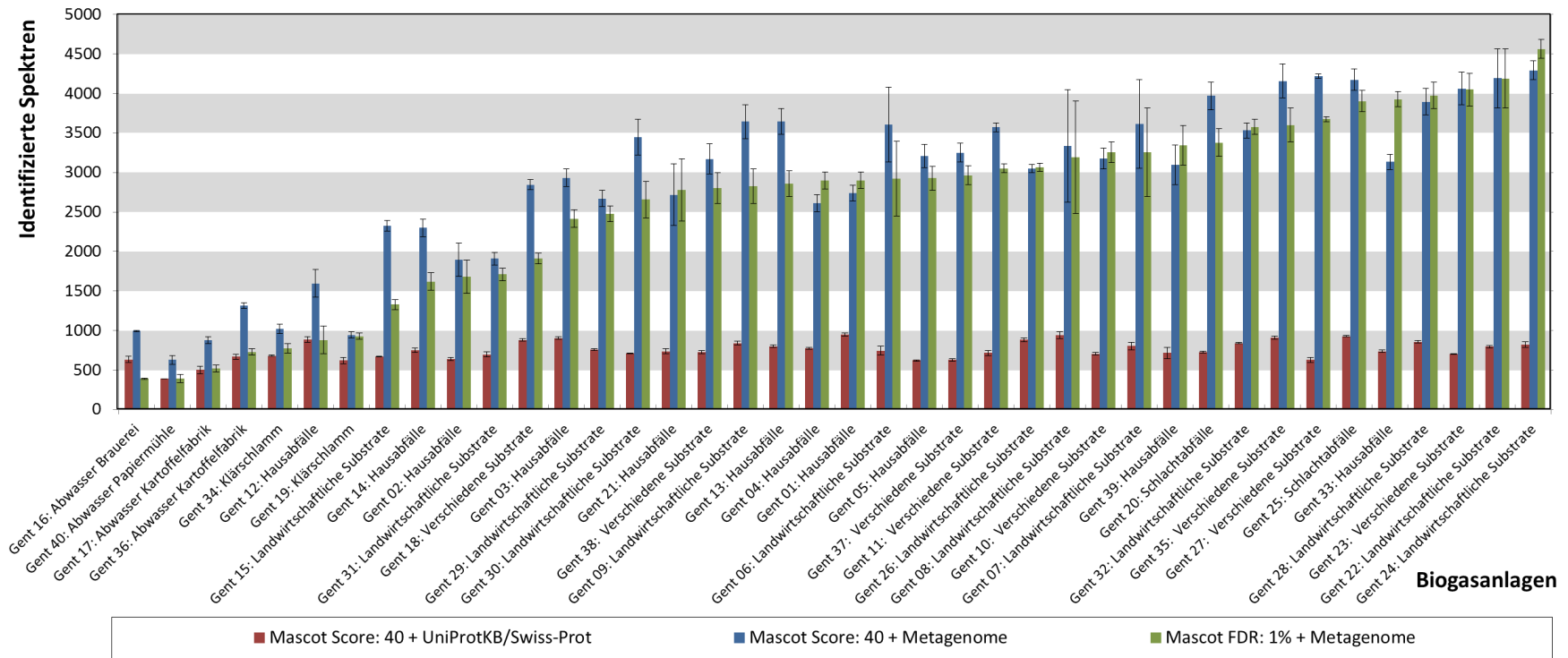


Abbildung 4.8: Einfluss verschiedener Schwellwerte auf die Anzahl an identifizierten Spektren

4.3.3 Multiple Suchalgorithmen

Die Anzahl und die Glaubwürdigkeit von Spektrenidentifikationen kann durch die Kombination verschiedener Datenbanksuchalgorithmen erhöht werden [Vaudel et al. (2011), Shteynberg et al. (2013), Muth et al. (2013b)].

Zur Evaluierung des Nutzens zusätzlicher Datenbanksuchalgorithmen wurde die mittels Elite Hybrid Ion Trap-Orbitrap MS/MS analysierte Probe Gent 24 (1D-Ansatz) [Anhang: 7.6] einmal mit Mascot [Perkins et al. (1999)], einmal mit Xtandem! [Craig and Beavis (2004)] und einmal mit OMMSA [Geer et al. (2004)] gegen UniProtKB/Swiss-Prot bei einer FDR von 1% gesucht.

Mascot identifiziert 580 Peptide, Xtandem! 346 und OMMSA 339 [Abbildung: 4.9]. Insgesamt konnten durch die Kombination der Datenbanksuchalgorithmen 737 verschiedene Peptide identifiziert werden, wovon 324 Peptide von mindesten zwei Datenbanksuchalgorithmen vorgeschlagen wurden.

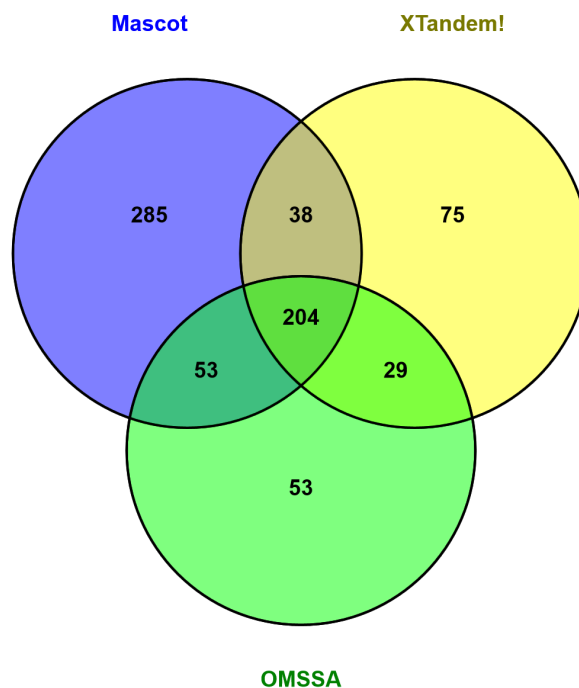


Abbildung 4.9: Einfluss verschiedener Datenbanksuchalgorithmen auf die Anzahl an identifizierten Peptiden, dargestellt als Venn-Diagramm [Oliveros (2007)]

4.3.4 Posttranslationale Modifikationen

Das Vorhandensein von PTMs war ein wichtiger Suchparameter für die Identifikation von Spektren mittels Datenbanksuchalgorithmen sowie für die spätere Interpretation der Da-

ten. Würden bei den Datenbanksuchen aber alle möglichen PTMs zugelassen, würde sich der Suchraum vervielfachen und dadurch auch die Anzahl an Zufallstreffern. Die Oxidation des Methionins und die Carbamidomethylierung (C) sind daher die einzigen PTMs welche standardmäßig bei Datenbanksuchen berücksichtigt werden. Beide entstehen bei der Probenvorbereitung [Boja and Fales (2001)].

Um zu überprüfen, ob noch weitere PTMs relevant waren, wurde die eingelaufene Probe Gent 24 (1D-Ansatz) [Anhang: 7.6] mittels Elite Hybrid Ion Trap-Orbitrap MS/MS analysiert und gegen UniProtKB/Swiss-Prot bei einem Mascot Score >40 und unter Zulassung der gängigsten Modifikationen gesucht.

Neben Peptiden mit Carbamidomethylierung und oxidiertem Methionin konnten mehrere amidierete und deamidierete Peptide, sowie Peptide mit Methyl-, Acetyl- und Formylgruppe identifiziert werden [Tabelle: 4.3]. Die meisten der gefundenen PTMs entsprachen allerdings genau dem Unterschied zu einer anderen Aminosäure [Tabelle: 4.4].

Tabelle 4.3: *Einfluss ausgewählter PTMs auf die Anzahl an identifizierten Peptiden.*
Zusätzlich sind die PTMs mit großem Einfluss auf die Anzahl der identifizierten Peptide hervorgehoben.

| PTM | Peptide | Peptide mit PTMs |
|----------------------------|---------|------------------|
| 1.) OHNE | 284 | 0 |
| 2.) Deamidated | 307 | 75 |
| 3.) OxidationM | 293 | 34 |
| 4.) AmidatedC | 288 | 23 |
| 5.) Carbamidomethyl | 304 | 20 |
| 6.) MethylC | 292 | 14 |
| 7.) AcetylK | 291 | 11 |
| 8.) FormylN | 291 | 8 |
| 9.) AcetylN | 286 | 3 |
| 10.) AcetylprotN | 286 | 3 |
| 11.) Pentose | 286 | 2 |
| 12.) AmidatedprotC | 284 | 2 |
| 13.) FormylprotN | 286 | 2 |
| 14.) BiotinN | 285 | 1 |
| 15.) Sulfo | 285 | 1 |
| 16.) AmmoniaLost | 284 | 0 |
| 17.) BiotinK | 284 | 0 |

| PTM | Peptide | Peptide mit PTMs |
|---------------|---------|------------------|
| 18.) PhosphST | 284 | 0 |
| 19.) PhosphY | 284 | 0 |

Tabelle 4.4: *Biochemischer Unterschied ausgewählter Aminosäuren,*
bei denen die Massendifferenz bestimmten PTMs entspricht.

| Modifizierung | Δ_{Masse} | Aminosäure 1 | Aminosäure 2 |
|---------------|------------------|----------------------------------|----------------------------------|
| Deamidierung | +1 | Asparagin (N) ($C_4H_6O_2N_2$) | Asparat (D) ($C_4H_5O_3N$) |
| | | Glutamin (Q) ($C_5H_8O_2N_2$) | Glutamat (E) ($C_5H_7O_3N$) |
| Amidierung | -1 | Asparat (D) ($C_4H_5O_3N$) | Asparagin (N) ($C_4H_6O_2N_2$) |
| | | Glutamat (E) ($C_5H_7O_3N$) | Glutamin (Q) ($C_5H_8O_2N_2$) |
| Methylierung | +14 | Glycin (G) (C_2H_3ON) | Alanin (A) (C_3H_5ON) |
| | | Serin (S) ($C_3H_5O_2N$) | Threonin (T) ($C_4H_7O_2N$) |
| | | Valin (V) (C_5H_9ON) | Isoleucin (I) (C_5H_9ON) |
| | | Valin (V) (C_5H_9ON) | Leucin (L) (C_5H_9ON) |
| | | Asparagin (N) ($C_4H_6O_2N_2$) | Glutamin (Q) ($C_5H_8O_2N_2$) |
| | | Asparat (D) ($C_4H_5O_3N$) | Glutamat (E) ($C_5H_7O_3N$) |
| | | Serin (S) ($C_3H_5O_2N$) | Asparat (D) ($C_4H_5O_3N$) |
| Formylierung | +28 | Threonin (T) ($C_4H_7O_2N$) | Glutamat (E) ($C_5H_7O_3N$) |
| | | Serin (S) ($C_3H_5O_2N$) | Glutamat (E) ($C_5H_7O_3N$) |
| Acetylierung | +42 | Serin (S) ($C_3H_5O_2N$) | Glutamat (E) ($C_5H_7O_3N$) |

4.3.5 *De novo* Sequenzierung und BLAST-Suche

Voraussetzung für die erfolgreiche Identifikation eines Proteins war das Vorhandensein der Aminosäuresequenz in der entsprechenden Proteindatenbank. Alternativ kann die Aminosäuresequenz aus den Spektren *de novo* sequenziert und durch BLAST-Suche [Shevchenko et al. (2001)] die Proteine identifiziert werden. Um das Potential dieser Strategie zu testen, wurden 39.071 Spektren der Probe Gent 24 (1D-Ansatz), welche mittels Elite Hybrid Ion Trap-Orbitrap MS/MS gemessen wurden, mit der Software DeNovoGUI (Version: 1.12.1) [Muth et al. (2014)] *de novo* sequenziert. Zusätzlich wurde für die zehn besten *de novo* sequenzierten Peptide das zugehörige Protein durch BLAST-Suche gegen NCBI mit einem Erwartungswert von kleiner 10^{-4} bestimmt.

Insgesamt lieferte die Software DeNovoGUI Peptidvorschläge zu 38.541 Spektren [Abbildung: 4.10]. Viele dieser Peptidvorschläge bestanden allerdings nur aus kurzen Teilfragmenten. 12.596 Peptidvorschläge waren aber valide (Scores über 100) und erklärten teilweise die gesamte Aminosäuresequenz des Peptids. Die BLAST-Suche der zehn besten komplett *de novo* sequenzierten Peptide lieferte nur für 3 Peptide entsprechende Protein-

4.3 Optimierung der Proteinidentifikation

identifikationen (BLAST-Score über 40). Dabei wurden die Proteine Glycerinaldehyde-3-phosphat Dehydrogenase, Methyl-Coenzym M Reduktase und Acetyl-CoA Decarboxylase/Synthase mit Scores von 41,8, 45,6 bzw. 44,8 identifiziert.

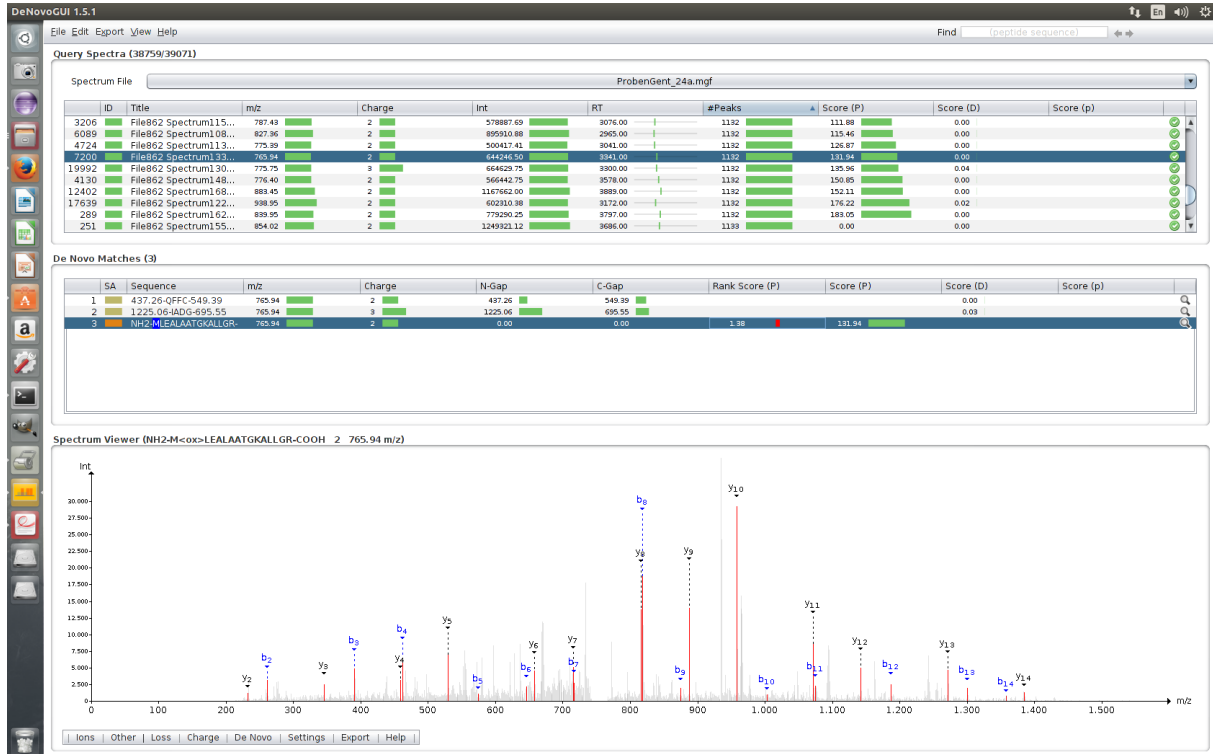


Abbildung 4.10: *De novo* Sequenzierung mittels DeNovoGUI für die Proben Gent 24, basierend auf 39.071 gemessenen Spektren.

4.3.6 Diskussion

Den größten Einfluss auf die Anzahl an identifizierten Spektren und damit auch auf die Anzahl an identifizierten Peptiden sowie Proteinen hatte die Auswahl der Datenbank [Tabelle: 4.2] [Muth et al. (2015b)]. Die besten Ergebnisse verbunden mit angemessenen Suchzeiten ermöglichte die Suche gegen Metagenome von ähnlichen Proben [Kapitel: 4.3.1, 4.3.2]. Allerdings mussten die identifizierten Proteine aus Metagenomdatenbanken für die taxonomische und funktionelle Auswertung noch mittels BLAST-Suche annotiert werden [Altschul et al. (1990)]. Im Gegensatz dazu waren diese Informationen für die annotierten und universellen Proteindatenbanken UniProtKB/Swiss-Prot, UniProtKB/TrEMBL und NCBIInr bereits vorhanden. Suchen gegen diese Datenbanken identifizierten aber deutlich weniger Spektren und benötigte teilweise zehnmal mehr Zeit. Zum Einen fehlen in diesen Datenbanken relevante Proteineinträge, zum Anderen beinhalten diese Datenbanken aber auch ein Vielfaches an irrelevanten Proteineinträgen. Dadurch waren nicht nur die Suchzeiten deutlich höher, sondern auch die Chance für Zufallstreffer und somit die FDR. Als Konsequenz wurden korrekte Identifikationen als falsch verworfen, weil sie unter die Signifikanzschwelle fielen. Ein Ansatz, um dieses Problem zu reduzieren wäre die Anzahl der Proteineinträge aus UniProtKB/Swiss-Prot, UniProtKB/TrEMBL und NCBIInr basierend auf taxonomischem Vorwissen oder durch eine zweistufige Suche zu reduzieren [Jagtap et al. (2013)]. Bei letzterer Strategie werden für die eigentliche Datenbanksuche alle Proteine aus der Datenbank entfernt, deren Taxa in der ersten Suchrunde nicht gefunden wurden.

Grundsätzlich ist für die Charakterisierung einer einzelnen Probe oder den Vergleich von Zeitreihen das Fehlen einiger Spektrenidentifikationen aufgrund der Datenbankabhängigkeit der FDR oder nicht optimalen Proteindatenbanken verkräftbar. Problematischer ist der Vergleich von Proben, von denen eine deutlich schlechter als die andere von der Datenbank abgebildet wird [Kapitel: 4.3.2]. Dann werden für diese Probe weniger Spektren identifiziert und die FDR ist höher. Folglich ist wiederum der Schwellwert für die korrekte Spektrenidentifikationen höher, wodurch deren Anzahl noch kleiner wird.

Die einfachste Lösung wäre es, in solchen Fällen einen festen Score (z. B. Mascot Score) zu verwenden [Kapitel: 4.3.2]. Allerdings wird in der wissenschaftlichen Gemeinschaft aktuell nur die Verwendung der FDR akzeptiert [Barnouin (2011)]. Einen vielversprechenden Ansatz für dieses Problem lieferte die Arbeit von *Gonnelli et al. (2015)*, die vorschlagen die Glaubwürdigkeit von Spektrenidentifikationen mittels vorher trainiertem Klassifikationsalgorithmus zu bestimmen.

Eine weitere Möglichkeit sowohl die Glaubwürdigkeit als auch die Anzahl von Spektreni-

identifikationen zu steigern, war die Identifikation mit verschiedenen Datenbanksuchalgorithmen [Kapitel: 4.3.3]. Jedoch konnten für Metaproteomanalysen die Datenbanksuchalgorithmen Xtandem! und OMMSA nur bedingt genutzt werden, weil sie zu jedem identifizierten Peptid nur das beste Protein ausgaben und dadurch homologe Proteine ignorierten. Für zukünftige Projekte wäre dennoch der Einsatz dieser Algorithmen sinnvoll, sofern alle Proteineinträge zu einem Peptid aus den Proteindatenbanken abgefragt werden. Zusätzlich könnten auch noch weitere Datenbanksuchalgorithmen wie MS Amanda [Dorfer et al. (2014)], MS-GFDB [Kim et al. (2010)], Myrimatch [Tabb et al. (2007)] oder Andromeda [Cox et al. (2011)] genutzt werden.

Parallel zu den verschiedenen Datenbanksuchalgorithmen wäre es auch möglich, die Peptididentifikationen durch *de novo* Sequenzierung zu validieren [Muth et al. (2015b)]. Leider eigneten sich die *de novo* sequenzierten Peptide nur begrenzt dazu, mittels BLAST-Suche Proteine zu identifizieren, welche in der Datenbank fehlten [Kapitel: 4.3.5]. Die einzelnen tryptischen Peptide waren meist zu kurz um sie eindeutig einem Protein oder zumindest einer Proteinfamilie zuzuordnen. Eine Möglichkeit, dieses Problem zu lösen, wäre die Nutzung alternativer Proteasen wie z. B. Lys-C oder Arg-C, welche deutlich längere Peptide erzeugen.

Eine eher untergeordnete Bedeutung für die Identifikation der Proteine hatten PTMs. Im Gegensatz zu Peptiden mit den standardmäßigen Modifikationen Carbamidomethylierung und Oxidation des Methionin, erschien das vermehrte Auftreten von Peptiden mit Amidierungen, Deamidierungen, Methyl-, Acetyl- und Formylgruppen zwar erstmal überraschend [Kapitel: 4.3.4]. Diese hätte ein Hinweis auf ungewollte Nebenreaktionen während des Trypsinverdaus [Hao et al. (2011)], den Abbau der Proteine oder Zellyse sein können [Justice et al. (2012)]. Viel wahrscheinlicher waren aber das Fehlen der korrekten Aminosäuresequenz in der Datenbank und die fälschliche Identifikation einer anderen Aminosäuresequenz. Zum Beispiel entspricht die Masse von Asparagin der von amidierter Asparaginsäure. Analog dazu könnten auch die Modifikationen mit Methyl-, Acetyl-, und Formylgruppen erklärt werden. Passend dazu waren die gefundenen PTMs nicht in UniProt/Swiss-Prot beschrieben und es gilt als unwahrscheinlich, bakterielle und archaeele Modifikationen ohne Anreicherungen zu identifizieren [Cain et al. (2014), Eichler and Maupin-Furlow (2013)]. Folglich wurde auch nur Carbamidomethyl und oxidiertes Methionin als PTMs für spätere Suchen beachtet.

Zusammenfassung: Durch den Einsatz von Metagenomdatenbanken konnte die Anzahl an identifizierten Spektren verfünffacht werden. Vielversprechende Ansätze, die Proteinidentifikation noch weiter zu verbessern, waren die Kombination von mehreren Datenbank-

suchalgorithmen und die Entwicklung von neuen bioinformatischen bzw. biostatistischen Konzepten, um die Glaubhaftigkeit von Spektrenidentifikationen zu validieren. Im Gegensatz dazu erbringt die Beachtung weiterer PTMs während der Datenbanksuche sowie die *de novo* Sequenzierung nur wenig Nutzen.

4.4 Bioinformatische Datenauswertung mit dem MetaProteomeAnalyzer

³Die besondere Herausforderung bei der Auswertung von Metaproteomanalysen ist die große Komplexität der Daten sowie die hohe Redundanz der Proteinidentifikationen aufgrund von homologen Proteinen. Die Konsequenz ist ein erheblicher Aufwand während der Datenauswertung. Da bisherige bioinformatische Workflows diese besonderen Probleme von Metaproteomanalysen nicht genügend beachten, war die Entwicklung eines eigenen bioinformatischen Workflows notwendig.

Kernstück dieses Workflows ist die Software MetaProteomeAnalyzer (MPA), welche die Beseitigung von Redundanzen sowie eine schnelle taxonomische und funktionelle Auswertung ermöglicht. Im Folgenden wird der gesamte entwickelte bioinformatische Workflow [Kapitel: 4.4.1] und anschließend das für die Beseitigung der Redundanzen entwickelte Metaproteinkonzept [Kapitel: 4.4.2] vorgestellt.

4.4.1 Workflow MetaProteomeAnalyzer

Nach der massenspektrometrischen Analyse der Proben erfolgt die Identifikation der Spektren mit dem Datenbanksuchalgorithmus Mascot [Perkins et al. (1999)] gegen UniProtKB/SwissProt und entsprechende Metagenomdatenbanken [Abbildung: 4.11]. Im Anschluss werden die Ergebnisdateien (.dat-Dateien) mit dem MPA-Client und einer FDR von 1% [Abbildung: 4.11 A] [Helsens et al. (2007)] über den MPA-Server eingeladen und in einer SQL-Datenbank [Abbildung: 4.11 B] gespeichert. Zusätzlich werden Metainformationen für die einzelnen identifizierten Proteine bezüglich Taxonomie, Funktion und Struktur von der UniProtKB/SwissProt-Datenbank automatisch abgefragt [Patient et al. (2008)]. Für Proteineinträge ohne UniProtKB-Metainformationen, z. B. aus Metagenomdatenbanken, ermöglicht der MPA die Metainformationen von homologen Proteinen durch BLAST-Suche gegen UniProtKB/SwissProt [Altschul et al. (1990)] einzubinden.

Für die Auswertung der Daten werden die Ergebnisse mit dem MPA-Client vom MPA-Server und der SQL-Datenbank heruntergeladen und redundante Proteine nach gleicher UniRef50 gruppiert [Kapitel: 4.4.2] [Abbildung: 4.11 C]. Die Taxonomie (NCBI-Taxonomie) [Federhen (2012)] der gruppierten Proteine wird dabei als der letzte gemeinsame Vorfahre im phylogenetischen Baum definiert. Neben der standardmäßigen Darstellung der identifizierten Spektren, Peptide und Proteine, umfasst der MPA-Client spezielle Ansichten für die Taxonomie, KEGG-Netzwerke, [Kanehisa and Goto (2000)], Funktionsprofile [UniProt Consortium (2012)] und EC-Nummern [Abbildung: 4.11 D]. Die entsprechenden Daten dieser Ansichten können im Anschluss flexibel als kommaseparierte Da-

³Basiert teilweise auf der Publikation Muth et al. (2015)

teien bzw. im Fall der Taxonomieprofile als Krona-Plot [Ondov et al. (2011)] exportiert werden [Abbildung: 4.11 E]. Basierend auf den identifizierten Metaproteinen (UniRef50), Taxonomien und biologischen Prozessen sowie deren Spektrenhäufigkeit [Lundgren et al. (2010)] erfolgt abschließend der Vergleich und die statistische Auswertung der verschiedenen Proben durch ein separates Matlab-Programm.

4.4.2 Metaproteinkonzept

Ein Hauptproblem von Metaproteomanalysen sind redundante Proteinidentifikationen aufgrund von homologen Proteinen. Gängige Ansätze, diese Redundanzen zu entfernen, sind das Gruppieren von Proteinen bei überlappenden Peptididentifikationen [Schneider et al. (2011), Keiblinger et al. (2012)] oder ähnlichen Proteinsequenzen [Lü et al. (2014)] zu sogenannten Metaproteinen [Muth et al. (2015a)]. Um diese Gruppierungsstrategien zu testen, wurden entsprechende Anforderungen definiert, vier Metaproteinregeln im MPA implementiert und für die Biogasanlage Ebendorf (2D-Ansatz) [Anhang: 7.6] getestet. Zusätzlich wurden die Ergebnisse der Metaproteinerzeugung für das Protein F420-abhängige Methylentetrahydromethanopterindehydrogenase visualisiert [Abbildung: 4.13].

Anforderungen für die Metaproteinerzeugung

Für die Definition von Anforderungen an das Metaproteinkonzept wurde vorab der Grad der Homologie für die Proteine Elongationsfaktor TU (EFTU) und Methyl Coenzym M Reduktase (MCR) Untereinheit β für repräsentative taxonomische Gruppen untersucht [Anhang: 7.9]. Folgende vier Punkte sollten bei der Erzeugung von Metaproteinen beachtet werden. Erstens ist die Aminosäuresequenz von verschiedenen, homologen Proteinen unterschiedlich stark konserviert. Zweitens gibt es auch innerhalb eines homologen Proteins Bereiche, die verschieden stark konserviert sind. Drittens sind die Aminosäuresequenzen umso ähnlicher, je näher zwei Spezies miteinander verwandt sind. Für Spezies aus verschiedenen Domänen kann der Unterschied in der Aminosäuresequenz aber auch über 50% betragen. Viertens ist der Nachweis von spezifischen Peptiden ein starkes Indiz für das Vorhandensein von bestimmten Taxonomien.

Implementierung des Metaproteinkonzeptes

Die Ausgangsbasis für die Implementierung des Metaproteinkonzeptes waren Ergebnisse von Datenbanksuchalgorithmen, welche für jedes identifizierte Spektrum die vorgeschlagenen Peptid- und Proteinidentifikation umfassten [Abbildung: 4.12 A]. Als erster Schritt der Metaproteinerzeugung werden die Ergebnisse der Datenbanksuchalgorithmen mit dem MPA in eine graphenähnliche Struktur umgewandelt, um mehrfach identifizierte Spektren und Peptide zu entfernen [Abbildung: 4.12 B]. Die Taxonomie von Peptiden,

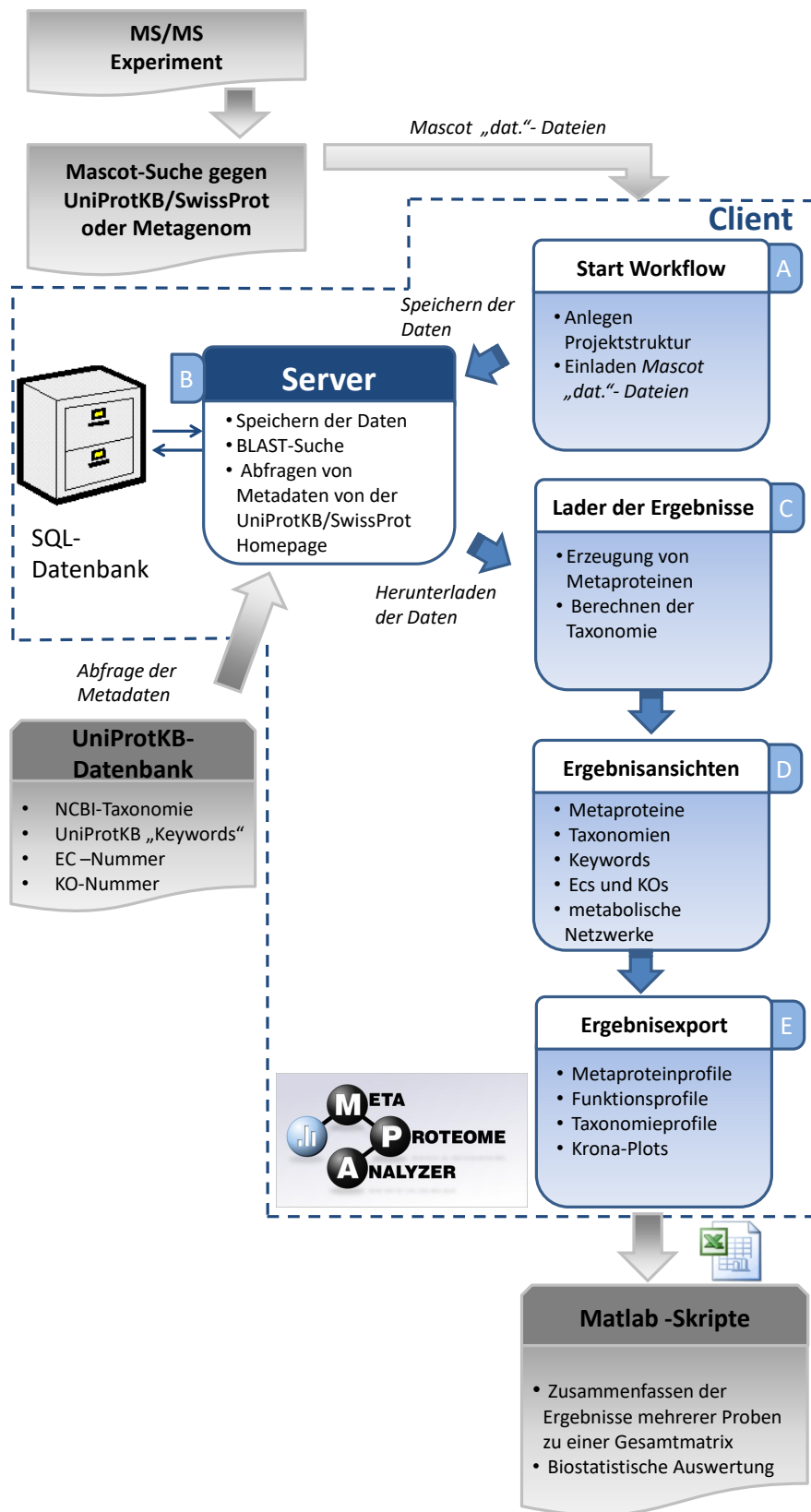
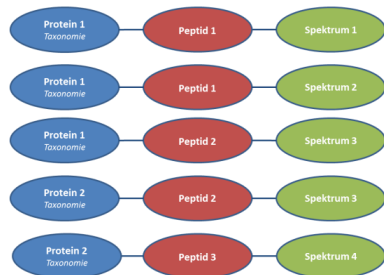


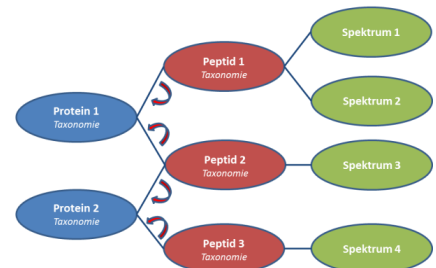
Abbildung 4.11: Bioinformatischer Datenauswerteworkflow mit dem MetaProteomeAnalyzer

die zu mehreren Proteinen gehören, werden dabei als kleinster gemeinsamer Vorfahre im phylogenetischen Baum definiert [Abbildung: 4.12 C] [Huson et al. (2011), Mesuere et al. (2012)]. Analog dazu wird anschließend die Proteintaxonomien mittels kleinstem gemeinsamen Vorfahre der einzelnen Peptidtaxonomien neu bestimmt [Abbildung: 4.12 D]. Die Gruppierung der homologen Proteine erfolgt dann nach den im nächsten Abschnitt vorgestellten Regeln. Zum Beispiel, wie in Abbildung 4.12 E gezeigt, bei geteilten Peptididentifikationen. Die Taxonomie des neu erzeugten Metaproteins wird analog zur Peptid- und Proteintaxonomie als kleinster gemeinsamer Vorfahre definiert [Abbildung: 4.12 F].

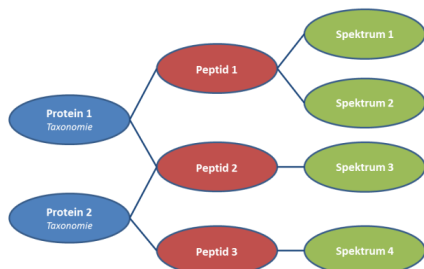
A. Ergebnisstruktur der Datenbanksuchalgorithmen



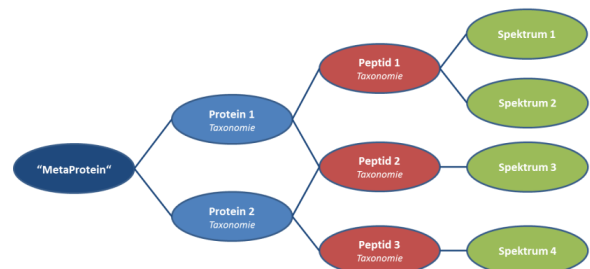
D. Aktualisieren der Proteintaxonomien aus Peptidtaxonomien



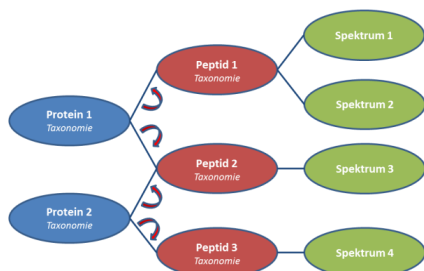
B. Überführung in nicht-redundante Ergebnisstruktur



E. Gruppieren der Proteine zu „Metaproteinen“ (überlappende Peptididentifikationen)



C. Definieren der Peptidtaxonomien aus Proteintaxonomien



F. Definieren der „MetaProtein“-Taxonomie

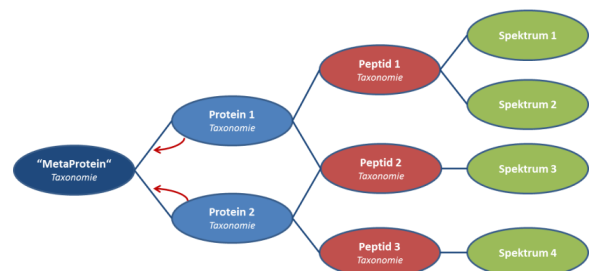


Abbildung 4.12: Workflow zu Erzeugung von Metaproteinen

Proteinregel

Redundante Proteinidentifikationen [Abbildung: 4.13 A] können bei ähnlicher Aminosäuresequenz bzw. Gensequenz zusammengefasst werden. Möglichkeiten, diese Sequenzähnlichkeiten rechenzeitsparend zu bestimmen, sind die Nutzung der Ergebnisse vorhandener Clusterprogramme wie UniRef Cluster [Suzek et al. (2007), Lü et al. (2014)] und KEGG

Ontologien (KO) [Mao et al. (2005), Gotelli et al. (2012)]. Während die UniRef Cluster auf einem Clustering der gesamten UniProtKB/SwissProt nach 100% (UniRef100), 90% (UniRef90), sowie 50% (UniRef50) Aminosäuresequenzähnlichkeit beruhen, werden bei KEGG Ontologien neue Gene nach ihrer Ähnlichkeit zu bereits klassifizierten Genen einsortiert.

Für den untersuchten Testdatensatz (Biogasanlage Ebendorf, 2D-Ansatz, Mascot-Suche gegen UniProtKB/SwissProt, FDR 1%) wurden die 1324 redundanten Proteinidentifikationen mittels UniRef50 Cluster und KO zu 677 bzw. 898 Metaproteinen gruppiert [Abbildung: 4.13 B, C]. Allerdings waren nicht für alle Proteine KO-Klassifikationen verfügbar. Ferner waren die UniRef100 und UniRef90 Cluster ungeeignet, da diese nur ermöglichten eng verwandte Proteine mit sehr kleinen Aminosäuresequenzunterschieden zusammenzufassen.

Peptidregel

Eine andere Möglichkeit, Proteinidentifikationen ohne Vorwissen von bestimmten Clustern zu gruppieren, ist das Zusammenfassen nach überschneidenden Peptididentifikationen. Dies erfolgte entweder sobald ein Peptid [Lü et al. (2014)] oder das gesamte Peptidset [Schneider et al. (2012), Keiblinger et al. (2012), Kolmeder et al. (2012)] gleich war. Insgesamt reduzierte das Gruppieren nach nur einer Peptidüberschneidung mit 749 Metaproteinen [Abbildung: 4.13 E] die Anzahl der 1324 redundanten Proteineinträge stärker als nach dem gleichen Peptidset mit 825 Metaproteinen [Abbildung: 4.13 D].

Taxonomieregel

Für bestimmte Experimente war die taxonomische Zusammensetzung durch andere Methoden oder die Gestaltung des Experimentes vorab bekannt. Um dieses Wissen in die Metaproteinerzeugung einfließen zu lassen, konnte die Protein- und die Peptidregel um die Forderung einer bestimmten taxonomischen Ähnlichkeit erweitert werden. So waren für das vorliegende Beispiel nur methanogene *Archaeen* relevant, wozu *Archaeoglobulus fulgidus* nicht zählte [Abbildung: 4.13 G]. Um diese Ordnung bei der Metaproteinerzeugung abzutrennen, wurde die Regel mit mindestens einer Peptidüberschneidung um die Forderung der gleichen taxonomischen Klasse erweitert. Dadurch wurde die Redundanz auf 757 Metaproteine reduziert.

Levenstheinregel

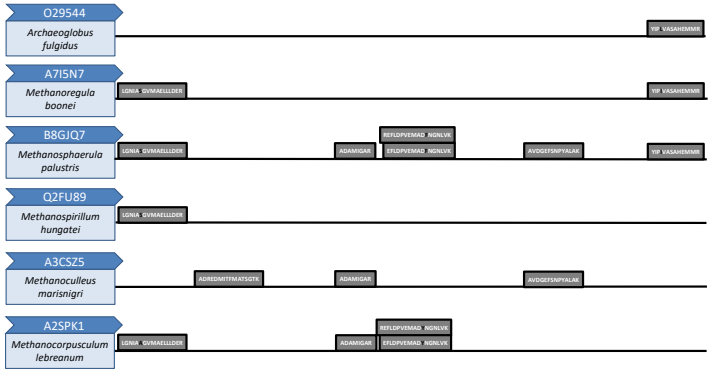
Um das Vorhandensein von zwei separaten Proteinisoformen zu beweisen, mussten zwei verschiedene Peptide für die gleiche Sequenzposition gefunden werden. Für die algorithmische Umsetzung dieser Forderung wurde die Regel mit mindestens einer Peptidüberschnei-

dung mit der Levensthein-Distanz [Levenshtein (1966)] erweitert. Durch die Levensthein-Distanz kleiner drei wurden Proteine nicht zu Metaproteinen fusioniert, wenn es zwei ähnliche Peptide gab, die sich um weniger als drei Punktmutationen unterschieden. So konnte für die untersuchte Probe nachgewiesen werden, dass 21 homologe Proteine von mindestens 2 Spezies exprimiert wurden [Abbildung: 4.13 F (770 Metaproteine) vs. E (749 Metaproteine)].

A. Redundante Proteinergebnisse (1324 "Metaproteine")

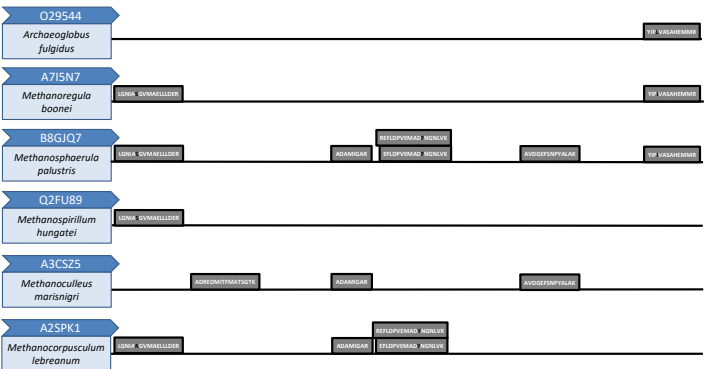


B. UniRef 50 Cluster (677 "Metaproteine")



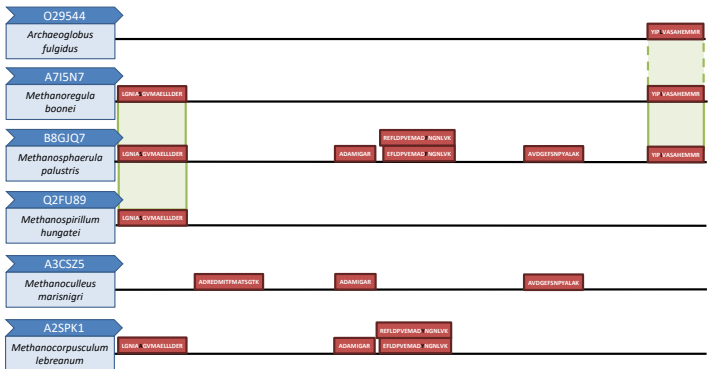
MP (p94951)
Metaprotein (A2SPK1)

C. KEGG Ontology (KO) (898 "Metaproteine")



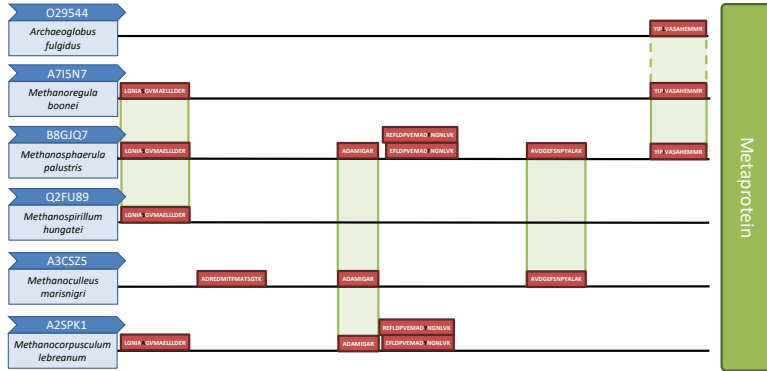
Metaprotein (K00319)

D. Peptidset ("825 Metaproteine")

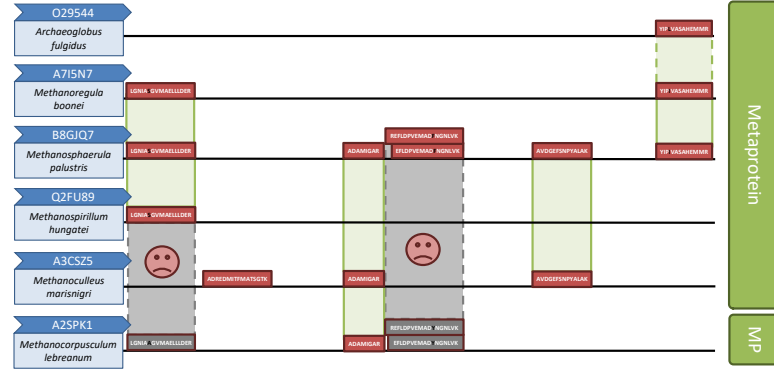


Metaprotein
MP
MP

E. Peptidüberschneidung (749 "Metaproteine")



F. Kombination Peptidüberschneidung + Levenstein-Distanz < 3 (770 "Metaproteine")



G. Kombination Peptidüberschneidung + gleiche Klasse ("757 Metaproteine")

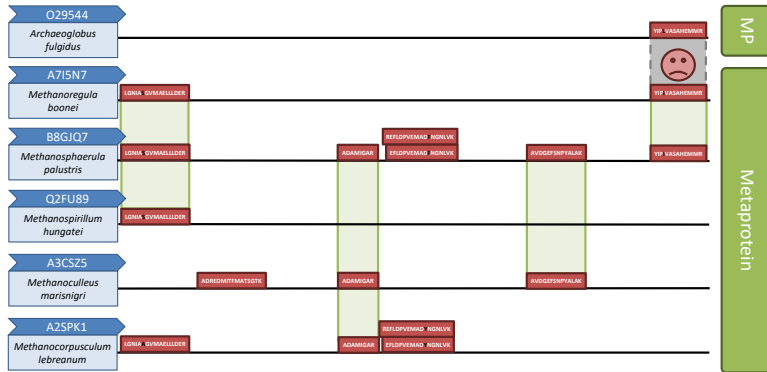


Abbildung 4.13: *Verschiedene Metaproteinstrategien.*

Diese Abbildung zeigt den Effekt verschiedener Gruppierungsstrategien auf die Erzeugung von Metaproteinen am Beispiel des Proteins F420-abhängige Methyltetrahydromethanopterindehydrogenase. Zusätzlich sind für die einzelnen Strategien die gesamte Anzahl an Metaproteinen dargestellt.

4.4.3 Diskussion

Die Entwicklung des MPAs ermöglichte die schnelle taxonomische und funktionelle Auswertung von Metaproteomanalysen. Die Vorteile des MPAs im Vergleich zu anderen Proteomauswertepattformen wie OpenMS [Sturm et al. (2008)], Proteome DiscovererTM oder ProteinScope [Thiele et al. (2010)] sind die Gruppierung von homologen Proteinen zu Metaproteinen und die automatische, taxonomische und funktionelle Annotation von unbekanntem Proteinen durch BLAST-Suche.

Für zukünftige Anwendungen wäre die direkte Gegenüberstellung von einzelnen Experimenten innerhalb des MPA z. B. in Form einer Matrix wünschenswert. Dies würde erlauben, Metaproteine übergreifend für verschiedene Experimente zu erzeugen und deren Häufigkeit zu vergleichen.

Ein weiterer Ansatz den MPA zu verbessern, wäre das Hinzufügen von zusätzlichen Datenbanksuchalgorithmen [Kapitel: 4.3.3] um die Anzahl und Glaubhaftigkeit der Proteinidentifikationen zu erhöhen. Zusätzlich könnten die Programmen Biocyc [Caspi et al. (2012)] und Reactome [Joshi-Tope et al. (2005)] in den Datenauswertungsworkflow integriert werden, um die Funktion der identifizierten Proteine direkt in ihren metabolischen Netzwerken zu visualisieren. Die finale Archivierung der durch den MPA prozessierten Daten könnte durch den Einbau einer Schnittstelle zu Onlineproteomarchiven wie z. B. PRIDE [Vizcaíno et al. (2013)] realisiert werden.

Alle Metaproteinstrategien eigneten sich dazu, die Redundanzen der homologen Proteine wirksam zu reduzieren. Der Vergleich der taxonomischen Profile für die verschiedenen Metaproteinstrategien im Krona-Plot zeigte keine wesentlichen Unterschiede (Daten nicht gezeigt). Grundsätzlich sind die Peptidregeln aber besser als die Proteinregeln, weil sie durch die Gruppierung nach identifizierten Peptiden den Konservierungsgrad der Proteinsequenz mit beachtet. Besonders vorteilhaft ist die Kombination der Gruppierung bei einer überschneidenden Peptididentifikation mit einer Levenstein-Distanz kleiner als drei. Diese Strategie ermöglicht es, das Vorhandensein von verschiedenen Proteinisoformen zu berücksichtigen. Für spezifische Fragestellungen, oder um entsprechendes taxonomisches Vorwissen in die Gruppierung einfließen zu lassen, kann auch die Einschränkung der Gruppierung bei bestimmten Taxonomieunterschieden sinnvoll sein.

Leider war der Vergleich von Proben, welche mit der Peptidregel gruppiert wurden, nicht möglich, da dafür Metaproteine experimentübergreifend erzeugt werden müssten. Daher war die Nutzung der Proteinregeln insbesondere der UniRef50 Cluster zu bevorzugen. Für zukünftige Experimente könnte dies aber durch entsprechende Erweiterungen des MPA realisiert werden.

Zusammenfassung: Die Proteinidentifikation sowie die taxonomische und funktionelle Auswertung der Metaproteomdaten konnte durch die Entwicklung des MPAs und des Metaproteinkonzepts deutlich vereinfacht werden.

Korrelationen von mikrobiellen Gemeinschaften mit Prozessparametern

Das langfristige Ziel der Optimierung des Biogasanlagendesigns und -betriebs erfordert es die wesentlichen Zusammenhänge zwischen den mikrobiellen Gemeinschaften und den Prozessparametern zu verstehen. Als erster Schritt wurde die Langzeitstabilität der mikrobiellen Gemeinschaften in robust arbeitenden Biogasanlagen geprüft [Kapitel: 5.1]. Anschließend wurde der Einfluss der Prozesstemperatur auf die mikrobiellen Gemeinschaften am Beispiel einer mesophilen und einer thermophilen Biogasanlage untersucht [Kapitel: 5.2]. Aufbauend auf diese Untersuchung wurde die Korrelationen weiterer Prozessparameter mit mikrobiellen Gemeinschaften analysiert [Kapitel: 5.3].

Die einzelnen Untersuchungen erfolgten über den Verlauf der Promotion und gingen folglich einher mit verschiedenen Entwicklungsstadien des Workflows [Kapitel: 4]. Daher ist die Auflösung des Metaproteoms zwischen den Experimenten verschieden.

5.1 Langzeitstabilität von mikrobiellen Gemeinschaften in Biogasanlagen

⁶Um die Stabilität der mikrobiellen Gemeinschaften in Biogasanlagen zu untersuchen, wurden sechs Biogasanlagen über neun Monate regelmäßig beprobt, die Proteine extrahiert und 100 μg Proteinextrakt mittels SDS-PAGE aufgetrennt. Für eine bessere Visualisierung der Unterschiede in der SDS-PAGE wurde ein Clustergram basierend auf der Kreuzkorrelation der SDS-PAGE Profile [Kapitel: 3.6.3] und dem ungewichteten arithmetischen Mittel erstellt [Ehlers and Cloete (1999)] [Kapitel: 5.1.1]. Zusätzlich wurden für je vier Zeitpunkte die SDS-PAGE Profile der Biogasanlagen BGP 3 und BGP 4 in 10 Fraktionen geteilt und mittels AmaZon ETD ESI-Ion Trap MS/MS und einem 120 min Gradient analysiert [Kapitel: 5.1.2]. Die Proteinidentifikation erfolgte mit Mascot gegen UniProt/SwissProt und einem Mascot Score >40 .

5.1.1 SDS-PAGE Profile von sechs landwirtschaftlichen Biogasanlagen

In allen SDS-PAGE Spuren der verschiedenen Biogasanlagen konnte neben einem starkem Hintergrund mehrere markante Banden für ähnliche Molekulargewichte beobachtet werden [Abbildung; 5.1 A rechts]. Während diese Bandenmuster für verschiedene Zeitpunkte oder parallele Fermenter (BGP 5_1/ BGP 5_2) ein und derselben Biogasanlage gleich aussahen, gab es für verschiedene Biogasanlagen kleinere Unterschiede. Besonders ausgeprägt waren diese Unterschiede zwischen der thermophilen Biogasanlage BGP 5 und den restlichen mesophilen Biogasanlagen. Passend dazu bildeten die thermophile und die mesophilen Biogasanlagen im Clustergram der SDS-PAGE Profile zwei separate Äste [Abbildung; 5.1 A].

Auffällig bei der Langzeituntersuchung der Biogasanlage BGP 3 war [Abbildung: 5.1 B], dass für zwei Zeitpunkte der Biogasanlage BGP 3_1 und BGP 3_2 wesentliche Banden fehlten. Im Clustergram fielen diese beiden Zeitpunkte auch nicht zusammen mit den restlichen Zeitpunkten der Biogasanlage BGP 3. Interessanterweise konnte bei Betrachtung der Prozessdaten [Abbildung: 5.2] für den zweiten Zeitpunkt eine Übersäuerung der Biogasanlage von pH 8,01 auf pH 6,34 und ein Einbruch der Gasproduktion von 2.000 m³/d auf 940 m³/d festgestellt werden.

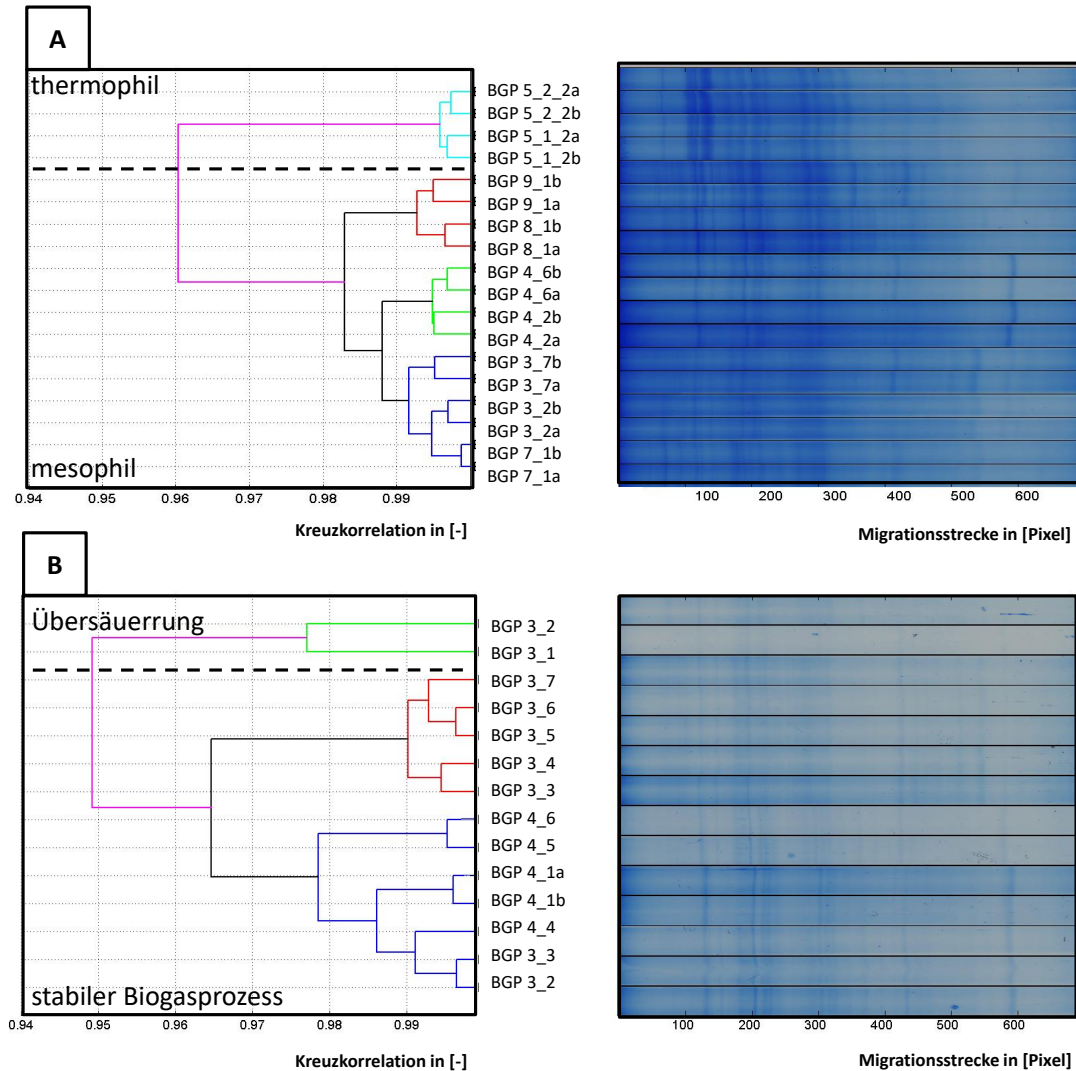


Abbildung 5.1: SDS-PAGE Profile von sechs verschiedenen Biogasanlagen.

In A sind sowohl die technischen als auch die biologischen Replikate aller sechs Biogasanlagen dargestellt. Dagegen zeigt B die Langzeituntersuchung der Biogasanlagen BGP 3 und BGP 4. Es wurden je 100 μg Proteinextrakt mittels SDS-PAGE aufgetrennt (Bild rechts, SDS-PAGE-Profile gedreht, Kontrast: +40%, Helligkeit: -40%) und die SDS-PAGEs mittels Clusteranalyse basierend auf der Kreuzkorrelation von deren Farbbintensitätsprofilen analysiert (Bild links).

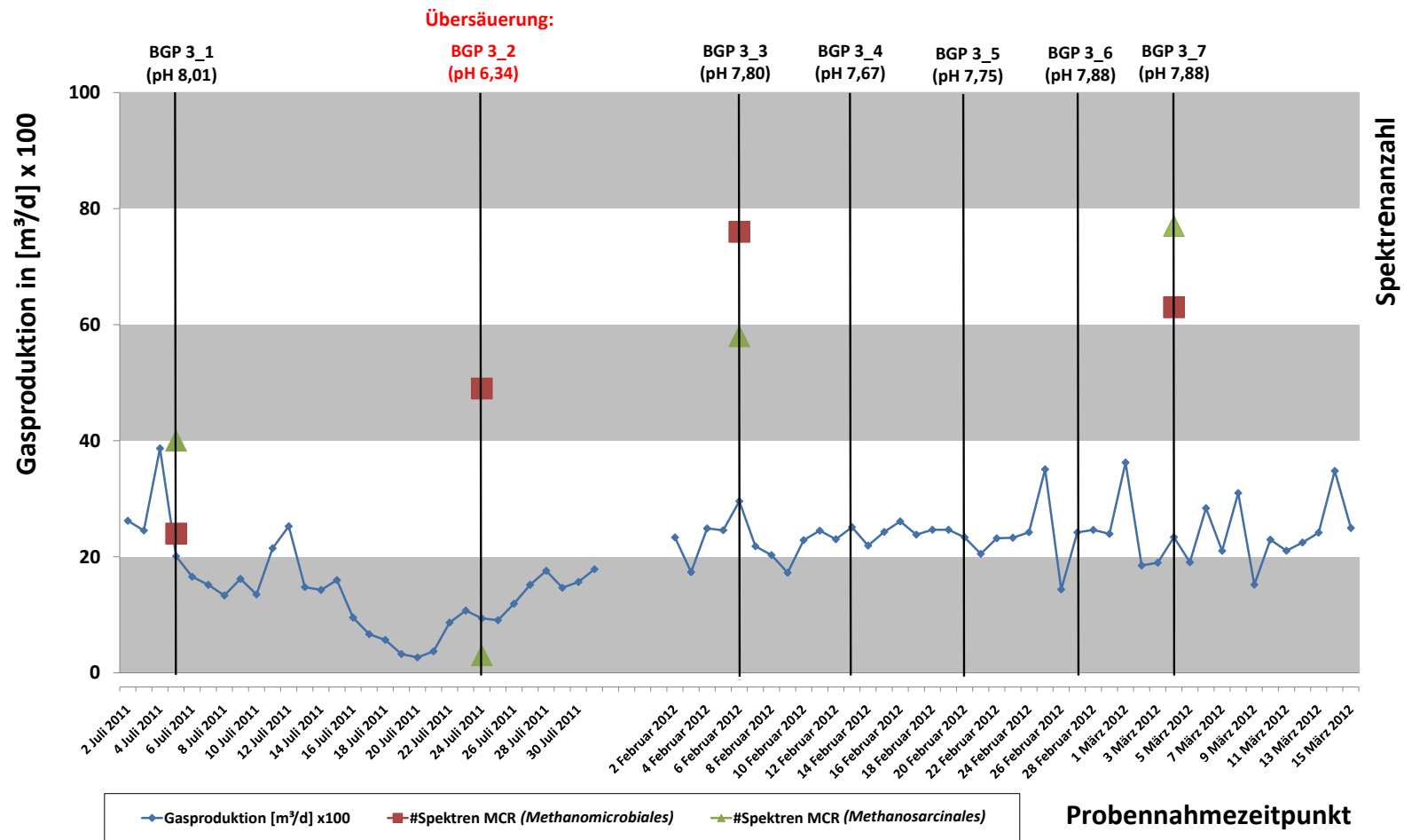


Abbildung 5.2: Prozessdaten Biogasanlage BGP 3

Dargestellt ist die Biogasproduktion über den gesamten untersuchten Zeitraum, die Probenahme, der pH-Wert und die Anzahl an Spektren, die für Methyl-Coenzym M Reduktase (MCR) identifiziert wurden.

5.1.2 Charakterisierung der mikrobiellen Taxonomien und Funktionen

Ausgehend von der hohen Ähnlichkeit der SDS-PAGE Profile wurden basierend auf der massenspektrometrischen Analyse die wesentlichen Taxonomien und Funktionen charakterisiert. Aufgrund limitierter MS-Messzeit erfolgte diese Untersuchung nur für die stabile Biogasanlage BGP 4 und die gestörte Biogasanlage BGP 3.

Insgesamt konnten dabei über 1.000 Proteine identifiziert werden, welche zum überwiegenden Teil von *Archaeen* und *Bakterien* stammten. Einige Proteine konnten aber auch *Pflanzen*, *Wirbeltieren* und *Pilzen* zugeordnet werden. Unter den *Bakterien* waren die dominierenden Klassen *Clostridia* (29%), *Gammaproteobacteria* (15%), *Actinobacteria* (10%), *Bacteroidia* (9%), *Bacilli* (7%) [Abbildung: 5.3 A]. Die meisten Proteine von *Archaeen* gehörten zu der Ordnung *Methanosarcinales* und einer unbekanntem Spezies. Entsprechend der MS-BLAST Suche von den identifizierten Proteinen ist diese unbekanntem Spezies am ehesten mit der hydrogenotrophen Spezies *Methanoculleus marisnigri* (90% Aminosäureidentität) verwandt.

Bei der Klassifizierung der Proteinfunktionen konnten die meisten identifizierten Proteine bei *Bakterien* dem Metabolismus (14%) und dem Transports (7%) zugeordnet werden [Abbildung: 5.3 B]. Diese Proteine dienten im Wesentlichen zur Zuckeraufnahme (ABC Zuckertransporter) und zum Zuckerabbau (Glyceraldehyde-3-phosphatdehydrogenase, Phosphopyruvathydratase, Phosphoglyceratkinase, Glycerolkinase). Ferner belegte die Identifikation von Peptidasen und entsprechenden Enzymen des Aminosäurestoffwechsels (Glutamatdehydrogenase, Aspartokinase, Dihydrodipicolinatereduktase) auch den Abbau von Proteinen durch die *Bakterien*.

Bezüglich der *Archaeen* konnten die meisten identifizierten Proteine der Methanogenese (34%) und der Energiekonservierung (6%) zugeordnet werden. Während es sich bei den meisten Proteinen der Energiekonservierung nur um verschiedene V-Typ ATPasen handelte, belegten die Methanogeneseenzyme sowohl die hydrogenotrophe, als auch die acetoclastische Methanogenese. Wesentliche Enzyme waren 5,10-Methanopterin S-Methyltransferase und Coenzyme F420-reduzierende Hydrogenase der Ordnung *Methanobacteriales* für den hydrogenotrophen Wege sowie Acetyl-CoA Decarboxylase/Synthase von der Ordnung *Methanosarcinales* für den acetoclastischen Weg. Das Vorhandensein der Enzyme Methanol Corrinoid Protein und Methanol Corrinoid Methyltransferase sowie Trimethylamine:Corrinoid Methyltransferase für die Ordnung *Methanosarcinales* belegte zusätzlich die Methanbildung aus Methanol [Sauer et al. (1997)] und Aminen [Ferguson and Krzycki (1997)]. Das häufigste identifizierte Methanogeneseenzym war die Methyl-Coenzym M Reduktase. Dieses katalysiert in allen methanogenen *Archaeen* den letzten Schritt von Methyl Coenzym M zu Methan.

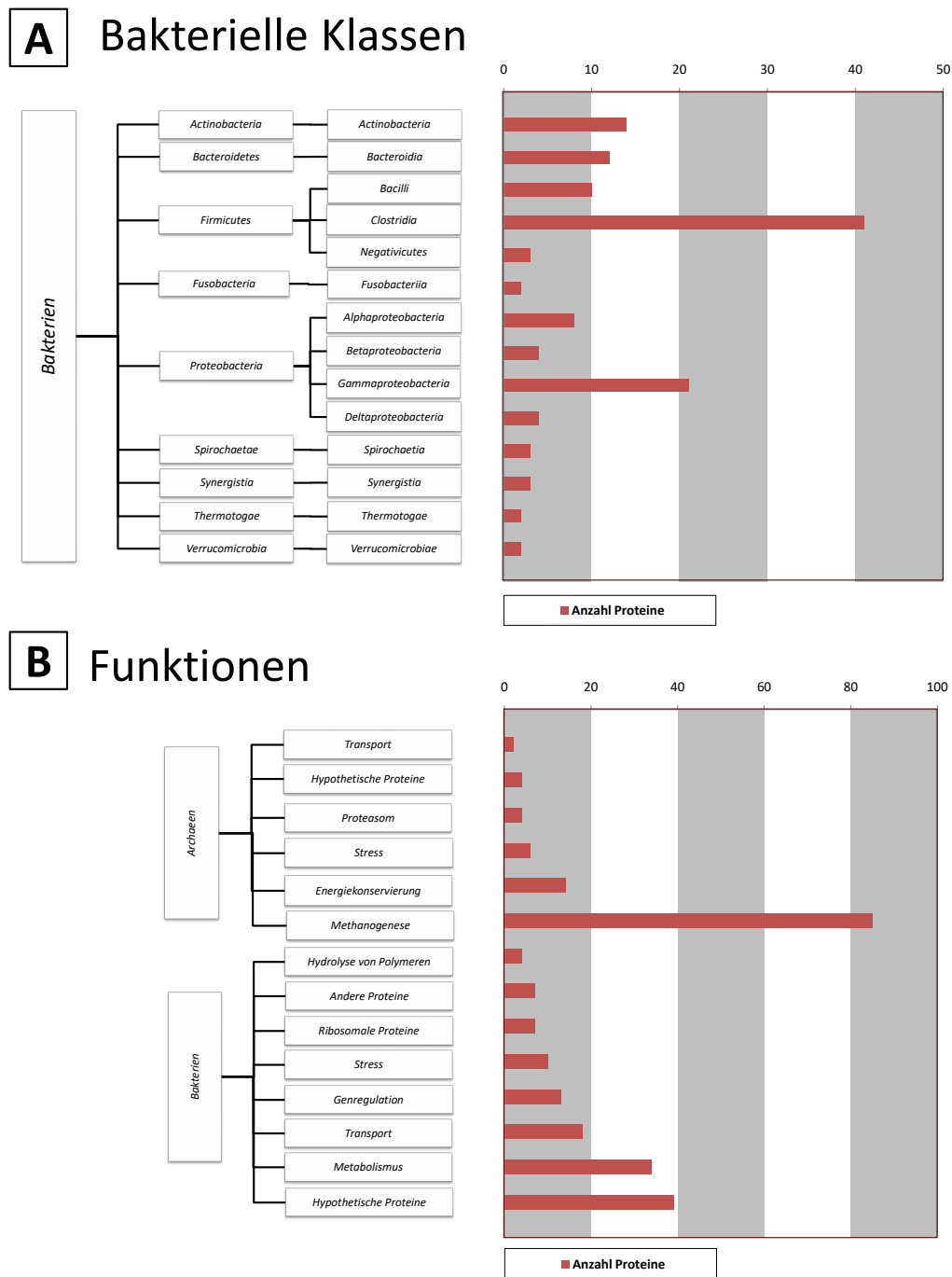


Abbildung 5.3: Profil der Taxonomien A und Funktionen B basierend auf allen identifizierten Proteinen. Es wurden dafür je 100 µg Proteinextrakt von vier Zeitpunkte der Biogasanlagen BGP 3 und BGP 4 mittels SDS-PAGE aufgetrennt. Anschließend wurde jede SDS-PAGE Spur in 10 Fraktionen geteilt, mittels AmaZon ETD ESI-Ion Trap MS/MS und einem 120 min Gradient analysiert und die Proteine durch Mascot-Suche gegen UniProt/SwissProt identifiziert (Mascot Score >40).

5.1.3 Diskussion

Die Stabilität der SDS-PAGE Muster der verschiedenen Biogasanlagen über die untersuchten Zeitpunkte [Abbildung; 5.1 B] zeigte eine hohe funktionelle Stabilität der mikrobiellen Gemeinschaften. Diese ging einher mit einer stetigen Biogasproduktion. Im Gegensatz dazu wurde in genetischen Arbeiten oft über die taxonomische Stabilität der mikrobiellen Gemeinschaften diskutiert [Lucas et al. (2015), Theuerl et al. (2015)]. Basierend auf den SDS-PAGE Profilen waren die größten funktionellen Unterschiede zwischen den mesophilen und der thermophilen Biogasanlage zu finden [Levén et al. (2007), Ziganshin et al. (2013)]. Alle sechs Biogasanlagen konnten darüber hinaus an ihrem SDS-PAGE Muster unterschieden werden, was auf kleinere funktionelle und taxonomische Unterschiede hindeutete. Wie die Biogasanlage BGP 3 zeigte, könnten Abweichungen vom anlagenspezifischen SDS-PAGE Muster ein Hinweis auf Prozessstörungen sein.

Die identifizierten Proteine belegten die wichtigsten Taxonomien und Funktionen des Biogasprozesses und deckten sich mit den Ergebnissen bisheriger Studien Abram et al. (2011), Hanreich et al. (2013). Das Fehlen von Cellulasen war allerdings auffällig, da diese essentiell für den Biomassenabbau sind. Vermutlich war die Anzahl an identifizierten Proteinen in dieser Metaproteomanalyse dafür nicht ausreichend. Eine andere mögliche Erklärung liefert die Arbeit von *Binner et al. (2011)* die zeigte, dass Cellulasen von den Mikroorganismen sehr schnell wieder abgebaut werden. Ein anderer wichtiger Punkt bezüglich des Biogasprozesses war der Aminosäureabbau. Mit der Glutamatdehydrogenase, Aspartokinase und Dihydrodipicolinatereduktase konnten sowohl Enzyme des Abbaus als auch der Synthese von Aminosäuren belegt werden. Insbesondere das Vorhandensein der letzten beiden lässt auf ein Mangel an Methionin, Lysin und Threonin schließen. Eine Erklärung dafür könnte ihr Mangel in pflanzlichen Substraten sein [Ridley et al. (2002)]. Neben der Betrachtung der Taxonomien und Funktionen im Allgemeinen war der Einfluss der Übersäuerung auf die mikrobielle Gemeinschaft relevant für den Biogasprozess. Wie in Abbildung 5.2 deutlich zu sehen ist, kam es während dieser Zeit zu einem massiven Einbruch der Spektrenanzahl für das Enzym Methyl-Coenzym M Reduktase von der Ordnung *Methanosarcinales*. Wahrscheinlich kam es durch den Abfall des pH-Wertes zum Absterben der *Methanosarcinen*, was auch schon von *Blume et al. (2010)* [Blume et al. (2010)] und *Munk et al. (2012)* [Munk et al. (2010)] beobachtet wurde. Dementsprechend könnte die Methyl-CoM Reduktase ein vielversprechendes Markerprotein für Störungen des Biogasprozesses sein.

Zusammenfassung: Die Proteinprofile der mikrobiellen Gemeinschaften in Biogasanlagen sind zeitstabil, sehr ähnlich aber dennoch anlagenspezifisch. Änderungen im Pro-

teilmuster können ein Hinweis auf geänderte Prozessbedingungen oder Prozessstörungen sein.

5.2 Vergleich einer mesophilen und einer thermophilen Biogasanlage

⁷Die Prozesstemperatur ist ein Parameter mit großem Einfluss auf die mikrobiellen Gemeinschaften in Biogasanlagen [Ziganshin et al. (2013), Levén et al. (2007)]. Im Gegensatz zu bisherigen genetischen Arbeiten [Ziganshin et al. (2013), Levén et al. (2007)] ermöglicht es die Untersuchung des Metaproteoms, neben taxonomischen auch funktionelle Unterschiede zu ermitteln. Die mikrobiellen Gemeinschaften der mesophilen Biogasanlage BGP 3 (43 °C) und der thermophilen Biogasanlage BGP 4 (52 °C) wurden mit einem 2D- und 3D-Ansatz und AmaZon ETD ESI-Ion Trap MS/MS analysiert [Kapitel: 4.1.3]. Die Proteinidentifikation erfolgte durch Mascot-Suche gegen die UniProtKB/Swiss-Prot Datenbank. Anschließend wurden die Ergebnisse mit einer 5% FDR-Schwelle in den MPA eingeladen und redundante Proteinidentifikationen bei einem überlappenden Peptid zu Metaproteinen gruppiert.

Auf diese Weise konnten für die mesophile und thermophile Biogasanlage 184 bzw. 383 Metaproteine für den 2D- und 757 bzw. 1663 Metaproteine für den 3D-Ansatz identifiziert werden. Basierend auf diesen Identifikationen erfolgte dann der Vergleich der Taxonomie- und Funktionsprofile [Kapitel: 5.2.1] sowie die Zuordnung der einzelnen Metaproteine zum Biogasprozess [Kapitel: 5.2.2].

5.2.1 Taxonomie- und Funktionsprofile

Das Vorhandensein von Schlüsseltaxa und -funktionen ist entscheidend für den stabilen Ablauf des Biogasprozesses [Kapitel: 2.2]. Basierend auf den kombinierten Ergebnissen des 2D- und 3D-Ansatzes konnten 18 verschiedene mikrobielle Ordnungen in der thermophilen und 13 in der mesophilen Biogasanlage [Abbildung: 5.4 A] identifiziert werden. Dabei zeigte das Taxonomieprofil der mesophilen Biogasanlage eine starke Dominanz der Ordnung *Methanosarcinales* und einige spezifische Ordnungen wie *Bacteriodales*, *Burkholderiales* und *Methanobacteriales*. Im Gegensatz dazu waren für die thermophile Biogasanlage die thermophilen *Bakterien* der Ordnungen *Thermoanaerobacteriales* und *Thermotogales* spezifisch sowie eine sehr große Häufigkeit der Ordnung *Clostridiales* und *Methanomicrobiales* charakteristisch. Ferner konnten in der thermophilen Biogasanlage mehrere nicht methanogene Archaeenordnungen wie *Thermococcales*, *Halobacteriales* und *Archaeoglobales* beobachtet werden.

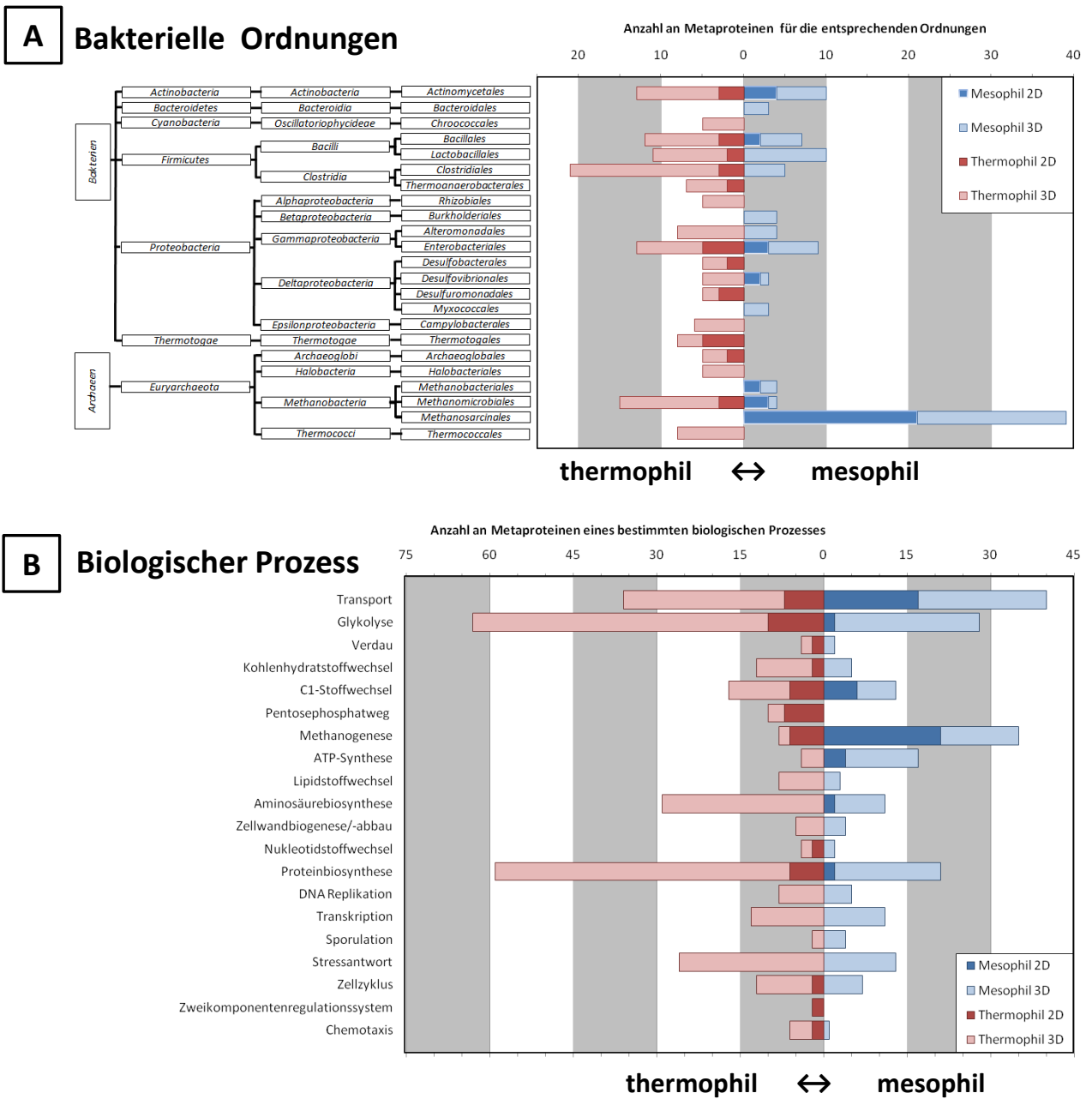


Abbildung 5.4: Taxonomie- und Funktionsprofile einer mesophilen (BGP 3: 43 °C) und einer thermophilen Biogasanlage (BGP 4: 52 °C), basierend auf identifizierten Metaproteinen des kombinierten 2D- und 3D-Ansatzes.

5.2.2 Zuordnung der identifizierten Proteine zum Biogasprozess

Für ein besseres Verständnis der real ablaufenden Prozesse in Biogasanlagen wurden die identifizierten Metaproteine des 2D- und des 3D-Ansatzes der mesophilen und der thermophilen Biogasanlageprobe den wesentlichen Schritten des Biogasprozesses zugeordnet [Tabelle: 5.1] [Anhang: 7.10].

Belegt werden konnte die Hydrolyse von Cellulose (EC: 3.2.1.91), Hemicellulose (EC: 3.2.1.155), Xylanen (EC: 3.2.1.8), Arabinogalaktan (EC: 3.2.1.55) und β -Galaktosiden (EC: 3.2.1.23).

Für die anschließende Gärung wurden die Enzyme für die Aufnahme von Glycerol-3-phosphat (EC: 3.6.3.20) und dem Mehrfachzucker Maltose (EC: 3.6.3.19) [Davidson and Chen (2004)] sowie die Glykolyse von Glukose-6-phosphat zu Phosphoenolpyruvat (EC: 2.7.1.11, EC: 4.1.2.13, EC: 1.2.1.12, EC: 1.2.1.59, EC: 2.7.2.3, EC: 5.4.2.1, EC: 4.2.1.11) identifiziert.

Weiterhin konnte mit den Enzymen NADP-abhängige Isopropanoldehydrogenase (EC: 1.1.1.80), Aldehyddehydrogenase (EC: 1.2.1.3) sowie Laktatdehydrogenase (EC: 1.1.1.27), Pyruvatflavodoxinoreduktase (EC: 1.2.7.-) [Kletzin and Adams (1996)] und Formiat-acetyltransferase (EC: 2.3.1.54) die weitere Gärung zu Alkohol und Milchsäure gezeigt werden. Umgekehrt können diese Enzyme bei kleinen Wasserstoffpartialdrücken auch für die Verwertung von Alkoholen und Milchsäuren genutzt werden. Passend dazu wurde in der mesophilen Biogasanlage das Enzym Tetrahydrofolatligase (EC: 6.3.4.3) gefunden, welches H_2 und CO_2 verbraucht, um Acetyl-CoA zu bilden [Lovell et al. (1990)].

Eine andere belegte Form der Wasserstoffnutzung war die hydrogenotrophe Methanogenese [Abbildung: 7.10]. So ermöglichen die Enzyme Methylentetrahydromethanopterinhydrogenase (EC: 1.5.99.9), Coenzym F420 Hydrogenase (EC: 1.12.98.1) und Coenzym F420-abhängige N₅,N₁₀-Methenyltetrahydromethanopterinreduktase (EC: 1.5.99.11) die Bildung von CH_4 aus H_2 und CO_2 . Zusätzlich zu Proteinen für den Wasserstoffweg konnten exklusive für die mesophile Biogasanlage Proteine für die Bildung von CH_4 aus Acetat (EC: 2.7.2.1, 1.2.99.2, 2.3.-), Aminen (EC: 2.1.1.249, 2.1.1.250) und Methanol (EC: 2.1.1.90) durch die Ordnung *Methanosarcinales* identifiziert werden [Kendall and Boone (2006)]. Weitere Methanogenesenzyme waren die Tetrahydromethanopterin S-Methyltransferase (EC: 2.1.1.86), Methyl CoM Reduktase M (EC: 2.8.4.1) sowie die energiekonservierende V-type ATPase (EC: 3.6.3.14).

5.2 Vergleich einer mesophilen und einer thermophilen Biogasanlage

Tabelle 5.1: *Identifizierte Metaproteine für die wesentlichen Abbaupfade des Biogasprozesses.*

| Protein | EC-number | BGP 3 | BGP 4 |
|--|------------------|--------------|--------------|
| Hydrolyse | | | |
| Cellulose-1,4- β -cellobiosidase | 3.2.1.91 | - | x |
| Xyloglukanase | 3.2.1.155 | x | - |
| Endo-1,4- β -xylanase | 3.2.1.8 | x | x |
| Intrazelluläre Exo- α -(1->5)-L-arabinofuranosidase | 3.2.1.55 | - | x |
| β -Galaktosidase | 3.2.1.23 | - | x |
| Transport | | | |
| Glycerol-3-phosphat-transportierende ATPase | 3.6.3.20 | x | x |
| Maltose-transportierende ATPase | 3.6.3.19 | x | x |
| Gärung | | | |
| Glukose-6-phosphatisomerase | 5.3.1.9 | - | x |
| Fruktose-6-phosphatkinase | 2.7.1.11 | - | x |
| Fruktose-1,6-bisphosphataldolase | 4.1.2.13 | x | x |
| Glycerinaldehyd-3-phosphatdehydrogenase | 1.2.1.12 | x | x |
| Phosphoglyceratkinase | 2.7.2.3 | x | x |
| Phosphoglyceratmutase | 5.4.2.1 | x | x |
| Phosphopyruvathydratase | 4.2.1.11 | x | x |
| NADP-abhängige Isopropanoldehydrogenase | 1.1.1.80 | - | x |
| Aldehyddehydrogenase | 1.2.1.3 | x | - |
| Laktatdehydrogenase | 1.1.1.27 | x | x |
| Pyruvatflavodoxinoxidoreduktase | 1.2.7.- | x | x |
| Formiatacetyltransferase | 2.3.1.54 | - | x |
| Tetrahydrofolatligase | 6.3.4.3 | - | x |
| Methanogenese | | | |
| Methylenetetrahydromethanopterindehydrogenase | 1.5.99.9 | - | x |
| Coenzyme F420 Hydrogenase | 1.12.98.1 | x | x |
| N5,N10-Methenyltetrahydromethanopterinreduktase | 1.5.99.11 | x | - |

| Protein | EC-number | BGP 3 | BGP 4 |
|---|-----------|-------|-------|
| Acetatkinase | 2.7.2.1 | x | - |
| Kohlenstoffmonoxiddehydrogenase | 1.2.99.2 | x | - |
| Acetyl-CoA Decarboxylase/Synthase-Komplex | 2.3.-.- | x | - |
| Di-/Trimethylamine-corrinoidprotein | 2.1.1.249 | x | - |
| Methanol:5-hydroxybenzimidazolylcobamide Co-methyltransferase | 2.1.1.90 | x | - |
| Methanopterin S-Methyltransferase | 2.1.1.86 | x | x |
| Methyl CoM Reduktase M | 2.8.4.1 | x | x |
| CoB-CoM Heterodisulfidreduktase | 1.8.98.1 | x | x |
| V-Typ ATPase | 3.6.3.14 | x | x |

5.2.3 Diskussion

Die beim Vergleich der taxonomischen Zusammensetzung beobachteten Unterschiede zwischen der mesophilen und thermophilen Biogasanlage deckten sich im Wesentlichen mit den Ergebnissen vorheriger Studien. So ist bekannt, dass die Ordnung *Methanosarcinales*, *Bacteroidales* und *Burkholderiales* nicht in thermophilen Biogasanlagen vorkommen [Pender et al. (2004), Krakat et al. (2010), Gunnigle et al. (2015)]. Hingegen treten die Ordnungen *Clostridiales* und *Thermoanaerobacteriales* in diesen vermehrt auf [Levén et al. (2007)]. Die höhere Anzahl an gefundenen taxonomischen Ordnungen in der thermophilen Biogasanlage widerspricht dagegen der Arbeit von *Leven et al. (2007)* [Levén et al. (2007)]. Die Auflösung der mikrobiellen Gemeinschaft in der thermophilen Biogasanlage mittels Metaproteomanalyse war mit doppelt so viel identifizierten Spektren allerdings deutlich höher als für die mesophile Biogasanlage.

Auffällig in den Taxonomieprofilen war auch, dass für beide Biogasanlagen etwa 50% der identifizierten Spektren *Archaeen* zugeordnet werden konnten. Im Gegensatz dazu beobachteten Genom-basierte Studien meist nur geringe Anteile an *Archaeen* (<4%) [Hanreich et al. (2013), Nettmann et al. (2010)]. Da dieser Anteil für vorhandene mathematische Prozessmodelle des Biogasprozesses ein wichtiger Parameter ist [Batstone et al. (2002), Lauwers et al. (2013)], wären weitere Untersuchungen mit vorheriger Validierung der Methoden für die Metagenom- bzw. Metaproteomuntersuchungen an definierten Mischkulturen [Stark et al. (2014)] wünschenswert.

Während die Taxonomieprofile der mesophilen und der thermophilen Biogasanlage größere Unterschiede aufwiesen, waren die Profile der Metaproteinfunktionen basierend auf dem UniProt Keyword „Biological Process“ [UniProt Consortium (2012)] sehr ähnlich [Abbil-

dung:5.4 B]. Für beide Anlagen wurden diese von Metaproteinen für den Transport, die Glykolyse, die Methanogenese, die Proteinbiosynthese und die Stressantwort dominiert. Kleinere Unterschiede waren die erhöhte Häufigkeit von Methanogenesemetaproteinen in der mesophilen Biogasanlage und mehr Glykolyse- und Proteinbiosynthesemetaproteinen in der thermophilen Biogasanlage. Außerdem wurden nur in der thermophilen Biogasanlage Metaproteine des Pentosephosphatweges nachgewiesen.

Im Grunde war die hohe funktionelle Ähnlichkeit der mikrobiellen Gemeinschaften in mesophilen und thermophilen Biogasanlagen wenig überraschend, da in beiden kontinuierlich CH_4 produziert wurde und folglich die mikrobiellen Gemeinschaften funktionell ähnliche Proteine benötigten.

Die anschließende Zuordnung der einzelnen Metaproteine aus dem kombinierten 3D-Ansatz zu den einzelnen Schritten des Biogasprozesses bestätigte die angenommenen Abbauwege der Biomasse zu CH_4 [Madigan and Martinko (2009), Seite 755]. Die aus der thermodynamischen Betrachtung [Schink (1997), Hattori (2008)] und aus genetischen Arbeiten [Levén et al. (2007)] vermutete Dominanz der hydrogenotrophen Methanbildung in thermophilen Biogasanlagen konnte auch auf Proteinlevel beobachtet werden. In der mesophilen Biogasanlage erfolgte die Methanbildung auch aus Acetat, Methanol und Aminen. Nicht gezeigt werden konnte die Umwandlung von Wasserstoff und Kohlenstoffdioxid zu Acetat (Homoacetogenese) bzw. deren Rückreaktion (syntrophe Acetatoxidation) [Kapitel: 2.2]. Trotzdem zeigte diese Untersuchung den Wert des 3D-Ansatzes für die umfassende taxonomische und funktionelle Charakterisierung der mikrobiellen Gemeinschaften in Biogasanlagen. Leider wurden für jede Probe 120 einzelne MS/MS Messung benötigt. Dies entsprach etwa 13 Tagen MS-Messzeit, weswegen nicht mehr Biogasanlagen mit dieser Methode untersucht werden konnten.

Zusammenfassung: Mittels 3D-Ansatz konnten einzelne Schritte des Biogasprozesses in sehr hoher Auflösung untersucht werden. Während in thermophilen Biogasanlagen die Methanbildung hauptsächlich aus Wasserstoff und Kohlenstoffdioxid zu erfolgen scheint, werden in mesophilen Biogasanlagen auch Acetat, Methanol und Amine genutzt. Markertaxa für die thermophile Biogasanlage waren die Ordnungen *Thermoanaerobacteriales* und *Thermotogales* und die Ordnungen *Bacteriales*, *Burkholderiales* und *Methanobacteriales* für die mesophile Biogasanlage.

5.3 Charakterisierung von mikrobiellen Gemeinschaften von 40 Biogasanlagen

Biogasanlagen unterscheiden sich in einer großen Anzahl an Prozessparametern, welche die mikrobiellen Gemeinschaften beeinflussen. Folglich muss eine hohe Anzahl an verschiedenen Biogasanlagen untersucht werden, um statistisch glaubwürdige Aussagen über die Zusammenhänge zwischen den Prozessparametern sowie der taxonomischen und funktionellen Zusammensetzung der mikrobiellen Gemeinschaften zu treffen. Dementsprechend wurde in diesem Experiment die taxonomische und funktionelle Zusammensetzung von 40 Biogasanlagenproben [Anhang: 7.6] analysiert und mit den Prozessparametern korreliert. Die Biogasanlagen unterschieden sich in wesentlichen Prozessparametern, wie dem Fermentertyp (CSTR/ UASB), in der Prozesstemperatur (33-55 °C), der Verweilzeit (0,38-150 d), dem pH-Wert (7,1-8,6) und in den Substraten (landwirtschaftliche Abfälle, Klärschlamm, Hausabfälle und industrielle Abfälle) [Anhang: 7.6, Anhang: CD „Prozessdaten 40 Biogasanlagen“]. Die Vorteile dieser Biogasanlagen waren die bereits durchgeführte Charakterisierung der chemischen und technischen Prozessparameter sowie die Bestimmung der taxonomischen Zusammensetzung basierend auf genetischer Methoden durch die Arbeit von *de Vrieze et al. (2015)*. Folglich konnte in dieser Studie nicht nur der Zusammenhang zwischen den Prozessparametern und der taxonomischen und funktionellen Zusammensetzung der mikrobiellen Gemeinschaft in Biogasanlage untersucht werden, sondern auch, ob Metaproteomanalysen bei gleichem Aufwand mehr bzw. hochwertigere Informationen liefern als genetische Methoden.

Die Untersuchung der 40 Biogasanlagenproben erfolgte anstelle des aufwendigeren 2D/3D-Ansatzes mittels 1D-Ansatz und hochauflösendem Elite Hybrid Ion Trap-Orbitrap MS/MS um den entsprechenden Probendurchsatz zu erreichen. Nach der Kontrolle der Phenolextraktion mittels SDS-PAGE [Abbildung: 5.5] wurden 200 µg Proteinextrakt durch Einlaufen in eine SDS-PAGE gereinigt, tryptisch verdaut und als Triplikat am Elite Hybrid Ion Trap-Orbitrap MS/MS gemessen. Für die Proteinidentifikation mittels Datenbanksuche wurde Mascot genutzt und gegen eine kombinierte Datenbank aus UniProt/Swiss-Prot und diverser Metagenome [Schlüter et al. (2008), Hanreich et al. (2013), Rademacher et al. (2012), Stolze et al. (2016)] gesucht. Abschließend wurden die Ergebnisse in den MPA mit einer FDR von 1% eingeladen, unbekannte Proteineinträge aus den Metagenomen durch BLAST Suche gegen UniProt/Swiss-Prot annotiert und die Proteineinträge nach UniRef50 gruppiert.

Nach der Überprüfung der Proteinextraktion und -identifikation [Kapitel: 5.3.1] erfolgte die Untersuchung der mikrobiellen Gemeinschaften. Dafür wurden die mikrobiellen Gemeinschaften zuerst mittels Clusteranalyse gruppiert [Kapitel: 5.3.2], anschließend die

taxonomische [Kapitel: 5.3.3, 5.3.4] und die funktionelle [Kapitel: 5.3.5] Zusammensetzung charakterisiert sowie Zusammenhänge zwischen dieser und den Prozessparametern identifiziert [Kapitel: 5.3.6].

5.3.1 Beurteilung der Proteinextrakte und der Proteinidentifikationen

Für Metaproteomanalysen sind die erfolgreiche Extraktion der Proteine und deren Identifikation essentiell. Die Trennung der extrahierten Proteine mittels SDS-PAGE zeigte, dass für alle Biogasanlagen Proteine im untersuchten Molekulargewichtsbereich von 10-180 kDa extrahiert werden konnten [Abbildung: 5.5]. Es gab einzelne dominierende Banden, die in allen Proben beobachtet wurden und spezifische Banden für bestimmte Biogasanlagen. Eine Ausnahme stellten drei Proben dar, in denen in einem Replikat (zweites Replikat nicht gezeigt) das SDS-PAGE Profil verschmiert (Gent 15, Gent 29) bzw. sehr schwach war (Gent 37).

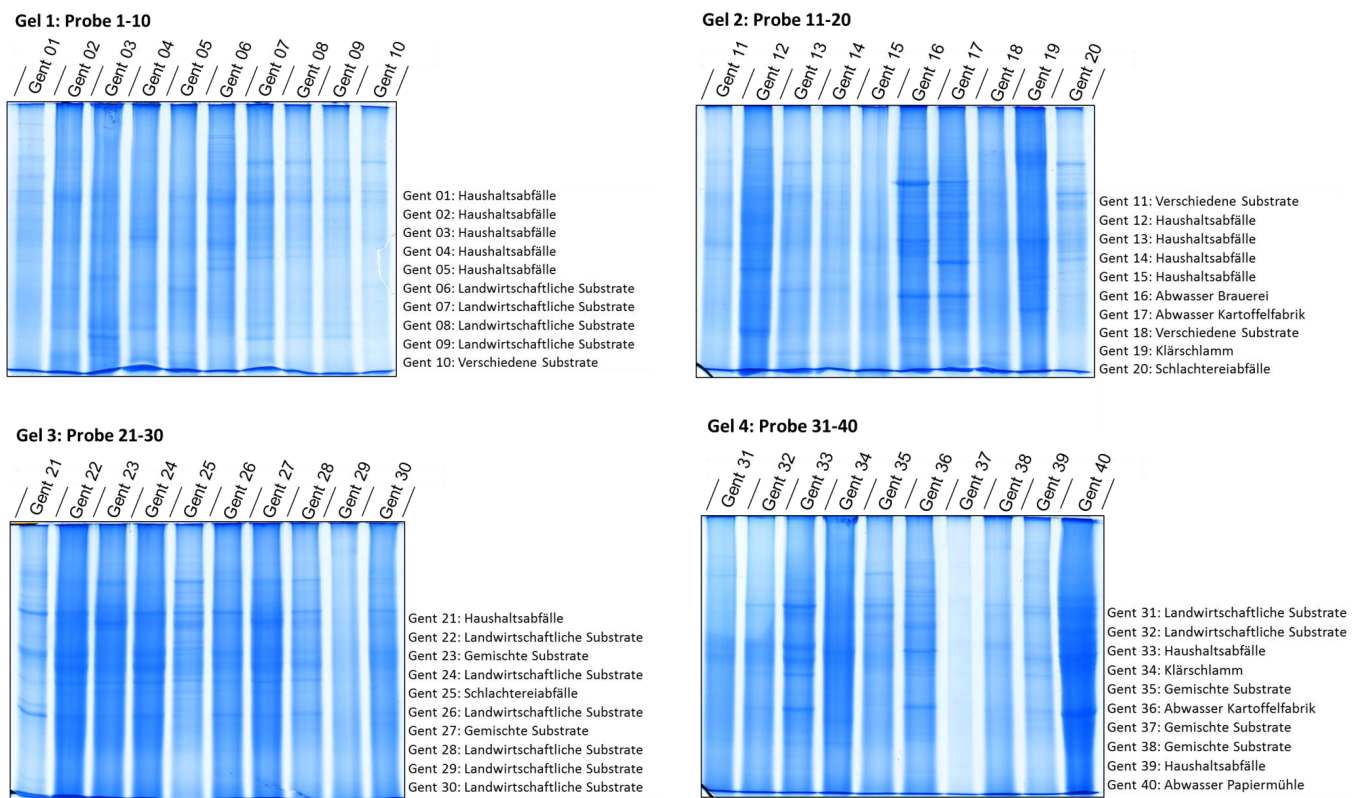


Abbildung 5.5: SDS-PAGE Profile von 40 Biogasanlagen. Die Proteine wurden mittels Phenolextraktion gereinigt, mittels Amidoschwarz-Assay quantifiziert und je 200 μ g Proteinextrakt pro Spur für die SDS-PAGE verwendet.

Die massenspektrometrische Untersuchung zeigte sowohl Schwankungen in der Anzahl an identifizierten Spektren als auch von Metaproteinen. So konnten aus jeweils ca. 40.000

gemessenen Spektren zwischen 389-4.565 Spektren [Abbildung: 4.8] bzw. 100-859 Metaproteine identifiziert werden. Eine detaillierte Betrachtung der unterschiedlichen Anzahl an Identifikationen kann Kapitel: 4.3.2 entnommen werden. Für den Vergleich der mikrobiellen Gemeinschaften der verschiedenen Biogasanlagenproben wurden die Ergebnisse für alle Proben auf Ebene der taxonomischen Ordnung (75 Ordnungen), der biologischen Prozesse (162 biologische Prozesse) und der Metaproteine (3138 Metaproteine) zusammengefasst [Anhang: CD „Metaproteine + Ordnungen + Biologische Prozesse 40 Biogasanlagen“]. Um den Einfluss von methodischen Schwankungen auf die Dateninterpretation [Tabb et al. (2010)] auszuschließen, wurden für die anschließende statistische Auswertung nur die Ergebnisse berücksichtigt, die in mindestens einer Probe mindestens 1% der Spektren repräsentierten [Anhang: 7.14 „Schwellenwerte für Metaproteine + Ordnungen + Biologische Prozesse“]. Die gefilterten Daten umfassten noch 35 Ordnungen, 65 biologische Prozesse (40%) und 187 Metaproteine (6%). Dies entsprach 46% der Ordnungen, 40% der biologische Prozesse und 6% der Metaproteine, repräsentierten aber immer noch über 65-99% der identifizierten Spektren [Abbildung: 5.6].

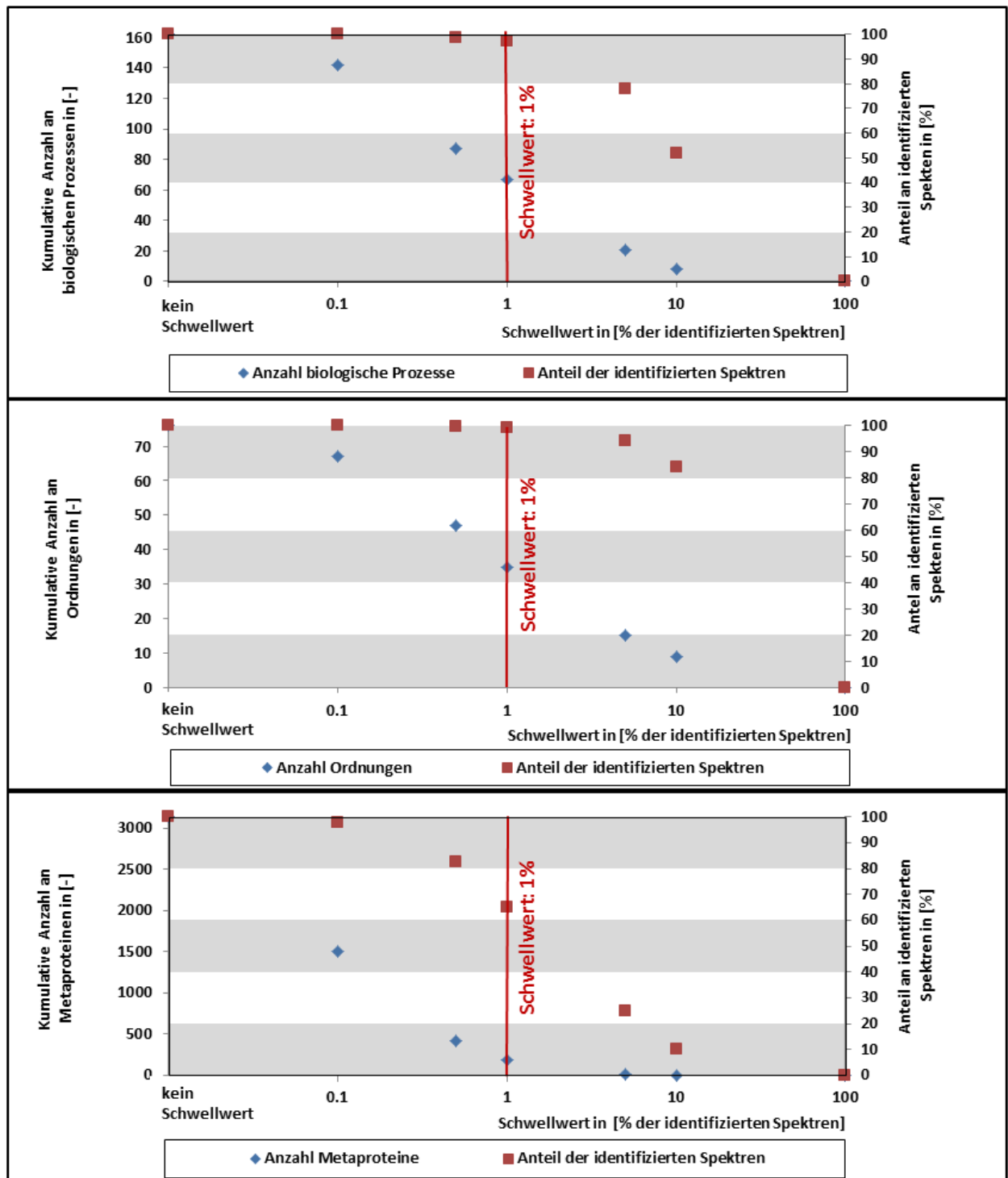


Abbildung 5.6: Schwellwerte für Metaproteine, taxonomischen Ordnungen und biologischen Prozesse.

Dargestellt ist die kumulative Summe an identifizierten Metaproteinen, taxonomischen Ordnungen bzw. biologischen Prozessen und der Anteil der identifizierten Spektren, welche diese Gruppen repräsentieren.

5.3.2 Clusteranalyse und Hauptkomponentenanalyse von verschiedenen mikrobiellen Gemeinschaften

Ein Ziel dieser Untersuchungen war es, zu prüfen, ob die Gruppierung der Biogasanlagen basierend auf den Metaproteomdaten ähnliche Ergebnisse liefert wie die Analyse der genetischen Daten von *De Vrieze et al. (2015)*. Basierend auf identifizierten Metaproteinen wurden dafür Clusteranalysen [Abbildung: 5.7] durchgeführt sowie PCAs basierend auf den taxonomischen Ordnungen, biologischen Prozessen und Metaproteinen [Abbildung: 5.8] [Anhang CD: 7.14 Faktorkoordinaten der Hauptkomponentenanalysen]. Triplikattmessungen der einzelnen Proben und unabhängige Proben von parallel arbeitenden Fermentern (SEH: Gent 20/ Gent 22) fielen als erstes bei der Clusteranalyse zusammen [Abbildung: 5.7]. Eine Ausnahme stellen die Biogasanlagen Gent 06 und Gent 07 dar, welche mit so ähnlichen Prozessparametern betrieben wurden, dass die Replikate beider Anlagen einen gemeinsamen Ast bildeten.

Eine hohe Ähnlichkeit konnte auch für verschiedene Zeitpunkte gleicher Biogasanlage (VCE, BCI, BOEYE, DRZ) festgestellt werden. Eine spezifische Gruppierung der Zeitpunkte bei der Clusteranalyse, wie von *Heyer et al. (2013)* oder *Werner et al. (2011)* beschrieben, wurde allerdings nicht beobachtet. Eventuell haben die stark unterschiedlichen Prozessbedingungen der einzelnen Zeitpunkte zu einer größeren Trennung geführt. Beispielsweise unterschieden sich die Mengen an kurzkettigen Fettsäuren für die Biogasanlagen VCE (Gent 15: 5.700 mg_{COD}/L, Gent 29: 0 mg_{COD}/L) und BCI (Gent 33: 5.600 mg_{COD}/L, Gent 39: 22.600 mg_{COD}/L) deutlich.

Besonders markant im Clustergram waren fünf große Hauptäste [Abbildung: 5.7]. Die manuelle Verknüpfung dieser mit den Prozessparametern zeigte, dass diese in Verbindung mit der Prozesstemperatur (mesophile/ thermophile), dem Reaktordesign (CSTR/ UASB-Reaktoren) und dem Substrat (z. B. Klärschlamm) standen [Abbildung: 5.7].

Anschließend wurde geprüft, ob die gefundenen Gruppen auch in den PCAs basierend auf den taxonomischen Ordnungen, den biologischen Prozessen und den Metaproteinen vorhanden waren [Abbildung: 5.8]. Die ersten zwei Hauptkomponenten dieser drei PCA erklärten zwar analog zur Arbeit von *De Vrieze et al. (2015)* (28,3% erklärte Varianz für die PCA des taxonomischen Profils) nur 25%, 33% bzw. 28% der Varianz, bestätigten aber die Ähnlichkeit der Triplikate und die vorgeschlagenen Cluster.

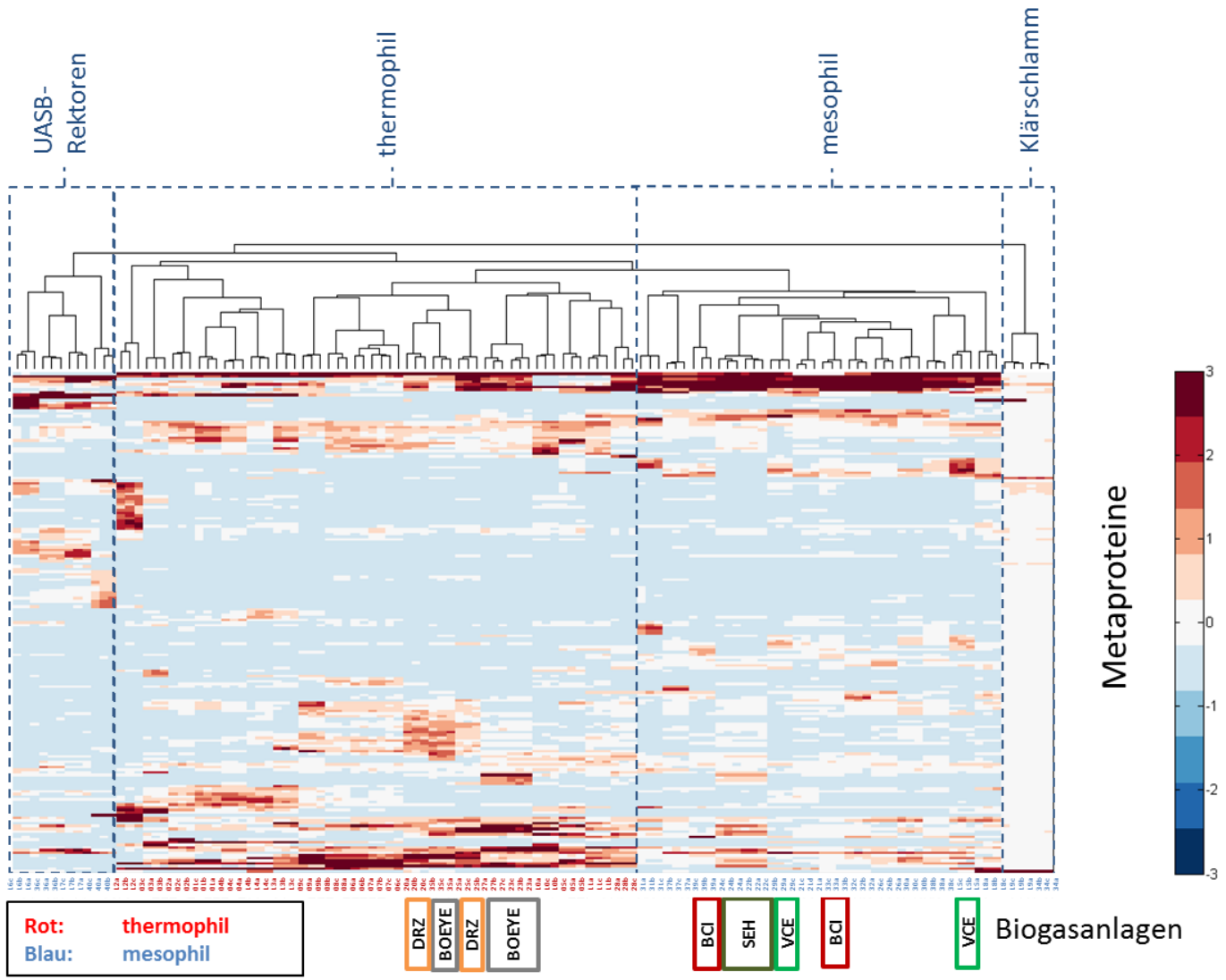


Abbildung 5.7: Clustergram von mikrobiellen Gemeinschaften der 40 Biogasanlagen basierend auf den identifizierten Metaproteinen. Die farbige markierten Boxen zeigen Proben von den gleichen Biogasanlagen zu verschiedenen Zeitpunkten (dunkelgrün: SEH (parallele Fermenter) [Gent 22 (2012-04-04), Gent 24 (2012-04-04)]; grün: VCE [Gent 15 (2011-10-10), Gent 29 (2012-04-10)]; grau: BOEYE [Gent 23 (2012-04-14), 27 (2012-04-11), Gent 35 (2011-09-29)]; orange: DRZ [Gent 20 (2011-09-29), Gent 25 (2012-04-04)]; rot: BCI [Gent 33 (2012-04-11) + Gent 39 (2011-09-29)].

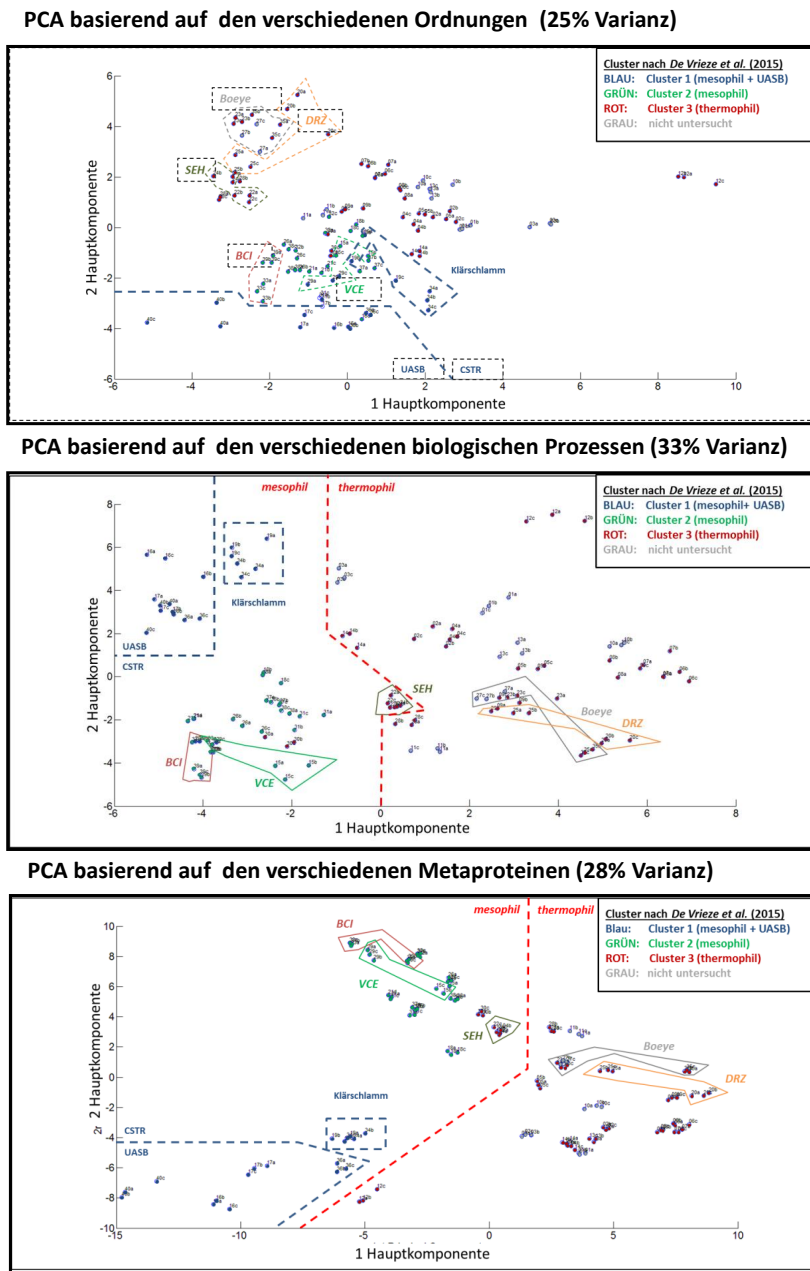


Abbildung 5.8: Hauptkomponentenanalyse der mikrobiellen Gemeinschaften von 40 Biogasanlagen basierend auf den identifizierten Ordnungen, biologischen Prozessen und Metaproteinen. Die farblich markierten Boxen zeigen Proben von den gleichen Biogasanlagen zu verschiedenen Zeitpunkten (dunkelgrün: SEH (parallele Fermenter) [Gent 22 (2012-04-04), Gent 24 (2012-04-04)]; grün: VCE [Gent 15 (2011-10-10), Gent 29 (2012-04-10)]; grau: BOEYE [Gent 23 (2012-04-14), 27 (2012-04-11), Gent 35 (2011-09-29)]; orange: DRZ [Gent 20 (2011-09-29), Gent 25 (2012-04-04)]; rot: BCI [Gent 33 (2012-04-11) + Gent 39 (2011-09-29)]. Zusätzlich wurden die Cluster aus der Arbeit von De Vrieze et al. (2015) als rote, blaue und grüne Punkte markiert.

Die Auswertung der Faktorkoordinaten der Hauptkomponentenanalyse [Anhang: CD „Faktorkoordinaten der Hauptkomponentenanalysen“] zeigte, dass in den mesophilen Biogasanlagen vermehrt Proteine der Methanogenese und des Transportes gefunden werden. Im Gegensatz dazu konnten in der thermophilen Biogasanlage mehr Proteine für die Aminosäurebiosynthese, die DNA-Rekombination sowie die DNA Reparatur identifiziert werden.

5.3.3 Untersuchung der ökologischen Indizes Richness und Evenness

Das Zählen der Arten (Richness) und die Betrachtung deren Gleichverteilung (Evenness) [Marzorati et al. (2008), Wittebolle et al. (2009), Mertens et al. (2005)] ist eine einfache Möglichkeit, die taxonomische Zusammensetzung in einer Biogasanlage zu beschreiben. Analog zur Bestimmung dieser Indizes aus genetischen Sequenzdaten werden diese hier erstmalig basierend auf Metaproteomdaten berechnet. Als taxonomische Ebene wurde dafür die Ordnungsebene gewählt und nur Ordnungen betrachtet, deren Spektrenanzahl mindestens 1% der Spektren in mindestens einer Biogasanlage [Marzorati et al. (2008)] repräsentierten.

Für die untersuchten Biogasanlagen variierte die Richness zwischen 18 und 35 mikrobiellen Ordnungen und die Evenness zwischen 69 und 87 (Gini-Koeffizient x 100) [Abbildung: 5.9]. Die Indizes zeigten eine starke Abhängigkeit von der Anzahl an identifizierten Spektren und vom verwendeten Schwellenwert [Anhang: 7.11]. Erst ab etwa 3.000 identifizierten Spektren konnte keine weitere Zunahme der Richness durch mehr Identifikationen beobachtet werden.

Der Vergleich verschiedener Zeitpunkte der Biogasanlagen SHE (Gent 22 Richness/Gini-Index: 34/70; Gent 24 Richness/Gini-Index: 35/69) und DRZ (Gent 20 Richness/Gini-Index: 33/73; Gent 25 Richness/ Gini-Index: 34/73) zeigte sehr ähnliche Werte für die Richness und die Evenness. Im Gegensatz dazu gab es größere Unterschiede für die Biogasanlagen BOEYE, BCI und VCE, wie auch schon in der Arbeit von *De Vrieze et al. (2015)* beobachtet. Für die einzelnen Zeitpunkte letzterer Biogasanlagen konnten aber auch größeren Unterschiede in der Prozesstemperatur und in der Menge an kurzkettigen Fettsäuren in der Gärflüssigkeit beobachtet werden. Beispielsweise nahm für die Biogasanlage VCE (Gent 15/Gent 29) die Prozesstemperatur von 38 °C auf 34 °C und die Menge von kurzkettigen Fettsäuren von 5735 mg_{COD}/L auf 0 mg_{COD}/L) ab. Allgemein konnte über alle Biogasanlagen der Trend beobachtet werden, dass bei höherer Prozesstemperatur die Richness etwas höher lag.

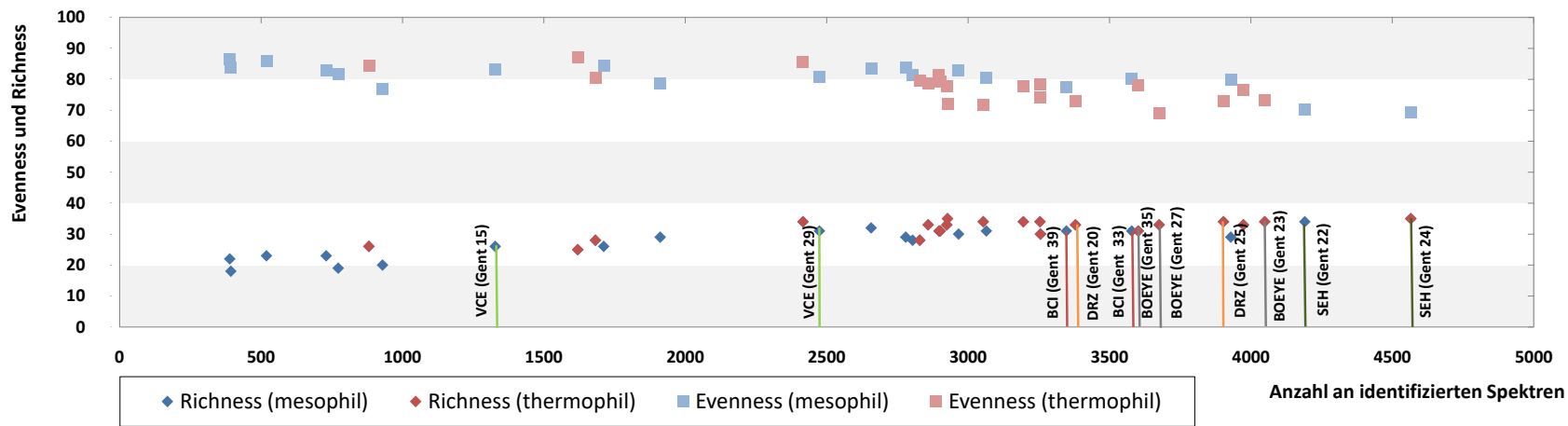


Abbildung 5.9: *Richness und Evenness der untersuchten Biogasanlagen.* Hervorgehoben wurde dabei die Richness der mesophilen und thermophilen Biogasanlagen sowie mehrfache Zeitpunkte der gleichen Biogasanlage (Balken). Um in dieser Abbildung die Abhängigkeit der Richness und Evenness von der Anzahl der identifizierten Spektren zu verdeutlichen, wurden diese in Abhängigkeit von der Anzahl identifizierter Spektren aufgetragen.

5.3.4 Taxonomische Zusammensetzung von mikrobiellen Gemeinschaften

Basierend auf den 40 verschiedenen Biogasanlagenproben sollte geprüft werden, welche taxonomischen Gruppen essentiell für den erfolgreichen Abbau der Biomasse zu CH₄ waren (*core microbiome* [Shade and Handelsman (2012)]. Zusätzlich sollten die Ergebnisse mit den 16S rRNA Gen-basierten Daten von *De Vrieze et al. (2015)* verglichen werden, um die taxonomische Aussagekraft von Metaproteomanalysen zu überprüfen. Als Vergleichskriterium wurde dafür der gegenüber Ausreißern robustere Median der jeweiligen Taxa gewählt.

Auf Ebene der Domänen konnten 34,2% der identifizierten Spektren den *Archaeen* und 67,8% den *Bakterien* zugeordnet werden. Der Archaeenanteil variierte allerdings für einzelne Biogasanlagen zwischen einem Minimum von 11,9% archaeelle Spektren in Biogasanlage Gent 05 und einem Maximum von 77,7% in Biogasanlage Gent 16. Im Vergleich dazu beobachteten *De Vrieze et al. (2015)* für die gleichen Proben mittels Real-Time PCR des 16S rRNA Gens Archaeenanteile von 0,2% bis 48% mit einem Median von 1%.

Um ein detaillierteren Einblick in die Schlüsseltaxa der mikrobiellen Gemeinschaften zu gewinnen, wurden zusätzlich Taxonomieprofile auf Ordnungsebene erstellt [Abbildung: 5.10]. Dominierend in den Metaproteomanalysen sind die Bakterienordnungen *Bacillales* (Median 23,8%), *Enterobacteriales* (Median 11,1%), *Bacteriodales* (Median 11,00%) und *Clostridiales* (Median 5,1%) sowie die Archaeenordnungen *Methanobacteriales* (Median 38,3%), *Methanosarcinales* (Median 30,1%) und *Methanococcales* (Median 8,4%). Im Vergleich dazu waren in der Arbeit von *De Vrieze et al. (2015)* die Ordnungen *Clostridiales* (Median 21,8%), *Bacteriodales* (Median 13,5%) und eine unklassifizierte *Clostridia*-Ordnung (MBA08 9,8%) sowie *Methanobacteriales* (Median 63,2%) am häufigsten.

5.3 Charakterisierung von mikrobiellen Gemeinschaften von 40 Biogasanlagen

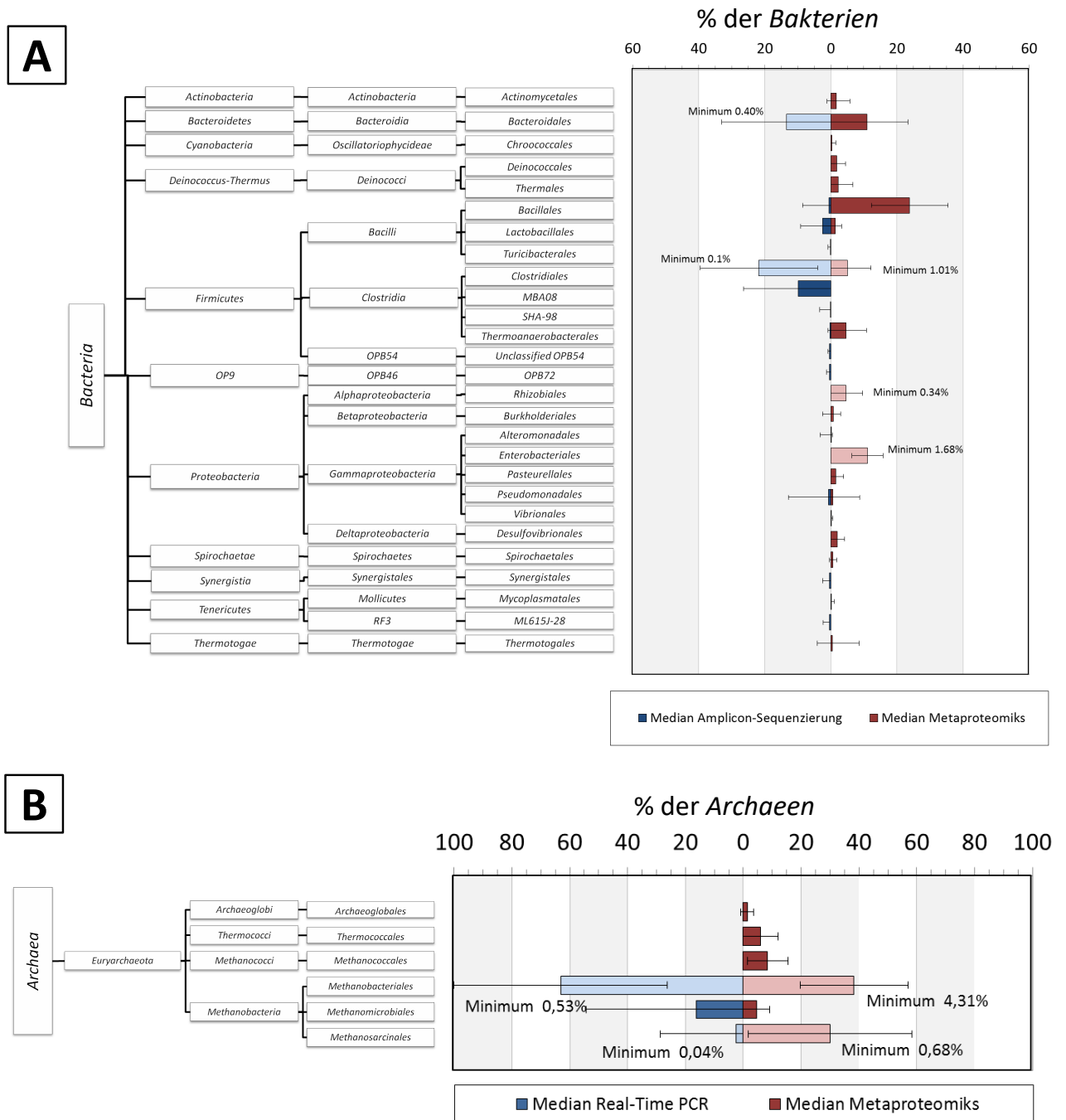


Abbildung 5.10: *Taxonomieprofile aller Biogasanlagen auf Ordnungsebene für Bakterien [A] und Archaeen [B] basierend auf den identifizierten Spektren und dem Median über alle Biogasanlagen mittels Analyse des Metaproteoms und des 16S-rRNA Gens. Es wurden für diesen Vergleich nur die 34 Biogasanlagenproben beachtet, welche identisch mit denen von De Vrieze et al. (2015) waren. Dargestellt ist der Median der jeweiligen Ordnung und die Standardabweichung über die untersuchten Biogasanlagen. Zusätzlich sind in Hellrot und Hellblau die Ordnungen und ihr Minimum dargestellt, welche in allen Biogasanlagen vorkamen.*

Wie schon im Theorieteil eingeführt [Kapitel: 2.2], besteht der Biogasprozess aus verschiedenen Stufen, für welche die entsprechenden Mikroorganismen vorhanden sein müssen. Dementsprechend stellte sich die Frage, ob bestimmte Taxonomien essentiell für bestimmte Stufen des Biogasprozess sind und folglich in allen Biogasanlagen vorkamen. Diese Hypothese traf basierend auf Metaproteindaten für die bakteriellen Ordnungen *Enterobacteriales*, *Clostridiales* und *Rhizobiales* zu, welche mit mindestens 1,7%, 1,0%, bzw. 0,3% gefunden wurden [Abbildung: 5.10]. Analog dazu zeigte die genetische Analyse das Vorhandensein der bakteriellen Ordnungen *Bacteriales* (Minimum 0,4%) und *Clostridiales* (Minimum 0,5%) in allen Biogasanlagen [De Vrieze et al. (2015)]. Für die *Archaeen* wiederum konnte basierend auf den Metaproteom- bzw. Genomdaten mindestens ein Anteil von 4,3% bzw. 0,5% der Ordnung *Methanobacteriales* und 0,7% bzw. 0,04% der Ordnung *Methanosarcinales* belegt werden. Die Ergebnisse der Clusteranalyse und die PCA hat-

ten das Vorhandensein von vier Typen von mikrobiellen Gemeinschaften vorgeschlagen [Kapitel: 5.3.2]. Möglicherweise gibt es analog zu den drei vorgeschlagen Typen von mikrobiellen Gemeinschaften im menschlichen Darm Arumugam et al. (2011) auch verschiedene Typen von mikrobiellen Gemeinschaften in Biogasanlagen. Deshalb sollte weiterführend untersucht werden, welche Ordnungen eventuell repräsentativ für einzelne Gruppen sind. Typisch für die thermophilen Biogasanlagen waren wie in Kapitel 5.2.1 schon gezeigt die thermophilen Ordnungen *Thermotogales* und *Thermococci* sowie ein geringerer Anteil an methanogenen *Archaeen* [Anhang CD: 7.14 Heatmap Metaproteine + Ordnungen + Biologische Prozesse]. Im Gegensatz dazu waren die Klasse *Clostridia* und die methanogenen Ordnungen *Methanobacteriales* und *Methanomicrobiales* charakteristisch für mesophile Biogasanlagen [Nettmann et al. (2010)]. Repräsentativ für UASB-Reaktoren und Biogasanlagen mit Klärschlamm als Substrat war die Ordnung *Methanosarcinales*, welche CH₄ aus Acetat produzieren kann. Ein weiteres Merkmal für UASB-Reaktoren war die Ordnung *Syntrophobacteriales*, welche im Zusammenhang mit syntrophen Interaktionen steht [Schink and Stams (2013)]. Komplementär dazu war die Ordnung *Nitrosomonadales*, welche Nitrifikation durchführen kann, stellvertretend für Biogasanlagen mit Klärschlamm als Substrat.

Einen großen Einfluss auf die taxonomische und funktionelle Zusammensetzung von mikrobiellen Gemeinschaften haben Interaktionen zwischen den einzelnen Taxonomien. Unter Einbeziehung der Prozessparameter wurde deshalb eine Korrelationsstudie [Anhang CD: 7.14 Korrelationsmatrizen der 40 Biogasanlagen] durchgeführt und zur besseren Visualisierung die Zusammenhänge als Graphennetz mittels der Software Gephi dargestellt [Abbildung: 5.11].

Ein Großteil der Taxa korrelierte dabei mit Prozessbedingungen, welche für sie optimal waren. Zum Beispiel erklärt eine höhere Prozesstemperatur eine höhere Abundanz der Ordnungen *Thermotogales*, *Deinococcales*, *Thermococcales* und *Spirochaetales*. Bei den *Archaeen* ist auffällig, dass die Ordnung *Methanomicrobiales* besonders viele positive Korrelationen zu anderen Ordnungen, wie z. B. *Clostridiales* oder *Thermoanaerobacteriales*, aufwies. Im Gegensatz dazu zeigte die Ordnung *Methanosarcinales* nur negative Korrelationen z. B. zu den Ordnungen *Clostridiales*, *Thermoanaerobacteriales* und *Methanomicrobiales*.

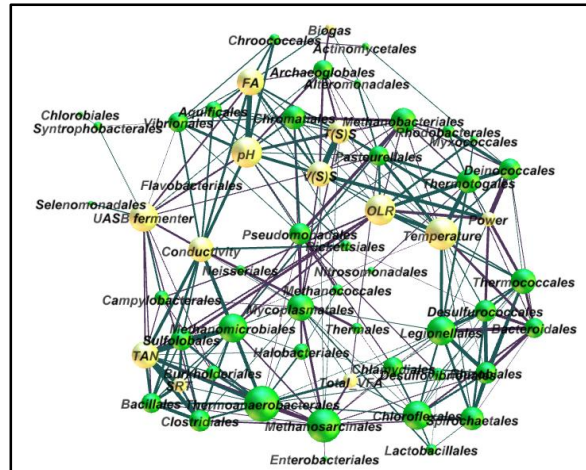
5.3.5 Funktionale Zusammensetzung von mikrobiellen Gemeinschaften

Neben der reinen Anwesenheit bestimmter Mikroorganismen ist auch das Vorhandensein von Proteinen für die einzelnen Schritte des Biogasprozesses wichtig für den erfolgreichen Abbau der Biomasse zu Methan. Folglich sollte auf Ebene der identifizierten Metaproteine [Abbildung: 5.12 A] und biologischen Prozesse [Abbildung: 5.12 B] die dominierenden sowie die in allen Biogasanlagen vorkommenden Funktionen charakterisiert werden. Die Häufigsten biologischen Prozesse waren die Methanogenese (Median: 21,0%; Minimum: 2,9%), gefolgt vom Transport (Median: 15,8%; Minimum 7,7%) und dem C1-Stoffwechsel (Median: 5,2 %; Minimum: 0,5%). Die Dominanz der Methanogenese zeigte sich auch in der Abundanz der Metaproteine. So waren Methyl Coenzyme M Reduktase (MCR) (Median α : 1,8% + 1,4%; β : 2,5%, γ : 1,4%; Minimum: 0,02%) und 5,10-Methylentetrahydromethanopterinreduktase (5,10-Methylen-H4MPT Reduktase) (Median: 1,3/ 2,1; Minimum: 0,04%) neben S-Schicht Proteinen (Median 2,9% + 0,7%, Minimum 0,1%) am häufigsten.

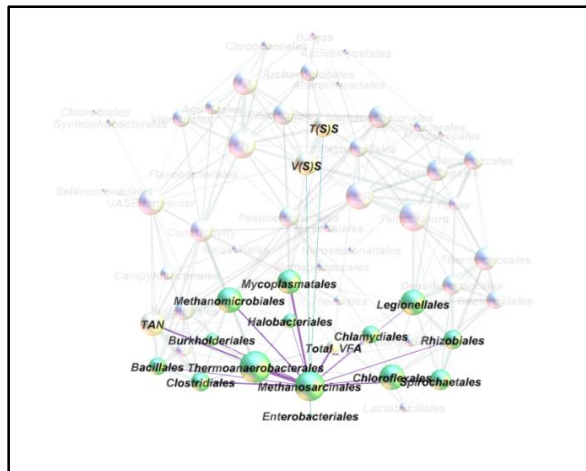
Insgesamt deckten die identifizierten Metaproteine nahezu alle wichtigen Schritte des Biogasprozesses ab [Anhang: CD „Metaproteine + Ordnungen + Biologische Prozesse 40 Biogasanlage“]. Zwar wurden einzelne, essentielle Metaproteine wie die β -Untereinheit des Schlüsselenzym MCR [Abbildung 5.12 A] nicht in allen Biogasanlagen gefunden, aber dafür konnten andere weniger abundante Metaproteine den gleichen Funktionen zugeordnet werden [Anhang: CD „Metaproteine + Ordnungen + Biologische Prozesse 40 Biogasanlage“].

Da die Clusteranalyse und die PCA vier verschiedene Gruppen von mikrobiellen Gemeinschaften gezeigt hatten [Kapitel: 5.3.2], stellte sich die weiterführende Frage, wie sich die biologischen Prozesse und Metaproteine für die einzelnen Gruppen unterscheiden. Allerdings war es schwierig, eindeutige Unterschiede im Profil der biologischen Prozesse bzw. Metaproteine festzustellen [Anhang: CD „Heatmap Metaproteine + Ordnungen + Biologische Prozesse“], da für UASB-Reaktoren und die mit Klärschlamm betrieb-

Alle Ordnungen und Prozessparameter



Fokus auf Methanosarcinales



Fokus auf Methanomicrobiales

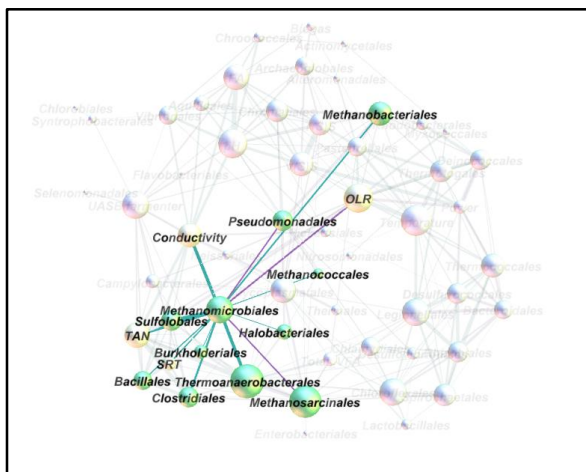


Abbildung 5.11: Zusammenhänge zwischen mikrobiellen Ordnungen und den Prozessparametern, visualisiert als Graphennetzwerk mit Gephi [Bastian et al. (2009)]. Ordnungen sind grüne Knoten und Prozessparameter gelbe. Blaue Kanten markieren positive und rote Kanten negative Korrelationen. Ferner ist die Knotengröße abhängig von der Anzahl der Kanten.

nen Biogasanlagen deutlich weniger Proteine gefunden wurden als für die thermophilen und restlichen mesophilen Biogasanlagen. Erkennbare Unterschiede waren mehr Proteine für DNA-Rekombination, DNA-Reparatur und Aminosäurebiosynthese in thermophilen Biogasanlagen, wie auch schon in der PCA beobachtet [Kapitel: 5.3.2]. Typisch für die mesophilen Biogasanlagen waren mehr Proteine für Fettsäureabbau, Lipidstoffwechsel und C1-Stoffwechsel. Für die weitere Klassifikation der mit Klärschlamm betriebenen Biogasanlagen eignete sich der biologische Prozess der Nitratassimilation und das Enzym Chymotrypsin-ähnliche Elastase IIIA (P09093, K01311).

Um ausgehend von den einzelnen Metaproteinen noch besser auf den tatsächlich ablaufenden Biogasprozess schließen zu können, wurden alle identifizierten Metaproteine mit einem Schwellenwert von 1% der Spektren in mindestens einer Biogasanlage einzelnen KEGG Stoffwechselkarten zugeordnet. Die beste Übereinstimmung zwischen den identifizierten Metaproteinen und einer KEGG Stoffwechselkarte konnte für die Karte des zentralen Kohlenstoffstoffwechsels („KEGG MAP 1200“) erreicht werden [Abbildung: 5.13].

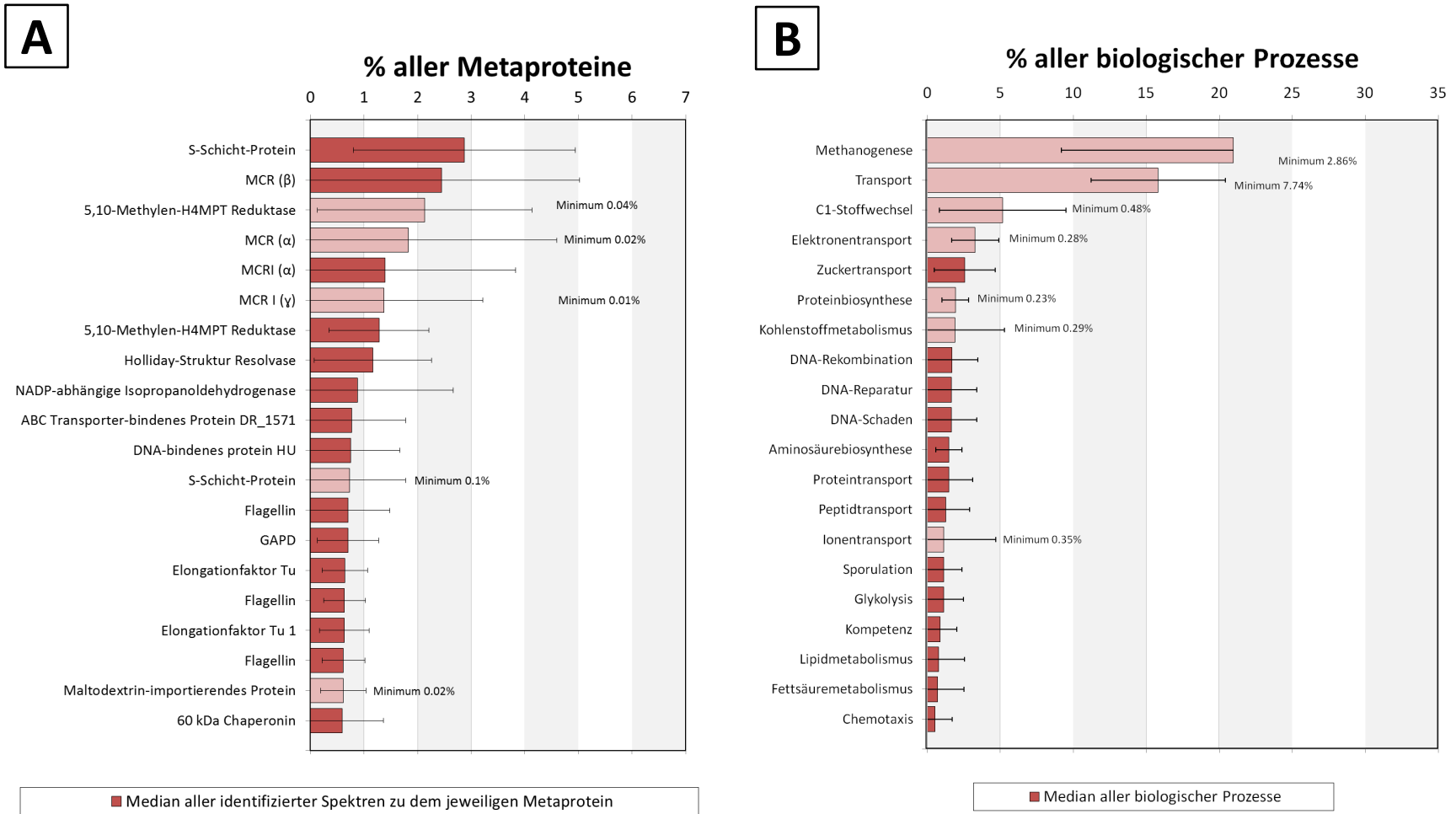


Abbildung 5.12: Häufigste biologische Prozesse und Metaproteine (UniRef50) in Biogasanlagen. Dargestellt sind die 20 häufigsten biologischen Prozesse [A] und 20 häufigsten Metaproteine [B] basierend auf der Anzahl der identifizierten Spektren und dem Median über alle Biogasanlagen. Die Fehlerbalken entsprechen der Standardabweichung. Zusätzlich sind Hellrot mit dem entsprechenden Minimum die biologischen Prozesse/ Metaproteine markiert, die in allen Biogasanlagen nachgewiesen wurden. MCR: Methyl Coenzym M Reduktase; GAPD: Glycerinaldehyd-3-phosphatdehydrogenase H4MPT: Tetrahydromethanopterin.

5.3 Charakterisierung von mikrobiellen Gemeinschaften von 40 Biogasanlagen

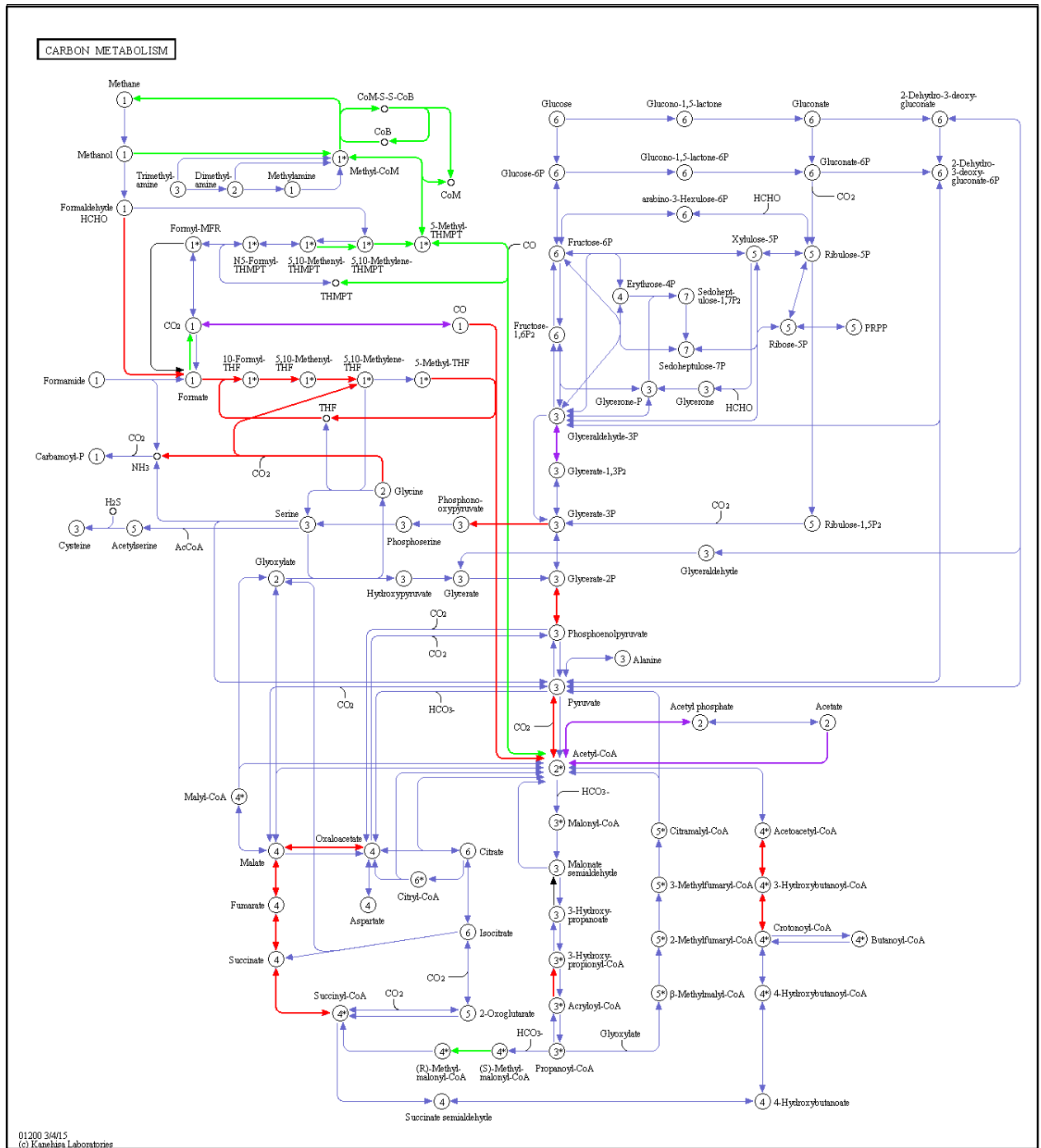


Abbildung 5.13: Zuordnung aller identifizierter Metaproteine zur KEGG Karte 1200 (Kohlenstoffstoffwechsel) Hervorgehoben sind grün Archaeen, rot Bakterien und lila Archaeen oder Bakterien.

Neben einer Vielzahl von bakteriellen Metaproteinen für die Glykolyse konnten Metaproteine für die Gärung bzw. Verwertung von Butyratgärung gefunden werden. Die identifizierten Metaproteine des reduktiven Citratzyklus gehören hauptsächlich zur Klasse

Proteobacter, insbesondere zur Spezies *Escherichia coli*. Dies könnte ein Hinweis auf Succinatgärung sein. Da für *Escherichia coli* aber auch das Enzym Laktatdehydrogenase (P52643, K03778) gefunden wurde, könnten diese *Bakterien* auch gemischte Säuregärung durchgeführt haben. Neben dem Abbau von Zuckern wurden auch mehrere Metaproteine für den Aminosäureabbau gefunden. Beispielsweise wurden für die Ordnung *Clostridiales* die Enzyme Glycinreduktasekomplex B (Q9R4G8, EC: 1.21.4.2) und Aminomethyltransferase (P54378, K00605) detektiert. Ersteres ermöglicht den anaeroben Glycinabbau über die Sticklandreaktionen (nicht in „KEGG MAP 1200“ enthalten) [Nisman (1954)] und das Zweite den Abbau von Glycin zu Kohlenstoffdioxide, Ammoniak und Methylentetrahydrofolat. Letzteres Produkt kann dann über den Wood-Ljungdahlweg weiter zu Acetat [Okamura-Ikeda et al. (1987)] umgewandelt werden. Passend dazu konnte eine Vielzahl von bakteriellen Proteinen des C1-Stoffwechsels gefunden werden. Es blieb allerdings offen, ob dieser Weg für die Acetogenese oder für die syntrophe Oxidation von Acetat genutzt wurde. Die meisten Metaproteine für diesen Weg gehörten zu homoacetogenen Spezies wie *Moorella thermoacetica* (*Clostridium thermoaceticum*) [Pierce et al. (2008)], aber zumindest das Metaprotein Formatetetrahydrofolatligase (Q3Z8K3, K01938) war spezifisch für den syntrophen Propionatoxidierer *Syntrophobacter fumaroxidans* [Schink and Stams (2013)]. Generell war die korrekte Zuordnung der Taxonomien für diese Metaproteine schwierig, da nur wenige Proteinsequenzen für syntrophe Acetatoxidierer bzw. Homoacetogene bekannt waren. Für den abschließenden Weg der Methanogenese konnten wiederum analog zu vorherigen Arbeiten [Kohrs et al. (2014), Lü et al. (2014), Kohrs et al. (2015)] nahezu alle Schritte der hydrogenotrophen und acetoclastischen Methanogenese beobachtet und den Ordnungen *Methanobacteriales* und *Methanosarcinales* zugeordnet werden.

5.3.6 Zusammenhang zwischen mikrobiellen Gemeinschaften und Prozessparametern

Aufbauend auf der reinen taxonomischen und funktionellen Beschreibung des Biogasprozesses sollte gezielt nach Metaproteinen, taxonomischen Ordnungen und biologischen Prozessen gesucht werden, welche mit bestimmten Prozessparametern korrelierten. Da die entsprechende Matrix aller Korrelationen 14,000 signifikante Korrelationen umfasste („Spearman Korrelation“, p-Wert <5%) [CD: Anhang „Korrelationsmatrizen der 40 Biogasanlagen“], wurden für die Korrelationsanalyse nur die Ordnungen, biologischen Prozesse und Metaproteine beachtet, welche in mindestens einer Probe 5% der Spektren ausmachten. Zusätzlich wurde der p-Wert auf kleiner 1% reduziert, wodurch die Korrelationsmatrix („Spearman-Korrelation“, p-Wert <1%) am Ende nur noch 75 Parameter

und 744 Korrelationen umfasste.

Schwierig bei der Interpretation der Korrelation zwischen den Prozessparameter und den mikrobiellen Gemeinschaften war, dass die einzelnen Größen voneinander abhängig waren [Anhang: 7.12]. Zum Beispiel führte ein hoher Ammoniumanteil im Substrat zu einem höheren pH-Wert [Abbildung: 5.14] und mehr Leitfähigkeit. Ferner wurden Biogasanlagen mit höheren Prozesstemperaturen auch höheren Raumbelastungen ausgesetzt, wodurch wiederum die Biogasproduktivität erhöht war.

Trotzdem konnten markante Zusammenhänge zwischen der mikrobiellen Gemeinschaft und der Prozesstemperatur, dem Gesamtstickstoff, der Raumbelastung und der Schlammverweilzeit beobachtet [Tabelle: 5.2] und anhand der „KEGG Karte 1200“ (zentraler Kohlenstoffmetabolismus) visualisiert werden [Anhang: 7.13]:

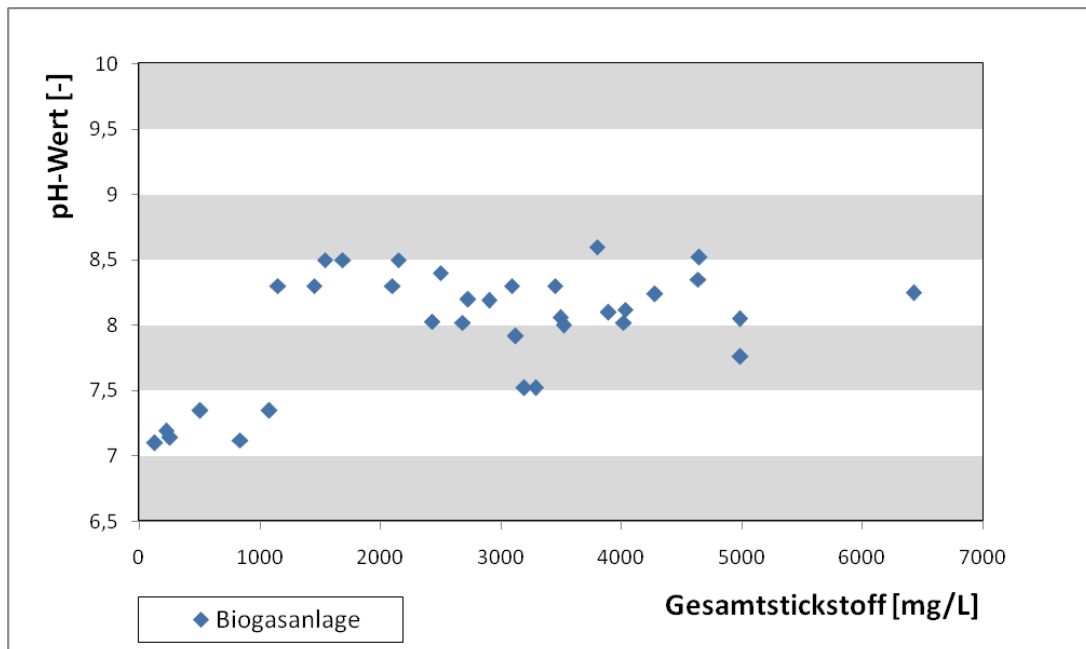


Abbildung 5.14: *Korrelation Gesamtstickstoff zu pH-Wert für 40 Biogasanlagenproben.*

Tabelle 5.2: *Effekt von Prozessparametern auf den Kohlenstoffmetabolismus.*

Aufgeführt sind alle Proteine bzw. Funktionen der „KEGG MAP 1200“, welche mit der Prozesstemperatur, der Gesamtammoniumstickstoffkonzentration, der Raumbelastung oder der Schlammverweilzeit korrelieren.

| Prozessparameter | Hoch | Niedrig |
|--|--|--|
| Prozesstemperatur | Enzyme der Glykolyse | Mehr Methanogenesenzyme (<i>Archaea</i>) |
| | Hydrogenotrophe Methanogenese (<i>Methanobacteriales</i>) | Fettsäurestoffwechsel |
| Gesamtammonium- stickstoffkonzentration | Hydrogenotrophe Methanogenese (<i>Archaea</i>) | Acetoclastische Methanogenese (<i>Methanosarcinales</i>) |
| | Bakterieller C1-Stoffwechsel | |
| | Glycinabbau im C1-Stoffwechsel | |
| Raumbelastung | Acetoclastische Methanogenese (<i>Archaea</i>) | Hydrogenotrophe Methanogenese (<i>Archaea</i>) |
| | Bakterielle Glycolysis | Bakterieller C1-Stoffwechsel |
| Schlammverweilzeit | Hydrogenotrophe Methanogenese (<i>Archaea</i>) | Acetoclastische Methanogenese (P16304 Adenylate kinase, <i>Methanosaeta concilii</i>) |
| | Bakterieller C1-Stoffwechsel | |

Prozesstemperatur

Bei höherer Prozesstemperatur konnten mehr Enzyme der Glykolyse beobachtet werden, während bei niedriger Prozesstemperatur vermehrt Enzyme der Methanogenese [Kohrs et al. (2014), Gunnigle et al. (2015)] und des Fettsäurestoffwechsels gefunden wurden. Die Zunahme des Enzymes Acetyl-CoA Decarboxylase/Synthase mit steigender Prozesstemperatur widersprach dagegen der gängigen Beobachtung, dass die acetoclastische Methanogenese bei höherer Temperatur ungünstiger ist als die hydrogenotrophe Methanogenese [Nozhevnikova et al. (2007)]. Allerdings wurde dieser Trend nur für die Δ -Untereinheit dieses Enzymes beobachtet.

Gesamtammoniumstickstoffkonzentration

Hohe TAN [Tabelle: 5.2] gelten als Stressfaktor für die mikrobiellen Gemeinschaften in Biogasanlagen [Müller et al. (2013)]. Für die untersuchten Biogasanlagen korrelierten hohe TAN mit dem bakteriellen C1-Stoffwechsel und der hydrogenotrophen Methanogenese der Familien *Methanobacteriaceae* und *Methanosarcinaceae*. Dies deutet sehr stark auf syntrophe Acetatoxidation hin. Weiterhin wurden positive Korrelationen vom TAN mit den Schlüsselordnungen der syntrophen Acetatoxidation *Clostridiales*, *Thermoanaerobacteriales* und *Methanomicrobiales* [Müller et al. (2013)] gefunden. Zwar könnten die Enzyme des C1-Stoffwechsels auch für die umgekehrte Reaktion der Homoacetogenese genutzt werden, aber dies ist aufgrund der negativen Korrelation mit der Ordnung *Methanosarcinales* und den Schlüsselenzymen der acetoclastischen Methanogenese (z. B. P26692 Acetyl-CoA Decarboxylase/Synthase Komplex A, *Methanosaeta concilii*) unwahrscheinlich. Passend dazu waren kleine TAN wiederum mit der acetoclastischen Methanogenese verbunden [De Vrieze et al. (2012)].

Raumbelastung

Eine höhere OLR ging hauptsächlich einher mit der acetoclastischen Methanogenese und der bakteriellen Glykolyse. Im Gegensatz dazu waren bei niedrigen OLRs die hydrogenotrophe Methanogenese und der bakterielle C1-Stoffwechsel erhöht.

Schlammverweilzeit

Bei hohen SRTs waren vermehrt Enzyme für die hydrogenotrophe Methanogenese (Archaea) und den bakterieller C1-Stoffwechsel beobachtbar. Im Gegensatz dazu korrelierten kurze SRTs mit Proteinen der acetoclastischen Spezies *Methanosaeta concilii* (P16304 Adenylatekinase) und des hydrogenotrophen *Methanobrevibacter smithii*. Prinzipiell waren die Korrelationen der SRT gegensätzlich zu denen der OLR. Dies ist aber nicht verwunderlich, da hohe OLRs meist zu niedrigen SRTs in Biogasanlagen führen [Anhang:

7.12].

Weiterführend zu den Korrelationen wurde mittels „Decision Tree Learning“-Algorithmus [Kapitel: 3.6.3] geprüft, ob einzelne Metaproteine, taxonomische Ordnungen oder biologische Prozesse als Marker zur Vorhersage von vorher klassifizierten Prozessparametern/-problemen verwendet werden können [Anhang: CD „Ergebnisse Decision Tree Learning“]. Vielversprechende Vorschläge des Lernalgorithmus für Marker waren, wie in früheren Publikationen angedeutet, die Zunahme des Enzyms 5,10-Methylentetrahydromethanopterinreduktase (Q8TXY4 *Euryarchaeota*) für hohe TANs, die Zunahme der Ordnung *Thermotogales* für hohe Prozesstemperatur [Huber and Hannig (2006)] sowie die Abnahme des Enzyms Methyl Coenzym M Reduktase (P07962, *Methanosarcina barkeri str. Fusaro*) für hohe OLRs [Munk et al. (2012), Heyer et al. (2013)].

5.3.7 Diskussion

Für alle untersuchten Biogasanlagenproben konnten erfolgreiche Proteine extrahiert [Abbildung: 5.5] und identifiziert werden [Abbildung: 4.8]. Die Ursache für die geringere Qualität der SDS-PAGE von drei Biogasanlagenproben bei einem Replikat waren wahrscheinlich verbleibende Huminstoffe, welche vor der SDS-PAGE nicht gut genug durch Zentrifugation abgetrennt wurden [Kuhn et al. (2011), Heyer et al. (2015)]. Nichtsdestotrotz zeigten die Proteinmuster der restlichen 37 Proben, dass der Workflow prinzipiell für die Untersuchung von verschiedensten Typen von Biogasanlagen geeignet war. Im Vergleich zum ersten SDS-PAGE basierten Screening von Biogasanlagen [Kapitel: 5.1.1], konnten hier auch größere Unterschiede innerhalb der Bandenmuster beobachtet werden. Dies deckte sich mit einer größeren Variation der Prozessparameter zwischen den einzelnen Biogasanlagen und deutete auf größere Unterschiede zwischen den mikrobiellen Gemeinschaften hin. Die Anzahl an identifizierten Spektren und Metaproteinen konnte dank eines besseren MS [Kapitel: 4.2.2] und der Metagenomdatenbanken [Kapitel: 4.3.2] im Vergleich zu früheren Experimenten [Kapitel: 5.1.1, 5.2.1] deutlich gesteigert werden. Dadurch konnten mit dem schnelleren 1D-Ansatz mehr Identifikationen erzielt werden als in vorherigen Studien mittels 2D-Ansatz [Kapitel: 5.2.1]. Dies ermöglichte erstmals die Untersuchung von Biogasanlagen im Hochdurchsatz (40 Biogasanlagen in einem Monat).

Durch die anschließende Clusteranalyse und PCAs [Abbildung: 5.8] konnte die Reproduzierbarkeit des Workflows bestätigt und das Vorkommen von vier Gruppen von mikrobiellen Gemeinschaften in Biogasanlagen beobachtet werden. Prinzipiell war die Gruppierung der Biogasanlagen sehr ähnlich zu der *De Vrieze et al. (2015)*. Lediglich der Einfluss des Stickstoffes auf die Gruppierung konnte basierend auf den Metaproteomdaten nicht beob-

achtet werden. Metaproteomanalysen ermöglichen folglich qualitativ ähnliche Ergebnisse wie genomische Untersuchungen.

Kleine Auffälligkeiten gab es bei der Zuordnung von charakteristischen Metaproteinen, taxonomischen Ordnungen und biologischen Prozessen zur Gruppe der Biogasanlagen, welche mit Klärschlamm betrieben wurden [Kapitel: 5.3.4, 5.3.5]. Chymotrypsin-ähnliche Elastase IIIA ist kein mikrobielles Protein, sondern ein ausgeschiedenes, menschliches Verdauungsenzym [Kuhn et al. (2011), Püttker et al. (2015)]. Ferner kann unter anaeroben Bedingungen in Biogasanlagen keine Nitrifikation durch die Ordnung *Nitrosomonadales* betrieben werden [Brenner et al. (2005)], da der dafür benötigte Sauerstoff fehlt. Folglich sind die Proteine der Ordnung *Nitrosomonadales* vermutlich nur Überreste aus dem Klärprozess. Nichtsdestotrotz scheinen Chymotrypsin-ähnliche Elastase IIIA und Proteine der Ordnung *Nitrosomonadales* verlässliche Marker für den Betrieb von Biogasanlagen mit Klärschlamm als Substrat, zu sein.

Die Ursache für die geringe erklärte Varianz von <34% bei den PCAs war vermutlich die große Anzahl an gemessenen taxonomischen Ordnungen, biologischen Prozessen und Metaproteinen. Eine Möglichkeit die erklärte Varianz auf Kosten der Visualisierung zu erhöhen und zusätzlich die Separation der Gruppen zu verbessern, wäre die Betrachtung von weiteren Hauptkomponenten.

Erstmalig wurden in dieser Arbeit die ökologischen Indizes Richness und Evenness auf Metaproteomdaten angewendet [Abbildung: 5.9]. Bei weniger als 3.000 identifizierten Spektren zeigte sich eine starke Abhängigkeit bei der Berechnung der Indizes von der Anzahl an identifizierten Spektren. Dementsprechend sollte die Bewertung dieser Indizes bei begrenzter Auflösung mit Vorsicht erfolgen, um falsche Ergebnisse zu vermeiden. So sollte auch das Ergebnis, dass basierend auf den Metaproteomdaten die Richness bei höheren Prozesstemperaturen zunimmt, durch weitere Experimente überprüft werden. Entsprechende genomische Analysen basierend auf dem 16S-rRNA von *Leven et al. (2007)* hatten vorher das Gegenteil beobachtet. Möglicherweise war die Auflösung der durchgeführten Metaproteomanalysen nicht ausreichend, um die höhere taxonomische Vielfalt in den mesophilen Biogasanlagen zu detektieren.

Zumindest waren die durch Metaproteomanalyse und genetische Untersuchung des 16S-rRNA Gens in Biogasanlagen identifizierten Taxa zum größten Teil identisch [Abbildung: 5.10] und deckten sich mit den Ergebnissen anderer Studien [Schlüter et al. (2008), Nettmann et al. (2010), Ziganshin et al. (2013), Kohrs et al. (2014), Stolze et al. (2015)]. Allerdings unterschieden sich die Häufigkeiten der einzelnen Taxa bei beiden Methoden deutlich. Große Unterschiede in der Häufigkeit der *Archaeen* und *Bakterien* bei der Ver-

wendung von verschiedenen Methoden wurden auch schon anderweitig in der Literatur beschrieben. So beobachteten *Hanreich et al. (2013)* 4% *Archaeen* mittels Metagenomanalyse und 20-30% mittels Metaproteomanalyse oder *Kim et al. (2014)* 0,2%/ 7% *Archaeen* mittels qPCR und 12%/ 6% mittels Mikroskopie. Wenn die Werte der Metaproteomanalyse mit den vorhergesagten Werten des „Anaerobic Digestion Model 1“ von 30-40% verglichen werden [Batstone et al. (2002), Anhang CD: „Anmerkungen zum Anaerobic Digestion Model 1“], erscheinen die Werte der Metaproteomanalyse allerdings durchaus realistisch. Mögliche Ursachen für die Unterschiede zwischen Metaproteomanalysen und genetische Untersuchung des 16S-rRNA Gens sind die Zellyse oder unterschiedliche Extraktionseffizienzen für die Proteine und DNA. Besondere Schwachstellen der genetischen Methoden waren die oft nicht beachtete Variation der Kopienzahl für das Zielgen, welche z. B. für das 16S-rRNA Gen zwischen 1-8 Kopien beträgt und die unterschiedliche Primeraffinität [Acinas et al. (2004)]. Letztere resultiert daher, dass die verschiedenen Taxonomien in der Probe unterschiedliche Versionen eines Gens herstellen, an welche der jeweilige Primer unterschiedlich gut binden kann. Die Nachteile von Metaproteomanalysen sind, dass nur in Datenbanken abgelegte Proteine identifiziert werden können und die taxonomische Charakterisierung durch den höheren Konservierungsgrad der Proteinsequenz und die kürzeren identifizierten Peptide schwierig ist. So konnten für dieses Experiment 48,6% der identifizierten Peptide keiner bestimmten Ordnung zugeordnet werden.

Trotz des Nachweises mehrerer mikrobieller Ordnungen in allen Biogasanlagen war allerdings keine Ordnung so häufig [Abbildung: 5.10], dass daraus auf eine Schlüsselrolle innerhalb des Biogasprozesses („Core Microbiome“) geschlossen werden konnte. Vermutlich können die einzelnen Schritte des Biogasprozesses von verschiedenen Taxa ausgeführt werden. Passend dazu zeigte die Korrelationsanalyse der mikrobiellen Ordnungen [Abbildung: 5.11] eine negative Korrelation zwischen den methanogenen Ordnungen *Methanosarcinales* und *Methanomicrobiales*, welche darauf hin deutet, dass in Biogasanlagen entweder die eine oder die andere Ordnung dominiert. Die große Streuung der Prozessparameter in den Biogasanlagen, z. B. 0,38-150 Tage Verweilzeit, Temperaturen zwischen 33 °C bis 55 °C und verschiedenste Substrate, trug wahrscheinlich zusätzlich zur großen Heterogenität der Taxonomieprofile bei.

Im Gegensatz zu den mikrobiellen Ordnungen konnte eine Vielzahl von Metaproteinen und biologischen Prozessen identifiziert werden, welche eine Schlüsselrolle in Biogasprozessen einnehmen, wie z. B. die biologischen Prozesse „Methanogenese“ oder „Transport“ [Abbildung: 5.12]. Allerdings konnte analog zu anderen Arbeiten [Stolze et al. (2015), Stolze et al. (2016), Campanaro et al. (2016)] gezeigt werden, dass für die einzelnen Schritte des Biogasprozesses verschiedene metabolische Wege genutzt werden können [Abbildung: 5.13].

So produzieren Vertreter der Ordnung *Methanomicrobiales* CH₄ aus H₂ und CO₂ [Karakashev et al. (2006)], während Vertreter der Ordnung *Methanosarcinales* auch Substrate wie Acetat, Methanol und Amine für die Methanproduktion nutzen können. Zusätzlich wurde die Häufigkeit der Proteine für die einzelnen Stoffwechselwege durch die Prozesstemperatur, die TAN, die OLR und die SRT beeinflusst. Die Ursache für den großen Einfluss dieser vier Prozessparameter auf die einzelnen Stoffwechselwege von Biogasprozessen ist wahrscheinlich darin begründet, dass diese die Ausbildung von syntrophen Interaktionen und den energetischen Nutzen der einzelnen Stoffwechselwege prägen [Schink (1997)]. Eine Möglichkeit die Bedeutung und die Aktivität der einzelnen Stoffwechselwege noch genauer zu ergründen, wären kombinierte Metaproteom- und Metabolomuntersuchungen mit C13-markierten Kohlenstoffquellen [von Bergen et al. (2013)].

Bei der Untersuchung der Korrelationen zwischen den Prozessparametern und den Metaproteinen, taxonomischen Ordnungen und biologischen Prozessen konnte eine Vielzahl an Zusammenhängen beobachtet werden. Es zeigte sich aber, dass die Korrelationsmatrix für die Suche nach Metaproteinen, taxonomischen Ordnungen und biologischen Prozessen, welche als Marker für Prozessparameter genutzt werden können, ungeeignet war. Die Gründe waren die Komplexität der Daten [CD: Anhang „Korrelationsmatrizen der 40 Biogasanlagen“] sowie die Abhängigkeit der einzelnen Größen voneinander [Abbildung: 5.14].

Eine einfachere und robustere Möglichkeit zur Identifikation von Markern für einzelne Prozessparameter war die Nutzung des „Decision Tree Learning“-Algorithmus. Aber auch bei diesem Ansatz konnten nicht für alle Prozessparameter Metaproteine, Ordnungen bzw. biologische Prozesse gefunden werden, welche als Marker geeignet sind [Anhang: CD „Ergebnisse Decision Tree Learning“]. Vermutlich waren die zugrunde liegenden Mechanismen entweder komplexer z. B. Kombinationen von Metaproteinen oder es hätten mehr Biogasanlagen für eine entsprechende statistische Absicherung untersucht werden müssen. Zusätzlich wurde der Erfolg des Lernalgorithmus maßgeblich durch die vorherige Klassifikation der Biogasanlagen in eine Gruppe mit hohen und eine Gruppe mit niedrigen Werten des jeweiligen untersuchten Prozessparameters beeinflusst [Kapitel: 3.6.3, Anhang: CD „Ergebnisse Decision Tree Learning“]. Die Grenzwerte für die Klassifikation entstammen der Literatur [Chen et al. (2008), Schnürer and Nordberg (2008), Weiland (2010) oder basierten auf dem Median, falls keine geeigneten Literaturwerte vorhanden waren. Aber selbst innerhalb der Literatur gab es Unterschiede in der jeweiligen Einordnung. Beispielsweise definierten *Chen et al. (2008)* [Chen et al. (2008)] die Schwelle für zu hohe TANs bei 4.200 mg/L und *Schnürer et al. (2008)* [Schnürer and Nordberg (2008)] bei 3.000 mg/L. Wurde der Schwellwert von *Schnürer et al. (2008)* genutzt, war

der Klassifikationsfehler kleiner.

Zusammenfassung: Die taxonomische Charakterisierung von mikrobiellen Gemeinschaften mittels Metaproteomanalyse lieferte ähnliche Ergebnisse wie Studien, die auf der Analyse des 16S rRNA Gens basierten. Bei der Charakterisierung der mikrobiellen Gemeinschaften der 40 Biogasanlagen zeigten sich vier Gruppen, welche mit mesophilen und thermophilen Prozessbedingungen, UASB-Reaktoren und Klärschlamm als Substrat assoziiert waren. Weiterhin wurden die mikrobiellen Gemeinschaften wesentlich von der Prozessstemperatur, der TAN, der OLR und der SRT geprägt.

Fazit und Ausblick

Im Rahmen dieser Promotionsarbeit sollte mittels Metaproteomanalyse ein grundlegendes Verständnis über den taxonomischen und funktionellen Aufbau der mikrobiellen Gemeinschaften in Biogasanlagen gewonnen werden. Die Metaproteomanalyse von mikrobiellen Gemeinschaften ist eine relativ junge Methode, die 2006 von *Wilmes et al.* erstmals beschrieben wurde. Im Laufe der Promotionsarbeit gab es eine dynamische Weiterentwicklung dieser Methode hinsichtlich der Probenvorbereitung, der Identifizierung der Proteine mit LC-MS/MS Systemen und der bioinformatischen Auswertung der erzeugten Daten. Infolgedessen musste der in dieser Arbeit genutzte Workflow wiederholt an die neuesten Entwicklungen angepasst werden. Dabei wurden in dieser Arbeit neben der Bearbeitung der inhaltlichen Fragen zum Biogasprozess Fortschritte erzielt, die das Feld der Metaproteomanalysen voranbrachten. Die Optimierung des Workflows der Metaproteomanalyse von Biogasanlagenproben wurde deshalb in diesem Abschnitt im Rahmen der Entwicklungen der letzten zehn Jahren zusammengefasst und diskutiert. Darüber hinaus wurde ein Ausblick auf mögliche Entwicklungen gegeben. Bei der Betrachtung wurden folgende Aspekte verglichen: die Zielsetzung der Metaproteomanalyse, die Probenvorbereitung, die Proteinseparation, verfügbare MS-Typen, der Umfang und die Qualität der Proteindatenbanken, die bioinformatische Auswertung und die Quantifizierung der Proteine [Tabelle: 6.1].

Tabelle 6.1: *Entwicklung der Metaproteomanalysen*

| | Metaproteomik vor "10 Jahren" | Metaproteomik "heute" | Metaproteomik vermutlich in "10 Jahren" |
|------------------------------------|--|---|--|
| Zielsetzung | Beschreibung der taxonomischen und funktionellen Zusammensetzung von mikrobiellen Gemeinschaften | Korrelation der taxonomischen und funktionellen Zusammensetzung mit Parametern/Lebensbedingungen | Prozessüberwachung/Diagnostik |
| Probenvorbereitung | Erste umfangreiche Protokolle für die Extraktion von Proteinen aus komplexen Umweltproben | Robuste Protokolle für Umweltproben und erste kommerzielle Kits für Probenvorbereitung | Gesamte Probenvorbereitung inklusive Messung in weniger als 12 h |
| Proteinseparation | 2D-PAGE + LC | SDS-PAGE + LC oder 2D-LC | nur LC |
| MS-Typ | MALDI-Tof-/ESI-Quadrupol-MS/MS | ESI-Orbitrap-MS/MS | Ionenmobilitäts-MS/MS |
| Datenbanken | UniprotKB/NCBI, erste Metagenome | Kombinierte Metagenome mit Annotation, ausgewählte Taxa aus UniprotKB/NCBI | Eigenes Metagenom für jede Probe |
| Bioinformatische Auswertung | Proteinidentifikation mit einer Datenbanksuchmaschine | Bessere Proteomiksoftware, kombinierte Datenbanksuchmaschinen und spezielle Metaproteomiksoftware | Weiterentwicklung der Software wegen Zunahme der Datenmengen |
| Quantifizierung | Intensität von 2D-Spots | Spektrenanzahl | Intensität der Peptidionen oder isotoopenmarkierter Standard |

Während erste Metaproteomanalysen deskriptiv die taxonomische und funktionelle Zusammensetzung analysierten und bestimmte biochemische Wege aufdeckten [Wilmes and Bond (2006), Ram et al. (2005), Kan et al. (2005), Abram et al. (2009)], zielten aktuelle Untersuchungen auf die Korrelation der mikrobiellen Gemeinschaften mit Lebensbedingungen bzw. Prozessparametern ab [Erickson et al. (2012), Kolmeder et al. (2016)]. Entsprechend wurde in der vorliegenden Promotionsarbeit zunächst auf die Zusammensetzung der mikrobiellen Gemeinschaften und die Abbauwege der Biomasse eingegangen, während später der Einfluss der Prozessparameter auf die mikrobiellen Lebensgemeinschaften untersucht wurde. Am Ende erfolgte basierend auf den Metaproteomdaten die Differenzierung der mikrobiellen Gemeinschaften nach Prozessparametern. Zukünftig könnten Metaproteomanalysen sogar für die Prozessüberwachung eingesetzt werden. Vielversprechend in diesem Zusammenhang ist das Konzept des „Proteotyping“ [Karlsson et al. (2015)], welches Proben entsprechend ihres Proteinprofils gruppiert und nach ihrer Ähnlichkeit zu bekannten Proben klassifiziert. Anwendung findet dieses Konzept bereits bei der Identifizierung von Mikroorganismen mittels MALDI-TOF-MS [Kallow et al. (2010), Welker and Moore (2011)].

Für den zukünftigen Einsatz von Metaproteomanalysen für Routineuntersuchungen ist eine robuste und schnelle Probenvorbereitung inklusive der Protein- und Peptidseparation nötig. Die erste Generation an Protokollen für die Proteinextraktion war sehr anfällig für Probenverunreinigungen, wodurch die anschließende Proteinseparation mittels 2D-PAGE erheblich gestört wurde und in vielen Fällen unmöglich war [Bastida et al. (2009), Kuhn et al. (2011), Hanreich et al. (2012)]. Darüber hinaus benötigte die 2D-PAGE und die anschließende Identifikation der Proteinspots mindestens eine Arbeitswoche. Zwischenzeitlich wurden verbesserte Protokolle etabliert, wie die Kombination aus Phenolextraktion und Methanol in Ammoniumacetatfällung [Benndorf et al. (2007)] oder kommerzielle Kits für die Proteinextraktion. Die Verfügbarkeit neuer, hoch auflösender MS, mit denen mehr Proteine identifizieren werden konnten, machte schließlich den Ersatz der aufwendigen 2D-PAGE durch die robustere eindimensionale SDS-PAGE möglich [Benndorf and Reichl (2014)]. Deshalb wurden im Rahmen dieser Arbeit die mikrobiellen Gemeinschaften von Anfang an mittels SDS-PAGE anstelle von 2D-PAGE untersucht [Heyer et al. (2013)]. Durch das im Verlauf der Arbeit verfügbar gewordene Elite Hybrid Ion Trap Orbitrap -MS/MS, konnte im Vergleich zu einem AmaZon ETD ESI-Ion Trap MS/MS die Anzahl an identifizierten Spektren vervierfacht werden [Kapitel: 4.2.2]. Die Messung der Proben als eine Fraktion mittels LC-MS/MS ermöglichte schließlich die umfassende Analyse von 40 Biogasanlagenproben mit der notwendigen Anzahl an Replikaten in einer vertretbaren Gesamtmesszeit von 12 Tagen. Durch einen kompletten Verzicht auf die Gelbasierte Proteinseparation, z. B. durch den FASP Verdau [Wiśniewski et al. (2009)] und

die damit verbundene Vermeidung auf die Messung mehrerer Fraktionen im LC-MS/MS, könnte die Analysezeit sogar unter 24 h reduziert werden. Eine moderate Verlängerung der LC-Gradienten oder Einsatz von MS mit einer zusätzlichen Ionenmobilitäts-Trennung [Silveira et al. (2014), Meier et al. (2015)] könnte die Auflösung und Sensitivität zusätzlich steigern.

Die bioinformatische Auswertung und die Verfügbarkeit von besseren Proteindatenbanken haben ebenfalls große Fortschritte gemacht. Anstelle einfacher Datenbanksuchmaschinen für die Proteinidentifikation wurde in den letzten Jahren komplexere Software für Proteomanalysen entwickelt, die es erlaubt, innerhalb eines Workflows mehrere Suchmaschinen zu kombinieren und Daten quantitativ und qualitativ auszuwerten [Jagtap et al. (2015), Vaudel et al. (2011), Sturm et al. (2008)]. Mit dem MPA konnte im Rahmen dieser Promotionsarbeit zur Entwicklung einer Software beigetragen werden, die auf die spezifischen Bedürfnisse der Metaproteomanalyse zugeschnitten ist [Muth et al. (2015a)]. Insbesondere die Gruppierung von redundanten Metaproteinen und ihre taxonomische Zuordnung soll hier erwähnt werden. Eine Herausforderung für diese Programme (auch den MPA) ist die ständig wachsende Datenmenge. Auf der einen Seite wächst der Umfang der experimentellen Daten durch immer sensitivere und höher auflösende MS. Auf der anderen Seite nimmt die Größe der Datenbanken durch die zunehmende Genomsequenzierung zu. Mit der Vervielfachung der bekannten Proteinsequenzen durch umfassende (Meta)-Genomsequenzierungen [Venter et al. (2004), Stolze et al. (2016)] können inzwischen nicht nur mehr Proteine identifiziert [Kapitel: 4.3.1], sondern es werden mittlerweile aus diesen Sequenzen Genome der einzelner Mikroorganismen zusammengesetzt und die Gene organismenspezifisch annotiert. Der rasanten Preisverfall für die Sequenzierung erleichtert zusätzlich die parallele Sequenzierung von Metagenomen und Metatranskriptomen, so dass Multi-Omiks-Studien möglich werden.

Einen wesentlichen Einfluss auf die Aussagekraft der Metaproteomuntersuchungen hat die Quantifizierbarkeit der identifizierten Proteine. Während die Quantifizierung von Proteinspots in 2D-PAGEs bei einer entsprechenden Anzahl an Replikaten zuverlässig war, ist die Quantifizierung der Proteine aus der Spektrenanzahlen bei der LC-MS/MS ungenauer. Wegen der Probenverunreinigungen haben sich metabolische oder chemische Markierungstechniken mit stabilen Isotopen für die Metaproteomanalyse meist als wenig zweckmäßig erwiesen. Für die absolute Quantifizierung der Intensität von Peptidionen können jedoch internen Standards aus isotoopenmarkierten Peptiden zugesetzt werden (AQUA, QCon-Cat) [Brun et al. (2009)]. Dies könnte besonders beim Routineeinsatz von Metaproteomanalysen bei der Prozessüberwachung für mehr Zuverlässigkeit sorgen.

Mit der Etablierung eines Workflows für Metaproteomanalysen wurde im Verlauf der

Promotionsarbeit eine umfassende Analyse von Proben aus Biogasanlagen möglich. So konnten die Metaproteome von über 40 Biogasanlagenproben mit den Prozessparametern Prozesstemperatur, TAN, OLR, SRT und Reaktordesign korreliert werden. Änderungen im taxonomischen und funktionellen Profil der mikrobiellen Gemeinschaften in Biogasanlagen könnten zukünftig als Indikator für veränderte Prozessparameter oder Prozessstörungen genutzt werden. Voraussetzung dafür sind weitere Metaproteomanalysen, wie sie momentan im Biogasmessprogramm 3 (FNR Projekt FKZ 22012513) erhoben werden. Detaillierte Metaproteomanalysen könnten aber auch zur Aufklärung syntropher Interaktionen, wie der syntrophen Acetatoxidation im Biogasprozess, beitragen [Westerholm et al. (2016)]. Für das Verständnis dieser am thermodynamischen Limit ablaufenden Reaktionen wäre eine ergänzende Bestimmung von Metaboliten, wie dem H_2 und den kurzkettigen Fettsäuren sinnvoll [Schink (1997)]. Taxonomische und funktionelle Informationen aus Metaproteomdaten könnten außerdem für die Modellierung mikrobieller Gemeinschaften genutzt werden [Hanemaaijer et al. (2015)]. Der Abgleich mit einem stöchiometrischen Netzwerkmodell des Biogasprozesses [Koch et al. (2016)] oder dem ADM1 Modell [Batstone et al. (2002)] könnte es wahrscheinlich ermöglichen, Störungen der mikrobiellen Gemeinschaften von Biogasanlagen vorherzusagen [Rabelo et al. (2011)].

Weiterführend könnten auch mikrobielle Gemeinschaften charakterisiert werden, die komplexe Biomassen effizienter aufschließen können als die mikrobiellen Gemeinschaften in Biogasanlagen. Vielversprechend wäre in diesem Zusammenhang die Untersuchung von mikrobiellen Gemeinschaft im Pansen von Wiederkäuern [Toyoda et al. (2009)], Termiten [Burnum et al. (2010)] oder Pandabären. Diesen ist gemeinsam, dass sie Cellulose bzw. sogar Lignocellulose effizient und mit deutlich kürzeren Verweilzeiten als in Biogasanlagen abbauen. Möglicherweise kann die Funktionsweise von mikrobiellen Gemeinschaften im Verdauungssystem von diesen Tieren auf Biogasanlagen übertragen.

Literatur

- Abram, F., Enright, A.-M., O'Reilly, J., Botting, C., Collins, G., and O'Flaherty, V. 2011. A metaproteomic approach gives functional insights into anaerobic digestion. *Journal of Applied Microbiology*, 110(6):1550–1560.
- Abram, F., Gunnigle, E., and O'Flaherty, V. 2009. Optimisation of protein extraction and 2-de for metaproteomics of microbial communities from anaerobic wastewater treatment biofilms. *Electrophoresis*, 30(23):4149–4151.
- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V., and Polz, M. F. 2004. Divergence and redundancy of 16s rRNA sequences in genomes with multiple rRNA operons. *Journal of Bacteriology*, 186(9):2629–2635.
- Acland, A., Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., Bryant, S. H., Canese, K., Church, D. M., et al. 2013. Database resources of the national center for biotechnology information. *Nucleic Acids Research*, 41(D1):D8–D20.
- Agentur für erneuerbare Energien 2014. Strommix in deutschland 2014. <http://www.unendlich-viel-energie.de/media/>.
- Aguilar, M.-I. 2004. *HPLC of Peptides and Proteins*. Springer.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.
- Amann, R. I., Ludwig, W., and Schleifer, K.-H. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, 59(1):143–169.
- Angelidaki, I. and Ellegaard, L. 2003. Codigestion of manure and organic wastes in centralized biogas plants. *Applied Biochemistry and Biotechnology*, 109(1-3):95–105.
- Angelidaki, I., Karakashev, D., Batstone, D. J., Plugge, C. M., and Stams, A. J. 2011. Biomethanation and its potential. *Methods in Enzymology*, 494:327–351.

- Appels, L., Baeyens, J., Degève, J., and Dewil, R. 2008. Principles and potential of the anaerobic digestion of waste-activated sludge. *Progress in Energy and Combustion Sciences*, 34(6):755–781.
- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H. B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E. G., Wang, J., Guarner, F., Pedersen, O., de Vos, W. M., Brunak, S., Doré, J., , M. I. T. C., Antolín, M., Artiguenave, F., Blottiere, H. M., Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariáz, G., Dervyn, R., Foerstner, K. U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le Roux, K., Maguin, E., Mérieux, A., Melo Minardi, R., M'rini, C., Muller, J., Oozeer, R., Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K., Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S. D., and Bork, P. 2011. Enterotypes of the human gut microbiome. *Nature*, 473(7346):174–180.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., et al. 2000. Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- Bairoch, A. 2000. The enzyme database in 2000. *Nucleic Acids Research*, 28(1):304–305.
- Barnouin, K. 2011. Guidelines for experimental design and data analysis of proteomic mass spectrometry-based experiments. *Amino Acids*, 40(2):259–260.
- Bastian, M., Heymann, S., Jacomy, M., et al. 2009. Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8:361–362.
- Bastida, F., Moreno, J., Nicolas, C., Hernandez, T., and Garcia, C. 2009. Soil metaproteomics: a review of an emerging environmental science. significance, methodology and perspectives. *European Journal of Soil Science*, 60(6):845–859.
- Batstone, D. J., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S., Rozzi, A., Sanders, W., Siegrist, H., and Vavilin, V. 2002. The iwa anaerobic digestion model no 1(adm 1). *Water Science & Technology*, 45(10):65–73.

- Becher, D., Bernhardt, J., Fuchs, S., and Riedel, K. 2013. Metaproteomics to unravel major microbial players in leaf litter and soil environments: challenges and perspectives. *Proteomics*, 13(18-19):2895–2909.
- Belasco, J. G. and Brawerman, G. 1993. *Control of messenger RNA stability*. Elsevier.
- Bell, A., Deutsch, E., Au, C., Kearney, R., Beavis, R., Sechi, S., Nilsson, T., and Bergeron, J. 2009. the human proteome organization (hupo) test sample working group. “a hupo test sample study reveals common problems in mass spectrometry-based proteomics.” *Nature Methods*, 6(6).
- Benjamini, Y. 1988. Opening the box of a boxplot. *The American Statistician*, 42(4):257–262.
- Benndorf, D., Balcke, G. U., Harms, H., and Von Bergen, M. 2007. Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *The ISME Journal*, 1(3):224–234.
- Benndorf, D. and Reichl, U. 2014. Proteomics in environmental and technical microbiology. *Engineering in Life Sciences*, 14(1):27–46.
- Bensmann, A., Hanke-Rauschenbach, R., Heyer, R., Kohrs, F., Benndorf, D., Reichl, U., and Sundmacher, K. 2014. Biological methanation of hydrogen within biogas plants: A model-based feasibility study. *Applied Energy*, 134:413–425.
- Bensmann, A., Hanke-Rauschenbach, R., and Sundmacher, K. 2013. Reactor configurations for biogas plants—a model based analysis. *Chemical Engineering Science*, 104:413–426.
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218):53–59.
- Berg, M., Parbel, A., Pettersen, H., Fenyő, D., and Björkesten, L. 2006. Reproducibility of lc-ms-based protein identification. *Journal of Experimental Botany*, 57(7):1509–1514.
- Bialek, K., Cysneiros, D., and O’Flaherty, V. 2013. Low-temperature (10 c) anaerobic digestion of dilute dairy wastewater in an egsb bioreactor: microbial community structure, population dynamics, and kinetics of methanogenic populations. *Archaea*, 2013.

- Biemann, K. 1988. Contributions of mass spectrometry to peptide and protein structure. *Biological Mass Spectrometry*, 16(1-12):99–111.
- Binner, R., Menath, V., Huber, H., Thomm, M., Bischof, F., Schmack, D., and Reuter, M. 2011. Comparative study of stability and half-life of enzymes and enzyme aggregates implemented in anaerobic biogas processes.
- Bio-Rad, Laboratory Inc. 2015. A guide to polyacrylamide gel electrophoresis and detection. Technical report, Bio-Rad, Laboratory Inc.
- Blagodatskaya, E. and Kuzyakov, Y. 2013. Active microorganisms in soil: Critical review of estimation criteria and approaches. *Soil Biology and Biochemistry*, 67:192–211.
- Blum, H., Beier, H., and Gross, H. J. 1987. Improved silver staining of plant proteins, rna and dna in polyacrylamide gels. *Electrophoresis*, 8(2):93–99.
- Blume, F., Bergmann, I., Nettmann, E., Schelle, H., Rehde, G., Mundt, K., and Klocke, M. 2010. Methanogenic population dynamics during semi-continuous biogas fermentation and acidification by overloading. *Journal of Applied Microbiology*, 109(2):441–450.
- Boja, E. S. and Fales, H. M. 2001. Overalkylation of a protein digest with iodoacetamide. *Analytical Chemistry*, 73(15):3576–3582.
- Boyle, W. 1976. Energy recovery from sanitary landfills—a review. *Microbial Energy Conversion*, (6).
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72(1):248–254.
- Bradshaw, R. A., Burlingame, A. L., Carr, S., and Aebersold, R. 2006. Reporting protein identification data the next generation of guidelines. *Molecular & Cellular Proteomics*, 5(5):787–788.
- Brenner, D., Krieg, N., and Staley, J. 2005. The proteobacteria (part c): the alpha-, beta-, delta-, and epsilonproteobacteria. *Garrity GM. Bergey's Manual of Systematic Bacteriology*, pages 1059–1144.
- Brun, V., Masselon, C., Garin, J., and Dupuis, A. 2009. Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, 72(5):740–749.

- Bundesministerium für Wirtschaft und Energie 2015. Entwicklung der erneuerbaren Energien in Deutschland im Jahr 2014. http://www.erneuerbare-energien.de/EE/Navigation/DE/Service/Erneuerbare_Energien_in_Zahlen/Entwicklung_de_erneuerbaren_Energien_in_Deutschland/entwicklung_der_erneuerbaren_energien_in_deutschland_im_jahr_2014.html.
- Burnum, K. E., Callister, S. J., Nicora, C. D., Purvine, S. O., Hugenholtz, P., Warnecke, F., Scheffrahn, R. H., Smith, R. D., and Lipton, M. S. 2010. Proteome insights into the symbiotic relationship between a captive colony of *Nasutitermes corniger* and its hindgut microbiome. *The ISME Journal*, 5(1):161–164.
- Buschhorn, H., Dürre, P., and Gottschalk, G. 1989. Production and utilization of ethanol by the homoacetogen acetobacterium *woodii*. *Applied and Environmental Microbiology*, 55(7):1835–1840.
- Buswell, A. and Hatfield, W. 1936. Anaerobic fermentations. *Bull.*, 32.
- Cain, J. A., Solis, N., and Cordwell, S. J. 2014. Beyond gene expression: The impact of protein post-translational modifications in bacteria. *Journal of Proteomics*, 97:265–286.
- Campanaro, S., Treu, L., Kougias, P. G., Francisci, D., Valle, G., and Angelidaki, I. 2016. Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. *Biotechnology for Biofuels*, 9(1):1.
- Caspi, R., Altman, T., Dreher, K., Fulcher, C. A., Subhraveti, P., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., et al. 2012. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Research*, 40(D1):D742–D753.
- Chen, Y., Cheng, J. J., and Creamer, K. S. 2008. Inhibition of anaerobic digestion process: a review. *Bioresource Technology*, 99(10):4044–4064.
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of Proteome Research*, 10(4):1794–1805.
- Craig, R. and Beavis, R. C. 2004. Tandem: matching proteins with tandem mass spectra. *Bioinformatics*, 20(9):1466–1467.
- Criquet, S., Farnet, A., and Ferre, E. 2002. Protein measurement in forest litter. *Biology and Fertility of Soils*, 35(5):307–313.

- Croft, D., O’Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., et al. 2011. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Research*, 39(suppl 1):D691–D697.
- Davidson, A. L. and Chen, J. 2004. Atp-binding cassette transporters in bacteria. *Annual Review of Biochemistry*, 73(1):241–268.
- De Vrieze, J., Hennebel, T., Boon, N., and Verstraete, W. 2012. Methanosarcina: The re-discovered methanogen for heavy duty biomethanation. *Bioresource Technology*, 112:1–9.
- De Vrieze, J., Saunders, A. M., He, Y., Fang, J., Nielsen, P. H., Verstraete, W., and Boon, N. 2015. Ammonia and temperature determine potential clustering in the anaerobic digestion microbiome. *Water Research*, 75:312–323.
- deRoy, K. 2014. *Microbial resource management: introducing new tools and ecological theories*. PhD thesis, Ghent University, Gent, Belgium.
- Desjardins, P., Hansen, J. B., and Allen, M. 2009. Microvolume protein concentration determination using the nanodrop 2000c spectrophotometer. *Journal of Visualized Experiments: JoVE*, 33:1610.
- Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., and Mechtler, K. 2014. Ms amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *Journal of Proteome Research*, 13(8):3679–3684.
- Dunn, M. 2000. *From Genome to Proteome*. Wiley-Blackwell.
- Dykstra, A. B., St Brice, L., Rodriguez, Jr, M., Raman, B., Izquierdo, J., Cook, K. D., Lynd, L. R., and Hettich, R. L. 2014. Development of a multipoint quantitation method to simultaneously measure enzymatic and structural components of the clostridium thermocellum cellulosome protein complex. *Journal of Proteome Research*, 13(2):692–701.
- Eckert, W. A. and Kartenbeck, J. 1997. *Proteine: Standardmethoden der Molekular-und Zellbiologie*. Springer DE.
- Egertson, J. D., Kuehn, A., Merrihew, G. E., Bateman, N. W., MacLean, B. X., Ting, Y. S., Canterbury, J. D., Marsh, D. M., Kellmann, M., Zabrouskov, V., et al. 2013. Multiplexed ms/ms for improved data-independent acquisition. *Nature Methods*, 10(8):744–746.

- Ehlers, M. M. and Cloete, T. E. 1999. Comparing the protein profiles of 21 different activated sludge systems after sds-page. *Water Research*, 33(5):1181–1186.
- Eichler, J. and Maupin-Furlow, J. 2013. Post-translation modification in archaea: lessons from haloferax volcanii and other haloarchaea. *FEMS Microbiology Reviews*, 37(4):583–606.
- Eikmeyer, F. G., Rademacher, A., Hanreich, A., Hennig, M., Jaenicke, S., Maus, I., Wibberg, D., Zakrzewski, M., Pühler, A., Klocke, M., et al. 2013. Detailed analysis of metagenome datasets obtained from biogas-producing microbial communities residing in biogas reactors does not indicate the presence of putative pathogenic microorganisms. *Biotechnology for Biofuels*, 6(1):49.
- Elias, J. E. and Gygi, S. P. 2010. Target-decoy search strategy for mass spectrometry-based proteomics. *Proteome Bioinformatics*, pages 55–71.
- Elias, J. E., Haas, W., Faherty, B. K., and Gygi, S. P. 2005. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nature Methods*, 2(9):667–675.
- Erickson, A. R., Cantarel, B. L., Lamendella, R., Darzi, Y., Mongodin, E. F., Pan, C., Shah, M., Halfvarson, J., Tysk, C., Henrissat, B., Raes, J., Verberkmoes, N. C., Fraser, C. M., Hettich, R. L., and Jansson, J. K. 2012. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of crohn’s disease. *PLoS One*, 7(11):e49138.
- Eschelbach, J. W. and Jorgenson, J. W. 2006. Improved protein recovery in reversed-phase liquid chromatography by the use of ultrahigh pressures. *Analytical Chemistry*, 78(5):1697–1706.
- Fachagentur Nachwachsende Rohstoffe 2010. *Leitfaden Biogas: von der Gewinnung zur Nutzung*; FNR.
- Fardeau, M.-L., Peillex, J.-P., and Belaich, J.-P. 1987. Energetics of the growth of methanobacterium thermoautotrophicum and methanococcus thermolithotrophicus on ammonium chloride and dinitrogen. *Archives of Microbiology*, 148(2):128–131.
- Federhen, S. 2012. The ncbi taxonomy database. *Nucleic Acids Research*, 40(D1):D136–D143.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. 1989. Electro-spray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64–71.

- Ferguson, D. and Krzycki, J. A. 1997. Reconstitution of trimethylamine-dependent coenzyme m methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase ii from methanosarcina barkeri. *Journal of Bacteriology*, 179(3):846–852.
- Ferry, J. G. 2011. Fundamentals of methanogenic pathways that are key to the biomechanation of complex biomass. *Current Opinion in Biotechnology*, 22(3):351–357.
- Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., et al. 2010. The pfam protein families database. *Nucleic Acids Research*, 38(suppl 1):D211–D222.
- Frank, A. and Pevzner, P. 2005. Pepnovo: de novo peptide sequencing via probabilistic network modeling. *Analytical Chemistry*, 77(4):964–973.
- Fronk, K. 2013. Stöchiometrische Netzwerkanalyse der Biogasbildung. Master’s thesis, OvGU.
- Gasch, C., Hildebrandt, I., Rebbe, F., and Röske, I. 2013. Enzymatic monitoring and control of a two-phase batch digester leaching system with integrated anaerobic filter. *Energy, Sustainability and Society*, 3(1):1–11.
- Gatto, L., Vizcaíno, J. A., Hermjakob, H., Huber, W., and Lilley, K. S. 2010. Organelle proteomics experimental designs and analysis. *Proteomics*, 10(22):3957–3969.
- Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. 2004. Open mass spectrometry search algorithm. *Journal of Proteome Research*, 3(5):958–964.
- Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem ms. *Proceedings of the National Academy of Sciences*, 100(12):6940–6945.
- Goldfarb, A. R., Saidel, L. J., and Mosovich, E. 1951. The ultraviolet absorption spectra of proteins. *Journal of Biological Chemistry*, 193(1):397–404.
- Gonnelli, G., Stock, M., Verwaeren, J., Maddelein, D., De Baets, B., Martens, L., and Degroeve, S. 2015. A decoy-free approach to the identification of peptides. *Journal of Proteome Research*, 14(4):1792–1798.
- Gotelli, N. J., Ellison, A. M., and Ballif, B. A. 2012. Environmental proteomics, biodiversity statistics and food-web structure. *Trends in Ecology & Evolution*, 27(8):436–442.

- Grotenhuis, J., Smit, M., Plugge, C., Xu, Y., Van Lammeren, A., Stams, A., and Zehnder, A. 1991. Bacteriological composition and structure of granular sludge adapted to different substrates. *Applied and Environmental Microbiology*, 57(7):1942–1949.
- Gunnigle, E., Siggins, A., Botting, C. H., Fuszard, M., O’Flaherty, V., and Abram, F. 2015. Low-temperature anaerobic digestion is associated with differential methanogenic protein expression. *FEMS Microbiology Letters*, 362(10):fnv059.
- Gupta, N. and Pevzner, P. A. 2009. False discovery rates of protein identifications: a strike against the two-peptide rule. *Journal of Proteome Research*, 8(9):4173–4181.
- Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology*, 17(10):994–999.
- Haange, S.-B., Oberbach, A., Schlichting, N., Hugenholtz, F., Smidt, H., von Bergen, M., Till, H., and Seifert, J. 2012. Metaproteome analysis and molecular genetics of rat intestinal microbiota reveals section and localization resolved species distribution and enzymatic functionalities. *Journal of Proteome Research*, 11(11):5406–5417.
- Haft, D. H., Selengut, J. D., Richter, R. A., Harkins, D., Basu, M. K., and Beck, E. 2013. Tigrfams and genome properties in 2013. *Nucleic Acids Research*, 41(D1):D387–D395.
- Hanemaaijer, M., Röling, W. F., Olivier, B. G., Khandelwal, R. A., Teusink, B., and Bruggeman, F. J. 2015. Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure. *Frontiers in Microbiology*, 6.
- Hanreich, A., Heyer, R., Benndorf, D., Rapp, E., Pioch, M., Reichl, U., and Klocke, M. 2012. Metaproteome analysis to determine the metabolically active part of a thermophilic microbial community producing biogas from agricultural biomass. *Canadian Journal of Microbiology*, 58(7):917–922.
- Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., Heyer, R., Rapp, E., Pühler, A., Reichl, U., and Klocke, M. 2013. Metagenome and metaproteome analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation. *Systematic and Applied Microbiology*, 36(5):330–338.
- Hao, P., Ren, Y., Alpert, A. J., and Sze, S. K. 2011. Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Molecular & Cellular Proteomics*, 10(10):O111–009381.

- Hartmann, H. and Ahring, B. K. 2005. Anaerobic digestion of the organic fraction of municipal solid waste: influence of co-digestion with manure. *Water Research*, 39(8):1543–1552.
- Hattori, S. 2008. Syntrophic acetate-oxidizing microbes in methanogenic environments. *Microbes and Environments*, 23(2):118–127.
- Hedderich, R. and Whitman, W. B. 2006. Physiology and biochemistry of the methane-producing archaea. In *The Prokaryotes*, pages 1050–1079. Springer.
- Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. 2007. Mascotdatfile: An open-source library to fully parse and analyse mascot ms/ms search results. *Proteomics*, 7(3):364–366.
- Hettich, R. L., Pan, C., Chourey, K., and Giannone, R. J. 2013. Metaproteomics: Harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Analytical Chemistry*, 85(9):4203–4214.
- Heyer, R., Benndorf, D., Kohrs, F., De Vrieze, J., Boon, N., Hoffmann, M., Rapp, E., Schlüter, A., Sczyrba, A., and Reichl, U. 2016. Proteotyping of biogas plant microbiomes separates biogas plants according to process temperature and reactor type. *Biotechnology for Biofuels*, 9:155.
- Heyer, R., Kohrs, F., Benndorf, D., Rapp, E., Kausmann, R., Heiermann, M., Klocke, M., and Reichl, U. 2013. Metaproteome analysis of the microbial communities in agricultural biogas plants. *New Biotechnology*, 30(6):614–622.
- Heyer, R., Kohrs, F., Reichl, U., and Benndorf, D. 2015. Metaproteomics of complex microbial communities in biogas plants. *Microbial Biotechnology*.
- Heyer, R. S. 2011. Diplomarbeit: Advanced metaproteome analysis of biogas plant samples. Master’s thesis, Otto-von-Guericke-Universität Magdeburg.
- Hofman-Bang, J., Zheng, D., Westermann, P., Ahring, B. K., and Raskin, L. 2003. Molecular ecology of anaerobic reactor systems. In *Biomethanation I*, pages 151–203. Springer.
- Holm-Nielsen, J. B., Al Seadi, T., and Oleskowicz-Popiel, P. 2009. The future of anaerobic digestion and biogas utilization. *Bioresource Technology*, 100(22):5478–5484.

- Huber, R. and Hannig, M. 2006. Thermotogales. In *The Prokaryotes*, pages 899–922. Springer.
- Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A., Bernard, T., Binns, D., Bork, P., Burge, S., et al. 2012. Interpro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Research*, 40(D1):D306–D312.
- Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. 2007. Megan analysis of metagenomic data. *Genome Research*, 17(3):377–386.
- Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N., and Schuster, S. C. 2011. Integrative analysis of environmental sequences using megan4. *Genome Research*, 21(9):1552–1560.
- Ibrahim, A. and Ahring, B. 1999. Extraction of intact ribosomal rna from anaerobic bioreactor samples for molecular ecological studies. *Biotechniques*, 27(6):1132–1138.
- Im, J.-F., McGuffin, M. J., and Leung, R. 2013. Gplom: the generalized plot matrix for visualizing multidimensional multivariate data. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2606–2614.
- Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. 2005. Exponentially modified protein abundance index (empai) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics*, 4(9):1265–1272.
- Jagtap, P., Goslinga, J., Kooren, J. A., McGowan, T., Wroblewski, M. S., Seymour, S. L., and Griffin, T. J. 2013. A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. *Proteomics*, 13(8):1352–1357.
- Jagtap, P. D., Blakely, A., Murray, K., Stewart, S., Kooren, J., Johnson, J. E., Rhodus, N. L., Rudney, J., and Griffin, T. J. 2015. Metaproteomic analysis using the galaxy framework. *Proteomics*, 15(20):3553–3565.
- Jehmlich, N., Hübschmann, T., Salazar, M. G., Völker, U., Benndorf, D., Müller, S., von Bergen, M., and Schmidt, F. 2010. Advanced tool for characterization of microbial cultures by combining cytomics and proteomics. *Applied Microbiology and Biotechnology*, 88(2):575–584.

- Jehmlich, N., Schmidt, F., Hartwich, M., von Bergen, M., Richnow, H.-H., and Vogt, C. 2008. Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (protein-sip). *Rapid Communications in Mass Spectrometry*, 22(18):2889–2897.
- Jørgensen, P. J. 2009. Biogas–green energy. *Process, Design, Energy supply, Environment. Faculty of Agricultural Sciences, Aarhus University*.
- Joshi-Tope, G., Gillespie, M., Vastrik, I., D’Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G. R., Wu, G. R., Matthews, L., Lewis, S., Birney, E., and Stein, L. 2005. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Research*, 33(Database issue):D428–D432.
- Justice, N. B., Pan, C., Mueller, R., Spaulding, S. E., Shah, V., Sun, C. L., Yelton, A. P., Miller, C. S., Thomas, B. C., Shah, M., VerBerkmoes, N., Hettich, R., and Banfield, J. F. 2012. Heterotrophic archaea contribute to carbon cycling in low-ph, suboxic biofilm communities. *Applied and Environmental Microbiology*, 78(23):8321–8330.
- Kabdasli, I., Tnay, O., Yilmaz, S., Arikan, O., et al. 2000. Ammonia removal from young landfill leachate by magnesium ammonium phosphate precipitation and air stripping. *Water Science and Technology*, 41(1):237–240.
- Kallow, W., Erhard, M., Shah, H. N., Raptakis, E., and Welker, M. 2010. Maldi-tof ms for microbial identification: Years of experimental development to an established protocol. *Mass Spectrometry for Microbial Proteomics*, page 255.
- Kalume, D. E., Molina, H., and Pandey, A. 2003. Tackling the phosphoproteome: tools and strategies. *Current Opinion in Chemical Biology*, 7(1):64–69.
- Kan, J., Hanson, T. E., Ginter, J. M., Wang, K., and Chen, F. 2005. Metaproteomic analysis of chesapeake bay microbial communities. *Saline Systems*, 1:7.
- Kanehisa, M. and Goto, S. 2000. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30.
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. 2004. The kegg resource for deciphering the genome. *Nucleic Acids Research*, 32(suppl 1):D277–D280.
- Karakashev, D., Batstone, D. J., Trably, E., and Angelidaki, I. 2006. Acetate oxidation is the dominant methanogenic pathway from acetate in the absence of methanosaetaceae. *Applied and Environmental Microbiology*, 72(7):5138–5141.

- Karlsson, R., Gonzales-Siles, L., Bouldund, F., Svensson-Stadler, L., Skovbjerg, S., Karlsson, A., Davidson, M., Hulth, S., Kristiansson, E., and Moore, E. R. 2015. Proteotyping: Proteomic characterization, classification and identification of microorganisms—a prospectus. *Systematic and Applied Microbiology*, 38(4):246–257.
- Kazda, M., Langer, S., and Bengelsdorf, F. R. 2014. Fungi open new possibilities for anaerobic fermentation of organic residues. *Energy, Sustainability and Society*, 4(1):6.
- Keiblinger, K. M., Wilhartitz, I. C., Schneider, T., Roschitzki, B., Schmid, E., Eberl, L., Riedel, K., and Zechmeister-Boltenstern, S. 2012. Soil metaproteomics—comparative evaluation of protein extraction protocols. *Soil Biology and Biochemistry*, 54:14–24.
- Kendall, M. M. and Boone, D. R. 2006. The order methanosarcinales. In *The Prokaryotes*, pages 244–256. Springer.
- Kim, S., Mischerikow, N., Bandeira, N., Navarro, J. D., Wich, L., Mohammed, S., Heck, A. J., and Pevzner, P. A. 2010. The generating function of cid, etd, and cid/etd pairs of tandem mass spectra: applications to database search. *Molecular & Cellular Proteomics*, 9(12):2840–2852.
- Kim, Y. S., Westerholm, M., and Scherer, P. 2014. Dual investigation of methanogenic processes by quantitative pcr and quantitative microscopic fingerprinting. *FEMS Microbiology Letters*, 360(1):76–84.
- Klappenbach, J. A., Saxman, P. R., Cole, J. R., and Schmidt, T. M. 2001. rrndb: the ribosomal rna operon copy number database. *Nucleic Acids Research*, 29(1):181–184.
- Kletzin, A. and Adams, M. W. 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *thermotoga maritima*. *Journal of Bacteriology*, 178(1):248–257.
- Klocke, M., Nettmann, E., Bergmann, I., Mundt, K., Souidi, K., Mumme, J., and Linke, B. 2008. Characterization of the methanogenic archaea within two-phase biogas reactor systems operated with plant biomass. *Systematic and Applied Microbiology*, 31(3):190–205.
- Klose, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Humangenetik*, 26(3):231–243.
- Koch, S. 2014. Erweiterung und Validierung eines stöchiometrischen Netzwerkmodells der Biogasbildung. Master’s thesis, OvGU.

- Koch, S., Benndorf, D., Fronk, K., Reichl, U., and Klamt, S. 2016. Predicting compositions of microbial communities from stoichiometric models with applications for the biogas process. *Biotechnology for Biofuels*, 9(1):1.
- Kohrs, F. 2010. Entwicklung einer schnellen und miniaturisierten Extraktion für Proteine und Peptide aus Belebtschlamm. Master's thesis, OvGU.
- Kohrs, F. 2011. Diplomarbeit: Weiterentwicklung und Optimierung von Methoden für die Metaproteomanalyse von Belebtschlamm. Master's thesis, Otto-von-Guericke-Universität Magdeburg.
- Kohrs, F., Heyer, R., Magnussen, A., Benndorf, D., Muth, T., Behne, A., Rapp, E., Kausmann, R., Heiermann, M., Klocke, M., and Reichl, U. 2014. Sample prefractionation with liquid isoelectric focusing enables in depth microbial metaproteome analysis of mesophilic and thermophilic biogas plants. *Anaerobe*.
- Kohrs, F., Wolter, S., Benndorf, D., Heyer, R., Hoffmann, M., Rapp, E., Bremges, A., Sczyrba, A., Schlüter, A., and Reichl, U. 2015. Fractionation of biogas plant sludge material improves metaproteomic characterization to investigate metabolic activity of microbial communities. *Proteomics*.
- Kolmeder, C. A., De Been, M., Nikkilä, J., Ritamo, I., Mättö, J., Valmu, L., Salojärvi, J., Palva, A., Salonen, A., and de Vos, W. M. 2012. Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions. *PloS One*, 7(1):e29913.
- Kolmeder, C. A., Salojärvi, J., Ritari, J., de Been, M., Raes, J., Falony, G., Vieira-Silva, S., Kekkonen, R. A., Corthals, G. L., Palva, A., Salonen, A., and de Vos, W. M. 2016. Faecal metaproteomic analysis reveals a personalized and stable functional microbiome and limited effects of a probiotic intervention in adults. *PLoS One*, 11(4):e0153294.
- Krakat, N., Westphal, A., Schmidt, S., and Scherer, P. 2010. Anaerobic digestion of renewable biomass: thermophilic temperature governs methanogen population dynamics. *Applied and Environmental Microbiology*, 76(6):1842–1850.
- Kramer, G., Sprenger, R. R., Nessen, M. A., Roseboom, W., Speijer, D., de Jong, L., de Mattos, M. J. T., Back, J., and de Koster, C. G. 2010. Proteome-wide alterations in escherichia coli translation rates upon anaerobiosis. *Molecular and Cellular Proteomics*, 9(11):2508–2516.

- Krzycki, J., Wolkin, R., and Zeikus, J. 1982. Comparison of unitrophic and mixotrophic substrate metabolism by an acetate-adapted strain of *Methanosarcina barkeri*. *Journal of Bacteriology*, 149(1):247–254.
- Kuhn, R., Benndorf, D., Rapp, E., Reichl, U., Palese, L. L., and Pollice, A. 2011. Metaproteome analysis of sewage sludge from membrane bioreactors. *Proteomics*, 11(13):2738–2744.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *Nature*, 227(5259):680–685.
- Lam, H., Deutsch, E. W., Eddes, J. S., Eng, J. K., King, N., Stein, S. E., and Aebersold, R. 2007. Development and validation of a spectral library searching method for peptide identification from ms/ms. *Proteomics*, 7(5):655–667.
- Larkin, M. A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., et al. 2007. Clustal w and clustal x version 2.0. *Bioinformatics*, 23(21):2947–2948.
- Lauwers, J., Appels, L., Thompson, I. P., Degève, J., Van Impe, J. F., and Dewil, R. 2013. Mathematical modelling of anaerobic digestion of biomass and waste: Power and limitations. *Progress in Energy and Combustion Science*.
- Leary, D. H., Hervey IV, W. J., Deschamps, J. R., Kusterbeck, A. W., and Vora, G. J. 2014. Which metaproteome? the impact of protein extraction bias on metaproteomic analyses. *Molecular and Cellular Probes*.
- Lettinga, G., Rebac, S., and Zeeman, G. 2001. Challenge of psychrophilic anaerobic wastewater treatment. *TRENDS in Biotechnology*, 19(9):363–370.
- Levén, L., Eriksson, A. R., and Schnürer, A. 2007. Effect of process temperature on bacterial and archaeal communities in two methanogenic bioreactors treating organic household waste. *FEMS Microbiology Ecology*, 59(3):683–693.
- Levenshtein, V. I. 1966. Binary codes capable of correcting deletions, insertions and reversals. In *Soviet Physics Doklady*, volume 10, page 707.
- Lienen, T., Kleyböcker, A., Brehmer, M., Kraume, M., Moeller, L., Görsch, K., and Würdemann, H. 2013. Floating layer formation, foaming, and microbial community structure change in full-scale biogas plant due to disruption of mixing and substrate overloading. *Energy, Sustainability and Society*, 3(1):20.

- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R. 1999. Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology*, 17(7):676–682.
- Link, A. J. and LaBaer, J. 2011. Solution protein digest. *Cold Spring Harbor Protocols*, 2011(2):pdb-prot5569.
- Liu, Y. and Whitman, W. B. 2008. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences*, 1125(1):171–189.
- Lorowitz, W. H., Zhao, H., and Bryant, M. 1989. *Syntrophomonas wolfei* subsp. *saponavida* subsp. nov., a long-chain fatty-acid-degrading, anaerobic, syntrophic bacterium; *Syntrophomonas wolfei* subsp. *wolfei* subsp. nov.; and emended descriptions of the genus and species. *International Journal of Systematic and Evolutionary Microbiology*, 39(2):122–126.
- Lossie, U. and Pütz, P. 2008. Targeted control of biogas plants with the help of fos/tac. *Practice Report Hach-Lange*.
- Lottspeich, F. and Kellermann, J. 2011. Icpl labeling strategies for proteome research. In *Gel-Free Proteomics*, pages 55–64. Springer.
- Lovell, C. R., Przybyla, A., and Ljungdahl, L. G. 1990. Primary structure of the thermostable formyltetrahydrofolate synthetase from *Clostridium thermoaceticum*. *Biochemistry*, 29(24):5687–5694.
- Lowry, O. H., Rosebrough, N. J., Farr, A. L., Randall, R. J., et al. 1951. Protein measurement with the folin phenol reagent. *Journal of Biological Chemistry*, 193(1):265–275.
- Lü, F., Bize, A., Guillot, A., Monnet, V., Madigou, C., Chapleur, O., Mazéas, L., He, P., and Bouchez, T. 2014. Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. *The ISME Journal*, 8(1):88–102.
- Lucas, R., Kuchenbuch, A., Fetzner, I., Harms, H., and Kleinstüber, S. 2015. Long-term monitoring reveals stable and remarkably similar microbial communities in parallel full-scale biogas reactors digesting energy crops. *FEMS Microbiology Ecology*, 91(3):fiv004.
- Lüders, S., Fallet, C., and Franco-Lara, E. 2009. Proteome analysis of the *Escherichia coli* heat shock response under steady-state conditions. *Proteome Science*, 7(1):1.

- Ludwig, W. and Klenk, H.-P. 2001. Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In *Bergey's manual® of systematic bacteriology*, pages 49–65. Springer.
- Lundgren, D. H., Hwang, S.-I., Wu, L., and Han, D. K. 2010. Role of spectral counting in quantitative proteomics. *Expert Review of Proteomics*, 7(1):39–53.
- Luo, G., Johansson, S., Boe, K., Xie, L., Zhou, Q., and Angelidaki, I. 2012. Simultaneous hydrogen utilization and in situ biogas upgrading in an anaerobic reactor. *Biotechnology and Bioengineering*, 109(4):1088–1094.
- Madigan, M. T. and Martinko, J. M. 2009. *Brock Mikrobiologie*. Pearson Deutschland GmbH.
- Maier, T., Schmidt, A., Güell, M., Kühner, S., Gavin, A.-C., Aebersold, R., and Serrano, L. 2011. Quantification of mrna and protein and integration with protein turnover in a bacterium. *Molecular Systems Biology*, 7(1).
- Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. 2005. Automated genome annotation and pathway identification using the kegg orthology (ko) as a controlled vocabulary. *Bioinformatics*, 21(19):3787–3793.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057):376–380.
- Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. 1998. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrometry Reviews*, 17(1):1–35.
- Marzorati, M., Wittebolle, L., Boon, N., Daffonchio, D., and Verstraete, W. 2008. How to get more out of molecular fingerprints: practical tools for microbial ecology. *Environmental Microbiology*, 10(6):1571–1581.
- Meier, F., Beck, S., Grassl, N., Lubeck, M., Park, M. A., Raether, O., and Mann, M. 2015. Parallel accumulation–serial fragmentation (pasef): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device. *Journal of Proteome Research*, 14(12):5378–5387.
- Melbinger, N., Donnellon, J., and Zablatzky, H. R. 1971. Toxic effects of ammonia nitrogen in high-rate digestion [with discussion]. *Journal (Water Pollution Control Federation)*, pages 1658–1670.

- Mertens, B., Boon, N., and Verstraete, W. 2005. Stereospecific effect of hexachlorocyclohexane on activity and structure of soil methanotrophic communities. *Environmental Microbiology*, 7(5):660–669.
- Mesuere, B., Devreese, B., Debysen, G., Aerts, M., Vandamme, P., and Dawyndt, P. 2012. Unipept: tryptic peptide-based biodiversity analysis of metaproteome samples. *Journal of Proteome Research*, 11(12):5773–5780.
- Michalski, A., Cox, J., and Mann, M. 2011. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent lc- ms/ms. *Journal of Proteome Research*, 10(4):1785–1793.
- Mitchell Wells, J. and McLuckey, S. A. 2005. Collision-induced dissociation (cid) of peptides and proteins. *Methods in Enzymology*, 402:148–185.
- Müller, B., Sun, L., and Schnürer, A. 2013. First insights into the syntrophic acetate-oxidizing bacteria—a genetic study. *Microbiologyopen*, 2(1):35–53.
- Müller, S., Hübschmann, T., Kleinstüber, S., and Vogt, C. 2012. High resolution single cell analytics to follow microbial community dynamics in anaerobic ecosystems. *Methods*, 57(3):338–349.
- Munk, B., Bauer, C., Gronauer, A., and Lebuhn, M. 2010. Population dynamics of methanogens during acidification of biogas fermenters fed with maize silage. *Engineering in Life Sciences*, 10(6):496–508.
- Munk, B., Bauer, C., Gronauer, A., Lebuhn, M., et al. 2012. A metabolic quotient for methanogenic archaea. *Water Science and Technology*, 66(11):2311.
- Muth, T., Behne, A., Heyer, R., Kohrs, F., Benndorf, D., Hoffmann, M., Lehteva, M., Reichl, U., Martens, L., and Rapp, E. 2015a. The metaproteome analyzer: a powerful open-source software suite for metaproteomics data analysis and interpretation. *Journal of Proteome Research*, 14(3):1557–1565.
- Muth, T., Benndorf, D., Reichl, U., Rapp, E., and Martens, L. 2013a. Searching for a needle in a stack of needles: challenges in metaproteomics data analysis. *Molecular Biosystems*, 9(4):578–585.
- Muth, T., Kolmeder, C. A., Salojärvi, J., Keskitalo, S., Varjosalo, M., Verdam, F. J., Rensen, S. S., Reichl, U., Vos, W. M., Rapp, E., et al. 2015b. Navigating through metaproteomics data: A logbook of database searching. *Proteomics*, 15(20):3439–3453.

- Muth, T., Peters, J., Blackburn, J., Rapp, E., and Martens, L. 2013b. Proteocloud: A full-featured open source proteomics cloud computing pipeline. *Journal of Proteomics*, 88:104–108.
- Muth, T., Weilnbo ck, L., Rapp, E., Huber, C. G., Martens, L., Vaudel, M., and Barsnes, H. 2014. Denovogui: an open source graphical user interface for de novo sequencing of tandem mass spectra. *Journal of Proteome Research*, 13(2):1143–1146.
- Myers, J. L., Well, A., and Lorch, R. F. 2010. *Research design and statistical analysis*. Routledge.
- Nader, W., Nebe, G., Nebe, C., and Birr, C. 1995. Process for the quantification of methane gas bacteria.
- Nagaraj, N., Kulak, N. A., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. 2012. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra hplc runs on a bench top orbitrap. *Molecular & Cellular Proteomics*, 11(3):M111–013722.
- Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., van Sluyter, S. C., and Haynes, P. A. 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics*, 11(4):535–553.
- Nelson, M. C., Morrison, M., and Yu, Z. 2011. A meta-analysis of the microbial diversity observed in anaerobic digesters. *Bioresource Technology*, 102(4):3730–3739.
- Nesvizhskii, A. I. and Aebersold, R. 2005. Interpretation of shotgun proteomic data the protein inference problem. *Molecular & Cellular Proteomics*, 4(10):1419–1440.
- Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry*, 75(17):4646–4658.
- Nettmann, E., Bergmann, I., Pramschüfer, S., Mundt, K., Plogsties, V., Herrmann, C., and Klocke, M. 2010. Polyphasic analyses of methanogenic archaeal communities in agricultural biogas plants. *Applied and Environmental Microbiology*, 76(8):2540–2548.
- Nettmann, E., Fröhling, A., Heeg, K., Klocke, M., Schlüter, O., and Mumme, J. 2013. Development of a flow-fluorescence in situ hybridization protocol for the analysis of microbial communities in anaerobic fermentation liquor. *BMC Microbiology*, 13(1):278.

- Neuhoff, V., Stamm, R., and Eibl, H. 1985. Clear background and highly sensitive protein staining with coomassie blue dyes in polyacrylamide gels: a systematic analysis. *Electrophoresis*, 6(9):427–448.
- Nikolausz, M., Walter, R., Sträuber, H., Liebetrau, J., Schmidt, T., Kleinsteuber, S., Bratfisch, F., Günther, U., and Richnow, H. 2013. Evaluation of stable isotope fingerprinting techniques for the assessment of the predominant methanogenic pathways in anaerobic digesters. *Applied Microbiology and Biotechnology*, 97(5):2251–2262.
- Nisman, B. 1954. The stickland reaction. *Bacteriological Reviews*, 18(1):16.
- Nozhevnikova, A. N., Nekrasova, V., Ammann, A., Zehnder, A. J. B., Wehrli, B., and Holliger, C. 2007. Influence of temperature and high acetate concentrations on methanogenesis in lake sediment slurries. *FEMS Microbiol Ecol*, 62(3):336–344.
- O’Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry*, 250(10):4007–4021.
- Okamura-Ikeda, K., Fujiwara, K., and Motokawa, Y. 1987. Mechanism of the glycine cleavage reaction. properties of the reverse reaction catalyzed by t-protein. *Journal of Biological Chemistry*, 262(14):6746–6749.
- Oliveros, J. 2007. Venny. an interactive tool for comparing lists with venn diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>.
- Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. 2007. Higher-energy c-trap dissociation for peptide modification analysis. *Nature Methods*, 4(9):709–712.
- Ondov, B. D., Bergman, N. H., and Phillippy, A. M. 2011. Interactive metagenomic visualization in a web browser. *BMC Bioinformatics*, 12(1):385.
- Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. 2002. Stable isotope labeling by amino acids in cell culture, silac, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics*, 1(5):376–386.
- Otto, A., Becher, D., and Schmidt, F. 2014. Quantitative proteomics in the field of microbiology. *Proteomics*, 14(4-5):547–565.

- Park, C. Y., Klammer, A. A., Käll, L., MacCoss, M. J., and Noble, W. S. 2008. Rapid and accurate peptide identification from tandem mass spectra. *Journal of Proteome Research*, 7(7):3022–3027.
- Patient, S., Wieser, D., Kleen, M., Kretschmann, E., Martin, M. J., and Apweiler, R. 2008. Uniprotjapi: a remote api for accessing uniprot data. *Bioinformatics*, 24(10):1321–1322.
- Patton, W. F. 2002. Detection technologies in proteome analysis. *Journal of Chromatography B*, 771(1):3–31.
- Pearson, K. 1895. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242.
- Pender, S., Toomey, M., Carton, M., Eardly, D., Patching, J. W., Colleran, E., and O’Flaherty, V. 2004. Long-term effects of operating temperature and sulphate addition on the methanogenic community structure of anaerobic hybrid reactors. *Water Research*, 38(3):619–630.
- Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18):3551–3567.
- Person, K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572.
- Piccolo, A. 2002. The supramolecular structure of humic substances: a novel understanding of humus chemistry and implications in soil science. *Advances in Agronomy*, 75:57–134.
- Pierce, E., Xie, G., Barabote, R. D., Saunders, E., Han, C. S., Detter, J. C., Richardson, P., Brettin, T. S., Das, A., Ljungdahl, L. G., and Ragsdale, S. W. 2008. The complete genome sequence of moorella thermoacetica (f. clostridium thermoaceticum). *Environmental Microbiology*, 10(10):2550–2573.
- Podwojski, K., Eisenacher, M., Kohl, M., Turewicz, M., Meyer, H. E., Rahnenführer, J., and Stephan, C. 2010. Peek a peak: a glance at statistics for quantitative label-free proteomics. *Expert Review of Proteomics*, 7(2):249–261.
- Poehlein, A., Cebulla, M., Ilg, M. M., Bengelsdorf, F. R., Schiel-Bengelsdorf, B., Whited, G., Andreesen, J. R., Gottschalk, G., Daniel, R., and Dürre, P. 2015. The complete genome sequence of clostridium aceticum: a missing link between rnf-and cytochrome-containing autotrophic acetogens. *mBio*, 6(5):e01168–15.

- Popov, N., Schmitt, M., Schulzeck, S., and Matthies, H. 1975. Eine störungsfreie mikromethode zur bestimmung des proteingehaltes in gewebehomogenaten. *Acta Biologica et Medica Germanica*, 34:1441–1446.
- Powell, S., Szklarczyk, D., Trachana, K., Roth, A., Kuhn, M., Muller, J., Arnold, R., Rattei, T., Letunic, I., Doerks, T., et al. 2012. eggnog v3. 0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research*, 40(D1):D284–D289.
- Püttker, S., Kohrs, F., Benndorf, D., Heyer, R., Rapp, E., and Reichl, U. 2015. Metaproteomics of activated sludge from a wastewater treatment plant—a pilot study. *Proteomics*.
- Quinlan, J. R. 1986. Induction of decision trees. *Machine Learning*, 1(1):81–106.
- Rabelo, S., Carrere, H., Maciel Filho, R., and Costa, A. 2011. Production of bioethanol, methane and heat from sugarcane bagasse in a biorefinery concept. *Bioresource Technology*, 102(17):7887–7895.
- Rademacher, A., Zakrzewski, M., Schlüter, A., Schönberg, M., Szczepanowski, R., Goemann, A., Pühler, A., and Klocke, M. 2012. Characterization of microbial biofilms in a thermophilic biogas system by high-throughput metagenome sequencing. *FEMS Microbiology Ecology*, 79(3):785–799.
- Rajoka, M., Tabassum, R., and Malik, K. 1999. Enhanced rate of methanol and acetate uptake for production of methane in batch cultures using *Methanosarcina mazei*. *Bioresource Technology*, 67(3):305–311.
- Ram, R. J., VerBerkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C., Shah, M., Hettich, R. L., and Banfield, J. F. 2005. Community proteomics of a natural microbial biofilm. *Science*, 308(5730):1915–1920.
- Ramsay, I. R. and Pullammanappallil, P. C. 2001. Protein degradation during anaerobic wastewater treatment: derivation of stoichiometry. *Biodegradation*, 12(4):247–256.
- Refai, S., Berger, S., Wassmann, K., and Deppenmeier, U. 2014. Quantification of methanogenic heterodisulfide reductase activity in biogas sludge. *Journal of Biotechnology*, 180:66–69.
- Rüger, M., Ackermann, M., and Reichl, U. 2014. Species-specific viability analysis of *Pseudomonas aeruginosa*, *Burkholderia cepacia* and *Staphylococcus aureus* in mixed culture by flow cytometry. *BMC Microbiology*, 14:56.

- Ridley, W. P., Sidhu, R. S., Pyla, P. D., Nemeth, M. A., Breeze, M. L., and Astwood, J. D. 2002. Comparison of the nutritional profile of glyphosate-tolerant corn event nk603 with that of conventional corn (*zea mays* L.). *Journal of Agricultural and Food Chemistry*, 50(25):7235–7243.
- Riedel, K. and Britz, T. 1993. Propionibacterium species diversity in anaerobic digesters. *Biodiversity & Conservation*, 2(4):400–411.
- Rieger, C. and Weiland, P. 2006. Prozessstörungen frühzeitig erkennen. *Biogas Journal*, 4(06):18–20.
- Rodriguez-Valera, F. 2004. Environmental genomics, the big picture? *FEMS Microbiology Letters*, 231(2):153–158.
- Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. 1992. In-gel digestion of proteins for internal sequence analysis after one-or two-dimensional gel electrophoresis. *Analytical Biochemistry*, 203(1):173–179.
- Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., et al. 2004. Multiplexed protein quantitation in *saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics*, 3(12):1154–1169.
- Rousseeuw, P. J. and Bassett Jr, G. W. 1990. The remedian: A robust averaging method for large data sets. *Journal of the American Statistical Association*, 85(409):97–104.
- Rutherford, S. M. and Gilani, G. S. 2009. Amino acid analysis. *Current Protocols in Protein Science*, pages 11–9.
- Sahlström, L. 2003. A review of survival of pathogenic bacteria in organic waste used in biogas plants. *Bioresource Technology*, 87(2):161–166.
- Santos, P. M., Benndorf, D., and Sá-Correia, I. 2004. Insights into *pseudomonas putida* kt2440 response to phenol-induced stress by quantitative proteomics. *Proteomics*, 4(9):2640–2652.
- Sauer, K., Harms, U., and Thauer, R. K. 1997. Methanol: coenzyme m methyltransferase from *methanosarcina barkeri*. *European Journal of Biochemistry*, 243(3):670–677.
- Schägger, H. 2006. Tricine-sds-page. *Nature Protocols*, 1:16–22.
- Schieder, D., Gronauer, A., Lebuhn, M., Bayer, K., Beck, J., Hiepp, G., and Binder, S. 2013. Prozessmodell Biogas.

- Schink, B. 1997. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiology and Molecular Biology Reviews*, 61(2):262–280.
- Schink, B. and Stams, A. J. 2013. *Syntrophism among prokaryotes*. Springer.
- Schlegel, H. G., Eitinger, T., and Fuchs, G. 2007. *Allgemeine Mikrobiologie*. Georg Thieme Verlag.
- Schlegel, T. 2013. Vergleichende Metaproteomanalyse von verschiedenen Belebtschlammproben. Master's thesis, Medizinischen Fakultät Charité – Universitätsmedizin Berlin.
- Schlüter, A., Bekel, T., Diaz, N. N., Dondrup, M., Eichenlaub, R., Gartemann, K.-H., Krahn, I., Krause, L., Krömeke, H., Kruse, O., et al. 2008. The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. *Journal of Biotechnology*, 136(1):77–90.
- Schmack, D. and Reuter, M. 2010. Methanogene Mikroorganismen zur Erzeugung von Biogas.
- Schmack, D. and Reuter, M. 2011. Microorganisms for liquefying biomasses.
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. 2016. The quantitative and condition-dependent escherichia coli proteome. *Nature Biotechnology*, 34(1):104–110.
- Schneider, T., Keiblinger, K. M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., Richter, A., Eberl, L., Zechmeister-Boltenstern, S., and Riedel, K. 2012. Who is who in litter decomposition? metaproteomics reveals major microbial players and their biogeochemical functions. *The ISME Journal*, 6(9):1749–1762.
- Schneider, T., Schmid, E., de Castro, J. V., Cardinale, M., Eberl, L., Grube, M., Berg, G., and Riedel, K. 2011. Structure and function of the symbiosis partners of the lung lichen (*lobaria pulmonaria* l. hoffm.) analyzed by metaproteomics. *Proteomics*, 11(13):2752–2756.
- Schnürer, A. and Nordberg, A. 2008. Ammonia, a selective agent for methane production by syntrophic acetate oxidation at mesophilic temperature. *Water Science & Technology*, 57(5):735–740.
- Schubert, O. T., Gillet, L. C., Collins, B. C., Navarro, P., Rosenberger, G., Wolski, W. E., Lam, H., Amodei, D., Mallick, P., MacLean, B., et al. 2015. Building high-quality assay libraries for targeted analysis of swath ms data. *Nature Protocols*, 10(3):426–441.

- Schweikl, H., Klein, U., Schindlbeck, M., and Wieczorek, H. 1989. A vacuolar-type atpase, partially purified from potassium transporting plasma membranes of tobacco hornworm midgut. *Journal of Biological Chemistry*, 264(19):11136–11142.
- Scigelova, M., Hornshaw, M., Giannakopoulos, A., and Makarov, A. 2011. Fourier transform mass spectrometry. *Molecular and Cellular Proteomics*, 10(7):M111.009431.
- Seifert, J., Herbst, F.-A., Halkjær Nielsen, P., Planes, F. J., Jehmlich, N., Ferrer, M., and Bergen, M. 2013. Bioinformatic progress and applications in metaproteogenomics for bridging the gap between genomic sequences and metabolic functions in microbial communities. *Proteomics*, 13(18-19):2786–2804.
- Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., and Harada, H. 1999. Fluorescence in situ hybridization using 16s rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. *Applied and Environmental Microbiology*, 65(3):1280–1288.
- Senesi, N., Miano, T., Provenzano, M., and Brunetti, G. 1989. Spectroscopic and compositional comparative characterization of ihss reference and standard fulvic and humic acids of various origin. *Science of the Total Environment*, 81:143–156.
- Shade, A. and Handelsman, J. 2012. Beyond the venn diagram: the hunt for a core microbiome. *Environmental Microbiology*, 14(1):4–12.
- Shevchenko, A., Sunyaev, S., Loboda, A., Shevchenko, A., Bork, P., Ens, W., and Standing, K. G. 2001. Charting the proteomes of organisms with unsequenced genomes by maldi-quadrupole time-of-flight mass spectrometry and blast homology searching. *Analytical Chemistry*, 73(9):1917–1926.
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry*, 68(5):850–858.
- Shin, S. G., Han, G., Lim, J., Lee, C., and Hwang, S. 2010. A comprehensive microbial insight into two-stage anaerobic digestion of food waste-recycling wastewater. *Water Research*, 44(17):4838–4849.
- Shoham, Y., Lamed, R., and Bayer, E. A. 1999. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends in Microbiology*, 7(7):275–281.

- Shteynberg, D., Nesvizhskii, A. I., Moritz, R. L., and Deutsch, E. W. 2013. Combining results of multiple search engines in proteomics. *Molecular & Cellular Proteomics*, 12(9):2383–2393.
- Siepen, J. A., Keevil, E.-J., Knight, D., and Hubbard, S. J. 2007. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *Journal of Proteome Research*, 6(1):399–408.
- Siggins, A., Gunnigle, E., and Abram, F. 2012. Exploring mixed microbial community functioning: recent advances in metaproteomics. *FEMS Microbiology Ecology*, 80(2):265–280.
- Silveira, J. A., Ridgeway, M. E., and Park, M. A. 2014. High resolution trapped ion mobility spectrometry of peptides. *Analytical Chemistry*, 86(12):5624–5627.
- Smet, E., Van Langenhove, H., and De Bo, I. 1999. The emission of volatile compounds during the aerobic and the combined anaerobic/aerobic composting of biowaste. *Atmospheric Environment*, 33(8):1295–1303.
- Smith, P., Krohn, R. I., Hermanson, G., Mallia, A., Gartner, F., Provenzano, M., Fujimoto, E., Goeke, N., Olson, B., and Klenk, D. 1985. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, 150(1):76–85.
- Sokal, R. R. 1958. A statistical method for evaluating systematic relationships. *The University of Kansas science bulletin*, 38:1409–1438.
- Stams, A. J. and Plugge, C. M. 2009. Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nature Reviews Microbiology*, 7(8):568–577.
- Stark, L., Giersch, T., and Wünschiers, R. 2014. Efficiency of rna extraction from selected bacteria in the context of biogas production and metatranscriptomics. *Anaerobe*, 29:85–90.
- Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., and Gilles, E. D. 2002. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420(6912):190–193.
- Stiehm, M. 2011. *Identifizierung einer Protease aus Stallstaubextrakten und ihr Einfluss auf die allergische Sensibilisierung und Atemwegsinfammation*. PhD thesis, Ruhr - Universität Bochum.

- Stolze, Y., Bremges, A., Ruming, M., Henke, C., Maus, I., Pühler, A., Sczyrba, A., and Schlüter, A. 2016. Identification and genome reconstruction of abundant distinct taxa in microbiomes from one thermophilic and three mesophilic production-scale biogas plants. *Biotechnology for Biofuels*, 9:156.
- Stolze, Y., Zakrzewski, M., Maus, I., Eikmeyer, F., Jaenicke, S., Rottmann, N., Siebner, C., Pühler, A., and Schlüter, A. 2015. Comparative metagenomics of biogas-producing microbial communities from production-scale biogas plants operating under wet or dry fermentation conditions. *Biotechnology for Biofuels*, 8:14.
- Sturm, M., Bertsch, A., Gröpl, C., Hildebrandt, A., Hussong, R., Lange, E., Pfeifer, N., Schulz-Trieglaff, O., Zerck, A., Reinert, K., et al. 2008. Openms—an open-source software framework for mass spectrometry. *BMC Bioinformatics*, 9(1):163.
- Suárez Quiñones, T., Plöchl, M., Budde, J., and Heiermann, M. 2012. Results of batch anaerobic digestion test—effect of enzyme addition. *Agricultural Engineering International: CIGR Journal*, 14(1):38–50.
- Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R., and Wu, C. H. 2007. Uniref: comprehensive and non-redundant uniprot reference clusters. *Bioinformatics*, 23(10):1282–1288.
- Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26):9528–9533.
- Tabb, D. L., Fernando, C. G., and Chambers, M. C. 2007. Myrimatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *Journal of Proteome Research*, 6(2):654–661.
- Tabb, D. L., Vega-Montoto, L., Rudnick, P. A., Variyath, A. M., Ham, A.-J. L., Bunk, D. M., Kilpatrick, L. E., Billheimer, D. D., Blackman, R. K., Cardasis, H. L., Carr, S. A., Clauser, K. R., Jaffe, J. D., Kowalski, K. A., Neubert, T. A., Regnier, F. E., Schilling, B., Tegeler, T. J., Wang, M., Wang, P., Whiteaker, J. R., Zimmerman, L. J., Fisher, S. J., Gibson, B. W., Kinsinger, C. R., Mesri, M., Rodriguez, H., Stein, S. E., Tempst, P., Paulovich, A. G., Liebler, D. C., and Spiegelman, C. 2010. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *Journal of Proteome Research*, 9(2):761–776.

- Takai, K. and Horikoshi, K. 2000. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Applied and Environmental Microbiology*, 66(11):5066–5072.
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2(8):151–153.
- Tanner, S., Shu, H., Frank, A., Wang, L.-C., Zandi, E., Mumby, M., Pevzner, P. A., and Bafna, V. 2005. Inspect: identification of posttranslationally modified peptides from tandem mass spectra. *Analytical Chemistry*, 77(14):4626–4639.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- Taylor, E. B. and Williams, M. A. 2010. Microbial protein in soil: influence of extraction method and C amendment on extraction and recovery. *Microbial Ecology*, 59(2):390–399.
- Thauer, R. K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriological Reviews*, 41(1):100.
- Thauer, R. K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology*, 6(8):579–591.
- Theuerl, S. and Klocke, M. 2014. Prozessmikrobiologie in landwirtschaftlichen Biogasanlagen Ermittlung der mikrobiellen Diversität sowie von hauptsächlichen verfahrenstechnischen Einflussfaktoren auf die Mikroflora. Bornimer Agrartechnische Berichte, Heft 84.
- Theuerl, S., Kohrs, F., Benndorf, D., Maus, I., Wibberg, D., Schlüter, A., Kausmann, R., Heiermann, M., Rapp, E., Reichl, U., et al. 2015. Community shifts in a well-operating agricultural biogas plant: how process variations are handled by the microbiome. *Applied Microbiology and Biotechnology*, pages 1–13.
- Thiele, H., Glandorf, J., and Hufnagel, P. 2010. Bioinformatics strategies in life sciences: from data processing and data warehousing to biological knowledge extraction. *Journal of Integrated Bioinformatics*, 7(1):141.

- Thiele, H., Glandorf, J., Hufnagel, P., Körting, G., and Blüggel, M. 2008. Managing proteomics data: from generation and data warehousing to central data repository. *Journal of Proteomics Bioinformatics*, 1:485–507.
- Toyoda, A., Iio, W., Mitsumori, M., and Minato, H. 2009. Isolation and identification of cellulose-binding proteins from sheep rumen contents. *Applied and Environmental Microbiology*, 75(6):1667–1673.
- Trinci, A. P., Davies, D. R., Gull, K., Lawrence, M. I., Bonde Nielsen, B., Rickers, A., and Theodorou, M. K. 1994. Anaerobic fungi in herbivorous animals. *Mycological Research*, 98(2):129–152.
- Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., Solovyev, V. V., Rubin, E. M., Rokhsar, D. S., and Banfield, J. F. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, 428(6978):37–43.
- UniProt Consortium 2012. Reorganizing the protein space at the universal protein resource (uniprot). *Nucleic Acids Research*, 40(Database issue):D71–D75.
- Ünlü, M., Morgan, M. E., and Minden, J. S. 1997. Difference gel electrophoresis. a single gel method for detecting changes in protein extracts. *Electrophoresis*, 18(11):2071–2077.
- Vanwonterghem, I., Jensen, P. D., Dennis, P. G., Hugenholtz, P., Rabaey, K., and Tyson, G. W. 2014. Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. *The ISME Journal*.
- Vaudel, M., Barsnes, H., Berven, F. S., Sickmann, A., and Martens, L. 2011. Searchgui: An open-source graphical user interface for simultaneous omssa and x! tandem searches. *Proteomics*, 11(5):996–999.
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., et al. 2004. Environmental genome shotgun sequencing of the sargasso sea. *Science*, 304(5667):66–74.
- Verstraete, W., Wittebolle, L., Heylen, K., Vanparys, B., De Vos, P., Van de Wiele, T., and Boon, N. 2007. Microbial resource management: the road to go for environmental biotechnology. *Engineering in Life Sciences*, 7(2):117–126.
- Vizcaíno, J. A., Côté, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., et al. 2013. The proteomics identifications (pride)

- database and associated tools: status in 2013. *Nucleic Acids Research*, 41(D1):D1063–D1069.
- Vogels, G. v. d. and Van der Drift, C. 1976. Degradation of purines and pyrimidines by microorganisms. *Bacteriological Reviews*, 40(2):403.
- von Bergen, M., Jehmlich, N., Taubert, M., Vogt, C., Bastida, F., Herbst, F.-A., Schmidt, F., Richnow, H.-H., and Seifert, J. 2013. Insights from quantitative metaproteomics and protein-stable isotope probing into microbial ecology. *The ISME Journal*, 7(10):1877–1885.
- Ward, A. J., Hobbs, P. J., Holliman, P. J., and Jones, D. L. 2008. Optimisation of the anaerobic digestion of agricultural resources. *Bioresource Technology*, 99(17):7928–7940.
- Washburn, M. P., Wolters, D., and Yates, J. R. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology*, 19(3):242–247.
- Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., and Humphery-Smith, I. 1995. Progress with gene-product mapping of the mollicutes: *Mycoplasma genitalium*. *Electrophoresis*, 16(7):1090–1094.
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191.
- Weiland, P. 2010. Biogas production: current state and perspectives. *Applied Microbiology and Biotechnology*, 85(4):849–860.
- Weiss, A., Jérôme, V., Freitag, R., and Mayer, H. K. 2008. Diversity of the resident microbiota in a thermophilic municipal biogas plant. *Applied Microbiology and Biotechnology*, 81(1):163–173.
- Welker, M. and Moore, E. R. B. 2011. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Systematic and Applied Microbiology*, 34(1):2–11.
- Werner, J. J., Knights, D., Garcia, M. L., Scalfone, N. B., Smith, S., Yarasheski, K., Cummings, T. A., Beers, A. R., Knight, R., and Angenent, L. T. 2011. Bacterial community structures are unique and resilient in full-scale bioenergy systems. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10):4158–4163.

- Westerholm, M. 2012. *Biogas production through the syntrophic acetate-oxidising pathway*. PhD thesis, Faculty of Natural Resources and Agricultural Sciences, Department of Microbiology, Uppsala.
- Westerholm, M., Moestedt, J., and Schnürer, A. 2016. Biogas production through syntrophic acetate oxidation and deliberate operating strategies for improved digester performance. *Applied Energy*, 179:124–135.
- Wöhlbrand, L., Trautwein, K., and Rabus, R. 2013. Proteomic tools for environmental microbiology—a roadmap from sample preparation to protein identification and quantification. *Proteomics*, 13(18-19):2700–2730.
- Wilkins, M. R., Appel, R. D., Van Eyk, J. E., Chung, M. C. M., Görg, A., Hecker, M., Huber, L. A., Langen, H., Link, A. J., Paik, Y.-K., Patterson, S. D., Pennington, S. R., Rabilloud, T., Simpson, R. J., Weiss, W., and Dunn, M. J. 2006. Guidelines for the next 10 years of proteomics. *Proteomics*, 6(1):4–8.
- Wilmes, P. and Bond, P. L. 2006. Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends in Microbiology*, 14(2):92–97.
- Wirth, R., Kovács, E., Maróti, G., Bagi, Z., Rákhely, G., Kovács, K. L., et al. 2012. Characterization of a biogas-producing microbial community by short-read next generation dna sequencing. *Biotechnology for Biofuels*, 5(4).
- Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. 2009. Universal sample preparation method for proteome analysis. *Nature Methods*, 6(5):359–362.
- Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W., and Boon, N. 2009. Initial community evenness favours functionality under selective stress. *Nature*, 458(7238):623–626.
- Wolters, D. A., Washburn, M. P., and Yates, J. R. 2001. An automated multidimensional protein identification technology for shotgun proteomics. *Analytical chemistry*, 73(23):5683–5690.
- Yang, Y., Zhang, S., Howe, K., Wilson, D. B., Moser, F., Irwin, D., and Thannhauser, T. W. 2007. A comparison of nlc-esi-ms/ms and nlc-maldi-ms/ms for gelc-based protein identification and itraq-based shotgun quantitative proteomics. *Journal of Biomolecular Techniques*, 18(4):226.

- Zakrzewski, M., Goesmann, A., Jaenicke, S., Jünemann, S., Eikmeyer, F., Szczepanowski, R., Al-Soud, W. A., Sørensen, S., Pühler, A., and Schlüter, A. 2012. Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing. *Journal of Biotechnology*, 158(4):248–258.
- Ziganshin, A. M., Liebetrau, J., Pröter, J., and Kleinsteuber, S. 2013. Microbial community structure and dynamics during anaerobic digestion of various agricultural waste materials. *Applied Microbiology and Biotechnology*, 97(11):5161–5174.
- Zybilov, B. L., Florens, L., and Washburn, M. P. 2007. Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Molecular BioSystems*, 3(5):354–360.

Tabellenverzeichnis

| | | |
|-----|---|------|
| 2.1 | Empfohlene Prozessparameter für den Biogasprozess | 6 |
| 2.2 | Substratzusätze für Biogasgasanlagen bei bestimmten Prozessproblemen . . | 10 |
| 4.1 | Reproduzierbarkeit der SDS-PAGE LC-MS/MS Methode | 60 |
| 4.2 | Anzahl an identifizierten Spektren für Datenbanken verschiedener Größe . | 64 |
| 4.3 | Einfluss ausgewählter PTMs auf die Anzahl an identifizierten Peptiden . . | 67 |
| 4.4 | Biochemischer Unterschied ausgewählter Aminosäuren | 68 |
| 5.1 | Identifizierte Metaproteine für die wesentlichen Abbauwege des Biogaspro- zesses | 94 |
| 5.2 | Effekt von Prozessparametern auf den Kohlenstoffmetabolismus | 116 |
| 6.1 | Entwicklung der Metaproteomanalysen | 124 |
| 7.1 | Hemmstoffe in Biogasprozessen | i |
| 7.2 | Generationszeiten methanogener <i>Archaeen</i> | ii |
| 7.3 | Chemikalienliste | vi |
| 7.4 | Geräteliste | xi |
| 7.5 | Auflistung untersuchter Biogasanlagen | xii |
| 7.6 | Auflistung untersuchter Biogasanlagen Gent | xv |
| 7.7 | Komponenten und Parameter für das Elite Hybrid Ion Trap-Orbitrap MS/MS | xvii |
| 7.8 | Komponenten und Parameter für das AmaZon ETD ESI-Ion Trap MS/MS | xx |

Abbildungsverzeichnis

| | | |
|------|--|----|
| 2.1 | Allgemeiner Aufbau einer Biogasanlage | 5 |
| 2.2 | Vier Stufen des Biogasprozesses | 12 |
| 2.3 | Mikrobielle Gleichverteilung (<i>Evenness/Community Organisation</i>) | 20 |
| 2.4 | Allgemeiner Workflow für Metaproteomanalysen | 22 |
| 2.5 | Schematischer Aufbau eines Massenspektrometers | 28 |
| 2.6 | Charakteristische Fragmentationen von Peptiden | 30 |
| 4.1 | SDS-PAGE mit verschiedene Probenmengen | 47 |
| 4.2 | SDS-PAGE Reproduzierbarkeit | 49 |
| 4.3 | Streuung der Proteinidentifikationen über die Anzahl an Fraktionen der SDS-PAGE | 50 |
| 4.4 | Identifizierte Proteine und Peptide einer mesophilen und einer thermophi- len Biogasanlage mit verschiedener Anzahl an Trennstufen. | 52 |
| 4.5 | Einfluss der Trypsinmenge auf die Anzahl an identifizierten Spektren und <i>missed cleavages</i> | 56 |
| 4.6 | Identifizierte Spektren bei verschiedener Probenbeladung und Gradienten- länge | 57 |
| 4.7 | Reproduzierbarkeit von LC-MS/MS Messungen | 59 |
| 4.8 | Einfluss verschiedener Schwellwerte auf die Anzahl an identifizierten Spektren | 65 |
| 4.9 | Einfluss verschiedener Datenbanksuchalgorithmen auf die Anzahl an iden- tifizierten Peptiden | 66 |
| 4.10 | <i>De novo</i> Sequenzierung mittels DeNovoGUI | 69 |
| 4.11 | Bioinformatischer Datenauswerteworkflow mit dem MetaProteomeAnalyzer | 75 |
| 4.12 | Workflow zu Erzeugung von Metaproteinen | 76 |
| 4.13 | Verschiedene Metaproteinstrategien | 80 |
| 5.1 | SDS-PAGE Profile von sechs verschiedenen Biogasanlagen | 85 |
| 5.2 | Prozessdaten Biogasanlage BGP 3 | 86 |
| 5.3 | Profil der Taxonomien und Funktionen basierend auf allen identifizierten Proteinen | 88 |
| 5.4 | Taxonomie- und Funktionsprofile einer mesophilen und einer thermophilen Biogasanlage | 92 |
| 5.5 | SDS-PAGE Profile von 40 Biogasanlagen | 98 |

| | | |
|------|--|--------|
| 5.6 | Schwellwerte für Metaproteine, taxonomischen Ordnungen und biologischen Prozesse | 100 |
| 5.7 | Clustergram von mikrobiellen Gemeinschaften der 40 Biogasanlagen | 102 |
| 5.8 | Hauptkomponentenanalyse der mikrobiellen Gemeinschaften von 40 Biogasanlagen | 103 |
| 5.9 | Richness und Evenness der untersuchten Biogasanlagen | 105 |
| 5.10 | Taxonomieprofile aller Biogasanlagen auf Ordnungsebene | 107 |
| 5.11 | Zusammenhänge zwischen mikrobiellen Ordnungen und den Prozessparametern | 110 |
| 5.12 | Häufigste biologische Prozesse und Metaproteine in Biogasanlagen | 112 |
| 5.13 | Zuordnung aller identifizierter Metaproteine zur KEGG Karte 1200 (Kohlenstoffstoffwechsel) | 113 |
| 5.14 | Korrelation Gesamtstickstoff zu pH-Wert | 115 |
| 7.1 | Gesamtübersicht von Biogasprozessen | iv |
| 7.2 | Säulenkonfiguration und Gradient des Elite Hybrid Ion Trap-Orbitrap MS/MS | xix |
| 7.3 | Säulenkonfiguration und Gradient des AmaZon ETD ESI-Ion Trap MS/MS | xxii |
| 7.4 | Multiples Sequenzalignment für das Protein EFTU | xxiii |
| 7.5 | Multiples Sequenzalignment für das Protein Methyl Coenzym M Reduktase Untereinheit β | xxiv |
| 7.6 | Methanogenesestoffwechsel für eine mesophile und eine thermophile Biogasanlage | xxv |
| 7.7 | Abhängigkeit der Richness und Evenness vom Schwellwert | xxvi |
| 7.8 | Zusammenhänge der Prozessparameter der Biogasanlagen | xxvii |
| 7.9 | Einfluss von hoher Prozesstemperatur auf den Kohlenstoffmetabolismus . . | xxviii |
| 7.10 | Einfluss von niedriger Prozesstemperatur auf den Kohlenstoffmetabolismus | xxix |
| 7.11 | Einfluss von hoher Gesamtammoniumstickstoffkonzentration auf den Kohlenstoffmetabolismus | xxx |
| 7.12 | Einfluss von niedriger Gesamtammoniumstickstoffkonzentration auf den Kohlenstoffmetabolismus | xxxi |
| 7.13 | Einfluss von hoher Raumbelastung auf den Kohlenstoffmetabolismus | xxxii |
| 7.14 | Einfluss von niedriger Raumbelastung auf den Kohlenstoffmetabolismus . . | xxxiii |
| 7.15 | Einfluss von langer Schlammverweilzeit auf den Kohlenstoffmetabolismus . | xxxiv |
| 7.16 | Einfluss von kurzer Schlammverweilzeit auf den Kohlenstoffmetabolismus . | xxxv |

Publikationsübersicht

Im Rahmen der vorliegenden Dissertation sind die nachfolgenden Artikel, Poster und Vorträge entstanden, sowie mehrere studentische Arbeiten betreut wurden. Einzelne Abbildungen und Ergebnisse der vorliegenden Arbeit sind aus diesen entnommen.

Artikel

- i. Hanreich, A., *Heyer, R.*, Benndorf, D., Rapp, E., Pioch, M., Reichl, U., Klocke, M. (2012). Metaproteom analysis to determine the metabolically active part of a thermophilic microbial community producing biogas from agricultural biomass. *Canadian Journal of Microbiology* 58, 917-922. DOI:10.1139/W2012-058.
- ii. Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., **Heyer, R.**, Rapp E., Pühler A., Reichl U., Klocke M. (2013). Metagenome and metaproteome analyses of microbial communities in mesophilic biogas producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation. *Systematic and Applied Microbiology*, 36(5), 330-338. DOI:10.1016/j.syapm.2013.03.006
- iii. **Heyer, R.**, Kohrs, F. (shared first), Benndorf, D., Rapp, E., Kausmann, R. , Heiermann, M., Klocke, M., Reichl, U. (2013). Metaproteome analysis of the microbial communities in agricultural biogas plants. *New Biotechnology* 30, 614-622. DOI:10.1016/j.nbt.2013.01.002
- iv. Kohrs, F., **Heyer, R.** (shared first), Magnussen, A., Benndorf, D., Muth, T., Behne, A., Rapp, E., Kausmann, R., Heiermann, M., Klocke, M., Reichl, U. (2014). Sample prefractionation with liquid isoelectric focusing enables in depth microbial metaproteome analysis of mesophilic and thermophilic biogas plants. *Anaerobe*, 29, 59-67, DOI:10.1016/j.anaerobe.2013.11.009.
- v. Bensmann, A., Hanke-Rauschenbach, R., **Heyer, R.**, Kohrs, F., Benndorf, D., Reichl, U., Sundmacher, K., (2014) Biological methanation of hydrogen within biogas plants - a model based feasibility study. *Applied Energy* 134 (2014) 413–425
- vi. Muth, T., Behne, A., **Heyer, R.**, Kohrs, F., Benndorf, D., Hoffmann, M., Lehtevä, M., Reichl, U., Martens, L., Rapp, E. (2015). The MetaProteomeAnalyzer: a powerful open-source software suite for metaproteomics data analysis and interpretation. *Journal of Proteome Research*, 14, 1557-1565. doi: 10.1002/pmic.201400559
- vii. **Heyer, R.**, Kohrs, F., Reichl, U., Benndorf, D. (2015). Metaproteomics of complex microbial communities in biogas plants. *Microbial Biotechnology*, 8, 749–763.
- viii. Püttker, S., Kohrs, F., Benndorf, D., **Heyer, R.**, Rapp, E., Reichl, U. (2015). Metaproteomics of activated sludge from a wastewater treatment plant - A pilot study. *Proteomics*, 5(20):3596-601. doi: 10.1002/pmic.201400559.

- ix. Kohrs, F., Wolter, S., Benndorf, D., **Heyer, R.**, Hoffmann, M., Rapp, E., Bremges, A., Sczyrba, A., Schlüter, A., Reichl, U. (2015). Fractionation of biogas plant sludge material improves metaproteomic characterization to investigate metabolic activity of microbial communities. *Proteomics*, 15(20):3585-9. doi: 10.1002/pmic.201400557
- x. Bensmann A., Hanke-Rauschenbach, R., **Heyer, R.**, Kohrs, F., Benndorf, D., Kausmann R., Plöchl M., Heiermann, M., Reichl, Sundmacher, K. (2016) Diagnostic concept for dynamically operated biogas production plants, *Renewable Energy*, 96, 479:489. DOI:10.1016/j.apenergy.2014.08.047
- xi. **Heyer, R.**, Benndorf, D., Kohrs, F., De Vrieze, J., Boon, N., Hoffmann, M., Rapp, E., Schlüter, A., Sczyrba, A., Reichl, U., (2016) Proteotyping of biogas plant microbiomes separates biogas plants according to process temperature and reactor type. *Biotechnology for Biofuels* ,9:155. doi: 10.1186/s13068-016-0572-4.

Poster

Heyer, R., Muth, T., Kohrs, F., Behne, A., Benndorf, D., Martens, L., Klocke, M., Rapp, E., Reichl, U. (2013). Biogas plant metaproteome analysis and biomarker search by using the software MetaProteomeAnalyser. Poster presented at 5th Congress of European Microbiologists (FEMS 2013), Leipzig, Germany.

Heyer, R., Muth, T., Behne, A., Kohrs, F., Benndorf, D., Kausmann, R., Heiermann, M., Klocke, M., Rapp, E., Reichl, U. (2013). MetaProteomeAnalyser: A software tool for functional and taxonomic characterization of metaproteome data. Poster presented at Annual Conference of the Association for General and Applied Microbiology (VAAM), Bremen, Germany.

Heyer, R., Kohrs, F., Benndorf D., De Vrieze J., Rapp E., Boon N., Reichl U., (2015). A metaproteomic approach for analysis of microbial community structure and function for improvement of biogas plant performance. Poster presented at 6th Congress of European Microbiologists (FEMS 2015), Maastricht, Netherlands.

Heyer, R., Kohrs, F., Benndorf, D., DeVrieze, J., Rapp, E., Schlüter, A., Sczyrba, A., Boon, N., Reichl, U. (2016), A metaproteomic approach for analysis of microbial community structure and function for improvement of biogas plant performance. Poster presented at Symposium on Advances and Applications in Metaproteomics , Magdeburg, Germany

Vorträge

Heyer, R., Muth, T., Behne, A., Kohrs, F., Benndorf, D., M., Rapp, E., Reichl, U. (2012) MetaProteomeAnalyser: A software tool specifically developed for the functional and taxonomic characterization of metaproteome data, 14. EMBL PhD Symposium, Heidelberg, Germany

- Heyer, R.**, Benndorf, D., Kohrs, F., Hanreich, A., Klocke, M., Rapp, E., Reichl, U. (2012) Investigation of the microbial community in biogas plants by metaproteome analysis, EMB2012 Bologna, Italy
- Heyer, R.**, Kohrs, F., Benndorf, D., Rapp, E., Reichl, U., (2014) Microbial population structure and metabolic function of biogas plants based on metaproteome data, BiogasScience2014, Wien, Österreich
- Heyer, R.**, (2011) Prozesskontrolle und Optimierung der Biogasproduktion mittels Metaproteomanalyse, DBU-Stipendiatenseminar, Bad Bederkesa, Deutschland
- Heyer, R.**, (2013) Prozesskontrolle und Optimierung der Biogasproduktion mittels Metaproteomanalyse, DBU-Stipendiatenseminar, Benediktbeuern, Deutschland
- Heyer, R.**, (2014) Prozesskontrolle und Optimierung der Biogasproduktion mittels Metaproteomanalyse, DBU-Stipendiatenseminar, Binz, Deutschland
- Heyer, R.**, (2014) Energie aus Abfällen - Wie uns Mikroorganismen helfen unsere Abfälle zu beseitigen, PerspektivForum „Green Jobs“, Osnabrück, Deutschland
- Heyer, R., Muth, T.**, (2016) Workshop about the MetaProteomeAnalyzer Software, Symposium on Advances and Applications in Metaproteomics , Magdeburg, Germany
- Heyer, R.**, Benndorf, D., Kohrs, F., De Vrieze, J., Boon, N., Hoffmann, M., Rapp, E., Schlüter, A., Sczyrba, A., Reichl, U., (2016) Proteotyping of biogas plant microbiomes, Annual Conference of the Association for General and Applied Microbiology (VAAM), Jena, Deutschland

**Betreute
Arbeiten**

- Rüdiger, J., (2013) Entwicklung eines proteinbasierten Schnelltests zur Charakterisierung von mikrobiellen Gemeinschaften in Umweltproben, Masterarbeit
- Behne, A., (2013) A performance evaluation of spectral library searching in metaproteomics, Studienarbeit
- To, D., (2014) Analyzing the microbial community in a lab-scale biogas reactor, Masterarbeit
- Düsedau, R., (2015) Antikörperbasierter Nachweis von Proteinen in Biogasanlagenproben mittels Western Blot, Bachelorarbeit

Anhang

7.1 Hemmstoffe

Tabelle 7.1: *Hemmstoffe in Biogasprozessen.*

Grenzwerte für bekannte Hemmstoff von Biogasprozessen [Jørgensen (2009), Seite 5]. Bei den dargestellten Werten handelt es sich um allgemeine Richtwerte. Für einzelne Mikroorganismen können diese Werte abweichen. Auch ist es möglich, dass sich nach entsprechenden Anpassungszeiten, die tatsächlichen inhibitorischen und toxischen Konzentrationen erhöhen.

| Substanz | Inhibition | Toxisch |
|--|-----------------------------|----------------------------|
| Ammoniak (NH_3) | 50-100 mg N/L | 100-200 mg N/L |
| Ammonium komplett (NH_4^+) | 1.000-6.000 mg N/L | 10.000 mg N/L (pH<7,5) |
| Chloride (Cl^-) | <8.000 mg/L | 10.000 mg/L |
| Cyanide (CN^-) | 2-20 mg/L | 30 mg/L |
| Formaldehyde (H_2CO) | 100-400 mg/L | 500-1.000 mg/L |
| Phenol ($\text{C}_6\text{H}_5\text{OH}$) | 100-200 mg/L | |
| Chloroform (CHCl_3) | >1 mg/L (einmalig) | > 50 mg/L (kontinuierlich) |
| Wasserstoff (H_2) | $p(\text{H}_2) 10^{-4}$ atm | |
| Kupfer (Cu^{3+}) | 10-250 mg/L | |
| Chrom (Cr^{3+}) | 50-100 mg/L | 200-400 mg/L |
| Nickel (Ni^{2+}) | 100-200 mg/L | 300-1.000 mg/L |
| Natrium (Na^+) | 3.000-10.000 mg/L | |
| Calcium (Ca^{2+}) | 8.000 mg/L | |
| Magnesium (Mg^{2+}) | 3.000 mg/L | |
| Zink (Zn^+) | 350-1.000 mg/L | |
| Sulphate (SO_4^{2-}) | 500-4.000 mg/L | |
| Schwefel (S) | 200 mg/L | |
| Schwefelwasserstoff (H_2S) | 250-1.000 mg/L | |

7.2 Generationszeiten methanogener Archaeen

| Substanz | Inhibition | Toxisch |
|----------------------|--|---------|
| Flüchtige Fettsäuren | 2.000 mg/l Essigsäureäquivalent (pH = 7,0) | |

7.2 Generationszeiten methanogener Archaeen

Tabelle 7.2: *Generationszeiten methanogener Archaeen.*

Diese Tabelle zeigt die Generationszeiten methanogener Archaeen basierend auf [Fardeau et al. (1987), Koch (2014), Rajoka et al. (1999), Krzycki et al. (1982), De Vrieze et al. (2012)].

| Mikroorganismus | Generationszeiten in [d] |
|---|---|
| <i>Methanobacterium thermoautotrophicus</i> | 0,25- 0,5 (CO ₂ + H ₂) |
| <i>Methanosarcina barkeri</i> | 10-70 (Acetat) |
| <i>Methanosarcina</i> | ca 1,66 |
| <i>Methanosaeta</i> | ca. 5 |

7.3 Biochemie von Biogasprozessen

In diesem Abschnitt sind die wichtigsten biochemischen Reaktionsgleichung von Biogasprozessen aufgeführt [Schlegel et al. (2007), Seite 354, Madigan and Martinko (2009) Kapitel 14] und ihr Zusammenhang in Form einer Abbildung dargestellt [Abbildung: 7.1] [Kanehisa and Goto (2000), Stelling et al. (2002), Hedderich and Whitman (2006) Seite 1065, Schlegel et al. (2007), Seite 354, Thauer et al. (2008), Madigan and Martinko (2009), Ferry (2011), Caspi et al. (2012), Fronk (2013), Koch (2014)]. Da der Biogasprozess sehr umfangreich ist, ist diese Abbildung nur in sehr starker Verkleinerung dargestellt. Die Abbildung in Originalgröße (A2) kann der beigefügten CD entnommen werden [CD: 7.14].

Allgemeiner Biogasprozess



Gärungen

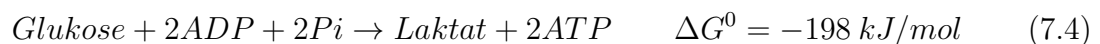
Acetatgärung



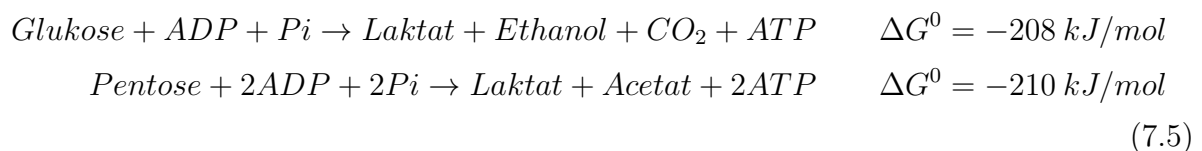
Ethanolgärung



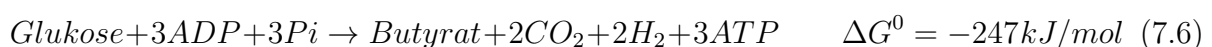
Homofermentative Laktatgärung



Heterofermentative Lactatgärung



Butyratgärung



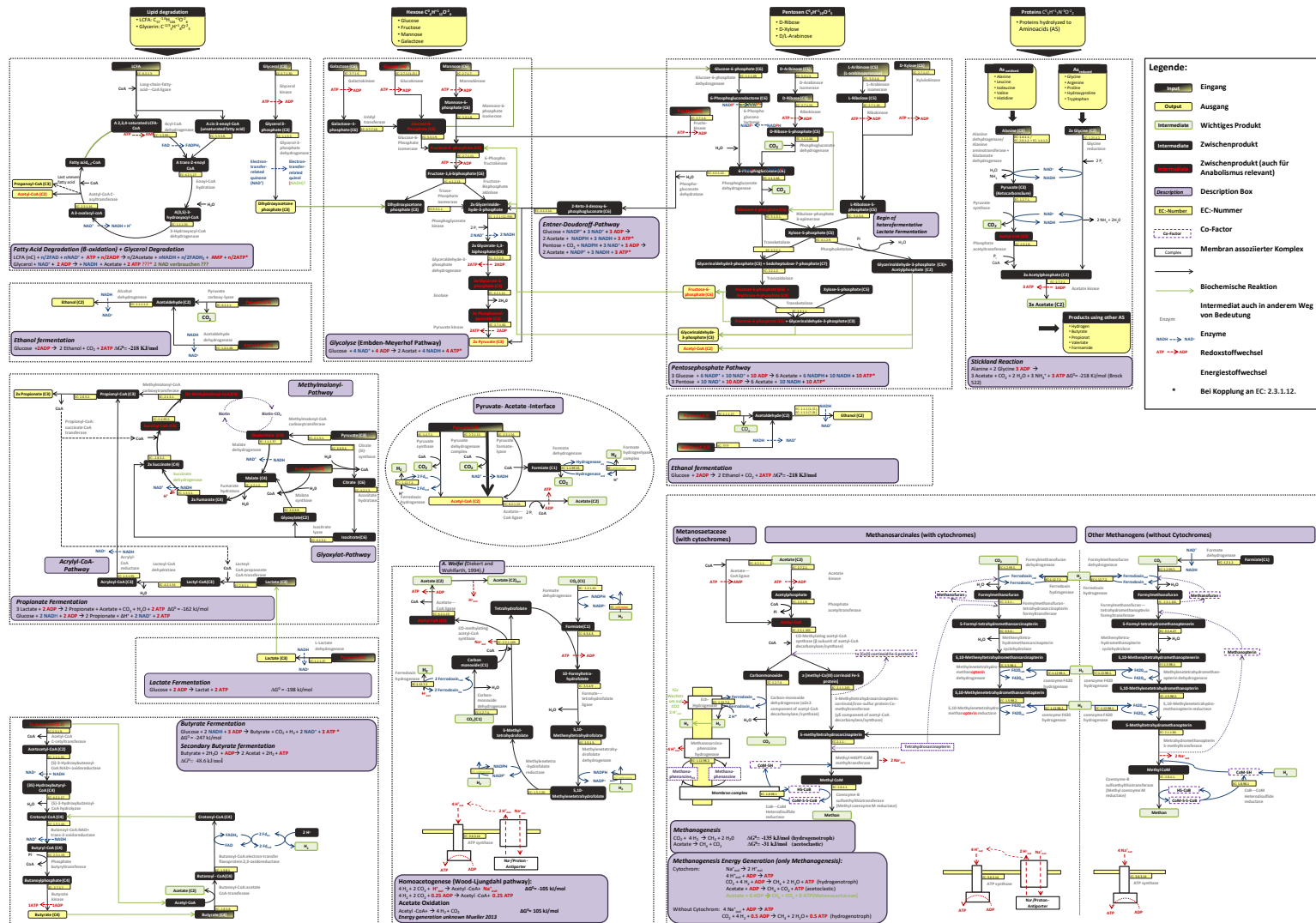
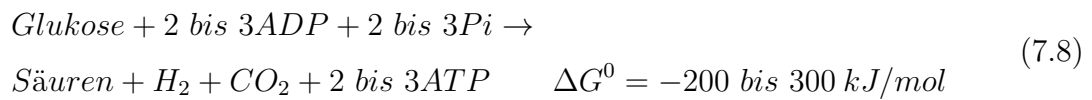


Abbildung 7.1: Gesamtübersicht von Biogasprozessen.

Propionatgärung



Gemischte Säuregärung



Acetogenese

Homoacetogenese



Acetatoxidation

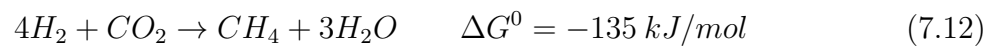


Methanogenese

Acetoclastische Methanogenese



Hydrogenotrophe Methanogenese



7.4 Chemikalienliste

Die für diese Arbeit verwendeten Chemikalien sind in der folgenden Tabelle aufgeführt. Die entsprechenden Lösungen wurden alle mit Reinstwasser (Millipore Q-POD) hergestellt. Für einzelne Chemikalien hatte sich im Laufe der Promotion aus kostengründen der Anbieter, aber nicht die Qualität geändert.

Tabelle 7.3: *Chemikalienliste*

| Chemikalie/ Lösung | Beschreibung/ Hersteller | Bestellnummer |
|---|--|---------------|
| 25 mM ABC-Puffer | 99 mg Ammoniumbicarbonat 50 mL Reinstwasser | |
| Aceton ($\geq 99\%$) | VWR International, Darmstadt, Deutschland | 20063.296 |
| Acetonitril (LC-MS) | Sigma-Alderich, St. Luis, USA | 34967 |
| Acrylamidlösung (30%) | Serva Heidelberg, Deutschland | 10688.01 |
| Ameisensäure (MS) | Sigma-Alderich, St. Luis, USA | 56302 |
| Amidoschwarz | Sigma-Alderich, St. Luis, USA | 195243 |
| Ammoniumacetat ($M = 77,08$ g/mol) | Carl Roth, Karlsruhe, Deutschland | Art. 7869.1 |
| 0,1 M Ammoniumacetat | 0,7708 g Ammoniumacetat Auffüllen auf 100 mL mit Methanol | |
| Ammoniumbicarbonat ($\geq 99,5\%$) | Sigma-Alderich, St. Luis, USA | 9830 |
| Ammoniumpersulfat | Merck, Darmstadt, Deutschland | 1.01201.0500 |
| Ammoniumsulfat | Carl Roth, Karlsruhe, Deutschland | 3746.3 |
| Ampholyt | GE Healthcare, Pittsburgh, USA | 17.0456-01 |
| APS-Lösung (0,1 g/mL) | 10 mL Reinstwasser 1 g APS | |
| Bovines Serum Albumin | Sigma-Alderich, St. Luis, USA | A3912 |
| Bromphenolblau | GE Healthcare, Pittsburgh, USA | 17-1329-01 |
| Bromphenolblaulösung (5 mg/ml) | 100 mL Reinstwasser 0,5 g Bromphenolblau | |
| Cellophanfolie (24 cm x 24 cm) | Carl Roth, Karlsruhe, Deutschland | K423.1 |
| Coomassie BB G 250 | AppliChem, Darmstadt, Deutschland | A3484.0100 |
| Coomassie Brilliant Blau | 5 g Coomassie Brilliant Blau G-250 | |

7.4 Chemikalienliste

| Chemikalie/ Lösung | Beschreibung/ Hersteller | Bestellnummer |
|--------------------------------|--|---------------|
| Stock Lösung | 100 mL Reinstwasser | |
| Cover Fluid | GE Healthcare, Pittsburgh, USA | 17-1335-01 |
| 1,4-Dithiothreitol | Carl Roth, Karlsruhe, Deutschland | 6908.2 |
| 45 mM 1,4-Dithiothreitollösung | 104 mg 1,4-Dithiothreitol 15 mL ABC-Puffer | |
| Essigsäure ($\geq 96\%$) | Carl Roth, Karlsruhe, Deutschland | T179.2 |
| Essigsäure (LC-MS) | Sigma-Aldrich, St. Luis, USA | 49199 |
| Ethanol (70 %) | Carl Roth, Karlsruhe, Deutschland | T913.3 |
| Extraktionspuffer I | 90% Reinstwasser 10% Ameisensäure | |
| Extraktionspuffer II | 50% Reinstwasser 49% Acetonitril 1% Trifluoressigsäure | |
| Färbelösung | 26 mg Amidoschwarz 100 mL Waschlösung | |
| Fixierungslösung | 329 mL Reinstwasser 571 mL Ethanol (70%) 100 mL Essigsäure (100%) | |
| Glycerol | Carl Roth, Karlsruhe, Deutschland | 3783.1 |
| Glycerollösung | 200 mL Glycerol 857 mL Ethanol (70%) Auffüllen auf 2 L mit Reinstwasser | |
| Glycin | VWR International, Darmstadt, Deutschland | 444495D |
| Harnstoff | AppliChem, Darmstadt, Deutschland | A1049.0500 |
| Harnstoffpuffer | 8,41 g Harnstoff 3,04 g Thioharnstoff 0,2 g 1,4-Dithiothreitol Auffüllen auf 20 mL mit Reinstwasser | |
| IEF-Puffer | DeStreak Rehydration Solution (3 mL) 15 μ L Ampholyt (Endkonzentration 0,5%) | |
| Iodacetamid | AppliChem, Darmstadt, Deutschland | A1666,0025 |
| 100 mM Iodacetamidlösung | 277 mg Iodacetamid 15 mL ABC-Puffer | |

7.4 Chemikalienliste

| Chemikalie/ Lösung | Beschreibung/ Hersteller | Bestellnummer |
|---|---|---------------|
| IPG Streifen pH 4-7, 13 cm | Serva Heidelberg, Deutschland | 943053.01 |
| IPG Streifen pH 4-7, 7 cm | GE Healthcare, Pittsburgh, USA | 17-6001-10 |
| Isobutanol | AppliChem, Darmstadt, Deutschland | A1150.1000 |
| Kolloidale Coomassie Stock Lösung | 50 g Ammoniumsulfat 484 mL Reinstwasser 6 mL Ortho-Phosphorsäure (85%) 10 mL Coomassie Brilliant Blau Stock Lösung | |
| Kolloidale Coomassie Färbelösung | 200 mL Kolloidale Coomassie Stock Lösung 50 mL Methanol | |
| Mercaptoethanol | Sigma-Aldrich, St. Luis, USA | M3148 |
| Methanol ($\geq 99,9\%$) | VWR International, Darmstadt, Deutschland | 20063.296 |
| Methanol (LC-MS) | Sigma-Aldrich, St. Luis, USA | 34966 |
| Natriumchlorid | Merck, Darmstadt, Deutschland | 1.06404.1000 |
| Natriumdodecylsulfat (SDS) | AppliChem, Darmstadt, Deutschland | A2572.0250 |
| Natriumhydroxid ($\geq 98\%$) | Carl Roth, Karlsruhe, Deutschland | P031.1 |
| 0,1 M Natriumhydroxid | 0,4 g Natriumhydroxid einwiegen Auffüllen auf 100 mL mit Reinstwasser | |
| OFFGEL Protein Sample | 1,44 mL OFFGEL Stock Solution (1,25X) 0,36 mL Probe | |
| 3100 OFFGEL Fractionator Starter Kit | Agilent Technologies, Santa Clara, USA | 5188-6444. |
| Phenol (M= 94,11 g/mol) | Merck, Darmstadt, Deutschland | 8.22296.1000, |
| Phenollösung | 10 g Phenol 1 mL Reinstwasser | |
| Phosphorsäure (85%) | Carl Roth, Karlsruhe, Deutschland | 6366.2 |
| Proteinstandard | Thermo Scientific, Waltham, USA | 26619 |
| Protein OFFGEL Stock Lösung (1,25X) | Inhalt der Flasche Thiourea (ausspülen mit 5 mL Reinstwasser) Inhalt der Flasche DTT (ausspülen mit 5 mL Reinstwasser) | |

7.4 Chemikalienliste

| Chemikalie/ Lösung | Beschreibung/ Hersteller | Bestellnummer |
|---|--|---------------|
| | 6 mL Glycerol Lösung 0,6 mL OFFGEL Puffer Auffüllen auf 50 mL mit Reinstwasser | |
| Protein IPG Strip Rehydration Lösung | 0,56 mL Protein OFFGEL Stock Lösung (1,25X) 0,14 mL Reinstwasser | |
| Saccharose (M = 342.29 g/mol) | Merck, Darmstadt, Deutschland | 1.07687.1000 |
| 2 M Saccharoselösung | 68,46 g Sucrose einwiegen Auffüllen auf 100 mL mit Reinstwasser | |
| 1 M Saccharoselösung | 34,23 g Sucrose einwiegen Auffüllen auf 100 mL mit Reinstwasser | |
| Salzsäure (37%) | VWR International, Darmstadt, Deutschland | 20252.42 |
| Sammelgel | 6,1 mL Reinstwasser 2,5 mL 0,5 M Tris-HCl (pH 6,8) 0,1 mL SDS 1,3 mL Polyacrylamid 0,075 mL APS 0,015 mL TEMED | |
| SDS-Laufpuffer (5x) | 15 g Tris 72 g Glycin 5 g SDS Auffüllen auf 1 L mit Reinstwasser | |
| SDS-Lösung (0,1 g/mL) | 1 g SDS Auffüllen auf 10 mL mit Reinstwasser | |
| SDS-Probenpuffer | 50 mL Reinstwasser 12,5 mL 0,5 M Tris-HCl (pH 6,8) 10 mL Glycerol 20 mL SDS-Lösung (0,1 g/mL) 5 mL Mercaptoethanol 1 mL Bromphenolblau-Lösung | |
| Silikaperlen (0,5 mm) | BioSpec Products Inc, Bartlesville, USA | 1179105z |
| Tetramethylethylendiamin | Carl Roth, Karlsruhe, Deutschland | 2367.1 |
| Thioharnstoff | Sigma-Alderich, St. Luis, USA | T8656 |

7.4 Chemikalienliste

| Chemikalie/ Lösung | Beschreibung/ Hersteller | Bestellnummer |
|---|---|---------------|
| Trenngel | 6,7 mL Reinstwasser 10 mL 1,5 M Tris-HCl (pH 8,8) 0,2 mL SDS 8 mL Polyacrylamid 0,15 mL APS 0,015 mL TEMED | |
| Triflouressigsäure | Sigma-Aldrich, St. Luis, USA | 73645 |
| Triflourethanol (MS) | Sigma-Aldrich, St. Luis, USA | 40967 |
| Tris | AppliChem, Darmstadt, Deutschland | A2264.1000 |
| 1,5 M Tris-HCl (pH 8,8) | 100 mL Reinstwasser 90,75 g Tris pH 8,8 mit 4 M HCl einstellen Auffüllen auf 500 mL mit Reinstwasser | |
| 0,5 M Tris-HCl (pH 6,8) | 100 mL Reinstwasser 15 g Tris pH 6,8 mit 1 M HCl einstellen Auffüllen auf 250 mL mit Reinstwasser | |
| Trypsin (Premium grade) | Serva Heidelberg, Deutschland | 111095 |
| Trypsin-Stammlösung (1 $\mu\text{g}/\mu\text{L}$) | 100 μg lyophilisiertes Trypsin Aufnehmen in 100 μL 50 mM Essigsäure | |
| Waschlösung (Amidoschwarz) | 180 mL Methanol 20 mL Essigsäure (LC-MS) | |
| Waschlösung (tryptischer Verdau) | 50% Methanol 45% Reinstwasser 5% Essigsäure | |
| Wassergesättigtes Butanol | 100 mL Reinstwasser 100 mL Butanol | |

7.5 Geräteliste

Tabelle 7.4: *Geräteliste*

| Gerät | Name | Hersteller |
|--------------------------|--------------------------------------|--|
| Anlage für Reinstwasser | MILLIPORE Q-POD | Merck, Darmstadt, Deutschland |
| Elektrophoresekammer | Mini-Protean Tetra System | BioRad, Hercules USA |
| Fractionator | 3100 OFFGEL Fractionator | Agilent Technologies, Santa Clara, USA |
| IEF Fokussierungseinheit | IPG-Phor | Pharmacia Biotech, Uppsala, Schweden |
| Kugelmühle | Kugelmühle MM400 | Retsch, Düsseldorf, Deutschland |
| Photometer | Spectrophotometer Genesys 10S UV-Vis | Thermo Scientific, Waltham, USA |
| Scanner | Biostep ViewPix900 | Biostep, Jahnsdorf, Deutschland |
| Schüttler | Thermomixer Comfort | Eppendorf, Wesseling-Berzdorf, Deutschland |
| Stromversorgungsgerät | PowerPac Basic | BioRad, Hercules USA |
| Vakuumentrifuge | Digital Series SpeedVac SPD121P | Thermo Scientific, Waltham, USA |
| Vortexer | Vortexer | VWR International, Darmstadt, Deutschland |
| Waage | Kern PRS 620-3 Waage | PK Elektronik Vertriebs GmbH, Karlsruhe, Deutschland |
| Zentrifuge | Biofuge primo R H7590 | Thermo Scientific, Waltham, USA |
| Zentrifuge | Biofuge fresco | Thermo Scientific, Waltham, USA |

7.6 Liste der beprobten Biogasanlagen

Tabelle 7.5: Auflistung aller untersuchten Biogasanlagen und ihrer wesentlichen Prozessparametern.

Diese Angaben beruhen im wesentlichen auf den Angaben der Biogasanlagenbetreiber und wurden teilweise im Rahmen des Biogas-Biocoenosis Projekts gesammelt [Theuerl and Klocke (2014)]. Auch können für einzelne Zeitpunkte die Prozessparameter der Biogasanlagen abweichen. Abkürzungen: BGP-Biogasanlage, HF-Hauptfermenter, Plug Flow-Propfenstromfermenter, CSTR-volldurchmischer Rührkesselreaktor, HTK-Hühnertrockenkot.

| BGP | Verfahrenstechnik | | Größe in [m ³] | Einsatzstoffe | | Tempera- turbereich | Raumbelas- tung HF in [kg _{oTS} /(m ³ ·d)] | Verweil- zeit HF in [d] | |
|---------------------|-------------------|--------------|-------------------------------|---------------|------------------|------------------------|--|-------------------------------|-----------------------------------|
| | Fermenter | Typ | | Substrate | Anteil in [%] | | | | Gesamt in [t _{FM} /a] |
| BGP1 (Bischheim) | 2x Haupt- | CSTR | 1.000 | Maissilage | 85 | 10.000 | mesophil | 4,1 | 76 |
| | 1x Nach- | CSTR | 1.300 | Sorghum | 15 | | | | |
| | 2x Endlager | CSTR | 1.800 | | | | | | |
| BGP2 (Neumann) | 1x Hydrolyse- | Ring in Ring | 500 | Maissilage | 65 | 10.400 | mesophil | 5,1 | 42 |
| | 1x Haupt- | CSTR | 1.130 | Rindergülle | 35 | | | | |
| | 1x Endlager | CSTR | 2.677 | | | | | | |
| BGP3 (Erchinger) | 1x Haupt- | Plug Flow | 270 | Maissilage | 35 | 6.200 | mesophil | 13,1 | 15 |
| | 1x Nach- | CSTR | 370 | Grassilage | 10 | | | | |
| | 1x Endlager | CSTR | 2.000 | Grünroggen | 10 | | | | |
| | | | | Ferkelmist | 5 | | | | |
| | | | Rindergülle | 40 | | | | | |
| BGP4 (Eckel) | 2x Haupt- | CSTR | 1.500 | Maissilage | 55 | 12.800 | mesophil | 2,7 | 43 |
| | 1x Endlager | CSTR | 1.500 | Grünroggen | 10 | | | | |
| | | | | Rindermist | 15 | | | | |
| | | | | Rindergülle | 20 | | | | |
| BGP5 (Kuhls) | 2x Haupt- | Plug Flow | 325 | Maissilage | 80 | 11.200 | thermophil | 14,6 | 21 |
| | 1x Nach- | CSTR | 1.600 | HTK | 20 | | | | |

| BGP | Verfahrenstechnik | | | Einsatzstoffe | | | Temperaturbereich | Raumbelastung HF in [kg _{oTS} /(m ³ ·d)] | Verweilzeit HF in [d] |
|--------------------------|--|-----------------------------------|-------------------------------|---|--------------------------|-----------------------------------|-------------------|---|--------------------------|
| | Fermenter | Typ | Größe in [m ³] | Substrate | Anteil in [%] | Gesamt in [t _{FM} /a] | | | |
| | 1x Nach- | CSTR | 2.250 | | | | | | |
| BGP6 (Meyer-Hochheim) | 1x Haupt- 1x Methan- 1x Nach- 1x Endlager | Plug Flow CSTR CSTR CSTR | 270 1.065 370 2.000 | Maissilage Gerstenkorn | 99 1 | 3700 | mesophil | 8,9 | 36 |
| BGP7 (Schwenker) | 1x Haupt- 1x Nach- 1x Endlager | CSTR CSTR CSTR | 1.750 1.750 4.400 | Maissilage Grünroggen Rindermist Rindergülle | 45 5 10 40 | 10.900 | mesophil | 3,9 | 59 |
| BGP8 (Demes) | 1x Haupt- 1x Nach- 1x Endlager | CSTR CSTR CSTR | 1.750 2.100 3.880 | Maissilage Grünroggen Grassilage Rindermist Rindergülle | 54 10 1 5 30 | 14.200 | mesophil | 5,0 | 45 |
| BGP9 (Bürger) | 1x Haupt- 1x Endlager | CSTR CSTR | 2.280 3.470 | Maissilage HTK Gülle | 45 20 35 | 11.800 | mesophil | 3,3 | 71 |
| BGP10 | 1x Haupt- 1x Endlager | CSTR CSTR | 1.400 3.380 | Maissilage Gras Mischgülle | 50 10 40 | 6.900 | mesophil | 2,9 | 74 |
| BGP11 (Ebendorf) | 2x Haupt- 4x Nachgärer 2x Endlager | CSTR CSTR CSTR | 3.500 — — | HTK Getreidesilage Süßhirse | 30 50 20 | 36.000 | mesophil | — | 35 |
| BGP12 | 1x Haupt- | CSTR | 1.557 | Maissilage | 55,4 | 38.900 | mesophil | 7,3 | 40 |

| BGP | Verfahrenstechnik | | | Einsatzstoffe | | | Temperaturbereich | Raumbelastung HF in [kg _{oTS} /(m ³ ·d)] | Verweilzeit HF in [d] |
|-------------|-------------------|------|-------------------------------|---------------|------------------|-----------------------------------|-------------------|---|--------------------------|
| | Fermenter | Typ | Größe in [m ³] | Substrate | Anteil in [%] | Gesamt in [t _{FM} /a] | | | |
| (Bielefeld) | 1x Endlager | | | Zuckerrübe | 7,6 | | | | |
| | | | | Kartoffeln | 5,2 | | | | |
| | | | | Gülle | 31,8 | | | | |
| Batchtest | Schottflasche | CSTR | 0,0005 | Gärrest | 300 g | — | mesophil | — | 5 bzw. 30 |
| | | | | Stroh und Heu | 9,2 g | | | | |

Tabelle 7.6: Auflistung aller untersuchten Biogasanlagen und ihrer wesentlichen Prozessparametern.

Diese Angaben beruhen im wesentlichen auf der Publikation *de Vrieze et al. (2015)* [De Vrieze et al. (2015)]. Abkürzungen: BGP-Biogasanlage, RT-Verweilzeit, FA-Freies Ammonium, Leit.-Leitfähigkeit, CSTR-volldurchmischer Rührkesselreaktor, UASB-Anaerobes Schlamm-aufstromrohr, HTK-Hühnertrockenkot, EPS-Energiepflanzen-silage, SA-Schlachtereiabfälle, VS-organische Trockengewicht, TS-Trockengewicht.

| Probe | Type | Größe in [m ³] | Substrate | pH [-] | TAN [mg/L] | Biogas [m ³ /(m ³ *d)] | Fettsäuren [mg _{COD} /L] | Raumbelastung [kg _{COD} /(m ³ *d)] | RT [d] | Temp. [°C] | FA [mg/L] | Leit. [mS/cm] | V(S)S [g/L] | T(S)S [g/L] |
|---------|------|-------------------------------|--|-----------|---------------|---|--------------------------------------|---|-----------|---------------|--------------|------------------|----------------|----------------|
| Gent 01 | CSTR | 20 | Haushaltsabfälle | 8,50 | 1540 | 6 | — | 15,0 | 20 | 50 | 743 | — | 132 | 155 |
| Gent 02 | CSTR | — | Haushaltsabfälle | 8,50 | 1690 | 6,25 | 130,1 | 13,8 | 22 | 50 | 816 | — | 131 | 153 |
| Gent 03 | CSTR | 30 | Haushaltsabfälle | 8,40 | 2500 | 5 | — | 8,8 | 30 | 42 | 776 | — | 137 | 160 |
| Gent 04 | CSTR | — | Haushaltsabfälle | 8,30 | 1450 | 3,25 | 3346,6 | 8,8 | 25 | 52 | 579 | — | 87 | 109 |
| Gent 05 | CSTR | — | Haushaltsabfälle | 8,20 | 2730 | 3,25 | 0,0 | 6,0 | 38 | 50 | 870 | — | 146 | 309 |
| Gent 06 | CSTR | — | Maissilage HTK | — | 1430 | 11 | 2369,0 | 13,8 | 124 | 55 | — | — | 138 | 319 |
| Gent 07 | CSTR | — | Maissilage HTK | — | 1430 | 11 | 49,0 | 13,8 | 33 | 55 | — | — | 170 | 389 |
| Gent 08 | CSTR | — | Maissilage HTK | — | 1430 | 11 | — | 13,8 | 23 | 55 | — | — | 119 | 418 |
| Gent 09 | CSTR | — | Maissilage Rindergülle | — | 2120 | 2,5 | 5693,1 | 6,8 | 23 | 42 | — | — | 154 | 349 |
| Gent 10 | CSTR | 150 | Gemischt Abfälle | 8,60 | 3800 | 3 | — | 5,0 | 150 | 50 | 2053 | — | 138 | 279 |
| Gent 11 | CSTR | 54 | Gemischte Abfälle | 8,30 | 3450 | 3 | — | 5,0 | 54 | 47 | 1134 | — | 176 | 304 |
| Gent 12 | CSTR | — | Haushaltsabfälle | 8,30 | 1150 | 2,5 | 428,7 | 5,0 | 23 | 47 | 378 | — | 131 | 372 |
| Gent 13 | CSTR | 34 | Haushaltsabfälle | 8,30 | 2100 | 3 | — | 8,8 | 34 | 47 | 691 | — | 111 | 254 |
| Gent 14 | CSTR | — | Haushaltsabfälle | 8,50 | 2150 | 4,75 | 501,7 | 8,8 | 45 | 47 | 940 | — | 157 | 623 |
| Gent 15 | CSTR | 1500 | Maissilage Gülle | 8,52 | 4647 | 5,6 | 5735,0 | 1,5 | 100 | 38 | 1460 | 29 | 77 | 127 |
| Gent 16 | UASB | 274 | Brauerreiabwasser | 7,14 | 253 | 1,5 | 0,0 | 3,3 | 0,38 | 34 | 4 | 6 | 48 | 57 |
| Gent 17 | UASB | — | Kartoffelabwasser | 7,12 | 836 | — | 325,0 | — | — | 34 | 12 | 8 | 20 | 30 |
| Gent 18 | CSTR | 1000 | Maissilage Fette Früchte | 8,19 | 2904 | 5,7 | 912,0 | — | 40 | 34 | 410 | 38 | 48 | 133 |
| Gent 19 | CSTR | 3255 | Klärschlamm Gülle | 7,35 | 508 | 2,7 | 0,0 | 3,0 | 20 | 33 | 11 | 8 | 18 | 36 |
| Gent 20 | CSTR | 1000 | SA | 8,03 | 2428 | 7,5 | 11110,0 | 11,0 | 20 | 54 | 696 | 31 | 17 | 24 |
| Gent 21 | CSTR | 1500 | Organische Abfälle | 8,10 | 3896 | 6,4 | 242,0 | 2,5 | 60 | 34 | 459 | 32 | 79 | 124 |
| Gent 22 | CSTR | 1200 | Maissilage Gülle | 8,06 | 3497 | 1,1 | 4467,0 | 4,0 | 40 | 34 | 380 | 40 | 33 | 59 |
| Gent 23 | CSTR | 3600 | Gülle (10%) SA (50-60%) EPS (30-40%) | 7,52 | 3288 | 1,5 | 6610,0 | 4,5 | 40 | 54 | 363 | 32 | 28 | 42 |
| Gent 24 | CSTR | 1200 | Maissilage Gülle | 8,00 | 3522 | 1,1 | 4688,0 | 4,0 | 40 | 34 | 338 | 40 | 33 | 58 |
| Gent 25 | CSTR | 1000 | SA | 8,03 | 2428 | 7,5 | 11110,0 | 11,0 | 20 | 54 | 696 | 25 | 35 | 51 |

| Probe | Type | Größe in [m ³] | Substrate | pH [-] | TAN [mg /L] | Biogas [m ³ /(m ³ *d)] | Fettsäuren [mg _{COD} /L] | Raumbelastung [kg _{COD} /(m ³ *d)] | RT [d] | Temp. [°C] | FA [mg/L] | Leit. [mS/cm] | V(S)S [g/L] | T(S)S [g/L] |
|---------|------|-------------------------------|--|-----------|----------------|---|--------------------------------------|---|-----------|---------------|--------------|------------------|----------------|----------------|
| Gent 26 | CSTR | 1250 | Gülle (hauptsächlich) | 8,05 | 4982 | 4,1 | 713,0 | 2,5 | 60 | 34 | 530 | 42 | 68 | 103 |
| Gent 27 | CSTR | 3600 | Gülle (10%) SA (50-60%) EPS (30-40%) | — | 3109 | 4,1 | 3087,0 | 4,5 | 40 | 54 | — | 25 | — | — |
| Gent 28 | CSTR | 2000 | Maissilage Gülle | 7,92 | 3123 | 7,4 | 8114,0 | 4,0 | 30 | 54 | 743 | 39 | 67 | 109 |
| Gent 29 | CSTR | 1500 | Maissilage Gülle | 8,24 | 4282 | 5,6 | 0,0 | 1,5 | 100 | 34 | 666 | 37 | 86 | 159 |
| Gent 30 | CSTR | 2500 | Maissilage Gülle | 8,12 | 4032 | 4,8 | 557,0 | 3,0 | 60 | 34 | 495 | 31 | 68 | 102 |
| Gent 31 | CSTR | 3000 | Maissilage Gülle | 7,76 | 4986 | 5,7 | 36760,0 | 5,0 | 40 | 34 | 287 | 25 | 50 | 74 |
| Gent 32 | CSTR | 3200 | Maissilage Gülle | 8,25 | 6427 | 2,1 | 434,0 | 4,0 | 40 | 34 | 1020 | 62 | 63 | 112 |
| Gent 33 | CSTR | 1500 | Organische Abfälle | 8,02 | 4019 | 2,8 | 5593,0 | 3,0 | 80 | 34 | 402 | 32 | 16 | 77 |
| Gent 34 | CSTR | 4000 | Klärschlamm | 7,35 | 1077 | 2,0 | 0,0 | 3,0 | 18 | 34 | 25 | 8 | 21 | 44 |
| Gent 35 | CSTR | 3600 | Gülle (10%) SA (50-60%) EPS (30-40%) | 7,52 | 3197 | 2,0 | 7609,0 | 5,6 | 40 | 54 | 353 | 33 | 29 | 45 |
| Gent 36 | UASB | — | Kartoffelabwasser | 7,10 | 128 | — | 0,0 | — | — | 34 | 2 | 9 | 43 | 54 |
| Gent 37 | CSTR | — | Unbekannt | 8,30 | 3091 | — | 0,0 | — | — | 34 | 540 | 39 | 13 | 21 |
| Gent 38 | CSTR | — | Gülle Essensreste | 8,35 | 4639 | — | 1096,0 | — | — | 34 | 890 | 62 | 55 | 95 |
| Gent 39 | CSTR | 1500 | Organische Abfälle | 8,02 | 2684 | 2,8 | 22601,0 | 3,0 | 80 | 34 | 268 | 32 | 24 | 44 |
| Gent 40 | UASB | 1210 | Abwässer Papiermühle | 7,19 | 222 | 1,2 | 251,0 | 5,6 | 0,46 | 35 | 4 | 7 | 85 | 165 |

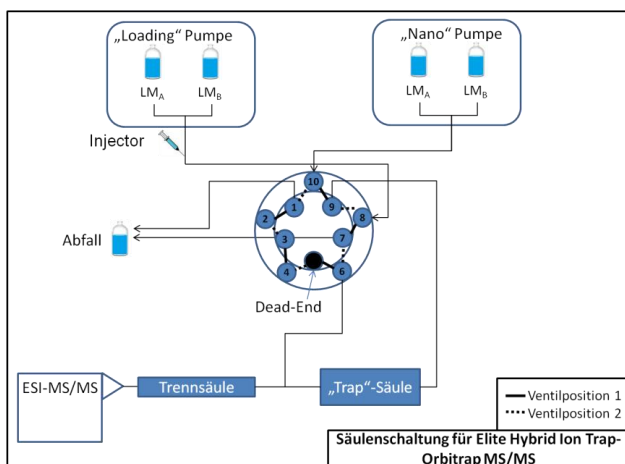
7.7 Komponenten und Parameter für das Elite Hybrid Ion Trap-Orbitrap MS/MS

Tabelle 7.7: Komponenten und Parameter für das Elite Hybrid Ion Trap-Orbitrap MS/MS

| HPLC Komponenten | |
|--------------------|---|
| HPLC (UPLC) | UltiMate 3000 RSLCnano splitless LC-System (Thermo Scientific, Waltham, USA) |
| Autosampler | WPS-3000TPL RS |
| Pumpensystem | NCS-3500RS NANO |
| Entgaser | SRD-3400 |
| Trennsäule | Acclaim PepMap RSLC, C18, 2 μm , 75 μm , 25 cm; 164536 (Thermo Scientific, Waltham, USA) |
| Vorsäule | Nano Trap Column, 100 μm i.d. x 2 cm, gepackt mit Acclaim Pep-Map100 C18, 5 μm , 100 Å, nanoViper; 164564 (Thermo Scientific, Waltham, USA) |
| Injektionsvolumen | 8 μL |
| Kapillarfluss | Lösungsmittel A: 98% VE Wasser, 2% Acetonitril, 0,05% Trifluoressigsäure Lösungsmittel B: 20% VE Wasser, 80% Acetonitril, 0,05% Trifluoressigsäure |
| Nanofluss | Lösungsmittel A: 98% VE Wasser, 2% Acetonitril, 0,1% Ameisensäure Lösungsmittel B: 10% VE Wasser, 10% Trifluorethanol, 80% Acetonitril, 0,1% Ameisensäure |
| Gradient | 30 min/120 min |
| MS-Komponenten | |
| MS | Elite Hybrid Ion Trap-Orbitrap Massenspektrometer (Thermo Fisher Scientific, Waltheim, USA) |
| Quelle | ESI-Nanoquelle |
| Auflösung | 200 |
| m/z-Bereich | 350-2.000 Da |
| Precursorauswahl | Intensivsten 20 Peaks |
| „Active Exclusion“ | 5 s für 2 gleiche Spektren |
| Fragmentierung | CID |

7.7 Komponenten und Parameter für das Elite Hybrid Ion Trap-Orbitrap MS/MS

| | |
|---|--|
| Modus | Positiver Ionen Modus |
| Parameter für Proteinidentifikation | |
| MS Toleranz | 10 ppm |
| MS/MS Toleranz | 0,5 Da |
| Masse | Monoisotopisch |
| Modifikationen | Carbamidomethyl (C) (variabel), Oxidation (M) (variabel) |
| ¹³ C | 1 |
| Ladung | +2, +3 |
| Enzyme | Trypsin |
| Methode | |
| CID_Top20_excl30s_ColSpray_C18pep30mingrad | |
| CID_Top20_excl30s_ColSpray_C18pep120mingrad | |



| Zeit in [min] | „Loading“ Pumpe Fluss in [$\mu\text{L}/\text{min}$] | „Loading“ Pumpe %B | „Nano“ Pumpe Fluss in [$\mu\text{L}/\text{min}$] | „Nano“ Pumpe %B | Ventil |
|---------------|---|--------------------|--|-----------------|--------|
| 0 | 7 | 0 | 0,3 | 4 | 10_1 |
| 5 | 7 | 0 | 0,3 | 4 | 1_2 |
| 120 | 7 | 0 | 0,3 | 55 | 1_2 |
| 120 | 7 | 0 | 0,3 | 90 | 1_2 |
| 125 | 7 | 0 | 0,3 | 90 | 1_2 |
| 125 | 7 | 0 | 0,3 | 4 | 1_2 |
| 150 | 7 | 0 | 0,3 | 4 | 1_2 |

Abbildung 7.2: Säulenkonfiguration und Gradient des Elite Hybrid Ion Trap-Orbitrap MS/MS

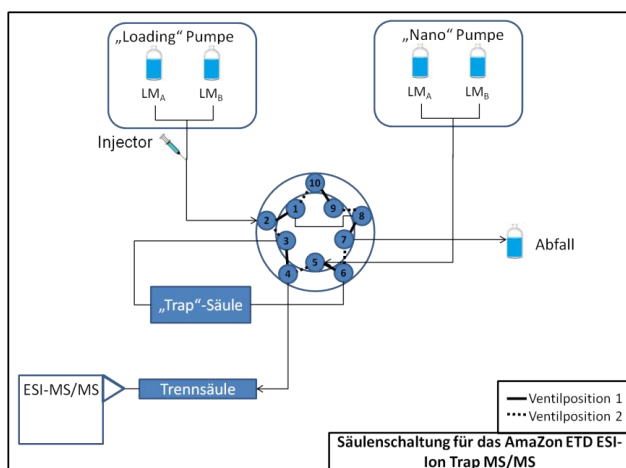
7.8 Komponenten und Parameter für das AmaZon ETD ESI-Ion Trap MS/MS

Tabelle 7.8: Komponenten und Parameter für das AmaZon ETD ESI-Ion Trap MS/MS

| HPLC Komponenten | |
|-------------------|---|
| UPLC | UltiMate 3000 nano splitted LC-System (Thermo Scientific, Waltham, USA) |
| Autosampler | WPS-3000TPL RS |
| Flussmanager | FLM -3x00 |
| Pumpensystem | NCS-3500RS NANO |
| Entgaser | SRD-3400 |
| Trennsäule | Acclaim PepMap RSLC, C18, 2 μm , 75 μm , 25 cm; 164536 (Thermo Scientific, Waltham, USA) |
| Vorsäule | Nano Trap Column, 100 μm i.d. x 2 cm, gepackt mit Acclaim Pep-Map100 C18, 5 μm , 100 Å, nanoViper; 164564 (Thermo Scientific, Waltham, USA) |
| Injektionsvolumen | 5 μL |
| Kapillarfluss | Lösungsmittel A: 98% VE Wasser, 2% Acetonitril, 0,05% Trifluoressigsäure Lösungsmittel B: 20% VE Wasser, 80% Acetonitril, 0,05% Trifluoressigsäure |
| Nanofluss | Lösungsmittel A: 98% VE Wasser, 2% Acetonitril, 0,1% Ameisensäure Lösungsmittel B: 10% VE Wasser, 10% Trifluorethanol, 80% Acetonitril, 0,1% Ameisensäure |
| Gradient | 30 min/90 min/120 min |
| MS-Komponenten | |
| MS | AmaZon ETD ESI-Ionenfalle MS (Bruker Daltonik, Bremen, Deutschland) |
| Quelle | ESI-Nanoquelle |
| Auflösung | 20,000 |
| m/z-Bereich | MS: 300 - 1.500 m/z, MS/MS: 100 - 2.400 m/z |
| Precursorauswahl | Intensivsten 10 Peaks + über 1% der Intensität |
| Ionenkapazität | 100,000 in max. 100 ms |

7.8 Komponenten und Parameter für das Amazon ETD ESI-Ion Trap MS/MS

| | |
|--|---|
| „Active Exclusion“ | 9 s für 2 gleiche Spektren |
| Fragmentierung | CID |
| Modus | Positiver Ionen Modus |
| Parameter für Proteinidentifikation | |
| MS Toleranz | 0,3 Da |
| MS/MS Toleranz | 0,3 Da |
| Masse | monoisotopisch |
| Modifikationen | Carbamidomethyl (C) (variabel), Oxidation (M) (variabel) |
| ¹³ C | 1 |
| Ladung | +2, +3 |
| Enzyme | Trypsin |
| Methoden | |
| LC-Methode | U3000Split_C18_Qstar_load8min_7μL_120min U3000Split_C18_Qstar_load8min_7μL_30min |
| MS-Methode | CaptiveSpray_Umweltproben_MSMS_CID_enh Resol |
| Datenauswertung | RH_KomplexProben_1MComp_FileRename_5-128min |



| Zeit in [min] | „Loading“ Pumpe Fluss in [μL/min] | „Loading“ Pumpe %B | „Nano“ Pumpe Fluss in [μL/min] | „Nano“ Pumpe %B | Ventil |
|---------------|-----------------------------------|--------------------|--------------------------------|-----------------|--------|
| 0 | 7 | 0 | 0,3 | 5 | 1_2 |
| 8 | --- | --- | --- | --- | 10_1 |
| 10 | 7 | 0 | 0,3 | 5 | 10_1 |
| 10,01 | 60 | 0 | --- | --- | 10_1 |
| 11 | 60 | 50 | --- | --- | 10_1 |
| 89 | 60 | 50 | --- | --- | 10_1 |
| 90 | 60 | 90 | --- | --- | 10_1 |
| 100 | --- | --- | 0,3 | 50 | 10_1 |
| 105 | --- | --- | 0,3 | 90 | 1_2 |
| 111 | 60 | 90 | --- | --- | 1_2 |
| 112 | 60 | 0 | --- | --- | 1_2 |
| 113 | --- | --- | 0,3 | 90 | 1_2 |
| 114 | --- | --- | 0,3 | 0 | 1_2 |
| 119 | --- | --- | 0,3 | 0 | 1_2 |
| 120 | --- | --- | 0,3 | 5 | 1_2 |
| 126 | 60 | 0 | --- | --- | 1_2 |
| 127 | 10 | 0 | --- | --- | 1_2 |
| 130 | 10 | 0 | 0,3 | 5 | 1_2 |

Abbildung 7.3: Säulenkonfiguration und Gradient des Amazon ETD ESI-Ion Trap MS/MS

7.9 Multiples Sequenzalignment

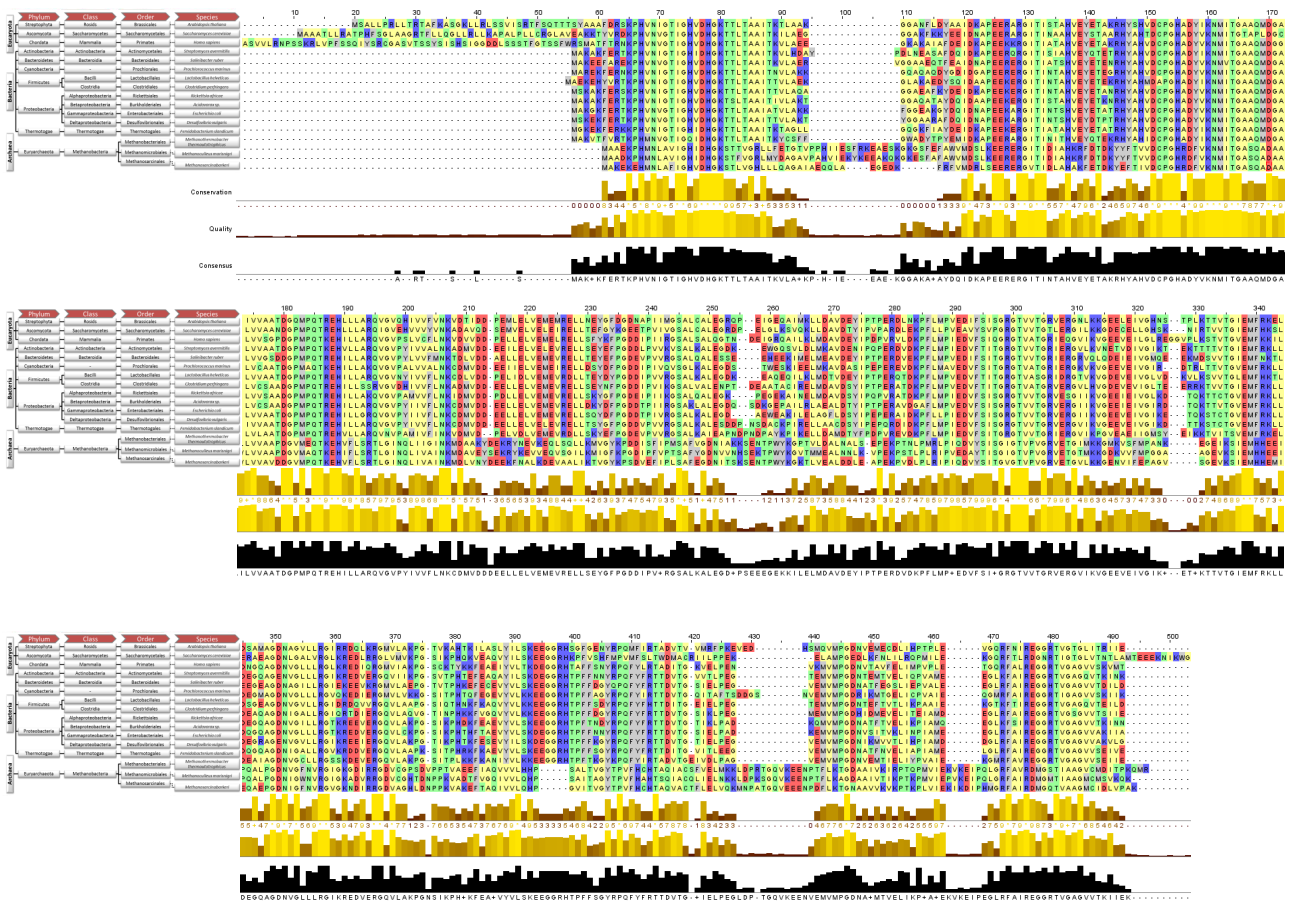


Abbildung 7.4: *Multiples Sequenzalignment* für das Protein *EFTU* basierend auf *EFTU*-Proteinsequenzen diverser Taxa, welche für Biogasprozesse repräsentativ sind. Für die Erstellung wurde die Software Jalview 2.0 genutzt („Average Distance using %Identity“) [Waterhouse et al. (2009)].

7.9 Multiples Sequenzalignment

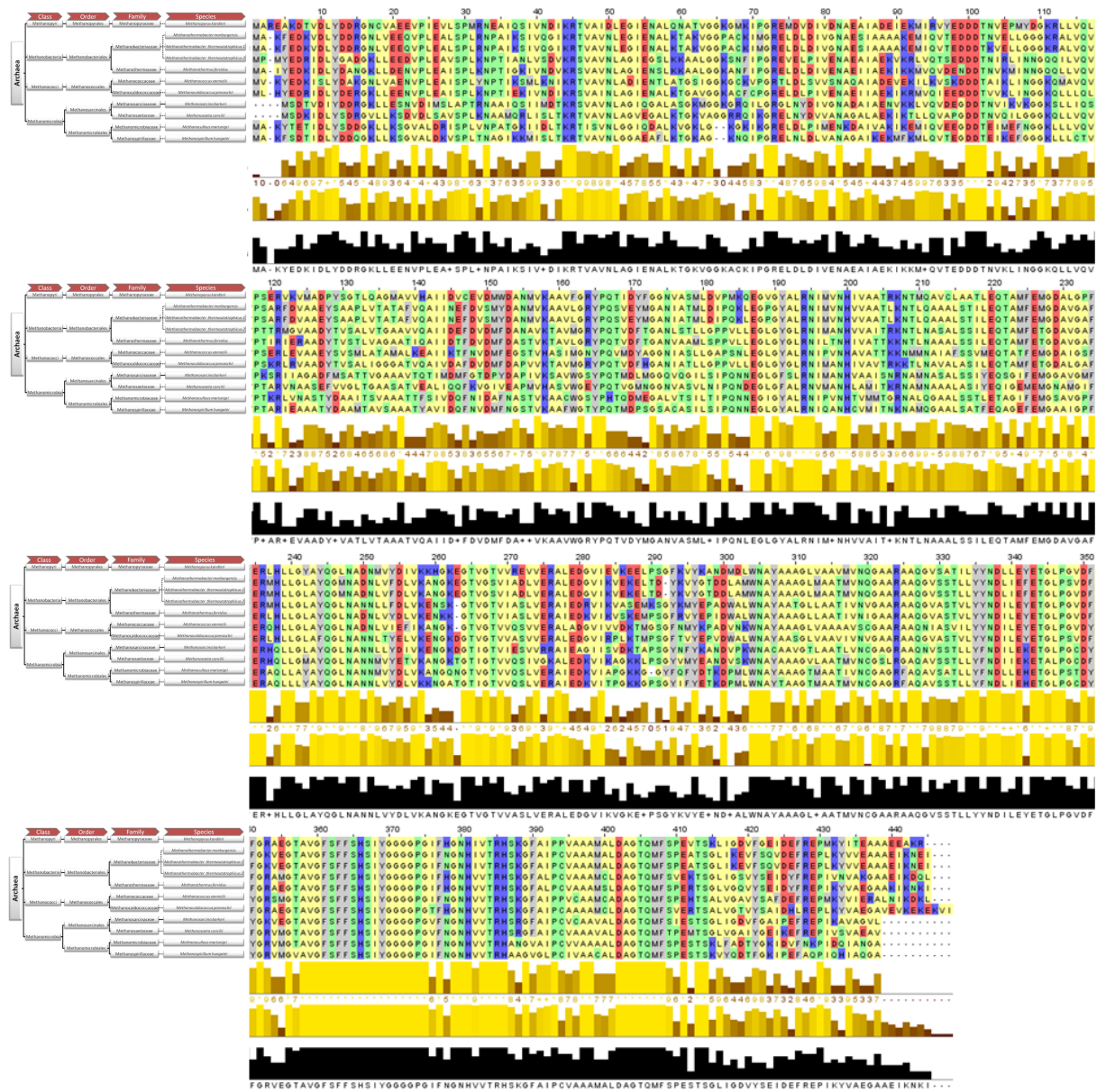


Abbildung 7.5: Multiples Sequenzalignment für das Protein Methyl Coenzym M Reduktase Untereinheit β

basierend auf Methyl CoM Reduktase Untereinheit β -Proteinsequenzen diverser Taxa, welche für Biogasprozesse repräsentativ sind. Für die Erstellung wurde die Software Jalview 2.0 genutzt („Average Distance using %Identity“) [Waterhouse et al. (2009)].

7.10 Methanogenesestoffwechsel für eine mesophile und eine thermophile Biogasanlage

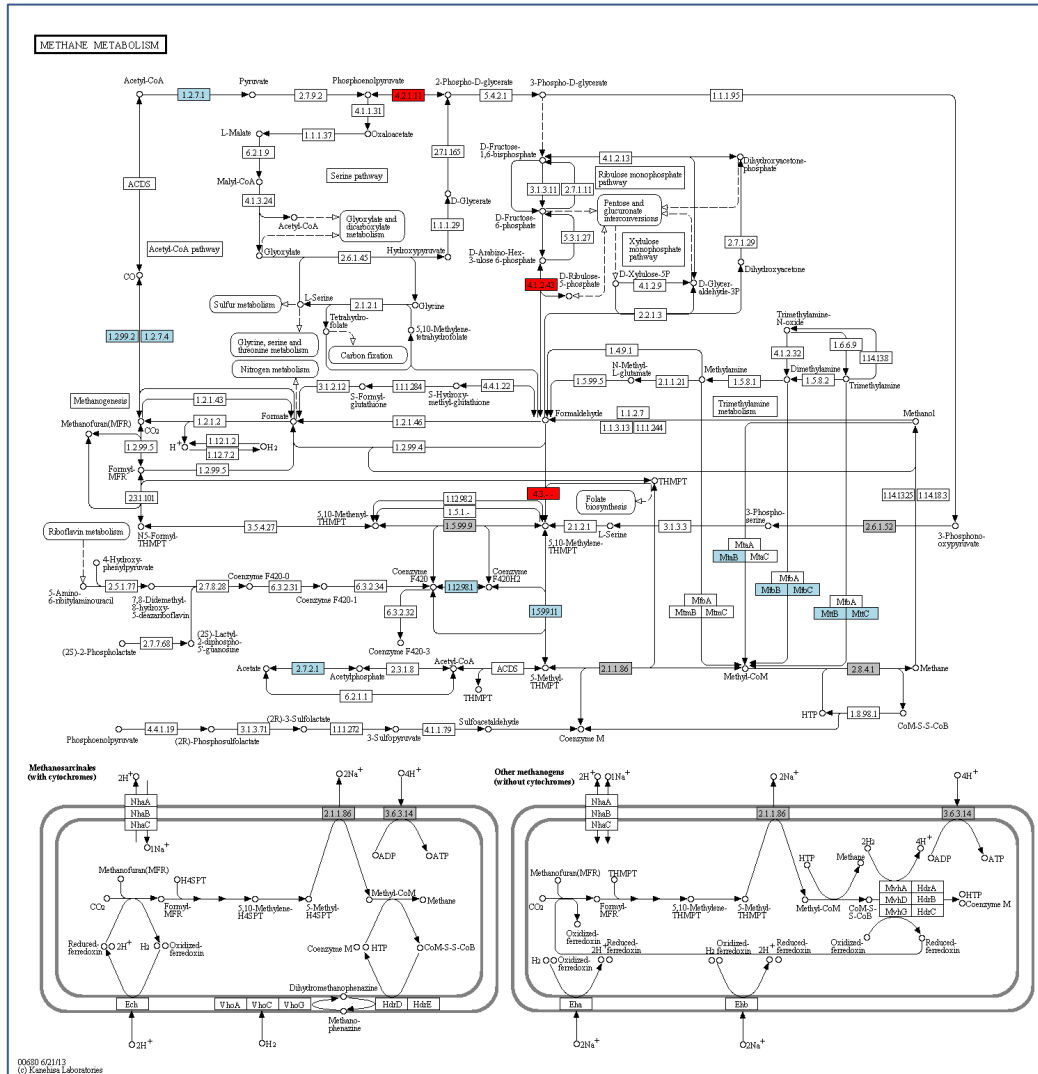


Abbildung 7.6: Methanogenesestoffwechsel für eine mesophile und eine thermophile Biogasanlage

dargestellt im KEGGnetzwerk „METHANE METABOLISM (map00680)“. Blau und Rot sind dabei die identifizierten Proteine der mesophilen bzw. der thermophilen Biogasanlage. Grau markierte Proteine konnten in beiden Biogasanlagen gefunden werden.

7.11 Abhängigkeit der Richness und Evenness vom Schwellwert

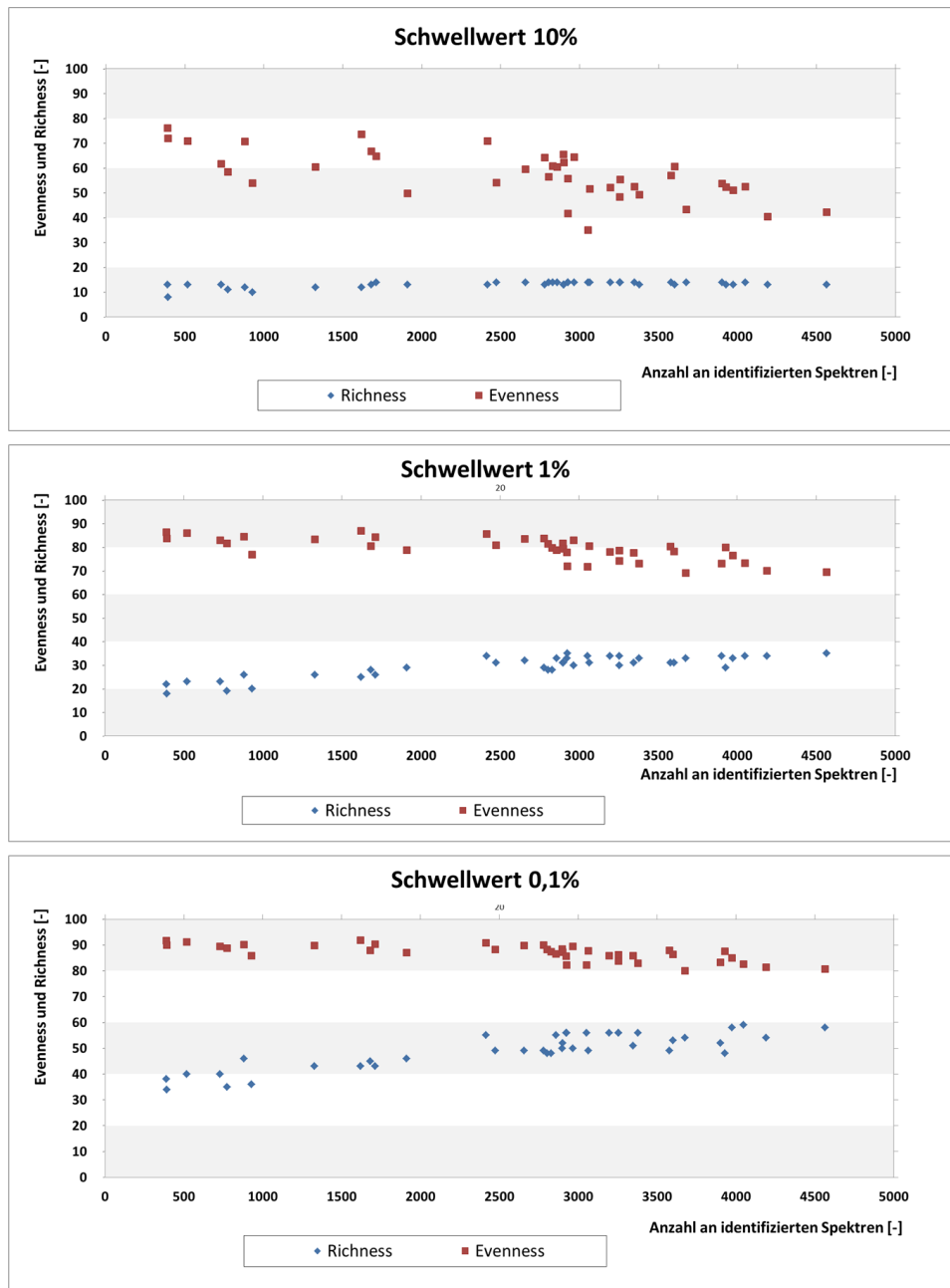


Abbildung 7.7: Abhängigkeit der Richness und Evenness vom Schwellwert.

Die Richness und Evenness wurde dafür auf Ordnungsebene berechnet, wobei nur Ordnungen beachtet wurden, welche in mindestens einer Biogasanlage mehr als 10%, 1% bzw. 0,1% der Spektren (Schwellwert) repräsentierten.

7.12 Zusammenhänge zwischen Prozessparametern visualisiert als GPLOM

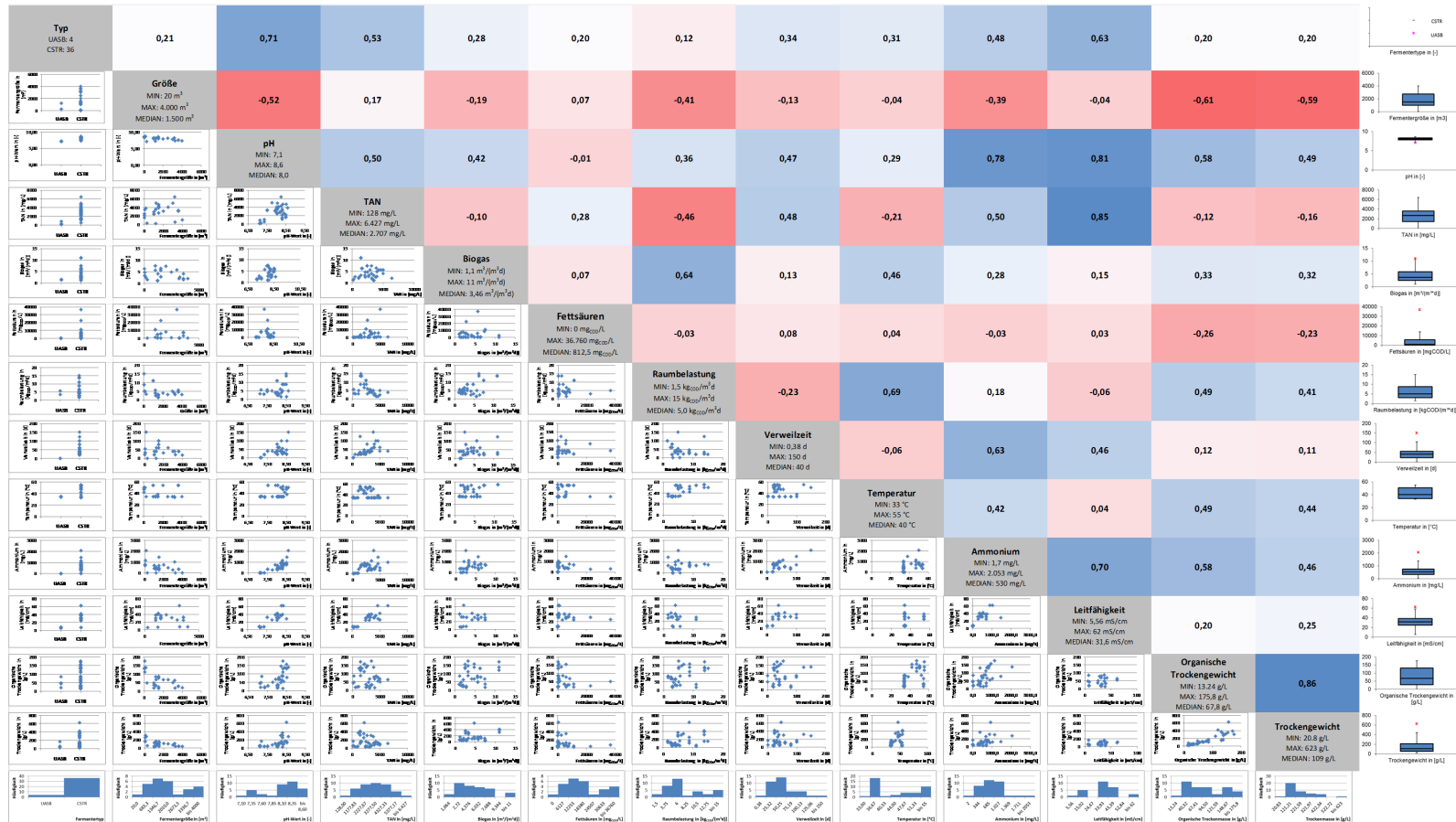


Abbildung 7.8: Zusammenhänge der Prozessparameter der Biogasanlagen, visualisiert als generalized plot matrix (GPLOM) [Im et al. (2013)]. In der unteren Matrix sind die Parameter der einzelnen Biogasanlagen als "Scatter-Plot" (jeweils Parameter gegen Parameter) und in der oberen die Korrelation („Spearman-Korrelation“, p-Wert: <1%) zwischen den untersuchten Prozessparametern dargestellt.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

Im Folgenden wird der Einfluss der Prozesstemperatur, der Raumbelastung, der Verweilzeit und des Gesamtstickstoffes auf den Kohlenstoffmetabolismus basierend auf der „KEGG Karte 1200“ dargestellt.

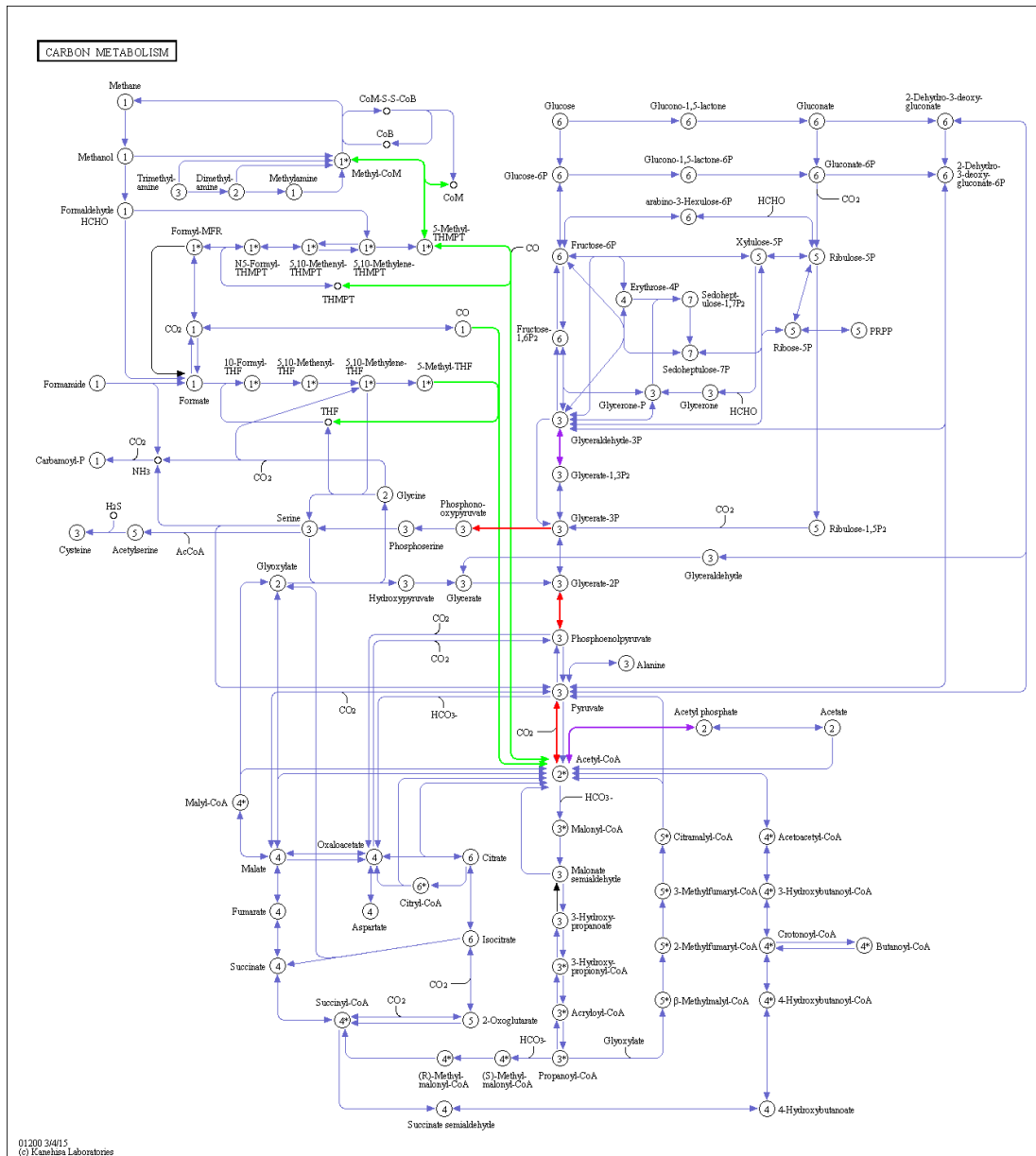


Abbildung 7.9: Einfluss von hoher Prozesstemperatur auf den Kohlenstoffmetabolismus. Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit erhöhter Prozesstemperatur korrelieren ($r > 0$, p -Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

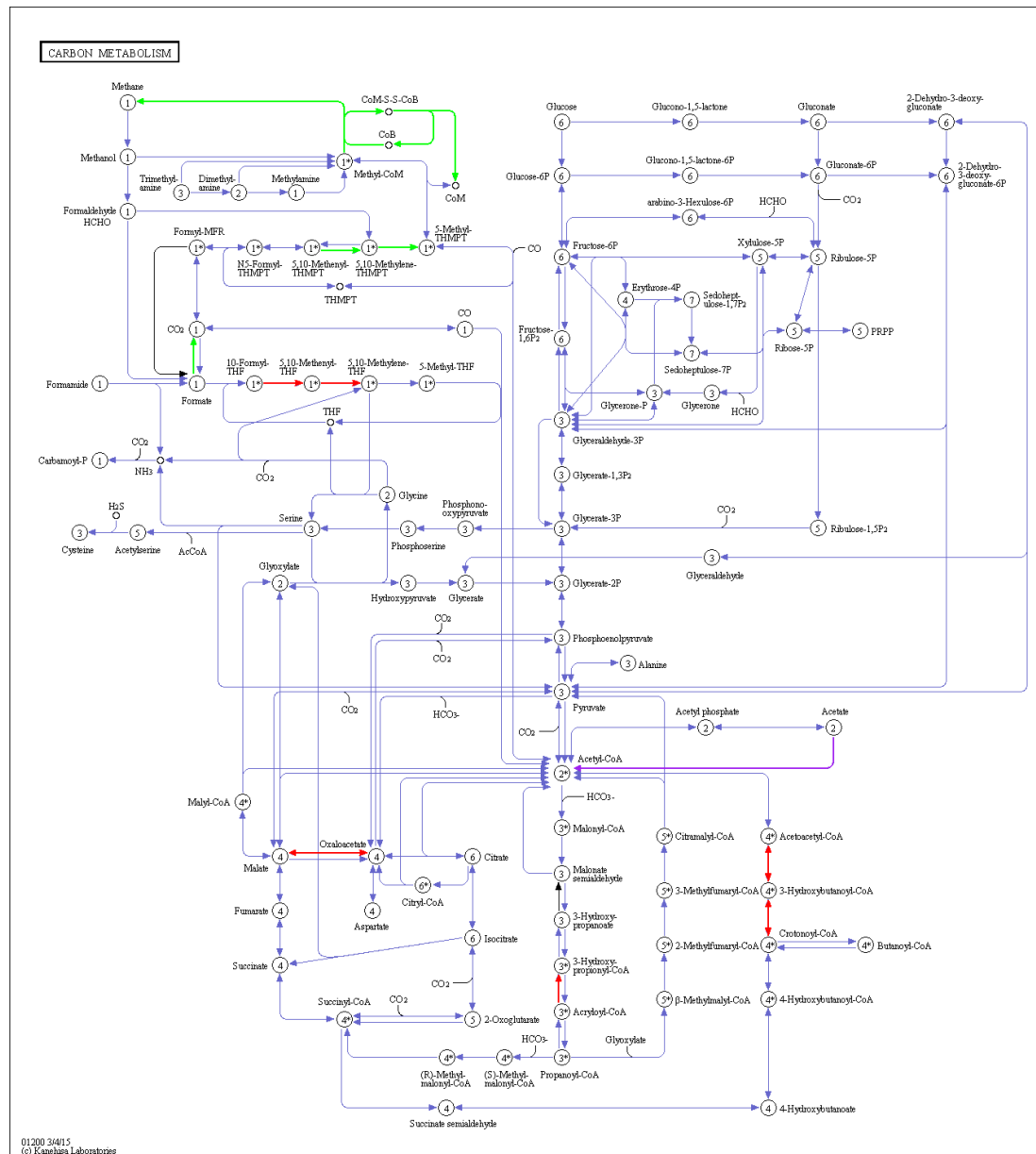


Abbildung 7.10: *Einfluss von niedriger Prozesstemperatur auf den Kohlenstoffmetabolismus.* Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit erhöhter Prozesstemperatur korrelieren ($r > 0$, p -Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

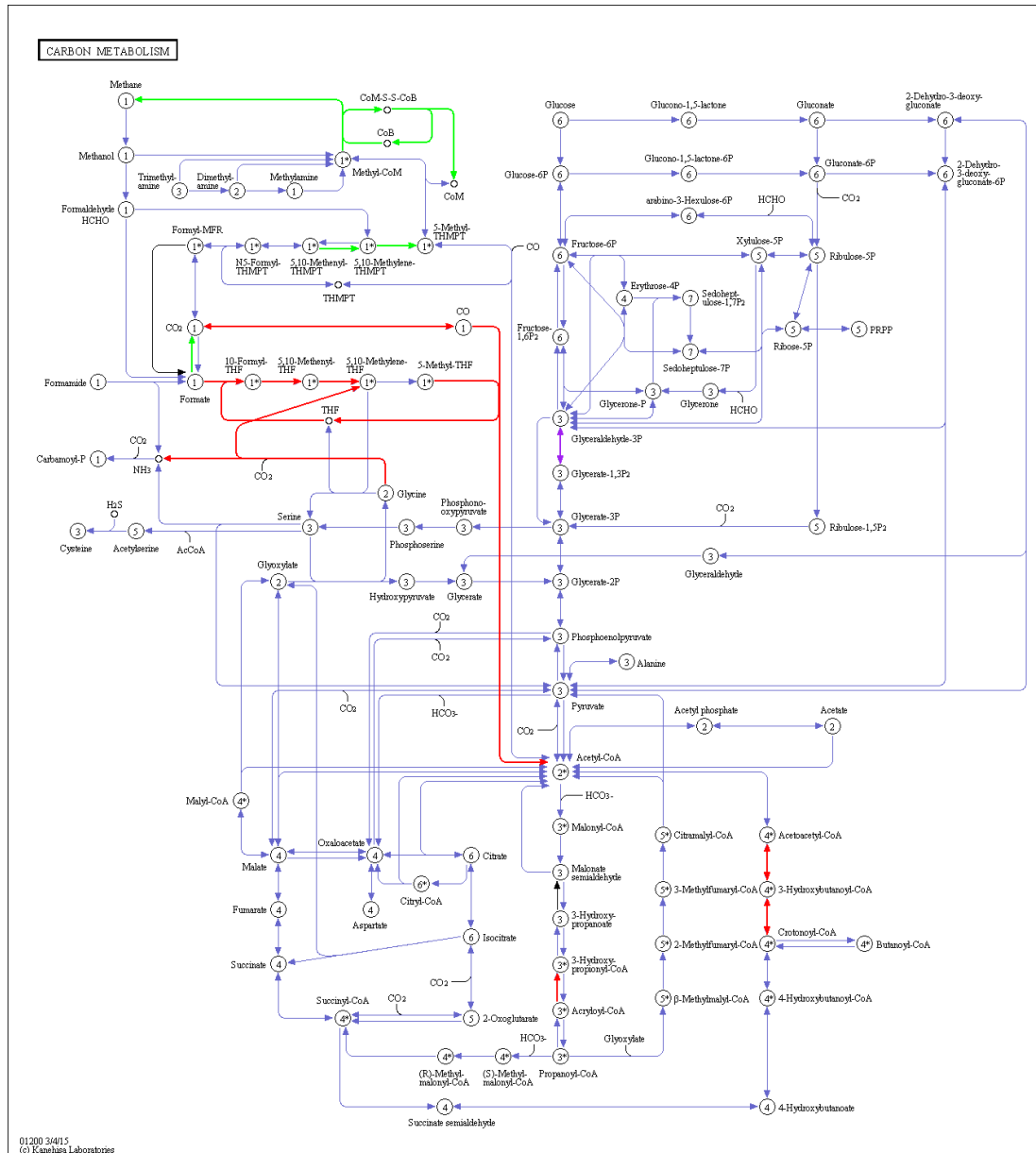


Abbildung 7.11: *Einfluss von hoher Gesamtammoniumstickstoffkonzentration auf den Kohlenstoffmetabolismus.* Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit viel Gesamtstickstoff korrelieren ($r > 0$, p-Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

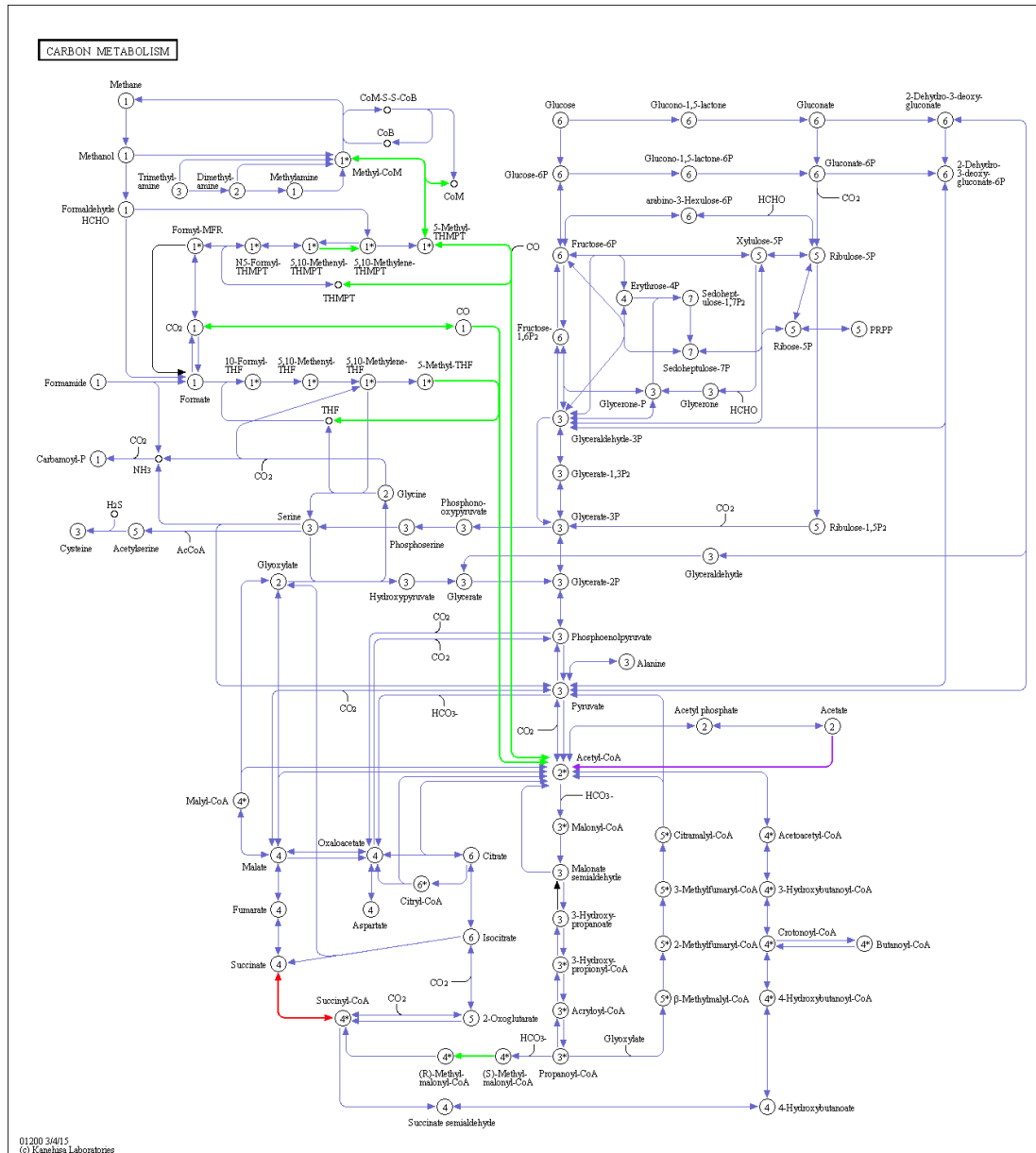


Abbildung 7.12: Einfluss von niedriger Gesamtammoniumstickstoffkonzentration auf den Kohlenstoffmetabolismus. Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit wenig Gesamtstickstoff korrelieren ($r > 0$, p-Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

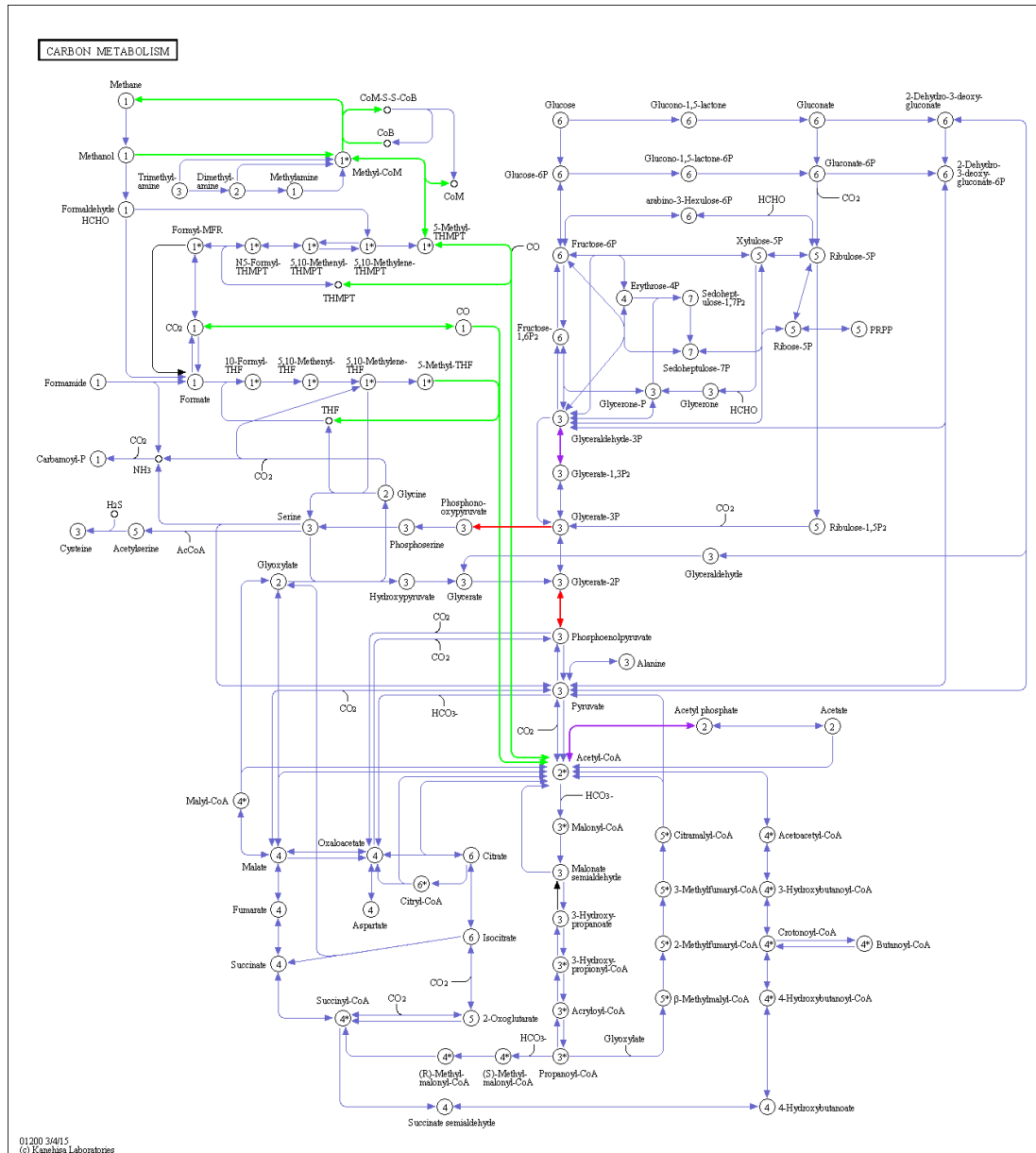


Abbildung 7.13: *Einfluss von hoher Raumbelastung auf den Kohlenstoffmetabolismus.* Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit erhöhter Raumbelastung korrelieren ($r > 0$, p-Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

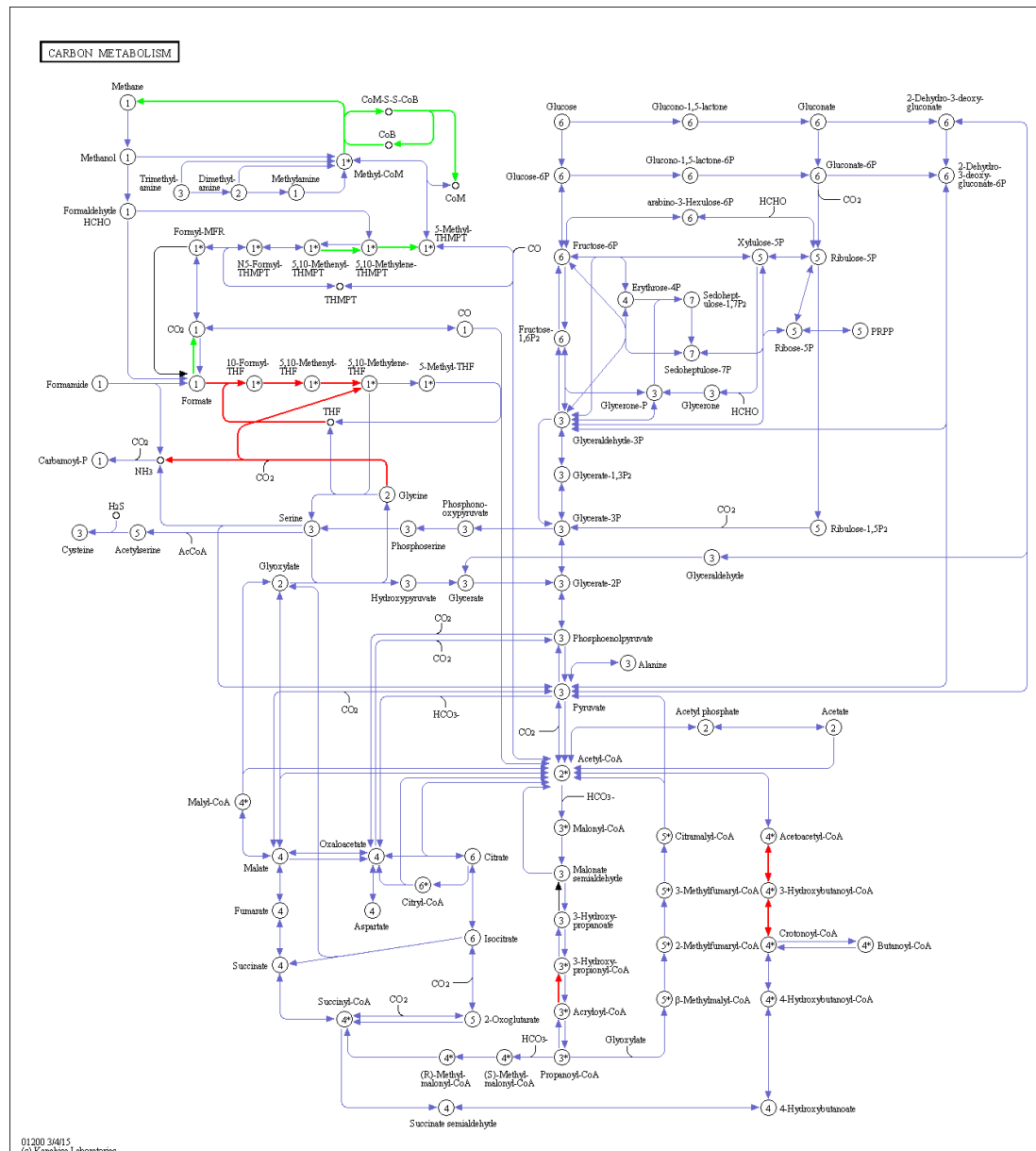


Abbildung 7.14: *Einfluss von niedriger Raumbelastung auf den Kohlenstoffmetabolismus.* Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit niedriger Raumbelastung korrelieren ($r > 0$, p -Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

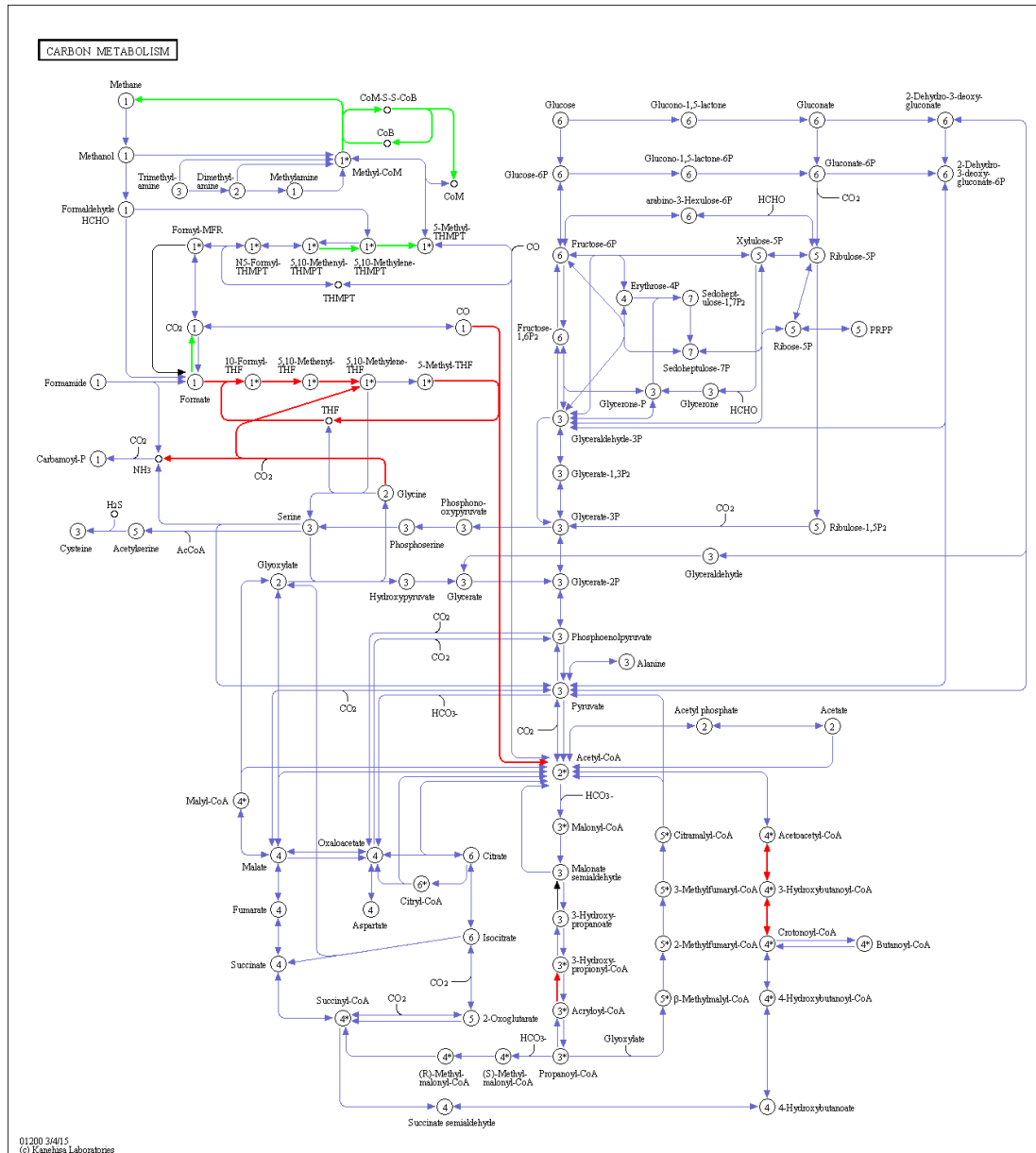


Abbildung 7.15: *Einfluss von langer Schlammverweilzeit auf den Kohlenstoffmetabolismus.* Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit langer Schlammverweilzeit korrelieren ($r: > 0$, p -Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.13 Zusammenhänge zwischen Kohlenstoffmetabolismus und Prozessparametern

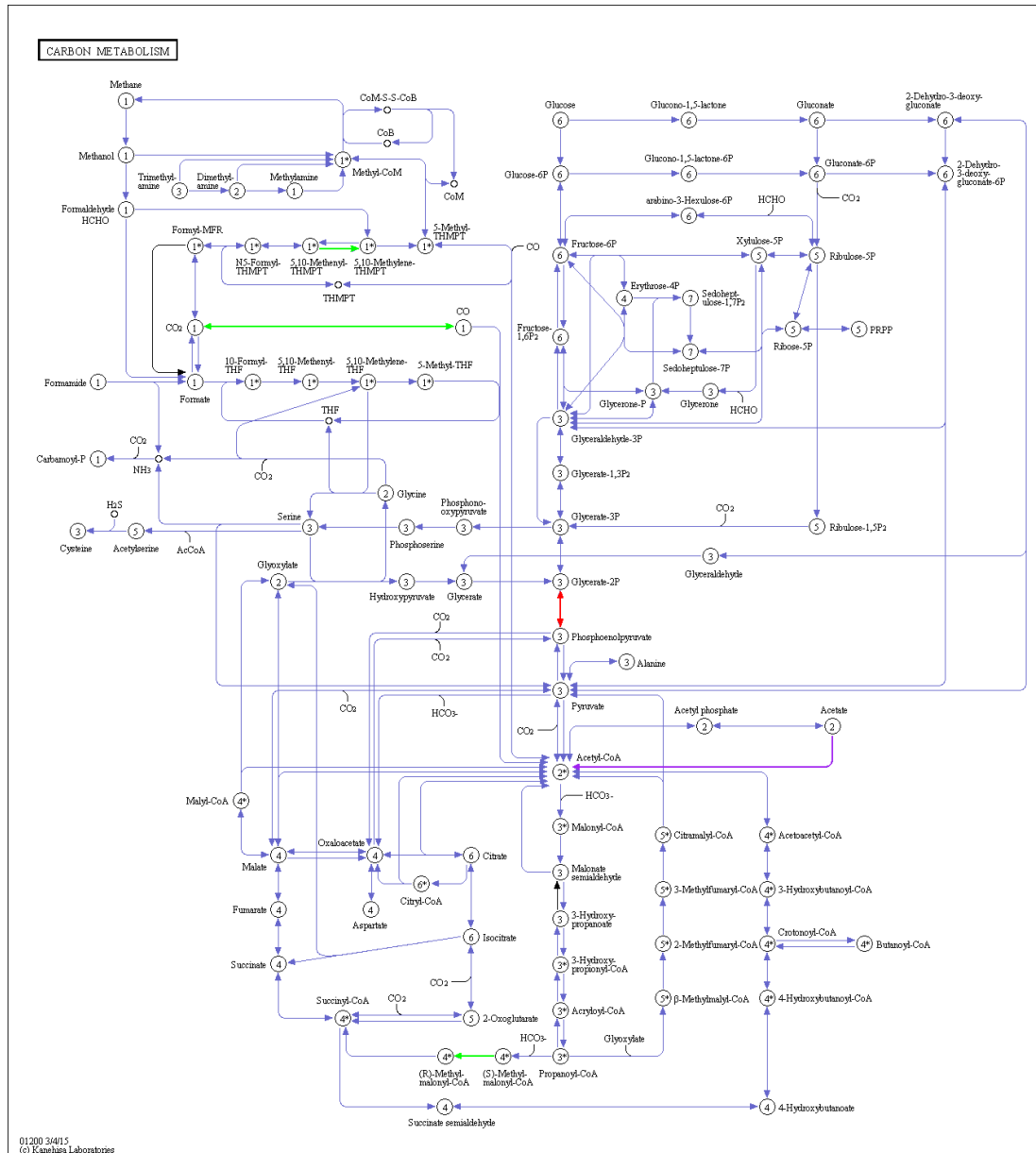


Abbildung 7.16: Einfluss von kurzer Schlammverweilzeit auf den Kohlenstoffmetabolismus. Hervorgehoben sind alle Proteine des Kohlenstoffmetabolismus („KEGG Karte 1200“), welche positiv mit kurzer Schlammverweilzeit korrelieren ($r: > 0$, p -Wert: $< 1\%$). Grün: Archaeenproteine, Rot: Bakterienproteine, Lila: Archaeen- oder Bakterienproteine.

7.14 CD

- Doktorarbeit
- Publikationen
- Betreute Arbeiten
- Standardarbeitsanweisungen
 1. Beprobung von Praxisbiogasanlagen (Arbeitsanweisung 10.5)
 2. Phenolextraktion mit Kugelmühle (Arbeitsanweisung 2.3)
 3. Proteinbestimmung mit Amidoschwarz (Arbeitsanweisung 3.1)
 4. Acetonfällungen (Arbeitsanweisung 2.4)
 5. SDS-PAGE BIORAD klein (Arbeitsanweisung 4.1)
 6. OFFGEL Fractionator 3100 Agilent (Arbeitsanweisung 4.7)
 7. Gelfaerbung Coomassie (Arbeitsanweisung 5.1)
 8. Trocknung von Gelen (Arbeitsanweisung 5.5)
 9. InGel-Verdau von zerschnittenen SDS-PAGE-Spuren (Arbeitsanweisung 6.3)
 10. Messung Bruker AmaZone ESI TRAP (Arbeitsanweisung 6.5)
 11. Messung Orbitrap Uni (Arbeitsanweisung 6.6)
 12. User-Manual MPA (Arbeitsanweisung 12.2)
 13. Jalview (Arbeitsanweisung 12.1)
- Zusätzliche Abbildungen und Tabellen
 1. Prozessdaten 40 Biogasanlagen
 2. Metaproteine + Ordnungen + Biologische Prozesse 40 Biogasanlagen
 3. Korrelationsmatrizen der 40 Biogasanlagen
 4. Schwellenwerte für Metaproteine + Ordnungen + Biologische Prozesse
 5. Faktorkoordinaten der Hauptkomponentenanalysen
 6. Heatmap Metaproteine + Ordnungen + Biologische Prozesse
 7. Ergebnisse „Decision Tree Learning“
 8. Anmerkungen zum Anaerobic Digestion Model 1
- Matlab Scripte

1. Clusteranalyse von SDS-PAGE Gelen
2. Statistische Auswertung von Biogasanlagenstudien