

**Identifizierung und strukturelle Charakterisierung
von Protein-Protein-Interaktionen des
Formiatstoffwechsels in *Escherichia coli* mittels
Massenspektrometrie-basierter Methoden**

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I -
Biowissenschaften

der Martin-Luther-Universität Halle-Wittenberg

von

Herrn Dipl.-Biol. Michael Zorn
geb. am 02.02.1983 in Halle (Saale)

Gutachter: 1. Prof. Dr. Andrea Sinz
 2. Prof. Dr. Gary Sawers
 3. Prof. Dr. Jens Meiler

Datum der öffentlichen Verteidigung: 27.02.2015

Inhaltsverzeichnis

Abbildungsverzeichnis.....	V
Tabellenverzeichnis.....	VIII
Abkürzungsverzeichnis.....	IX
Zusammenfassung.....	XII
1 Einleitung.....	1
1.1 Formiatmetabolismus in <i>Escherichia coli</i>	1
1.1.1 Bildung von Formiat.....	1
1.1.2 Transport von Formiat über die Cytoplasmamembran.....	2
1.1.3 Differentielle Regulation der Formiatdehydrogenasen in <i>E. coli</i>	5
1.1.4 Strukturen der respirativen Formiatdehydrogenasen Fdh-O und Fdh-N.....	6
1.1.5 Synthese von Selenoproteinen in <i>E. coli</i>	8
1.2 Massenspektrometrie.....	10
1.2.1 Elektrospray-Ionisation.....	10
1.2.1.1 Modell des geladenen Rückstands.....	12
1.2.1.2 Modell der Ionenverdampfung.....	12
1.2.2 Nano-ESI.....	13
1.2.3 LTQ-Orbitrap-Massenspektrometrie.....	13
1.3 Untersuchung von Proteinen und Peptiden mittels Massenspektrometrie....	16
1.3.1 Massenspektrometrische Analyse von co- und posttranslationalen Modifikationen.....	16
1.3.2 Peptid- und Protein-Quantifizierung mittels Massenspektrometrie.....	18
1.3.2.1 Quantifizierung durch Isotopenmarkierung von Peptiden und Proteinen.....	18
1.3.2.2 Markierungsfreie Quantifizierung.....	19
1.4 Charakterisierung von Protein-Protein-Interaktionen.....	19
1.4.1 Identifizierung und Charakterisierung von Protein-Protein- Interaktionen mittels chemischer Quervernetzung.....	20

1.4.1.1	Reaktivität von Cross-Linking-Reagenzien	22
1.4.1.2	Analyse und Identifizierung von Cross-Linking-Produkten.....	24
2	Zielstellung.....	26
3	Material und Methoden	28
3.1	Material.....	28
3.1.1	Chemikalien	28
3.1.2	Cross-Linking-Reagenzien.....	29
3.1.3	Verbrauchsmaterialien	30
3.1.4	Geräte	30
3.1.5	LC/MS-Systeme.....	31
3.1.6	Elektronische Datenverarbeitung.....	31
3.1.7	Puffer und Lösungen.....	32
3.1.8	Bakterienstämme	34
3.1.9	Proteine und Enzyme.....	34
3.1.10	Nährmedien und Zusätze für die Bakterienkultur	34
3.2	Zellbiologische Methoden	35
3.2.1	Herstellung von <i>E. coli</i> -Vorkulturen.....	35
3.2.2	Analyse des Zellwachstums.....	35
3.2.3	Stammhaltung.....	36
3.2.4	Kultivierung	36
3.3	Proteinbiochemische Methoden.....	37
3.3.1	Reinigung von Enzymen der formiatabhängigen Nitratatmung	37
3.3.2	Herstellung eines Zellwand-Subproteoms.....	39
3.3.3	2,6-Dichlorphenolindophenol (DCPIP)-Enzymaktivitätstest	39
3.3.4	Proteinkonzentrationsbestimmung mittels Bradford-Assay	39
3.3.5	Chemisches Cross-Linking	40
3.3.5.1	SDAD-Cross-Linking für die Identifizierung der Bindungs- partner von Fdh-N und Fdh-O.....	40
3.3.5.2	SDAD-Cross-Linking für die Identifizierung der Bindungs- partner von FocA	40

3.3.5.3	BS ² G-Cross-Linking für die Strukturanalyse des FocA-PfIB-Komplexes	41
3.3.6	Anreicherung von Cross-Linking-Produkten.....	41
3.3.6.1	Reinigung von FocA/Protein-Komplexen	41
3.3.6.2	Anreicherung von Fdh/Protein-Komplexen	42
3.4	Bioanalytische Methoden.....	42
3.4.1	SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)	42
3.4.2	Enzymatische Proteinspaltung	43
3.4.2.1	<i>In-Gel</i> -Proteinspaltung	43
3.4.2.2	<i>In-Lösungs</i> -Proteinspaltung	44
3.4.3	Nano-HPLC/Nano-ESI-LTQ-Orbitrap-Massenspektrometrie.....	45
3.4.4	Auswertung der MS- und MS/MS-Daten	45
3.5	Bioinformatische Methoden	46
3.5.1	Modellierung eines <i>N</i> -terminalen Bereichs von FocA mittels RosettaScripts.....	46
3.5.2	Clusteranalyse der <i>N</i> -terminal modellierten Region von FocA.....	47
3.5.3	Rekonstruktion der FocA-Pentamere	48
3.5.4	Modellierung des FocA/PfIB-Komplexes	48
4	Ergebnisse.....	50
4.1	Untersuchung des unspezifischen Seleneinbaus in Enzyme der formiatabhängigen Nitratatmung.....	50
4.1.1	Reinigung der Selenoenzyme Fdh-O und Fdh-N.....	50
4.1.2	Identifizierung von Selen-haltigen Peptiden aus Fdh-O und Fdh-N	52
4.1.3	Unspezifischer Einbau von Selenocystein in NarG und OmpA	57
4.1.4	<i>selC</i> -unabhängiger Einbau von Selenocystein in NarG	58
4.1.5	Vergleichende Quantifizierung von Modifizierungsstellen	59
4.2	Identifizierung von zellwandständigen Interaktionspartnern der Formiatdehydrogenasen Fdh-O und Fdh-N	61
4.2.1	Isolierung und Verringerung der Komplexität eines Zellwand-Subproteoms.....	63

4.2.2	Chemisches Cross-Linking	66
4.2.3	Anreicherung der Fdh/Protein-Komplexe	67
4.2.4	Identifizierung von Bindungspartnern der Fdhs	68
4.2.5	Aufklärung der Interaktionsregionen von Fdh/Protein-Komplexen	69
4.3	Charakterisierung der FocA/PfIB-Interaktion	73
4.3.1	Analyse der FocA/PfIB-Interaktion mittels chemischem Cross- Linking	73
4.3.1.1	Identifizierung der Interaktionspartner von FocA	73
4.3.1.2	Ableitung von Distanzbeschränkungen im FocA/PfIB-Komplex mittels chemischem Cross-Linking	77
4.3.2	Erstellung eines 3D-Strukturmodells des FocA/PfIB-Komplexes mittels Rosetta	81
4.3.2.1	Charakterisierung der Konformation des <i>N</i> -Terminus von FocA mittels <i>de novo</i> -Faltung	81
4.3.2.2	Rekonstruktion der cytoplasmatischen Interaktionsfläche zwischen FocA und PfIB	84
4.3.2.3	Modellierung des FocA/PfIB-Komplexes	87
5	Diskussion und Ausblick	90
5.1	Selektiver unspezifischer Einbau von Selenocystein in Proteine der formiatabhängigen Nitratatmung von <i>E. coli</i>	90
5.2	Chemisches Cross-Linking zur Identifizierung und strukturellen Charakterisierung von Protein-Protein-Interaktionen	93
5.3	Zellwandständige Interaktionspartner der Fdhs	94
5.3.1	FdnH/Skp-Komplex	94
5.3.2	FdnG/OppA-Komplex	96
5.3.3	FdnG/ErfK-Komplex	97
5.3.4	FdnGH/ToIB-Komplex	98
5.4	Charakterisierung der FocA/PfIB-Interaktion	99
Anhang	XVIII
Literaturverzeichnis	LV
Danksagung	LXXXV

Abbildungsverzeichnis

Abbildung 1: Röntgenkristallstruktur von PflB aus <i>E. coli</i>	2
Abbildung 2: Röntgenkristallstruktur von FocA aus <i>E. coli</i>	3
Abbildung 3: Röntgenkristallstruktur von FocA aus <i>S. typhimurium</i>	4
Abbildung 4: Differentielle Formiatmetabolisierung in Abhängigkeit der Wachstumsbedingungen von <i>E. coli</i>	6
Abbildung 5: Struktur der Formiatdehydrogenase N aus <i>E. coli</i> mit enthaltenen Kofaktoren.	7
Abbildung 6: Schematischer Aufbau einer Elektrospray-Ionenquelle nach Fenn	11
Abbildung 7: Elektrospray-Ionisation	12
Abbildung 8: Schematischer Aufbau des LTQ-Orbitrap XL-Hybridmassenspektrometers	13
Abbildung 9: Schematische Darstellung der segmentierten zweidimensionalen, linearen Ionenfalle (LTQ)	14
Abbildung 10: Schematische Darstellung der Orbitrap	15
Abbildung 11: Nomenklatur der Fragmentionen nach Roepstorff, Fohlmann und Biemann	16
Abbildung 12: Umwandlung von Selenocystein zu Dehydroalanin und Carbamidomethylierung von Selenocystein durch Iodacetamid	17
Abbildung 13: Schematische Darstellung des Arbeitsablaufes für die Identifizierung und Strukturaufklärung von Protein-Protein-Interaktionen mittels chemischen Cross-Linkings	21
Abbildung 14: Reaktionsmechanismus homo- und heterobifunktioneller Cross-Linker	23
Abbildung 15: SDS-PAGE-Analyse des letzten Reinigungsschrittes mittels Anionenaustausch-chromatographie (MonoQ 5/50 GL-Säule)	51
Abbildung 16: Identifizierung eines Selenocystein enthaltenden Peptids von FdoG	54
Abbildung 17: Aminosäuresequenzen der katalytischen Untereinheiten von Fdh-O und Fdh-N.	55
Abbildung 18: Röntgenkristallstruktur der katalytischen Untereinheit von Fdh-N aus <i>E. coli</i>	59

Abbildung 19: Arbeitsablauf für die Identifizierung von Interaktionspartnern der Formiatdehydrogenasen.	61
Abbildung 20: SDS-PAGE-Analyse des mittels Größenausschlusschromatographie fraktionierten Zellwand-Subproteoms.	63
Abbildung 21: Qualitative Analyse der SDAD-Modifikation der Fdhs und der chemischen Quervernetzung.	66
Abbildung 22: Anreicherung von Quervernetzungsprodukten mittels Größenausschlusschromatographie.	67
Abbildung 23: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FdnH und Skp sowie FdnG und OppA.	71
Abbildung 24: Intermolekulare SDAD-Cross-Links innerhalb der FdnH/Skp- und FdnG/OppA-Komplexe.	72
Abbildung 25: Arbeitsablauf für die Identifizierung der Interaktionspartner von FocA.	74
Abbildung 26: SDS-PAGE-Analyse von FocA/Protein-Komplexen.	75
Abbildung 27: Massenspektrometrische Identifizierung eines Peptids der PflB.	77
Abbildung 28: SDS-PAGE-Analyse von FocA/Protein-Komplexen.	78
Abbildung 29: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FocA und PflB.	80
Abbildung 30: Strukturell ungeordneter Bereich des <i>N</i> -Terminus von FocA.	82
Abbildung 31: Strukturvorhersage des <i>N</i> -Terminus (Aminosäuren 1 – 32) von FocA.	84
Abbildung 32: Energetische Minimierung rekonstruierter FocA-Pentamere.	86
Abbildung 33: BCL::Clusteranalyse.	88
Abbildung 34: Modelle des FocA/PflB-Komplexes.	89

Abbildungsverzeichnis Anhang

Abbildung A 1: Primärstruktur von FocA aus <i>E. coli</i>	XVIII
Abbildung A 2: Span-Datei für die Modellierung des <i>N</i> -terminalen Bereichs (Aminosäuren 1 - 32) von FocA.	XIX
Abbildung A 3: Flag-Datei für die Modellierung des <i>N</i> -terminalen Bereichs (Aminosäuren 1 - 32) von FocA.	XIX

Abbildung A 4:	XML-Datei für die Modellierung des <i>N</i> -terminalen Bereichs (Aminosäuren 1 - 32) von FocA.....	XXII
Abbildung A 5:	Flag-Datei für die Clusteranalyse des <i>N</i> -terminalen Bereichs (Aminosäuren 1 – 32) der α -Untereinheit von FocA.....	XXII
Abbildung A 6:	Anwendung <i>make_FocA_pentamer.py</i> für die Rekonstruktion von FocA-Pentameren.....	XXIV
Abbildung A 7:	XML-Datei für die Relaxierung rekonstruierter FocA-Pentamere.....	XXV
Abbildung A 8:	Flag-Datei für die Relaxierung rekonstruierter FocA-Pentamere.....	XXV
Abbildung A 9:	Span-Datei für die Relaxierung rekonstruierter FocA-Pentamere.....	XXVI
Abbildung A 10:	XML-Datei für die Modellierung des FocA/PfIB-Komplexes.....	XXVII
Abbildung A 11:	Flag-Datei für die Modellierung des FocA/PfIB-Komplexes.....	XXVII
Abbildung A 12:	<i>Docking_minimal_rmsd_calaculator_parallel.py</i>	XXIX
Abbildung A 13:	<i>Cluster_energies.py</i>	XXX
Abbildung A 14:	Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FdnH und Skp, FdnG und OppA, FdnG und ErfK, FdnH und TolB sowie FdnG und TolB.	XXXIV
Abbildung A 15:	Identifizierung intermolekularer Cross-Links zwischen FdnGH und TolB.....	XXXV
Abbildung A 16:	SDS-PAGE-Analyse quervernetzter FocA/Protein-Komplexe.....	XXXVI
Abbildung A 17:	Identifizierung eines Peptids von PflA.	XXXVII
Abbildung A 18:	Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FocA und PfIB.	XLII

Tabellenverzeichnis

Tabelle 1: Verwendete Bakterienstämme	34
Tabelle 2: Verwendete Medienzusätze	35
Tabelle 3: Zusammensetzung der Polyacrylamidgele.....	43
Tabelle 4: Massenspektrometrische Analyse der Proteinreinigung.....	52
Tabelle 5: Identifizierte Selenocysteine der Fdh-O und Fdh-N.	56
Tabelle 6: Identifizierte Selenocysteine in Nar und OmpA.	57
Tabelle 7: Identifizierte Selenocysteine in NarG, die aus einer <i>E. coli selC</i> - Mutante isoliert wurden.	58
Tabelle 8: Vergleichende Quantifizierung von modifizierten und nicht modifizierten Peptiden.....	60
Tabelle 9: Klassifizierung identifizierter Proteine des Zellwand-Subproteoms.	64
Tabelle 10: Hochabundante Proteine des Zellwand-Subproteoms.	65
Tabelle 11: Identifizierte Bindungspartner der Fdhs.....	69
Tabelle 12: Identifizierte intermolekulare Vernetzungsprodukte zwischen Fdhs und Bindungspartnern.....	70
Tabelle 13: Auswahl an identifizierten Bindungspartnern von FocA.	76
Tabelle 14: Identifizierte intermolekulare Cross-Linking-Produkte des FocA/PfIB-Komplexes.	79
Tabelle 15: Clusteranalyse.	83
Tabelle 16: BCL::Clusteranalyse.....	87

Tabellenverzeichnis Anhang

Tabelle A 1: Zusammenfassung identifizierter Proteine des Zellwand- Subproteoms.....	XLIII
Tabelle A 2: Identifizierte Bindungspartner der Fdhs.....	LI
Tabelle A 3: Zusammenfassung der potentiellen FocA-Bindungspartner.....	LII

Abkürzungsverzeichnis

Die verwendeten Drei- und Ein-Buchstaben-Codes der proteinogenen Aminosäuren entsprechen den Vorschlägen der IUPAC-IUB-Kommission für biochemische Nomenklatur.

a	Dehydroalanin
ACN	Acetonitril
AGC	<i>Automatic gain control</i>
ArcA	<i>Anoxic redox control</i>
API	<i>Atmospheric pressure ionization</i>
APS	Ammoniumpersulfat
Å	Ångström
bp	Basenpaare
BS ² G	Bis(sulfosuccinimidyl) glutarat
BSA	Bovines Serumalbumin
c	Carbamidomethyliertes Cystein
CID	<i>Collision induced dissociation</i>
CoA	Coenzyme A
Co-IP	Co-Immunopräzipitation
CRM	<i>Charged residue model</i>
C-Trap	<i>Curved ion trap</i>
Da	Dalton
DC	Direct current
DDM	<i>n</i> -Dodecyl- β -D-maltosid
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DOC	Desoxycholat
DTT	Dithiothreitol
EDTA	Ethylendiamintetraessigsäure
EIC	<i>Extracted ion chromatogram</i>
ELISA	Enzyme-linked immunosorbent assay
ESI	Elektrospray-Ionisation
FA	Ameisensäure
Fdh-H	Formiatdehydrogenase H

Fdh-N	Formiatdehydrogenase N
Fdh-O	Formiatdehydrogenase O
Fdhs	Formiatdehydrogenasen
FhIA	Formiat-sensierendem Transkriptionsfaktor
FHL	Formiat-Hydrogen-Lyase
FocA	<i>Formate channel A</i>
FFT	<i>Fast Fourier-Transformation</i>
Fnr	<i>Fumarat nitrate regulator</i>
FNT	Formiat-Nitrit-Transporter
FRET	Förster-Resonanz-Energietransfer
FT	Fouriertransformation
HCD	<i>Higher energy collisional dissociation</i>
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
HPLC	<i>High performance liquid chromatography</i>
Hyd-3	Hydrogenase 3
ICAT	<i>Isotope-coded affinity tag</i>
ICPL	<i>Isotope-coded protein labels</i>
iTRAQ	<i>Isotope tags for relative and absolute quantification</i>
IEM	<i>Ion evaporation model</i>
K _d	Gleichgewichtsdissoziationskonstante
K _m	Michaelis-Menten-Konstante
LC	<i>Liquid chromatography</i>
LTQ	<i>Linear trap quadrupole</i> (Thermo Fisher Scientific)
MGD	Molybdopterin-Guanin-Dinukleotid
MilliQ-H ₂ O	MilliQ-Wasser
ModE	Molybdat-sensierender Transkriptionsfaktor
MOPS	3-(N-Morpholino)Propansulfonsäure
MS	Massenspektrometrie
MS/MS	Tandem-Massenspektrometrie
<i>m/z</i>	Masse-zu-Ladungsverhältnis
NAD	Nicotinamidadenindinukleotid
Nar	Nitratreduktase
NHS	<i>N</i> -Hydroxysuccinimid
NMR	<i>Nuclear magnetic resonance</i>

Pa	Pascal
PAGE	Polyacrylamid-Gelelektrophorese
PBS	<i>Phosphate buffered saline</i>
PDB	Proteindatenbank
PDH	Pyruvatdehydrogenase
PflA	Pyruvat-Formiat-Lyase-aktivierendes Enzym
PflB	Pyruvat-Formiat-Lyase
PSM	<i>Peptide spectral match</i>
PTM	Posttranslationale Modifikation
pH	Dekadischer Logarithmus des Zahlenwertes der Oxoniumionenkonzentration
pKs	Säurekonstante
PPI	Protein-Protein-Interaktion
RMSD	<i>root mean square deviation</i>
SDAD	Succinimidyl-2-([4,4'-azipentanamido]ethyl)-1,3'-dithiopropionat
SDS	Natriumdodecylsulfat
Sec	Selenocystein
SEV	Sekundärelektronenvervielfacher
SILAC	<i>stable isotope labeling by amino acids in cell culture</i>
Sulfo-NHS	<i>N</i> -Hydroxysulfosuccinimid
Tat	<i>Twin-arginine translocation pathway</i>
TFA	Trifluoressigsäure
TFE	Trifluorethanol
TIC	<i>Total ion current</i>
TMTs	<i>Tandem mass tags</i>
TRIS	Tris(hydroxymethyl)aminomethan
U	Selenocystein
u	Carbamidomethyliertes Selenocystein
UAS	<i>Upstream activating sequence</i>
UE	Untereinheit
URS	<i>Upstream regulatory sequence</i>
v/v	<i>Volume per volume</i> (Volumen/Volumen)
w/v	<i>Weight per volume</i> (Masse/Volumen)

Zusammenfassung

Im Rahmen dieser Doktorarbeit erfolgte die Identifizierung und strukturelle Charakterisierung von Protein-Protein-Interaktionen des Formiatstoffwechsels in *Escherichia coli* mittels Massenspektrometrie (MS)-basierter Methoden. Im ersten Abschnitt wurde der Einfluss eines Überangebots an Selen während des Zellwachstums auf dessen Einbau in Proteine von *E. coli* untersucht. Die Formiatdehydrogenasen Fdh-O, Fdh-N und Fdh-H sind die einzigen Proteine in *E. coli*, welche Selenocystein durch die Dekodierung des Kodons UGA an jeweils einer spezifischen Position einbauen. Ein Überangebot an Selen kann durch den unspezifischen Einbau von Selenocystein in Proteine toxisch wirken. Die gemeinsame Reinigung der homologen Enzyme Fdh-N und Fdh-O und die Reinigung der Nitratreduktase (Nar) erfolgte aus dem *E. coli*-Wildtypstamm MC4100 nach anaerober Kultivierung in Gegenwart von 2 μM bzw. 100 μM Selenit. Die MS-basierte Analyse der katalytischen Untereinheiten beider Fdhs führte zur Identifizierung des spezifischen UGA-kodierten Selenocysteins und zusätzlicher „unspezifischer“ Selenocysteine, die bestimmte Cysteine ersetzen. Der Einbau weiterer Selenocysteine erfolgte trotz einer gewissen Variabilität nicht zufällig und war unabhängig von der verwendeten Selenitkonzentration. Die MS-Analysen zeigten keinen Einbau von Selenomethionin. Die Untersuchung von Nar aus einer *E. coli*-*seI*C-Mutante wies ein ähnliches Einbaumuster an Selenocystein wie im *E. coli*-Wildtypstamm MC4100 auf. Dies deutet darauf hin, dass der unspezifische Einbau von Selenocystein unabhängig von der spezifischen Selenocystein-einbaumaschinerie abläuft. Während der Biosynthese von Cystein kann Sulfid durch Selenid ersetzt werden. Eine anschließende Acetylierung der tRNA^{Cys} mit Selenocystein führt zur Dekodierung der Cystein-spezifischen Kodone UGU und UGC mit Selenocystein. Die Ergebnisse lassen vermuten, dass ein gewisser Grad eines Fehleinbaus von Selenocystein in Enzyme durch einen Austausch bestimmter nicht-essentieller Cysteine toleriert wird. Dies könnte ein Puffersystem für die Bewältigung hoher intrazellulärer Selenkonzentrationen darstellen.

Fdh-N und Fdh-O sind mit der inneren Cytoplasmamembran assoziiert, wobei ihre aktiven Zentren dem Periplasma zugewandt sind. Fdh-N erstreckt sich 90 Å in das periplasmatische Kompartiment, welches in *E. coli* eine Distanz von 100-150 Å zwischen innerer und äußerer Membran aufweist.

Ziel des zweiten Teils dieser Arbeit war die Aufklärung von Interaktionen der Fdh-N und Fdh-O mit Proteinen innerhalb des periplasmatischen Raumes, welche eine Stabilisierung der respiratorischen Enzyme nach Reifung und Translokation über die Cytoplasmamembran gewährleisten. Hierfür wurde eine chemische Quervernetzung unter Verwendung eines heterobifunktionellen amin/photo-reaktiven Reagens und eine anschließende MS-Analyse der Quervernetzungsprodukte durchgeführt. Dies führte zur Identifizierung von sechs potentiellen Bindungspartnern. Für vier ausgewählte Fdh-Bindungspartner Chaperon Skp, L, D-Transpeptidase ErfK, Oligopeptid-Bindeprotein OppA, und dem Chaperon TolB erfolgte die Aufklärung der Grenzflächenregionen der Fdh/Protein-Komplexe. Dadurch wurden detaillierte strukturelle und funktionelle Einblicke in die Mechanismen gewonnen, die eine Nachreifung der komplexen, aus mehreren Untereinheiten bestehenden Enzyme, Fdh-N und Fdh-O im Periplasma von *E. coli* unterstützen.

Der dritte Teil dieser Arbeit behandelt die strukturelle Charakterisierung des FocA/PfIB-Komplexes und dessen Einfluss auf die Regulation des Formiattransportes. Die Formiat-Nitrat-Transporter (FNT) bilden eine Superfamilie pentamerer Membrankanäle, die einwertige Anionen über biologische Membranen translozieren. FocA erlaubt die bidirektionale Translokation von Formiat über die Cytoplasmamembran. Die zugrunde liegenden Kontrollmechanismen der Translokation und Substratspezifität sind jedoch weitestgehend unbekannt. Zur Identifizierung potentieller Bindungspartner von FocA wurde ein Quervernetzungs-experiment auf Grundlage des amin/photo-reaktiven Reagenses durchgeführt, welches bereits für die Interaktionsstudien der Fdhs verwendet wurde. Die Experimente bestätigten die Ausbildung eines Komplexes zwischen dem pentameren FocA-Kanal und dimerem PfIB. Für den Erhalt detaillierter 3D-Strukturinformationen des FocA/PfIB-Komplexes wurde ein weiteres chemisches Quervernetzungs-experiment unter Verwendung eines aminreaktiven homobifunktionellen Reagenses durchgeführt. Im Anschluss daran erfolgte die Auftrennung des Reaktionsgemisches durch SDS-PAGE, die enzymatische Spaltung der Proteine und die Analyse der erhaltenen Peptidmischungen durch LC/MS/MS. Dies führte zur Identifizierung von 14 einzigartigen intermolekularen Quervernetzungsprodukten zwischen der cytoplasmatische Seite von FocA und PfIB, wodurch die spezifische Interaktion beider Bindungspartner bestätigt wurde. Die Tatsache, dass 11 von 14 Quervernetzungsprodukten Modifikationen N-terminaler Aminosäuren in FocA zeigten, deutet darauf

hin, dass der *N*-Terminus von FocA eine hohe Flexibilität besitzt, welche es ihm erlaubt, mit verschiedenen Regionen von PfIB zu interagieren. Zur Aufklärung der Bindungsorientierung von PfIB und FocA erfolgte die Erstellung von Modellen des Komplexes durch ein bioinformatisch-gestütztes Protein-Protein-*Docking* beider Interaktionspartner mit Hilfe des Software-Pakets Rosetta. Ca. 50.000 generierte Modelle wurden auf ihre strukturelle Übereinstimmung mit experimentell erstellten Distanzbeschränkungen überprüft. Die Mehrheit der Distanzbeschränkungen erfüllte die C_{α} - C_{α} -Ausschlußgrenze von 25 Å. Diese Ergebnisse zeigen die spezifische Interaktion von PfIB mit FocA und lassen vermuten, dass PfIB über die Interaktion mit FocA die Translokation des Substrates kontrolliert und möglicherweise die Richtung der Substratbewegung durch die einzelnen Poren der FocA-Monomere bestimmt. Die Kombination aus chemischer Quervernetzung von Protein-Protein-Interaktionen mit einer MS-Analyse der Quervernetzungsprodukte und einer bioinformatischen Modellierung von Proteinkomplexen durch Rosetta stellt somit einen vielversprechenden Ansatz zur Identifizierung und strukturellen Charakterisierung großer heterogener Membranproteinkomplexe dar.

Summary

In this thesis, protein-protein-interactions in the formate metabolism of *Escherichia coli* were identified and structurally characterized by mass spectrometry (MS)-based methods. In the first part of this work, the influence of excessive selenium during cell growth on the incorporation into *E. coli* proteins was investigated. The formate dehydrogenases Fdh-O, Fdh-N, and Fdh-H, are the only proteins in *E. coli* that incorporate selenocysteine at a specific position by decoding a UGA codon. However, an excess of selenium can lead to toxicity through misincorporation of selenocysteine into proteins. The highly homologous enzymes Fdh-N and Fdh-O, were co-purified along with nitrate reductase (Nar) from the wild type *E. coli* strain MC4100 after anaerobic growth with nitrate in the presence of 2 μM or 100 μM selenite. MS-based analysis of the catalytic subunits of both Fdhs identified the UGA-specified selenocysteine residue and revealed the incorporation of additional, 'non-specific' selenocysteinyI residues, which always replaced particular cysteinyI residues. Although variable, their incorporation was not random and was independent of the selenite concentration used. Selenomethionine was never observed in our analyses. Nar isolated from an *E. coli* *selC* mutant showed a similar selenocysteine incorporation pattern as in the wild-type strain indicating that non-specific selenocysteine incorporation is independent of the specific selenocysteine pathway. Thus, selenide replaces sulfide in the biosynthesis of cysteine and misacylated selenocysteyl-tRNA^{Cys} decodes either UGU or UGC codons, which usually specify cysteine. Nevertheless, not every UGU or UGC codon was decoded as selenocysteine. Together, the results suggest that a certain degree of misincorporation of selenocysteines into enzymes through replacement of particular, non-essential cysteines, is tolerated. This might act as a buffering system to cope with excessive intracellular selenium.

Fdh-N and Fdh-O are associated with the inner membrane and have their respective active site located towards the periplasm. Fdh-N extends 90 Å into the periplasmic compartment, which in *E. coli* ranges between 100-150 Å from the inner to the outer membrane leaflet. The aim of this second part of this thesis was to shed light on the interactions of Fdh-N and Fdh-O in the periplasmic space that might be involved in stabilizing these respiratory enzymes after maturation and translocation across the cytoplasmic membrane. For this purpose, chemical cross-linking using a heterobifunctional amine/photo-reactive reagent followed by MS analysis of the

cross-linked products was performed. This resulted in the identification of six potential binding partners. For four selected Fdh-binding partners, namely the chaperone Skp, the L, D-transpeptidase ErfK, the oligopeptide binding protein OppA, and the chaperone TolB, interface regions within the Fdh/protein complexes were mapped. Detailed structural and functional insights into the mechanisms that support the post-maturation of the multi-subunit enzymes Fdh-N and Fdh-O in the periplasm of *E. coli* were obtained.

The third part of this work deals with the structural characterization of the FocA/PfIB complex and its regulatory effect on formate transport. The formate-nitrate transporters (FNT) form a superfamily of pentameric membrane channels that translocate monovalent anions across biological membranes. FocA translocates formate bidirectionally, but the mechanisms underlying the control of translocation and substrate specificity remain unclear. In order to identify potential binding partners of FocA, an initial cross-linking experiment was performed using the same cross-linker as for the interaction studies of Fdhs. These experiments confirmed the formation of a complex between the pentameric channel FocA and the dimeric PfIB. Chemical cross-linking with a homobifunctional cross-linker was performed for deriving detailed 3D-structural data of the FocA/PfIB-complex. After cross-linking, reaction mixtures were separated by SDS-PAGE, enzymatically digested, and the resulting peptide mixtures were analyzed by LC/MS/MS. Fourteen unique intermolecular cross-links were identified between the cytoplasmic face of FocA and PfIB, indicating a specific interaction of both proteins. The fact that 11 out of 14 cross-links are with amino acids located in the *N*-terminus of FocA suggests a high flexibility of FocA's *N*-terminus, allowing it to react with diverse regions of PfIB. In order to determine a binding orientation of PfIB to FocA, models of the complexes were generated by computational docking of PfIB to FocA using the software suite Rosetta. Approximately 50,000 models were evaluated for their agreement with the experimentally derived cross-linking restraints. The majority of the cross-links identified were satisfied below a C_α-C_α cutoff distance of 25 Å. The results show that PfIB interacts specifically with FocA and suggest that PfIB controls whether the substrate is translocated and possibly also the directionality of the substrate movement through the pore of each FocA monomer.

In conclusion, chemical cross-linking and MS in combination with computational modeling using Rosetta proved to be a highly promising approach for the

identification and structural characterization of multisubunit membrane protein complexes.

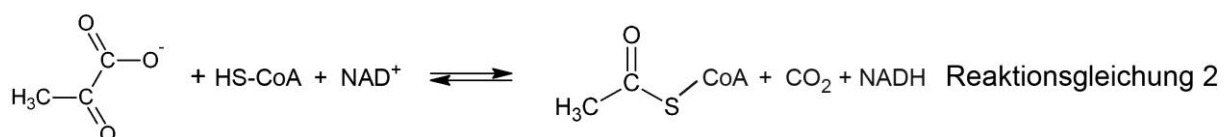
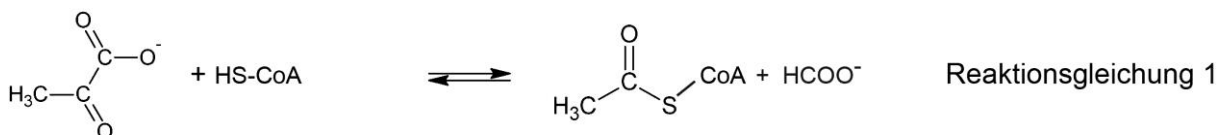
1 Einleitung

1.1 Formiatmetabolismus in *Escherichia coli*

Formiat ist ein Schlüsselmetabolit im anaeroben Energiestoffwechsel von *Escherichia coli* und anderen Enterobakterien. Die Synthese und Verstoffwechslung von Formiat ist in ein fein abgestimmtes Netzwerk metabolischer Stoffwechselwege integriert.

1.1.1 Bildung von Formiat

Das Glycylradikalenzym Pyruvat-Formiat-Lyase (PflB) ist mit einem Anteil von bis zu 3% der löslichen Proteine ein unter anaeroben Wachstumsbedingungen hochabundantes Enzym [1]. PflB katalysiert die Synthese von Formiat und Acetyl-Coenzym A (CoA) aus Pyruvat und CoA [2, 3] (Reaktionsgleichung 1). Die Aktivierung des Enzyms erfolgt posttranslational ausschließlich unter anaeroben oder Sauerstoff-limitierenden Wachstumsbedingungen und unterliegt einer strengen metabolischen Kontrolle [4]. Zur Aktivierung wird durch das PflB-aktivierende Enzym (PflA) ein Glycylradikal in das PflB-Apoenzym eingeführt [1, 2, 5], wobei die Aktivität von PflB auf eine Untereinheit des Homodimers (Abbildung 1) beschränkt ist [2, 6-8]. PflB ist unter anaeroben Bedingungen das analoge Enzym zur respiratorischen Pyruvatdehydrogenase (PDH), die Pyruvat oxidativ decarboxyliert [9]. PDH katalysiert die heterolytische Spaltung von Pyruvat, wobei Reduktionsäquivalente auf NAD^+ übertragen werden (Reaktionsgleichung 2). Die unter nicht-oxidativen Bedingungen katalysierte homolytische Spaltung von Pyruvat durch PflB führt dagegen zum Verbleib der Reduktionsäquivalente auf CO_2 , wodurch das Redoxgleichgewicht trotz des Fehlens externer Elektronenakzeptoren erhalten bleibt.



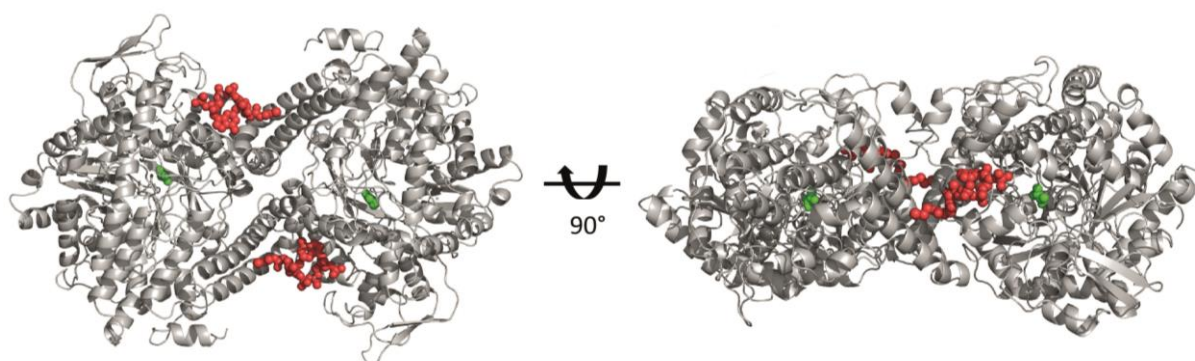


Abbildung 1: Röntgenkristallstruktur von PflB aus *E. coli*. Die Bänderstruktur zeigt das Homodimer PflB (PDB-Eintrag: 1H17) im Komplex mit Coenzym A (rote Sphären) und dem isosteren Pyruvatanalogen Oxalat (grüne Sphären) [8].

Unter fermentativen Stoffwechselbedingungen wird durch PflB bis zu einem Drittel des Kohlenstoffs aus Glucose zu Formiat umgewandelt [10, 11]. Formiat hat einen pK_s -Wert von 3,75 und liegt bei neutralen pH-Werten als nicht-membrangängiges Anion vor. Eine durch PflB-Aktivität induzierte cytoplasmatische Anreicherung von Formiat würde zur Absenkung des intrazellulären pH-Wertes und zur Entkopplung des Protonengradienten führen [10], weshalb Formiat abhängig von den physiologischen Bedingungen weiter metabolisiert oder sekretiert wird [12-14].

1.1.2 Transport von Formiat über die Cytoplasmamembran

Die Identifizierung des Formiatkanals (*formate channel*, FocA) erfolgte durch eine *E. coli*-Mutante mit einer Deletion im *focA*-Gen, die einen um 50% reduzierten Export sowie eine cytoplasmatische Anreicherung von Formiat zeigte [12]. Die Studie bestätigte zudem, dass unter anaeroben Bedingungen Formiat durch FocA importiert wird.

Die Gene, die für FocA und PflB kodieren, sind gemeinsam im *focA-PflB*-Operon lokalisiert [15, 16]. Die Transkription dieses bicistronischen Operons ist komplex [15, 17, 18] und wird unter anaeroben Bedingungen durch eine ca. 450 bp stromaufwärts gelegene regulatorische Sequenz (*upstream regulatory sequence*, URS) kontrolliert, welche die drei Promotoren 6, 6a und 7 enthält. Die Promotoren 6 und 7 sind von den Transkriptionsfaktoren Fnr (*fumarat nitrate regulator*) bzw. Fnr in Kombination mit ArcA (*anoxic redox control*) abhängig [19-21]. Die Transkription von Promotor 6a ist schwach konstitutiv [16, 22]. Stromaufwärts des *focA-pflB*-Operons liegt das für

PflA kodierende Gen *pflA* [15, 23], dessen Co-Transkription durch einen starken Terminator verhindert wird [24].

FocA gehört zur Familie der Formiat-Nitrit-Transporter (FNT), einer funktionell diversen Klasse von Membranproteinen [12, 25-28]. Diese sind unter den Archaeobakterien, Eubakterien und bestimmten Pilzen verbreitet und haben unterschiedliche zelluläre Funktionen [28-30]. Charakteristisch für FNT-Kanäle ist die Translokation kleiner einwertiger Anionen, wie z. B. Formiat, Nitrit oder Hydrosulfid, durch die bakterielle Cytoplasmamembran [25-28, 31].

Die Röntgenkristallstruktur von FocA aus *E. coli* zeigt ein zylindrisches Homopentamer mit einer Höhe von 55 Å sowie einem Durchmesser von 80 Å (Abbildung 2) [25]. Für eine optimale Kristallpackung und Röntgenbeugung erfolgte eine N-terminale Verkürzung um 21 Aminosäuren. Jedes Monomer besteht aus einer asymmetrischen Anordnung von sechs α -helikalen Transmembransegmenten (TM1-6), die eine axiale Passage für einen selektiven Transport von Formiat ausbilden (Abbildung 2). Die N- und C-Termini sowie die TM2- und TM3-verbindenden Ω -Schleifen sind auf der cytoplasmatischen Seite lokalisiert. Zusätzlich besitzt FocA eine zentrale Pore (Abbildung 2), welche einen minimalen Durchmesser von 3,8 Å aufweist und sowohl polare als auch unpolare Bereiche enthält. Eine Funktion dieser zentralen Pore für die Translokation von Molekülen kann nicht ausgeschlossen werden [25].

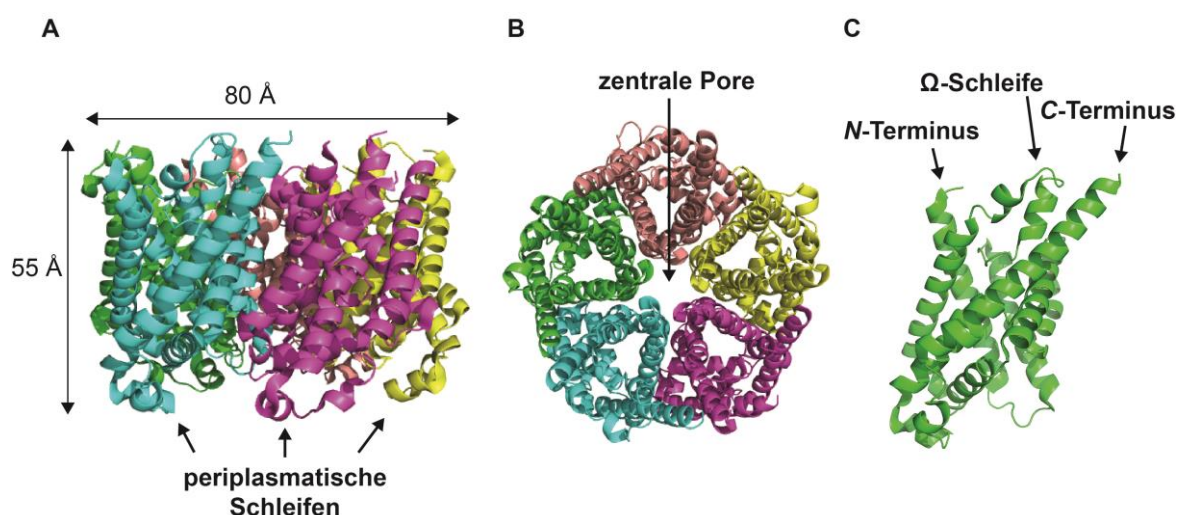


Abbildung 2: Röntgenkristallstruktur von FocA aus *E. coli*. Gezeigt ist die Bänderstruktur von FocA aus *E. coli* (PDB-Eintrag: 3KCU) [25]. (A) FocA-Pentamer mit Blick parallel zur Membran; (B) cytoplasmatische Ansicht des FocA-Pentamers; (C) FocA-Monomer (α -Untereinheit).

Die Struktur von FocA aus *Vibrio cholerae* deutet auf einen Zusammenhang zwischen Flexibilität der Ω -Schleife und dem substratvermittelten Transport von Formiat hin [26]. In einer „oberen“ Position der Ω -Schleife bildet Thr90 eine Wasserstoffbrückenbindung mit His208 der S-Schleife im selektiven Filter des Formiatkanals aus, wodurch der Eintritt von Formiat verhindert wird [26]. Formiat konkurriert mit Thr90 um die Wasserstoffbrückenbindung zu His208. In einer „unteren“ Position der Ω -Schleife nimmt Formiat die Position von Thr90 ein, wodurch es zur Öffnung des selektiven Filters im Formiatkanal kommt.

Strukturelle und elektrophysiologische Untersuchungen von FocA aus *Salmonella typhimurium* führten zum Postulat einer pH-regulierten Formiattranslokation [27]. Die Röntgenkristallstruktur eines bei pH 4 erhaltenen Proteinkristalls zeigt erstmals drei verschiedene Konformationen des N-Terminus innerhalb eines FocA-Pentamers (Abbildung 3). Diese werden entsprechend ihrer morphologischen Eigenschaft, die cytoplasmatische Seite der Substratkanäle zu blockieren, einzuengen oder zu öffnen als geschlossene, intermediäre bzw. offene Konformation differenziert. Der Konformationstypus benachbarter Monomere eines FocA-Pentamers ist dabei genau definiert (Abbildung 3). Die geordneten α -helikalen N-Termini intermediärer und geschlossener Untereinheiten interagieren zur gegenseitigen Stabilisierung ihrer Konformationen miteinander (Abbildung 3). Der N-Terminus der offenen Konformation repräsentiert eine ungeordnete Struktur, welche die cytoplasmatische

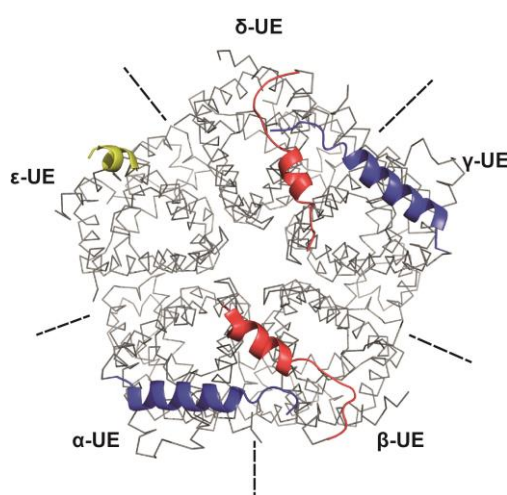


Abbildung 3: Röntgenkristallstruktur von FocA aus *S. typhimurium*. FocA wurde bei pH 4 kristallisiert. Die Röntgenkristallstruktur zeigt die verschiedenen Konformationen des N-Terminus innerhalb eines FocA-Pentamers (PDB-Eintrag: 3Q7K) [27]: Gezeigt ist die cytoplasmatische Seite von FocA. Die intermediäre (blau) und geschlossene Konformation (rot) der jeweiligen Untereinheiten (UE) sind als farbige Bänderstrukturen hervorgehoben. Die ungeordnete N-terminale Konformation des offenen Stadiums (gelb) kann nicht vollständig dargestellt werden. Das Proteinrückgrat der Gesamtstruktur von FocA ist als *ribbon* (grau) dargestellt.

Seite des Substratkanals nicht blockiert. Es wird vermutet, dass eine pH-regulierte kooperative Konformationsänderung der *N*-Termini einen kontrollierten Ex- bzw. Import von Formiat ermöglicht [27].

1.1.3 Differentielle Regulation der Formiatdehydrogenasen in *E. coli*

Das Genom von *E. coli* kodiert drei cytoplasmamembranständige Formiatdehydrogenasen (Fdhs), die unter verschiedenen Wachstumsbedingungen synthetisiert werden (Abbildung 4) [11, 14, 32]. Es handelt sich dabei um die Selenhaltigen Isoenzyme Fdh-H, Fdh-N und Fdh-O, die einen Molybdän-bis-Molybdopterin-Guanin-Dinukleotid (Mo-bis-MGD)-Kofaktor zur Oxidation von Formiat koordinieren [32].

Durch PflB produziertes Formiat wird zunächst über FocA aus fermentierenden *E. coli*-Zellen exportiert, wodurch eine cytoplasmatische Ansäuerung verhindert wird [12] (Abbildung 4). Sinkt der extrazelluläre pH-Wert unter 6,8, wird Formiat zurück importiert und über die Fdh-H metabolisiert, die eine Komponente des cytoplasmatischen Formiat-Hydrogen-Lyase (FHL)-Komplexes ist [10-12]. Die Transkription der Gene des FHL-Komplexes ist abhängig von einem saueren pH-Wert, einem Sigma-54-kontrollierten Promotor, einem Molybdät-sensierenden Transkriptionsfaktor (ModE) und dem Formiat-sensierenden Transkriptionsfaktor (FhIA) [33-35]. Die Synthese des FHL-Komplexes wird durch die Präsenz externer Elektronenakzeptoren, wie z. B. Sauerstoff oder Nitrat, unterdrückt [4, 35].

In Anwesenheit dieser erfolgt die Oxidation von Formiat zu CO₂ und H⁺ durch die periplasmatisch orientierte und energiekonservierende Fdh-O bzw. Fdh-N [32] zur Erzeugung einer protonenmotorischen Kraft [36, 37]. Die Transkription der Gene *fdnGHI* zur Synthese von Fdh-N erfolgt nur unter anaeroben Bedingungen in der Anwesenheit von Nitrat als Elektronenakzeptor [38]. Die Transkription der Gene *fdoGHI* zur Synthese von Fdh-O erfolgt sowohl unter aeroben [39, 40] als auch anaeroben Bedingungen mit Nitrat [39, 41, 42]. Die hohe Affinität der Fdh-N für Formiat (K_m-Wert: 0,12 mM) [36] ermöglicht die Aufrechterhaltung einer niedrigen Konzentration des Metaboliten. Dadurch wird die zur Aktivierung des Transkriptionsfaktors FhIA (K_m-Wert: 5 mM) [43] nötige Formiatkonzentration nicht erreicht und die Synthese des FHL-Komplexes verhindert [35] (Abbildung 4).

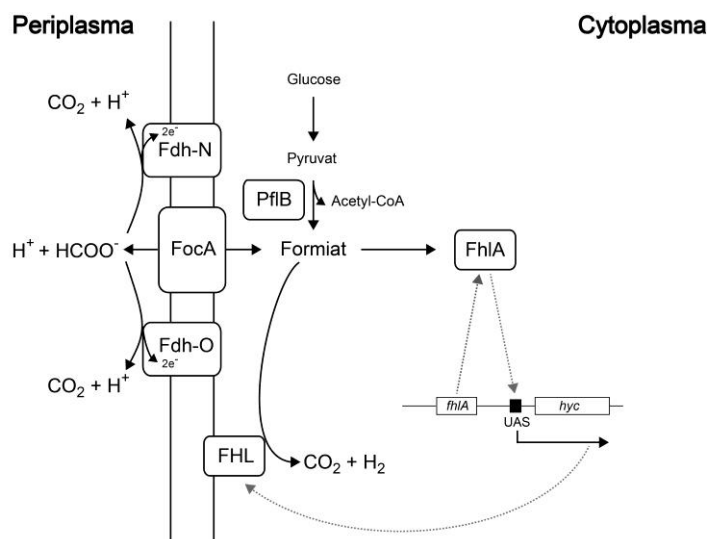


Abbildung 4: Differentielle Formiatmetabolisierung in Abhängigkeit der Wachstumsbedingungen von *E. coli*. Entnommen und verändert nach [44]. Unter fermentativen Bedingungen wird Formiat vom cytoplasmatischen Fhl-Komplex zu CO_2 und H_2 metabolisiert. Die Induktion der Komponenten des FHL-Komplexes ist u. a. abhängig vom Formiat-sensierenden Transkriptionsaktivator FhlA und einem sauren pH-Wert. In Anwesenheit der externen Elektronenakzeptoren Nitrat oder Sauerstoff wird Formiat durch die periplasmatische und energiekonservierende Fdh-N bzw. Fdh-O zu CO_2 oxidiert, wobei Protonen und Elektronen entstehen und eine protonenmotorische Kraft erzeugt wird. FocA ermöglicht den Ex- bzw. Import von Formiat in Abhängigkeit der Wachstumsbedingungen. UAS: *upstream activating sequence*; Die *hyc*-Gene kodieren die Komponenten der Hydrogenase 3, welche zusammen mit Fdh-H den FHL-Komplex bildet.

1.1.4 Strukturen der respirativen Formiatdehydrogenasen Fdh-O und Fdh-N

Die jeweiligen Untereinheiten der Isoenzyme Fdh-N und Fdh-O haben eine hohe Ähnlichkeit ihrer Primärstrukturen, wobei die α -Untereinheiten (FdnG/FdoG) eine Aminosäuresequenzidentität von 76% und die β -Untereinheiten (FdnH/FdoH) von 75% besitzen [45]. Beide Enzyme zeigen eine identische Stöchiometrie ihrer Untereinheiten ($\alpha_3\beta_3\gamma_3$ -Struktur, Abbildung 5). Die α -Untereinheit (FdnG/FdoG) enthält das katalytisch aktive Selenocystein, ein Eisenschwefel [4Fe-4S]-Zentrum sowie einen Mo-bis-MGD-Kofaktor [36-38]. Die elektronenübertragende β -Untereinheit (FdnH/FdoH) enthält vier [4Fe-4S]-Zentren. Die γ -Untereinheit (FdnI/FdoI) enthält zwei Häm *b*-Gruppen und bildet eine Bindestelle für Menachinon [37]. Als integrales Membranprotein ist die γ -Untereinheit der Membrananker für die Stabilisierung der beiden periplasmatischen membran-assoziierten Untereinheiten (G und H).

Zur Reifung der komplexen Enzyme werden die α - und β -Untereinheit zusammen als Kofaktoren enthaltendes Heterodimer (FdnGH, FdoGH) durch das Twin-Arginin-Translokations (Tat)-System über die Cytoplasmamembran transportiert [46]. Dieses

vermittelt den Transport von vollständig gefalteten Proteinen [47, 48], wofür eine charakteristische *N*-terminale Signalsequenz der Substrate notwendig ist. Diese besteht aus einer polaren *N*-terminalen Region (n-Region) variabler Länge, gefolgt von einer relativ hydrophoben Region (h-Region), die eine Sequenz von 15-25 Aminosäuren umfasst. Die n- und h-Regionen flankieren die konservierte Konsensussequenz SRRxFLK, die das charakteristische Argininpaar enthält [49, 50]. Häufig dient eine positiv geladene Region (c-Region) innerhalb der h-Region als Erkennungsmotiv für die Abspaltung der Signalsequenz während des Substrattransports [49]. Der Transport der heterodimeren Substrate FdnGH und FdoGH wird durch das *N*-terminale Signalpeptid der katalytischen Untereinheiten (FdnG/FdoG) vermittelt. Die β -Untereinheit (FdnH/FdoH), welche kein eigenes Signalpeptid hat, wird vermutlich zusammen mit der α -Untereinheit (FdnG/FdoG) co-exportiert [46]. Nach dem Transport des Heterodimers (FdnGH, FdoGH) über die Cytoplasmamembran erfolgt die Assemblierung mit der integralen Membranuntereinheit (FdnI, FdoI).

Die Struktur der vollständig assemblierten Fdh-N zeigt einen pilzähnlichen Multiproteinkomplex mit einer Molekularmasse von 510 kDa [37]. Das Enzym bildet

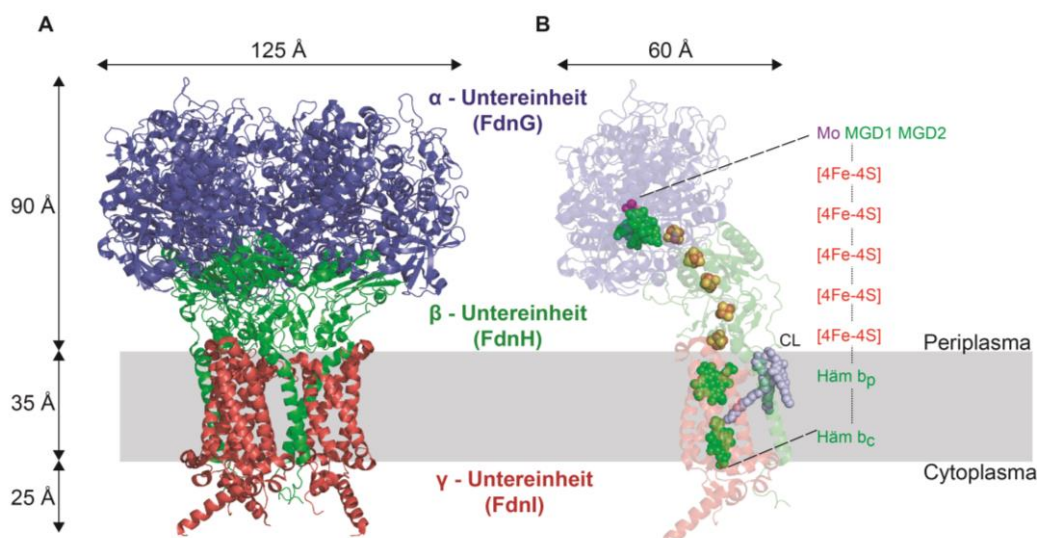
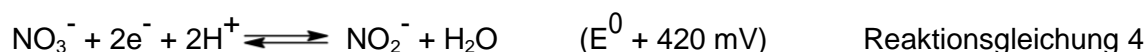
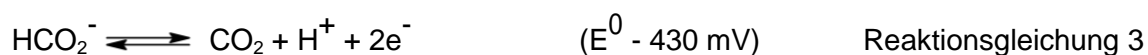


Abbildung 5: Struktur der Formiatdehydrogenase N aus *E. coli* mit enthaltenen Kofaktoren. Dargestellt sind die Bänderstrukturen (PDB-Eintrag: 1KQF) [37] (A) der vollständigen Fdh-N sowie (B) eines Heterotrimers in der Ansicht parallel zur Membran. Das Heterotrimer enthält einen Mo-bis-MGD-Kofaktor, fünf [4Fe-4S]-Zentren und zwei Häm *b*-Gruppen (Häm-*b_p*, Häm-*b_c*: periplasmatisch und cytoplasmatisch orientiertes Häm *b*) und eine Menachinonbindestelle (nicht gezeigt), die eine direkte Weiterleitung der Elektronen von Formiat auf Menachinon ermöglichen. Molybdän (violett), MGD-Kofaktoren und Häm *b*-Gruppen (grün), Eisen (rot), Schwefel (gelb) und Cardiolipin (hellblau) sind als Sphären hervorgehoben.

ein Trimer aus Heterotrimeren, das sich 90 Å in das Periplasma erstreckt und einen Durchmesser von 125 Å hat. Die strickleiterartig angeordneten Kofaktoren erlauben die Aufnahme der Elektronen von Formiat über den Mo-bis-MGD-Kofaktor und eine effiziente Weiterleitung zur Menachinonbindestelle über fünf [4Fe-4S]-Zentren und zwei Häm-Gruppen. Die strukturelle Charakterisierung der Fdh-N und der Nitratreduktase A (Nar) [51] unterstützen das chemiosmotische Modell der Energiekopplung nach Peter Mitchell [52, 53]. Fdh-N und die diametral membranassoziierte Nar bilden während der Nitratatmung eine Redoxschleife zur Erzeugung einer protonenmotorischen Kraft zur Energiekonservierung [36, 37, 51, 53-55]. Die periplasmatische Oxidation von Formiat (Reaktionsgleichung 3) und die cytoplasmatische Reduktion von Nitrat (Reaktionsgleichung 4) sowie ein durch Menachinon vermittelter Elektronentransfer von Fdh-N auf Nar führen zur Translokation von Protonen über die Cytoplasmamembran [53, 55]. Dabei repräsentiert Formiat mit seinem niedrigen Standardredoxpotential von -430 mV einen idealen Elektronendonator für die Reduktion von Nitrat [55].



1.1.5 Synthese von Selenoproteinen in *E. coli*

Das Spurenelement Selen ist von essentieller Bedeutung für biologische Funktionen in Prokaryoten und Eukaryoten [40, 56]. Der spezifische Einbau von Selen erfolgt in Form der natürlich vorkommenden nicht-kanonischen Aminosäure Selenocystein (Sec bzw. U) [57]. Selenoproteine sind Oxidoreduktasen, deren katalytische Mechanismen anhand der eukaryotischen Glutathionperoxidase, der bakteriellen Glycinreduktase und der Fdh-H von *E. coli* aufgeklärt wurden [58-60]. Die Fdh-Isoenzyme stellen die einzigen Selenoproteine in *E. coli* dar und enthalten jeweils ein katalytisch aktives Selenocystein.

Der Einbau von Selenocystein erfolgt co-translational [61] und wird durch das *in-frame* Kodon TGA (UGA) kodiert [62, 63]. Für die Synthese von Selenocystein und dessen spezifischen Einbau werden u. a. die vier Genprodukte von *selABCD* benötigt [64]. Eine Mutation in einem dieser Gene verhindert die Translation des Kodons UGA und führt zum vorzeitigen Abbruch der Proteinbiosynthese, wodurch die

vollständige Reifung zu aktiven Fdhs verhindert wird [65]. Das im Mittelpunkt stehende Gen *seI/C* kodiert für die tRNA^{Sec}, deren Antikodon das Kodon UGA erkennt und den spezifischen Einbau von Selenocystein vermittelt [66]. Für die Dekodierung des Kodons UGA am Ribosom und zur Verhinderung eines Abbruchs der Translation wird ein Komplex, bestehend aus dem Translationsfaktor SelB [67], GTP, der Selenocysteyl-tRNA^{Sec} und einem strukturellen Motiv der mRNA (*selenocystein insertion sequence*, SECIS), ausgebildet [64].

Die Selenolgruppe des Selenocysteins hat im Vergleich zur Thiolgruppe des Cysteins einen niedrigeren pK_s-Wert von 5,3 (pK_s_{Thiol}: 8,3) und liegt bei einem neutralen pH-Wert negativ geladen vor. Die höhere Reaktivität der nukleophilen Selenolgruppe [68] bietet einen Vorteil für die Enzymkatalyse. Eine Variante des Enzyms Fdh-H, in der das Selenocystein gegen ein Cystein ausgetauscht wurde, zeigt eine 300-fach reduzierte katalytische Aktivität [60].

Ein Überangebot von Selen während des Zellwachstums führt trotz zahlreicher Kontrollmechanismen [64] zu einem unspezifischen Einbau in Proteine. Entgegen früherer Hypothesen, dass ein Überangebot von Selen zum Einbau von Selenomethionin führt [69], deuten spätere Untersuchungen auf den bevorzugten Einbau von Selen in Form einer Substitution von Cystein durch dessen isosteres Analogon Selenocystein [70, 71]. Diese Beobachtungen werden von den Erkenntnissen gestützt, dass die O-Acetylserin-Synthetase von *E. coli* Selenid als Substrat akzeptiert [72] und eine Aminoacetylierung der Cysteinyl-tRNA mit Selenocystein möglich ist [70, 73].

Ein optimaler spezifischer Einbau von Selenocystein in Proteine erfolgt bei der Ergänzung des Kulturmediums mit 100 nM Selenit. Dabei sollte die Cysteinkonzentration mit 80 µM einen ca. 800-fachen Überschuss gegenüber Natriumselenit aufweisen [64]. Selenitkonzentrationen über 1 µM resultieren in einem unspezifischen Seleneinbau [74]. Die zusätzlichen reaktiven Selenolgruppen können zur Störung zellulärer Prozesse führen.

1.2 Massenspektrometrie

Die Massenspektrometrie (MS) ist ein technisches Verfahren zur Massenanalyse von Atomen und Molekülen. Ein Massenspektrometer besteht grundsätzlich aus Ionenquelle, Analysator und Detektor. Zur Untersuchung wird der Analyt zunächst in der Ionenquelle in die Gasphase überführt und ionisiert. Die erzeugten Ionen werden im Massenanalysator entsprechend ihrer Masse-zu-Ladungsverhältnisse (m/z) getrennt und detektiert [75, 76].

1.2.1 Elektrospray-Ionisation

Die Atmosphärendruck-Ionisation (*atmospheric pressure ionization*, API) ermöglichte als erste Technik die Überführung eines Analyten von der Flüssigphase in die Gasphase und dessen Transfer in einen Massenanalysator [77]. Die Weiterentwicklung zur Elektrospray-Ionisation (*electrospray ionization*, ESI) und deren Kopplung mit Massenanalysatoren basiert auf Arbeiten von Dole und Fenn [78-80].

Die kontinuierlich arbeitende ESI-Quelle ermöglicht die Verbindung der Flüssigkeitschromatographie mit der Massenspektrometrie (Abbildung 6). Die Zufuhr der Probenlösung erfolgt durch eine Sprühkapillare bei Flussgeschwindigkeiten von 5 – 20 $\mu\text{l}/\text{min}$, an die ein elektrisches Potential von 3 – 4 kV relativ zur Gegenelektrode angelegt ist. Das dadurch erzeugte Aerosol aus mikrometergroßen elektrisch geladenen Tröpfchen wird durch einen N_2 -Trockengasstrom geleitet, der als Wärmequelle für die Verdampfung des leicht flüchtigen Lösungsmittels dient. Ein kleiner Anteil des versprühten Analyten gelangt durch die Öffnung einer beheizten Transferkapillare, welche die unter Atmosphärendruck stehende Sprühzone mit einer Kammer niederen Drucks (10^{-2} bar) verbindet. Das in einer Überschallexpansion entstehende Gas des evaporierenden Analyten wird durch eine erste Pumpstufe entfernt. Es folgt der Transfer des vollständig desolvatisierten und ionisierten Analyten durch einen Skimmer in die Hochvakuumregion (ca. 10^{-7} – 10^{-8} bar) des Massenspektrometers. Diese „sanfte“ Ionisation des Analyten unter Atmosphärendruck und dessen schrittweise Überführung in den Hochvakuumbereich des Massenspektrometers [79, 81-83] ermöglicht die Analyse intakter Makromoleküle [84].

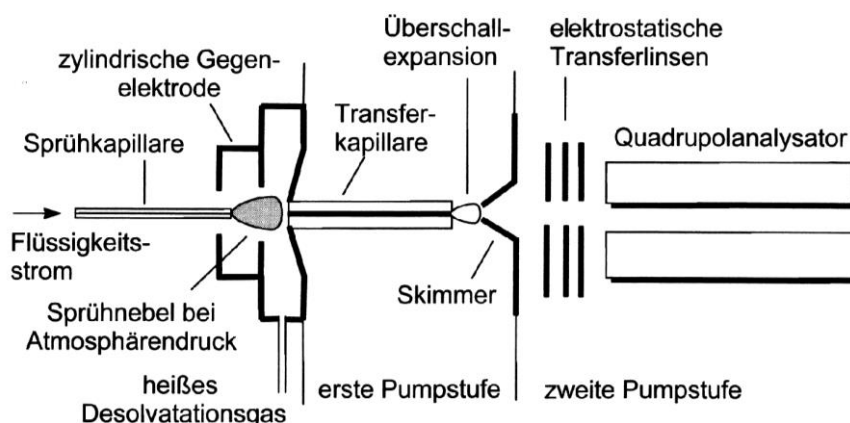


Abbildung 6: Schematischer Aufbau einer Elektrospray-Ionenquelle nach Fenn [79].

Am Auslass der Sprühkapillare ist die austretende Flüssigkeit einem elektrischen Feld von ca. 10^6 V/m^2 ausgesetzt, was zu einer Ladungstrennung in der Elektrolytlösung und zur Ausbildung eines Taylor-Konus führt (Abbildung 7A und B) [85, 86]. Aufgrund der hohen elektrostatischen Kräfte wird an diesem ein Flüssigkeitsstrahl in Richtung Gegenelektrode emittiert, welcher zerfällt und in ein Spray aus elektrisch geladenen Tröpfchen übergeht [87].

Die während der Ionisierung hervorgerufene Schrumpfung und der Zerfall der Tröpfchen bis hin zur Freisetzung der geladenen Ionen wird durch die Theorie der Coulomb-Explosion beschrieben. Die am Taylor-Konus erzeugten Tröpfchen besitzen anfangs Durchmesser von wenigen Mikrometern. Durch die Evaporation des Lösungsmittels verkleinert sich der Tropfendurchmesser, wodurch die Ladungsdichte steigt. Am Raleigh-Limit übersteigt die Ladungsabstoßung die Kohäsionskräfte der Oberflächenspannung und die Tröpfchen zerfallen in Mikro-Tröpfchen [88]. Mittels Blitzlicht-Mikrofotografiebildern wurde die Deformation der Tröpfchen während des Fluges in Richtung Gegenelektrode gezeigt und somit das Konzept der *Droplet Jet Fission* für die Erzeugung von Mikrotropfen gestützt (Abbildung 7C) [89-92]. Verjüngte Tropfenbereiche mit einer starken Krümmung tragen höhere Ladungsanteile [93], welche die Absoßung kleinerer Mikro-Tröpfchen induzieren [89, 93, 94]. Die Freisetzung vollständig desolvatisierter Ionen aus Mikro-Tröpfchen wird gegenwärtig durch zwei Modelle erklärt.

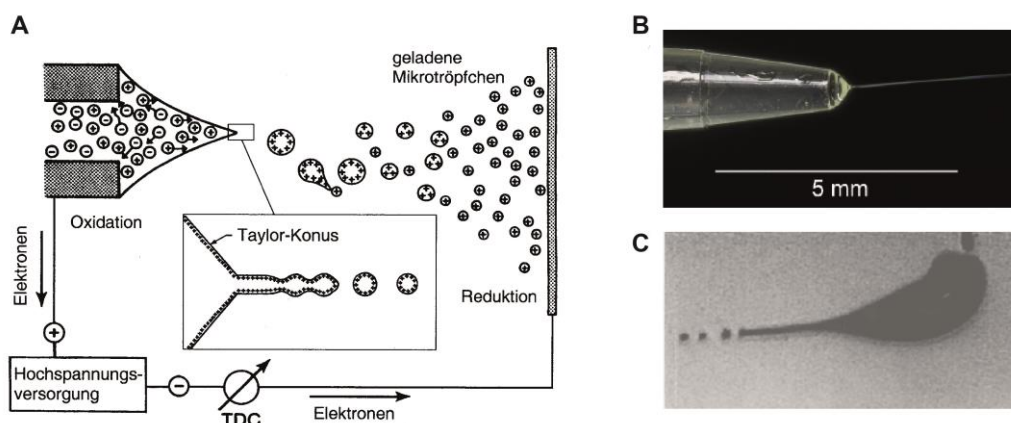


Abbildung 7: Elektrospray-Ionisation (A) Darstellung der Bildung des Taylor-Konus und Erzeugung eines Aerosols elektrisch geladener Tröpfchen [94]. **(B)** Fotografische Abbildung des Taylor-Konus [86, 95]. **(C)** Mittels Blitzlicht-Mikrofotografie erzeugtes Schattenbild eines deformierten Tröpfchens. Stark gekrümmte Bereiche tragen höhere Ladungsanteile, wodurch Mikro-Tröpfchen abgestoßen werden [89].

1.2.1.1 Modell des geladenen Rückstands

Das von Dole beschriebene Modell des geladenen Rückstands (*charged residue model*, CRM) beschreibt den sukzessiven Verlust aller Lösungsmittelmoleküle durch mehrere Coulomb-Explosionen bis hin zur Entstehung von Nano-Tröpfchen, die hinreichend klein sind, um nur noch ein einziges Analytmolekül zu enthalten. Die vollständige Desolvatisierung führt zur Übertragung der Ladungen auf den Analyten und somit zu dessen Ionisierung [78, 96, 97]. Die durch das CRM-Modell beschriebenen niedrigen Ladungszustände von Makromolekülen konnten experimentell bestätigt werden [98].

1.2.1.2 Modell der Ionenverdampfung

Das Modell der Ionenverdampfung (*ion evaporation model*, IEM) postuliert die Bildung desolvatisierter Ionen durch direkte Ionenemission von der Oberfläche hochgeladener Mikrotröpfchen mit einem Durchmesser von ca. 10 nm [99]. Das IEM-Modell unterstützt die Annahme, dass Analytmoleküle entsprechend ihrer räumlichen Ausdehnung unterschiedlich viele Ladungen tragen können. Flache und planare Moleküle besitzen einen höheren durchschnittlichen Ladungszustand als kompakte Moleküle. Dies geht mit der Beobachtung einher, dass entfaltete Formen globulärer Proteine unter identischen ESI-Bedingungen höhere Ladungszustände aufweisen [98, 100].

1.2.2 Nano-ESI

Die Entwicklung der Nano-ESI erhöhte die Sensitivität und verringerte die zur Analyse benötigten Probenvolumina. Sprühkapillaren mit geringerem Durchmesser ermöglichen niedrigere Flussraten und führen zur Bildung eines Aerosols kleinerer Tröpfchen [87]. Die heutzutage verwendeten Borosilikatglas-Kapillaren besitzen Nadelspitzen mit Öffnungsdurchmessern von 5 – 30 μm und erlauben Flussraten von 100 - 300 nl/min [101]. Bereits vergleichsweise niedrige Spannungen von 0,7 - 1,5 kV ermöglichen die Ionisierung des Analyten. Während in den konventionellen ESI Tröpfchen mit 1 – 2 μm Durchmesser erzeugt werden, ist die Tröpfchengröße der Nano-ESI mit 200 nm etwa 100- bis 1000-fach kleiner.

1.2.3 LTQ-Orbitrap-Massenspektrometrie

Die Analyse der desolvatisierten Ionen wird im Folgenden an dem in dieser Arbeit verwendeten Hybridmassenspektrometer LTQ-Orbitrap XL (Thermo Fisher Scientific, Abbildung 8) beschrieben. Nach dem Verlassen der ESI-Quelle werden die Ionen über eine Reihe von Radiofrequenz-Multipolen in die lineare Ionenfalle (LTQ) überführt, welche aus einem dreiteilig segmentierten Quadrupol-Analysator (Abbildung 9) besteht. Dieser ermöglicht den Aufbau eines elektrischen Felds frei von Verzerrungen. Das Anlegen einer leicht höheren Gleichstromspannung am vorderen und hinteren Segment ermöglicht die Speicherung der Ionen in axialer Richtung (z-Achse) [102]. Ein Radiofrequenz-Potential speichert die Ionen in radialer Dimension (xy-Ebene).

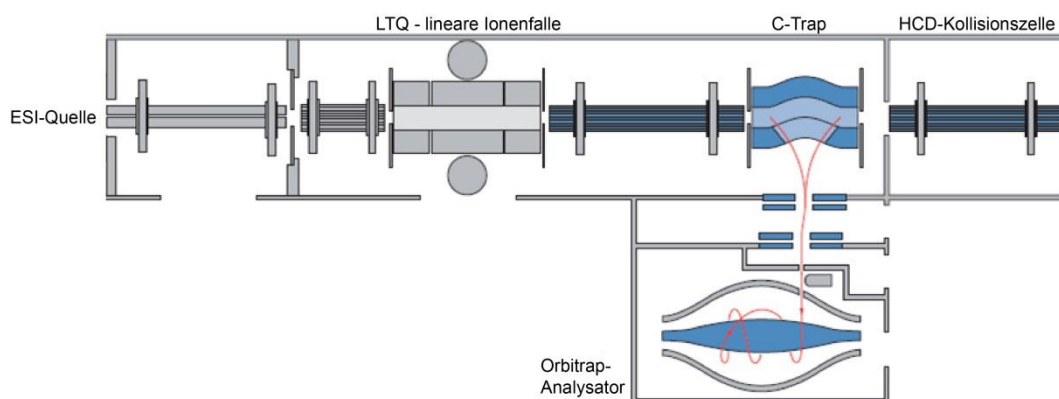


Abbildung 8: Schematischer Aufbau des LTQ-Orbitrap XL-Hybridmassenspektrometers [103]. Die Hauptbestandteile lineare Ionenfalle (LTQ), C-Trap, HCD-Kollisionszelle und Orbitrap sind gezeigt. Nach Injektion in den Orbitrap-Analysator kreisen die Ionen um die spindelförmige Zentralelektrode und oszillieren gleichzeitig entlang der z-Achse.

Durch Kollision mit einem Inertgas verlieren die gespeicherten Ionen kinetische Energie, was zu einer höheren Massenauflösung führt. Die lineare Fokussierung der Ionen in axialer Richtung vermindert Raumladungseffekte und ermöglicht eine höhere Speicherkapazität gegenüber dreidimensionalen Ionenfallen (Paul-Fallen), in denen die Ionenwolke punktuell fokussiert ist [104]. Ein Vorscan im Millisekundenbereich ermittelt den Totalionenstrom (*total ion current*, TIC) innerhalb eines gewünschten Massenbereichs. Zur Detektion der gespeicherten Ionen wird das Radiofrequenz-Potential so verändert, dass die Ionen beim Erreichen der m/z -spezifischen Resonanzfrequenz die Ionenfalle über seitliche Spalte in Richtung zweier Sekundärelektronenvervielfacher (SEV) verlassen und anschließend detektiert werden [102]. Durch die automatische Verstärkungsregelung (*automatic gain control*, AGC) werden Injektionszeiten für anschließende Analysenscans optimiert und eine effektive Ladungsdichte erzielt [105].

Um eine hohe Massengenauigkeit zu erzielen, werden die Vorläuferionen zur Analyse in die Orbitrap weitergeleitet (Abbildung 8 und Abbildung 10). Der axiale Auswurf der Ionen aus der linearen Ionenfalle wird durch Anlegen einer zusätzlichen Wechselspannung an den Elektroden des vorderen und hinteren Segmentes erreicht. Dabei entsteht ein elektrostatisches Potential entlang der z-Achse [102], wodurch Ionen bestimmter m/z -Werte mit einer Resonanzfrequenz oszillieren und die Ionenfalle in Richtung C-Trap verlassen (Abbildung 8). In der C-Trap verlieren die Ionen kinetische Energie durch Kollision mit Stickstoffmolekülen und werden als verdichtete „Ionenpakete“ tangential in den Orbitrap-Analysator injiziert [103]. Der von Alexander Makarov entwickelte Orbitrap-Analysator [106, 107] stellt eine Weiterentwicklung der Kingdon-Falle dar [108].

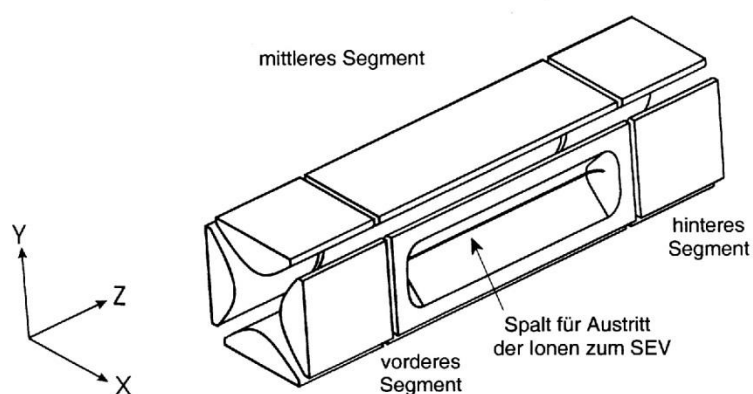


Abbildung 9: Schematische Darstellung der segmentierten zweidimensionalen, linearen Ionenfalle (LTQ) [102].

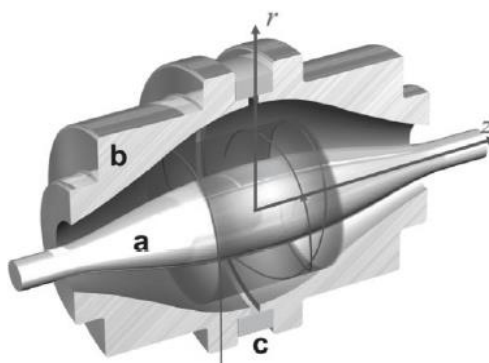


Abbildung 10: Schematische Darstellung der Orbitrap [109]. Ionen rotieren um die spindelförmige Zentralelektrode (a) und oszillieren dabei in Richtung der z-Achse. Die äußere Elektrode (b) ist durch einen Keramikring (c) in zwei Hälften geteilt. Die sich bewegenden Ionen erzeugen in der äußeren Elektrode einen Bildstrom, welcher verstärkt und aufgezeichnet wird. Die m/z -Werte der oszillierenden Ionen werden anschließend durch schnelle Fourier-Transformation (FFT) des Transienten ermittelt.

Die Orbitrap besteht aus einer spindelförmigen Zentralelektrode (Durchmesser 8 mm) und einer faßförmigen Außenelektrode (30 mm) (Abbildung 10). Während die Zentralelektrode auf ein Potential von -3,4 kV (Positiv-Ionen-Modus) eingestellt ist, verbleibt die Außenelektrode geerdet [110]. Die Ionenpakete werden mit kinetischen Energien von ca. 1,3 keV senkrecht zur z-Achse und tangential zur Innenelektrode injiziert. Die Ionenpakete rotieren um die Zentralelektrode und oszillieren entlang der z-Achse [111]. Die Außenelektrode detektiert den durch die axialen Schwingungen erzeugten Bildstrom, welcher der Oszillationsfrequenz ω entspricht [111]. Die axiale Oszillationsfrequenz (ω) verhält sich indirekt proportional zur Wurzel des Masse/Ladungs (m/z)-Verhältnisses (Gleichung 1; k: Gerätekonstante). Mittels schneller Fourier-Transformation (*Fast Fourier-Transformation*, FFT) wird aus dem Frequenzspektrum des Bildstroms ein Massenspektrum erzeugt [107]. Nach der Detektion eines hochaufgelösten Übersichtsspektrums im Orbitrap-Analysator werden ausgewählte Vorläuferionen mittels Tandem-Massenspektrometrie (MS/MS) analysiert. Die Ionen mit entsprechenden m/z -Werten werden innerhalb der LTQ isoliert und durch kollisions-induzierte Dissoziation (*collision induced dissociation*, CID) fragmentiert. Die entstehenden Fragmentionen können sowohl in der LTQ als auch in der Orbitrap analysiert werden [111].

$$\omega = \sqrt{\frac{z \cdot k}{m}}$$

Gleichung 1

1.3 Untersuchung von Proteinen und Peptiden mittels Massenspektrometrie

Die massenspektrometrische Identifizierung von Proteinen wird hauptsächlich mittels „Bottom-up“-Strategien realisiert [112]. Dabei werden Proteine durch die Wahl einer geeigneten spezifischen Protease in definierte Peptide gespalten. In heterogenen Gemischen sind Peptide meist nicht durch alleinige Bestimmung ihres Molekulargewichts identifizierbar, da isobare Moleküle mit unterschiedlicher Aminosäurezusammensetzung vorhanden sein können [113]. Das durch MS/MS erzeugte Fragmentierungsmuster gibt Aufschluss über die Aminosäuresequenz des Vorläuferions [112]. Je nach Art der Fragmentierung werden unterschiedliche Bindungen bevorzugt gespalten. Die in der Proteinanalytik häufig verwendete Fragmentierung durch CID führt zur Spaltung der Peptidbindung, wodurch charakteristische b- und y-Fragmentionen erzeugt werden (Abbildung 11) [114]. Die erhaltenen Massen der Vorläufer- und Fragmentionen werden durch die Verwendung von Suchmaschinen mit Informationen aus Proteinsequenz-Datenbanken verglichen und identifiziert.

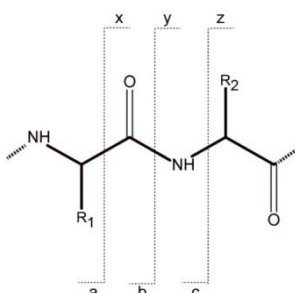


Abbildung 11: Nomenklatur der Fragmentionen nach Roepstorff, Fohlmann und Biemann [114, 115].

Durch Bindungsbrüche des Proteinerückgrats entstehen charakteristische N- und C-terminale Fragmentionen, wobei es je nach Art der Fragmentierung zur Bildung von korrespondierenden a- und x-, b- und y- bzw. c- und z-Ionen kommt.

1.3.1 Massenspektrometrische Analyse von co- und posttranslationalen Modifikationen

Co- und posttranslationale Modifikationen (PTMs) können die Reifung, Aktivität oder den Umsatz von Proteinen beeinflussen. Eine stetig wachsende Anzahl neuer Identifizierungen führte bisher zur Charakterisierung von mehr als 200 verschiedenen Modifikationen [116]. Die Veränderung der physikochemischen Eigenschaften von Proteinen durch PTMs [117] ist Grundlage der dynamischen Regulation zahlreicher

zellulärer Prozesse, wie z. B. Proteindegradation, subzelluläre Lokalisierung, Konformationsänderung, Protein-Protein-Interaktion und Signaltransduktion [118-120].

Traditionelle biologische Experimente, wie z. B. die Radiomarkierung von post-translational modifizierten Proteinen [121], immunologische Nachweise spezifischer PTMs [122] oder die ortsgerichtete Mutagenese potentieller Modifikationsstellen [123] sind zeitaufwendig und resultieren in einer vergleichsweise langsamen Datenakkumulation. Moderne LC/MS-Strategien hingegen ermöglichen eine eindeutige Identifizierung hunderter Modifikationsstellen verschiedenster PTMs [124-128] im Hochdurchsatzverfahren. Experimentelle Anreicherungsstrategien PTM-tragender Proteine, wie z. B. die TiO₂-Phosphopeptidanreicherung, spielen eine wesentliche Rolle für den Erfolg der anschließenden Identifizierungen [129].

LC/MS-Ansätze zeigen dennoch Limitierungen für die Analyse bestimmter Modifikationen. Phosphorylierungen von Serinen und Threoninen, Glykosylierungen [117], Sulfonierungen [117] und Nitrosylierungen [130] gehen häufig während der CID-Fragmentierung als Neutralverluste verloren. Dies verringert die Sequenzinformation und verhindert eine eindeutige Zuordnung von Modifikationsstellen im Protein.

Ebenfalls erschwert ist die LC/MS-Analyse von Selenocystein-haltigen Peptiden aufgrund der Umwandlung des co-translational eingebauten Selenocysteins zu Dehydroalanin, womit ein Massenverlust von 82 u verbunden ist (Abbildung 12) [131]. Im Gegensatz zu CID-labilen Modifikationen findet die Umwandlung von Selenocystein zu Dehydroalanin schon während der LC/MS-Probenvorbereitung statt.

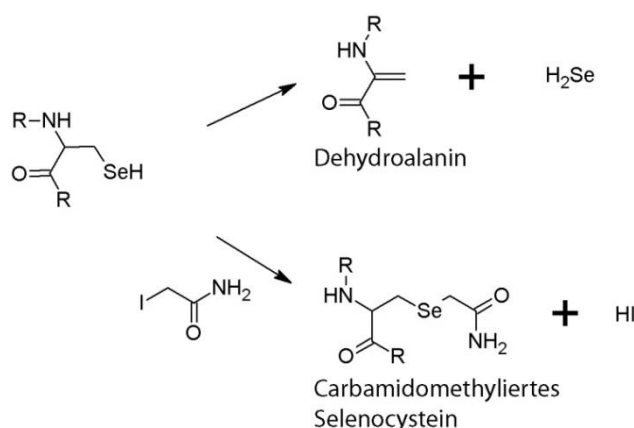


Abbildung 12: Umwandlung von Selenocystein zu Dehydroalanin und Carbamidomethylierung von Selenocystein durch Iodacetamid [131].

1.3.2 Peptid- und Protein-Quantifizierung mittels Massenspektrometrie

Klassische Quantifizierungsansätze in der Proteomik [132-135] liefern trotz einer guten Sensitivität, Linearität und dynamischen Breite keine Information über die Identität der untersuchten Proteine [136]. Eine Auftrennung des Proteoms durch 2D-Gelelektrophorese [137, 138] bietet eine unzureichende Auflösung und bleibt weitestgehend auf die Quantifizierung von löslichen und abundanten Proteinen beschränkt. Diese Probleme werden durch die Anwendung moderner LC/MS-Techniken gelöst. In der Proteomik proteolytisch erzeugte Peptide zeigen jedoch ein breites Feld an physikochemischen Eigenschaften, wie Größe, Ladung und Hydrophobizität. Eine exakte massenspektrometrische Quantifizierung benötigt den Vergleich jedes individuellen Peptids zwischen verschiedenen Proben.

1.3.2.1 Quantifizierung durch Isotopenmarkierung von Peptiden und Proteinen

Ein Quantifizierungsansatz basiert auf der Tatsache, dass stabil isotopen-markierte Peptide ein identisches Verhalten während der chromatographischen und massenspektrometrischen Analyse zeigen. Anhand der Massendifferenzen zwischen isotopenmarkierten und nicht-isotopenmarkierten Peptiden kann eine Quantifizierung durch den Vergleich der Signalintensitäten erfolgen. Die Einführung stabiler Isotope kann als interner Standard in Aminosäuren (I) metabolisch, (II) chemisch, (III) enzymatisch oder (IV) als extern zugesetzter Standard in Form eines synthetischen Peptids erfolgen [136, 139-146]. Erste metabolische Isotopen-Markierungen für die Quantifizierungen verschiedener Entwicklungsstadien von Mikroorganismen erfolgte durch die Verwendung von ^{15}N -substituiertem Kulturmedium [147, 148]. Größere Verbreitung erlangte die stabile Isotopenmarkierung durch Aminosäuren in Zellkultur (*stable isotope labeling by amino acids in cell culture*, SILAC) [149].

Eine post-biosynthetische Markierung von Proteinen und Peptiden erfolgt durch chemische oder enzymatische Derivatisierung *in vitro*. Ein Beispiel ist die durch Trypsin und GluC katalysierte Einführung von ^{18}O in Peptide [150, 151] für deren optimierte *de novo*-Sequenzierung mittels Massenspektrometrie [152]. Isotopen-Markierungs-Reagenzien für relative Quantifizierungen von Peptiden und Proteinen, wie z. B. *isotope-coded affinity tag* (ICAT) [153], *isotope-coded protein labels* (ICPL) [154], *isotope tags for relative and absolute quantification* (iTRAQ) [155], *tandem mass tags* (TMTs) [156] sowie Essigsäure- und Bernsteinsäureanhydrid [157-160]

nutzen die Reaktivität von Amin- und Thiolgruppen als Angriffspunkte für die chemische Derivatisierung.

1.3.2.2 Markierungsfreie Quantifizierung

Zur relativen Quantifizierung ohne Isotopenmarkierung („*label-free quantification*“) werden derzeit zwei verschiedene Techniken angewandt.

Zum einen werden die detektierten Signalintensitäten von Peptid-Vorläuferionen eines bestimmten Proteins verglichen. Dazu werden die Ionenchromatogramme bestimmter *m/z*-Werte von Vorläuferionen der LC/MS/MS-Analyse extrahiert (*extracted ion chromatogram*, EIC) und die massenspektrometrischen Peak-Flächen über einen bestimmten chromatographischen Zeitraum integriert [161]. Die Signalintensitäten individueller Peptide aus verschiedenen Proben können relativ zueinander quantifiziert werden [161-166].

Andererseits basiert das Zählen von Massenspektren (*spectral counting*) auf der empirischen Beobachtung, dass die Abundanz eines Proteins in einer Probe mit der Anzahl der erhaltenen MS/MS-Spektren für zugehörige Peptide (*peptide spectral matches*, PSMs) korreliert [167-169]. Eine relative Quantifizierung erfolgt durch den Vergleich der erhaltenen Spektrenanzahl verschiedener Proben. Im Gegensatz zur Quantifizierung durch Betrachtung der Signalintensitäten von Vorläuferionen, profitiert die Methode des *spectral counting* durch die Aufnahme von MS/MS-Daten über den gesamten chromatographischen Zeitraum.

1.4 Charakterisierung von Protein-Protein-Interaktionen

Zahlreiche zelluläre Prozesse werden von Netzwerken interagierender Proteine, dem sogenannten „Interaktom“ [170, 171] vermittelt, wobei über 80% eines Gesamtproteoms involviert sein können [172]. Nicht-kovalente Wechselwirkungen bilden die Basis für Protein-Protein-Interaktionen (PPIs) [173], welche nach der Stabilität und Dauer ihrer Interaktion klassifiziert werden. Stabile PPIs sind häufig die Grundlage von Proteinassamblierungen mit zellulär-strukturellen Funktionen, wie z. B. der Assoziation des α/β -Tubulin-Heterodimers zur Bildung von Mikrotubuli [174-176]. Transiente und schwach-affine Interaktionen ermöglichen u. a. eine situationsbedingte Signaltransduktion in intra- und extrazellulären Signalwegen. Die

meisten Transkriptionsfaktoren und PPIs auf Zelloberflächen zeigen hohe Dissoziationsraten [177, 178].

Die Charakterisierung nicht-aufgeklärter Proteinnetzwerke [179] erfolgt mit Hilfe von genetischen, biochemischen, biophysikalischen und bioinformatischen Ansätzen. Etablierte Methoden, wie das *yeast-two-hybrid*-System, Co-Immunopräzipitation (Co-IP), *pull-down*-Assays und Förster-Resonanz-Energietransfer (FRET), haben Vor- und Nachteile. Das *yeast-two-hybrid*-System ist oft mit hohen Falsch-Positiv-Raten assoziiert [180] und nur bedingt für die Aufklärung der PPIs von Membranproteinen anwendbar, die häufig Zielstrukturen für Arzneistoffe darstellen [181]. Immunologisch basierte Nachweise setzen die Verfügbarkeit von spezifischen Antikörpern voraus. Co-IP und *pull-down*-Assays nutzen häufig Zellysate, was zur Identifizierung von hochabundanten Proteinen als falsch-positiven Interaktionspartnern führen kann [182]. Transiente und schwach bindende Interaktionspartner können während der Waschschriffe verloren gehen und werden so häufig nicht detektiert.

Die hochauflösenden Methoden der NMR-Spektroskopie [95] und Röntgenkristallographie [26] haben sich für die Stukturaufklärung von Proteinen und Proteinkomplexen bewährt. Nachteile dieser Techniken sind die Notwendigkeit hoher Proteinenmengen, nicht-physiologische Analysenbedingungen und die geringe Zugänglichkeit für die Stukturaufklärung von Membranproteinen [183]. Die Interpretation der erhaltenen Strukturdaten ist zudem sehr aufwändig.

Eine alternative Technik zur Identifizierung und strukturellen Charakterisierung von Proteinen und Proteinkomplexen ist chemisches Cross-Linking¹ in Verbindung mit MS.

1.4.1 Identifizierung und Charakterisierung von Protein-Protein-Interaktionen mittels chemischer Quervernetzung

Das seit den 1960er Jahren verwendete chemische Cross-Linking [184] basiert auf der Einführung kovalenter Bindungen zwischen funktionellen Gruppen von Aminosäuren der Proteine mittels Quervernetzungsreagenzien [185-188]. Chemisches Cross-Linking und massenspektrometrische Analyse können in Kombination mit biochemischen Methoden wie z. B. Co-IP und

¹ In dieser Arbeit wird zusätzlich der englische Begriff Cross-Linking für die Quervernetzung von Proteinen verwendet.

Affinitätschromatographie verwendet werden [189, 190], wodurch u. a. die Identifizierung von transienten und schwach-affinen Interaktionspartnern ermöglicht wird (Abbildung 13A). Die kovalente Fixierung der Interaktionspartner eines Proteins erlaubt intensivere Waschbedingungen während affinitätschromatographischer Schritte, wodurch die Depletion hochabundanter Kontaminanten optimiert wird. Nach der Reinigung des quervernetzten Proteinkomplexes erfolgt die Identifizierung der Interaktionspartner mittels Massenspektrometrie.

Chemisches Cross-Linking und die massenspektrometrische Analyse der quervernetzten Peptide eröffnet zudem neue Möglichkeiten für die Ableitung nieder aufgelöster 3D-Strukturinformation von Proteinen und Protein-Komplexen (Abbildung 13B) [185]. Die experimentell ermittelten Distanzbeschränkungen zwischen quervernetzten Aminosäuren unterstützen die bioinformatische Erstellung von Modellen der Proteine bzw. Proteinkomplexe und stellen einen alternativen Ansatz für die Charakterisierung dreidimensionaler Proteinstrukturen dar [191].

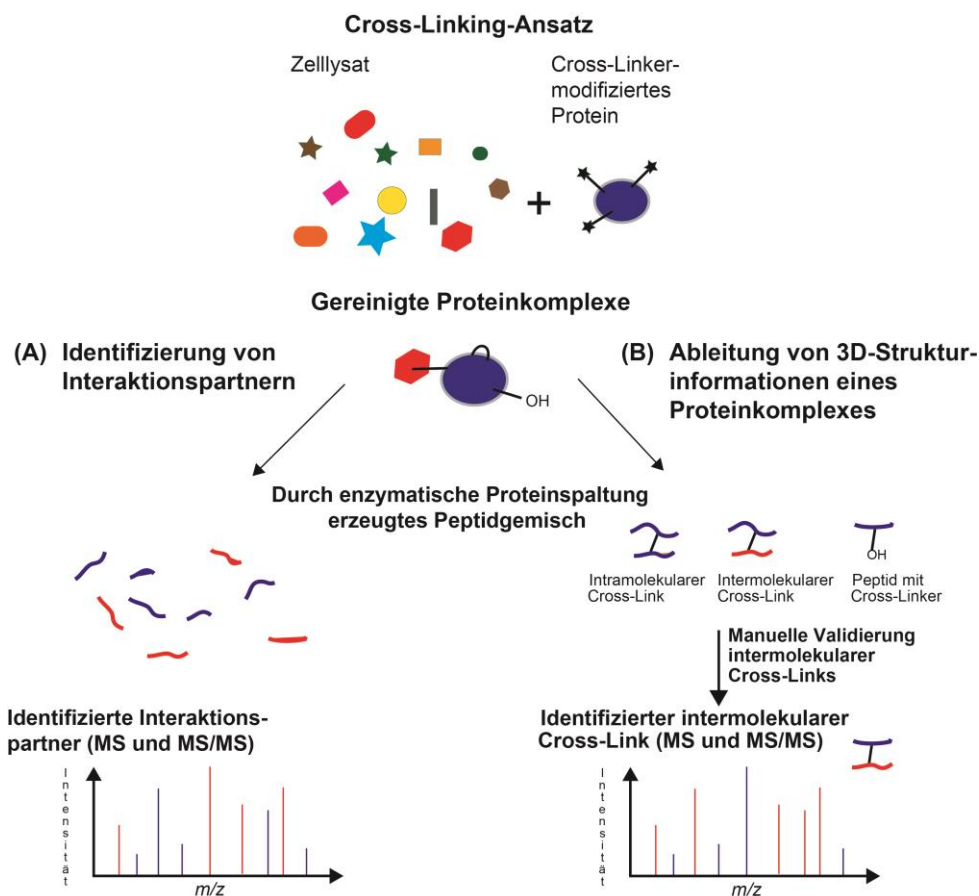


Abbildung 13: Schematische Darstellung des Arbeitsablaufes für die Identifizierung und Strukturaufklärung von Protein-Protein-Interaktionen mittels chemischen Cross-Linkings [187].

Da die Ableitung von Strukturinformationen auf Peptidebene erfolgt, ist die Größe der untersuchten Proteine und Proteinkomplexe kein limitierender Faktor [192] und führte bereits in vorangegangenen Studien zur erfolgreichen Charakterisierung von Multiproteinkomplexen [193, 194]. Zusätzlich ermöglichen unter physiologischen Bedingungen durchgeführte Cross-Linking-Experimente die Charakterisierung flexibler Proteinbereiche sowie die Aufzeichnung dynamischer Konformationsänderungen [191].

1.4.1.1 Reaktivität von Cross-Linking-Reagenzien

Cross-Linker weisen in Regel zwei reaktive Gruppen auf, die identisch oder verschieden sein können (homo- bzw. heterobifunktionelle Cross-Linker, Abbildung 14). Einige Cross-Linker enthalten zusätzlich eine dritte Funktionalität, die meist der Affinitätsanreicherung von Cross-Linking-Produkten dient [185]. Chemisches Cross-Linking macht sich hauptsächlich die Reaktivität von Amin-, Carboxyl- und Thiolgruppen zu Nutze. *N*-Hydroxysuccinimidester (NHS-Ester), wie z. B. Bis(sulfosuccinimidyl)glutarat (BS²G), reagieren unter physiologischen pH-Bedingungen (pH 7,0 – 7,5) mit primären Aminen (*N*-Terminus und ϵ -Aminogruppe von Lysin), wodurch unter Freisetzung der NHS- oder Sulfo-NHS-Gruppe stabile Amidbindungen entstehen (Abbildung 14A) [185]. Zudem wurden Umesterungsreaktionen mit Hydroxylgruppen an Tyrosinen, Threoninen und Serinen beobachtet [195, 196]. Diese auf der Oberfläche von Proteinen abundant vorkommenden hydrophilen Aminosäuren stellen zahlreiche und leicht zugängliche Reaktionstellen für Cross-Linking-Reagenzien dar.

Eine Alternative zur spezifischen Verknüpfung funktioneller Gruppen stellt das photoreaktive Cross-Linking dar [187]. Als photoaktivierbare Gruppen dienen vor allem Arylazide, Diazirine oder Benzophenone, die reaktive Gruppen in heterobifunktionellen Cross-Linkern [197] darstellen können. Chemische Quervernetzungsexperimente mit heterobifunktionalen Cross-Linkern wie z. B. Succinimidyl-2-([4,4'-azipentanamido]ethyl)-1,3'-dithiopropionat (SDAD), laufen in einer Zweischnittreaktion ab (Abbildung 14B). Dafür erfolgt zunächst die Modifikation des Proteins A mit der aminreaktiven Gruppe des Cross-Linkers. Nach dem Abstoppen der Reaktion und der Abtrennung überschüssigen Cross-Linkers wird der Interaktionspartner (Protein B) hinzugegeben und die Quervernetzung durch

Aktivierung der photoreaktiven Gruppe erzeugt. Durch UV-A-Bestrahlung ($\lambda_{\text{max}} \sim 365 \text{ nm}$) von Diazirinen wird, neben anderen Produkten [198], ein reaktives Carben gebildet, das sich mittels seines freien Elektronenpaares kovalent in räumlich nahe C-H- oder Heteroatom-H-Bindungen einfügen kann. Die Verwendung photoreaktiver Gruppen bietet den Vorteil einer besser zu kontrollierenden Reaktion, da eine Quervernetzung erst durch UV-A-Bestrahlung erzeugt wird. Das durch Photoaktivierung erzeugte Carben reagiert zudem unspezifisch und erhöht die Anzahl an möglichen Reaktionsstellen im Protein [187].

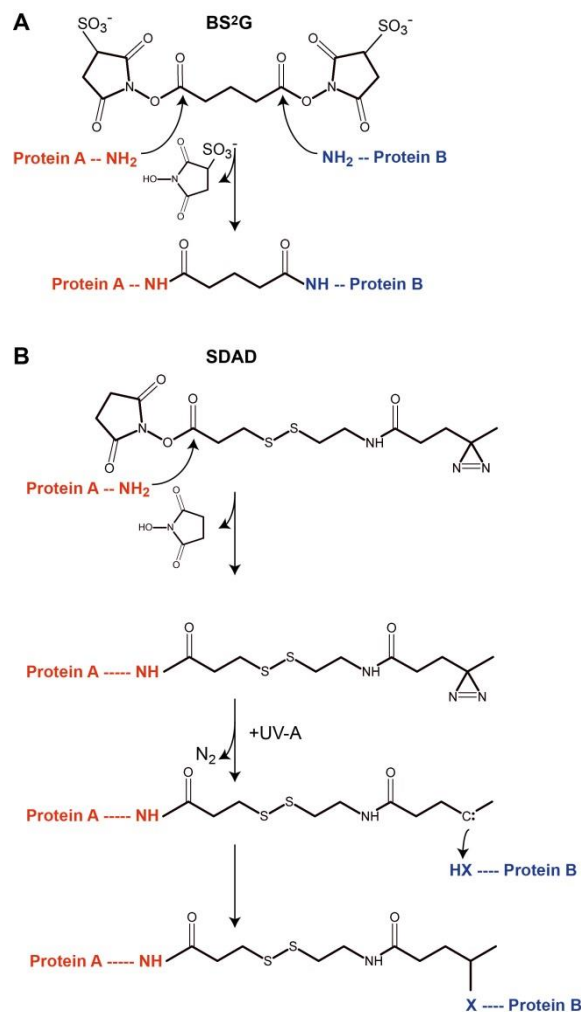


Abbildung 14: Reaktionsmechanismus homo- und heterobifunktionaler Cross-Linker [185].

(A) Chemische Quervernetzung von zwei Proteinen durch das homobifunktionelle Reagenz BS²G. Die Aminogruppen der Lysine greifen nukleophil die Carbonyl-C-Atome des Diesters an. Unter Abspaltung der NHS-Gruppe kommt es durch Ausbildung von zwei Säureamidbindungen zur Verknüpfung beider Proteine (B) Chemische Quervernetzung von zwei Proteinen durch das heterobifunktionelle Reagenz SDAD. In einer Zweischnittreaktion erfolgt zunächst die Modifikation von Protein A mit der aminreaktiven Gruppe von SDAD. Anschließend erfolgt die Quervernetzung mit Protein B durch Aktivierung der photoreaktiven Gruppe mittels UV-Bestrahlung.

1.4.1.2 Analyse und Identifizierung von Cross-Linking-Produkten

Die Auswertung früherer Quervernetzungs-Experimente war auf die Nutzung von gelelektrophoretischen Methoden beschränkt [199]. Heute erfolgt die Identifizierung von Cross-Linking-Produkten mittels Massenspektrometrie durch „*bottom-up*“- oder „*top-down*“-Strategien. Aufgrund der Limitierungen der „*top-down*“-Strategie für die Analyse größerer Proteinkomplexe [187] und die Identifizierung von Cross-Linking-Produkten in komplexen Proteingemischen wird bevorzugt die „*bottom-up*“-Strategie für die Ableitung von Distanzbeschränkungen verwendet.

Nach enzymatischer Spaltung („*bottom-up*“-Strategie) des Cross-Linking-Ansatzes werden mit dem Cross-Linker modifizierte Peptide unterschiedlich klassifiziert [200]: Typ 0 beschreibt Peptide, die mit partiell hydrolysierten Linkern modifiziert sind. Das Abstoppen der Cross-Linking-Reaktion mit NH_4HCO_3 kann zur Aminolyse des Quervernetzungsreagenzes führen. Typ 1 entspricht intrapeptidalen Cross-Links und Typ 2 beschreibt interpeptidale Cross-Links. Die Cross-Linking-Produkte vom Typ 2 liefern für die Strukturaufklärung von PPIs wichtige Distanzbeschränkungen, wobei das quervernetzte Peptid mit dem größeren Molekulargewicht als α -Peptid und das mit dem kleineren Molekulargewicht als β -Peptid bezeichnet wird.

Im komplexen Peptidgemisch des proteolytisch gespaltenen Cross-Linking-Ansatzes liegen die interpeptidal quervernetzten Peptide (Typ 2) meist niedrig abundant vor. Trotz der Optimierung moderner Massenspektrometer für eine Selektion möglichst vieler Vorläuferionen eines Massenspektrums [191], werden für die meist schwachen Signale von Cross-Linking-Produkten oftmals keine Fragmentationenspektren akquiriert. Strategien für die Anreicherung quervernetzter Peptide vor der massenspektrometrischen Analyse, wie starke Kationenaustausch-chromatographie [201] oder die Verwendung biotinylierter trifunktionaler Cross-Linker für eine Affinitätsanreicherung [202] haben sich etabliert. Die Verwendung von isotopenmarkierten Cross-Linkern [203] oder fragmentierbaren Cross-Linkern [204] stellen durch die Einführung von Reportersignalen [187] weitere Möglichkeiten für die verbesserte Identifizierung von Cross-Linking-Produkten dar.

Die Identifizierung der Cross-Linking-Produkte kann mit Hilfe von Programmen, wie z. B. GPMAW [205], CoolToolBox [206], xQuest [207] und StavroX [208], erfolgen. Dazu werden in einem virtuellen Cross-Linking-Experiment theoretische Massen von intermolekular quervernetzten Peptiden mit experimentell ermittelten Massen verglichen. Bewertungsalgorithmen erleichtern eine initiale Identifizierung von Cross-

Linking-Produkten. Eine anschließende manuelle Validierung und die Durchführung von *decoy*-Analysen (mit invertierten Proteinsequenzen oder einem veränderten Molekulargewicht des Cross-Linkers) [207] vermindern die Wahrscheinlichkeit falsch-positiver Identifizierungen.

2 Zielstellung

Bisherige molekularbiologische und biochemische Untersuchungen zum unspezifischen Einbau von Selen in *E. coli* [69, 71, 209] erlaubten keine Aussagen zur Lokalisation unspezifisch eingeführter Selenocysteine in Proteine. Ein Ziel dieser Arbeit war daher, die Auswirkungen eines Überangebotes von Selen auf den unspezifischen Einbau von Selenocystein in *E. coli* mittels Massenspektrometrie (MS) zu untersuchen und somit die Positionen des Einbaus von Selenocystein in Selenoenzymen zu identifizieren. Hierzu sollte zunächst die Reinigung der nativen Formiatdehydrogenasen Fdh-O und Fdh-N aus dem *E. coli*-Wildtypstamm MC4100 erfolgen. Im Anschluss an eine enzymatische Proteinspaltung sollten Selenopeptide mit Hilfe der MS identifiziert und die Häufigkeit des Einbaus von Selenocystein an den jeweiligen Positionen in Fdh-O und Fdh-N quantifiziert werden. Die zusätzliche MS-Analyse weiterer Proteine aus der *E. coli*-*selC*-Mutante FM460 sollte die Unabhängigkeit des unspezifischen Einbaus von Selenocystein von der spezifischen Selenocystein-Einbaumaschinerie zeigen.

Ein weiteres Ziel dieser Arbeit bestand in der Identifizierung und Charakterisierung von Protein-Protein-Interaktionen des Formiatstoffwechsels in *E. coli*. Vorausgegangene Studien zur Identifizierung von Bindungspartnern der respiratorischen Fdhs [210] hatten zu keiner Aufklärung von Proteininteraktionen geführt, welche möglicherweise eine Stabilisierung dieser komplexen Enzyme nach Reifung und Translokation über die Cytoplasmamembran gewährleisten. Die Durchführung eines chemischen Quervernetzungsexperimentes unter Verwendung eines heterobifunktionellen amin/photo-reaktiven Reagenzes und eine anschließende MS-Analyse der Quervernetzungsprodukte sollte zur Identifizierung von möglichen Bindungspartnern führen. Die erhaltenen Quervernetzungsprodukte sollten außerdem detaillierte Informationen über die Interaktionsflächen der Fdh/Protein-Interaktionen liefern.

Als weiteres Ziel dieser Arbeit sollte die strukturelle Charakterisierung der Interaktion von FocA und dem bereits bekannten Bindungspartner PflB erfolgen, wobei die Kombination aus chemischer Quervernetzung mit massenspektrometrischer Analytik Distanzinformationen zur dreidimensionalen Struktur des FocA/PflB-Komplexes

liefern sollte. Diese sollten als Grundlage für die bioinformatische Modellierung des Komplexes zur Aufklärung der Bindungsorientierung von FocA und PfIB im Komplex dienen, um so Einblicke in die einer regulierten Translokation von Formiat zugrundeliegenden Mechanismen zu erhalten.

3 Material und Methoden

3.1 Material

3.1.1 Chemikalien

3-N-(Morpholino)propansulfonsäure (MOPS)	Merck
Aceton	Roth
Acetonitril (für die Flüssigkeitschromatographie)	Merck
Acrylamid/Bisacrylamid-Lösung 37,5:1, 40% (w/v)	Merck
Ameisensäure (FA)	Sigma
Ammoniumhydrogencarbonat	Sigma
Ammoniumpersulfat (APS)	Sigma
Ammoniumsulfat	Sigma
Coomassie-Brillantblau G250	Fluka
Dichlorphenolindophenol (DCPIP)	Roth
Dikaliumhydrogenphosphat	Roth
Dimethylsulfoxid (DMSO)	Sigma
Dithiothreitol (DTT)	Sigma
Dodecyl- β -D-maltosid (DDM)	Roth
Essigsäure	Roth
Ethanol	Roth
Ethylendiamintetraessigsäure (EDTA)	Sigma
Glucosemonohydrat	Roth
Glycerin	Merck
Harnstoff	Roth
Hefeextrakt	Roth
2-(4-(2-Hydroxyethyl)-1-piperazinyl)ethansulfonsäure (HEPES)	Sigma
Iodacetamid	Applichem
Isopropanol	Roth
Kaliumdihydrogenphosphat	VWR
Kaliumnitrat	Grüssing

Laemmli-Probenpuffer	Bio-Rad
β -Mercaptoethanol	Sigma
Methanol	Roth
Natriumchlorid	Grüssing
Natriumdesoxycholat	Fluka BioChemica
Natriumdodecylsulfat, 10% (w/v) (SDS)	Bio-Rad
Natriumformiat	Sigma
Natriumhydroxid	Roth
Natriummolybdat	Merck
Natriumselenit	Sigma
ortho-Phosphorsäure	Roth
Pepton (tryptisch verdaut)	Roth
Phenazinmethosulfat (PMS)	AppliChem
Protease-Inhibitor <i>complete</i> EDTA-frei	Roche
Rotiphorese 10x SDS PAGE (10-fach konzentrierter Laufpuffer für die Gelelektrophorese)	Roth
Phenylmethylsulfonylfluorid (PMSF)	Roth
Salzsäure (Endotoxin-frei)	Sigma
Sucrose (Saccharose)	Roth
<i>N, N, N', N'</i> -Tetramethylethyldiamin (TEMED)	Fluka
Trifluoressigsäure (TFA)	Sigma
Trifluorethanol (TFE)	Fluka
Tris(hydroxymethyl)aminomethan (Tris)	Sigma
Triton X-100	Roth

3.1.2 Cross-Linking-Reagenzien

Bis(sulfosuccinimidyl)glutarat (BS ₂ G)	Thermo Fisher Scientific
Succinimidyl-2-([4,4'-azipentanamido]ethyl)- 1,3'-dithiopropionat (SDAD)	Thermo Fisher Scientific

3.1.3 Verbrauchsmaterialien

15 ml-Reaktionsgefäße	Greiner bio-one
50 ml-Reaktionsgefäße	Greiner bio-one
Kulturschalen	SPL Life Sciences
Membranfilter (Porengröße 0,2 µm)	Sartorius
Precise™ Protein Gel 4-20% (w/v) Acrylamid, 12 Taschen	Thermo Fisher Scientific
Spritzen Injekt Solo (5, 10 und 20 ml)	B. Braun
Zentrifugationsfiltereinheiten Ultra-0,5 und Ultra-15 (10 kDa Ausschlussgrenze)	Millipore

3.1.4 Geräte

Äkta FPLC	GE Healthcare
Analysenwaage ALC	Acculab
Analysenwaage LE225D	Sartorius
Autoklav VX-75	Systemc
Biostat ED Reaktor	Sartorius
Brutschrank ED 53 I	Binder
Gaulin Homogenisator MICRON LAB 20 CONT	Invensys
Geldokumentationssystem Gel Doc XR	Bio-Rad
Heizblock Thermostat	HLC
HPLC-Vorsäule Acclaim PepMap100 Nano-Trap, C18, 100 µm x 2 cm, 5 µm, 100 Å	Dionex
HPLC-Trennsäule Acclaim PepMap100, C18, 75 µm x 25 cm, 3 µm, 100 Å	Dionex
Magnetrührer M21/1	Framo
Magnetrührer R3T	mLw
Minishaker MS1	IKA
Minizentrifuge	Kisker Biotech GmbH
NanoDrop ND-1000	Peqlab
Personal Thermocycler	Biometra
pH-Meter PH211	Hanna Instruments
Pacific PW (Reinstwasseranlage)	TKA

Rotoren (JA 20, JLA 14, JS 13.1, JLA 16.250, 45-Ti)	Beckman-Coulter
Schütteltisch see-saw rocker SSL4	Stuart
SpeedVac Concentrator Savant SPD1010	Thermo Fisher Scientific
Spektrophotometer NanoVue Plus	GE Healthcare
Sterilbank HeraSafe KS	Thermo Fisher Scientific
Thermomixer Biometra TS 1	Analytik Jena
Thermomixer MKR 10/13	HLC
Trockenschrank Heraeus-T6	Kendro
Thermoshake VTO500	Gerhardt
Ultraschall Ultrasonic cleaner	VWR
Ultraschall Vibra Cell 75185	Bioblock Scientific
Ultrazentrifuge (Optima L Serie)	Beckman-Coulter
UV-A-Bestrahlungskammer (CLEO Performance R UV-A-Fluoreszenzlampe, Phillips)	Institut für Pharmazie, Eigenbau Prof. Rüttinger
Vakuumkonzentrator Duoconcentrator	miVac
Vortex-Mixer 7-2020	neolab
Waage MXX-2001	Denver Instrument
Waage SLR	Schott Instruments
Zentrifuge 5415 R	Eppendorf
Zentrifuge 5804 R	Eppendorf
Zentrifuge MiniSpin plus	Eppendorf
Zentrifuge J2-21	Beckman-Coulter
Zentrifuge J2-HS	Beckman-Coulter

3.1.5 LC/MS-Systeme

Massenspektrometer LTQ-Orbitrap XL	Thermo Fisher Scientific
Nano-ESI-Quelle	Proxeon
Nano-HPLC-System Ultimate 3000	Dionex

3.1.6 Elektronische Datenverarbeitung

GENTle 1.9.4	Magnus Manske [211]
	- Verwaltung von DNA- und Proteinsequenzen

KiNG	Kinemage, Next Generation [212] - Darstellung von Proteinstrukturen
MFCSwin	Sartorius AG - Steuerung des Biostat ED-Reaktors
Molprobrity	[213] - Validierung von Proteinstrukturen
Pymol 0.99	DeLano Scientific LLC - Darstellung von Proteinstrukturen
Proteome Discoverer 1.3	Thermo Fisher Scientific - Prozessierung von LC/MS/MS-Rohdaten - Mascot Datenbanksuche [214]
ProtParam	Expasy [215] - Berechnung von Extinktionskoeffizienten
StavroX 2.0.6	Michael Götze [208] - Identifizierung von Cross-Linking-Produkten
Xcalibur 2.0.7	Thermo Fisher Scientific - Steuerung der Nano-HPLC und des Nano-ESI- LTQ-Orbitrap-Massenspektrometers - Datenaufnahme und Bearbeitung der LC/MS/MS-Rohdaten

3.1.7 Puffer und Lösungen

Proteinreinigung:

- Aufschlusspuffer: 50 mM MOPS (pH 7,5), 1 μ M DNase I, 2 mM PMSF
- Membranpuffer I: 50 mM MOPS (pH 7,5), 1% (w/v) DDM
- Membranpuffer II: 50 mM MOPS (pH 7,5), 0,5% (w/v) DOC
- Waschpuffer I: 50 mM MOPS (pH 7,5), 1 M Ammoniumsulfat
- Waschpuffer II: 50 mM MOPS (pH 7,5)
- Waschpuffer III: 50 mM MOPS (pH 7,5), 0,1% (v/v) Triton X-100
- Elutionspuffer I : 0,5% (v/v) Triton X-100
- Elutionspuffer II: 50 mM MOPS (pH 7,5), 0,1% (v/v) Triton X-100, 1 M NaCl
- Fließmittel I: 50 mM MOPS (pH 7,5), 1 M NaCl, 0,5% (v/v) Triton X-100
- Fließmittel II: 50 mM HEPES (pH 7,5), 250 mM NaCl

Cross-Linking-Reaktionen:

- Reaktionspuffer I: 50 mM HEPES (pH 8,0), 100 mM NaCl, 0,007% (w/v) DDM
- Reaktionspuffer II: 100 mM MOPS (pH 7,0), 150 mM NaCl, 0,03% (w/v) DDM
- Reaktionspuffer III: 100 mM MOPS (pH 7,0), 150 mM NaCl, 0,1% (w/v) DDM

Anreicherung von Cross-Linking-Produkten:

- Waschpuffer IV: 100 mM Tris-HCl (pH 8,0), 150 mM NaCl, 1 mM EDTA, 0,1% DDM
- Elutionspuffer III: 100 mM Tris-HCl (pH 8,0), 150 mM NaCl, 1 mM EDTA, 0,1% DDM, 2,5 mM Desthiobiotin
- Fließmittel III: 50 mM HEPES (pH 8,0), 100 mM NaCl, 0,007% (w/v) DDM

SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE):

- Bio-Rad Laemmli-Probenpuffer I: 62,5 mM Tris-HCl (pH 6,8), 25% (v/v) Glycerin, 5% (v/v) β -Mercaptoethanol, 2% (w/v) SDS, 0,01% (w/v) Bromphenolblau
- Bio-Rad Laemmli-Probenpuffer II: 62,5 mM Tris-HCl (pH 6,8), 25% (v/v) Glycerin, 2% (w/v) SDS, 0,01% (w/v) Bromphenolblau
- Kolloidale Coomassie *Blue silver*-Färbelösung [216]: 10% (w/v) $(\text{NH}_4)_2\text{SO}_4$, 5% (w/v) Coomassie Brilliant Blue G250, 2% (v/v) ortho-Phosphorsäure
- Fixierlösung: Fixierlösung 40% (v/v) Methanol, 10% (v/v) Essigsäure

LC/MS/MS:

- Fließmittel A: 5% (v/v) Acetonitril, 0,1% (v/v) FA
- Fließmittel B: 80% (v/v) Acetonitril, 0,08% (v/v) FA

Enzymatische Proteinspaltung:

- Glu-C (Promega, *Staphylococcus aureus* V8): 0,5 $\mu\text{g}/\mu\text{l}$ Glu-C in 1 mM HCl
- Trypsin (Promega, porcine): 0,5 $\mu\text{g}/\mu\text{l}$ Trypsin in 1 mM HCl
- DTT-Lösung: 10 mM DTT in 100 mM NH_4HCO_3
- Iodacetamid-Lösung: 55 mM Iodacetamid in 100 mM NH_4HCO_3
- Peptidextraktions-Lösung: 25% (v/v) Acetonitril, 2,5% (v/v) TFA

3.1.8 Bakterienstämme

Tabelle 1: Verwendete Bakterienstämme

Bakterienstamm	Genotyp	Resistenz	Referenz/Herkunft
MC4100	<i>F</i> -, [<i>araD139</i>]B/r, Δ (<i>argF-lac</i>)169, λ -, <i>e14</i> -, <i>flhD5301</i> , Δ (<i>fruK-yeiR</i>) 725(<i>fruA25</i>), <i>relA1</i> , <i>rpsL150(strR)</i> , <i>rbsR22</i> , Δ (<i>fimB-fimE</i>)632(::IS1), <i>deoC1</i>	-	[217]
FM460	MC4100 Δ (<i>selC</i>) 400::Kan	Kanamycin	[39]

3.1.9 Proteine und Enzyme

Die Fusionsproteine FocA-Strep-tag II und PflB-His₆-tag sowie die lösliche Proteinfraction des *E. coli*-Stammes REK702 wurden am Institut für Biologie (AG Sawers) der Naturwissenschaftlichen Fakultät I der Martin-Luther-Universität Halle-Wittenberg hergestellt.

FocA-Strep-tag II (C-terminaler Strep-tag)	AG Sawers
PflB-His ₆ -tag (C-terminaler His-tag)	AG Sawers
Lösliche Proteinfraction des <i>E. coli</i> - Stammes REK702 [12]	AG Sawers
PageRuler ProteinLadder, gefärbt und ungefärbt	Fermentas
Lysozym (Hühnereiweiß)	Sigma, Taufkirchen

3.1.10 Nährmedien und Zusätze für die Bakterienkultur

Die pH-Werte der Nährmedien wurden bei Raumtemperatur eingestellt. Alle Medien wurden, wenn nicht anders angegeben, bei 121 °C für 20 min autoklaviert. Die Medienzusätze KH₂PO₄/K₂HPO₄, KNO₃, Glucose, Na₂SeO₃, Na₂MoO₄ und Kanamycin wurden separat sterilisiert und dem bereits autoklavierten Medium hinzugefügt. Zur Herstellung von Agarplatten wurde dem Medium 1,5% (w/v) Agar-Agar zugesetzt.

<u>LB-Medium [218]:</u>	1% (w/v) Pepton
	0,5% (w/v) Hefeextrakt
	1% (w/v) NaCl

<u>TGYEP-Medium [219]:</u>	1% (w/v) Pepton
	0,5% (w/v) Hefeextrakt
nach dem Autoklavieren	
hinzugegeben:	100 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ (pH 6,5), (autoklaviert)
	100 mM KNO_3 (autoklaviert)
	0,8% (w/v) Glucose (autoklaviert)
	2 - 100 μM Na_2SeO_3 (sterilfiltriert)
	2 - 100 μM Na_2MoO_4 (sterilfiltriert)

Tabelle 2: Verwendete Medienzusätze. Die Medienzusätze wurden in MilliQ-H₂O gelöst.

Medienzusatz	Stammkonzentration	Endkonzentration
Glucose	40% (w/v)	0,8% (w/v)
$\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$ (pH 6,5)	1 M	100 mM
Kaliumnitrat	1 M	100 mM
Kanamycin	30 mg/ml	30 $\mu\text{g}/\text{ml}$
Natriummolybdat	100 mM	2 – 100 μM
Natriumselenit	100 mM	2 – 100 μM

3.2 Zellbiologische Methoden

Die Kultivierung der *E. coli*-Stämme erfolgte unter sterilen Bedingungen.

3.2.1 Herstellung von *E. coli*-Vorkulturen

Zur Herstellung von *E. coli*-Vorkulturen wurden sterile Kultivierungsgefäße genutzt. Das Füllvolumen mit LB-Medium betrug ca. 30% des Gefäßvolumens. Die Anzucht erfolgte bei 37 °C und 125 Upm. Die Wachstumsdauer war versuchsabhängig und betrug 8 - 15 h.

3.2.2 Analyse des Zellwachstums

Das Zellwachstum wurde anhand der optischen Dichte der *E. coli*-Kulturen verfolgt. Die Messung der optischen Dichte erfolgte bei einer Wellenlänge von 595 nm.

3.2.3 Stammhaltung

Die Stammhaltung über kurze Zeiträume erfolgte durch Lagerung der Bakterienkulturen auf Agar-LB-Platten bei 4°C. Über längere Zeiträume wurde die Stammhaltung durch das Anlegen von Kryokulturen gewährleistet. Dafür wurden die Zellen in LB-Medium zu einer hohen Dichte herangezogen, sedimentiert und schließlich in LB-Medium mit 15% (v/v) Glycerin aufgenommen. Die Lagerung erfolgte bei - 80°C in 2 ml-Kryogefäßen mit Schraubdeckel.

3.2.4 Kultivierung

Die Kultivierung des *E. coli*-Stammes MC4100 erfolgte in einem Bioreaktor (Biostat ED-Reaktor). Die zentrale Prozessverwaltung wurde durch das Programm MFCSwin gesteuert. Zuerst wurden 5 l Medium, bestehend aus 1% (w/v) Pepton und 0,5% (w/v) Hefeextrakt, innerhalb des Reaktors für 45 min bei 121°C autoklaviert. Nach dem Abkühlen des Mediums auf Raumtemperatur erfolgte die Einstellung des pH-Wertes auf 7,5 durch Zugabe von sterilem 1M K₂HPO₄/KH₂PO₄. Anschließend wurde das Medium mit 0,8% (w/v) Glucose, 100 µM Natriummolybdat und 100 µM Natriumselenit ergänzt. Das automatische Kontrollsystem des Reaktors erlaubte einen kontinuierlichen und gleichmäßigen Eintrag von molekularem Sauerstoff in das Medium. Der Sauerstoffpartialdruck wurde auf 30% reguliert. Zellen einer 2 l-Vorkultur wurden durch Zentrifugieren pelletiert und zur Inokulation des Bioreaktors in 100 ml frischem Medium resuspendiert. Die Kultivierung erfolgte bei 37 °C und einer Sauerstoffsättigung von mindestens 30%. Die automatisierte Titration des Mediums auf pH 7,5 erfolgte durch Zugabe von 10% (v/v) Phosphorsäure oder 10% (w/v) Kaliumhydroxid. Das bakterielle Wachstum wurde durch die kontinuierliche Zugabe einer Nährlösung, bestehend aus 7% (w/v) Pepton, 3,5% (w/v) Hefeextrakt, 0,8% (w/v) Glucose, 100 µM Natriummolybdat und 100 µM Natriumselenit, unterstützt. Nach einer Kultivierungsdauer von 22 h wurde der Sauerstoffpartialdruck auf 0% gesenkt und 100 mM Kaliumnitrat hinzugegeben. Die Kultivierung wurde für weitere 2,5 h fortgesetzt, die Zellen anschließend durch Zentrifugieren pelletiert, in flüssigem Stickstoff eingefroren und bei - 20 °C bis zur weiteren Verwendung gelagert.

Alternativ erfolgte die Kultivierung der *E. coli*-Stämme MC4100 und FM460 in 5 bzw. 20 l-Gefäßen bei 22 °C oder 37 °C. Diese wurden zu 75% ihres Nettovolumens mit

TGYEP-Medium befüllt. Nach der Inokulation des Mediums mit einer 500 ml Vorkultur des entsprechenden *E. coli*-Stammes wurde das Gefäß mit Parafilm luftdicht verschlossen. Durch leichtes Rühren mittels Magnetrührer wurde eine gleichmäßige Durchmischung des Mediums gewährleistet. Die Kultivierung erfolgte über einen Zeitraum von 8 - 10 h. Die Kultur wurde gegen Ende des exponentiellen Wachstums mittels Eis-Wasserbad auf 0 - 4 °C abgekühlt. Die Zellen wurden anschließend durch Zentrifugation bei 5 000 × g für 30 min sedimentiert.

3.3 Proteinbiochemische Methoden

Die Arbeitsschritte der Proteinreinigung und Herstellung des Zellwand-Subproteoms erfolgten, insofern nicht anders vermerkt, bei 0 - 4°C.

3.3.1 Reinigung von Enzymen der formiatabhängigen Nitratatmung

Zur Reinigung von Proteinen der formiatabhängigen Nitratatmung wurden zunächst Zellen mit einer Feuchtmasse von 50 g in 150 ml Aufschlusspuffer resuspendiert. Anschließend erfolgte der Zellaufschluss mittels dreimaliger Passage durch eine French-Press und Abtrennung nicht aufgeschlossener Zellen und Zelltrümmern durch Zentrifugation des Zelllysates für 30 min bei 28 000 × g. Anschließend wurde der zellfreie Überstand erneut für 2 h bei 40 000 Upm zentrifugiert (Beckman-Coulter, Optima L Series; Type 45 Ti Rotor). Die sedimentierte Membranfraktion wurde in Membranpuffer I resuspendiert und eine Proteinkonzentration von 5 mg/ml eingestellt. Die Solubilisierung der Membranproteine erfolgte durch leichtes Schwenken bei 8°C über einen Zeitraum von 1 h. Alternativ wurde die Solubilisierung mit Membranpuffer II durchgeführt. Unlösliche Membranbestandteile wurden durch Zentrifugation für 1 h bei 40 000 Upm abgetrennt. Der lösliche Überstand wurde anschließend einer schrittweisen Ammoniumsulfatpräzipitation unterzogen. Durch eine langsame Zugabe von festem Ammoniumsulfat wurde eine Konzentration von 50% (w/v) eingestellt. Die Suspension wurde für 1 h gerührt und das entstandene Präzipitat durch Zentrifugation für 30 min bei 28 000 × g pelletiert. Der lösliche Überstand wurde anschließend durch abwechselndes Präzipitieren und Zentrifugieren für die Herstellung von 60, 70 und 80% (w/v) Ammoniumsulfat-enthaltenden Proteinfractionen verwendet. Das Präzipitat der 70% Ammoniumsulfat-enthaltenden Proteinfraction wurde in 5 ml Waschpuffer I resuspendiert.

Die weitere Proteinreinigung erfolgte mit Hilfe des Flüssigkeitschromatographiesystems ÄKTA FPLC, (GE Healthcare). Dazu wurde die 70% Ammoniumsulfatenthaltende Proteinfraction über eine Butylsepharose-Matrix (4,9 ml, GE Healthcare) geleitet, die zuvor mit zehn Säulenvolumina Waschpuffer I äquiliert wurde. Der erste Waschschrift erfolgte mit 10 Säulenvolumina Waschpuffer I. Für die Elution weiterer Kontaminanten wurde die Ammoniumsulfatkonzentration durch die Erhöhung der Konzentration des Waschpuffers II von 0% auf 100% innerhalb von zehn Säulenvolumina auf 0 M abgesenkt. Anschließend erfolgte die Elution der Zielproteine mit Elutionspuffer I. Fraktionen, in denen eine Fdh-Aktivität nachgewiesen werden konnte, wurden vereint und mittels Amicon-Zentrifugationsfiltereinheiten auf 2 ml konzentriert (Ausschlussgrenze für Molekulargewichte >30 kDa), wobei simultan ein Pufferaustausch zu Fließmittel I durchgeführt wurde. Die weitere Reinigung erfolgte mittels Größenausschlusschromatographie. Hierfür wurde zunächst eine Sephacryl S500-Säule (320 ml, GE Healthcare) mit mindestens zwei Säulenvolumina Fließmittel I äquiliert. Die Proteinlösung wurde injiziert und bei einer Flussrate von 0,5 ml/min aufgetrennt. Elutionsfraktionen mit detektierbarer Fdh-Aktivität wurden vereint und mittels Amicon-Zentrifugationsfiltereinheiten (Ausschlussgrenze für Molekulargewichte >30 kDa) gegen Waschpuffer II umgepuffert. Für die weitere Reinigung der Zielproteine wurde eine starke Anionenaustauschchromatographie (MonoQ 5/50 GL-Säule, Amersham Biosciences) durchgeführt. Die Säule wurde zunächst mit zwei Volumina Waschpuffer III äquiliert. Nach der Beladung der Säule mit der Proteinlösung erfolgte ein Waschschrift mit 10 Säulenvolumina Waschpuffer III. Die Zielproteine wurden anschließend durch eine Erhöhung der Konzentration des Elutionspuffers II von 0 auf 100% innerhalb von 10 Säulenvolumina eluiert.

Die qualitative Beurteilung der Proteinreinigung erfolgte mittels SDS-PAGE. Fdh-enthaltende Fraktionen wurden konzentriert, in flüssigem Stickstoff eingefroren und bis zur weiteren Verwendung bei - 20°C gelagert.

3.3.2 Herstellung eines Zellwand-Subproteoms

Zur Herstellung eines Zellwand-Subproteoms wurden Zellen aus einer 15 l-Kultur des *E. coli*-Stammes MC4100 durch Zentrifugation bei 5 000 × g für 30 min sedimentiert. Das gewonnene Zellpellet wurde in 800 ml 50 mM HEPES (pH 8,0) resuspendiert und durch Zentrifugation bei 5 000 × g für 30 min pelletiert. Dieser Waschschrift wurde dreimal wiederholt. Anschließend wurde das Pellet in 140 ml 50 mM HEPES (pH 8,0), 20% (w/v) Saccharose aufgenommen und 70 ml 0,1% (w/v) Lysozym in 100 mM EDTA hinzugegeben. Nach einer Inkubationszeit von 5 min wurden die Sphäroplasten durch Zentrifugation bei 13 000 × g für 20 min sedimentiert. Der gewonnene Überstand wurde dekantiert, in flüssigem Stickstoff eingefroren und bis zur weiteren Verwendung bei - 20°C gelagert.

Zur Verringerung der Komplexität des Zellwand-Subproteoms wurde eine Größenausschlusschromatographie durchgeführt. Dazu wurde zunächst eine Superdex S75 16/600 GL-Säule (GE Healthcare) mit zwei Säulenvolumina Fließmittel II NaCl äquilibriert. Anschließend wurde die Proteinfraction (190 mg Protein) des Zellwand-Subproteoms injiziert und bei einer Flussrate von 0,75 ml/ml aufgetrennt. Unterfraktionen wurden in flüssigem Stickstoff eingefroren und bis zur weiteren Verwendung bei - 20°C gelagert.

3.3.3 2,6-Dichlorphenolindophenol (DCPIP)-Enzymaktivitätstest

Die Identifizierung Fdh-enhaltender Fraktionen erfolgte durch die formiatabhängige Reduktion von DCPIP in Anwesenheit von PMS bei 37°C [36]. Die Nachweisreaktion wurde durch die Zugabe von 5 – 50 µl Proteinlösung zum Reaktionsansatz (50 mM MOPS (pH 7,5), 80 mM Natriumformiat, 260 µM PMS, 120 µM DCPIP) initiiert. Proteinfractionen mit Fdh-Aktivität wurden durch einen Farbumschlag des DCPIPs von blau nach gelb detektiert.

3.3.4 Proteinkonzentrationsbestimmung mittels Bradford-Assay

Zur Bestimmung der Proteinkonzentration wurde eine Standardreihe von 0 – 1 mg/ml bovinem Serum Albumin (BSA) in entsprechendem Puffer der zu bestimmenden Proteinfraction hergestellt. Der Reaktionsansatz für die Messung bestand aus 980 µl Bradford-Reagenz und 20 µl Proteinlösung. Die Messung der Absorption erfolgte bei

595 nm. Mit der aus der Standardreihe resultierenden linearen Kalibriergerade wurde die Proteinkonzentration der Proben berechnet.

3.3.5 Chemisches Cross-Linking

Die Cross-Linker-Stammlösungen wurden unmittelbar vor Beginn der Cross-Linking-Reaktion durch Lösen in DMSO hergestellt. In allen Cross-Linking-Experimenten erfolgte der Pufferaustausch oder die Abtrennung von überschüssigem Cross-Linker mittels Amicon Ultra-Zentrifugationsfiltereinheiten (Ausschlussgrenze für Molekulargewichte >10 kDa).

3.3.5.1 SDAD-Cross-Linking für die Identifizierung der Bindungspartner von Fdh-N und Fdh-O

Eine Proteinfraktion co-gereinigter Fdh-N und Fdh-O (siehe 3.3.1) wurde mit Reaktionspuffer I auf eine Proteinkonzentration von 35,7 μM verdünnt. SDAD wurde in DMSO gelöst und dem Reaktionsansatz in einem 140-fachen molaren Überschuss gegenüber den Fdhs hinzugefügt. Die Modifikation der Fdhs mit der aminreaktiven Gruppe von SDAD erfolgte für 2 h auf Eis unter Ausschluss von Licht. Die Reaktion wurde durch Zugabe von NH_4HCO_3 in einer Endkonzentration von 50 mM abgestoppt. Anschließend erfolgte die Abtrennung von überschüssigem Cross-Linker durch einen dreimaligen Pufferaustausch mit Reaktionspuffer I. Im nächsten Schritt erfolgte das Umpuffern einer durch Größenausschlusschromatographie hergestellten Unterfraktion des Zellwand-Subproteoms (siehe 3.3.2) gegen Reaktionspuffer I. Diese wurde anschließend in einem fünffachen molaren Überschuss (178,5 μM) gegenüber den Fdhs zum Reaktionsansatz hinzugefügt. Anschließend wurde die Quervernetzung zwischen den Fdhs und Interaktionspartnern durch UV-A-Bestrahlung ($\lambda_{\text{max}} = 365 \text{ nm}$, Strahlungsleistung = 8.000 mJ/cm^2) induziert.

3.3.5.2 SDAD-Cross-Linking für die Identifizierung der Bindungspartner von FocA

Gereinigtes FocA-Strep-tag II-Fusionsprotein wurde mit Reaktionspuffer II auf eine Konzentration von 10 μM verdünnt. Anschließend wurde SDAD in DMSO gelöst und dem Reaktionsansatz in einem 50-fachen molaren Überschuss gegenüber FocA

hinzugefügt. Die Modifikation von FocA mit der aminreaktiven Gruppe von SDAD erfolgte für 1 h bei Raumtemperatur unter Ausschluss von Licht. Die Reaktion wurde durch Zugabe von NH_4HCO_3 in einer Endkonzentration von 50 mM abgestoppt. Anschließend wurde ein Pufferaustausch zur Abtrennung von überschüssigem Cross-Linker mit Reaktionspuffer II durchgeführt. Eine lösliche Proteinfraction des *E. coli*-Stammes REK702 wurde dem Reaktionsansatz in einer Endproteinkonzentration von 100 μM hinzugefügt. Das Cross-Linking zwischen FocA und seinen Bindungspartnern wurde durch UV-A-Bestrahlung ($\lambda_{\text{max}} = 365 \text{ nm}$, Strahlungsleistung = 8.000 mJ/cm^2) induziert.

3.3.5.3 BS²G-Cross-Linking für die Strukturanalyse des FocA-PfIB-Komplexes

Für das Cross-Linking zwischen FocA und PfIB erfolgte zunächst ein Pufferaustausch der Proteinfractionen von FocA und PfIB mit Reaktionspuffer III. Die Interaktionspartner wurden in Endkonzentrationen von 20 μM (FocA-Pentamer) und 13,1 μM (PfIB-Dimer) eingesetzt. BS²G wurde in DMSO gelöst und dem Reaktionsansatz in einer Endkonzentration von 25 mM hinzugefügt. Die Cross-Linking-Reaktion erfolgte über einen Zeitraum von 30 min bei Raumtemperatur. Die Reaktion wurde durch Zugabe von NH_4HCO_3 in einer Endkonzentration von 50 mM abgestoppt.

3.3.6 Anreicherung von Cross-Linking-Produkten

3.3.6.1 Reinigung von FocA/Protein-Komplexen

FocA/Protein-Komplexe wurden mittels Streptactin-Affinitätschromatographie gereinigt. Dazu wurde eine Gravitationsflusssäule mit 1 ml Streptactin-Sepharose befüllt. Nach Äquilibrierung der Matrix mit fünf Säulenvolumina Waschpuffer IV wurde der Cross-Linking-Ansatz aufgetragen und mit 15 Säulenvolumina Waschpuffer IV gewaschen. Die gereinigten FocA/Protein-Komplexe wurden mit Elutionspuffer III in sechs Fraktionen à 0,5 ml eluiert.

3.3.6.2 Anreicherung von Fdh/Protein-Komplexen

Die Anreicherung von Fdh/Protein-Komplexen erfolgte mittels Größenausschlusschromatographie. Dazu wurde eine Superdex 200 16/300-Säule (GE Healthcare) verwendet, die zunächst mit zwei Säulenvolumina Fließmittel III äquilibriert wurde. Anschließend wurden 100 µl des Cross-Linking-Ansatzes (Proteinkonzentration 110 µM) auf die Säule geladen. Die Trennung der Proteine entsprechend ihrer hydrodynamischen Radien erfolgte bei einer Flussrate von 0,5 ml/min.

3.4 Bioanalytische Methoden

3.4.1 SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Die Auftrennung von Proteinen erfolgte mittels SDS-Polyacrylamid-Gelelektrophorese [220] in einer vertikalen Elektrophorese-Apparatur. Die Zusammensetzung der Gele ist in Tabelle 3 aufgeführt. Die einzelnen Bestandteile des Trenngels wurden gemischt und zwischen zwei Glasplatten gegossen. Nach Polymerisation des Trenngels und der Entfernung des Isopropylalcohols wurde das Sammelgel auf das Trenngel gegossen. Vor der Elektrophorese wurden die Proteinproben 1 : 1 (v/v) mit BioRad Laemmli-Probenpuffer I (mit β-Mercaptoethanol) oder II (ohne β-Mercaptoethanol) versetzt und für 3 min bei 98°C inkubiert. Von den Probenlösungen wurden 5 bis 15 µl Probe pro Geltasche aufgetragen. Die Elektrophorese erfolgte bei einer konstanten Stromstärke von 30 mA in Rotiphoese SDS-Laufpuffer, bis die Bromphenolblau-Lauffront das untere Ende des Gels erreicht hatte.

Nach Beendigung der Elektrophorese erfolgte eine Coomassie *Blue Silver*-Färbung [216]. Dazu wurden die Gele für 1 h in Fixierlösung geschwenkt, danach mit Wasser gespült und über Nacht in Coomassie *Blue Silver*-Färbelösung inkubiert. Die Hintergrundfärbung wurde durch anschließendes Waschen in Wasser entfernt. Die Dokumentation der Gele erfolgte mithilfe des Gel Doc XR-Systems (Bio-Rad).

Tabelle 3: Zusammensetzung der Polyacrylamidgele.

Bestandteil	Trenngel		Sammelgel
	12%	10%	5%
40% (w/v) Acrylamid/ Bisacrylamid-Lösung	3000 µl	2500 µl	650 µl
1,5 M Tris-HCl, pH 8,8	2500 µl	2500 µl	—
0,5 M Tris-HCl, pH 6,8	—	—	1250 µl
MilliQ-H ₂ O	4340 µl	4840 µl	3000 µl
10% (v/v) SDS	100 µl	100 µl	50 µl
10% (v/v) APS	50 µl	50 µl	25 µl
TEMED	10 µl	10 µl	10 µl

3.4.2 Enzymatische Proteinspaltung

Zur Vermeidung von Keratineintrag wurden die folgenden Schritte der enzymatischen Proteinspaltung unter Verwendung von Handschuhen und Ärmelschonern unter einer Sicherheitswerkbank (Vertical Laminar Airflow Polaris, Steril) durchgeführt.

3.4.2.1 *In-Gel*-Proteinspaltung

Für die *In-Gel*-Proteinspaltung wurden die entsprechenden Proteinbanden mit einem Skalpell aus dem gefärbten Polyacrylamidgel ausgeschnitten und in Würfel mit einer Kantenlänge von ca. 1 mm zerkleinert. Diese wurden in 1,5 ml-Reaktionsgefäße überführt und für 10 min unter Schütteln in 500 µl ACN gewaschen. Nach Entfernung des ACN erfolgte die Reduktion von Thiolgruppen durch Zugabe von 50 µl einer 10 mM DTT-Lösung. Dieser Reaktionsansatz wurde unter schütteln für 30 min bei 56°C inkubiert. Es folgte ein weiterer Waschschrift mit ACN. Anschließend erfolgte die Carbamidomethylierung der reduzierten Thiolgruppen durch Zugabe von 50 µl einer 55 mM Iodacetamid-Lösung und 20-minütiger Inkubation unter Lichtausschluss. Im Anschluss wurden die Gelstücke durch Zugabe von 100 µl eines 1:1 (v/v)-Gemischs aus Acetonitril und 100 mM NH₄HCO₃ vollständig entfärbt. Vor und nach diesem Schritt wurde erneut, wie oben beschrieben, mit ACN gewaschen. Zur Proteinspaltung wurden zunächst 2 µl-Enzym-Aliquote mit 100 mM NH₄HCO₃ auf 80 µl aufgefüllt. In Abhängigkeit von der Menge der Gelstücke wurden jeweils 20 - 40 µl Trypsin-Lösung bzw. Trypsin- und GluC-Lösung hinzugegeben. Für die Proteolyse mit Trypsin wurden die Gelstücke zunächst für 2h bei 4°C und anschließend über

Nacht bei 37°C inkubiert. Zur Proteolyse mit einer Kombination aus Trypsin und GluC wurden die Gelstücke zunächst für 3 h bei 4 °C mit GluC inkubiert. Danach wurde Trypsin hinzugegeben und für eine weitere Stunde bei 4°C inkubiert. Die proteolytische Spaltung erfolgte schließlich über Nacht bei 37°C.

Zur Extraktion der Peptide aus den Gelstücken erfolgte die Zugabe von Peptidextraktions-Lösung, deren Volumen dem zweifachen Volumen des Reaktionsansatzes entsprach. Nach 10-minütigem Schütteln bei 37°C wurden die Überstände in neue 1,5 ml-Reaktionsgefäße überführt. Die Extraktion wurde anschließend wiederholt. Vor der massenspektrometrischen Analyse (LC/MS/MS) wurden die vereinten Überstände im Vakuumkonzentrator auf ein Drittel des Ausgangsvolumens eingeeengt.

3.4.2.2 In-Lösungs-Proteinspaltung

Zunächst erfolgte die Präzipitation der Proteinfractionen durch Zugabe eines fünffachen Überschusses an eiskaltem Aceton und einer Inkubation über Nacht bei -20°C. Die ausgefällten Proteine wurden anschließend durch Zentrifugation bei 16 000 x g und 0°C pelletiert. Der Überstand wurde verworfen und das Proteinpellet im Vakuumkonzentrator getrocknet. Die Proteine wurden in 25 µl 8 M Harnstoff, 400 mM NH₄HCO₃ resuspendiert und für 20 min im Ultraschallbad inkubiert. Nach der Verdünnung des Ansatzes durch Zugabe von 170 µl MilliQ-H₂O wurde dieser erneut für 5 min im Ultraschallbad inkubiert. Die Proteolyse erfolgte durch Zugabe von Trypsin-Lösung bzw. Trypsin- und Glu-C-Lösung. Je nach Proteinmenge wurde ein Enzym/Protein-Verhältnis von 1:50 - 1:70 (w/w) eingesetzt. Für die Proteolyse mit Trypsin wurde der Ansatz nach der Zugabe des Enzyms über Nacht bei 37°C inkubiert. Für die Proteolyse mit einer Kombination aus Trypsin und GluC wurde die Probe zunächst für 6 h mit GluC bei Raumtemperatur inkubiert, woran sich die Zugabe von Trypsin und eine weitere Inkubation über Nacht bei 37°C anschloss. Abschließend wurden die Proben durch Zugabe von 1% (v/v) TFA auf einen stark sauren pH-Wert gebracht.

3.4.3 Nano-HPLC/Nano-ESI-LTQ-Orbitrap-Massenspektrometrie

Die Analyse von Proben erfolgte über ein Nano-HPLC-System (Ultimate), das *online* über eine Nano-ESI-Ionisationsquelle (Proxeon) an das LTQ-Orbitrap XL-Hybridmassenspektrometer (Thermo Fisher Scientific) gekoppelt wurde. Peptidmischungen wurden mittels Autosampler aufgenommen und in Gegenwart von 0,1% (v/v) TFA mit einer Flussrate von 20 µl/min auf die Vorsäule (Acclaim PepMap, C18, 300 µm x 5 mm, 5µm, 100 Å, LC Packings) geleitet und über einen Zeitraum von 15 min gewaschen. Anschließend erfolgte die Trennung der Peptide über eine Trennsäule (Acclaim PepMap C18, 75 µm x 250 mm, 3 µm, 100Å, LC Packings), die zuvor mit Fließmittel A äquilibriert wurde. Die Trennung erfolgte mit einer Flussrate von 300 nl/min, wobei die Säulentemperatur auf 30°C reguliert wurde. Für die Elution wurde in einem 90-minütigen bzw. 180-minütigen Gradienten die Konzentration des Fließmittels B von 0% auf 40% erhöht.

Die Erfassung von MS- und MS/MS-Daten erfolgte über den gesamten Zeitraum der Elution. Hierfür wurden zunächst hochaufgelöste Massenspektren ($R = 60000$ bei m/z 400) im Orbitrap-Analysator über einen m/z -Bereich von 300 – 2000 aufgenommen. Datenabhängig wurden die fünf intensivsten Signale nacheinander in der LTQ mit einem Massenfenster von $\Delta m/z = 2,5$ isoliert und mittels CID fragmentiert. Dabei wurden nur Ionen isoliert, die mindestens zweifach geladen waren und eine minimale Signalintensität von 500 *Counts* aufwiesen. Die relative Kollisionsenergie wurde auf 35% reguliert. Die gebildeten Fragmentationen wurden aufgrund der höheren Sensitivität in der LTQ detektiert. Die analysierten Vorläuferionen wurden anschließend für 180 s von einer erneuten Fragmentierung ausgeschlossen (*dynamic exclusion*). Die Datenaufnahme wurde mittels XCalibur 2.07 und DCMS link 2.0 kontrolliert.

3.4.4 Auswertung der MS- und MS/MS-Daten

Die Auswertung von massenspektrometrischen Rohdaten (RAW-Dateien) erfolgte mit Hilfe der Software Proteome Discoverer 1.3 (Thermo Fisher Scientific). Zur Identifizierung von Proteinen wurde ein Abgleich der MS- und MS/MS-Daten mit der UniProtKB/SwissProt-Datenbank [221] unter Verwendung der Suchmaschine *Mascot* 2.2.06 durchgeführt [214].

Die Identifizierung von Cross-Linking-Produkten erfolgte mit Hilfe der hauseigenen Software StavroX 2.0.6 [208]. Dazu wurden die massenspektrometrischen Rohdaten zunächst mittels der Software Proteome Discoverer 1.3 in StavroX-kompatible MGF (*Mascot generic format*)-Dateien umgewandelt. Neben MGF-Dateien benötigt StavroX die Primärstrukturen der kovalent verknüpften Proteine für die Erstellung einer Massenliste theoretisch verknüpfter Peptide. Hierfür wurden die Aminosäuresequenzen im FASTA-Dateiformat bereitgestellt. Carbamidomethylierungen von Thiol-Seitenketten wurden als statische, Oxidationen von Methioninen als dynamische Modifikationen angenommen. Bei der Verwendung der Proteasen Trypsin und GluC wurden die Schnittstellen C-terminal von Lysin, Arginin, Glutamat und Aspartat in Betracht gezogen. Für jede mögliche Spaltstelle innerhalb eines Peptids wurden maximal zwei überlesene Schnittstellen zugelassen. Es erfolgte ein automatischer Abgleich der theoretischen mit den experimentell ermittelten MS- und MS/MS-Daten durch StavroX. Hierfür wurde die Massentoleranz für Vorläuferionen auf maximal 10 ppm begrenzt. Für die Zuordnung der Fragmentationen wurde eine Massentoleranz von 0,8 Da zugelassen. Potentielle Cross-Linking-Produkte wurden anschließend manuell bewertet. Zur Quantifizierung von identifizierten Peptiden und Proteinen wurden zugehörige Fragmentationenspektren (*peptide spectral matches*) ausgezählt (*spectral counting*).

3.5 Bioinformatische Methoden

Die bioinformatische Erstellung des FocA/PfIB-Komplexes erfolgte mit Hilfe des Software-Pakets Rosetta [222] auf Basis des Betriebssystems Linux Redhat. Die Kristallstrukturen von FocA aus *E. coli* (PDB-Eintrag: 3KCU), FocA aus *V. cholerae* (PDB-Eintrag: 3KLY) und PfIB aus *E. coli* (PDB-Eintrag: 1H17) wurden der Proteindatenbank [223, 224] entnommen.

3.5.1 Modellierung eines N-terminalen Bereichs von FocA mittels RosettaScripts

Die Strukturvorhersage des N-Terminus von FocA aus *E. coli* erfolgte mit Hilfe von RosettaScripts [225]. Hierfür wurde die Modellierung des N-terminalen Bereichs (Aminosäuren 1 - 32) der α -Untereinheit innerhalb des FocA-Pentamers vorgenommen.

Zunächst erfolgte die Beschränkung des *N*-terminal möglichen Faltungsraumes durch die Rekonstruktion fehlender Bereiche (Abbildung A 1) der unvollständigen Kristallstruktur von FocA aus *E. coli*. Für die Ergänzung der fehlenden Struktur der Ω -Schleife (Aminosäuren 102 - 110) der γ -Untereinheit wurden die Atomkoordinaten dieses Bereichs von der α -Untereinheit übernommen. Dazu wurden zunächst beide Untereinheiten mittels Pymol übereinander gelegt und anschließend die Atomkoordinaten für die Aminosäuren 102 – 110 der α -Untereinheit in die PDB-Datei der γ -Untereinheit übertragen. Die gleiche Vorgehensweise erfolgte für die Rekonstruktion des fehlenden C-terminalen Bereichs (Aminosäuren 281 – 285) der α -Untereinheit. Als strukturelle Vorlage wurden hierfür die C-terminalen Atomkoordinaten der α -Untereinheit von FocA aus *V. cholerae* verwendet.

Für die Erstellung einer weiteren Beschränkung des *N*-terminal möglichen Faltungsraumes wurde die Voraussage der Membrantopologie von FocA durch die Erstellung einer Span-Datei (Abbildung A 2) durchgeführt. Dies erfolgte mit Hilfe der Anwendung *octopus2span.pl* [226]. Anschließend wurde eine Fragment-Bibliothek über die Anwendungen *runsam.py*, *clean_up_sam_rdb.src* und *make_fragments.pi* generiert.

Die Modellierung der *N*-terminalen Region wurde unter Verwendung von *rosetta_scriptsdefault.linuxgccrelease* (zugehörige Flag- und XML-Dateien siehe Abbildung A 3 und Abbildung A 4) durchgeführt. Hierbei erfolgte die zusätzliche Beschränkung des Faltungsraumes des *N*-terminalen Bereichs der α -Untereinheit durch die gleichzeitige Modellierung des angrenzenden C-terminalen Bereichs (Aminosäuren 279 – 285) der ε -Untereinheit.

3.5.2 Clusteranalyse der *N*-terminal modellierten Region von FocA

Zur Beurteilung der erhaltenen Konformationen für die *N*-terminal modellierten Aminosäuren 1 – 32 der α -Untereinheit erfolgte eine Clusteranalyse mit Hilfe des Rosetta-Algorithmus. Zunächst wurden 5% der Modelle ausgewählt, welche die niedrigsten Energiewerte für die *N*-terminale Region (Aminosäuren 1 - 32) aufwiesen. Die energetische Beurteilung der *N*-terminalen Region erfolgte durch die Anwendung *score_by_specific_residue.py*. Eine zuvor erstellte Liste definierte die zu bewertenden Aminosäuren 1 - 32 der α -Untereinheit. Anschließend wurde die Clusteranalyse mit Hilfe der Anwendung *cluster.default.linuxgccrelease* (zugehörige

Flag-Datei siehe Abbildung A 5) mit einem Clusterradius von 1,5 Å RMSD (C_{α} - C_{α} -Distanzen) durchgeführt, wobei die strukturelle Gruppierung auf die Aminosäuren 10 - 285 der α -Untereinheit beschränkt wurde.

3.5.3 Rekonstruktion der FocA-Pentamere

Die Rekonstruktion vollständiger FocA-Pentamere erfolgte durch die Anwendung *make_FocA_pentamer.py* (siehe Abbildung A 6). Dabei wurden die Aminosäuren 10 - 285 N-terminal modellierter α -Untereinheiten ausgewählt und durch zufällige Rekombination zu FocA-Pentameren assembliert. Die rekonstruierten FocA-Pentamere wurden anschließend durch die Anwendung *rosetta_scripts.default.linuxgccrelease* energetisch minimiert [227] (zugehörige XML-Datei und Flag-Dateien siehe Abbildung A 7 und Abbildung A 8). Eine zuvor erstellte Span-Datei (Abbildung A 9) definierte die Transmembranbereiche von FocA.

3.5.4 Modellierung des FocA/PfIB-Komplexes

Zur Modellierung des FocA/PfIB-Komplexes erfolgte zunächst die Relaxierung der Kristallstruktur von PfIB aus *E. coli*. Das anschließende *Docking* von FocA mit PfIB erfolgte durch RossettaDock [228]. Hierfür wurden zunächst beide Strukturen mittels Pymol neu orientiert. PfIB wurde in einem Abstand, welcher seiner halben Länge entsprach, der cytoplasmatischen Seite von FocA gegenübergestellt. Die Erstellung der FocA/PfIB-Modelle erfolgte über die Anwendung *rosetta_scripts.default.linuxgccrelease* (zugehörige XML- und Flag-Datei siehe Abbildung A 10 und Abbildung A 11).

Zur Darstellung des FocA/PfIB-Modells, das die native Konformation repräsentiert, erfolgte zunächst die strukturelle Gruppierung mittels *BCL::Cluster* [229]. Hiervon wurden 10% der energetisch ungünstigsten FocA/PfIB-Komplexe ausgeschlossen. Die verbleibenden FocA/PfIB-Modelle wurden entsprechend ihrer Interaktionsflächenenergien sortiert. Ungefähr 1% der Modelle, welche über die energetisch günstigsten Interaktionsflächen verfügten, wurden für die strukturelle Gruppierung durch *BCL::Cluster* ausgewählt.

Für die Clusteranalyse wurde ein Algorithmus gewählt, welcher die durchschnittliche Distanz des Proteinrückgrats (C_{α} - C_{α} -Distanzen) aller Modelle von benachbarten Clustern berechnet (*average pair-wise distance*). Eine Liste der durchschnittlichen

paarweisen Distanzen wurde zunächst mit Hilfe der Anwendung *docking_minimal_rmds_calculator_parallel.py* (Abbildung A 12) erstellt. Die anschließende Clusteranalyse erfolgte zunächst durch die Auswertung der Clusterumfänge. Hierfür wurden die erhaltenen Cluster durch die Erstellung eines Dendrogramms mit Hilfe der Anwendungen *bcl_executable.exe* visualisiert. Anschließend erfolgte die Auswertung der Clusterumfänge innerhalb Pymols.

Um das Modell des FocA/PfIB-Komplexes zu finden, das die native Konformation am besten repräsentiert, erfolgte der Vergleich der Interaktionsflächenenergie von Gruppen, welche einen durchschnittlichen paarweisen RMSD-Wert von 11 Å aufweisen. Die Darstellung der der Interaktionsflächenenergien von Modellen einzelner Cluster erfolgte mit Hilfe der Anwendung *output_cluster_energies.py* (Abbildung A 13).

4 Ergebnisse

4.1 Untersuchung des unspezifischen Seleneinbaus in Enzyme der formiatabhängigen Nitratatmung

4.1.1 Reinigung der Selenoenzyme Fdh-O und Fdh-N

Die Produktion und Isolierung der nativen Formiatdehydrogenasen Fdh-N und Fdh-O erfolgte aus dem *E. coli*-Wildtypstamm MC4100. Die Kultivierung der Zellen wurde unter anaeroben Wachstumsbedingungen in TGYEP-Medium [219] durchgeführt, welches durch den enthaltenen Phosphatpuffer einen stabilen pH-Wert gewährleistet. Für die Untersuchung des unspezifischen Seleneinbaus wurden die Zellen bei einer Konzentration von 2 μM oder 100 μM Natriumselenit kultiviert, was einem 500-fachen bzw. 10-fachen molaren Überschuss an Cystein (1 mM) [230] gegenüber Natriumselenit entsprach. Beide Konzentrationen erlaubten die Beurteilung der Auswirkungen eines relativen Überangebotes von Selenit gegenüber Cystein für den spezifischen und unspezifischen Seleneinbau in die Enzyme Fdh-O und Fdh-N.

Die Reinigung der Fdhs aus den mit 2 μM und 100 μM Natriumselenit kultivierten Zellen erfolgte nach identischen Protokollen, wobei die Zellen kurz vor dem Ende der exponentiellen Wachstumsphase geerntet wurden. Nach dem Zellaufschluss mit Hilfe eines Hochdruckhomogenisators wurden Zelltrümmer durch Zentrifugation abgetrennt. Anschließend wurden die Membranfraktionen durch Ultrazentrifugation gewonnen. Die Membranproteine wurden solubilisiert und einer schrittweisen Ammoniumsulfatpräzipitation unterzogen. Die durch Zugabe von 70% (w/v) Ammoniumsulfat erhaltene Proteinfraction wurde in drei aufeinanderfolgenden chromatographischen Schritten, bestehend aus hydrophober Interaktionschromatographie, Größenausschlusschromatographie und starker Anionenaustauschchromatographie, gereinigt.

Für die qualitative Kontrolle der Proteinreinigung wurde eine SDS-PAGE-Analyse durchgeführt (Abbildung 15). Aus dem erhaltenen Gel wurden ausgewählte Proteinbanden (Abbildung 15, I - VIII) ausgeschnitten und mit Trypsin proteolytisch gespalten. Danach erfolgte die Analyse der Peptidgemische mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS (Tabelle 4). Die Untersuchung

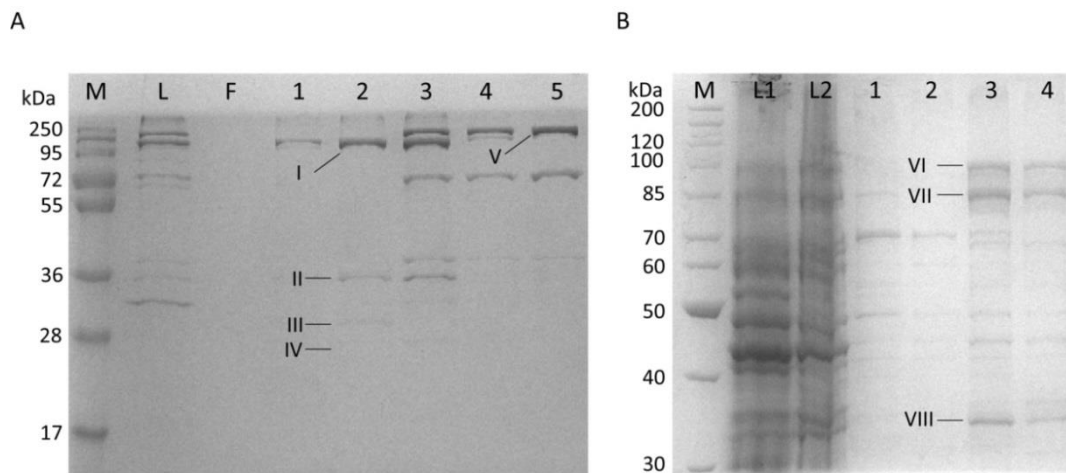


Abbildung 15: SDS-PAGE-Analyse des letzten Reinigungsschrittes mittels Anionenaustauschchromatographie (MonoQ 5/50 GL-Säule). Die Färbung der Proteine erfolgte mit kolloidaler Coomassie-Brillantblau-Lösung. **(A)** Proteinreinigung aus mit 2 µM Natriumselenit kultivierten MC4100-Zellen. Die Auftrennung der Proteine wurde in einem 12%-igen Trenngel durchgeführt. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur L: Proteinlösung vor der Reinigung; Spur F: Durchlaufraktion; Spuren 1 – 5: Elutionsfraktionen. **(B)** Proteinreinigung aus mit 100 µM Natriumselenit kultivierten MC4100-Zellen. Die Auftrennung der Proteine wurde in einem 8%-igen Trenngel durchgeführt. Spur M: Proteinmarker (*PageRuler Unstained Protein Ladder*); Spur L1: Proteinlösung vor der Proteinreinigung durch Größenausschlusschromatographie (Sephacryl S500-Säule); Spur L2: Proteinlösung vor der Proteinreinigung mittels Anionenaustauschchromatographie; Spuren 1 – 4: Elutionsfraktionen. Die Ergebnisse der MS-Analyse der Gelbänder I – VIII sind in Tabelle 4 dargestellt.

führte zur Identifizierung der drei Untereinheiten (G, H und I) der gemeinsam gereinigten Isoenzyme Fdh-O und Fdh-N. Die jeweiligen Untereinheiten wurden mit einer ähnlichen Sequenzabdeckung identifiziert. Die Bestätigung der katalytischen Untereinheiten erfolgte mit absoluten Sequenzabdeckungen von 70% (FdoG) bzw. 77% (FdnG), der elektronenübertragenden Untereinheiten mit 50% (FdnH) bzw. 60% (FdoH) und der integralen Membranuntereinheiten mit 18% (Fdoi bzw. Fdni). Die vergleichsweise geringe Sequenzabdeckung von Fdoi und Fdni ist in Übereinstimmung mit deren schwacher Anfärbung im SDS-Polyacrylamidgel (Abbildung 15, Gelbänder III und IV) und könnte in den hydrophoben Eigenschaften dieser integralen Membranproteine [32, 36] sowie dem Vorliegen nur weniger Spaltstellen für Trypsin und GluC begründet liegen. Die Abundanz der Proteine wurde durch *spectral counting* abgeschätzt. Die erhaltenen *peptide spectral matches* (PSMs) für identifizierte Peptide der Untereinheiten von Fdh-O und Fdh-N (Tabelle 4) aus den mit 2 µM und 100 µM Natriumselenit kultivierten Zellen deuten auf eine deutlich variierende Produktion beider Enzyme in *E. coli* während der Nitratatmung hin.

Tabelle 4: Massenspektrometrische Analyse der Proteinreinigung. Untersucht wurde der Reinigungsschritt der Fdhs mittels Anionenaustauschchromatographie (Abbildung 15). Die Proteinidentifizierung erfolgte durch *In-Gel*-Proteinspaltung der Banden I – VIII und anschließender Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS. Die Proteinbanden I – V stammen von der Proteinreinigung aus mit 2 μ M Natriumselenit, die Proteinbanden VI – VIII von der Proteinreinigung aus mit 100 μ M Natriumselenit kultivierten MC4100-Zellen.

Gelbande	Protein	UniProtKB-Eintrag	Identifizierte Peptide	Peptide spectral matches (PSMs)	Sequenz-abdeckung [%]	Mascot Score
I	FdnG	P24183	64	428	75	9352
	FdoG	P32176	52	279	66	6298
II	FdnH	P0AAJ3	18	95	45	1179
	FdoH	P0AAJ5	15	58	44	521
III	FdnI	P0AEK7	2	19	15	55
IV	FdoI	P0AEL0	3	6	12	173
V	NarG	P09152	73	382	60	7963
VI	FdoG	P32176	48	110	48	2727
	FdnG	P24183	28	40	31	836
VII	FdoG	P32176	44	133	48	3045
	FdnG	P24183	21	32	23	781
VIII	FdoH	P0AAJ5	19	83	57	794
	FdnH	P0AAJ3	8	13	26	235
	OmpA	P0A910	11	15	34	431

Weiterhin erfolgte durch die Anionenaustauschchromatographie die vollständige Reinigung und Abtrennung der Nitratreduktase A von den Fdhs. Die Proteinreinigung aus den mit 100 μ M Natriumselenit kultivierten Zellen zeigte die Anwesenheit von Kontaminanten, wobei u. a. das in *E. coli* hochabundant auftretende Protein OmpA identifiziert wurde.

4.1.2 Identifizierung von Selen-haltigen Peptiden aus Fdh-O und Fdh-N

Nach der Identifizierung der gereinigten Proteine sollten die Auswirkungen eines Überangebotes von Selen während des Zellwachstums auf dessen unspezifischen Einbau in die Enzyme der formiatabhängigen Nitratatmung untersucht werden. Im Fokus lag die Untersuchung der Selenocystein-enthaltenden Enzyme Fdh-O und Fdh-N.

Die katalytischen Untereinheiten von Fdh-O und Fdh-N beinhalten jeweils ein Selenocystein an Position 196 und 15 bzw. 14 Cysteine [38, 41]. Der Einbau des Selenocysteins U196 wird durch die tRNA^{Sec} vermittelt [65]. Die Identifizierung des

spezifisch eingebauten Selenocysteins diente als Kontrolle für die MS-Analysen und ermöglichte die Nutzung beider Isoenzyme als geeignete Systeme für die Untersuchung des unspezifischen Seleneinbaus.

Zunächst wurde untersucht, ob ein Überangebot von Selen in einem unspezifischen Einbau von Selenocystein oder Selenomethionin in die Fdhs resultiert. Dabei konnte kein Einbau von Selenomethionin detektiert werden, was frühere Vermutungen bestätigt, dass ein unspezifischer Seleneinbau hauptsächlich in Form von Selenocystein stattfindet [64, 70, 71].

Die Untersuchung von Selenocystein-enthaltenden Peptiden wird durch dessen Umwandlung zu Dehydroalanin während der Probenaufarbeitung erschwert, womit ein Massenverlust von 82 u verbunden ist (Abbildung 12) [131]. Für die Identifizierung Selenocystein-haltiger Peptide wurden zwei verschiedene analytische Ansätze etabliert. Zum einen wurden isolierte Proteine aus mit 2 μ M Natriumselenit kultivierten Zellen reduziert und alkyliert, was zur Carbamidomethylierung aller zugänglichen Cysteine und Selenocysteine und einer Massenzunahme von 57 u führte. Zum anderen wurden isolierte Proteine aus mit 100 μ M Natriumselenit kultivierten Zellen für die Etablierung einer alternativen Identifizierungsstrategie von Selenocysteinen nicht reduziert und alkyliert. Dies führte zur ausschließlichen Identifizierung von Dehydroalaninen, anhand derer auf die Anwesenheit von Selenocysteinen in Peptiden geschlossen werden konnte.

Entsprechende MS- und MS/MS-Daten belegen die eindeutige Identifizierung von carbamidomethylierten Selenocysteinen und korrespondierenden Dehydroalaninen für bestimmte Positionen in den untersuchten Peptiden der Fdh-N und Fdh-O. Abbildung 16A zeigt die Identifizierung eines selenocysteinenthaltenden Peptids der FdoG mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS-Analyse. Das Signal bei m/z 993,428 zeigt das charakteristische Isotopenmuster eines Selenocystein-haltigen Peptids. Dieses wird durch die sechs natürlich vorkommenden Isotope des Selens und ihren jeweiligen Häufigkeiten hervorgerufen: ^{74}Se (0,9%), ^{76}Se (9,0%), ^{77}Se (7,6%), ^{78}Se (23,6%), ^{80}Se (49,7%) und ^{82}Se (9,2%). MS/MS-Daten der Peptide LPSTu₆₁₈FAEENGSIIVNSGR (Abbildung 16B) und LPSTa₆₁₈FAEENGSIIVNSGR (Abbildung 16C) belegen jeweils die Identifizierung eines carbamidomethylierten

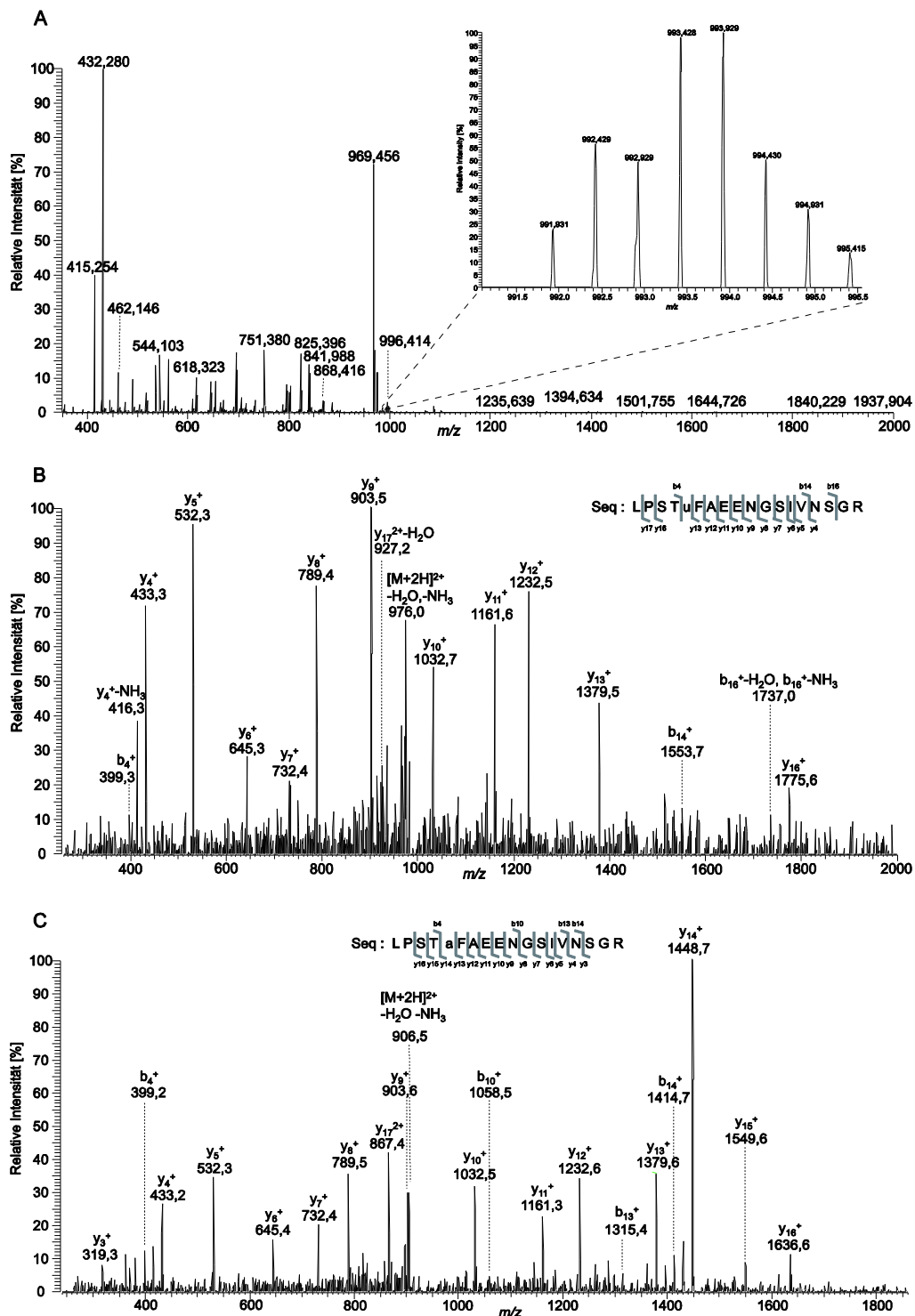


Abbildung 16: Identifizierung eines Selenocystein enthaltenden Peptids von FdoG. (A) Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS-Analyse. Das Selenopeptid wurde bei m/z 993,428 als doppelt geladene Spezies detektiert. Die Massenabweichung zwischen theoretischer und experimentell bestimmter Masse beträgt 1,1 ppm. Das charakteristische Isotopenmuster des Selenocystein-enthaltenden Peptids ist vergrößert dargestellt. (B, C) Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS/MS-Analyse der Peptide LPST_{u618}FAEENGSIIVNSGR (B, m/z 993,428) und LPST_{a618}FAEENGSIIVNSGR (C, m/z 923,950). Vorläuferionen wurden innerhalb der linearen Ionenfalle (LTQ) selektiert und mittels CID fragmentiert. (u) Carbamidomethyliertes Selenocystein; (a) Dehydroalanin.

Selenocysteins (u) und eines Dehydroalanins (a) an Position 618 der FdoG (Abbildung 17).

Durch die massenspektrometrische Analyse der Untereinheiten FdoG und FdnG wurden fünf Selenocysteine an identischen Positionen identifiziert, wobei die Cysteine C165, C380, C391, C554 und C618 substituiert wurden (Abbildung 17 und Tabelle 5). Ein Einfluss der Natriumselenitkonzentration während der Kultivierung auf die Anzahl und Positionen dieser ausgetauschten Cysteine konnte nicht nachgewiesen werden. Für fünf weitere Positionen innerhalb der FdoG und FdnG (C50, C53, C57, C359 und C572) wurde trotz Kultivierung bei einer höheren Natriumselenitkonzentration von 100 µM kein Austausch von Cysteinen durch Selenocysteine detektiert (Abbildung 17). Die Cysteine an den Positionen 50, 53 und 57 sind zusammen mit dem in dieser Analyse nicht identifizierten Cystein an Position 92 an der Koordination des [4Fe-4S]-Zentrums der FdoG bzw. FdnG beteiligt [37].

1	MQVSRROFFKICAGGMAGTTAAALGFAPSVALAETRYKLLRTRTRNTCTYCSVGGGLLMYSLGDGAKNAKASIFHIEGDDPHVNRGALCPKGAGLVD MDVSRROFFKICAGGMAGTTAAALGFAPKQALAQARNYKLLRAKEIRNTCTYCSVGGGLLMYSLGDGAKNAREAIYHIEGDDPHVNRGALCPKGAGLLD	FdoG FdnG
101	FIHSESRLLKFPPEYRAPGSDKWQIISWEEAFDRIAKLMKEDRDANYIAQNAEGVTVNRWLSTGMLCASASSNETGYLTQKFSFALGMLAVDNQARVTHGPT YVNSENLRYPEYRAPGSDKWQIRISWEEAFSRIAKLMKADRDANFIEKNEQGVTVNRWLSTGMLCASGASNETGMLTQKFSFALGMLAVDNQARVTHGPT	FdoG FdnG
201	VASLAPTFGRGAMTNHWVDIKNANLVVMGGNAEAHPVGRWAMEAKIHNGAKLIVIDPRFTRTAADVADYAPIRSGTDIAFLSGVLLYLLNNEKFNRE VASLAPTFGRGAMTNHWVDIKNANVVMGGNAEAHPVGRWAMEAKNNNDATLIVVDFPRFTRTASVADIYAPIRSGTDITFLSGVLRYLNIENKINAE	FdoG FdnG
301	YTEAYTNASLIVFEDYGFEDGLFTGYDAEKRYDKSSWYIELDENGFAKRDTTLQHPRVWNLLKHVSRYTPDVVENICGTPKDAFLKVEEYIAETSAH YVKHYTNASLLVRDDFAFEDGLFSGYDAEKRYDKSSWNYQLDENGYAKRDETLTHPRVWNLLKHVSRYTPDVVENICGTPKDAFLKVEVLASTSAP	FdoG FdnG
401	DKTASFLYALGWTQHSVGAQNIRTMAMIQLLLGNMGMAGGGVNALRGHSNIQGLTDLGLLSQSLPGYMLPSEKQTDLQTYLTANTPKPLLEGQVNYWGN DRTTTTFLYALGWTQHTVGAQNIRTMAMIQLLLGNMGMAGGGVNALRGHSNIQGLTDLGLLSTSLPGYMLPSEKQVDLQSYLEANTPKATLADQVNYWSN	FdoG FdnG
501	YPKFFVSMMAFFGDKATAENSWGFDWLPKWDKGYDVLQYFEMMKEGVNGYICQGFNPVASFPPKNKVIKGLSKLKLFLVTIDPLNTEISNFWQNHGELN YPKFFVSLMKSFGDAAQKENNWGYDWPKWDTYDVIKYFNMMDEKVTGYFCQGFNPVASFPPDKNKVVSGLSKLYMVMVIDPLVTETSTFWQNHGESN	FdoG FdnG
601	EVDSSKIQTVEVFRLPSTFAEENGSIIVNSGRWLQWHHWKADAPGIALTDGEILSGIFLRLRKMYAEQGGANPDQVLNMTWNYAIPHEPSSSEEVAMESNGK DVDPASIQTEVFRLPSTFAEEDGSIANSGRWLQWHHWKQDAPGEARNDEILAGIYHHLRELYQSEGGKGVPELKMMSWNYKQPHEPQSDVAKENNGY	FdoG FdnG
701	ALADITDPATGAVIVKKGQQLSSFAQLRDDGTTSCGCIWIFAGSWTPEGNQMARRDNDPSGLGNTLGAWAWPLNRRILYNRASADPQGNPWPDPKRLQLLK ALELDYDANGVLIKKGQQLSSFAHLRDBGTTASSCWIYTGSWTEQGNQMANRDNDSPSGLGNTLGAWAWPLNRRVLYNRASADINGKPWDPKRMLIQW	FdoG FdnG
801	WDGKTWTGWDIPDYSAAPPSSGSGVGFIMQEQGMGRLFALDKMAEGPFPEHYEPFETPLGTNPLHPNVISNPAARIFKDDAEALGKADKFPYVGTYYRLTE NGSKWTGNDIPDFGNAAPGTPPTGPFIMQPEGMGRFAINKMAEGPFPEHYEPIETPLGTNPLHPNVVSNPVVRLYEQDALRMGKKEQFPYVGTYYRLTEH	FdoG FdnG
901	HFHYWTKHALLNAILQPEQFVEIGESLANKLGIAQGDTVKVSNSRGYIKAKAVVTKRIRTLKANGKDIDTIGIPIHWGYEGVAKGFIANTLTPFVGDAN FHTWTKHALLNAILQPEQFVEISETLAAAKGINNGDRVTVSSKRGFIHFAVAVVTRRIKPLNVNGQVETVGIPIHWGFEGVAKGYIANTLTPVNGDANS	FdoG FdnG
1001	TQTPEFKSFLVNVEKV QTPEYKAFLVNIEKA	FdoG FdnG

Abbildung 17: Aminosäuresequenzen der katalytischen Untereinheiten von Fdh-O und Fdh-N. Die Aminosäuresequenzen von FdoG und FdnG wurden gegenübergestellt. Durch MS/MS-Daten bestätigte Peptide sind schwarz, nicht abgedeckte Sequenzbereiche sind grau dargestellt. Nicht modifizierte und nicht identifizierte Cysteine sind orange, identifizierte Selenocysteine sind grün und das natürlich vorkommende Selenocystein an Position 196 ist rot hervorgehoben.

Tabelle 5: Identifizierte Selenocysteine der Fdh-O und Fdh-N. Die Kultivierung des *E. coli*-Stammes MC4100 erfolgte in TGYEP-Medium, das mit 2 μM oder 100 μM Natriumselenit supplementiert wurde. Die Identifizierung der Selenocystein-enthaltenden Peptide erfolgte durch *In-Gel*-Proteinspaltung der Proteine in den Banden I – VIII (Abbildung 15) und anschließender Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS-Analyse.

Protein	Aminosäuresequenz und Modifikationsstelle ^{*1}	<i>m/z</i>		<i>z</i> ^{*2}	[M+H] ⁺	Δm [ppm]	100 μM ^{*3}		2 μM ^{*3}	
		experimentell	theoretisch							
FdoG	LPSTu ₆₁₈ FAEENGSIIVNSGR	993,428		+2	1985,848	1,12	+			
	LPSTa ₆₁₈ FAEENGSIIVNSGR	923,950		+2	1846,893	0,19	+		+	
	Vu ₃₉₁ EYIAETSADHK	785,818		+2	1570,629	0,30	+			
	Va ₃₉₁ EYIAETSADHK	716,341		+2	1431,674	0,09	+		+	
	WLSTGMLa ₁₆₅ ASASSNETGYLTQK	1165,553		+2	2330,099	0,92			+	
	VNGYIa ₅₅₄ QGFNPVASFPNK	960,984		+2	1920,962	1,22			+	
FdnG	WLSTGMLa ₁₆₅ ASGASNETGMLTQK	1134,536		+2	2268,065	0,71			+	
	VTGYFa ₅₅₄ QGFNPVASFPDK	971,971		+2	1942,934	0,66			+	
	LPSTa ₆₁₈ FAEEDGSIANSR	910,426		+2	1819,846	0,29			+	
	Va ₃₉₁ EVLASTSAPDR	657,338		+2	1313,669	0,30			+	
FdnG/	Vu ₁₉₆ HGPTVASLAPTFGR	859,400		+2	1717,793	0,61	+			
FdoG	Va ₁₉₆ HGPTVASLAPTFGR	789,923		+2	1578,840	0,82	+		+	
	YTPDVVENIu ₃₈₀ GTPK	820,857		+2	1640,708	0,87	+			
	YTPDVVENIa ₃₈₀ GTPK	751,380		+2	1501,752	0,34	+		+	
FdoH	LIDVTTa ₃₉₁ GcK	594,817		+2	1188,628	0,87	+			
	LIDVTTcIGu ₄₂ K	664,294		+2	1327,582	0,26	+			
	LIDVTTcIGa ₄₂ K	566,307		+2	1131,607	0,43	+		+	
FdnH	DEVGHa ₆₂ VGVDNPADLSAK	977,961		+2	1954,915	0,83			+	
FdoH/	Au ₄₅ QVAcSEWNDIR	828,821		+2	1656,635	1,10	+			
FdnH	AcQVAa ₄₉ SEWNDIR	759,343		+2	1517,679	0,25	+			
	cTLa ₁₆₀ VDR	416,702		+2	832,398	0,15	+			
	Ta ₁₇₉ PTGAIHFGTK	599,814		+2	1198,620	0,62	+		+	

^{*1} Modifikationen; c: carbamidomethyliertes Cystein; u: carbamidomethyliertes Selenocystein; a: Dehydroalanin;

^{*2} z: Ladungszustand des Vorläuferions; ^{*3} Es wurden zwei Proteinreinigungen aus *E. coli*-MC4100-Zellen durchgeführt, die in Anwesenheit von 2 μM oder 100 μM Natriumselenit kultiviert wurden.

Cystein C12 ist Bestandteil des Tat-Signalpeptids, weshalb dessen Detektion in den vollständig prozessierten und gereinigten Fdhs ausgeschlossen war.

Die Elektronen-übertragenden Untereinheiten FdoH und FdnH enthalten jeweils vier [4Fe-4S]-Zentren, welche durch 16 bzw. 17 Cysteine koordiniert sind [37]. Wie bei der Untersuchung der katalytischen Untereinheiten FdoG und FdnG wurde durch massenspektrometrische Analyse beider Polypeptide auch hier die Substitution bestimmter Cysteine durch Selenocysteine (Tabelle 5) detektiert. FdnH und FdoH zeigen Aminosäureaustausche an den Positionen C45, C49, C160 und C179. Zusätzlich weist FdoH Substitutionen an den Positionen C39 und C42 auf, welche die [4Fe-4S]-Zentren eins bis vier koordinieren [37]. Während bei Kultivierung mit 2 μM Selenit nicht alle Cysteine durch Selenocysteine ersetzt wurden, führte die

Kultivierung mit 100 μM Selenit mit einer Ausnahme (FdnH, C62) zur vollständigen Substitution der enthaltenen Cysteine durch Selenocysteine (Tabelle 5).

Die Analyse des Selenocysteineinbaus in die integralen Membranuntereinheiten FdI und FdN war aufgrund der geringen Sequenzabdeckung von 18% nicht realisierbar. Aufgrund hoher Sequenzidentitäten zwischen Fdh-O und Fdh-N konnte für mehrere Positionen keine eindeutige Zuordnung identifizierter Substitutionsstellen zu den Proteinen FdnG/FdoG (Positionen 196 und 380) und FdnH/FdoH (Positionen 45, 49, 160 und 179) erfolgen (Tabelle 5).

4.1.3 Unspezifischer Einbau von Selenocystein in NarG und OmpA

Nach der Identifizierung entsprechender Modifikationen in Fdh-O und Fdh-N wurden die Polypeptide NarG und OmpA (Tabelle 4) analysiert, was die Beurteilung des unspezifischen Selenocysteineinbaus in Nicht-Selenoproteine ermöglichte.

Das Protein NarG besteht aus 1247 Aminosäuren, wobei die Sequenz 12 Cysteine enthält, von denen drei (C54, C58 und C59) das [4Fe-4S]-Zentrum des Proteins koordinieren [51, 54]. Die Sequenzabdeckung von NarG lag bei ca. 60%. Für vier Cysteine (C292, C493, C838 und C875) wurde ein Austausch gegen Selenocystein detektiert (Tabelle 6). Die Identifizierung dieser Substitutionen war auf gereinigtes Protein aus mit 2 μM Natriumselenit kultivierten Zellen beschränkt.

Tabelle 6: Identifizierte Selenocysteine in Nar und OmpA. Die Kultivierung des *E. coli*-Stammes MC4100 erfolgte in TGYEP-Medium, welches jeweils mit 2 μM oder 100 μM Natriumselenit supplementiert wurde. Die Identifizierung erfolgte durch *In-Gel*-Proteinspaltung der Banden I – VIII (Abbildung 15) und Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS.

Protein	Aminosäuresequenz und Modifikationsstelle* ¹	<i>m/z</i>	<i>z</i> * ²	[M+H] ⁺	Δm	100	2
		experimentell		theoretisch	[ppm]	μM * ³	μM * ³
OmpA	GMGESNPVTGNTu ₃₁₁ DNVK	922,355	+2	1843,704	0,56	+	
	GMGESNPVTGNTa ₃₁₁ DNVK	852,878	+2	1704,749	0,17	+	+
	AALIDa ₃₂₃ LAPDR	562,308	+2	1123,609	0,89		+
NarG	GLNDVNa ₄₉₃ ATSYDDVK	790,366	+2	1579,724	0,84		+
	KFSEVa ₈₃₈ VGHLGK	635,351	+2	1269,695	0,01		+
	FSEVa ₈₃₈ VGHLGK	381,204	+2	1141,598	1,12		+
	La ₂₉₂ DLWLAPK	512,794	+2	1024,582	0,52		+
NarH	IEAGQPTVCSETa ₂₆₃ VGR	808,391	+2	1615,774	0,18		+
NarG/	KGEa ₈₇₅ DLIPGK	513,284	+2	1025,561	0,82		+
NarZ	GEa ₈₇₅ DLIPGK	449,237	+2	897,466	0,80		+

*¹ Modifikationen; u: carbamidomethyliertes Selenocystein; a: Dehydroalanin.

*² *z*: Ladungszustand des Vorläuferions; *³ Es wurden zwei identische Proteinreinigungen aus *E. coli*-MC4100-Zellen durchgeführt, die mit 2 μM oder 100 μM Natriumselenit kultiviert wurden.

Wie in den katalytischen Untereinheiten der Fdhs konnten keine Substitutionen der [4Fe-4S]-Zentren koordinierenden Cysteine gefunden werden. Auffallend ist ebenfalls eine ähnliche Anzahl identifizierter Selenocysteine im Polypeptid NarG im Vergleich zu FdoG und FdnG (4 bzw. 5 Substitutionen, Tabelle 5 und Tabelle 6). Aufgrund der hohen Sequenzidentität von NarG und NarZ konnte für zwei Peptide keine eindeutige Zuordnung der identifizierten Substitutionsstelle zum jeweiligen Protein erfolgen.

Die Analyse von OmpA ergab eine Sequenzabdeckung von ca. 55%. Das 153 Aminosäuren umfassende Polypeptid besitzt zwei Cysteine an den Positionen 311 und 323. Für beide Positionen wurde eine Substitution mit Selenocystein detektiert (Tabelle 6).

4.1.4 *selC*-unabhängiger Einbau von Selenocystein in NarG

Für die Aufklärung des *selC*-unabhängigen unspezifischen Selenocysteineinbaus erfolgte die Analyse von NarG, welches aus der *E. coli selC*-Deletionsmutante FM460 isoliert wurde. Aufgrund einer Mutation in *selC* ist dieser Stamm nicht zur Ausbildung der tRNA^{Sec} fähig, wodurch die vollständige Synthese der Selenoproteine in *E. coli* verhindert wird [65, 66]. Im Gegensatz dazu bleibt die Synthese von NarG unbeeinflusst.

Tabelle 7 zeigt die Identifizierung von vier Selenocystein-haltigen Peptiden in NarG. Diese bestätigen den *selC*-unabhängigen unspezifischen Einbau von Selenocystein in *E. coli*. Zudem stimmen die gefundenen Substitutionen an den Positionen 292, 493 und 838 mit denen von NarG aus dem *E. coli*-Wildtypstamm MC4100 überein (Tabelle 6).

Tabelle 7: Identifizierte Selenocysteine in NarG, die aus einer *E. coli selC*-Mutante isoliert wurden. Die Identifizierung erfolgte durch *In-Gel*-Proteinspaltung der Banden I – VIII (Abbildung 15) und anschließender Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS.

Protein	Aminosäuresequenz und Position der Selenocysteine* ¹	<i>m/z</i>	<i>z</i> ^{*2}	[M+H] ⁺	Δm
		experimentell		theoretisch	[ppm]
NarG	GLNDVNu ₄₉₃ ATSYDDVK	859,843	+2	1718,679	1,20
	GLNDVNa ₄₉₃ ATSYDDVK	790,3666	+2	1579,724	0,78
	KFSEVu ₈₃₈ VGHLGK	704,827	+2	1408,648	0,06
	Lu ₂₉₂ DLWLAPK	582,271	+2	1163,536	0,10

*¹ u: carbamidomethyliertes Selenocystein; a: Dehydroalanin; *² z: Ladungszustand des Vorläuferions

4.1.5 Vergleichende Quantifizierung von Modifizierungsstellen

Für weitere Untersuchungen der dem unspezifischen Seleneinbau zugrunde liegenden Mechanismen wurde eine Quantifizierung von Cysteinen und Selenocysteinen in den bereits untersuchten Polypeptiden durchgeführt. Dazu wurden die PSMs von modifizierten und nicht modifizierten Peptiden ausgezählt und verglichen (*spectral counting*). Selenocysteinmodifikationen wurden durch carbamidomethylierte Selenocysteine und Dehydroalanine repräsentiert. Für das spezifisch eingebaute Selenocystein an Position 196 der FdoG und FdnG konnte kein Cysteineinbau detektiert werden (Abbildung 18). Diese Beobachtung diente als Bestätigung für die verwendete MS-Methode.

Tabelle 8 verdeutlicht die Häufigkeit der auftretenden Selenocysteinmodifikationen. So wurden Selenocystein/Cystein-Verhältnisse von 0,14 bis 2 (unter Ausschluss von Position 196 in FdoG und FdnG) beobachtet. Dabei konnte keine Abhängigkeit der Häufigkeit des unspezifischen Selenocysteineinbaus von einer UGU oder UGC-Kodierung verzeichnet werden.

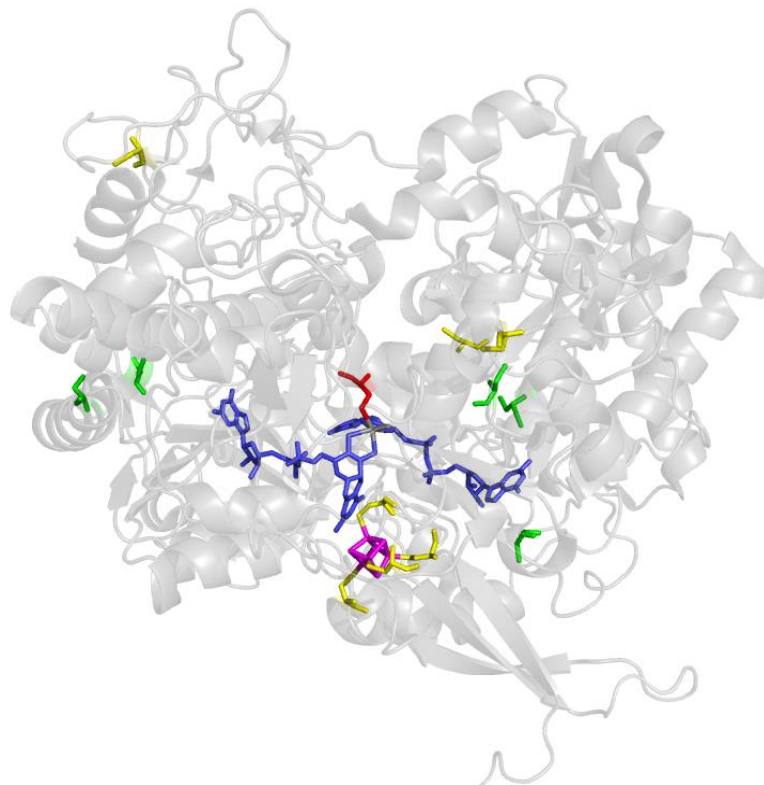


Abbildung 18: Röntgenkristallstruktur der katalytischen Untereinheit von Fdh-N aus *E. coli* (PDB-Eintrag 1KQF). Identifizierte Cysteine sind gelb, Selenocysteine sind grün und rot dargestellt. Das natürlich vorkommende Selenocystein (rot) an Position 196 koordiniert den Mo-bisMGD-Kofaktor (blau). Das [4Fe-4S]-Zentrum wurde in magenta hervorgehoben. Cysteine, Selenocysteine und Kofaktoren sind als *sticks* dargestellt.

Tabelle 8: Vergleichende Quantifizierung von modifizierten und nicht modifizierten Peptiden. Die Identifizierung erfolgte durch *In-Gel*-Proteinspaltung der Banden I – VIII (Abbildung 15) und anschließendem *spectral counting* der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS/MS.

Protein	Aminosäuresequenz und Position der Modifikation* ¹	peptide spectral matches Cystein	peptide spectral matches Selenocystein u. Dehydroalanin	peptide spectral match-Verhältnis* ²
FdoG	LPSTX ₆₁₈ FAEENGSIIVNSGR	4	4	1
	VX ₃₉₁ EYIAETSAHDK	9	4	0,45
	WLSTGMLX ₁₆₅ ASASSNETGYLTQK	-	1	-
	VNGYIX ₅₅₄ QGFNPVASFPNK	2	1	0,5
FdnG	WLSTGMLX ₁₆₅ ASGASNETGMLTQK	-	2	-
	VTGYFX ₅₅₄ QGFNPVASFPDK	-	2	-
	LPSTX ₆₁₈ FAEEDGSIANSGR	2	3	1,5
	VX ₃₉₁ EVLASTSAPDR	1	2	2
FdnG/ FdoG	<u>VU₁₉₆HGPTVASLAPTFGR</u>	-	8	-
FdoG	YTPDVVENIX ₃₈₀ GTPK	9	9	1
FdoH	LIDVTTX ₃₉ IGcK	7	3	0,43
	LIDVTTcIGX ₄₂ K	9	2	0,22
FdnH	DEVGHX ₆₂ VGYYDNPADLSAK	-	2	-
FdoH/ FdnH	AX ₄₅ QVAcSEWNDIR	2	2	1
FdnH	AcQVAX ₄₉ SEWNDIR	1	1	1
	cTLX ₁₆₀ VDR	1	1	1
	TX ₁₇₉ PTGAIHFGTK	6	2	0,33
OmpA	GMGESNPVTGNTX ₃₁₁ DNVK	1	4	4
	AALIDX ₃₂₃ LAPDR	7	1	0,14
NarG	GLNDVNX ₄₉₃ ATSYDDVK	1	2	2
	KFSEVX ₈₃₈ VGHLGK	-	2	-
	FSEVX ₈₃₈ VGHLGK	1	2	2
	LX ₂₉₂ DLWLAPK	1	2	2
NarH	IEAGQPTVCSETX ₂₆₃ VGR	-	2	-
NarG/ NarZ	KGEX ₈₇₅ DLIPGK	-	2	-
NarZ	GEX ₈₇₅ DLIPGK	-	1	-

*¹ Durch Fragmentationenspektren bestätigte Modifikationen werden durch ein X repräsentiert. Der Index definiert die jeweilige Position des Cysteins, carbamidomethylierten Cysteins (c), Dehydroalanins (a) oder carbamidomethylierten Selenocysteins (u) innerhalb der Aminosäuresequenz. U196 ist die natürlich vorkommende Aminosäure in FdoG bzw. FdnG, die Sequenz ist unterstrichen; *² Selenocystein bzw. Dehydroalanin/Cystein-Verhältnis.

Nachdem eine Quantifizierung des unspezifischen Selenocysteineinbaus an UGU- und UGC-kodierten Positionen erfolgt war, wurde eine mögliche Substitution dieser Cysteine durch weitere Aminosäuren untersucht. Hierbei wurde der Austausch eines Cysteins durch Serin an Position 380 in FdoG oder FdnG gefunden, wobei ein einziger PSM für das Peptid YTPDVVENIS₃₈₀GTPK erhalten wurde. Aufgrund der hohen Sequenzidentität beider Proteine konnte für diese identifizierte Substitutionsstelle keine eindeutige Zuordnung zu FdoG oder FdnG erfolgen. Zudem konnten keine Substitutionen durch andere Aminosäuren identifiziert werden.

4.2 Identifizierung von zellwandständigen Interaktionspartnern der Formiatdehydrogenasen Fdh-O und Fdh-N

Die Identifizierung von Bindungspartnern der beiden respiratorischen Formiatdehydrogenasen in *E. coli* erfolgte durch *in vitro*-Cross-Linking und anschließender massenspektrometrischer Analyse. Der experimentelle Ablauf ist in Abbildung 19 schematisch dargestellt. Durch das Einführen chemischer Quervernetzungen in Protein-Protein-Komplexe können niedrig abundante, schwach affine und transiente Interaktionspartner kovalent gebunden und anschließend mittels Massenspektrometrie analysiert werden. *In vitro*-Cross-Linking-Experimente werden häufig durch Inkubation eines gereinigten rekombinanten Fusionsproteins mit Zelllysat durchgeführt [192, 231].

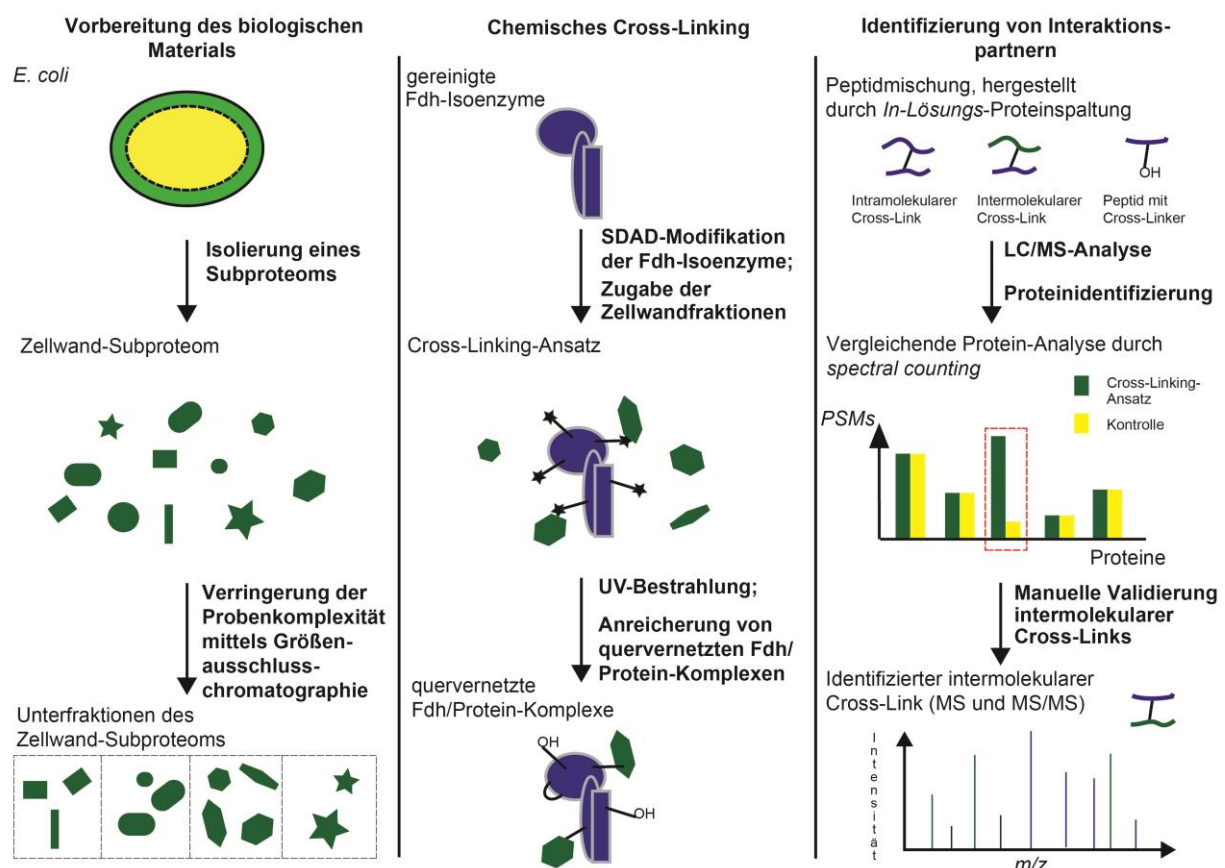


Abbildung 19: Arbeitsablauf für die Identifizierung von Interaktionspartnern der Formiatdehydrogenasen.

Die Anwesenheit hochabundanter Proteine innerhalb dieser sehr heterogenen Proteingemische kann dabei zu falsch positiven Identifizierungen oder einer erschwerten Detektion der meist gering abundant vorliegenden Cross-Linking-Produkte führen. Deshalb wurde ein Zellwand-Subproteom von *E. coli*, welches potentielle Bindungspartner enthält, eingesetzt. Dieses wurde zunächst durch Isolierung des periplasmatischen Kompartiments und der äußeren Membran der Zellwand von *E. coli* gewonnen. Anschließend erfolgte eine weitere Verringerung der Komplexität dieses Zellwand-Subproteoms mittels Größenausschlusschromatographie.

Für die Cross-Linking-Reaktionen wurden die gereinigten Isoenzyme Fdh-O und Fdh-N mit der aminreaktiven Gruppe des heterobifunktionalen Cross-Linkers SDAD modifiziert. Die native Reingung der Fdhs hat den Vorteil, dass zelluläre Prozesse nicht durch die Überproduktion von Proteinen gestört und potentielle Protein-Protein-Interaktionen nicht durch einen Affinitäts-*tag* beeinflusst werden. Anschließend wurde die durch Größenausschlusschromatographie erzeugte Unterfraktion des Zellwand-Subproteoms mit den SDAD-modifizierten Fdhs inkubiert und die Quervernetzung durch UV-A-Bestrahlung induziert.

Die Anreicherung der quervernetzten Fdh/Protein-Komplexe erfolgte mittels Größenausschlusschromatographie unter nativen Bedingungen. Die Elutionsfraktionen wurden gesammelt und Fdh/Protein-Komplex-enthaltende Fraktionen einer Proteinspaltung in Lösung unterzogen. Die komplexen Peptidmischungen wurden durch Nano-HPLC getrennt und die Cross-Linking-Produkte mittels *online* gekoppelter Nano-ESI-LTQ-Orbitrap-Massenspektrometrie analysiert. Die Identifizierung potentieller Interaktionspartner erfolgte durch eine vergleichende Proteinanalyse zwischen dem Cross-Linking-Ansatz und einer nicht-quervernetzten Kontrolle. Identifizierte Proteine wurden mittels *spectral counting* quantifiziert und potentielle Interaktionspartner anhand einer höheren Anzahl erhaltener PSMs im Cross-Linking-Ansatz gegenüber der Kontrolle detektiert.

Anschließend erfolgte die Identifizierung und Validierung intermolekularer Vernetzungsprodukte, mit deren Hilfe zum einen Interaktionspartner identifiziert und zum anderen Interaktionsregionen innerhalb von Fdh/Zellwandprotein-Komplexen aufgeklärt wurden.

4.2.1 Isolierung und Verringerung der Komplexität eines Zellwand-Subproteoms

Für die Herstellung eines Zellwand-Subproteoms wurden *E. coli* MC4100-Zellen einer Kombination aus osmotischem Schock und Lysozym-EDTA-Behandlung unterzogen. Es folgte eine weitere Verringerung der Komplexität durch Größenausschlusschromatographie für die anschließende Quervernetzung mit den gereinigten Fdhs. Dieses Zellwand-Subproteom sowie die durch Größenausschlusschromatographie gewonnenen Unterfraktionen wurden mittels SDS-PAGE analysiert (Abbildung 20), welche eine eindeutige Trennung der Proteine entsprechend ihrer hydrodynamischen Radien zeigt.

Es folgte eine qualitative und quantitative Beurteilung des Zellwand-Subproteoms. Dazu wurden die Proteine mittels SDS-PAGE getrennt (Abbildung 20, Spur L), einer *In-Gel*-Proteinspaltung unterzogen und die erhaltenen Peptidgemische durch Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS analysiert. Die Abundanz identifizierter Proteine wurde mit Hilfe von *spectral counting* abgeschätzt (Tabelle 9).

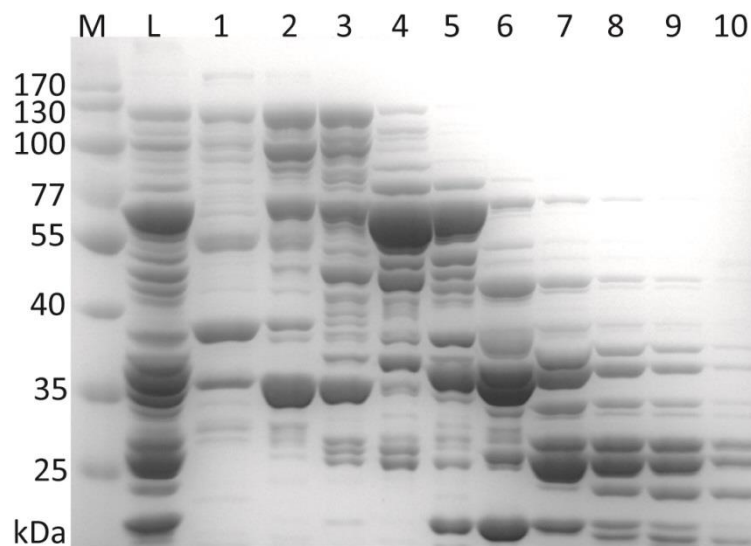


Abbildung 20: SDS-PAGE-Analyse des mittels Größenausschlusschromatographie fraktionierten Zellwand-Subproteoms. Die Fraktionierung erfolgte unter Verwendung einer Superdex S75 16/600 GL-Säule. Das Zellwand-Subproteom wurde aus *E. coli* MC4100-Zellen gewonnen. Die Auftrennung der Proteine erfolgte mittels eines 12%-igen Trenngels. Die Färbung der Proteine erfolgte mit kolloidaler Coomassie-Brillantblau-Lösung. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur L: Proteinlösung vor der chromatographischen Trennung; Spuren 1 - 10: Elutionsfraktionen 1 - 10.

Die massenspektrometrische Analyse des Zellwand-Subproteoms führte zur Identifizierung von 129 periplasmatischen Proteinen, 12 Proteinen der äußeren Membran, 20 Proteinen der inneren Membran, 158 cytoplasmatischen Proteinen und 17 Proteinen deren zelluläre Lokalisation derzeit nicht charakterisiert ist (Tabelle 9 und Tabelle A1).

Die Anzahl der PSMs an cytoplasmatischen Proteinen beträgt 22,1% (1721). Mit 74,4% (5783) stammte ein großer Anteil der PSMs von Proteinen der Zellwand (periplasmatische, innere und äußere Membranproteine). 3,4% der PSMs (266) konnten keinem zellulären Kompartiment zugeordnet werden. Diese Daten belegen die Anwesenheit hochabundanter periplasmatischer Proteine innerhalb der Zellwandpräparation und sind vergleichbar mit dem Ergebnis einer früheren Studie [232], in der die Zellwandpräparation mit Hilfe der TSE (Tris-Sucrose supplementiert mit EDTA)-Methode durchgeführt wurde.

Identifizierte hochabundante Proteine des Zellwand-Subproteoms sind in Tabelle 10 exemplarisch aufgelistet. Unter den 100 häufigsten Proteinen wurden nur 19 Proteine dem cytoplasmatischen Kompartiment zugeordnet (Tabelle A1).

Tabelle 9: Klassifizierung identifizierter Proteine des Zellwand-Subproteoms. Die Proteine wurden mittels SDS-PAGE getrennt und durch *In-Gel*-Proteinspaltung mit anschließender Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS-Analyse identifiziert (Abbildung 20, Spur L). Identifizierte Proteine wurden ihrer subzellulären Lokalisation zugeordnet. Die Proteinmengen wurden durch *spectral counting* abgeschätzt.

Subzelluläre Lokalisation	Anzahl identifizierter Proteine	Peptide spectral matches (PSMs)	Verteilung PSMs [%]
Periplasma	129	5381	69,3
Äußere Membran	12	122	1,6
Innere Membran	20	280	3,6
Cytoplasma	158	1721	22,1
Nicht charakterisiert	17	266	3,4
Gesamt	336	7776	100

81 dieser Proteine konnten dem Periplasma sowie der inneren und äußeren Membran der Zellwand zugeordnet werden. Diese Daten zeigen die erfolgreiche Anreicherung von Zellwandproteinen mit einem Hauptanteil periplasmatischer Proteine und einem geringen Anteil niedrig abundanter cytoplasmatischer Kontaminanten.

Tabelle 10: Hochabundante Proteine des Zellwand-Subproteoms. Die Proteine wurden mittels SDS-PAGE (Abbildung 20, Spur L) getrennt und durch *In-Gel*-Proteinspaltung mit anschließender Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS-Analyse identifiziert. Die UniProtKB-Einträge identifizierter Proteine sind angegeben.

UniProt- KB- Eintrag	Protein	Gen	Lokalisation	PSMs
P37329	Periplasmatisches Molybdat-bindendes Protein	<i>modA</i>	Periplasma	357
P0AFH9	Osmotisch-induzierbares Protein Y	<i>osmY</i>	Periplasma	267
P23843	Periplasmatisches Oligopeptid-bindendes Protein	<i>oppA</i>	Periplasma	249
P24183	Formiatdehydrogenase-N, α -Untereinheit	<i>fdnG</i>	Periplasma	166
P33363	Periplasmische β -Glucosidase	<i>bglX</i>	Periplasma	149
P0AFM3	Periplasmatisches Glycin/Betain-bindendes Protein	<i>proX</i>	Periplasma	144
P45523	FKBP-Peptidyl-Prolyl cis-trans-Isomerase	<i>fkpA</i>	Periplasma	141
P23865	Schweif-spezifische Protease	<i>prc</i>	Periplasma	127
P0C0V1	Periplasmische Serinendoprotease	<i>degP</i>	Periplasma	122
P16700	Thiosulfat-bindendes Protein	<i>cysP</i>	Periplasma	119
P09373	Pyruvat-Formiatlyase	<i>pflB</i>	Cytosol	110
P0AEN0	Periplasmatisches Cystin-bindendes Protein	<i>fliY</i>	Periplasma	109
P07024	Protein UshA	<i>ushA</i>	Periplasma	107
P0ABZ8	Chaperon SurA	<i>surA</i>	Periplasma	106
A7ZKF2	Glucan-Biosyntheseprotein	<i>mdoG</i>	Periplasma	95
P05458	Protease 3	<i>ptrA</i>	Periplasma	89
A1A7M1	Äußerer Membranprotein-Assemblierungsfaktor	<i>bamA</i>	Periplasma	86
A7ZSL4	Elongationsfaktor Tu 1	<i>tuf1</i>	Cytosol	82
P39099	Periplasmische pH-abhängige Serinendoprotease	<i>degQ</i>	Periplasma	82
C4ZU48	Periplasmatische Serinendoprotease	<i>napA</i>	Periplasma	80
A7ZJC2	Protein TolB	<i>tolB</i>	Periplasma	80

4.2.2 Chemisches Cross-Linking

Nach der erfolgreichen Herstellung eines Zellwand-Subproteoms mit hochangereicherten periplasmatischen Proteinen und mittels Größenausschlusschromatographie erzeugten Unterfraktionen erfolgte das chemische Cross-Linking mit dem heterobifunktionalen Cross-Linker SDAD. Dazu wurden die gereinigten Fdhs zunächst mit der aminreaktiven Gruppe des Cross-Linkers an Lysinen, aber auch Serinen, Threoninen und Tyrosinen modifiziert. Um eventuell durch die Modifikationen hervorgerufenen strukturellen Veränderungen der Fdhs auszuschließen, erfolgte die Analyse der Formiat:DCPIP-Oxidoreduktase-Enzymaktivität. Die Enzymaktivität der SDAD-modifizierten Fdhs (Abbildung 21A) bestätigte die strukturelle Integrität der katalytischen Untereinheiten der Fdhs.

Anschließend wurde eine Fraktion der Zellwandproteine in einem Molekulargewichtsbereich zwischen 10 – 70 kDa hergestellt. Hierfür wurden die Fraktionen 5 und 6 der Größenausschlusschromatographie (Abbildung 20; Spuren 5 und 6) vereint und für das anschließende Cross-Linking zu den SDAD-modifizierten Fdhs zugegeben. Nach der Quervernetzung durch UV-A-Bestrahlung wurde auf die Anwesenheit von Fdh/Protein-Cross-Linking-Produkten mittels SDS-PAGE geprüft.

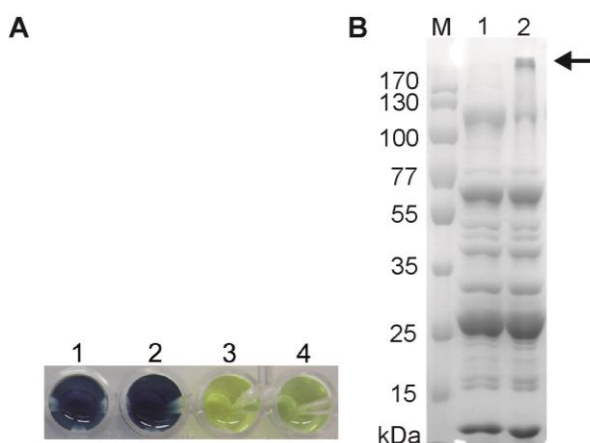


Abbildung 21: Qualitative Analyse der SDAD-Modifikation der Fdhs und der chemischen Quervernetzung.

(A) Formiat:DCPIP-Oxidoreduktase-Aktivitätstest. (1) Reaktionsansatz ohne Formiat; (2) Reaktionsansatz ohne Enzym; (3) Reaktionsansatz mit gereinigten Fdhs; (4) Reaktionsansatz mit SDAD-modifizierten Fdhs nach Abtrennung von überschüssigem SDAD-Cross-Linker. (B) SDS-PAGE-Analyse der chemischen Quervernetzung nach UV-A Bestrahlung (8.000 mJ/cm^2). Die Auftrennung der Proteine wurde in einem 12%-igen Trenngel durchgeführt. Die Färbung der Proteine erfolgte mit mit kolloidaler Coomassie-Brillantblau-Lösung. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur 1: Cross-Linking-Ansatz mit 50 mM β -Mercaptoethanol; Spur 2: Cross-Linking-Ansatz ohne Reduktionsmittel. Das durch die Quervernetzungsreaktion erzeugte Signal ist durch einen Pfeil hervorgehoben.

Wie die Analyse (Abbildung 21B) zeigt, führte das Cross-Linking zur Entstehung eines eindeutigen Signals im Molekulargewichtsbereich oberhalb von 170 kDa, das Fdh/Protein-Cross-Linking-Produkten entsprechen könnte (Spur 2). Die gebildeten Komplexe konnten durch Reduktion der Disulfidbrücke des spaltbaren Cross-Linkers wieder getrennt werden (Spur 1).

4.2.3 Anreicherung der Fdh/Protein-Komplexe

Die Anreicherung der kovalent verknüpften Fdh/Protein-Komplexe erfolgte durch Größenausschlusschromatographie mit einer Superdex S200 16/300 GL-Säule. Die Elutionsprofile des SDAD-Cross-Linking-Ansatzes und der nicht-quervernetzten Kontrolle wurden für die Bestimmung der Retentionszeiten von Quervernetzungsprodukten übereinander gelegt. Das Chromatogramm des SDAD-Cross-Linking-Ansatzes zeigt gegenüber der Kontrolle höhere Absorptionswerte bei 280 nm innerhalb des Retentionsvolumens zwischen 8 und 13 ml (Abbildung 22). Dieses Ergebnis weist auf die Entstehung von Molekülen mit höheren Molekulargewichten, wie z. B. Fdh/Protein-Komplexen, durch chemisches Cross-Linking hin.

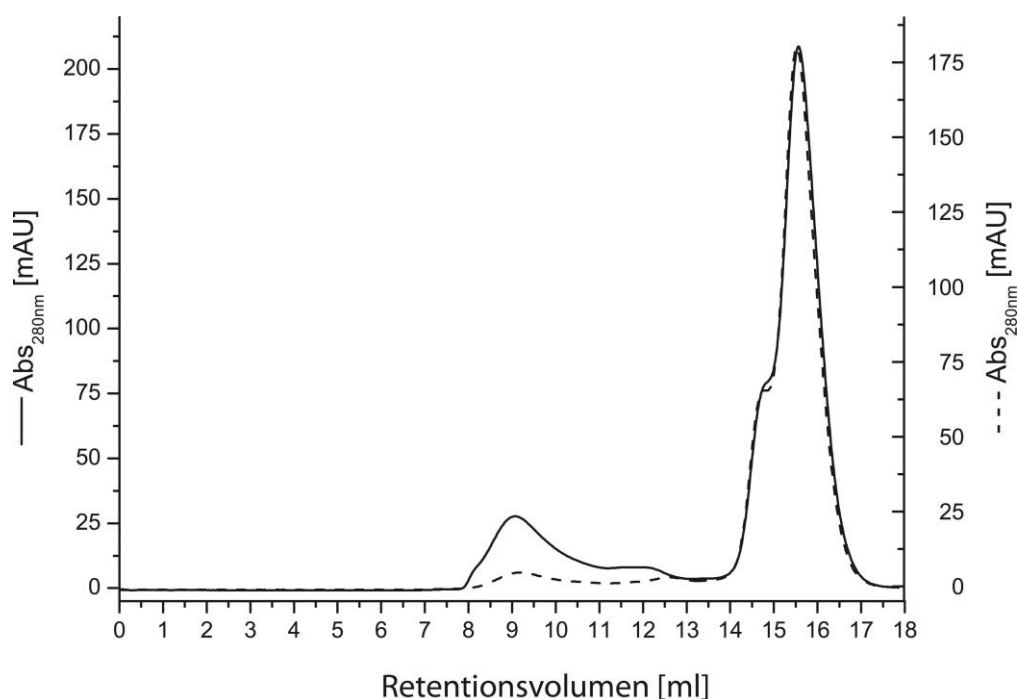


Abbildung 22: Anreicherung von Quervernetzungsprodukten mittels Größenausschlusschromatographie. Die Abbildung zeigt die Chromatogramme des SDAD-Reaktionsansatzes (schwarze durchgängige Linie) und der nicht-quervernetzten Probe (Kontrolle, gestrichelte Linie).

Das Hauptsignal im Chromatogramm (Retentionsvolumen von 14 – 17 ml) entspricht Proteinen der Zellwandpräparation, welche nicht kovalent an die Fdhs gebunden wurden. Das Elutionsprofil des Cross-Linking-Ansatzes zeigt eine klare Trennung der Fdh/Protein-Komplexe von nicht-quervernetzten Proteinen. Die Anreicherung von Fdh/Protein-Komplexen bildete die Basis für die erfolgreiche Identifizierung von intermolekularen Cross-Links zwischen Fdhs und Bindungspartnern. Zudem wurde durch die Anreicherung der Komplexe das Risiko der Maskierung durch nicht-quervernetzte Proteine während der anschließenden LC/MS-Analyse vermindert.

4.2.4 Identifizierung von Bindungspartnern der Fdhs

Für die Identifizierung von Bindungspartnern wurde eine *In-Lösungs*-Proteinspaltung der angereicherten Fdh/Protein-Komplexe durchgeführt. Quervernetzte Peptide wurden mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS/MS analysiert. Bindungspartner der Fdhs wurden durch eine vergleichende Proteinanalyse des Cross-Linking-Ansatzes und der Kontrolle identifiziert. Quervernetzte Fdh-Interaktionspartner wurden hierbei durch eine höhere Anzahl detektierter PSMs vergleichbarer Proteine im Cross-Linking-Ansatz gegenüber der Kontrolle identifiziert. Die Mehrheit der identifizierten Proteine zeigte eine einheitliche oder nur leicht variierende Anzahl PSMs zwischen dem Cross-Linking-Ansatz und der Kontrolle und diente somit als nützlicher Vergleich (Tabelle 11 und Tabelle A2). Die α -Untereinheit der respiratorischen Nitratreduktase I (NarG), der gering leitfähige mechanosensitive Kanal MscS und das periplasmatische Chaperon Skp zeigten höhere Verhältnisse der PSMs von Cross-Linking-Ansatz, zur Kontrolle (Tabelle 11). Für das lösliche periplasmatische Protein ErfK (eine putative L, D-Transpeptidase) und das Oligopeptid-bindende Protein OppA wurden ebenfalls höhere Verhältnisse der PSMs im Cross-Linking-Ansatz beobachtet. Aufgrund der Anreicherung dieser periplasmatischen und inneren Membranproteine im Cross-Linking-Ansatz wurde eine detaillierte Untersuchung der Fdh/Protein-Interaktionsregionen dieser Komplexe durchgeführt. Das Protein TolB zeigte ebenfalls ein erhöhtes PSM-Verhältnis zwischen dem Cross-Linking-Ansatz und der Kontrolle innerhalb des Retentionsvolumens zwischen 8 und 10 ml der Größenausschlusschromatographie (Daten nicht gezeigt).

Tabelle 11: Identifizierte Bindungspartner der Fdhs. Die Elutionsfraktionen der Größenausschlusschromatographie (Superdex S200 16/300 GL-Säule) innerhalb des Retentionsvolumens von 8 - 13 ml (Abbildung 22) wurden vereint und in Lösung gespalten. Die Peptidgemische wurden mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS/MS analysiert. Die UniProtKB-Einträge identifizierter Proteine sind angegeben.

UniProtKB- Eintrag	Protein	PSMs		PSM- Verhältnis*
		Cross- Linking- Ansatz	Kontrolle	
P24183	Formiatdehydrogenase-N, α -Untereinheit	158	160	0,98
P32176	Formiatdehydrogenase-O, α -Untereinheit	77	77	1
P0C0S2	Gering leitfähiger mechanosensitiver Kanal	75	23	3,26
	MscS			
P0A9M8	Phosphatacetyltransferase	54	57	0,94
A7ZTU4	ATP Synthetase, β -Untereinheit	53	58	0,91
P09152	Respiratorische Nitratreduktase I, α-Untereinheit NarG	45	10	4,5
A7ZHA4	Chaperon DnaK	35	32	1,09
P0AAJ3	Formiatdehydrogenase-N, β -Untereinheit	34	34	1
P23843	Periplasmatisches Oligopeptid-bindendes Protein OppA	31	19	1,63
P0AEU9	Chaperon Skp	31	0	-
P63388	ATP-bindendes Phospholipidimportprotein	28	24	1,16
P39176	Putative L, D-Transpeptidase ErfK	27	14	1,92
A1AJ51	60 kDa Chaperon I	26	23	1,13
P0A9Q8	Alkoholdehydrogenase	23	28	0,82

* Es ist das Verhältnis der gemessenen *peptide spectral matches* vom Cross-Linking-Ansatz zur Kontrolle angegeben.

4.2.5 Aufklärung der Interaktionsregionen von Fdh/Protein-Komplexen

Nachdem mögliche Bindungspartner der Fdhs identifiziert wurden, erfolgte deren Bestätigung durch die massenspektrometrische Analyse intermolekularer Cross-Links. Es wurden 13 intermolekulare Cross-Linking-Produkte zwischen den Fdhs und den Zellwandinteraktionspartnern Skp, OppA, ErfK und TolB identifiziert (Tabelle 12), wobei es sich um vier Cross-Links zwischen FdnH und Skp, drei zwischen FdnG und OppA und einen Cross-Link zwischen FdnG und ErfK handelt. Weiterhin wurden drei Cross-Links zwischen FdnH und TolB und zwei Cross-Links zwischen FdnG und TolB identifiziert. Aminosäuresequenzidentitäten von 76% bzw. 75% zwischen den G- und H-Untereinheiten der Fdh-O und Fdh-N führten bei drei Cross-Linking-Produkten nicht zu einer eindeutigen Aussage, ob Fdh-N oder Fdh-O mit dem jeweiligen Interaktionspartner quervernetzt wurde.

Tabelle 12: Identifizierte intermolekulare Vernetzungsprodukte zwischen Fdhs und Bindungspartnern.

<i>m/z</i>	[M+H] ⁺	<i>z</i> ^{*1}	Δm	Fdhs	modifizierte AS ^{*2}	Interaktionspartner	modifizierte AS ^{*3}
experimentell	theoretisch		[ppm]				
645,886	1290,764	+2	+0,5	FdnH	K202	Skp	T118, R119
652,093	2605,351	+4	-0,1	FdnH	K204	Skp	R59
733,354	2198,044	+3	+1,4	FdnH/FdoH	K110	Skp	M74
1260,684	2520,350	+2	+4,4	FdnH/FdoH	K97	Skp	T81, K82
483,229	1929,897	+4	+0,8	FdnG/FdoG	S125	OppA	A381, D382
835,401	2504,188	+3	+0,8	FdnG	S183	OppA	Y121
877,082	2629,239	+3	+2,6	FdnG	Y76	OppA	D167
479,570	1436,687	+3	+6,1	FdnH	K32	TolB	Q282
637,834	2548,305	+4	+2,8	FdnG	S783	TolB	L262
674,687	2022,043	+3	+1,7	FdnH	K190	TolB	L156
740,657	2959,586	+4	+6,6	FdnH	K32	TolB	Q229,T230,L231
885,997	1770,987	+2	-0,4	FdnG	K44	TolB	A412, R413
567,289	2266,139	+4	-2,7	FdnG	K138	ErfK	V235, K236

*¹ Ladungszustand des Vorläuferions*² Reaktionstellen in Fdhs*³ Reaktionstellen im Interaktionspartner

Mit Ausnahme von TolB wurden die Quervernetzungen von Skp, OppA und ErfK entweder zur G- oder H-Untereinheit von Fdh-O und Fdh-N identifiziert. Es wurden keine Cross-Linking-Produkte mit MscS und NarG gefunden.

Nano-ESI-LTQ-Orbitrap-MS/MS-Daten der intermolekularen Cross-Linking-Produkte von FdnH mit Skp und FdnG mit OppA sind in Abbildung 23 exemplarisch gezeigt. Die MS/MS-Daten des Cross-Linking-Produktes von FdnH/Skp (*m/z* 1260,684) umfassen die Aminosäuren L91 – K97 (FdnH) und T81 - R91 (Skp), wobei die Quervernetzung zwischen K97 (FdnH) und T81 oder K82 (Skp) eindeutig identifiziert wurde (Abbildung 23A). Die MS/MS-Daten des Cross-Linking-Produktes von FdnG/OppA (*m/z* 877,082) umfassen die Aminosäuren G67 – E79 (FdnG) und A164 – E171 (OppA). Dabei wurde die Quervernetzung zwischen Y76 (FdnG) mit D167 (OppA) eindeutig identifiziert (Abbildung 23B). Eine Zusammenfassung der MS/MS-Daten aller intermolekularer Cross-Linking-Produkte der Fdh/Protein-Komplexe befindet sich im Anhang (Abbildung A 14).

Abbildung 24 zeigt die durch SDAD modifizierte Aminosäuren der FdnH/Skp- und FdnG/OppA-Komplexe. Die nähere Betrachtung der quervernetzten Aminosäuren für die Proteinstrukturen von Skp und FdnH zeigt die definierte Anordnung der modifizierten Bereiche (Abbildung 24A). Die Proteinstruktur von FdnH zeigt vier modifizierte Lysine (K97, K110, K202 und K204), im zentralen Bereich von Skp,

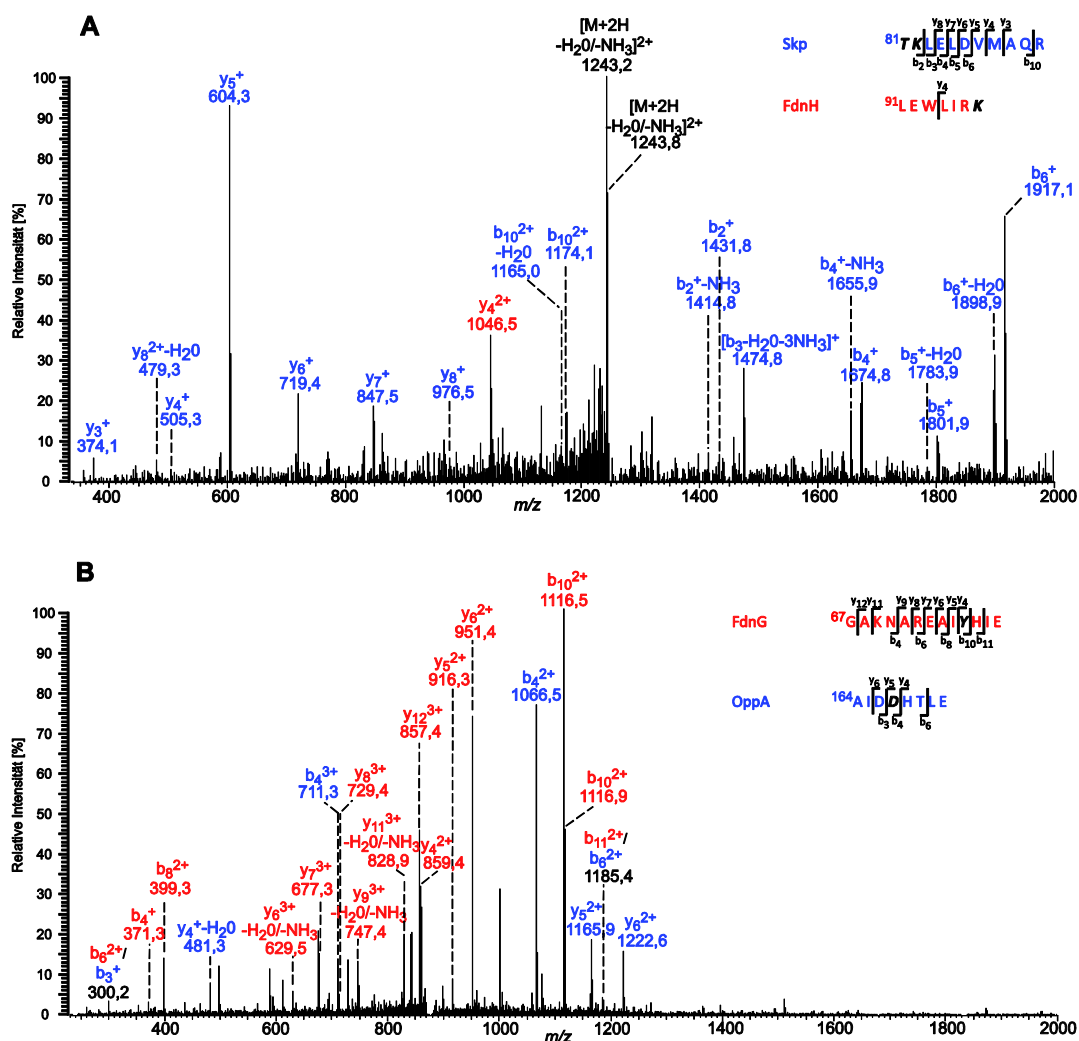


Abbildung 23: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FdnH und Skp sowie FdnG und OppA. Vorläuferionen wurden innerhalb der linearen Ionenfalle selektiert und mittels CID fragmentiert. Peptide und Fragmentationen von FdnG und FdnH sind in rot dargestellt. Peptide und Fragmentationen von Skp und OppA sind in blau dargestellt. **(A)** Fragmentationenspektrum eines doppelt geladenen Vorläuferions (m/z 1260,684). Das Cross-Linking-Produkt umfasst die Aminosäuren 91 – 97 von FdnH und die Aminosäuren 81 – 91 von Skp, wobei K97 von FdnH und T81 (oder K82) von Skp mit SDAD quervernetzt sind. **(B)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 877,082). Das Cross-Linking-Produkt umfasst die Aminosäuren 67 – 79 von FdnG und die Aminosäuren 164 – 171 von OppA, wobei Y76 von FdnG und D167 von OppA mit SDAD quervernetzt ist.

die mit SDAD modifiziert wurden. Hingegen wurden keine modifizierten Aminosäuren im α -helikalen membranintegralen C-Terminus (*C-tail*) von FdnH gefunden. Für das Homotrimer Skp konnte nicht differenziert werden, in welcher der drei Untereinheiten die jeweiligen Aminosäuren modifiziert wurden. Die Proteinstrukturen des FdnG/OppA-Komplexes zeigen ebenfalls eine gerichtete Orientierung der modifizierten Aminosäuren. Im Gegensatz dazu zeigt der FdnGH/ToIB-Komplex eine

breitere Verteilung der modifizierten Aminosäuren über die jeweiligen Proteinstrukturen (Abbildung A 15). Allerdings befinden sich die modifizierten Aminosäuren nur auf einer Seite des FdhGH-Dimers.

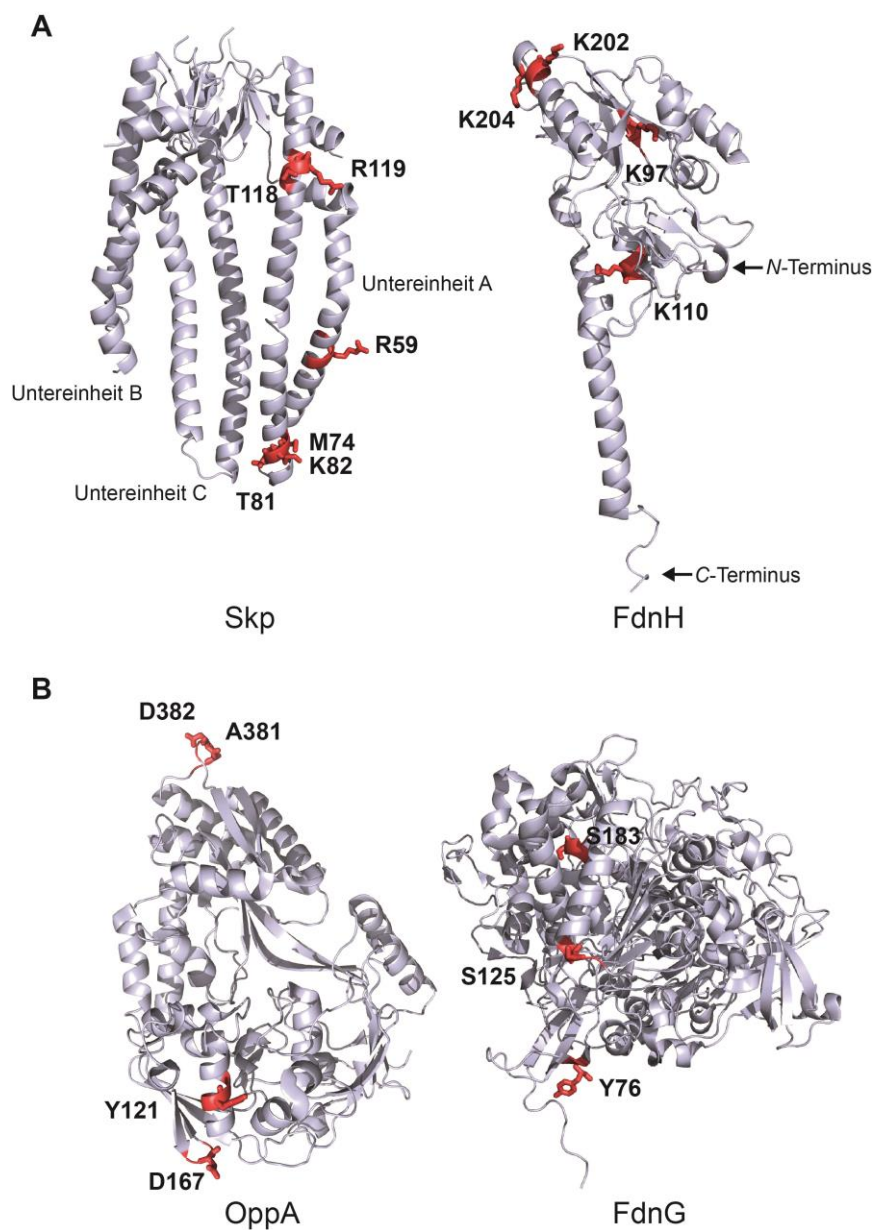


Abbildung 24: Intermolekulare SDAD-Cross-Links innerhalb der FdnH/Skp- und FdnG/OppA-Komplexe.

Die Strukturen von FdnG, FdnH, Skp-Trimer und OppA sind als Bänderdiagramme dargestellt (PDB-Einträge: 1KQF, 1SG2 und 3TCF). Identifizierte Quervernetzungsstellen sind als rote *sticks* hervorgehoben. **(A)** FdnH/Skp-Cross-Linking-Produkte; vier Cross-Links zwischen Skp (T118 oder R119, R59, M74 und T81 oder K82) und FdnH (K202, K204, K110 und K97) wurden identifiziert. Die identifizierten Quervernetzungsstellen innerhalb von Skp sind für Untereinheit A dargestellt. **(B)** OppA/FdnG-Cross-Linking-Produkte; drei Cross-Links zwischen OppA (A381 oder D382, Y121 und D167) und FdnG (S125, S183 und Y76) wurden identifiziert. Die Sequenzpositionen der quervernetzten Aminosäuren sind in Tabelle 12 aufgeführt.

4.3 Charakterisierung der FocA/PflB-Interaktion

Das Formiat-bildende Glycylradikalenzym PflB und der Formiatkanal FocA sind genetisch und physiologisch miteinander gekoppelt [12, 15, 16]. *In vivo*- und *in vitro*-Experimente, welche am Institut für Biologie in der Abteilung Mikrobiologie der Martin-Luther-Universität Halle-Wittenberg durchgeführt wurden, identifizierten eine Interaktion dieser beiden Schlüsselproteine des Formiatmetabolismus. Die Experimente deuten auf eine große Bedeutung des *N*-Terminus von FocA für die Interaktion mit PflB hin. Desweiteren wurde eine regulatorische Funktion von PflB für den Formiat-Transport nachgewiesen. Zur Aufklärung der zugrunde liegenden Mechanismen einer regulierten Formiattranslokation erfolgte in der vorliegenden Arbeit die strukturelle Untersuchung des FocA/PflB-Komplexes.

4.3.1 Analyse der FocA/PflB-Interaktion mittels chemischem Cross-Linking

4.3.1.1 Identifizierung der Interaktionspartner von FocA

Zur Bestätigung von PflB als Interaktionspartner von FocA und für die Identifizierung weiterer Bindungspartner wurden *in vitro*-Cross-Linking-Experimente durchgeführt. Der experimentelle Ablauf ist in Abbildung 25 schematisch dargestellt. Zunächst wurde das gereinigte FocA-Strep-tag-II-Fusionsprotein mit der aminreaktiven Gruppe des Cross-Linkers SDAD modifiziert. Nach der Abtrennung des überschüssigen Cross-Linkers wurde eine Fraktion löslicher Proteine des *E. coli*-Stammes REK702 hinzugegeben, der eine erhöhte Synthese von FocA und PflB zeigt [12]. Nach Induktion der Cross-Linking-Reaktion an der photoreaktiven Diazirinegruppe des SDAD durch UV-Bestrahlung wurden FocA/Protein-Komplexe mittels Streptactin-Affinitätschromatographie gereinigt. Als Kontrolle wurde die Affinitätschromatographie mit einer nicht-quervernetzten Fraktion bestehend aus FocA-Strep-tag-II-Fusionsprotein (nicht SDAD-modifiziert) und löslichen Proteinen des *E. coli*-Stammes REK702 durchgeführt. Für die Beurteilung der Cross-Linking-Reaktion wurden die gereinigten FocA/Protein-Komplexe und die Kontrollprobe mittels SDS-PAGE unter nicht-reduzierenden Bedingungen analysiert.

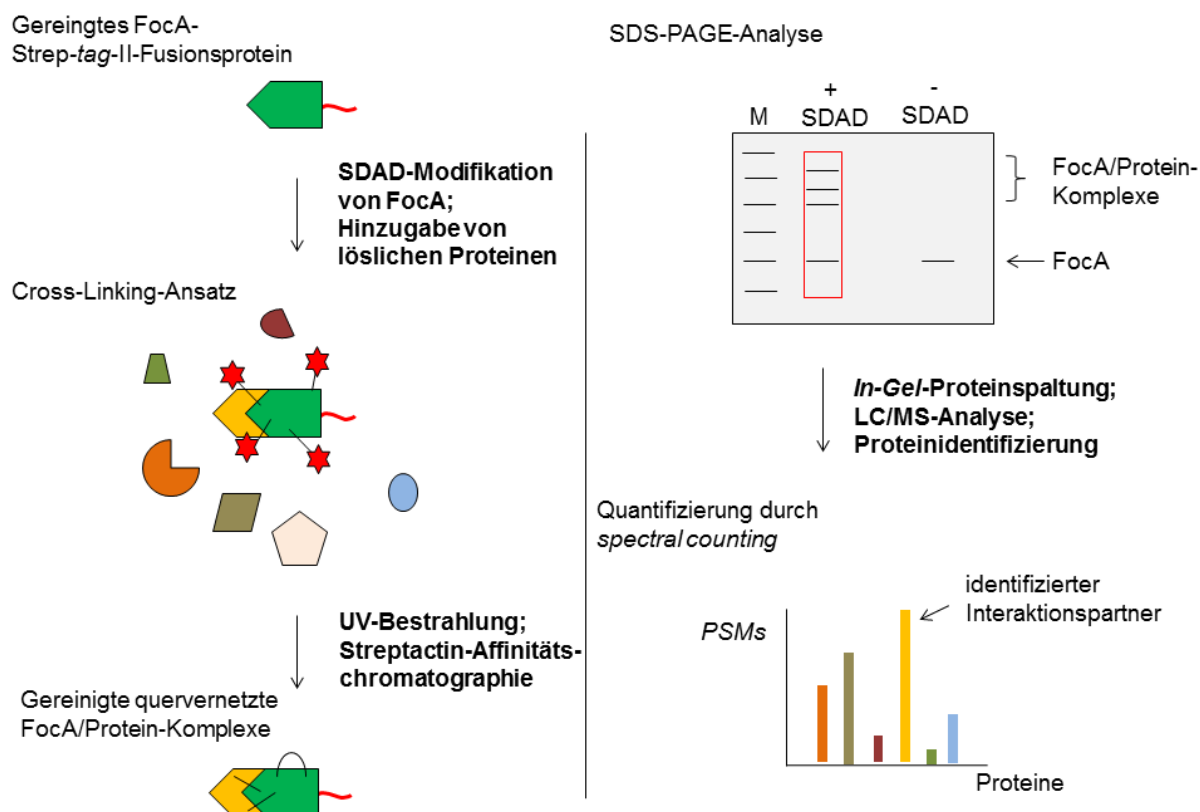


Abbildung 25: Arbeitsablauf für die Identifizierung der Interaktionspartner von FocA.

Die Cross-Linking-Reaktion zwischen FocA und löslichen *E. coli*-Proteinen führte zu zahlreichen Proteinbanden (Abbildung 26, Spur 4), welche über einen Molekularmassenbereich von ungefähr 28 bis 250 kDa detektiert wurden. Die im Cross-Linking-Ansatz detektierten Signale bei apparenten Molekularmassen von ca. 28, 55 und 72 kDa implizieren eine Oligomerisierung von FocA. Die visuelle Beurteilung von Proteinbanden im höheren Molekularmassenbereich wurde durch die unzureichende elektrophoretische Trennung der FocA/Protein-Komplexe erschwert. Die elektrophoretische Trennung unter reduzierenden Bedingungen führte zu einer geringfügigen Abschwächung der Proteinbanden im Molekularmassenbereich zwischen 36 und 250 kDa und zu einer Verstärkung des Signals bei ca. 28 kDa, das dem monomeren FocA entspricht (Abbildung A 16). Durch die Spaltung der Disulfidbrücke im SDAD wurden die quervernetzten FocA/Protein-Komplexe bzw. oligomerisiertes FocA vermutlich gespalten. Im Gegensatz zum Cross-Linking-Ansatz zeigt die Kontrolle eine klare Proteinbande mit einer apparenten Molekularmasse von ungefähr 28 kDa. Zusätzlich ist eine diffuse und schwache Bande oberhalb von

250 kDa zu erkennen. Die massenspektrometrische Analyse dieser Proteinbande führte u. a. zur Identifizierung von PflB und FocA.

Die Identifizierung der Bindungspartner von FocA erfolgte zunächst durch die *In-Gel*-Proteinspaltung der gesamten Spur 4 (Abbildung 26) und ihrer enthaltenen Proteine. Die Analyse der Peptidgemische wurde mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS durchgeführt. Eine grobe Quantifizierung der identifizierten Proteine erfolgte durch *spectral counting*.

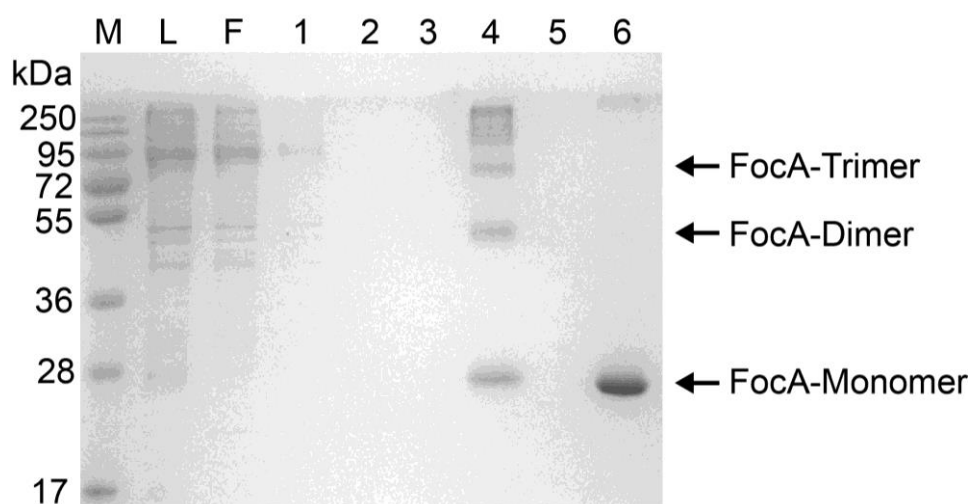


Abbildung 26: SDS-PAGE-Analyse von FocA/Protein-Komplexen. Die Reinigung der FocA/Protein-Komplexe erfolgte durch Streptactin-Affinitätschromatographie. Die Auftrennung der Proteine erfolgte unter nicht-reduzierenden Bedingungen in einem 12%-igen Trenngel. Die Färbung der Proteine erfolgte mit kolloidaler Coomassie-Brillantblau-Lösung. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur L: Proteinlösung vor der Reinigung (*Load*); Spur F: Durchlaufraction (*Flow through*); Spuren 1 - 3: Waschfraktionen; Spur 4: Quervernetzte FocA/Protein-Komplexe nach Affinitätschromatographie; Spur 5: Spur nicht beladen; Spur 6: FocA-Strep-tag-II-Fusionsprotein und lösliche Proteine des *E. coli*-Stammes REK702 (ohne Cross-Linker).

Eine hohe Anzahl an PSMs wurde für FocA, PflB, Alkoholdehydrogenase und Glutamatdecarboxylase erhalten (Tabelle 13 und Tabelle A3). Auffällig für FocA ist, dass ausschließlich Peptide des *N*-terminalen Bereichs identifiziert wurden. Infolgedessen ergibt sich die relativ geringe Sequenzabdeckung von 14,4%. Die fehlende Sequenzabdeckung für weitere Bereiche von FocA liegt in den hydrophoben Eigenschaften des Transmembrananteils und einem daraus resultierenden Mangel an positiv-ionisierbaren Aminosäuren für die massenspektrometrische Analyse sowie einer geringen Anzahl tryptischer Spaltstellen begründet. Dies beeinflusste nicht die Identifizierung von Interaktionspartnern. PflB wurde mit der zweithöchsten Anzahl erhaltener PSMs als Interaktionspartner von FocA identifiziert. Die massenspektrometrische Analyse führte mit sieben detektierten PSMs weiterhin zur Identifizierung des Pyruvat-Formiatlyase-aktivierenden Proteins (PflA) (Tabelle 13 und Abbildung A 17), das dafür bekannt ist, PflB post-translational zu modifizieren [2].

Abbildung 27 zeigt die Identifizierung eines Peptids der PflB durch Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS-Analyse. Das Signal des zweifach-geladenen Vorläuferions wurde bei m/z 976,495 detektiert.

Tabelle 13: Auswahl an identifizierten Bindungspartnern von FocA. Die Proteinidentifizierung erfolgte durch *In-Gel*-Proteinspaltung (Abbildung 26, Spur 4) und anschließender Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS.

UniProtKB- Eintrag	Protein	Sequenz- abdeckung [%]	Anzahl der identifizierten Peptide	Anzahl der <i>peptide</i> <i>spectral</i> <i>matches</i>
P0AC25	Formiatkanal A (FocA)	14,4	5	273
P09373	Pyruvat-Formiatlyase (PflB)	49	40	117
P0A9Q8	Alkoholdehydrogenase (AdhE)	38,2	28	56
P69911	Glutamatdecarboxylase (GadA)	18,9	10	32
P0A9N6	Pyruvat-Formiatlyase-aktivierendes Enzym (PflA)	18,2	4	7

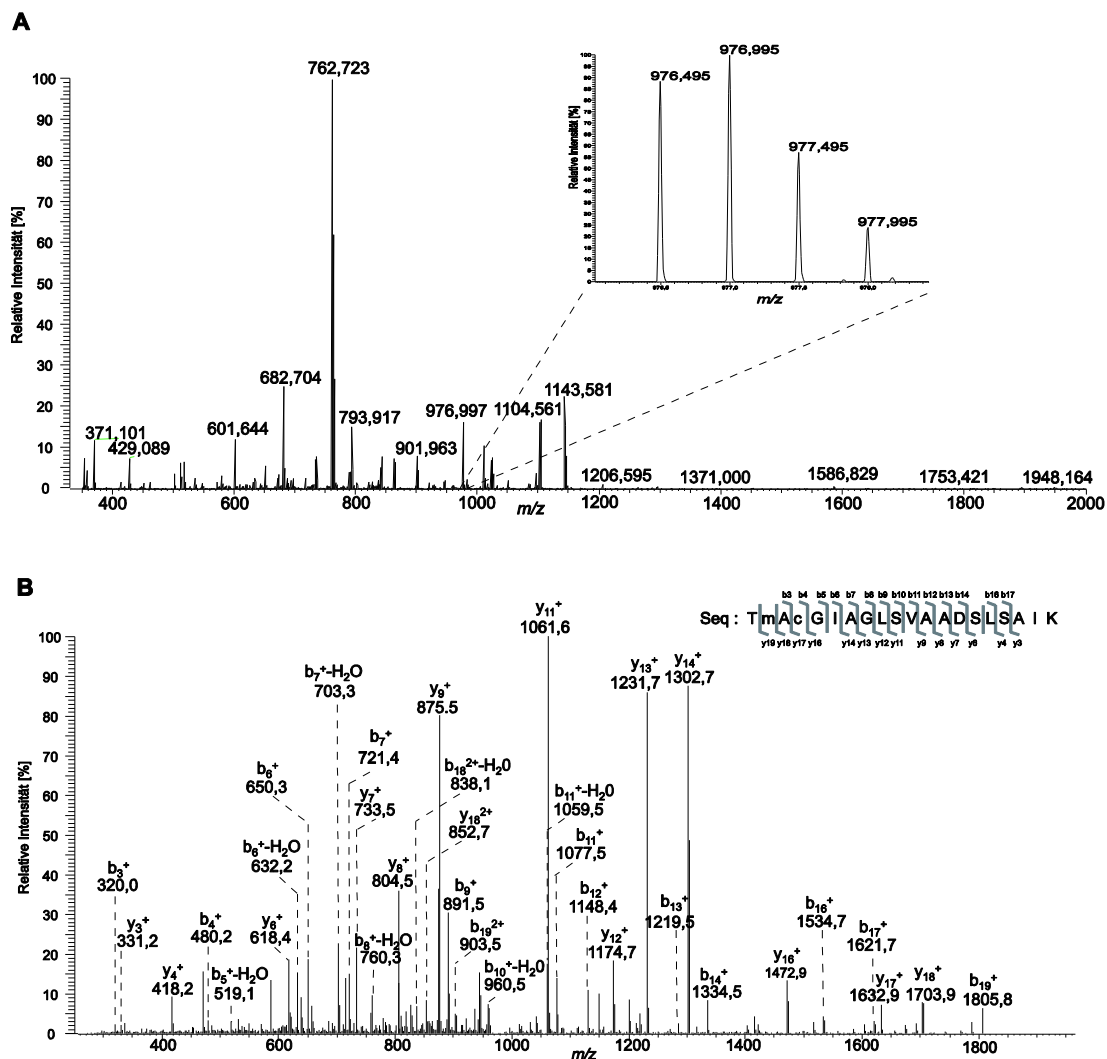


Abbildung 27: Massenspektrometrische Identifizierung eines Peptids der PfIB. (A) Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS-Analyse des Peptids TmAcGIAGLSVAADSLSAIK. Das Massenspektrum wurde bei einer Retentionszeit von 74,08 min aufgenommen. Das Isotopenmuster des zweifach geladenen Vorläuferions bei m/z 976,495 ist vergrößert dargestellt. Die Abweichung zwischen theoretischer und experimentell bestimmter Masse beträgt 0,36 ppm. **(B)** Zugehöriges Fragmentationenspektrum (CID-MS/MS). Das Vorläuferion wurde in der linearen Ionenfalle (LTQ) selektiert und fragmentiert. M2 ist oxidiert (m) und C4 carbamidomethyliert (c).

4.3.1.2 Ableitung von Distanzbeschränkungen im FocA/PfIB-Komplex mittels chemischem Cross-Linking

Nach der Bestätigung der Interaktion von FocA und PfIB erfolgte die strukturelle Charakterisierung des Proteinkomplexes anhand von Distanzbeschränkungen, die aus chemischen Cross-Linking-Experimenten erhalten wurden. Hierzu wurde der homobifunktionelle aminreaktive Cross-Linker BS²G genutzt. Die Verwendung von NHS-Estern, wie BS²G, hatte in früheren Studien bereits zur strukturellen Aufklärung von Proteinkomplexen geführt [194, 231, 233].

Nachdem das chemische Cross-Linking mit gereinigtem FocA und PflB durchgeführt worden war, erfolgte die Analyse des Cross-Linking-Ansatzes mittels SDS-PAGE. Diese zeigte Proteinbanden bei apparenten Molekularmassen größer als 200 kDa, die vermutlich FocA/PflB-Komplexen entsprechen (Abbildung 28). Im Vergleich zur Kontrolle (FocA und PflB ohne Cross-Linker) führte das Cross-Linking weiterhin zu einem Signal bei einer apparenten Molekularmasse von ca. 200 kDa. Die Proteinbanden bei ca. 28 und 85 kDa, die den Monomeren von FocA und PflB entsprechen, waren nicht detektierbar. Diese Beobachtung impliziert die Quervernetzung von FocA und PflB durch chemisches Cross-Linking. Die unterschiedlichen Quervernetzungsprodukte, welche durch das chemische Cross-Linking von pentameren FocA und dimeren PflB induziert werden, erschweren allerdings die visuelle Beurteilung der FocA/PflB-Komplexbildung mittels SDS-PAGE. Für die Identifizierung von intermolekularen Cross-Linking-Produkten zwischen FocA und PflB wurden die Gelbanden oberhalb der Molekularmasse von monomerem PflB (ca. 85 kDa) zunächst einer *In-Gel*-Proteinspaltung unterzogen und anschließend mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS analysiert. Die Auswertung führte zur Identifizierung von 14 intermolekularen Cross-Linking-Produkten zwischen FocA und PflB (Tabelle 14).

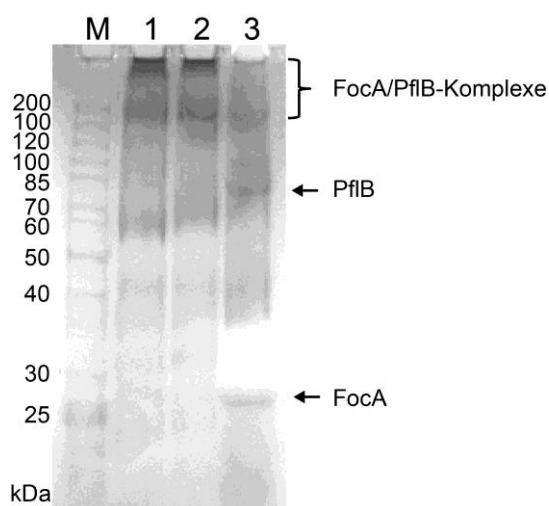


Abbildung 28: SDS-PAGE-Analyse von FocA/Protein-Komplexen. Die Auftrennung der Proteine wurde in einem 10%-igen Trenngel durchgeführt. Die Färbung der Proteine erfolgte mit kolloidaler Coomassie-Brillantblau-Lösung. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur 1 – 2: Ansatz nach chemischem Cross-Linking zwischen FocA und PflB; Spur 3: Kontrolle, FocA und PflB (ohne BS²G).

Diese wurden ausschließlich zwischen den cytoplasmatischen Regionen von FocA und PflB beobachtet. Dabei zeigten elf Cross-Linking-Produkte die Quervernetzung der N-terminalen Region von FocA mit PflB, wobei die Aminosäuren K17, Y25 oder K26, T28 und K29 aus FocA involviert waren.

Tabelle 14 zeigt die durch BS²G vernetzten Aminosäuren und ihre jeweiligen Positionen in den Aminosäuresequenzen von FocA und PflB. Neben der dominierenden Modifikation von Lysinen wurde für acht Positionen die Modifikation von Serinen, Threoninen und Tyrosinen in FocA und PflB detektiert. Für zwei weitere Positionen (Y25 von FocA und T298 von PflB) kann eine Modifikation von Tyrosin und Threonin nicht ausgeschlossen werden.

Die Fragmentationsspektren von zwei intermolekularen Cross-Linking-Produkten zwischen FocA und PflB sind in Abbildung 29 exemplarisch dargestellt.

Tabelle 14: Identifizierte intermolekulare Cross-Linking-Produkte des FocA/PflB-Komplexes. Die Identifizierung erfolgte durch *In-Gel*-Proteinspaltung und anschließender Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS.

<i>m/z</i> experimentell	<i>[M+H]⁺</i> theoretisch	<i>z</i> ^{*1}	Δm [ppm]	Modifizierte Aminosäure in FocA ^{*2}	Modifizierte Aminosäure in PflB ^{*2}
1318,342	3953,016	3	+0,9	K17	K312
1021,527	3062,576	3	-2,8	K26	K195, K197
883,116	2647,332	3	+0,1	K17	K32
838,437	2513,296	3	+0,6	K17	K441
830,753	2490,240	3	+1,5	K112	S463
795,067	2383,182	3	+1,5	K17	Y126
750,040	2248,101	3	+2,1	K191	K296, T298
744,708	2232,114	3	-1,4	K17	S338
712,704	2136,097	3	+0,5	K17	Y192
1036,523	2072,037	2	+0,7	K17	K643
675,336	2023,997	3	-0,9	S187	K135, K136
642,367	1925,086	3	+0,3	T28	K96
508,280	1522,826	3	-0,4	Y25, K26	Y726
621,298	1241,590	2	0	K29	T395

*¹ Ladungszustand des gemessenen Vorläuferions.

*² Vernetzungsstelle in der Aminosäuresequenz von FocA bzw. PflB.

Die MS/MS-Daten des Cross-Linking-Produktes (m/z 621,298) in Abbildung 29A umfassen die Aminosäuren A27 – K29 (FocA) und T395 – E401 (PflB), wobei die Quervernetzung zwischen K29 (FocA) und T395 (PflB) eindeutig identifiziert wurde. Die MS/MS-Daten des Cross-Linking-Produktes (m/z 830,753) in Abbildung 29B umfassen die Aminosäuren I105 – K112 (FocA) und S463 – E475 (PflB). Hier wurde die eindeutige Quervernetzung zwischen K112 (FocA) und S463 (PflB) eindeutig identifiziert. Eine vollständige Darstellung der MS/MS-Daten aller identifizierten intermolekularen Cross-Linking-Produkte des FocA/PflB-Komplexes ist der Abbildung A 18 zu entnehmen.

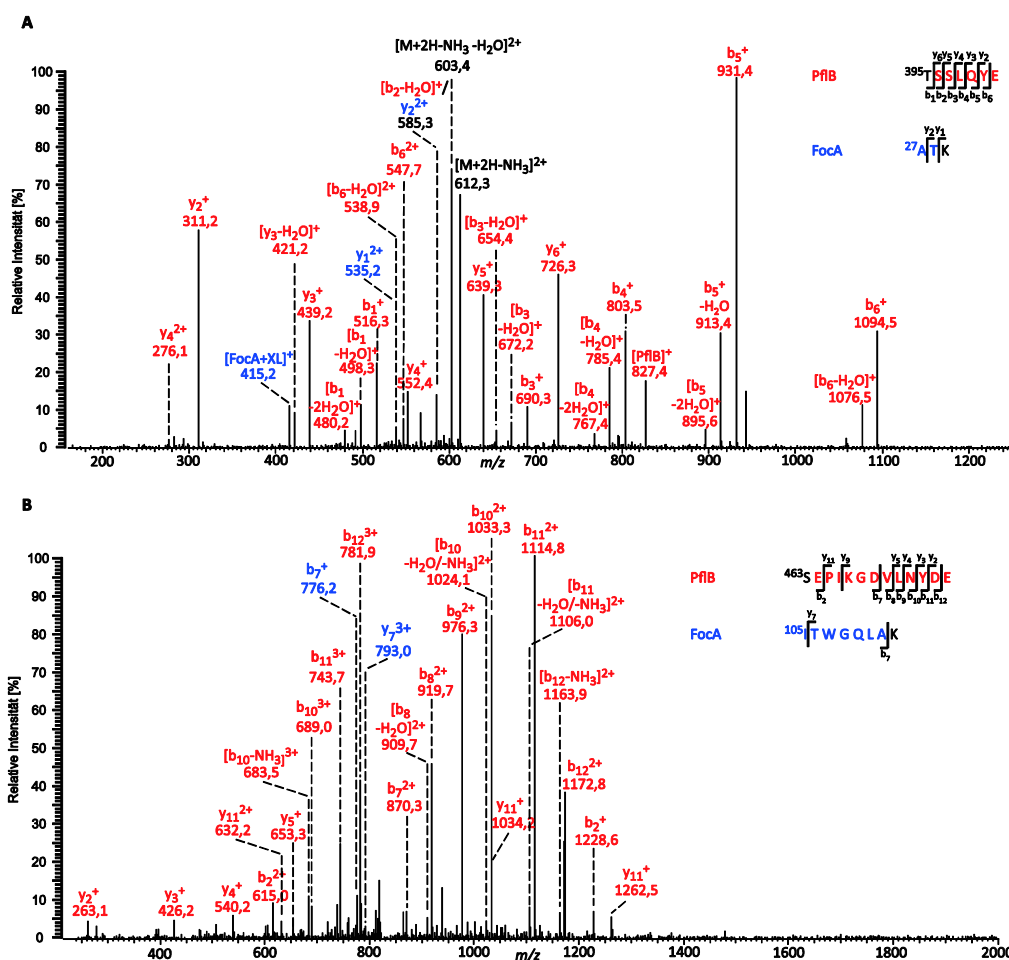


Abbildung 29: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FocA und PflB. Vorläuferionen wurden innerhalb der linearen Ionenfalle selektiert und mittels CID fragmentiert. Fragmentationen von FocA sind in blau, von PflB in rot dargestellt. **(A)** Fragmentationenspektrum eines zweifach geladenen Vorläuferions bei m/z 621,298. Das Cross-Linking-Produkt umfasst die Aminosäuren 27 – 29 von FocA und die Aminosäuren 395 – 401 von PflB, wobei K29 von FocA über BS²G mit T395 von PflB quervernetzt ist. **(B)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions bei m/z 830,753. Das Cross-Linking-Produkt umfasst die Aminosäuren 105 – 112 von FocA und die Aminosäuren 463 – 475 von PflB, wobei K112 von FocA über BS²G mit S463 von PflB quervernetzt ist.

4.3.2 Erstellung eines 3D-Strukturmodells des FocA/PfIB-Komplexes mittels Rosetta

Nach der Bestätigung einer Interaktion von PfIB mit der cytoplasmatischen Region von FocA erfolgte die Erstellung eines 3D-Strukturmodells des FocA/PfIB-Komplexes durch Rosetta [222] unter Zuhilfenahme der mittels chemischen Cross-Linkings erstellten Distanzbeschränkungen.

4.3.2.1 Charakterisierung der Konformation des *N*-Terminus von FocA mittels *de novo*-Faltung

Der *N*-Terminus von FocA aus *S. typhimurium* zeigt ein α -helikales Peptid, welches in verschiedenen Konformationen vorliegen kann (Abbildung 3) [27]. Demgegenüber fehlen Strukturinformationen für die *N*-terminale Region von FocA aus *E. coli* [25]. Aufgrund dieser unvollständigen Struktur sollte in Vorbereitung der Modellierung des FocA/PfIB-Komplexes von *E. coli* zunächst die Rekonstruktion einer vollständigen FocA-Struktur für die Ableitung der cytoplasmatischen Interaktionsfläche für PfIB erfolgen.

Hierfür erfolgte die Strukturvorhersage des *N*-Terminus mittels RosettaScripts [225]. Dazu wurde zunächst die Modellierung einer einzelnen *N*-terminalen Region (Aminosäuren 1 - 32, α -Untereinheit) in der Umgebung eines FocA-Pentamers durchgeführt. Mit Hilfe von RosettaScripts wurden 19 400 Modelle erstellt. Zur Beurteilung der erhaltenen Konformationen für die *N*-terminalen Aminosäuren von FocA wurde eine Clusteranalyse durchgeführt. Dafür wurden zunächst 5% (950) der Modelle mit niedrigsten Energiewerten für die *N*-terminale Region selektiert. Diese wurden anschließend mit Hilfe des Rosetta-Algorithmus einer Clusteranalyse unterzogen.

Die Mehrzahl der erstellten Modelle zeigte ein strukturell ungeordnetes Peptid im Bereich der Aminosäuren 1 - 9 (Abbildung 30). Dies erzeugte eine hohe Clusteranzahl und erschwerte die strukturbasierte Gruppierung der erstellten Modelle. Deshalb wurde die Clusteranalyse auf die Aminosäuren 10 - 285 der α -Untereinheit begrenzt. Hierbei wurde ein Clusterradius von 1,5 Å *root-mean-square deviation* (RMSD, C α -C α -Distanzen) verwendet. Die Analyse führte zur Erstellung von 77 Clustern. Die beiden umfangreichsten Cluster grenzten sich mit 267 (28,1%) und 160 (16,8%) Modellen deutlich von kleineren Gruppen ab.

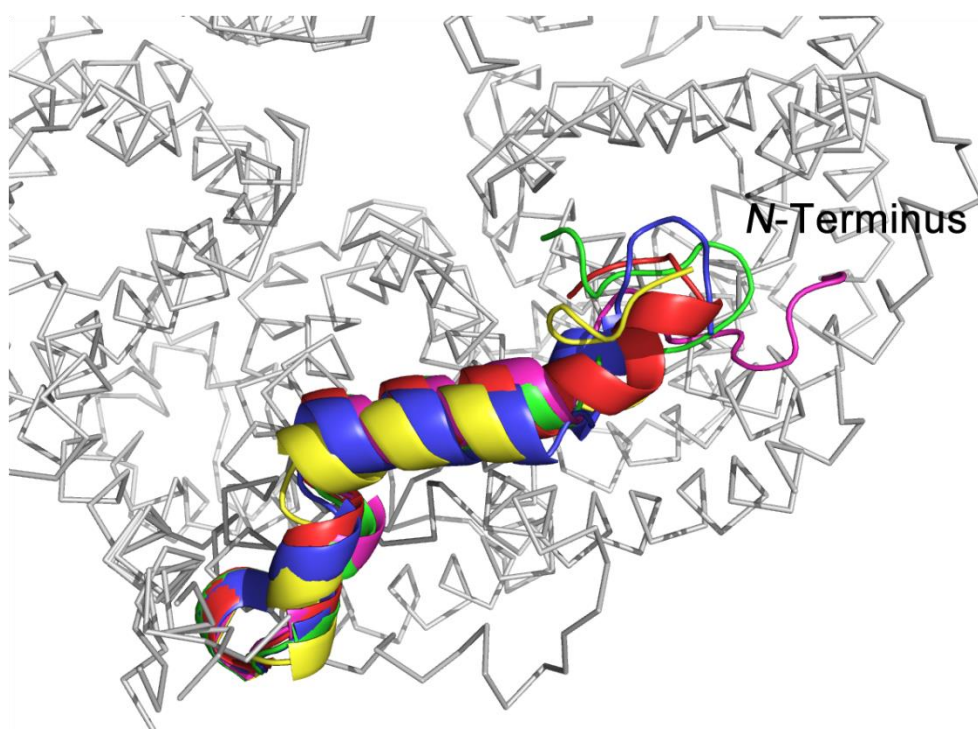


Abbildung 30: Strukturell ungeordneter Bereich des N-Terminus von FocA. Die Strukturvorhersage des *N*-Terminus (Aminosäuren 1 - 32) von *E. coli* erfolgte mittels RosettaScripts. Exemplarisch gezeigt sind die Konformationen von fünf übereinander gelegten Strukturen, die einen RMSD-Wert von 1,5 Å (Ca-Ca-Distanzen) aufweisen. Die Aminosäuren 1 – 9 sind unstrukturiert.

Die energetische Betrachtung der 18 umfangreichsten Cluster (Tabelle 15) zeigte die Korrelation des Clusterumfangs mit korrespondierenden Energiewerten. Für die beiden umfangreichsten Cluster wurden Modelle mit niedrigen Energiewerten für die *N*-terminalen 32 Aminosäuren (- 61,64 und - 62,40) und für das gesamte FocA-Pentamer (- 2054,23 und - 2040,20) beobachtet. Jedoch verfügten Modelle mit einer energetisch günstigen *N*-terminalen Region nicht zwingend über die niedrigsten Energiewerte des gesamten FocA-Pentamers. Der zweitgrößte Cluster enthält das Modell mit dem energetisch günstigsten *N*-terminalen Bereich (- 62,40). Die Gesamtenergie dieses Modells ist mit einem Energiewert von - 1417,95 jedoch sehr hoch. Diese Daten implizieren eine unzureichende energetische Minimierung der FocA-Pentamere durch RosettaScripts. Die teilweise hohen Energiewerte der FocA-Pentamere erschwerten die Beurteilung der *N*-terminalen Bereiche. Im Weiteren wurde deshalb für die Beurteilung der *N*-terminalen Konformationen die energetische Betrachtung ausschließlich auf die Aminosäuren 1 - 32 beschränkt.

Die Konformationen der *N*-terminalen Region der beiden umfangreichsten Cluster sind in Abbildung 31 dargestellt. Der modellierte *N*-Terminus des Clusters mit einem Umfang von 267 Modellen zeigte eine Konformation, welche die cytoplasmatische Seite des Formiatkanals nicht bedeckt. Im Gegensatz dazu repräsentierte das zweitumfangreichste Cluster mit 160 Modellen eine *N*-terminale Konformation, die den cytoplasmatischen Trichter des Formiatkanals verengt.

Tabelle 15: Clusteranalyse. Die Clusteranalyse wurde mit Rosetta durchgeführt. Hierfür wurden 5% (950) der Modelle verwendet, welche die niedrigsten Energiewerte des *N*-Terminus aufwiesen. Die Gruppierung wurde mit einem Clusterumfang von 1,5 Å (*root-mean-square deviation*, Ca-Ca-Distanzen) durchgeführt. Die niedrigsten Energiewerte der beiden Modelle (FocA-Pentamer und FocA-*N*-Terminus) des jeweiligen Clusters wurden gegenübergestellt.

Cluster-umfang	niedrigster Energiewert des FocA-Pentamers im Cluster	niedrigster Energiewert des <i>N</i> -Terminus im Cluster* (zugehöriger Energiewert des Pentamers)
267	- 2054,23	- 61,64 (- 2038,08)
160	- 2040,20	- 62,40 (- 1417,95)
28	- 2051,65	- 60,43 (- 2051,65)
28	- 2045,28	- 56,93 (- 2045,28)
32	- 2043,55	- 58,26 (- 2000,04)
42	- 2041,95	- 57,46 (- 2022,81)
43	- 2041,25	- 57,52 (- 2004,88)
27	- 2036,68	- 60,78 (- 2035,27)
15	- 2029,59	- 55,08 (- 2023,30)
11	- 2029,86	- 56,41 (- 1970,56)
10	- 2028,34	- 56,40 (- 2024,10)
19	- 2026,39	- 56,80 (- 2022,73)
17	- 2026,06	- 58,12 (- 2002,13)
19	- 2024,99	- 59,48 (- 2014,98)
11	- 2025,75	- 55,20 (- 1981,19)
20	- 2019,44	- 58,27 (- 2016,74)
11	- 2019,54	- 57,58 (- 1987,45)
10	- 2009,48	- 55,51 (- 2009,48)

* Für die Evaluierung der Energie des *N*-Terminus wurden nur die Aminosäuren 1 - 32 der α -Untereinheit in Betracht gezogen.

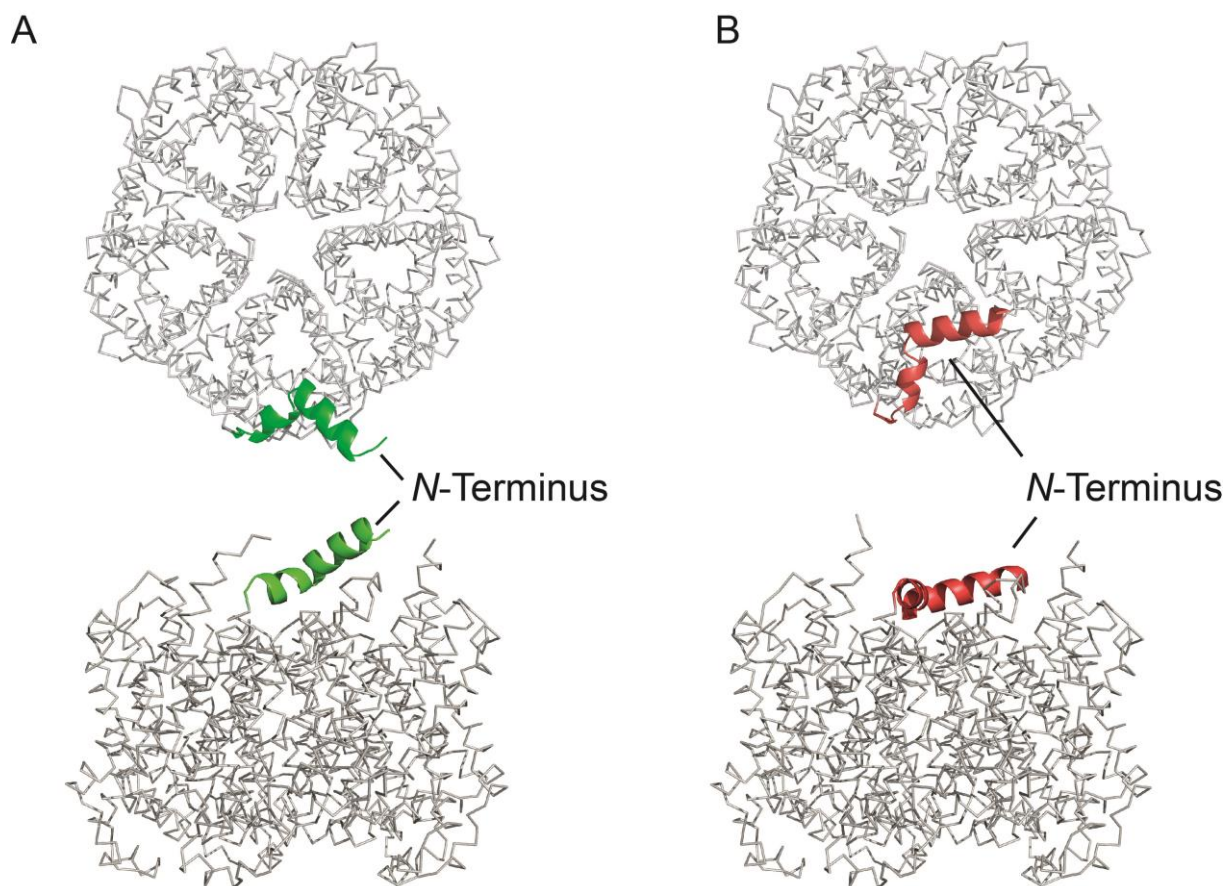


Abbildung 31: Strukturvorhersage des N-Terminus (Aminosäuren 1 – 32) von FocA. Die Strukturvorhersage erfolgte mittels RosettaScripts. Hierfür wurde der N-Terminus der α -Untereinheit innerhalb des FocA-Pentamers *de novo* gefaltet. Anschließend erfolgte eine Clusteranalyse mit 5% der energetisch günstigsten Modelle. Gezeigt sind die energetisch günstigsten N-terminalen Regionen der beiden umfangreichsten Cluster. Die Darstellungen zeigen die Strukturen vom Cytoplasma in Richtung Periplasma (oben) und parallel zur Membran (unten). (A) N-terminale Konformation des umfangreichsten Cluster mit einer Anzahl von 267 Modellen. (B) N-terminale Konformation des zweitumfangreichsten Cluster mit einer Anzahl von 160 Modellen.

Die hohen Modellumfänge dieser beiden Cluster und ihre niedrigen Energiewerte deuten darauf hin, dass es sich bei den dargestellten Strukturvorhersagen um Konformationen handelt, die höchstwahrscheinlich die nativen Faltungen des N-Terminus von FocA repräsentieren.

4.3.2.2 Rekonstruktion der cytoplasmatischen Interaktionsfläche zwischen FocA und PfIB

Nachdem die Struktur für den N-Terminus von FocA modelliert worden war, sollte die vollständige cytoplasmatische Interaktionsfläche zwischen FocA und PfIB erstellt

werden. Die strukturellen Untersuchungen der FocA/PfIB-Interaktion mittels chemischem Cross-Linking und durch die *de novo*-Faltung des *N*-Terminus deuteten auf eine hohe Flexibilität des *N*-Terminus von FocA hin und unterstützen die Hypothese der in der Kristallstruktur von *S. typhimurium* beobachteten multiplen Konformationen [27]. Desweiteren gibt der Grad der Oligomerisierung von pentameren FocA und dimeren PfIB keinen Grund zur Annahme einer absolut symmetrischen Interaktionsfläche beider Proteine. Deshalb wurden für die Erstellung einer vollständigen cytoplasmatischen Interaktionsfläche von FocA mehrere Konformationen in Betracht gezogen. Hierfür wurden die vervollständigten FocA-Monomer-Modelle der 18 umfangreichsten Cluster ausgewählt und durch zufällige Rekombination zu FocA-Pentameren assembliert. Dabei richtete sich die Anwesenheit einer bestimmten Konformation der *N*-terminalen Region innerhalb der rekonstruierten FocA-Pentamere nach dem prozentualen Anteil des jeweiligen Clusterumfangs von der Gesamtmodellanzahl aller 18 Cluster. FocA-Monomer-Modelle aus umfangreicheren Clustern wurden häufiger für die Rekonstruktion von vollständigen FocA-Pentameren verwendet als Modelle aus weniger umfangreichen Clustern. Aufgrund der ungeordneten Struktur im Bereich der Aminosäuren 1 – 9 erfolgte für die Rekonstruktion der FocA-Pentamere ausschließlich die Verwendung der Strukturinformationen der Aminosäuren 10 – 285 der α -Untereinheit.

In Vorbereitung des FocA/PfIB-*Dockings* und für die nachfolgende energetische Beurteilung der erhaltenen Modelle wurde eine Relaxierung [227] der rekonstruierten FocA-Pentamere durchgeführt. Diese ermöglichte die energetische Minimierung von 14922 Modellen. Ein rekonstruiertes FocA-Pentamer zeigt vor der Relaxierung zahlreiche Überlagerungen von Aminosäureseitenketten und Proteinrückgratstrukturen (Abbildung 32A, links). Diese wurden durch die zufällige Rekombination einzelner Untereinheiten induziert und traten bevorzugt an den Grenzflächen der einzelnen Untereinheiten auf. Durch die Relaxierung konnten diese Überlagerungen nahezu vollständig aufgelöst werden (Abbildung 32A, rechts), wobei sich minimale Veränderungen der globalen Konformation der rekonstruierten FocA-Pentamere ergaben (Abbildung 32B, links). Die lokale Auswirkung der Relaxierung auf die Konformationsänderung einzelner Rotamere ist am Beispiel des Phenylalanins an Position 193 exemplarisch dargestellt (Abbildung 32B, rechts), das sich im Zentrum des FocA-Pentamers befindet. Die *in silico*-Assemblierung der Untereinheiten induzierte eine Überlagerung von vier F193-Seitenketten (blau). Die Relaxierung

fürte zur Neuorientierung und optimalen räumlichen Verteilung der einzelnen Rotamere (grün).

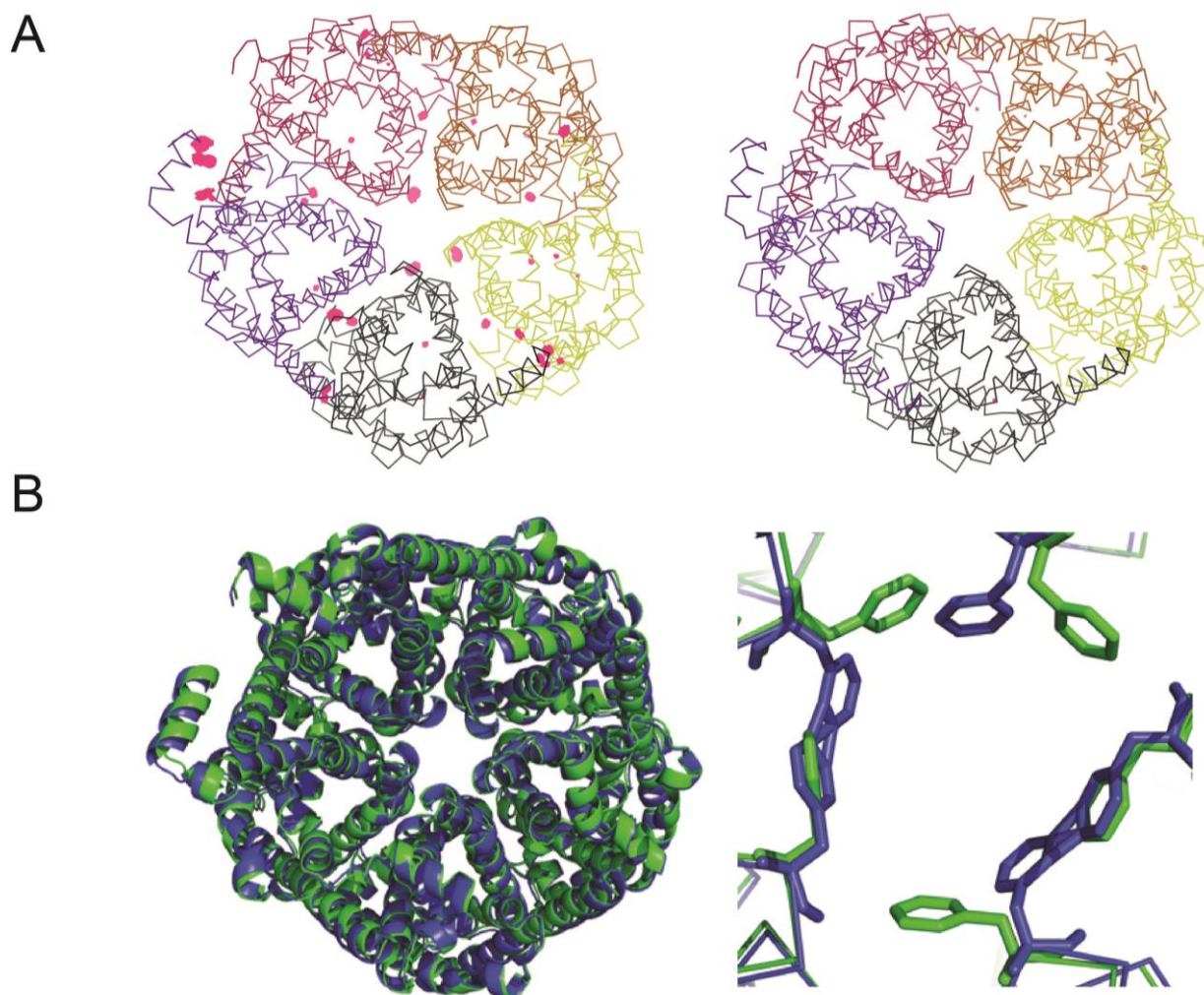


Abbildung 32: Energetische Minimierung rekonstruierter FocA-Pentamere. (A) Rekonstruiertes FocA-Pentamer vor (links) und nach (rechts) der energetischen Minimierung mittels Rosetta. Bereiche, in denen sich Aminosäureseitenketten und Proteinrückgratstrukturen überlagern, sind als rote Punkte hervorgehoben. Die Analyse und Darstellung der Interaktionen erfolgte mit Molprobit [213] und King [212]. (B) Rekonstruiertes FocA-Pentamer vor (blau) und nach (grün) der energetischen Minimierung mittels Rosetta. Die Bänderstruktur (links) zeigt die globale Veränderung der Proteinrückgratstruktur durch die energetische Minimierung. Die Überlagerung von Aminosäureseitenketten ist am Beispiel von F193 gezeigt (rechts).

4.3.2.3 Modellierung des FocA/PfIB-Komplexes

Nach der Rekonstruktion vollständiger FocA-Pentamere erfolgte die Modellierung des FocA/PfIB-Komplexes. Hierfür wurde zunächst die Relaxierung der Kristallstruktur von PfIB mittels Rosetta durchgeführt. Anschließend erfolgte das *Docking* von FocA und PfIB, was zur Erstellung von 53340 Modellen führte. Für die Darstellung des FocA/PfIB-Komplexes, der die native Konformation repräsentiert, erfolgte die strukturelle Gruppierung mittels *BCL::Cluster* [229]. Dabei wurden 10% der Modelle mit den höchsten Energiebeträgen nicht in Betracht gezogen. Die ungefähr 48000 verbleibenden Modelle wurden entsprechend ihrer FocA/PfIB-Interaktionsflächenenergiewerte sortiert, und die 500 energetisch günstigsten Modelle für das *BCL::Cluster* ausgewählt. Für die Clusteranalyse wurde ein Algorithmus gewählt, welcher die durchschnittliche Distanz des Proteinrückgrats (C α -C α -Distanzen) aller Modelle von benachbarten Clustern berechnet (*average pair-wise distance*). Die Anzahl der Modelle und Energiewerte der Interaktionsflächen von Gruppen mit einem durchschnittlich paarweisen RMSD-Wert von 11 Å wurden weiter untersucht (Tabelle 16).

Tabelle 16: BCL::Clusteranalyse. Die Gruppierung erfolgte mit 500 Modellen, welche die niedrigsten Werte der FocA/PfIB-Interaktionsflächenenergie aufwiesen. Die Clusteranalyse basierte auf dem strukturellen Vergleich der durchschnittlich paarweisen RMSD-Werte von FocA/PfIB-Modellen. Für die Berechnung der Durchschnitte wurden jeweils zehn Modelle eines Clusters mit den niedrigsten Interaktionsflächenenergiewerten in Betracht gezogen.

Gruppen- umfang	Niedrigster Interaktions- flächen- energiewert	Durchschnittlicher Interaktionsflächen- energiewert	Standardabweichung des durchschnittlichen Interaktionsflächen- energiewertes	Durchschnittlicher RMSD [Å]
71	-16,81	-16,44	0,27	7,96
54	-18,83	-15,90	1,44	12,30
53	-16,28	-15,53	0,51	11,86
45	-16,70	-14,87	0,89	8,83
43	-17,55	-14,60	1,18	6,88
43	-18,12	-15,42	1,20	12,46
35	-15,92	-14,74	0,62	9,45
31	-19,32	-15,86	1,55	8,62
31	-18,24	-15,69	1,19	11,48
30	-15,02	-13,87	0,60	6,68
26	-18,53	-15,39	1,58	8,52
14	-19,42	-14,80	1,70	7,39
13	-16,60	-13,90	1,51	7,04

Dabei wurden 13 Cluster mit Umfängen von 13 bis 71 Mitgliedern identifiziert. Der umfangreichste Cluster beinhaltete 71 FocA/PflB-Modelle und hob sich mit einer relativ geringen Differenz von nachfolgenden Clustern ab. Mit einem durchschnittlichen Energiewert von -16,44 weisen die Modelle des umfangreichsten Clusters die energetisch günstigsten Interaktionsflächen auf und verfügen mit 0,27 über die geringste Standardabweichung. Zusätzlich zeigen diese Modelle einen niedrigen durchschnittlichen RMSD-Wert von 7,96 Å, was die hohe strukturelle Ähnlichkeit dieser modellierten FocA/PflB-Komplexe widerspiegelt. Auffällig sind die niedrigeren Energiewerte der Interaktionsflächen von einzelnen Modellen umfangärmerer Cluster. Das Cluster mit einem Umfang von 14 Modellen beinhaltet z. B. das FocA/PflB-Modell mit dem niedrigsten Interaktionsflächenenergiewert von -9,42. Die Cluster mit Umfängen von 54, 43, 31, 26 verfügen vereinzelt ebenfalls über Modelle mit geringeren Energiewerten für die Interaktionsflächen. Abbildung 33 zeigt die Konformationen der energetisch (FocA/PflB-Interaktionsflächenenergie) repräsentativsten Modelle der 13 Cluster (Tabelle 16).

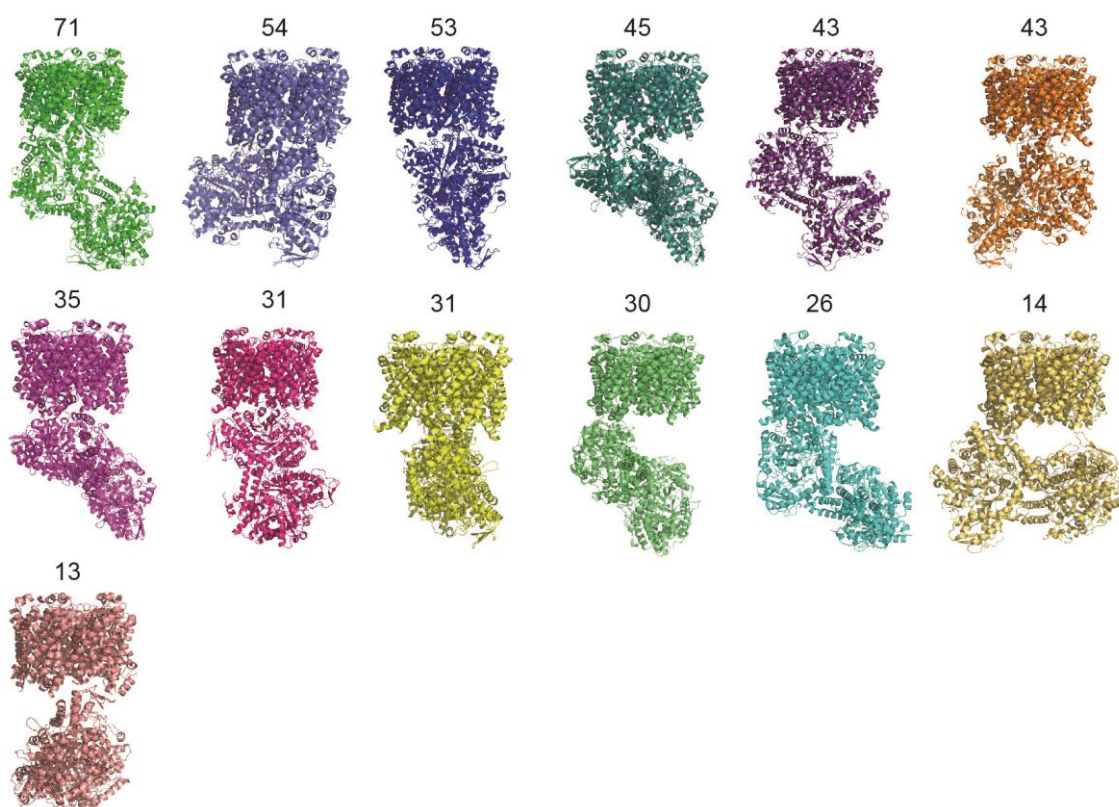


Abbildung 33: BCL::Clusteranalyse. Gezeigt sind repräsentative FocA/PflB-Modelle der 13 Cluster, welche paarweise *RMSD*-Werte ($C\alpha$ - $C\alpha$ -Distanzen) von 11 Å verfügten. Die Clusterumfänge der jeweiligen FocA/PflB-Modelle sind angegeben.

Die dargestellten Modelle zeigen unterschiedlichen Bindungsorientierungen von PfIB zu FocA, wobei verschiedene Regionen von PfIB in die Interaktion involviert sind. Die Interaktion basiert vorwiegend auf der Bindung von einer PfIB-Untereinheit an FocA, wobei hauptsächlich der *N*-Terminus von FocA das Bindemotiv darstellt. Die FocA/PfIB-Komplexe von Clustern mit Umfängen von 54, 53, 31 und 14 Modellen zeigen die Interaktion beider PfIB-Untereinheiten mit FocA. Eine vollständige Abdeckung aller fünf formiatleitenden Kanäle von FocA durch die Interaktion mit PfIB wurde in keinem der Modelle beobachtet.

Für die endgültige Erstellung eines Modells des FocA/PfIB-Komplexes wurde das beschriebene Protokoll von Dr. David Nannemann in der Arbeitsgruppe von Prof. Dr. Jens Meiler (Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA) überarbeitet. Die Modellierung des Komplexes erfolgte unter Zuhilfenahme der mittels chemischem Cross-Linking erstellten Distanzbeschränkungen. Die zwei möglichen Konformationen des FocA/PfIB-Komplexes sind in Abbildung 34 zu sehen.

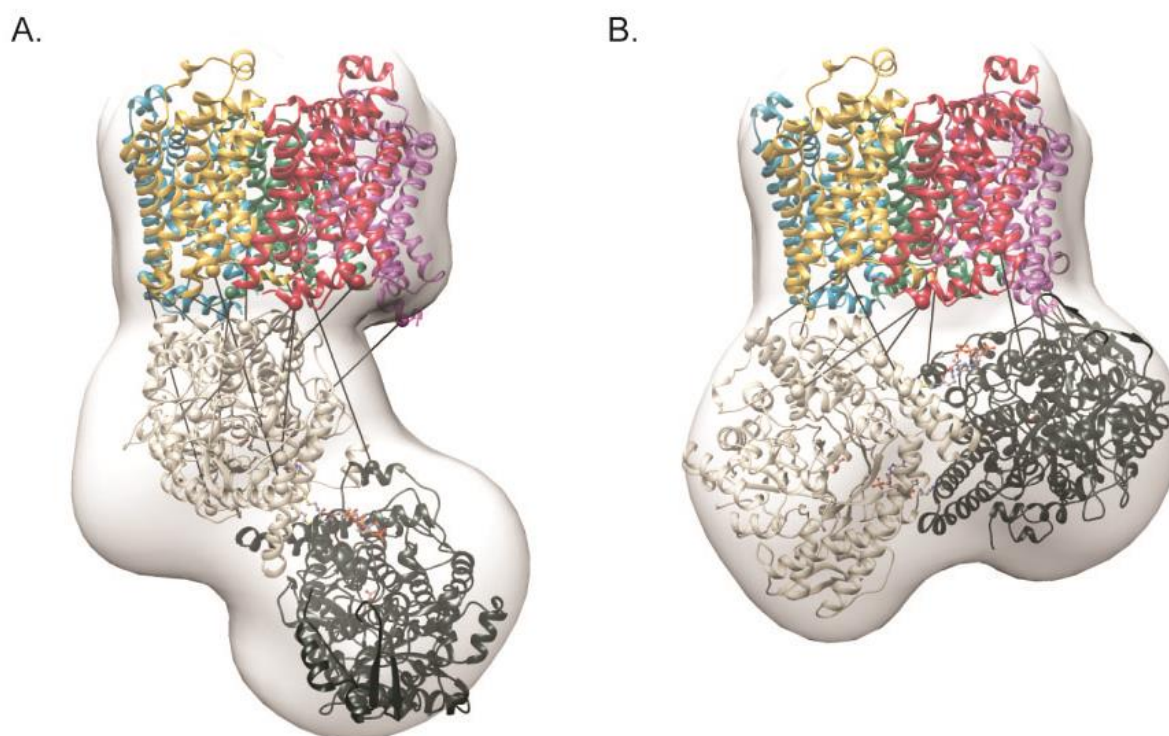


Abbildung 34: Modelle des FocA/PfIB-Komplexes. Dargestellt sind die Konformationen A und B, welche mit Hilfe der Cross-Linking-Distanzbeschränkungen ausgewählt wurden. Die FocA- und PfIB-Untereinheiten sind farblich hervorgehoben. Die intermolekularen Cross-Links sind durch Linien dargestellt. Der konformative Raum, welcher durch die Modelle eingenommen wird, die die mittels Cross-Linking erstellten Distanzbeschränkungen erfüllen, ist als graue Hülle hervorgehoben.

5 Diskussion und Ausblick

5.1 Selektiver unspezifischer Einbau von Selenocystein in Proteine der formiatabhängigen Nitratatmung von *E. coli*

Selen ist für zahlreiche Organismen ein essentielles Spurenelement [64, 68]. Der spezifische Einbau von physiologisch aktivem Selen erfolgt co-translational in Form der 21. Aminosäure Selenocystein [61] und wird auf genetischer, biochemischer und physiologischer Ebene streng reguliert [64].

Eine Ergänzung des Kulturmediums mit Selenitkonzentrationen über 1 μM führt trotz zahlreicher Kontrollmechanismen [64] zu einem unspezifischen Einbau von Selen in Proteine von *E. coli* [74]. Während der unspezifische Einbau von Selen anfänglich als eine Anreicherung von Selenomethionin im Proteom postuliert wurde [69], deuten neuere Erkenntnisse auf den unspezifischen Einbau von Selenocystein [70, 71].

Die in dieser Arbeit durchgeführten Experimente zur Untersuchung des unspezifischen Einbaus von Selen in *E. coli* zeigen, dass dieser zu einem Austausch von Cystein durch Selenocystein an bestimmten Positionen in Proteinen führt. Zur Analyse erfolgte zunächst die Proteinreinigung des Selenoenzyms Fdh-N aus dem *E. coli*-Wildtypstamm MC4100, welcher bei Selenitkonzentrationen von 2 μM und 100 μM kultiviert wurde. Fdh-N enthält ein spezifisches Selenocystein an Position 196 (U196) der katalytischen Untereinheit [37] und stellt dadurch ein geeignetes Beispiel zur Charakterisierung von Selenoproteinen mittels Massenspektrometrie dar. Die Proteinreinigung der nativen Fdh-N aus der Membranfraktion von *E. coli* führte zur Koisolierung des homologen Isoenzym Fdh-O sowie Nar. Zusammen stellen diese Enzyme die Schlüsselkomponenten der Formiat-Nitrat-Atmungskette in *E. coli* dar [36]. Vorausgegangene Studien hatten eine deutlich niedrigere Proteinbiosynthese und Enzymaktivität von Fdh-O gegenüber Fdh-N [39, 40] gezeigt, welche während der Nitratatmung stark induziert wird [38]. Die katalytischen und Elektronen-übertragenden Untereinheiten zeigen eine hohe Aminosäuresequenzidentität (FdoG/FdnG: 76%; FdoH/FdnH: 75%) [45]. Die relative Quantifizierung der beiden gemeinsam gereinigten Fdh-O und Fdh-N, mittels *spectral counting* zeigte eine bis zu ca. dreifach höhere Menge an Fdh-O gegenüber Fdh-N und impliziert im Kontext mit den Erkenntnissen einer kürzlich veröffentlichten Studie [42] eine

bedeutendere Rolle der wenig beschriebenen Fdh-O während der formiatabhängigen Nitratatmung in *E. coli*.

Die massenspektrometrische Analyse der Fdh-O und Fdh-N führte zur Identifizierung des spezifischen UGA-kodierten Selenocysteins an Position 196 sowie fünf Substitutionen von Cystein gegen Selenocystein an identischen Positionen (C165U, C380U, C391U, C554U und C618U) in den großen katalytischen Untereinheiten (FdoG und FdnG) (Abbildung 17). Dabei wurde auch bei der Ergänzung des Kulturmediums mit einer hohen Selenitkonzentration von 100 μM kein Austausch der [4Fe-4S]-Zentren koordinierenden Cysteine beobachtet. Interessanterweise konnte für die Positionen 165, 380, 391, 554 und 618 eine Variabilität im Einbau von Cystein und Selenocystein beobachtet werden (Tabelle 8), während an Position 196 ausschließlich Selenocystein detektiert wurde. Die massenspektrometrische Analyse der Elektronen-übertragenden Untereinheiten der Fdhs (FdoH und FdnH) sowie Nar zeigte ebenfalls die Substitution von Cystein durch Selenocystein an definierten Positionen.

Während für FdnG und FdoG ausschließlich Substitutionen von Cysteinen auftraten, welche nicht an der Koordinierung des [4Fe-4S]-Zentrums beteiligt sind, konnte für FdnH und FdoH ein Austausch von Cysteinen, welche die beiden proximal zur Membran gelegenen [4Fe-4S]-Zentren koordinieren, verzeichnet werden.

Das Ergebnis eines zusätzlichen Einbaus von Selenocystein an den für Cystein kodierenden Kodonen UGU und UGC bekräftigt die Erkenntnis einer früheren Studie, dass der unspezifische Einbau von Selenocystein auf die Schritte der Cysteinbiosynthese angewiesen ist [71]. Desweiteren belegt die Beobachtung des Austausches bestimmter Cysteine durch Selenocysteine in NarG aus der *E. coli*-*selC*-Deletionsmutante FM460, welche identisch mit den identifizierten Austauschpositionen in NarG des *E. coli*-Wildtypstamms MC4100 sind, dass sich der unspezifische Einbau von Selenocystein unabhängig von SelC (tRNA^{Sec}) vollzieht. Diese Beobachtung wird dadurch gestützt, dass eine effiziente Acetylierung der tRNA^{Cys} auch mit dem isosteren Cysteinanalogon Selenocystein erfolgen kann [70, 73], wodurch ein Einbau von Selenocystein an den Kodonen UGU und UGC möglich ist. Der unspezifische Einbau von Selen in die Biosynthese von Cystein erfolgt vermutlich durch die O-Acetylserinsulfhydrylase, welche neben Sulfid auch Selenid als Substrat akzeptiert [72]. Die dadurch resultierende Synthese von Selenocystein und dessen unspezifischer Einbau kann durch eine Erhöhung der Cystein-

konzentration während des Zellwachstum und einer dadurch hervorgerufenen Inhibierung der O-Acetylserinsulfhydrylase unterbunden werden [71, 234], was die Unabhängigkeit des unspezifischen Einbaus des Selenocysteins vom spezifischen Mechanismus des Selenocysteineinbaus bekräftigt.

Der Austausch von identischen Cysteinen durch Selenocystein in FdnG und FdoG sowie der Austausch an definierten Positionen in FdnH, FdoH und NarG deutet auf einen unspezifischen Einbau von Selenocystein hin, welcher dennoch selektiv erfolgt. Dies führt zu der Frage nach dem zugrunde liegenden Mechanismus. Eine Untersuchung, ob ein selektiver unspezifischer Selenocysteineinbau bevorzugt am Kodon UGU oder UGC erfolgt bzw. ob die jeweilige Positionen innerhalb der mRNA einen Einfluss auf die Rate des Selenocysteineinbaus hat, lieferte keine schlüssigen Ergebnisse. Die Abhängigkeit eines selektiven Selenocysteineinbaus von der Struktur der mRNA wurde durch die Veränderung der Wachstumstemperatur von 37°C auf 22°C überprüft. Dabei konnten jedoch keine Veränderungen in der Rate des Cysteinaustausches detektiert werden.

Eine alternative Erklärung für die Translation der Kodone UGU und UGC mit Cystein oder Selenocystein bietet die Variation der Translationsgeschwindigkeit unterschiedlicher Bereiche der mRNA, welche durch den unterschiedlichen Gebrauch der Kodone (*codon usage*) in *E. coli* [235] hervorgerufen werden könnte. So könnte der vermehrte Fehleinbau von Selenocystein in gewisse Bereiche des Proteins mit einer höheren Translationsgeschwindigkeit des jeweiligen Kodonbereichs assoziiert sein.

Der Einbau von Selenocystein an UGU und UGC-kodierten Positionen könnte der Zelle die Möglichkeit geben, einen Überschuss an Selen an bestimmte Positionen in Proteinen aufzunehmen. Ein derartiger Mechanismus könnte erlauben, dass bestimmte, Fe-S-Zentren koordinierende, Thiolgruppen oder katalytisch wichtige Cysteine von einer Substitution unberührt bleiben, um die Inaktivierung dieser Enzyme zu vermeiden.

Der Austausch von Cystein durch das reaktivere Selenocystein ($pK_{s_{\text{Thiol}}}$: 8,3; $pK_{s_{\text{Selenol}}}$: 5,2) [68] könnte zu instabilen Varianten der Fdhs führen, welche *in vivo* vermutlich abgebaut würden und folglich mit der in dieser Arbeit durchgeführten Untersuchung der vollständig prozessierten und gereinigten Enzyme nicht detektiert werden könnten. Diese Möglichkeit darf nicht außer Acht gelassen werden und sollte in weiterführenden Untersuchungen überprüft werden.

5.2 Chemisches Cross-Linking zur Identifizierung und strukturellen Charakterisierung von Protein-Protein-Interaktionen

In dieser Arbeit wurde chemisches Cross-Linking zur Identifizierung von Interaktionspartnern der Fdhs sowie zur Erstellung eines Modells des FocA/PflB-Komplexes genutzt.

Trotz seines großen Potentials stellt chemisches Cross-Linking in Kombination mit MS eine bislang wenig genutzte Methode zur Identifizierung von Protein-Protein-Interaktionen dar [236]. Neben der Identifizierung von Interaktionspartnern ist die Ableitung nieder aufgelöster 3D-Strukturinformationen der Proteinkomplexe möglich [237, 238]. Die Methode liefert Distanzinformationen durch die Identifizierung kovalent quervernetzter Aminosäuren und wurde bereits erfolgreich an gereinigten homogenen Proteinkomplexen bis zur Größe des mitochondrialen Ribosoms angewandt [193, 239-242].

Für die Vermeidung falsch-positiver Identifizierungen von Quervernetzungsprodukten müssen hohe Anforderungen an die Qualität der MS-Daten gestellt werden. So muss die komplexe Datenauswertung von Quervernetzungsprodukten photoaktivierbarer Cross-Linker aufgrund einer erhöhten Anzahl von möglichen Reaktionsstellen durch hoch-aufgelöste und hochgenaue Fragmentationsspektren gestützt werden [187]. Desweiteren sollte die Qualität der Daten durch die Bestimmung der *false discovery rate* ermittelt werden.

Die Vorteile des Cross-Linkings liegen in der Verwendung geringer Probenmengen. Die kovalente Fixierung der Interaktionspartner eines Proteins erlaubt intensivere Waschbedingungen während affinitätschromatographischer Schritte, wodurch eine Depletion hochabundanter Kontaminanten erzielt wird. Im Gegensatz zu den hoch-auflösenden Techniken der Proteinstrukturanalyse, wie z. B. Röntgenkristallographie und NMR-Spektroskopie, bietet chemisches Cross-Linking in Kombination mit bioinformatischen Methoden zusätzlich die Möglichkeit der Strukturaufklärung von Proteinkomplexen heterogener Proteinproben [231]. Dabei ist die Größe der zu analysierenden Komplexe nicht limitiert [192] und die Charakterisierung flexibler Proteinbereiche sowie die Aufzeichnung dynamischer Konformationsänderungen ist möglich [191].

5.3 Zellwandständige Interaktionspartner der Fdhs

Die komplexen Molybdoenzyme Fdh-N und Fdh-O benötigen die Unterstützung akzessorischer Reifungsproteine, welche den korrekten Einbau prosthetischer Gruppen sowie eine optimale Proteinfaltung im Cytoplasma koordinieren [210]. Derzeit existieren nur wenige Informationen über zellwandständige Interaktionspartner beider Enzyme, die eine Stabilisierung der Proteinfaltung und Membraninsertion auf periplasmatischer Seite zur vollständigen Reifung der Atmungskettenproteine vermitteln.

Das chemische Quervernetzungsexperiment der gemeinsam gereinigten Isoenzyme Fdh-O und Fdh-N mit einem Zellwand-Subproteom und die nachfolgende massenspektrometrische Analyse der Quervernetzungsprodukte führte zur Identifizierung der vier potentiellen zellwandständigen Bindungspartner Skp, OppA, ErfK und TolB. Die detaillierte Betrachtung der Quervernetzungsstellen lieferte strukturelle Informationen über die möglichen Grenzflächen zwischen den beiden Fdhs und ihren Interaktionspartnern.

5.3.1 FdnH/Skp-Komplex

Das Protein Skp hat eine Molekularmasse von 17 kDa und ist ein ATP-unabhängiges periplasmatisches Chaperon [243, 244]. Die Röntgenkristallstruktur von Skp zeigt ein „Quallen“-artiges Homotrimer, an dessen Grundkörper sechs α -Helices eine flexible Substratbindestelle bilden (Abbildung 24) [245, 246]. Skp bindet Substrate, wie z. B. äußere Membranproteine, mit Dissoziationskonstanten im nanomolaren Bereich [247]. Die Substratbindung erfolgt domänenselektiv, wobei im Fall von Substraten mit zwei Domänen die Transmembrandomäne in die Substratbindestasche aufgenommen wird und die lösliche nativ-gefaltete Domäne außerhalb der Bindestasche verbleibt [248]. Die Skp-Substrat-Wechselwirkung wurde u. a. am Beispiel des Skp-OMP-Komplexes beschrieben [244]. Die Funktion von Skp liegt in der Aufnahme von Proteinen an der inneren Membran und der Faltungsstabilisierung während der Überführung durch das periplasmatische Kompartiment zur anschließenden Insertion in die äußere Membran in gram-negativen Bakterien [243, 249, 250]. *In vitro*-Experimente zeigen, dass Skp auch die Faltung löslicher Proteine durch eine Inhibierung der Aggregation unterstützt [251].

Die Auswertung des Quervernetzungsexperimentes führte zur Identifizierung von vier Quervernetzungsprodukten zwischen Fdh-N und Skp (Tabelle 12). Dabei wurden die Reaktionsstellen des Quervernetzungsreagenz in Fdh-N ausschließlich am zentralen Bereich der Elektronen-übertragenden Untereinheit (FdnH) identifiziert. Die detektierten Reaktionsstellen der photoreaktiven Gruppe des Quervernetzungsreagenz im Skp-Trimer konnten aufgrund der drei identischen Aminosäuresequenzen nicht einer bestimmten Skp-Untereinheiten zugeordnet werden, wodurch eine Verteilung der erhaltenen Cross-Links über alle drei Untereinheiten in Betracht gezogen werden muss. Die Cross-Linker-Reaktionsstellen in Skp konnten, ähnlich wie in einer vorausgegangenen Studie [248], ausschließlich an den flexiblen α -Helices verzeichnet werden, welche die Substratbindestelle bilden. Dadurch kann eine spezifische Interaktion zwischen Fdh-N und Skp angenommen werden.

Im Gegensatz zu äußeren Membranproteinen ist wenig über einen Mechanismus bekannt, über den Skp eine Faltung von löslichen oder inneren Membranproteinen, wie z. B. Tat-Substraten, vermittelt und stabilisiert. Aufgrund des breiten Substratspektrums von Skp [252] und dessen Funktion, bestehende Proteinfaltungen durch Inhibierung der Aggregation zu stabilisieren [253], muss die Frage gestellt werden, ob Skp ebenfalls eine Aufnahme, Stabilisierung und die Insertion von membranassoziierten Tat-Substraten, wie z. B. FdnGH, vermitteln kann. Dies wird von der Beobachtung gestützt, dass Skp mit Fdh-N interagiert.

Die meisten inneren integralen Membranproteine werden durch das allgemeine sekretorische System (Sec) co-translational assembliert [254], wohingegen nur wenige über das Tat-System eingebaut werden [255]. Zur Reifung der Fdhs werden die α - und β -Untereinheiten vermutlich zusammen als Kofaktoren enthaltendes Heterodimer (FdnGH, FdoGH) durch das Tat-System über die Cytoplasmamembran transportiert [46] und anschließend mit den Sec-abhängigen integralen Membranuntereinheiten (FdnI, FdoI) [254] assembliert. Die Assoziation des $\alpha\beta$ -Heterodimers mit der Membran während oder nach der Tat-abhängigen Translokation ist auf die Hilfe eines C-terminalen hydrophoben Transmembranbereichs (*C-tail*) von FdnH bzw. FdoH [256] angewiesen. Vorstellbar ist, dass Skp eine Funktion in der Stabilisierung einer membrangebundenen oder periplasmatischen Zwischenstufe der $\alpha\beta$ -Heterodimere ausüben könnte. Da keine Vernetzungsprodukte an der α -helikalen Transmembrandomäne von FdnH identifiziert wurden (Abbildung 24A), liegt diese während der Interaktion des FdnGH-

Dimers mit Skp vermutlich außerhalb der Substratbindestelle vor. Ein möglicher Grund hierfür wäre, dass die Assoziation des FdnGH-Dimers mit der Cytoplasmamembran bereits während des Tat-Transports über den *C-tail* durchgeführt wird und eine Stabilisierung des FdnGH-Dimers durch Skp zur anschließenden Interaktion mit FdnI für die Reifung des vollständigen Enzyms an der Membran erfolgt. Eine andere Möglichkeit ist die vollständige Entlassung des FdnGH-Dimers in das Periplasma, wobei der *C-tail* des FdnGH-Dimers nicht in die Substratbindetasche aufgenommen wird, damit eine anschließende Membraninsertion und Assoziation mit FdnI erfolgen kann. Beide Modelle werden von der Tatsache gestützt, dass Skp in Phospholipidschichten penetrieren kann [257].

5.3.2 FdnG/OppA-Komplex

Die Genprodukte von *oppABCDF* bilden einen ATP-abhängigen Oligopeptid-Transporter, der zur Superfamilie der ABC (*ATP-binding cassette*)-Transportersysteme gehört [258] und dessen Funktion die Aufnahme von Oligopeptiden und Wiederverwertung von Zellwandpeptiden ist [259, 260]. Das Protein OppA ist eine periplasmatische Komponente des Transporters, welche für die Substrataufnahme verantwortlich ist. Die Bindung der Substrate erfolgt dabei mit Dissoziationskonstanten im nanomolaren Bereich [261]. Die Genprodukte OppBCDF stellen membranassoziierte Komponenten des Oligopeptidtransporters dar. Die Kristallstruktur von OppA wurde mit 2,3 Å aufgelöst und zeigt ein zweilappiges, über drei Domänen enthaltendes Protein, das vorwiegend β -Faltblätter aufweist [262].

Die massenspektrometrische Analyse des Cross-Linking-Ansatzes führte zur Identifizierung von drei Quervernetzungsprodukten zwischen OppA und der katalytischen Untereinheit der Fdh-N (Tabelle 12). Die Reaktionsstellen des verwendeten Quervernetzungsreagenzes SDAD in FdnG befinden sich an zwei Serinen und einem Tyrosin, was bereits beschriebene Umesterungsreaktionen aminreaktiver Quervernetzungsreagenzien mit Hydroxylseitenketten bestätigt [195]. Wie bereits im FdnH/Skp-Komplex beobachtet wurde, sind die Reaktionsstellen des Quervernetzungsreagenzes in FdnG und OppA gerichtet orientiert und nicht zufällig über die globulären Proteine verteilt (Abbildung 24B). In OppA ist eine Reaktionsstelle (A381 oder D382) räumlich relativ weit von den beiden übrigen

Reaktionsstellen (Y121 und D167) getrennt, welche in der Struktur nahe nebeneinander liegen.

Die definierte Anordnung der quervernetzten Aminosäuren lässt auf eine direkte und spezifische Interaktion zwischen FdnG und OppA schließen (Abbildung 24B). Neben einer direkten Interaktion von FdnG und OppA muss eine durch Skp vermittelte Interaktion in Betracht gezogen werden, da dieses ebenfalls einen Interaktionspartner von OppA [253] und FdnH darstellt.

5.3.3 FdnG/ErfK-Komplex

Das Peptidoglykan der bakteriellen Zellwand ist eine hochadaptive und streng regulierte Komponente [263], welche der Zelle als Exoskelett dient [264]. Das engmaschige Netzwerk besteht aus alternierenden Einheiten des Dissacharids *N*-Acetylglucosaminsäure-*N*-Acetylmuraminsäure, welche durch Pentapeptide quervernetzt sind [263]. ErfK stellt eine von fünf in *E. coli* vorkommenden L-, D-Transpeptidasen (Ldt) dar [265]. Diese haben unterschiedliche Funktionen, wobei ErfK eine kovalente Verknüpfung des Braun-Lipoprotein der äußeren Membran mit Peptidoglykan erzeugt, wodurch die Integrität der Zellwand gewährleistet wird [266]. Ein Verlust der Aktivität sämtlicher Ldt-Enzyme in *E. coli* unter einem Mangel an 2,6-Diaminopimelinsäure (DAP) zeigt ein eingeschränktes Zellwachstum und deutet auf eine Funktion von ErfK für die Bewältigung von Peptidoglykanstress hin [265].

Die massenspektrometrische Analyse führte zur Identifizierung eines Quervernetzungsprodukts zwischen ErfK und der katalytischen Untereinheit von Fdh-N (Tabelle 12). Die Beobachtung einer Quervernetzung zu K138 von FdnG indiziert, dass sich die Interaktionsfläche von Fdh-N mit ErfK von denen mit Skp, OppA und TolB unterscheidet. Das Fehlen weiterer Distanzbeschränkungen erschwert jedoch eine genaue Charakterisierung der Grenzfläche des FdnG/ErfK-Komplexes. Die membranassoziierten Untereinheiten der Fdh-N erstrecken sich 90 Å in den periplasmatischen Raum hinein [37], wodurch ein Kontakt der membranassoziierten katalytischen Untereinheiten mit den an die äußere Membran angrenzenden Peptidoglykanschichten erfolgen könnte [264]. Eine Stabilisierung der Fdhs durch ihre Einbindung in strukturelle Komponenten der Zellwand, wie z. B. durch ErfK, welches die strukturelle Integrität der Membran gewährleistet [266], ist vorstellbar und muss in weiterführenden Experimenten näher untersucht werden.

5.3.4 FdnGH/TolB-Komplex

Das Tol/Pal-System von *E. coli* liegt im periplasmatischen Kompartiment und ist vorwiegend der inneren und äußeren Membran zugewandt, wobei seine Funktion der Aufrechterhaltung der Membranintegrität dient [267-270]. Das System erstreckt sich durch den periplasmatischen Raum und stellt eine Verbindung der inneren und äußeren Membran dar [271]. Das System setzt sich aus den Proteinen TolQ, TolR und TolA zusammen, die einen Komplex innerhalb der inneren Membran bilden [267]; das periplasmatische Protein TolB interagiert mit dem Peptidoglykan-assoziierten Lipoprotein (Pal). Pal ist in der äußeren Membran verankert und interagiert mit der C-terminalen Domäne von TolA [272]. Eine *in vivo*-Studie zeigte, dass TolB in einer Pal-abhängigen Weise mit dem großen Lipoprotein der äußeren Membran und OmpA interagiert [273]. Neben einer Funktion für die Stabilisierung der Membran dient die Tol-Maschinerie für die Aufnahme von Colizinen [274].

Im Gegensatz zu den Interaktionspartnern Skp, OppA und ErfK, für welche ausschließlich Quervernetzungsprodukte zur großen katalytischen Untereinheit (FdnG) oder Elektronen-übertragenden Untereinheit (FdnH) identifiziert wurden, konnten für den Fdh-N/TolB-Komplex sowohl Quervernetzungsprodukte zu FdnG als auch zu FdnH identifiziert werden (Tabelle 12 und Abbildung A 15). Auffällig ist, dass die Quervernetzungsprodukte von FdnG und TolB bzw. FdnH und TolB unterschiedliche Regionen von TolB betreffen. Wie im Fall der Interaktion zwischen Fdh-N und dem Chaperon Skp wurden auch für den Fdh-N/TolB-Komplex drei räumlich benachbarte Quervernetzungsprodukte im zentralen Bereich von FdnH gefunden.

Das im Vergleich zum Cytoplasma viskosere Periplasma [275] enthält spezielle Chaperone, die eine Fehlfaltung und Aggregation von Proteinen verhindern [276]. Die Interaktion mit dem Chaperon TolB könnte eine Stabilisierung von Fdh-N innerhalb des gelartigen Kompartiments [277, 278] bewirken. Dies wird von der Beobachtung gestützt, dass Fdh-N ebenfalls mit dem faltungsstabilisierenden Chaperon Skp interagiert. Auffällig ist, dass sowohl für die Interaktion mit Skp als auch TolB keine Quervernetzung mit der α -helikalen Transmembrandomäne von FdnH detektiert werden konnte, welche die Bindung zur Membran vermittelt [256]. Ebenfalls vorstellbar ist die Stabilisierung von Proteinkomplexen, wie z. B. den Fdhs, durch deren strukturelle Einbindung in die Tol/Pal-Nanomaschinerie, welche der inneren und äußeren Membran zugewandt ist [267-269]. Die Stabilisierung der Fdhs

durch ihre Einbindung dieser großen Enzyme in strukturelle Komponenten der Zellwand wird zusätzlich durch die Interaktion von Fdh-N mit Erfk gestützt.

5.4 Charakterisierung der FocA/PflB-Interaktion

Vorausgegangene *in vivo*- und *in vitro*-Protein-Protein-Interaktionsstudien, welche am Institut für Biologie in der Abteilung Mikrobiologie (AG Sawers) der Martin-Luther Universität Halle-Wittenberg durchgeführt wurden, hatten eine Interaktion von PflB mit FocA gezeigt. In dieser Arbeit konnten durch die massenspektrometrische Analyse eines heterobifunktionellen amin-/photoreaktiven Quervernetzungsreagenzes neben PflB weitere mögliche Bindungspartner von FocA identifiziert werden. Aktives PflB enthält ein Glycylradikal, das Sauerstoff-empfindlich ist und in dessen Anwesenheit eine Halbwertszeit von ca. 10 s besitzt [1]. Die in Anwesenheit von Sauerstoff durchgeführten Interaktionsstudien belegen somit die Interaktion von inaktivem Glycylradikal-freiem PflB mit FocA. Chemische Quervernetzungs-experimente führten zudem zur Identifizierung der FocA-Bindungspartner AdhE, PflA sowie Proteinen des fermentativen Stoffwechsels (Tabelle 13 und Tabelle A3). AdhE ist unter anaeroben Bedingungen ein abundantes Enzym und modifiziert PflB post-translational, wodurch dessen Aktivität inhibiert werden kann [279]. Das PflB-aktivierende Enzym PflA führt ebenfalls zu einer post-translationen Modifikation von PflB, wodurch dessen Aktivität kontrolliert wird. Dieses Radikalenzym [280] ermöglicht die Einführung eines Glycylradikals in PflB an Position 734 [2]. Zusätzlich wurden verschiedene Aminosäuredecarboxylasen sowie die Pyruvatdehydrogenase als Bindungspartner von FocA identifiziert, was eine Verbindung zwischen FocA und dem Kohlendioxid/Bicarbonat-Metabolismus impliziert.

Die hohe Abundanz bestimmter Proteine birgt das Risiko der Identifizierung nicht-spezifischer Interaktionspartner während der Untersuchung von Protein-Protein-Wechselwirkungen [236]. So liegt z. B. PflB unter anaeroben Wachstumsbedingungen hochabundant vor [1]. Für die Cross-Linking-Experimente zur Identifizierung der Bindungspartner von FocA erfolgte zudem die Verwendung von löslichen Proteinen des *E. coli*-Stammes REK702, welcher aufgrund einer Mutation im initiierenden Translationskodon von *focA* (GUG → AUG) höhere Mengen an FocA und PflB produziert [12, 281]. Da verschiedene biochemische Methoden, wie z. B. die bakterielle *two hybrid*-Analyse und *pull down*-Experimente zur Analyse der

FocA/PflB-Interaktion angewandt wurden, kann eine unspezifische Interaktion beider Proteine weitgehend ausgeschlossen werden.

In vitro-Experimente zeigten die Translokation verschiedener Metabolite der gemischten Säuregärung durch gereinigtes FocA [282]. *In vivo*-Studien hingegen belegen die ausschließliche Translokation von Formiat durch FocA [281]. Zusätzlich wurde beobachtet, dass eine Inaktivierung von PflB in einer *E. coli*-PflA-Mutante eine Inhibierung des Imports von Formiat zur Folge hat [AG Sawers]. Diese Experimente lassen vermuten, dass die spezifische Interaktion von PflB mit FocA die bidirektionelle Formiattranslokation an der Membran reguliert [29, 283] und gewährleistet, dass nur Formiat als Substrat für FocA dient. So könnte das physiologisch bedingte Ausbleiben der Formiatbildung unter aeroben Wachstumsbedingungen eine Interaktion von inaktivem PflB mit FocA zur Folge haben, wodurch der Import von Formiat über FocA verhindert wird. Dies ist physiologisch relevant, da die Anwesenheit externer Elektronenakzeptoren, wie Sauerstoff, die Synthese des FHL-Komplexes unterdrücken [4, 35] und eine Oxidation von Formiat über die respiratorischen Fdhs des periplasmatischen Kompartiments erfolgt [32].

Eine Interaktion von aktivem PflB mit FocA kann momentan nicht ausgeschlossen und muss in zukünftigen Studien genauer charakterisiert werden. Diese könnte beispielsweise eine andere Konformation des FocA/PflB-Komplexes bzw. einen veränderten Import bzw. Export von Formiat zeigen.

Zur strukturellen Charakterisierung der FocA/PflB-Interaktion erfolgte die massenspektrometrische Analyse eines Quervernetzungsexperimentes mit einem homobifunktionellen amin-reaktiven Cross-Linker in Kombination mit einer bioinformatischen Modellierung des FocA/PflB-Komplexes mittels Rosetta. Elf von 14 erhaltenen intermolekularen Distanzbeschränkungen verfügten über Reaktionsstellen (K17, Y25 oder K26, T28 und K29) im *N*-terminalen Bereich von FocA (Tabelle 14). Dieses Ergebnis und die Erkenntnisse der in der AG Sawers durchgeführten bakteriellen *two-hybrid*-Interaktionsstudie zeigen die große Bedeutung der *N*-terminalen Region von FocA für die Interaktion mit PflB. Die Cross-Linking-Ergebnisse deuten auch auf eine hohe Flexibilität des *N*-Terminus von FocA hin. Diese Flexibilität wurde ebenfalls durch die Strukturvorhersage mittels Rosettascripts gezeigt (Abbildung 31) und wird durch die Beobachtung multipler Konformationen des *N*-Terminus gestützt [27]. Die Reaktionsstellen der verbleibenden

Quervernetzungsprodukte in FocA wurden in direkter Nähe zur zentralen Pore beobachtet. In Anbetracht der möglichen cytoplasmatischen Interaktionsfläche von FocA und der Größe von PflB sowie der Cross-Linker-Reaktionsstellen an der zentralen Pore von FocA wurde eine 1 : 1-Stöchiometrie von pentamerem FocA und dimerem PflB angenommen. Die Notwendigkeit der Zugabe eines Detergenz zur Verbesserung der Löslichkeit von FocA beeinträchtigte allerdings die direkte Bestimmung der FocA/PflB-Stöchiometrie. Diese sollte in weiteren Experimenten, wie z. B. nativer MS oder analytischer Ultrazentrifugation geklärt werden.

Die maximal durch BS²G überbrückbare Ca-Ca-Distanz beträgt 25 Å [284]. Aufgrund der beobachteten Flexibilität des *N*-Terminus von FocA wurde die Ca-Ca-Distanzbeschränkung auf maximal 40 Å erhöht. Die Aufklärung der Bindungsorientierung von PflB an FocA mit Hilfe der Distanzbeschränkungen in der AG Meiler führte zur Identifizierung der zwei biologisch relevanten Konformationen A und B (Abbildung 34). Diese könnten während des chemischen Cross-Linking-Experiments in Lösung vorhanden gewesen sein und spiegeln somit eine dynamische Konformationsänderung des FocA/PflB-Komplexes wieder. Es wird postuliert, dass die Konformation A eine entropisch begünstigte Zwischenstufe einer initialen Komplexbildung zwischen FocA und PflB darstellt. In dieser bildet PflB über nur eine Untereinheit einen Kontakt zu FocA aus. Ausgehend von Konformation A kommt es zur Ausbildung der enthalpisch bevorzugten Konformation B, in welcher PflB über die Bindung beider Unterheiten einen engeren Kontakt mit einer vermutlich höheren Affinität zu FocA ausbildet. Die Konformationsänderung geht mit einer Neuorientierung von PflB einher, wobei die Eintrittsstelle des CoA-Moleküls von PflB [8] zentral vor die cytoplasmatische Seite von FocA positioniert wird.

Die Austrittsstelle von Formiat an PflB ist momentan nicht aufgeklärt [6, 8]. Wenn der Austritt von Formiat an der Eintrittsstelle des CoA-Moleküls erfolgen sollte, könnte aktives PflB zu einer Konzentrationserhöhung von Formiat an FocA beitragen, welches schnell über offene Substratkanäle exportiert werden könnte. Ein lokaler Gradient von Formiat an der Cytoplasmamembran könnte zu einer optimalen Translokation des Metaboliten über den Formiatkanal beitragen und würde vielmehr eine Ansäuerung des gesamten Cytoplasmas verhindern. Diese Hypothese kann jedoch aufgrund der geringen Auflösung des FocA/PflB-Modells nicht bestätigt werden. Zudem sind keine Aminosäuren in PflB bekannt, welche aktiv die Leitung von Formiat zu den Substratkanälen von FocA ermöglichen könnten. In Abwesenheit

von PflB könnte die Formiattranslokation durch eine geschlossene Konformationen der *N*-Termini des FocA-Pentamers verhindert werden. Eine kürzlich veröffentlichte Studie zur Moleküldynamik an FocA zeigte jedoch, dass die geschlossene Konformation des *N*-Terminus den Im- und Export von Formiat nicht zum Erliegen bringt [285]. Diese Erkenntnisse deuten auf die Notwendigkeit einer Interaktion von FocA mit PflB für eine kontrollierte Translokation von Formiat hin.

Die genaue Unterscheidung zwischen den Konformationen A und B wird durch die Flexibilität der *N*-Termini von FocA erschwert. Die Kristallstrukturen zeigen die fehlenden Beugungsdaten der *N*-terminalen α -Helices [25-27]. Die in dieser Arbeit durchgeführte Strukturvorhersage des *N*-Terminus konnte ebenfalls keine Informationen über die Struktur ersten neun Aminosäuren des *N*-Terminus liefern (Abbildung 30). Diese fehlende Information kann die Konformation des FocA/PflB-Modells beeinflussen. Der in der AG Meiler durchgeführte rigide Austausch der *N*-Termini an der cytoplasmatischen Seite von FocA kann zudem die Interaktion von FocA und PflB limitieren. Die vollständige Strukturvorhersage der *N*-Termini von FocA könnte die Orientierung beider Interaktionspartner verbessern und die Herabsetzung der maximalen C α -C α -Distanz für die Erfüllung der Cross-Linking-Distanzbeschränkung ermöglichen.

Eine mögliche Interpretation der Ergebnisse könnte auch sein, dass eine der beiden Konformationen ein Artefakt des Quervernetzungsexperiments oder der Modellierung bzw. Filterung der FocA/PflB-Modelle darstellt. Die Tatsache, dass nur die Konformation B, trotz höherer C α -C α -Distanzen, alle 14 Beschränkungen erfüllt, könnte bedeuten, dass ausschließlich diese Konformation von Relevanz ist. Eine experimentelle Bestätigung der erhaltenen Konformationen könnte z. B. durch Ionenmobilitäts-MS erfolgen. Zudem muss in zukünftigen Experimenten die Bestimmung der Assoziationskonstante zwischen FocA und PflB ermittelt werden.

Anhang

>3KCU:A | PDBID | CHAIN | SEQUENCE

MKADNPF~~DL~~LLPAAMAKVAEEAGVYKATKHPLKTFYLAITAGVFISIAFVFYITATTGTGTMPPFGMAKLVGGICF
SLGLILCVVCGADLFTSTVLIVVAKASGRITWQQLAKNWLNVYFGNLVGALLFVLLMWLSGEYMTANGQWGLNVL
QTADHKVHHTFIEAVCLGILANLMVCLAVWMSYSGRSLMDKAFIMVLPVAMFVASGFEHSIANMFMI PMGIVIRD
FASPEFWTAVGSAPENFSHLTVMNFITDNLIPVTIGNI IGGLLVGLTYWVIYLR~~END~~H

>3KCU:B | PDBID | CHAIN | SEQUENCE

MKADNPF~~DL~~LLPAAMAKVAEEAGVYKATKHPLKTFYLAITAGVFISIAFVFYITATTGTGTMPPFGMAKLVGGICF
SLGLILCVVCGADLFTSTVLIVVAKASGRITWQQLAKNWLNVYFGNLVGALLFVLLMWLSGEYMTANGQWGLNVL
QTADHKVHHTFIEAVCLGILANLMVCLAVWMSYSGRSLMDKAFIMVLPVAMFVASGFEHSIANMFMI PMGIVIRD
FASPEFWTAVGSAPENFSHLTVMNFITDNLIPVTIGNI IGGLLVGLTYWVIYLR~~END~~H

>3KCU:C | PDBID | CHAIN | SEQUENCE

MKADNPF~~DL~~LLPAAMAKVAEEAGVYKATKHPLKTFYLAITAGVFISIAFVFYITATTGTGTMPPFGMAKLVGGICF
SLGLILCVVCGADLFTSTVLIVVAKASGRITWQQLAKNWLNVYFGNLVGALLFVLLMWLSGEYMTANGQWGLNVL
QTADHKVHHTFIEAVCLGILANLMVCLAVWMSYSGRSLMDKAFIMVLPVAMFVASGFEHSIANMFMI PMGIVIRD
FASPEFWTAVGSAPENFSHLTVMNFITDNLIPVTIGNI IGGLLVGLTYWVIYLR~~END~~H

>3KCU:D | PDBID | CHAIN | SEQUENCE

MKADNPF~~DL~~LLPAAMAKVAEEAGVYKATKHPLKTFYLAITAGVFISIAFVFYITATTGTGTMPPFGMAKLVGGICF
SLGLILCVVCGADLFTSTVLIVVAKASGRITWQQLAKNWLNVYFGNLVGALLFVLLMWLSGEYMTANGQWGLNVL
QTADHKVHHTFIEAVCLGILANLMVCLAVWMSYSGRSLMDKAFIMVLPVAMFVASGFEHSIANMFMI PMGIVIRD
FASPEFWTAVGSAPENFSHLTVMNFITDNLIPVTIGNI IGGLLVGLTYWVIYLR~~END~~H

>3KCU:E | PDBID | CHAIN | SEQUENCE

MKADNPF~~DL~~LLPAAMAKVAEEAGVYKATKHPLKTFYLAITAGVFISIAFVFYITATTGTGTMPPFGMAKLVGGICF
SLGLILCVVCGADLFTSTVLIVVAKASGRITWQQLAKNWLNVYFGNLVGALLFVLLMWLSGEYMTANGQWGLNVL
QTADHKVHHTFIEAVCLGILANLMVCLAVWMSYSGRSLMDKAFIMVLPVAMFVASGFEHSIANMFMI PMGIVIRD
FASPEFWTAVGSAPENFSHLTVMNFITDNLIPVTIGNI IGGLLVGLTYWVIYLR~~END~~H

Abbildung A 1: Primärstruktur von FocA aus *E. coli*. In der Kristallstruktur von FocA aus *E. coli* (PDB-Eintrag 3KCU) sind strukturell fehlende Bereiche in rot dargestellt.

TM region prediction for 3KCU5,topo predicted using OCTOPUS
30 1289
antiparallel
n2c

34	54	34	54
66	96	66	96
115	135	115	135
162	182	162	182
193	223	193	223
256	276	256	276
291	311	291	311
323	353	323	353
372	392	372	392
419	439	419	439
450	480	450	480
513	533	513	533
543	563	543	563
575	605	575	605
624	644	624	644
671	691	671	691
702	732	702	732
765	785	765	785
791	811	791	811
823	853	823	853
872	892	872	892
919	939	919	939
950	980	950	980
1013	1033	1013	1033
1038	1058	1038	1058
1070	1100	1070	1100
1119	1139	1119	1139
1166	1186	1166	1186
1197	1227	1197	1227
1260	1280	1260	1280

Abbildung A 2: Span-Datei für die Modellierung des N-terminalen Bereichs (Aminosäuren 1 - 32) von FocA.

```
-parser:protocol build_3KCU_termini.xml  
-database  
/blue/meilerlab/home/nannemdp/rosetta_clean/rosetta/rosetta_database/  
-mute core,scoring,MembranePotential  
protocols,rosetta_scripts,ParsedProtocol,REPORT  
-ex1  
-ex2  
-use_input_sc  
-linmem_ig 10
```

Abbildung A 3: Flag-Datei für die Modellierung des N-terminalen Bereichs (Aminosäuren 1 - 32) von FocA.

Overview:

Stage 1 : 9mers, score0,wts_patch, 2000 cycles, uniform policy
 Stage 2 : 9mers, score1,wts_patch, 2000 cycles, uniform policy
 Stage 3a: 9mers, score2,wts_patch, 2000 cycles, uniform policy
 Stage 3b: 9mers, score5,wts_patch, 2000 cycles, uniform policy
 Stage 4 : 3mers, score3,wts_patch, 4000 cycles, smooth policy

```
bin/rosetta_scripts,macosgccrelease -fasta 1a19a.fasta -parser:protocol
classic_abinitio.xml -parser:view -database ~/minirosetta_database/
```

```
<dock_design>
  <SCOREFXNS>
    <score0
weights="/rosetta_clean/rosetta/rosetta_database/scoring/weights/score0,wts
" />
      <score_membrane_2
weights="rosetta_clean/rosetta/rosetta_database/scoring/weights/score1,wts"
patch="/blue/meilerlab/home/zornm0/scripter_fragment_insertion/score_membra
ne_s2,wts_patch" />
      <score_membrane_3a
weights="rosetta_clean/rosetta/rosetta_database/scoring/weights/score2,wts"
patch="/blue/meilerlab/home/zornm0/scripter_fragment_insertion/score_membra
ne_s3a,wts_patch" />
      <score_membrane_3b
weights="rosetta_clean/rosetta/rosetta_database/scoring/weights/score5,wts"
patch="/blue/meilerlab/home/zornm0/scripter_fragment_insertion/score_membra
ne_s3b,wts_patch" />
      <score_membrane_4
weights="rosetta/rosetta_database/scoring/weights/score3,wts"
patch="score_membrane_s4,wts_patch" />
    </SCOREFXNS>
    <TASKOPERATIONS>
      <InitializeFromCommandline name=ifcl />
      <RestrictToRepacking name=rtr />
    </TASKOPERATIONS>
    <FILTERS>
  </FILTERS>
  <MOVERS>
    Defines the base mover for small fragments (typically 3-mers)
    <SingleFragmentMover name=sfm_small_N-term
fragments=aa3KCU503_05,200_v1_3 policy=smooth>
      <MoveMap>
        <Span begin=33 end=1289 chi=1 bb=0 />
      </MoveMap>
    </SingleFragmentMover>
    <SingleFragmentMover name=sfm_small_A-C-term
fragments=aa3KCU503_05,200_v1_3 policy=smooth>
      <MoveMap>
        <Span begin=1 end=275 chi=1 bb=0 />
        <Span begin=286 end=1289 chi=1 bb=0 />
      </MoveMap>
    </SingleFragmentMover>
    <SingleFragmentMover name=sfm_small_E-C-term
fragments=aa3KCU503_05,200_v1_3 policy=smooth>
      <MoveMap>
        <Span begin=1 end=1279 chi=1 bb=0 />
      </MoveMap>
    </SingleFragmentMover>
    Defines the base mover for large fragments (typically 9-mers)
```

```

    <SingleFragmentMover name=sfm_large_N-term
fragments=aa3KCU509_05,200_v1_3 policy=uniform>
    <MoveMap>
        <Span begin=33 end=1289 chi=1 bb=0 />
    </MoveMap>
</SingleFragmentMover>
    <SingleFragmentMover name=sfm_large_A-C-term
fragments=aa3KCU509_05,200_v1_3 policy=uniform>
    <MoveMap>
        <Span begin=1 end=275 chi=1 bb=0 />
        <Span begin=286 end=1289 chi=1 bb=0 />
    </MoveMap>
</SingleFragmentMover>
    <SingleFragmentMover name=sfm_large_E-C-term
fragments=aa3KCU509_05,200_v1_3 policy=uniform>
    <MoveMap>
        <Span begin=1 end=1279 chi=1 bb=0 />
    </MoveMap>
</SingleFragmentMover>

    <ParsedProtocol name=small mode=single_random >
        <Add mover_name=sfm_small_N-term apply_probability=0,8/>
        <Add mover_name=sfm_small_A-C-term apply_probability=0,1/>
        <Add mover_name=sfm_small_E-C-term
apply_probability=0,2/>
    </ParsedProtocol>
    <ParsedProtocol name=large mode=single_random >
        <Add mover_name=sfm_large_N-term
apply_probability=0,8/>
        <Add mover_name=sfm_large_A-C-term
apply_probability=0,1/>
        <Add mover_name=sfm_large_E-C-term
apply_probability=0,2/>
    </ParsedProtocol>

    Wrap the base movers in GenericMonteCarlo objects for scoring
    <GenericMonteCarlo name=stage1 scorefxn_name=score0
mover_name=large temperature=2,0 trials=20/>
    <GenericMonteCarlo name=stage2 scorefxn_name=score_membrane_2
mover_name=large temperature=2,0 trials=20/>
    <GenericMonteCarlo name=stage3a scorefxn_name=score_membrane_3a
mover_name=large temperature=2,0 trials=40/>
    <GenericMonteCarlo name=stage3b scorefxn_name=score_membrane_3b
mover_name=large temperature=2,0 trials=40/>
    <GenericMonteCarlo name=stage4 scorefxn_name=score_membrane_4
mover_name=small temperature=2,0 trials=60/>
    Converts the centroid-level pose to fullatom for scoring and
vice-versa
    <SwitchResidueTypeSetMover name=fullatom set=fa_standard/>
    <SwitchResidueTypeSetMover name=centroid set=centroid/>

    Allows use of membrane scoring terms
    <MembraneTopology name=membrane span_file=3KCU5,span/>
    <FastRelax name=relax scorefxn=score12 repeats=1 >
        <MoveMap>
            <Span begin=1 end=34 chi=1 bb=1 />
            <Span begin=35 end=1278 chi=1 bb=0 />
            <Span begin=1279 end=1289 chi=1 bb=1 />
        </MoveMap>
    </FastRelax>

```

```
</MOVERS>
<APPLY_TO_POSE>
</APPLY_TO_POSE>
<PROTOCOLS>
  <Add mover_name=membrane />
  <Add mover_name=centroid />
  <Add mover_name=stage1 />
  <Add mover_name=stage2 />
  <Add mover_name=stage3a />
  <Add mover_name=stage3b />
  <Add mover_name=stage4 />
  <Add mover_name=fullatom />
  <Add mover_name=relax />
</PROTOCOLS>
</dock_design>
```

Abbildung A 4: XML-Datei für die Modellierung des N-terminalen Bereichs (Aminosäuren 1 - 32) von FocA.

```
-database /home/nannemdp/rosetta_clean/rosetta/rosetta_database
-score:weights membrane_highres_Menv_smooth
-in:file:fullatom
-cluster:radius 1,5
-cluster:sort_groups_by_energy
-cluster:exclude_res 1 2 3 4 5 6 7 8 9 286 ... 1289
```

Abbildung A 5: Flag-Datei für die Clusteranalyse des N-terminalen Bereichs (Aminosäuren 1 – 32) der α -Untereinheit von FocA. Die Clusteranalyse erfolgte mit Hilfe des Rosetta-Algorithmus. Der Clusterradius wurde mit 1,5 Å RMSD (C α -C α -Distanzen) definiert. Die Clusteranalyse wurde auf die Aminosäuren 10 – 285 der α -Untereinheit von FocA begrenzt.

```
#!/usr/bin/env python2,5

from random import shuffle
from rosettautil,protein import util
from Bio,PDB import *
from optparse import OptionParser
import warnings

usage = "%prog [options] pdb_input_files"
parser=OptionParser(usage)
parser.add_option("--number_strucs", dest="number_strucs",help="number of
structures to build")
(options,args)= parser.parse_args()
warnings.simplefilter('ignore',PDBExceptions,PDBConstructionWarning)

template = PDBParser(),get_structure('3KCU','3KCU5_termini,pdb')

template_list=['A','B','C','D','E']

template_A_atoms = []
template_B_atoms = []
template_C_atoms = []
template_D_atoms = []
template_E_atoms = []

template_chain_dict=[93]
for model in template:
    for chain in model:
        if chain.get_id() == 'A':
            for residue in chain:
                template_A_atoms.append(residue['N'])
                template_A_atoms.append(residue['CA'])
                template_A_atoms.append(residue['C'])
                template_A_atoms.append(residue['O'])
            template_chain_dict['A']=template_A_atoms
        if chain.get_id() == 'B':
            for residue in chain:
                template_B_atoms.append(residue['N'])
                template_B_atoms.append(residue['CA'])
                template_B_atoms.append(residue['C'])
                template_B_atoms.append(residue['O'])
            template_chain_dict['B']=template_B_atoms
        if chain.get_id() == 'C':
            for residue in chain:
                template_C_atoms.append(residue['N'])
                template_C_atoms.append(residue['CA'])
                template_C_atoms.append(residue['C'])
                template_C_atoms.append(residue['O'])
            template_chain_dict['C']=template_C_atoms
        if chain.get_id() == 'D':
            for residue in chain:
                template_D_atoms.append(residue['N'])
                template_D_atoms.append(residue['CA'])
                template_D_atoms.append(residue['C'])
                template_D_atoms.append(residue['O'])
            template_chain_dict['D']=template_D_atoms
        if chain.get_id() == 'E':
            for residue in chain:
                template_E_atoms.append(residue['N'])
                template_E_atoms.append(residue['CA'])
                template_E_atoms.append(residue['C'])
```

```

        template_E_atoms,append(residue['O'])
        template_chain_dict['E']=template_E_atoms

protomer_list = []

for structure_file in args:
    tag = structure_file,split("/"),pop(),split(",")[0]
    protomer_list,append(tag)

working_list=[]
for i in range(int(options,number_strucs)):
    subunit_list=[]
    for template_id in template_list:
        if len(working_list) < 5:
            working_list = protomer_list[:]
            shuffle(working_list)

            current_tag = working_list,pop()
            print i, template_id, current_tag
            structure =
PDBParser(),get_structure(current_tag,str(current_tag+'.pdb'))
            #structure_atoms = structure,get_atoms()
            protomer_atoms=[]
            for model in structure:
                for chain in model:
                    chain,id =str(template_id)
                    # print chain,id
                    for residue in chain:
                        if residue,get_id()[1] in range(30,277):
                            protomer_atoms,append(residue['N'])
                            protomer_atoms,append(residue['CA'])
                            protomer_atoms,append(residue['C'])
                            protomer_atoms,append(residue['O'])

            superpose = Superimposer()

superpose,set_atoms(template_chain_dict[template_id],protomer_atoms)
#print superpose,rotran
#print superpose,rms
structure_atoms = structure,get_atoms()
#for atom in structure_atoms:
#    print atom,get_id()
superpose,apply(structure,get_atoms())

#for atom in structure_atoms:
#    print atom,get_id()

name = str(str(i)+template_id+'.pdb')
subunit_list,append(name)
w = PDBIO()
w,set_structure(structure)
w,save(name)

for pdb in subunit_list:
    pdb_file = open(pdb,'r')
    combined_pdb = open(str(i)+'.pdb','a')
    for line in pdb_file,readlines():
        line,strip()
        combined_pdb,write(line)
    pdb_file,close()
    combined_pdb,close()

```

Abbildung A 6: Anwendung *make_FocA_pentamer.py* für die Rekonstruktion von FocA-Pentameren.

```

<dock_design>
  <SCOREFXNS>
    <mem_cen_sf weights=score_membrane />
    <mem_fa_sf weights=membrane_highres_Menv_smooth/>
  </SCOREFXNS>
  <TASKOPERATIONS>
    <InitializeFromCommandline name=ifcl/>
    <RestrictToRepacking name=rtr />
  </TASKOPERATIONS>
  <FILTERS>
  </FILTERS>
  <MOVERS>
    <MembraneTopology name=membrane span_file=3KCU,span />
    <PackRotamersMover name=repack scorefxn=mem_fa_sf
task_operations=ifcl,rtr/>
    <TaskAwareMinMover name=min scorefxn=mem_fa_sf bb=1 chi=1
jump=0 type=dfpmin_armijo_nonmonotone tolerance=0,00025 />
  </MOVERS>
  <APPLY_TO_POSE>
  </APPLY_TO_POSE>
  <PROTOCOLS>
    <Add mover=membrane />
    <Add mover=repack />
    <Add mover=min />
  </PROTOCOLS>
</dock_design>

```

Abbildung A 7: XML-Datei für die Relaxierung rekonstruierter FocA-Pentamere.

```

-in
  -file
    -fullatom
    -spanfile
/blue/meilerlab/home/zornm0/relax/cluster_reconstruction/3KCU,span
-out:file:fullatom
-database
/blue/meilerlab/home/nannemdp/rosetta_clean/rosetta/rosetta_database/
-nstruct 1
-linmem_ig 10
-use_input_sc
-ex1
-ex2

```

Abbildung A 8: Flag-Datei für die Relaxierung rekonstruierter FocA-Pentamere.

TM region prediction for octopusnew,topo predicted using OCTOPUS
30 1380
antiparallel
n2c

25	45	25	45
57	87	57	87
106	126	106	126
153	173	153	173
184	214	184	214
247	267	247	267
301	321	301	321
333	363	333	363
382	402	382	402
429	449	429	449
460	490	460	490
523	543	523	543
577	597	577	597
609	639	609	639
658	678	658	678
705	725	705	725
736	766	736	766
799	819	799	819
853	873	853	873
885	915	885	915
934	954	934	954
981	1001	981	1001
1012	1042	1012	1042
1075	1095	1075	1095
1129	1149	1129	1149
1161	1191	1161	1191
1210	1230	1210	1230
1257	1277	1257	1277
1288	1318	1288	1318
1351	1371	1351	1371

Abbildung A 9: Span-Datei für die Relaxierung rekonstruierter FocA-Pentamere.

```

<dock_design>
  <SCOREFXNS>
</SCOREFXNS>
  <TASKOPERATIONS>
    <InitializeFromCommandline name=ifcl/>
    <RestrictToRepacking name=rtr/>
    <RestrictToInterface name=rti />
  </TASKOPERATIONS>
  <FILTERS>
</FILTERS>
  <MOVERS>
    <Docking name=dock score_low=score_docking_low
score_high=score12 fullatom=0 local_refine=0 optimize_fold_tree=1
conserve_foldtree=1 design=0 task_operations=ifcl/>
    PackRotamersMover name=pr scorefxn=score12 task_operations=rtr
  />
    <TaskAwareMinMover name=min scorefxn=score12 chi=1 bb=1 jump=5
tolerance=0,001 task_operations=ifcl,rti/>
  </MOVERS>
  <APPLY_TO_POSE>
</APPLY_TO_POSE>
  <PROTOCOLS>
    <Add mover_name=dock/>
    Add mover_name=pr/>
    Add mover_name=min/>
  </PROTOCOLS>
</dock_design>

```

Abbildung A 10: XML-Datei für die Modellierung des FocA/PfIB-Komplexes. Das *Docking* von FocA und PfIB erfolgte mit Hilfe der Anwendung *rosetta_scripts.de-fault.linuxgccrelease*.

```

-docking
  -dock_pert 3 8
  -spin 1
  -randomize2
  -dock_mcm_trans_magnitude 0,1
  dock_mcm_rot_magnitude 1
  -partners ABCDE_FG
-database
/blue/meilerlab/home/nannemdp/rosetta_clean/rosetta/rosetta_database/
-nstruct 2
-linmem_ig 10
-packing:repack_only
-use_input_sc
-ex1
-ex2
-jd2:ntrials 10
-in:file:fullatom
-out:file:fullatom

```

Abbildung A 11: Flag-Datei für die Modellierung des FocA/PfIB-Komplexes. Das *Docking* von FocA und PfIB erfolgte mit Hilfe der Anwendung *rosetta_scripts.de-fault.linuxgccrelease*.

```
#!/usr/bin/env python2,7

import sys
sys,path,append("/blue/meilerlab/apps/Linux2/x86_64/python/2,5,2/lib/python
2,5/site-packages")
from Bio,PDB import *
from optparse import OptionParser
import warnings
import math
from multiprocessing import Pool
import time

usage = "%prog [options] pdb_input_files"
parser=OptionParser(usage)
parser.add_option("--protomer_order", dest="protomer_order",help="the new
chain name order as a string",default="")
parser.add_option("--pairwise_rmsd_list_name",
dest="listed_pairwise_rmsds",help="file name for a list of pairwise
rmsds",default="tmp_rmsd_list.txt")
parser.add_option("--pairwise_rmsd_table_name",
dest="pairwise_rmsd_table",help="file name for the pairwise rmsd
table",default="tmp_rmsd_table.csv")
parser.add_option("--bcl",dest="bcl",help="add bcl matrix-type tag to first
row of file,",default=False,action="store_true")
parser.add_option("--number_procs","--num_processors","--
number_processors","--num_procs",dest="num_procs",help="number of worker
processors to use",default=1,type="int")
(options,args)= parser.parse_args()
warnings.simplefilter('ignore',PDBExceptions,PDBConstructionWarning)

####GLOBAL DATA####
protomer_order_list = []
if options.protomer_order == "":
    protomer_order_list.append('ABCDEFGF')
    protomer_order_list.append('BCDEAFG')
    protomer_order_list.append('CDEABFG')
    protomer_order_list.append('DEABCFG')
    protomer_order_list.append('EABCFG')
    protomer_order_list.append('ABCDEGF')
    protomer_order_list.append('BCDEAGF')
    protomer_order_list.append('CDEABGF')
    protomer_order_list.append('DEABCGF')
    protomer_order_list.append('EABCDGF')
else:
    protomer_order_list.append(options.protomer_order)
x=len(args)

####Calculatate rmsd between two strucs####
def chain_rmsd_calculator(rmsd_square):
    i = rmsd_square[0]
    j = rmsd_square[1]
    pdb1 = args[i]
    pdb2 = args[j]
    pdb1_structure = PDBParser().get_structure(str(pdb1),pdb1)
    pdb2_structure = PDBParser().get_structure(str(pdb2),pdb2)
    pdb1_atom_list = []
    for model in pdb1_structure:
        for chain in model:
            for residue in chain:
                pdb1_atom_list.append(residue['CA'])
    min_rms = 100000
```

```

min_order = ''
for chain_order in protomer_order_list:
    pdb2_atom_list = []
    for model in pdb2_structure:
        for chain_order_id in chain_order:
            for chain in model:
                if chain.get_id() == chain_order_id:
                    for residue in chain:

pdb2_atom_list.append(residue['CA'])
superpose = Superimposer()
superpose.set_atoms(pdb1_atom_list,pdb2_atom_list)
if superpose.rms < min_rms:
    min_rms = superpose.rms
    min_order = chain_order
order_rmsd = (min_order,min_rms)
return (i,j,min_order,min_rms)

if __name__ == "__main__":

    combination_list = []
    for i in range(0,x-1):
        for j in range(1,x):
            if j<=i:
                continue
            else:
                combination_list.append((i,j))
    #####Parallelization#####
    pool = Pool(processes=options.num_procs)
    rmsd_list = pool.map(chain_rmsd_calculator, combination_list)
    pool.close()

    #####INPUT/OUTPUT#####
    best_fit_file = open(options.listed_pairwise_rmsds,'w')
    rmsd_table = open(options.pairwise_rmsd_table,'w')

    if options.bcl:
        rmsd_table.write("bcl::storage::Table<upper> ")
    for dock_model in args:
        rmsd_table.write(dock_model+" ")
    rmsd_table.write("\n")

    for i in range(0,x-1):
        rmsd_table.write(args[i]+" 0 ")
        for j in range(1,x):
            if j<=i:
                rmsd_table.write("0 ")
            else:
                for square in rmsd_list:
                    if square[0] == i and square[1] == j:
                        pdb1 = args[i]
                        pdb2 = args[j]
                        min_order = square[2]
                        min_rmsd = square[3]
                        best_fit_file.write(str('%s %s %s %s \n'
% (pdb1, pdb2, min_order, min_rmsd)))
                        rmsd_table.write(str(min_rmsd)+" ")
                rmsd_table.write("\n")
    rmsd_table.close()
    best_fit_file.close()

```

Abbildung A 12: Docking_minimal_rmsd_calculator_parallel.py.

```
#!/usr/bin/env python2,7

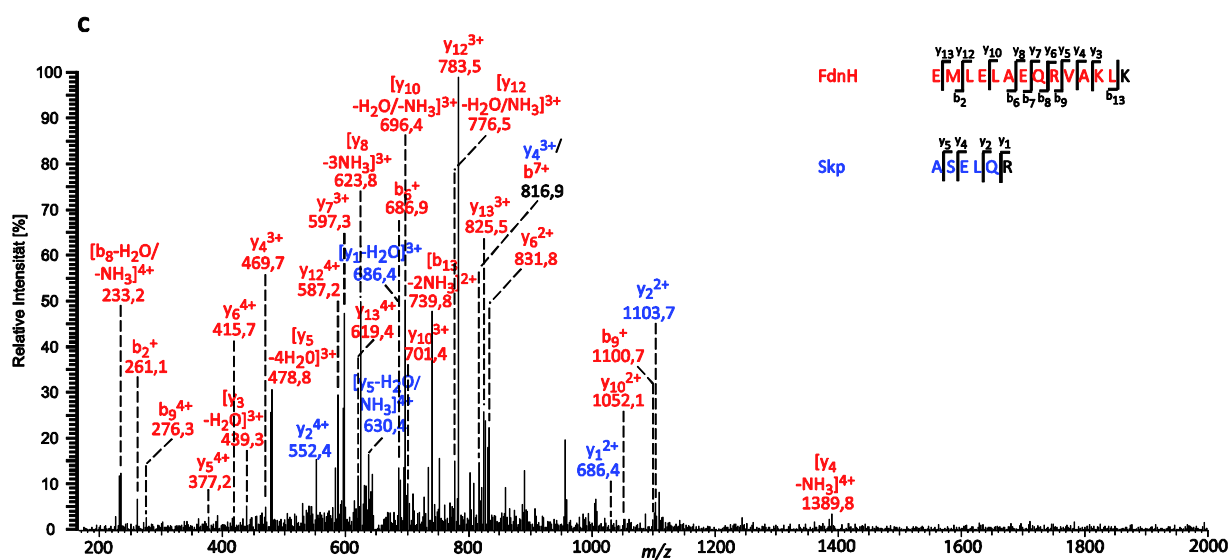
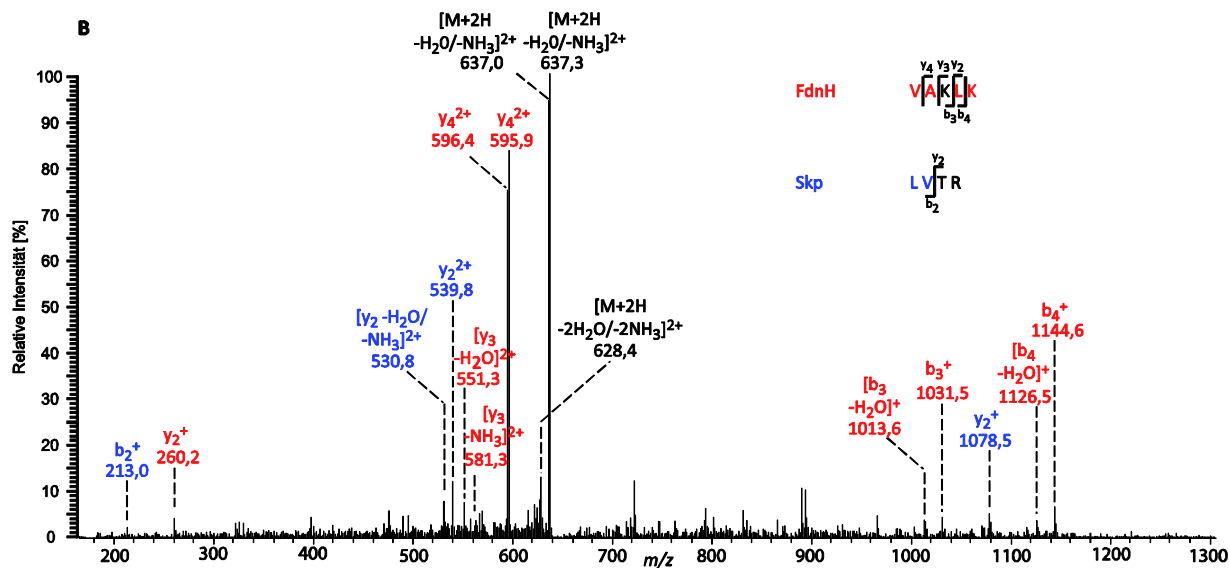
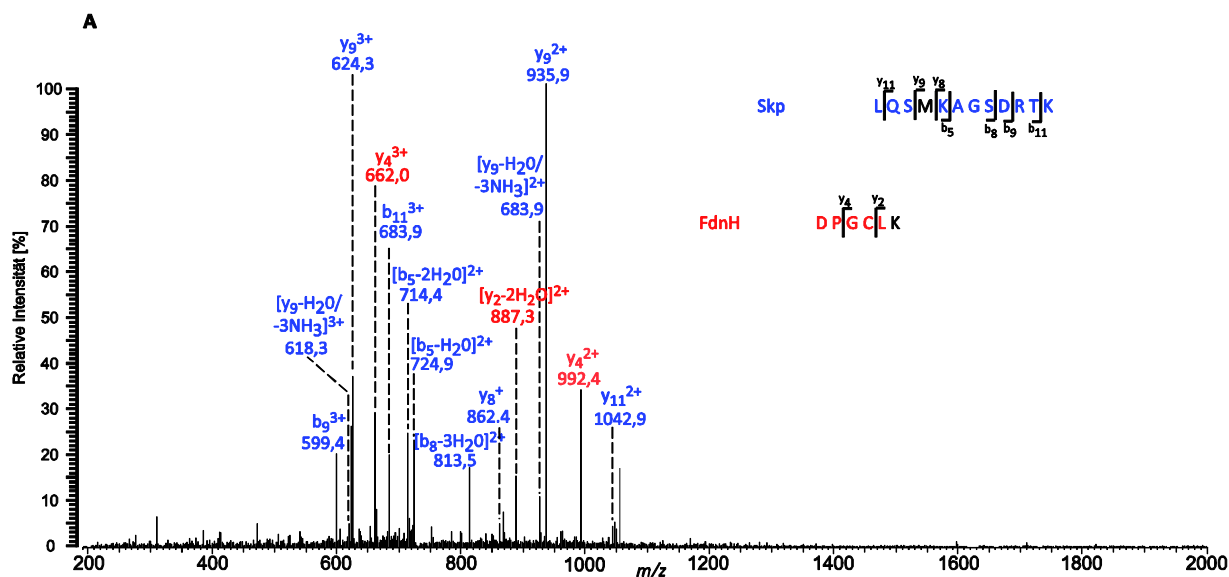
import sys

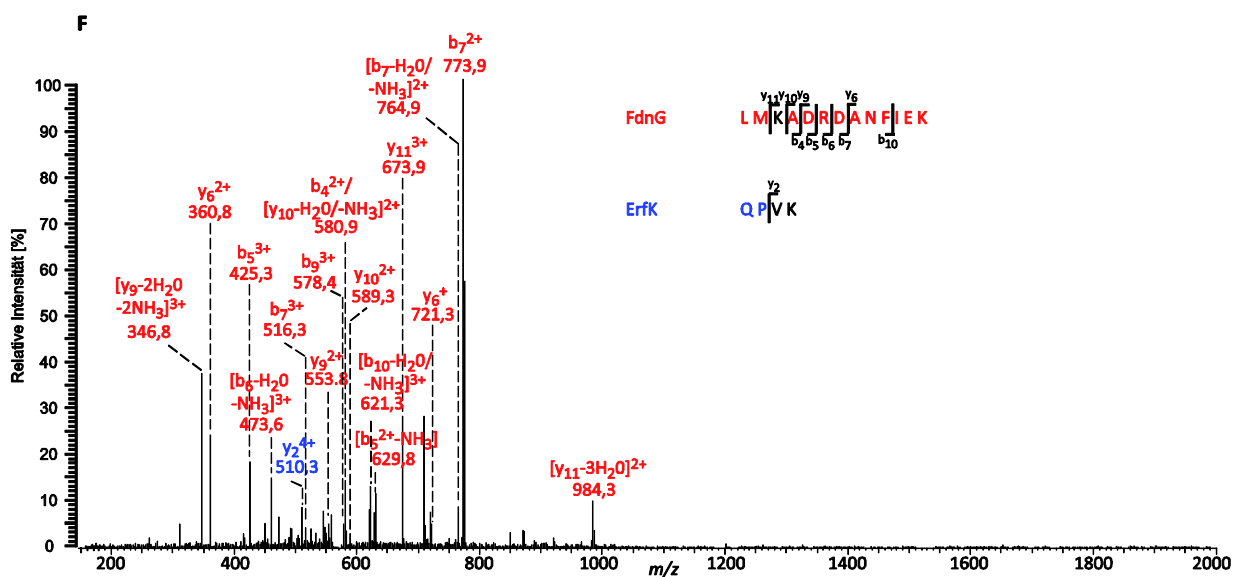
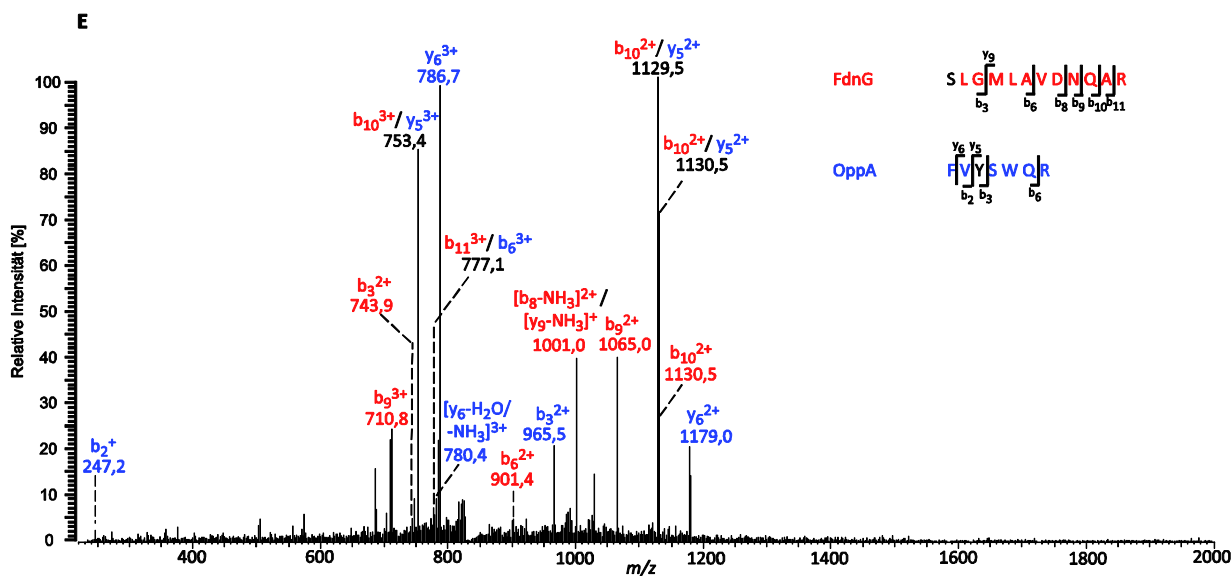
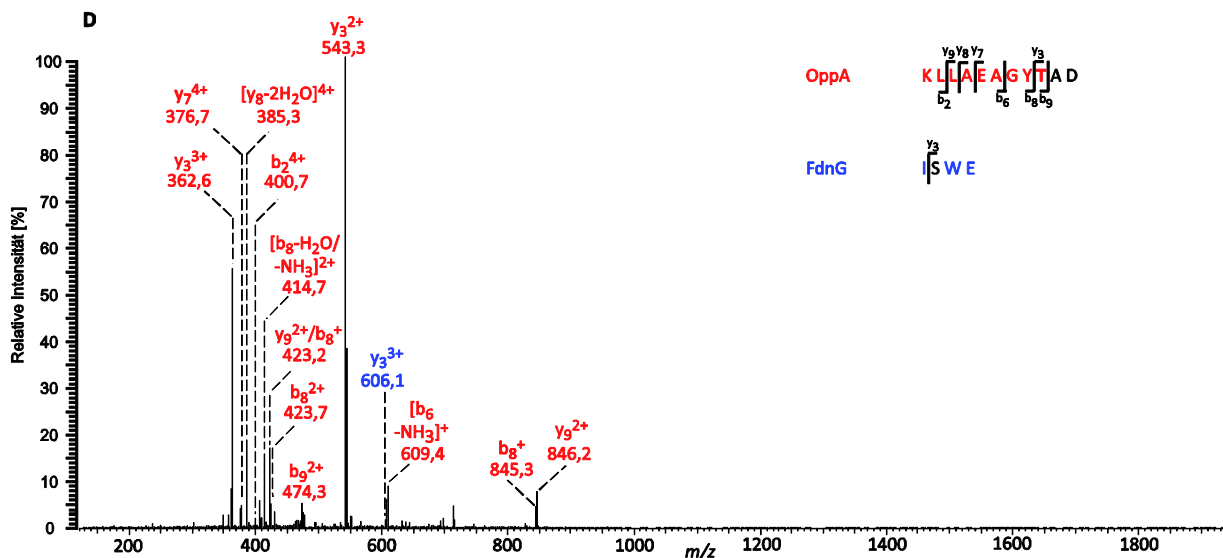
cluster_file = open(sys,argv[1],'r')

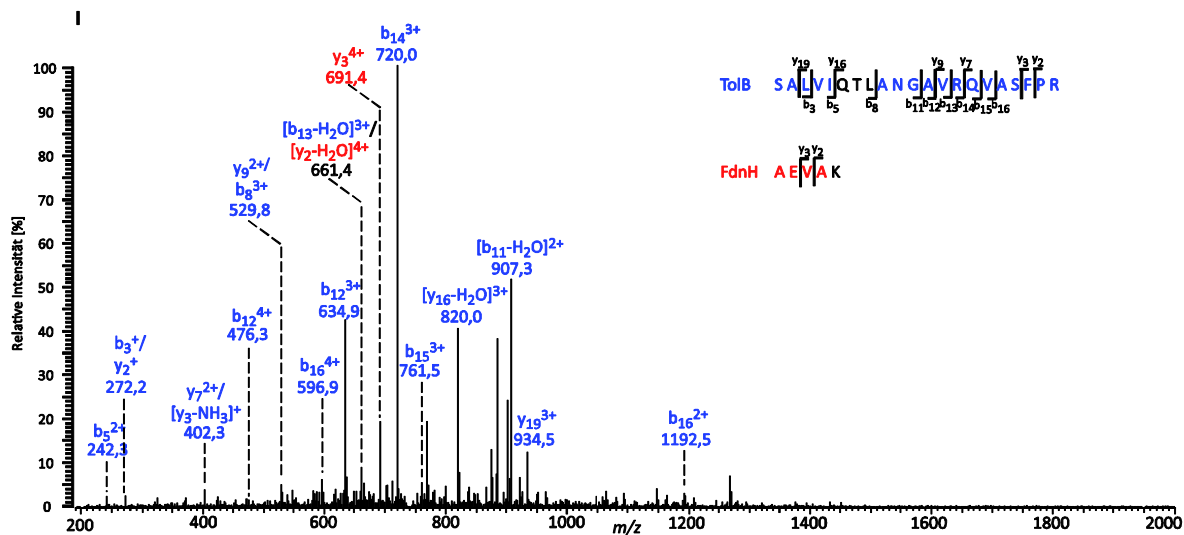
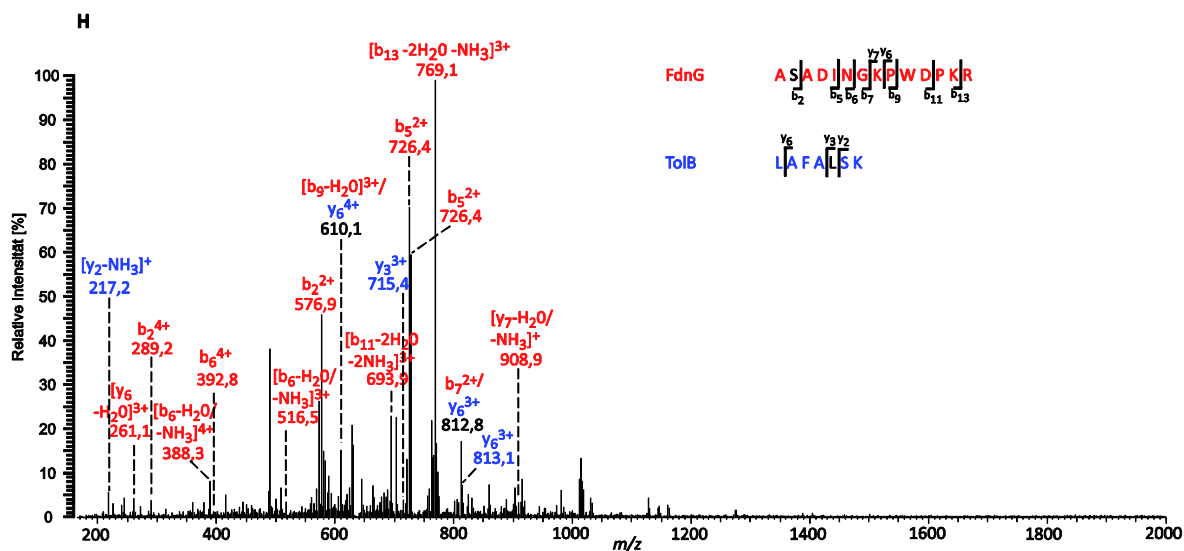
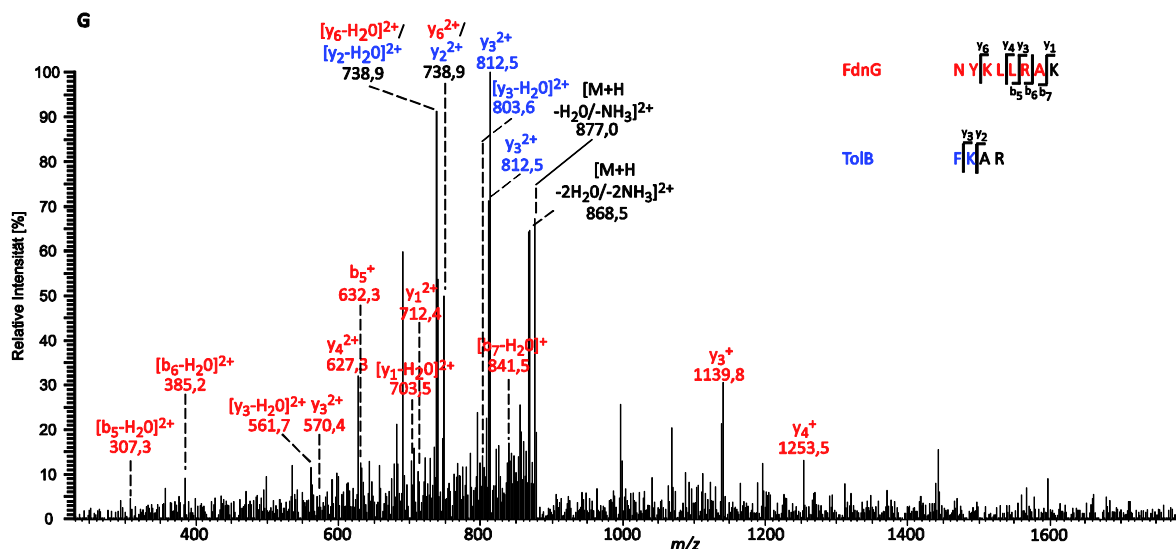
score_file = open(sys,argv[2],'r')
score_dict = {}
for line in score_file.readlines():
    line = line.strip()
    model_name = line.split()[0]
    score = line.split()[1]
    score_dict[model_name] = score

for line in cluster_file.readlines():
    line = line.strip()
    print line,score_dict[line]
```

Abbildung A 13: Cluster_energies.py.







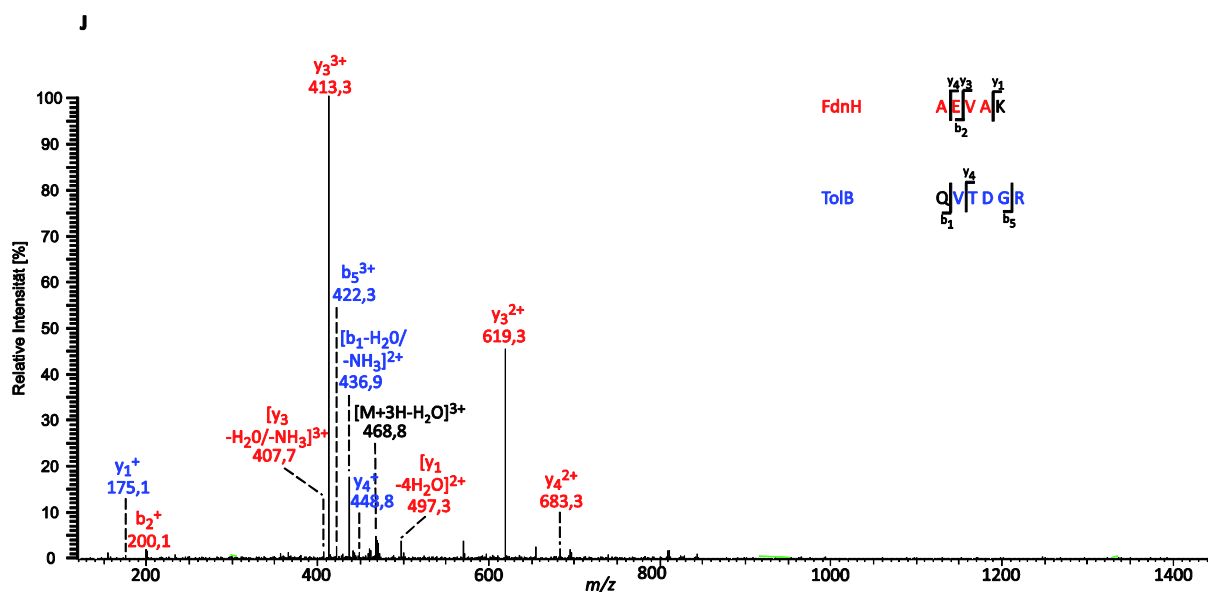


Abbildung A 14: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FdnH und Skp, FdnG und OppA, FdnG und ErfK, FdnH und TolB sowie FdnG und TolB. Vorläuferionen wurden innerhalb der linearen Ionenfalle selektiert und mittels CID fragmentiert. Peptide und Fragmentionen von FdnG und FdnH sind in rot und von Skp, OppA, TolB und ErfK in blau dargestellt. Oxidierte Methionine wurden als m und carbamidomethylierte Cysteine als B definiert. (A) Fragmentionenspektrum eines dreifach geladenen Vorläuferions (m/z 733,354). Das Cross-Linking-Produkt umfasst die Aminosäuren 105 – 110 von FdnH/FdoH und die Aminosäuren 716 – 82 von Skp, wobei K110 von FdnH/FdoH und M74 von Skp über SDAD quervernetzt sind. (B) Fragmentionenspektrum eines zweifach geladenen Vorläuferions (m/z 645,886). Das Cross-Linking-Produkt umfasst die Aminosäuren 200 – 204 von FdnH und die Aminosäuren 116 – 119 von Skp, wobei K202 von FdnH und T118 oder R119 von Skp über SDAD quervernetzt sind. (C) Fragmentionenspektrum eines vierfach geladenen Vorläuferions (m/z 652,093). Das Cross-Linking-Produkt umfasst die Aminosäuren 191 – 204 von FdnH und die Aminosäuren 54 – 59 von Skp, wobei K204 von FdnH und R59 von Skp über SDAD quervernetzt sind. (D) Fragmentionenspektrum eines vierfach geladenen Vorläuferions (m/z 483,229). Das Cross-Linking-Produkt umfasst die Aminosäuren 124 – 127 von FdnG/FdoG und die Aminosäuren 372 – 382 von OppA, wobei S125 von FdnG/FdoG und A381 oder D382 von OppA über SDAD quervernetzt sind. (E) Fragmentionenspektrum eines dreifach geladenen Vorläuferions (m/z 835,401). Das Cross-Linking-Produkt umfasst die Aminosäuren 183 – 194 von FdnG und die Aminosäuren 119 – 125 von OppA, wobei S183 von FdnG und Y121 von OppA über SDAD quervernetzt sind. (F) Fragmentionenspektrum eines vierfach geladenen Vorläuferions (m/z 567,289). Das Cross-Linking-Produkt umfasst die Aminosäuren 136 – 148 von FdnG und die Aminosäuren 233 – 236 von ErfK, wobei K138 von FdnG und V235 oder K236 von ErfK über SDAD quervernetzt sind. (G) Fragmentionenspektrum eines zweifach geladenen Vorläuferions (m/z 885,997). Das Cross-Linking-Produkt umfasst die Aminosäuren 37 – 44 von FdnG und die Aminosäuren 410 – 413 von TolB, wobei K44 von FdnG und A412 oder R413 von TolB über SDAD quervernetzt sind. (H) Fragmentionenspektrum eines vierfach geladenen Vorläuferions (m/z 637,834). Das Cross-Linking-Produkt umfasst die Aminosäuren 782 – 795 von FdnG und die Aminosäuren 258 – 264 von TolB, wobei S783 von FdnG und L262 von TolB über SDAD quervernetzt sind. (I) Fragmentionenspektrum eines vierfach geladenen Vorläuferions (m/z 740,657). Das Cross-Linking-Produkt umfasst die Aminosäuren 28 – 32 von FdnH und die Aminosäuren 224 – 244 von TolB, wobei K32 von FdnH und Q229, T230 oder L231 von TolB über SDAD quervernetzt sind. (J) Fragmentionenspektrum eines dreifach geladenen Vorläuferions (m/z 479,570). Das Cross-Linking-Produkt umfasst die Aminosäuren 28 – 32 von FdnH und die Aminosäuren 282 – 287 von TolB, wobei K32 von FdnH und Q282 von TolB über SDAD quervernetzt sind.

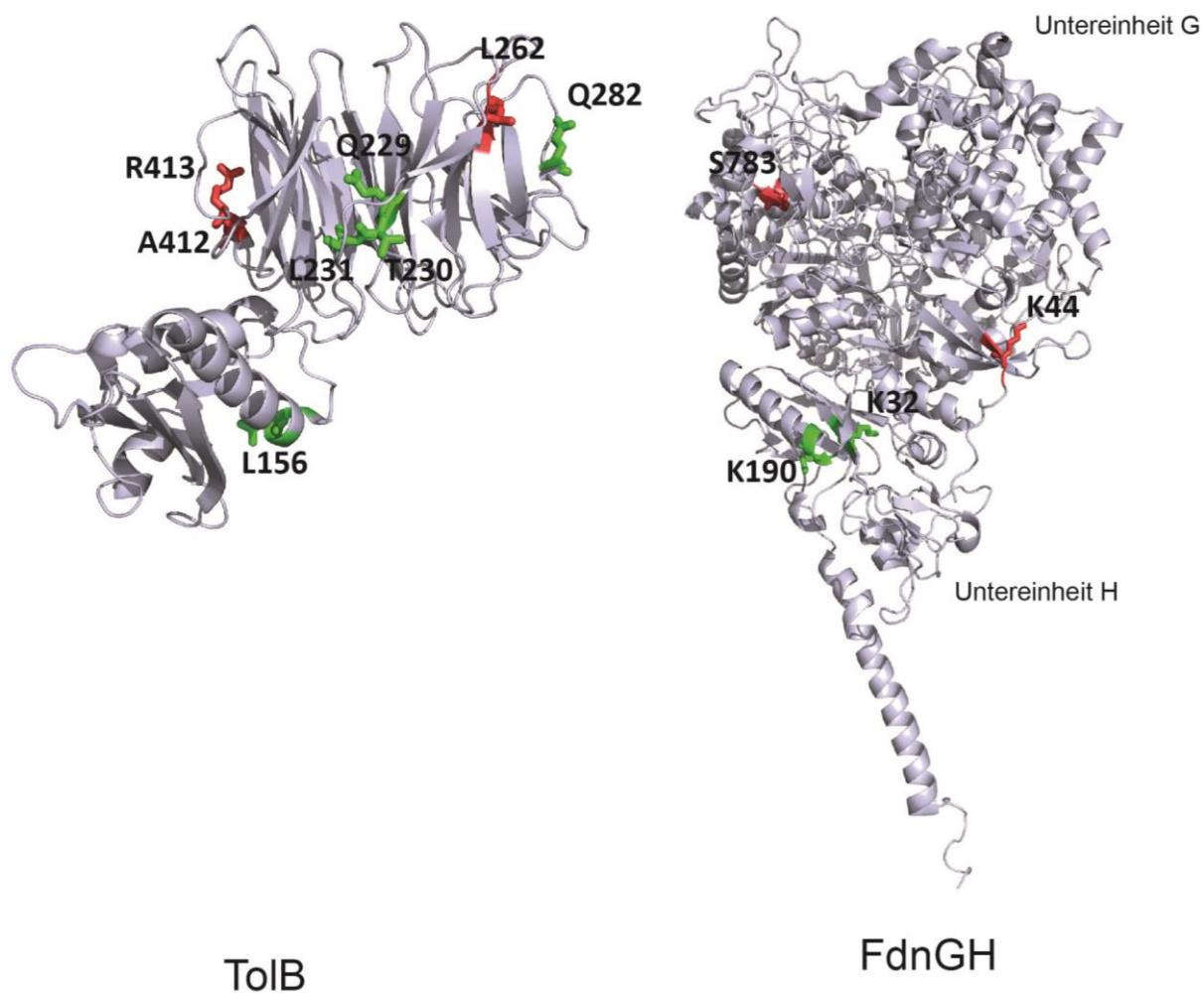


Abbildung A 15: Identifizierung intermolekularer Cross-Links zwischen FdnGH und TolB. Die Strukturen von FdnGH und TolB sind als Bänderstrukturen dargestellt (Proteindatenbankeinträge: 1KQF und 1CRZ). (A) FdnG/TolB-Cross-Linking-Produkte sind als rote *sticks*, FdnH/TolB-Cross-Linking-Produkte sind als grüne *sticks* dargestellt. Fünf Cross-Linking-Produkte zwischen FdnGH (2 × K32, K44, K190 und S783) und TolB (L156, Q229 oder T230 oder L231, L262, Q282 und A412 oder R413) wurden identifiziert.

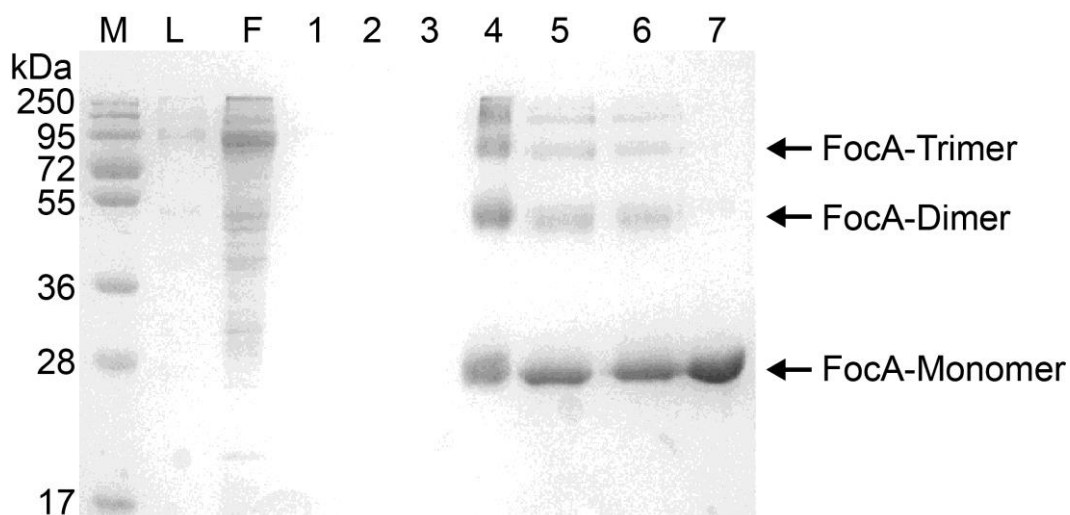


Abbildung A 16: SDS-PAGE-Analyse quervernetzter FocA/Protein-Komplexe. Die Reinigung der FocA/Protein-Komplexe erfolgte durch Streptavidin-Affinitätschromatographie. Die Auftrennung der Proteine wurde in einem 12%-igen Trenngel durchgeführt. Die Färbung der Proteine erfolgte mit kolloidalen Coomassie-Brillantblau-Lösung. Spur M: Proteinmarker (*PageRuler Prestained Protein Ladder*); Spur L: Proteinlösung vor der Reinigung (Load); Spur F: Durchlaufraction (*Flow through*); Spuren 1 – 3: Waschfraktionen; Spur 4: Elutionsfraktion von quervernetzten FocA-Komplexen, 1 : 1 (v:v) mit Laemmli-Probenpuffer verdünnt, welcher kein Reduktionsmittel enthält; Spur 5: Elutionsfraktion von quervernetzten FocA-Komplexen, 1 : 1 (v:v) mit Laemmli-Probenpuffer, welcher mit 50 mM β -Mercaptoethanol versetzt wurde; Spur 6: Elutionsfraktion von quervernetzten FocA-Komplexen, 1 : 1 (v:v) mit Laemmli-Probenpuffer, welcher mit 100 mM DTT versetzt wurde; Spur 7: FocA-Strep-tag-II-Fusionsprotein.

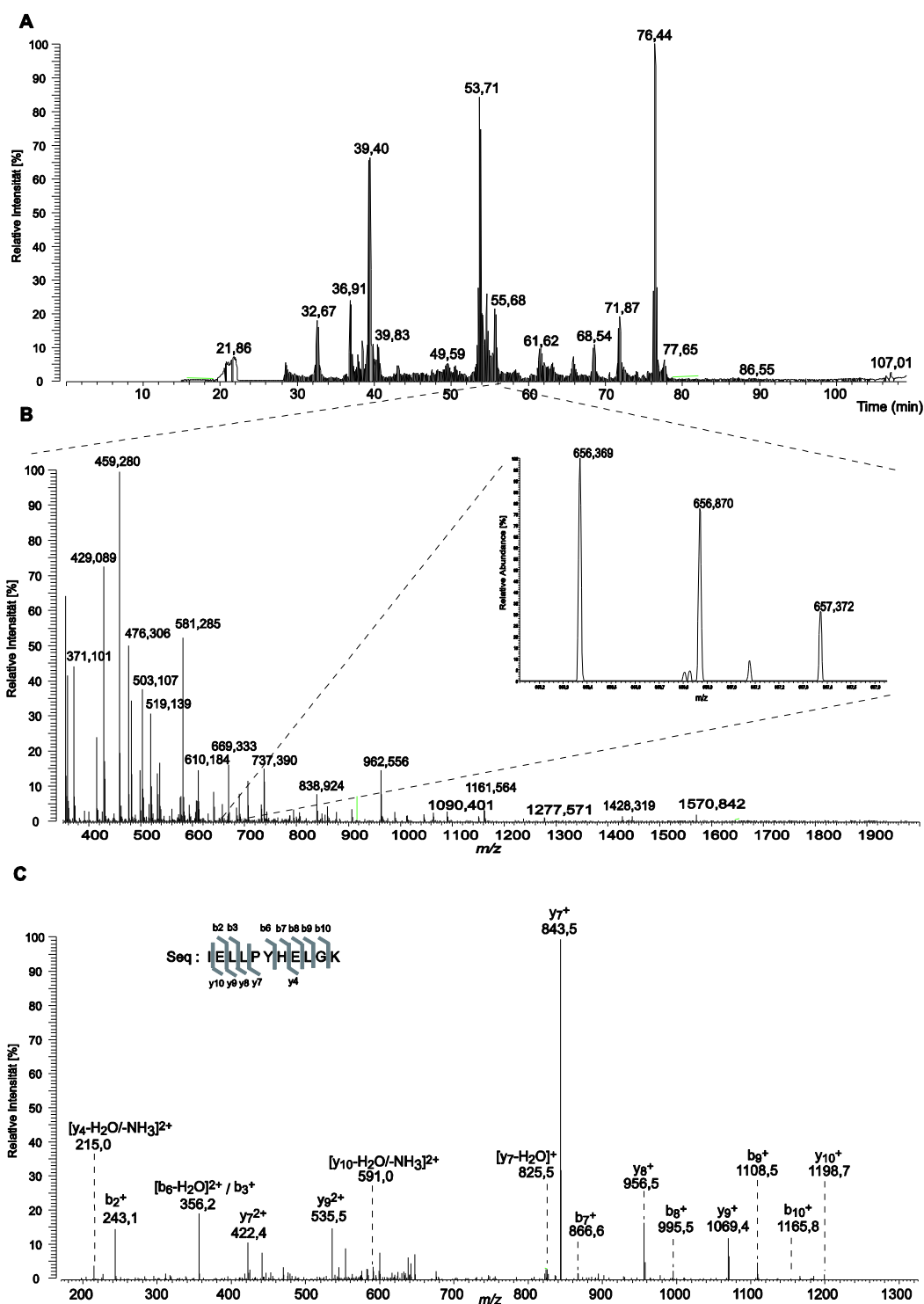
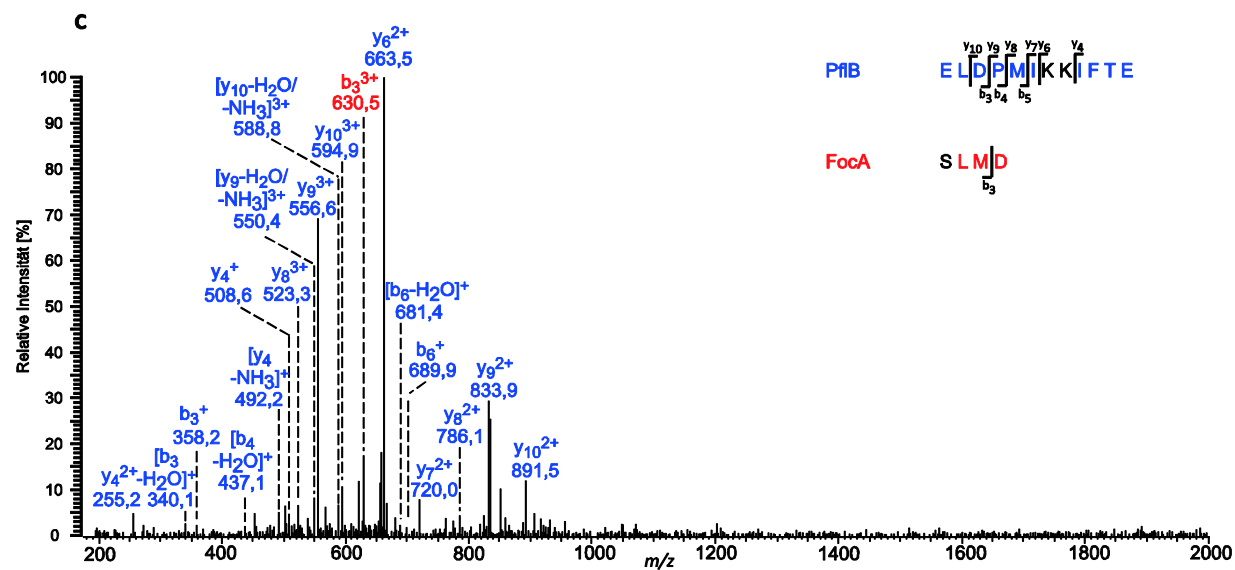
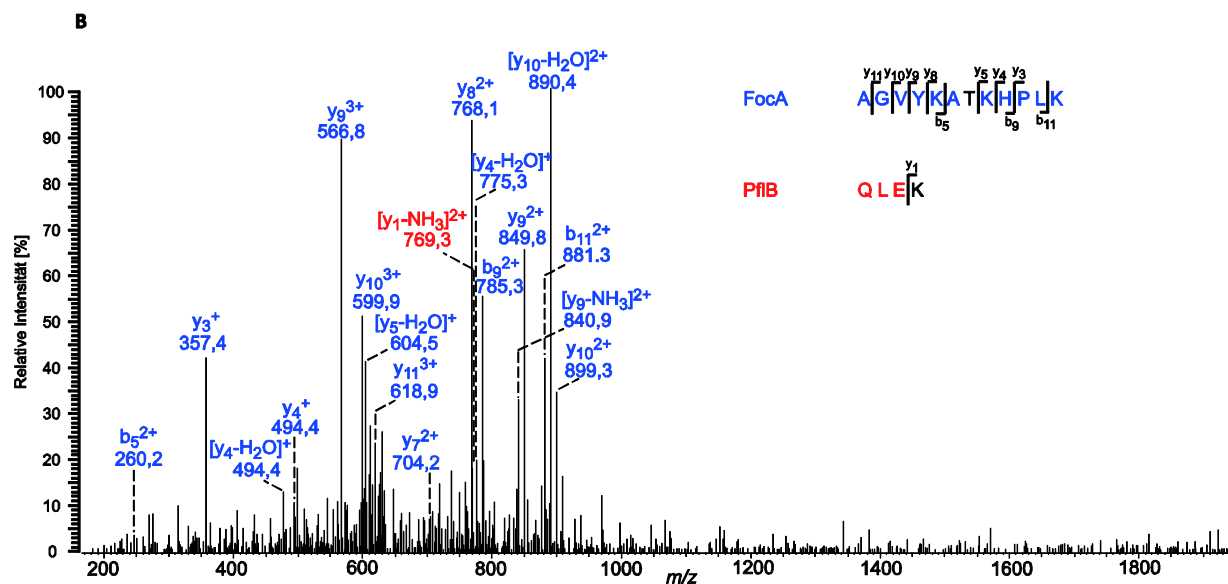
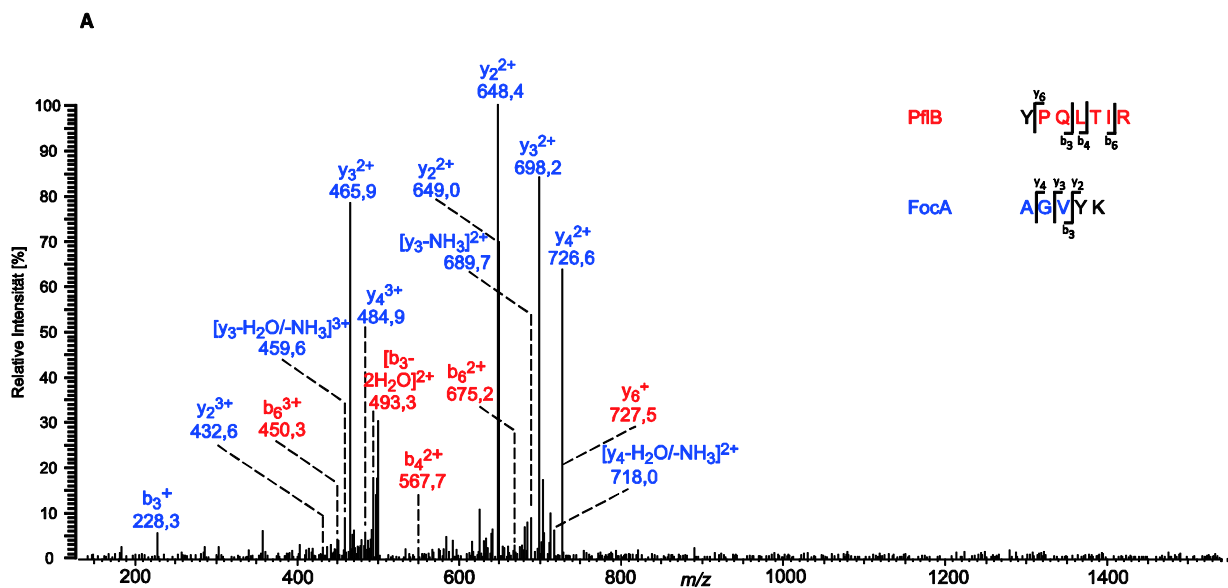
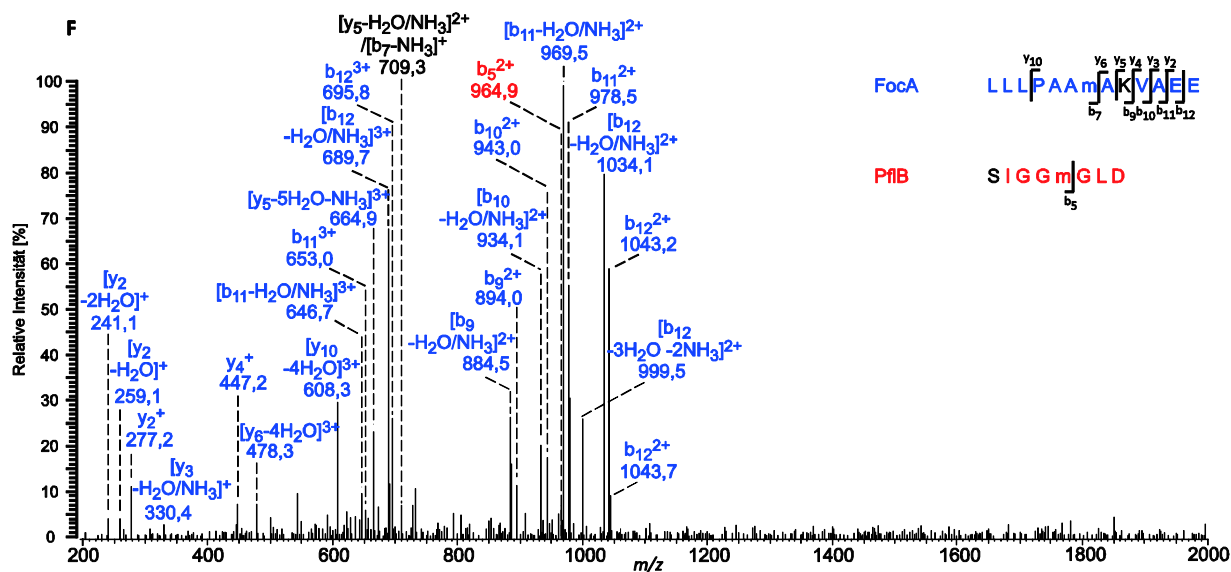
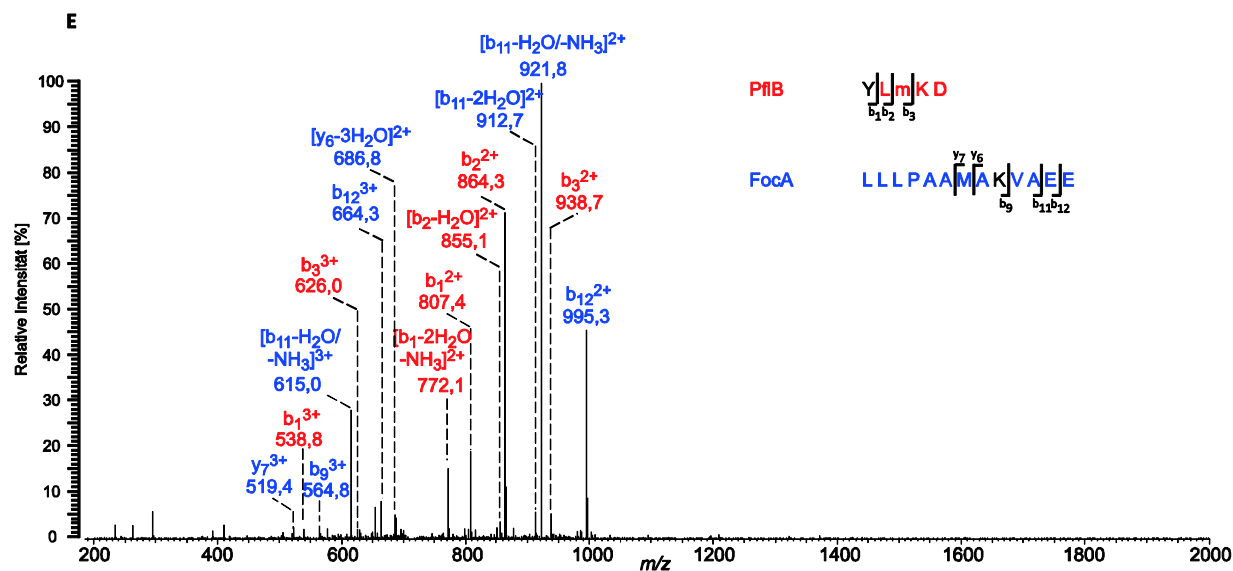
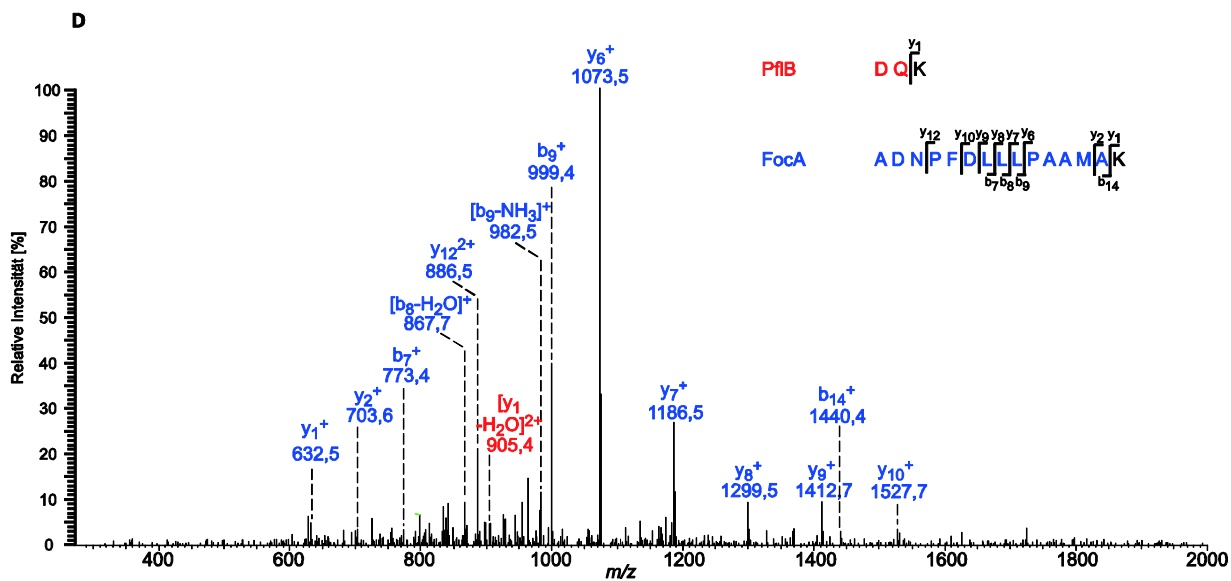
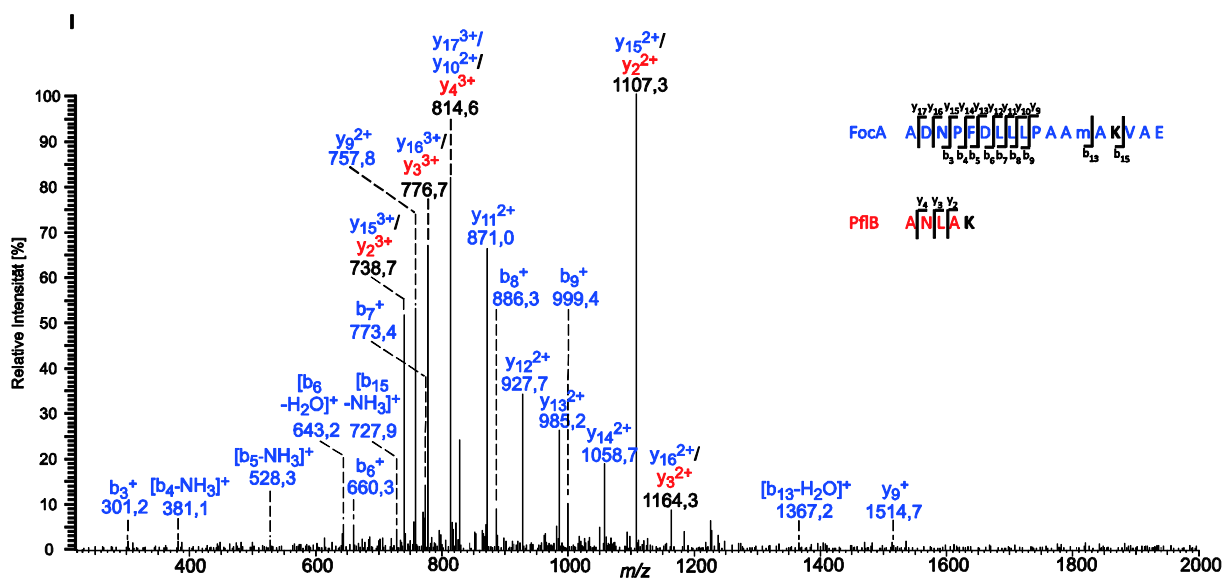
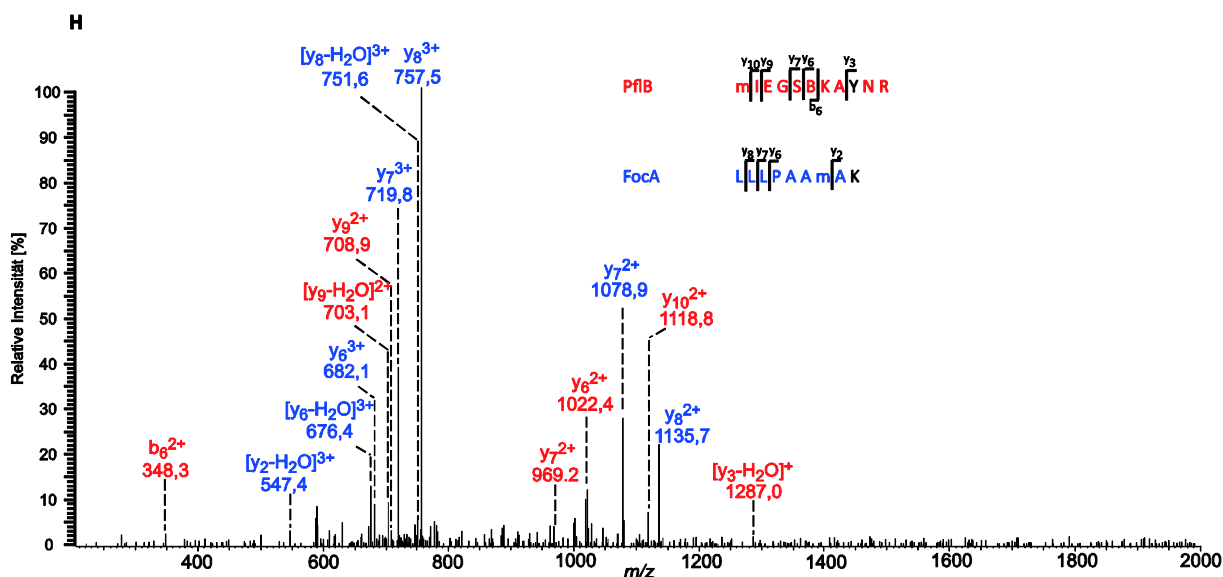
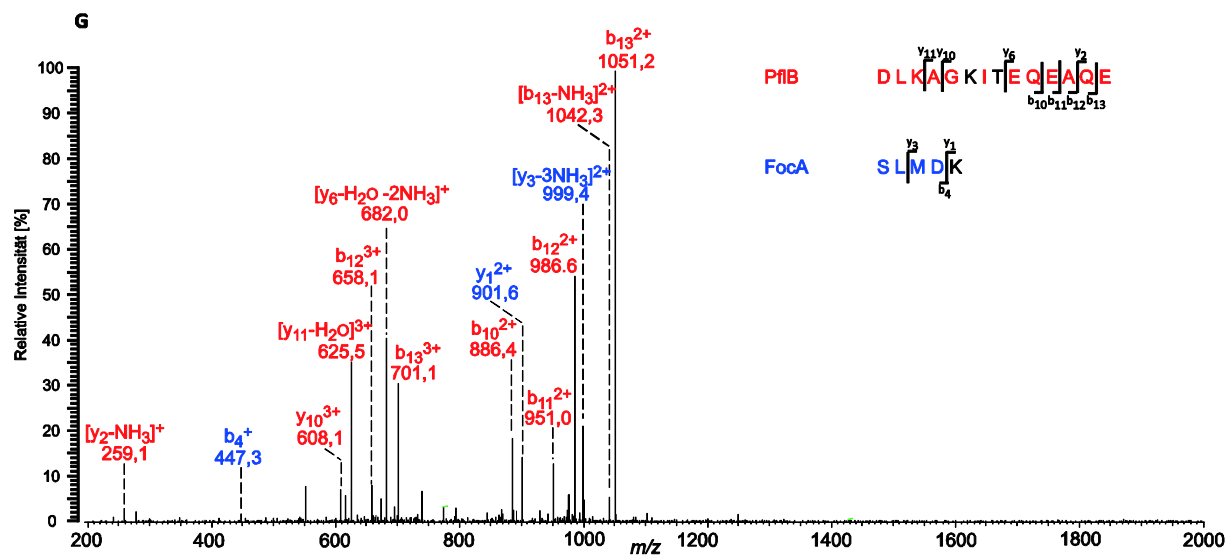


Abbildung A 17: Identifizierung eines Peptids von PflA. (A) Chromatogramm des Totalionenstroms (TIC). (B) Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS-Analyse. Das Massenspektrum wurde bei einer Retentionszeit von 58,12 min aufgenommen. Das Isotopenmuster des Peptids (IELLPYHELGK) ist vergrößert dargestellt. (C) Fragmentationenspektrum eines zweifach geladenen Vorläuferions bei m/z 656,368. Die Abweichung zwischen theoretischer und experimentell bestimmter Masse beträgt 0,39 ppm. Das Vorläuferion wurde in der linearen Ionenfalle (LTQ) selektiert und mittels CID fragmentiert.







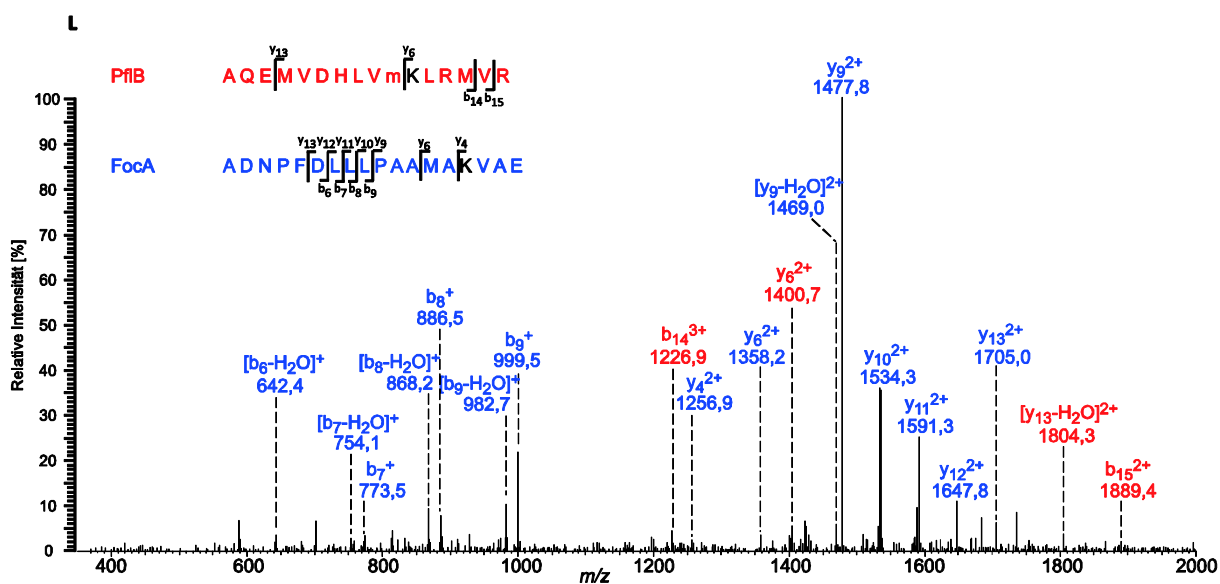
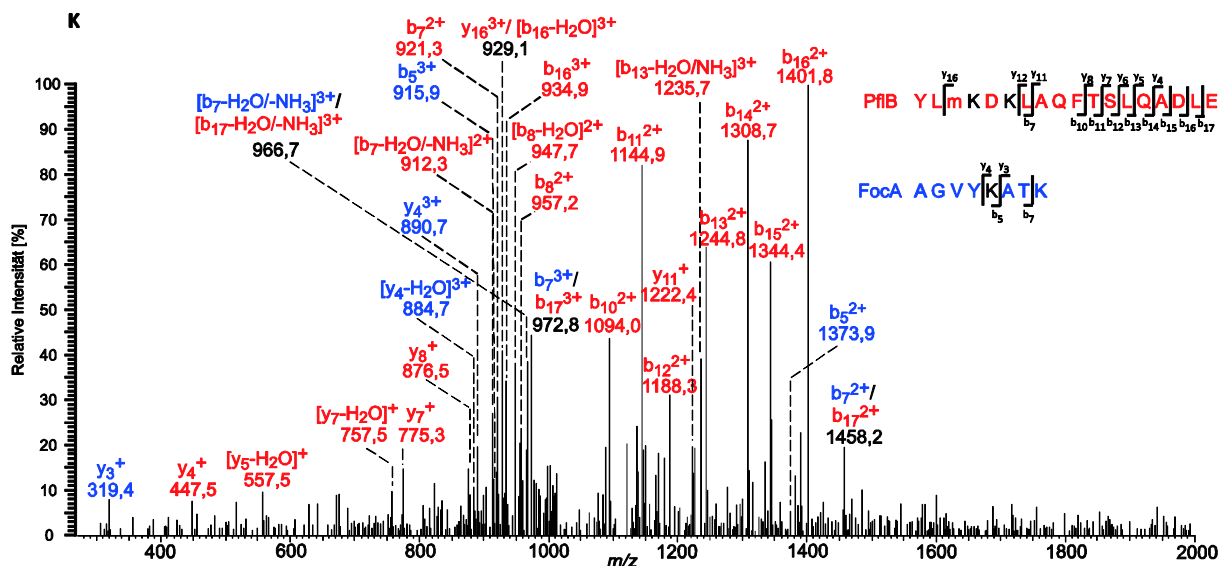
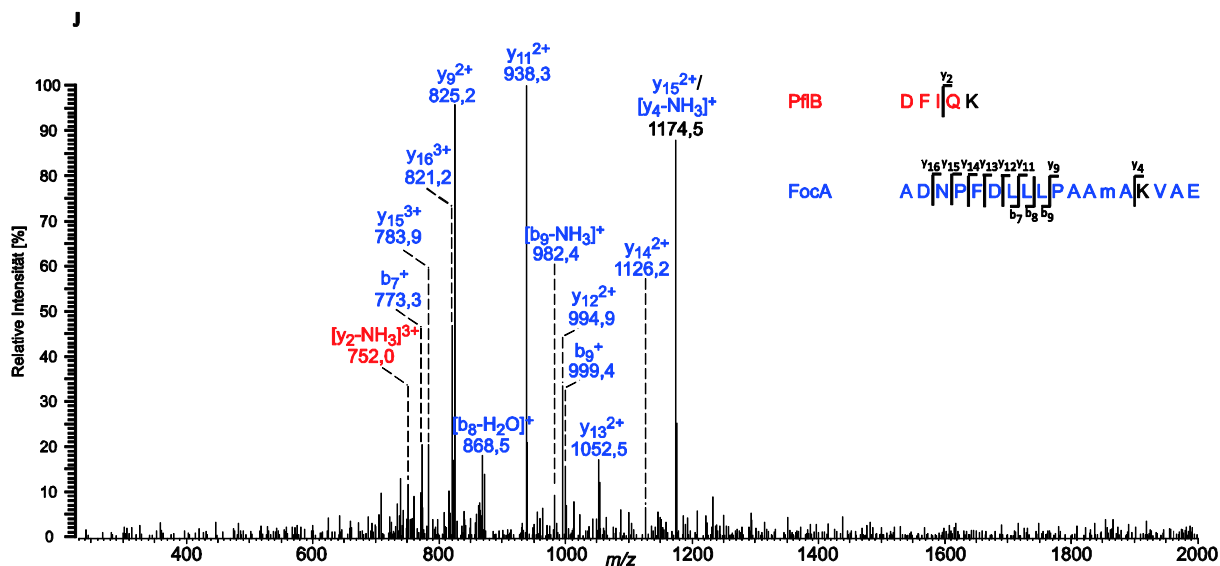


Abbildung A 18: Nano-ESI-LTQ-Orbitrap-MS/MS-Daten intermolekularer Cross-Links zwischen FocA und PflB. Vorläuferionen wurden innerhalb der linearen Ionenfalle selektiert und mittels CID fragmentiert. Peptide und Fragmentationen von FocA sind in blau und von PflB in rot dargestellt. Oxidierte Methionine wurden als m und Carbamidomethylierte Cysteine als B definiert. **(A)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 621,298). Das Cross-Linking-Produkt umfasst die Aminosäuren 22 – 26 von FocA und die Aminosäuren 726 – 732 von PflB, wobei Y25 oder K26 von FocA und Y726 von PflB über BS²G quervernetzt sind. **(B)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 642,367). Das Cross-Linking-Produkt umfasst die Aminosäuren 22 – 33 von FocA und die Aminosäuren 93 – 96 von PflB, wobei T28 von FocA und K96 von PflB über BS²G quervernetzt sind. **(C)** Fragmentationenspektrum eines dreifach geladenen m/z 675,336 Vorläuferions (m/z 675,336). Das Cross-Linking-Produkt umfasst die Aminosäuren 187 – 190 von FocA und die Aminosäuren 129 – 140 von PflB, wobei S187 von FocA und K135 oder K136 von PflB über BS²G quervernetzt sind. **(D)** Fragmentationenspektrum eines zweifach geladenen Vorläuferions (m/z 1036,523). Das Cross-Linking-Produkt umfasst die Aminosäuren 3 – 17 von FocA und die Aminosäuren 641 – 643 von PflB, wobei K17 von FocA und K643 von PflB über BS²G quervernetzt sind. **(E)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 712,704). Das Cross-Linking-Produkt umfasst die Aminosäuren 9 – 21 von FocA und die Aminosäuren 192 – 196 von PflB, wobei K17 von FocA und Y192 von PflB über BS²G quervernetzt sind. **(F)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 744,708). Das Cross-Linking-Produkt umfasst die Aminosäuren 9 – 21 von FocA und die Aminosäuren 338 – 345 von PflB, wobei K17 von FocA und S338 von PflB über BS²G quervernetzt sind. **(G)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 750,040). Das Cross-Linking-Produkt umfasst die Aminosäuren 187 – 191 von FocA und die Aminosäuren 291 – 304 von PflB, wobei K191 von FocA und K296 oder T298 von PflB über BS²G quervernetzt sind. **(H)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 795,067). Das Cross-Linking-Produkt umfasst die Aminosäuren 9 – 17 von FocA und die Aminosäuren 118 – 128 von PflB, wobei K17 von FocA und Y126 von PflB über BS²G quervernetzt sind. **(I)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 838,437). Das Cross-Linking-Produkt umfasst die Aminosäuren 3 – 20 von FocA und die Aminosäuren 437 – 441 von PflB, wobei K17 von FocA und K441 von PflB über BS²G quervernetzt sind. **(J)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 883,116). Das Cross-Linking-Produkt umfasst die Aminosäuren 3 – 20 von FocA und die Aminosäuren 28 – 32 von PflB, wobei K17 von FocA und K32 von PflB über BS²G quervernetzt sind. **(K)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 1021,527). Das Cross-Linking-Produkt umfasst die Aminosäuren 22 – 29 von FocA und die Aminosäuren 192 – 209 von PflB, wobei K26 von FocA und K195 oder K197 von PflB über BS²G quervernetzt sind. **(L)** Fragmentationenspektrum eines dreifach geladenen Vorläuferions (m/z 1318,342). Das Cross-Linking-Produkt umfasst die Aminosäuren 3 – 20 von FocA und die Aminosäuren 302 – 317 von PflB, wobei K17 von FocA und K312 von PflB über BS²G quervernetzt sind.

Tabelle A 1: Zusammenfassung identifizierter Proteine des Zellwand-Subproteoms. Die präparierten Proteine wurden durch SDS-PAGE getrennt. Die Identifizierung der Proteine in Spur L (Abbildung 20) erfolgte durch *In-Gel*-Proteinspaltung und einer anschließenden Analyse der Peptide mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS- und MS/MS. Proteine mit mindestens zwei PSMs sind aufgeführt.

UniProtKB- Eintrag	Protein	Sequenz- abdeckung [%]	Peptide spectral matches (PSMs)
P37329	Periplasmatisches Molybdat-bindendes Protein	87,16%	357
P0AFH9	Osmotisch-induzierbares Protein Y	70,15%	267
P23843	Periplasmatisches Oligopeptid-bindendes Protein OppA	74,59%	249
P24183	Formiatdehydrogenase-N, α -Untereinheit	61,48%	166
P33363	Periplasmatische beta-Glucosidase	66,93%	149
P0AFM3	Periplasmatisches Glycine-Betain-bindendes Protein	71,52%	144
P45523	FKBP-Typ cis-trans-Isomerase FkpA	51,48%	141
P23865	Schweif-spezifische Protease	47,95%	127
P0C0V1	Periplasmatische Serinendoprotease	60,34%	122
P16700	Thiosulfat-bindendes Protein	78,70%	119
P09373	Pyruvat-Formiatlyase PflB	54,61%	110
P0AEN0	Periplasmatisches Cystin-bindendes Protein	77,44%	109
P07024	Protein UshA	59,27%	107
P0ABZ8	Chaperon SurA	56,31%	106
A7ZKF2	Glucan Biosynthese Protein G	63,21%	95
P05458	Protease 3	51,04%	89
A1A7M1	Äußeres Membranprotein-Assemblierungsfaktor BamA	40,99%	86
A7ZSL4	Elongationsfaktor Tu 1	80,46%	82
P39099	Periplasmatische pH-abhängige Serinendoprotease	58,46%	82
C4ZU48	Periplasmatische Nitratreduktase	46,50%	80
A7ZJC2	Protein TolB	62,79%	80
P0AFK9	Periplasmatisches Spermidin/Putrescin Protein	56,03%	80
P0ADV7	Vermutliches Phospholipid Bindeproteine MlaC	76,78%	75
P02925	Periplasmatische D-Ribose Bindeprotein	67,57%	72
P0AGC4	Lösliche Murein-Transglykosidase	55,35%	71
P76177	Protein YdgH	75,16%	69
P76193	Vermutliche L, D-Transpeptidase	79,64%	68
P37648	Protein YhjJ	68,67%	67
P32176	Formiatdehydrogenase-O, α -Untereinheit	43,01%	66
P30859	Putativer ABC-Transporter, Arginine-bindendes Protein 2	63,37%	63
A7ZHA4	Chaperon protein DnaK	65,52%	59
C4ZVG4	Glukan Biosyntheseprotein D	55,35%	58
P0AEY0	Periplasmatisches Maltose-bindendes Protein	75,25%	57
P37902	Periplasmatisches Glutamat/Aspartat-bindendes Protein	71,85%	56
P0AFG9	Pyruvatdehydrogenase E1-Komponente	43,74%	53
P08331	2',3'-zyklische-Nucleotid 2'-Phosphodiesterase/3'- Nucleotidase	59,66%	53
B7LE10	Phosphoglyceroltransferase 1	43,91%	53

P18956	Gammaglutamyltranspeptidase	53,97%	53
P77202	Thioldisulfid-austauschendes Protein DsbG	74,60%	52
P69909	Glutamatdecarboxylase, α -Untereinheit	51,29%	51
P69911	Glutamatdecarboxylase, β -Untereinheit	51,29%	51
P0ADS7	Nicht-Charakterisiertes Protein YggE	48,78%	51
P0A9B4	Glyceraldehyde-3-Phosphatdehydrogenase A	62,54%	50
P06959	Dihydrolipoyllysineacetyltransferase	59,05%	50
P63285	Chaperon ClpB	41,31%	48
A7ZR34	Phosphoglyceratkinase	66,93%	48
P0AEG7	Thiol:Disulfid-austauschendes Protein	58,90%	48
P31550	Periplasmatisches Thiamin-bindendes Protein	55,96%	46
P28629	Arginindecarboxylase	34,57%	45
P0AEQ5	Periplasmatisches Glutamin-bindendes Protein	70,56%	45
P78067	Thiosulfat-Schwefeltransferase	56,32%	44
P00811	Beta-Lactamase	63,93%	44
P26648	Zellteilungsprotein FtsP	64,26%	41
P18775	Dimethylsulfoxidreductase DmsA	32,43%	40
P75797	Glutathione-bindendes Protein	41,02%	40
P0AAY0	Vermutliche L,D-Transpeptidase	45,10%	39
P0AG69	30S ribosomales Protein S1	51,89%	38
A1AJ51	60 kDa Chaperon	52,92%	38
P77348	Periplasmisches Murein-bindendes Protein	51,40%	38
P66948	TPR wiederholdendes-Peptid-Protein	49,90%	38
P33590	Periplasmatisches Nickel-bindendes Protein	55,92%	37
P0A911	Äußeres Membranprotein A	51,73%	37
P00805	L-Asparaginase 2	62,07%	36
P19926	Glucose-1-Phosphatase	61,74%	35
A7ZKW9	Periplasmische Trehalase	44,78%	35
P76116	Nicht-charakterisiertes Protein YncE	54,67%	35
P0AEG5	Thiol:Disulfid-austauschendes Protein DsbA	80,77%	35
P0ADT1	Nicht-charakterisiertes Protein YggN	52,72%	34
P0A9Q8	Alkoholdehydrogenase	35,02%	33
A7ZSL5	Elongationsfaktor G	47,44%	32
P39371	N-Acetylneuraminatpimerase	62,23%	32
P0AET3	Säurestress-Chaperon HdeB,	48,15%	31
P07102	Periplasmatisches Protein AppA	45,83%	31
P0AG82	Phosphat-bindendes Protein PstS	72,54%	31
P76223	Protein YnjB	59,79%	31
A7ZS61	Polyribonucleotidtransferase	33,47%	30
A7ZHR0	Elongationsfaktor Ts	72,79%	29
P09394	Glycerolphosphoryl-diester-Phosphodiesterase	72,07%	29
P36649	Blaue Kupferoxidase CueO	39,73%	28
B7MD95	Trigger-Faktor	41,96%	28
P0AGD2	Superoxiddismutase	80,35%	28
B7NAT2	Protein Ycel	69,63%	28

P0AB24	Eisenaufnahmesystem, Komponente EfeO	61,33%	27
A7ZSJ7	50S ribosomales Protein L5	73,18%	27
A7ZJW0	Lipoprotein-Trägerprotein der äußeren Membran	67,49%	27
P0ACE1	Hydrogenase-2, große Untereinheit	38,62%	26
P39176	Vermutliche L,D-Transpeptidase ErfK	50,65%	26
P0AFY0	Sigmafaktor-E-regulierendes Protein RseB	54,40%	26
P06996	Äußeres Membranprotein C	50,14%	25
P00350	6-Phosphogluconatdehydrogenase	47,22%	25
P0ABK6	Cysteinesynthase A	51,70%	25
P39172	Zinkaufnahmesystem-zugehöriges Protein ZnuA	56,77%	25
P33913	Nicht-charakterisiertes Protein YejA	39,40%	24
Q47622	Periplasmatisches Peptidtransportprotein SapA	42,41%	24
Q46863	Putatives-bindendes Protein YgiS	45,61%	24
P0AEU2	Periplasmatisches Histidin-bindendes Protein	61,92%	24
A7ZUJ7	50S ribosomales Protein L1	57,26%	24
P45955	Nicht-charakterisiertes Protein YbgF	35,36%	24
A7ZSI5	30S ribosomales Protein S4	50,97%	23
P22525	Vermutliche L,D-Transpeptidase YcbB	32,85%	22
P0AD62	Pyruvatkinase 1	45,11%	21
B7UIL4	30S ribosomales Protein S2	48,13%	21
A7ZJM7	DNA-Schutzprotein	71,26%	21
P0AFL5	Peptidyl-prolyl cis-trans-Isomerase A	65,26%	21
P0AG81	Periplasmatisches sn-Glycerol-3-Phosphat-bindendes Protein UgpB	45,21%	20
P0A871	Transaldolase B	61,83%	20
P33362	Vermutliches Substrat-bindendes Protein OsmF	44,59%	20
P0A8N6	Lysin-tRNA-Ligase	37,03%	19
P0A9P2	Dihydrolipoyldehydrogenase	39,45%	19
A7ZQM2	Enolase	34,03%	19
A7ZSK6	50S ribosomales Protein L2	57,51%	19
P30860	ABC-Transporter, Arginin-bindendes Protein 1	58,85%	19
P26365	N-acetylmuramoyl-L-Alanin-Amidase AmiB	40,45%	18
A7ZSI4	DNA-RNA-Polymerase, α -Untereinheit	41,95%	18
P0AB72	Fructose-bisphosphat-Aldolase Klasse 2	37,33%	18
P31133	Periplasmatisches Putrescin-bindendes Protein	41,35%	18
P21338	Ribonuklease 1	53,36%	18
P27302	Transketolase 1	22,93%	17
P69778	Äußeres Membranlipoprotein	48,72%	17
P0AG78	Sulfat-bindendes Protein	45,90%	17
P0A869	Transaldolase A	51,58%	17
P08660	Lysin-sensitive Aspartokinase 3	19,38%	16
P0AD71	D-alanyl-D-alanin-Carboxypeptidase/Endopeptidase AmpH	38,96%	16
P37661	Phosphoethanolamintransferase EptB	22,38%	16
P0A9M8	Phosphatacetyltransferase	17,51%	15

P08839	Phosphoenolpyruvat-Phosphotransferase	22,78%	15
P0A3B2	GTP-Bindeprotein TypA/BipA	25,70%	15
A7ZPE4	Penicillin-insensitive Mureinendopeptidase	33,21%	15
A7ZSK3	30S ribosomales Protein S3	53,65%	15
P39187	Nicht-charakterisiertes Protein YtfJ	48,91%	15
P07395	Phenylalanine-tRNA-Ligase, β -Untereinheit	21,64%	14
P17169	Glutamin-Fruktose-6-Phosphate-Aminotransferase	27,26%	14
A1A8D9	Chaperon HtpG	24,04%	14
P0ADV3	Lipopolysaccharid-Exportsystem Protein LptA	40,00%	14
P0AE22	Säurephosphatase (Klasse B)	33,76%	14
A1A8D9	Chaperon HtpG	24,04%	14
P0AE22	Säurephosphatase (Klasse B)	33,76%	14
P0ADA2	Acyl-CoA Thioesterase 1	50,00%	14
P32160	Nicht-charakterisiertes Protein YiiQ	63,32%	14
P37325	Nicht-charakterisiertes Protein YbcH	42,23%	13
P0ADU6	Protein YgiW	26,15%	13
A7ZU36	5-Methyltetrahydropteroyltriglutamat-Homocystein- Methyltransferase	10,23%	12
P0AFF8	Transkriptionselongationsprotein NusA	27,07%	12
P23847	Periplasmisches Dipeptid-Transportprotein	24,49%	12
P09551	Periplasmatisches Lysin-Arginin-Ornithin-bindendes Protein	46,15%	12
P37028	Vitamin B12-bindendes Protein	47,37%	12
P08201	Nitritreduktase [NAD(P)H] große Untereinheit	14,29%	11
C4ZQF0	Aspartat-tRNA-Ligase	19,66%	11
P0A6A5	Acetatkinase	36,25%	11
A7ZPZ4	Serin-Hydroxymethyltransferase	23,50%	11
A7ZPX4	Cysteindesulfurase	29,70%	11
A7ZR64	S-Adenosylmethioninsynthase	31,77%	11
A7ZV47	Adenylosuccinatesynthetase	28,24%	11
P0AEE6	Periplasmatisches D-Galactose-bindendes Protein	31,63%	11
A7ZTH2	ADP-L-Glycerol-D-Manno-Heptose-6-Epimerase	40,00%	11
P0A797	6-Phosphofruktokinase Isozyme 1	29,38%	11
P75804	Lösliche Aldosedehydrogenase YliL	29,38%	11
A7ZP31	Ecotin	27,16%	11
P03841	Periplasmatisches Maltoseprotein	33,01%	11
P0AE10	Alkylhydroperoxidreductase γ -Untereinheit	54,55%	11
A1AEN7	Alanin-tRNA-Ligase	13,70%	10
A7ZS65	Translation Initiierungsfaktor IF-2	11,12%	10
P0A8N4	Lysine-tRNA-Ligase	19,80%	10
P08200	Isocitratdehydrogenase [NADP]	26,20%	10
P77619	Protein YfeW	25,35%	10
A7ZJD0	2,3-bisPhosphoglycerate-abhängige Phosphoglycerat- mutase	39,20%	10
P0AEU9	Chaperon Skp	27,33%	10

P0AGD5	Superoxiddismutase	40,93%	10
P0ACP6	HTH-tTyp Transcriptionalregulator GntR	2,42%	9
A7ZHB5	Isoleucine-tRNA-Ligase	10,34%	9
C4ZXE9	Glycin-tRNA-Ligase, β -Untereinheit	16,98%	9
B7LAX0	Glucose-6-Phosphat Isomerase	20,04%	9
P15288	Cytosol-unspezifische Dipeptidase	22,27%	9
P02930	Äußeres Membranprotein TolC	21,10%	9
P39265	Periplasmatisches D-Allose-bindendes Protein	41,16%	9
P0AEL7	Periplasmatisches Enterobactin-bindendes Protein	22,33%	9
P39325	Periplasmatisches ABC-Transporter-bindendes Protein	25,79%	9
A7ZKA9	Flavoprotein WrbA	44,95%	9
P36683	Aconitathydratase 2	9,94%	8
A7ZUH4	Phosphoenolpyruvatcarboxylase	8,49%	8
P13656	Vermutliche bifunktionales Chitinase/Lysozym	9,92%	8
P46923	Trimethylamin-N-oxid-Reduktase 2	11,12%	8
A1AIC1	Katalase-Peroxidase	12,12%	8
C4ZRT7	Prolin-tRNA-Ligase	18,88%	8
P0A954	3-Oxoacyl-[acyl-carrier-protein] Synthase 1	22,17%	8
P0A992	Fruktose-bisphosphat-Aldolase Klasse 1	24,86%	8
P76422	Hydroxymethylpyrimidin-/Phosphomethylpyrimidin-Kinase	13,16%	8
P38683	Periplasmatisches Protein TorT	21,05%	8
P76108	Vermutliches ABC-Transporter-bindendes Protein YdcS	23,10%	8
A7ZIN4	Adenylatkinase, GN= <i>adk</i>	35,05%	8
A7ZUD3	Triosephosphatisomerase	32,16%	8
A7ZV87	Eisen-Schwefel-Zentrum Reperaturprotein YtfE	26,36%	8
P0A7B0	Anorganische Pyrophosphatase	33,52%	8
P0A9M1	Lonprotease	7,78%	7
P25718	Alpha-Amylase	10,80%	7
B7NPH0	Hydroxylaminreductase	18,18%	7
P63884	N-acetylmuramoyl-L-Alanin-Amidase	21,34%	7
A7ZPW6	Peptidase B	18,03%	7
P00634	Alkalische Phosphatase	17,83%	7
A1ADS2	Glutamat-tRNA-Ligase	16,99%	7
P76176	Nicht-charakterisierte Serin Protease YdgD	20,88%	7
A7ZK57	3-Hydroxydecanoyldehydratase	34,30%	7
P0A919	Äußeres Membranprotein X	46,78%	7
C4ZUJ6	30S ribosomales Protein S7	35,20%	7
P33343	Nicht-charakterisiertes Protein YehD	42,78%	7
P21179	Katalase HP11	10,49%	6
P33570	Transketolase 2	10,49%	6
A7ZTU6	ATP-Synthase, α -Untereinheit	13,26%	6
P0AC54	Glucose-6-Phosphatdehydrogenase	12,42%	6
A7ZTU4	ATP Synthase, β -Untereinheit	16,96%	6
A7ZMI2	Phenylalanin-tRNA-Ligase, α -Untereinheit	17,74%	6
P0AAJ3	Formiatdehydrogenase, β -Untereinheit	15,99%	6

A7ZN92	Sulfoxidreduktase, katalytische Untereinheit	22,16%	6
A7ZSJ4	50S ribosomales Protein L6	29,94%	6
P64598	Nicht-charakterisiertes Protein YraP	40,31%	6
P25516	Aconitathydratase 1	6,51%	5
P04825	Aminopeptidase N	7,47%	5
C4ZWC9	Leucin-tRNA-Ligase	4,19%	5
P07118	Valin-tRNA-Ligase	5,36%	5
A7ZQC8	Protein RecA	13,60%	5
A7ZJW2	Serin-tRNA-Ligase	9,30%	5
A1AJV0	Phosphopentomutase	16,46%	5
P0AET0	Säurestress-Chaperone HdeA	25,45%	5
P24232	Flavoheмоprotein	10,86%	5
A7ZSR4	3-Dehydroquinatsynthase	14,64%	5
P52643	D-Lactatdehydrogenase	14,59%	5
P04816	spezifisches Leucin-bindendes Protein	11,92%	5
P0ADA3	Mureinhydrolase-Aktivator	12,14%	5
P0A957	KHG/KDPG-Aldolase	19,72%	5
P0AEK2	3-oxoacyl-[acyl-carrier-protein]-Reductase	29,92%	5
B7USA1	Translation-Initiationsfaktor IF-3	32,78%	5
Q0TBG1	Guanylatkinase	21,26%	5
P33342	Nicht-charakterisiertes Chaperon YehC	28,03%	5
P69784	Spezifische Glucose-Phosphotransferase, Enzym IIA	34,91%	5
P0A864	Thiolperoxidase	48,21%	5
A7ZNT3	Methionine-tRNA-Ligase	7,53%	4
P35340	Alkylhydroperoxidreduktase, Untereinheit F	9,40%	4
A1AHE9	2,3-bisPhosphoglycerate-unabhängige Phosphoglycerat- mutase	9,73%	4
P11349	Respiratorische Nitratereductase 1, β -Untereinheit	10,35%	4
A7ZJA8	Succinyl-CoALigase, β -Untereinheit	10,82%	4
P37659	Protein BcsG	11,81%	4
A7ZJE2	6-Phosphogluconolactonase	16,01%	4
P0A9A8	Zellteilungsprotein FtsZ	12,79%	4
P0AD98	Leu/Ile/Val-bindendes Protein	12,53%	4
P77258	N-Ethylmaleimidreduktase	11,78%	4
A7ZJZ6	Phosphoserinaminotransferase	16,85%	4
P00954	Tryptophan-tRNA-Ligase	12,28%	4
A7ZVS7	Purinnukleosidphosphorylase	13,81%	4
P15028	Periplasmatisches Fe(3+)-Zitrat-bindendes Protein	8,00%	4
P28635	D-Methionine-bindendes Lipoprotein	19,56%	4
P12758	Uridinphosphorylase	18,97%	4
P0ACK0	cAMP-Rezeptorprotein	20,00%	4
A7ZV18	Elongationfaktor P	14,89%	4
P25736	Endonuklease-1	14,89%	4
A7ZHR2	Ribosom-Recyclingfaktor	25,41%	4
P76344	Metal-bindendes Protein ZinT	15,28%	4

P0A6W7	Transkriptionselongationsfaktor GreA	26,58%	4
P37194	Äußeres Membranprotein Slp	20,21%	4
P0AB26	Nicht-charakterisiertes Lipoprotein YceB	34,41%	4
P0ACY2	Vermutliche NAD(P)H-Nitroreduktase YdjA	25,68%	4
P00562	Bifunktionale Aspartkinase/Homoserindehydrogenase 2	3,95%	3
P0AES0	Bifunktionale Glutathionyl/Spermidin-Synthetase/Amidase	6,14%	3
P00579	Sigmafaktor RpoD	6,04%	3
A1AJP2	Ribosomale RNA-Methyltransferase, kleine Untereinheit	9,33%	3
A7ZJ63	Glutamine-tRNA-Ligase	5,60%	3
A7ZMI6	Threonine-tRNA-Ligase	4,05%	3
A1AIE8	Vitamin B12 Transporter	6,84%	3
P76128	Vermutliche periplasmatisches D, D-diPeptide-bindendes Protein	6,59%	3
P25526	Succinatsemialdehyddehydrogenase	6,64%	3
A7ZUU5	Cytochrome C-552	5,86%	3
A7ZQM3	CTP-Synthase	6,97%	3
P0AG32	Transcriptionterminationsfaktor Rho	6,21%	3
P77318	Nicht-charakterisierte Sulfatase YdeN	7,50%	3
A7ZIJ6	ATP-abhängige Clp-Protease	8,02%	3
P0AEP5	UTP-Glucose-1-phosphate-Uridyltransferase	11,26%	3
P02931	Äußeres Membranprotein F	9,94%	3
P28304	Quinonoxidoreduktase 1	9,17%	3
A7ZTA6	Glycine-tRNA-Ligase, α -Untereinheit	14,19%	3
P75785	Putative Phosphoethanolamintransferase	7,02%	3
P75954	Vermutliche L, D-Transpeptidase	5,63%	3
P27250	Nicht-charakterisiertes Alkoholdehydrogenase-ähnliches Protein (Zinktyp)	10,62%	3
Q46858	Nicht-charakterisiertes-Protein YqhG	9,42%	3
P0A9Q3	Aerobe-Atmung-Kontrollprotein ArcA	14,71%	3
A1AJU8	Deoxyribosephosphataldolase	14,29%	3
P0ABU1	1,4-Dihydroxy-2-naphthoyl-CoA-Synthase	8,77%	3
P0ACA5	Stringentes Hungerprotein A	18,40%	3
P0AF57	Nicht-charakterisiertes Protein YjcO	13,10%	3
P25397	Telluritmethyltransferase	14,21%	3
A7ZPU1	Urazilphosphoribosyltransferase	13,94%	3
P77717	Nicht-charakterisiertes Lipoprotein YbaY	22,11%	3
A7ZTU8	ATP-Synthase, β -Untereinheit	19,23%	3
P0AF00	Molybdenumkofaktor Biosyntheseprotein B	19,41%	3
A7ZMI4	50S ribosomales Protein L20	22,88%	3
P0ADA6	Nicht-charakterisiertes Lipoprotein YajG	17,71%	3
A7ZII0	Nicht-charakterisiertes Protein YajQ	21,47%	3
A7ZQ24	Autonomer Glycylradikalkofaktor	14,17%	2
P31828	Vermutliche Zinkprotease	2,47%	2
A7ZHI9	Proteintranslokase SecA	1,89%	2
P0ADY2	Peptidylprolyl-cis-trans-Isomerase D	1,28%	2

A7ZHC0	4-Hydroxy-Tetrahydrodipicolinatreductase	7,33%	2
P0A9C7	Glutaminsynthetase	3,84%	2
P06988	Histidinoldehydrogenase	3,00%	2
P0ADG8	Inosine-5'-monophosphatdehydrogenase	6,56%	2
P21599	Pyruvatkinase 2	7,08%	2
A7ZLS1	NAD-abhängiges Äpfelsäure-Enzym	3,19%	2
A7ZK21	Asparagin-tRNA-Ligase	3,22%	2
P33015	Nicht-charakterisiertes inneres Membranprotein YeeE	1,70%	2
P15034	Xaa-Pro-Aminopeptidase	4,54%	2
P0AFK1	Protein PmbA	6,00%	2
P37759	dTDP-glucose 4,6-dehydratase 1	3,60%	2
P00509	Aspartataminotransferase	9,09%	2
A7ZHA5	Chaperone DnaJ	5,32%	2
P37747	UDP-Galactopyranosemutase	4,90%	2
A7ZSD0	Malatdehydrogenase	3,53%	2
P0A9K5	PhoH-ähnliches Protein	7,23%	2
Q46856	Alkoholdehydrogenase YqhD	6,98%	2
P33013	D-Alanyl-D-Alanincarboxypeptidase	6,44%	2
Q46857	2,5-diketo-D-Buttersäurereductase A	8,00%	2
A7ZQ54	Protein GrpE	10,71%	2
P40710	Lipoprotein NlpE	9,32%	2
P23836	Transcriptionales Regulatorprotein PhoP	12,56%	2
P0A9L1	FKBP-Typ peptidyl-prolyl-cis-trans-Isomerase SlyD	10,20%	2
Q0TGH3	Vermutliches transkriptionales Regulatorprotein YeeN	7,98%	2
P02924	Periplasmatische L-Arabinose-bindendes Protein	8,51%	2
P26646	Vermutliche Acrylyl-CoA-Reductase	7,10%	2
P0AET9	7-alpha-hydroxysteroiddehydrogenase	7,45%	2
P58345	Elektronentransportkomplex Protein RnfG	11,17%	2
P0ABS3	RNA-Polymerase-bindender Transkriptionsfaktor	13,91%	2
P18776	Anaerobe Dimethylsulfoxidreduktase, β -Untereinheit	10,73%	2
P0A9A0	Ferritin-1	10,91%	2
A7ZJ30	LPS-assemblierendes Lipoprotein LptE	9,84%	2
A7ZQC0	S-Ribosylhomocysteinylase	15,79%	2
P23869	Peptidyl-prolyl-cis-trans-Isomerase B	10,98%	2
A7ZKL2	Protein YcfP	11,11%	2
P76296	Nicht-charakterisiertes Protein YecT	11,11%	2
A7ZVB3	Protein YjgA	9,84%	2
P64583	Nicht-charakterisiertes Protein YqjD	25,74%	2

Tabelle A 2: Identifizierte Bindungspartner der Fdhs. Die Elutionsfraktionen der Größenausschlusschromatographie (Superdex S200 16/300 GL-Säule) innerhalb des Retentionsvolumens von 7 – 13 ml (Abbildung 22) wurden vereint und in Lösung gespalten. Die Peptidgemische wurden mittels Nano-HPLC/Nano-ESI-LTQ-Orbitrap-MS und -MS/MS analysiert. Die UniProtKB-Einträge identifizierter Proteine sind angegeben.

UniprotKB- Eintrag	Protein	PSMs	
		Cross-Linking-Ansatz	Kontrolle
P0AFG9	Pyruvatdehydrogenase E1-Komponente	190	216
A7ZUJ8	50S ribosomales Protein L10	182	184
P24183	Formatdehydrogenase-N, α -Untereinheit	158	160
P07395	Phenylalanin-tRNA-Ligase, β -Untereinheit	85	79
P32176	Formiatdehydrogenase-O, α -Untereinheit	77	77
P0C0S2	Gering-leitender mechanosensitiver Kanal	75	23
A7ZUK0	50S ribosomales Protein L7/L12	69	66
P0A9M8	Phosphatacetyltransferase	54	57
A7ZTU4	ATP-Synthase, β -Untereinheit	53	58
A7ZMI2	Phenylalanin-tRNA-Ligase, α -Untereinheit	50	29
P09152	Respiratorische Nitratereduktase 1, α -Untereinheit	45	10
A7ZHA4	Chaperon DnaK	35	32
P0AAJ3	Formiatdehydrogenase-N, β -Untereinheit	34	34
P23843	Periplasmisches Oligopeptid-bindendes Protein	31	19
P0AEU9	Chaperon Skp	31	0
P0A719	Ribosephosphat-Pyrophosphatkinase	30	23
P63388	Vermutliches ATP-abhängiges Phospholipid-Importprotein	28	24
P0AG69	30S ribosomales Protein S1	28	47
P39176	Vermutliche L, D-Transpeptidase ErfK	27	14
A1AJ51	60 kDa Chaperon 1	26	23
A7ZJM7	DNA-Schutzprotein	24	14
P0A9Q8	Alkoholdehydrogenase	23	28
P0AAJ6	Formiatdehydrogenase-O, β -Untereinheit	22	15

Tabelle A 3: Zusammenfassung der potentiellen FocA-Bindungspartner. Dargestellt sind identifizierte Proteine, für welche bis zu 7 *peptide spectral matches* (PSMs) detektiert wurden.

UniProtKB-Eintrag	Protein	MW [kDa]	Peptide	PSMs	Mascot score
P0AC25	Vermutlicher Formiatkanal 1	31,0	5	273	1271
P09373	Pyruvat-Formiatlyase 1	85,3	40	117	1680
A7ZSL4	Elongationsfaktor 1	43,3	17	59	800
P0A9B4	Glycerinaldehyd-3-Phosphatdehydrogenase A	35,5	14	59	417
P0A9Q8	Alkoholdehydrogenase	96,1	28	56	1571
P06959	Dihydrolipoyllysinacetyltransferase	66,1	21	51	647
A7ZHA4	Chaperon DnaK	69,1	27	47	251
Q8FB83	DNA-RNA-Polymerase, β -Untereinheit	155,1	33	42	1097
A7ZUK2	DNA-RNA Polymerase, β -Untereinheit	155,1	33	40	1097
A7ZSK6	50S ribosomales Protein L2	29,8	11	35	300
P0A911	Äußeres Membranprotein A	37,2	9	34	151
P69911	Glutamatdecarboxylase, β -Untereinheit	52,6	10	32	450
A7ZUK1	DNA-RNA-Polymerase, β -Untereinheit	150,5	28	31	913
P69909	Glutamatdecarboxylase, α -Untereinheit	52,7	10	30	428
Q46829	6-phospho-beta-Glucosidase BglA	55,3	11	28	138
A7ZSI5	30S ribosomales Protein S4	23,5	6	28	140
B1IU2	Alanyl-tRNA-Synthetase	96,0	18	28	761
P62577	Chloramphenicolacetyltransferase	25,6	4	26	254
P24182	Biotincarboxylase	49,3	9	25	326
A7ZQM2	Enolase	45,6	12	25	322
A1AJ51	60 kDa Chaperon 1	57,3	16	23	115
P0A9P2	Dihydrolipoyldehydrogenase	50,7	13	23	266
P0A9H4	Induzierbare Lysindecaboxylase	81,2	14	23	412
P08839	Phosphoenolpyruvatphosphotransferase	63,5	13	22	471
P00393	NADH-Dehydrogenase	47,3	8	21	465
A7ZR64	S-Adenosylmethioninsynthase	41,9	12	21	449
A7ZSJ4	50S ribosomales Protein L6	18,9	6	21	195
A7ZTR3	Tryptophanase	52,7	10	21	220
P28629	Arginindecaboxylase	84,4	14	20	467
P0A797	6-phosphofruktokinase Isozyme 1	34,8	7	20	203

A7ZSI4	DNA-RNA-Polymerase, α -Untereinheit	36,5	7	18	87
B7UIL4	30S ribosomales Protein S2	26,7	6	18	177
A7ZSI3	50S ribosomales Protein L17	14,4	4	17	62
A7ZSK3	30S ribosomales Protein S3	26,0	7	16	80
P00350	6-phosphogluconatdehydrogenase	51,4	7	14	87
B1XBE0	Chaperon DnaJ	41,1	8	14	118
A7ZSL5	Elongationsfaktor G	77,5	11	14	337
P0A954	3-Oxoacylsynthase 1	42,6	5	14	206
P0A9M8	Phosphatacetyltransferase	77,1	8	13	239
A7ZV47	Adenylosuccinatsynthetase	47,3	7	13	256
A7ZQM3	CTP-Synthase	60,3	8	13	227
A7ZUJ8	50S ribosomales Protein L10	17,7	4	12	81
A7ZV73	30S ribosomales Protein S18	9,0	3	12	83
P0AE10	Alkylhydroperoxidreductase, γ -Untereinheit	20,7	3	11	84
P0AES4	DNA-Gyrase, α -Untereinheit	96,9	4	11	103
A7ZR34	Phosphoglyceratkinase	41,1	5	11	190
P00562	Aspartokinase/Homoserine-Dehydrogenase 2	88,8	7	10	392
P0A9H8	Cyclopropan Phospholipidsynthase	43,9	6	10	116
P0ACG0	DNA-bindendes Protein H-NS	15,5	4	10	102
P07003	Pyruvatdehydrogenase [Ubiquinon]	62,0	7	10	172
A7ZSK8	50S ribosomales Protein L4	22,1	5	10	52
A7ZSJ7	50S ribosomales Protein L5	20,3	3	10	134
P0A6A5	Acetatkinase	43,3	4	9	125
A7ZHC0	Dihydrodipicolinatreductase	28,7	3	9	106
P21599	Pyruvatkinase 2	51,3	8	9	44
A7ZUK0	50S ribosomales Protein L7/L12	12,3	3	9	40
A7ZSI6	30S ribosomales Protein S11	13,8	3	9	139
A7ZNR4	D-Tagatose-1,6-bisphosphat-Aldolase	47,1	4	8	112
Q8X8V9	D-Methionin-bindendes Lipoprotein	29,4	6	8	44
P0A7T1	30S ribosomales Protein S13	13,1	4	8	76
P07395	Phenylalanyl-tRNA-Synthetase, β -Untereinheit	87,3	6	8	197
B7L7S9	Arginyl-tRNA-Synthetase	64,7	6	8	138
P00954	Tryptophanyl-tRNA-Synthetase	37,4	4	8	74
P0AB72	Fruktose-bisphosphate-Aldolase Klasse 2	39,1	4	7	38

A7ZUH4	Phosphoenolpyruvatcarboxylase	99,0	7	7	223
A1AJV0	Phosphopentomutase	44,4	4	7	136
A7ZVS7	Purinnucleosidphosphorylase	25,9	5	7	116
P14407	Fumarathydratase Klasse 1	60,1	6	7	51
P0ACD8	Hydrogenase 1, große Untereinheit	66,2	6	7	222
A7ZK14	Chromosom partitionierendes Protein	170,1	5	7	116
P0A9N6	Pyruvat-Formatelyase 1-aktivierendes Enzym	28,2	4	7	101
A7ZHI9	Protein Translokase, SecA	102,0	7	7	211
A7ZTA5	Glycyl-tRNA-Synthetase, β -Untereinheit	76,7	6	7	224

Literaturverzeichnis

1. Knappe, J., and Sawers, G. (1990) A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. *FEMS Microbiol. Rev.* 6, 383-398
2. Wagner, A. F., Frey, M., Neugebauer, F. A., Schafer, W., and Knappe, J. (1992) The free radical in pyruvate formate-lyase is located on glycine-734. *Proc. Natl. Acad. Sci. USA* 89, 996-1000
3. Knappe, J., Neugebauer, F. A., Blaschkowski, H. P., and Ganzler, M. (1984) Post-translational activation introduces a free radical into pyruvate formate-lyase. *Proc. Natl. Acad. Sci. USA* 81, 1332-1335
4. Sawers, G., and Suppmann, B. (1992) Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. *J. Bacteriol.* 174, 3474-3478
5. Sawers, G., and Watson, G. (1998) A glycy radical solution: oxygen-dependent interconversion of pyruvate formate-lyase. *Mol. Microbiol.* 29, 945-954
6. Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S., and Volker Wagner, A. F. (1999) Structure and mechanism of the glycy radical enzyme pyruvate formate-lyase. *Nat. Struct. Biol.* 6, 969-975
7. Sawers, R. G., and Clark, D. P. (2004) Fermentative pyruvate and acetyl CoA metabolism. July 2004, posting date. Chapter 3.5.3. In R. Curtiss III (Editor in Chief), *EcoSal--Escherichia coli and Salmonella: Cellular and Molecular Biology*. [Online.] <http://www.ecosal.org>. ASM Press, Washington, D.C
8. Becker, A., and Kabsch, W. (2002) X-ray structure of pyruvate formate-lyase in complex with pyruvate and CoA. How the enzyme uses the Cys-418 thiyl radical for pyruvate cleavage. *J. Biol. Chem.* 277, 40036-40042

9. Alexeeva, S., de Kort, B., Sawers, G., Hellingwerf, K. J., and de Mattos, M. J. (2000) Effects of limited aeration and of the ArcAB system on intermediary pyruvate catabolism in *Escherichia coli*. *J. Bacteriol.* 182, 4934-4940
10. Sawers, R. G. (2005) Formate and its role in hydrogen production in *Escherichia coli*. *Biochem. Soc. Trans.* 33, 42-46
11. Sawers, G., Blokesch, M., and Böck, A. (2004) Anaerobic formate and hydrogen metabolism. September 2004, posting date. Chapter 3.5.4. In R. Curtiss III (Editor in Chief), *EcoSal-Escherichia coli and Salmonella: Cellular and Molecular Biology*. [Online.] <http://www.ecosal.org>. ASM Press, Washington, D.C
12. Suppmann, B., and Sawers, G. (1994) Isolation and characterization of hypophosphite-resistant mutants of *Escherichia coli*: identification of the FocA protein, encoded by the *pfl* operon, as a putative formate transporter. *Mol. Microbiol.* 11, 965-982
13. Belaich, A., and Belaich, J. P. (1976) Microcalorimetric study of the anaerobic growth of *Escherichia coli*: growth thermograms in a synthetic medium. *J. Bacteriol.* 125, 14-18
14. Stephenson, M., and Stickland, L. H. (1932) Hydrogenlyases: Bacterial enzymes liberating molecular hydrogen. *Biochem. J.* 26, 712-724
15. Sawers, G., and Bock, A. (1989) Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. *J. Bacteriol.* 171, 2485-2498
16. Sawers, R. G. (2006) Differential turnover of the multiple processed transcripts of the *Escherichia coli* *focA-pflB* operon. *Microbiology* 152, 2197-2205
17. Sawers, G., and Bock, A. (1988) Anaerobic regulation of pyruvate formate-lyase from *Escherichia coli* K-12. *J. Bacteriol.* 170, 5330-5336
18. Sawers, G. (1993) Specific transcriptional requirements for positive regulation of the anaerobically inducible *pfl* operon by ArcA and FNR. *Mol. Microbiol.* 10, 737-747

19. Drapal, N., and Sawers, G. (1995) Promoter 7 of the *Escherichia coli pfl* operon is a major determinant in the anaerobic regulation of expression by ArcA. *J. Bacteriol.* 177, 5338-5341
20. Drapal, N., and Sawers, G. (1995) Purification of ArcA and analysis of its specific interaction with the *pfl* promoter-regulatory region. *Mol. Microbiol.* 16, 597-607
21. Kaiser, M., and Sawers, G. (1995) Fnr activates transcription from the P6 promoter of the *pfl* operon *in vitro*. *Mol. Microbiol.* 18, 331-342
22. Reyes-Ramirez, F., and Sawers, R. G. (2006) Aerobic activation of transcription of the anaerobically inducible *Escherichia coli focA-pfl* operon by fumarate nitrate regulator. *FEMS Microbiol. Lett.* 255, 262-267
23. Rodel, W., Plaga, W., Frank, R., and Knappe, J. (1988) Primary structures of *Escherichia coli* pyruvate formate-lyase and pyruvate formate-lyase-activating enzyme deduced from the DNA nucleotide sequences. *Eur. J. Biochem. / FEBS* 177, 153-158
24. Sauter, M., and Sawers, R. G. (1990) Transcriptional analysis of the gene encoding pyruvate formate-lyase-activating enzyme of *Escherichia coli*. *Mol. Microbiol.* 4, 355-363
25. Wang, Y., Huang, Y., Wang, J., Cheng, C., Huang, W., Lu, P., Xu, Y. N., Wang, P., Yan, N., and Shi, Y. (2009) Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel. *Nature* 462, 467-472
26. Waight, A. B., Love, J., and Wang, D. N. (2010) Structure and mechanism of a pentameric formate channel. *Nat. Struct. Mol. Biol.* 17, 31-37
27. Lu, W., Du, J., Wacker, T., Gerbig-Smentek, E., Andrade, S. L., and Einsle, O. (2011) pH-dependent gating in a FocA formate channel. *Science* 332, 352-354
28. Czyzewski, B. K., and Wang, D. N. (2012) Identification and characterization of a bacterial hydrosulphide ion channel. *Nature* 483, 494-497

-
29. Waight, A. B., Czyzewski, B. K., and Wang, D. N. (2013) Ion selectivity and gating mechanisms of FNT channels. *Curr. Opin. Struct. Biol.* 23, 499-506
 30. Unkles, S. E., Symington, V. F., Kotur, Z., Wang, Y., Siddiqi, M. Y., Kinghorn, J. R., and Glass, A. D. (2010) Physiological and biochemical characterization of AnNitA, the *Aspergillus nidulans* high affinity nitrite transporter. *Eukaryot. cell* 10 1724-1732
 31. Lu, W., Schwarzer, N. J., Du, J., Gerbig-Smentek, E., Andrade, S. L., and Einsle, O. (2012) Structural and functional characterization of the nitrite channel NirC from *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* 109, 18395-18400
 32. Sawers, G. (1994) The hydrogenases and formate dehydrogenases of *Escherichia coli*. *Antonie van Leeuwenhoek* 66, 57-88
 33. Birkmann, A., and Bock, A. (1989) Characterization of a cis regulatory DNA element necessary for formate induction of the formate dehydrogenase gene (*fdhF*) of *Escherichia coli*. *Mol. Microbiol.* 3, 187-195
 34. Hasona, A., Self, W. T., Ray, R. M., and Shanmugam, K. T. (1998) Molybdate-dependent transcription of *hyc* and *nar* operons of *Escherichia coli* requires MoeA protein and ModE molybdate. *FEMS Microbiol. Lett.* 169, 111-116
 35. Rossmann, R., Sawers, G., and Bock, A. (1991) Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. *Mol. Microbiol.* 5, 2807-2814
 36. Enoch, H. G., and Lester, R. L. (1975) Purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. *J. Biol. Chem.* 250, 6693-6705
 37. Jormakka, M., Tornroth, S., Byrne, B., and Iwata, S. (2002) Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. *Science* 295, 1863-1868
 38. Berg, B. L., Li, J., Heider, J., and Stewart, V. (1991) Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. I. Nucleotide sequence of the *fdnGHl*

operon and evidence that opal (UGA) encodes selenocysteine. *J. Biol. Chem.* 266, 22380-22385

39. Sawers, G., Heider, J., Zehelein, E., and Bock, A. (1991) Expression and operon structure of the *sel* genes of *Escherichia coli* and identification of a third selenium-containing formate dehydrogenase isoenzyme. *J. Bacteriol.* 173, 4983-4993

40. Pinsent, J. (1954) The need for selenite and molybdate in the formation of formic dehydrogenase by members of the *coli-aerogenes* group of bacteria. *Biochem. J.* 57, 10-16

41. Abaibou, H., Pommier, J., Benoit, S., Giordano, G., and Mandrandberthelot, M. A. (1995) Expression and characterization of the *Escherichia coli fdo* locus and a possible physiological role for aerobic formate dehydrogenase. *J. Bacteriol.* 177, 7141-7149

42. Soboh, B., Pinske, C., Kuhns, M., Waclawek, M., Ihling, C., Trchounian, K., Trchounian, A., Sinz, A., and Sawers, G. (2011) The respiratory molybdo-selenoprotein formate dehydrogenases of *Escherichia coli* have hydrogen:benzyl viologen oxidoreductase activity. *BMC Microbiol.* 11

43. Hopper, S., and Bock, A. (1995) Effector-mediated stimulation of ATPase activity by the sigma 54-dependent transcriptional activator FHLA from *Escherichia coli*. *J. Bacteriol.* 177, 2798-2803

44. Leonhartsberger, S., Korsas, I., and Bock, A. (2002) The molecular biology of formate metabolism in enterobacteria. *J. Mol. Microbiol. Biotechnol.* 4, 269-276

45. Sousa, P. M., Videira, M. A., and Melo, A. M. (2013) The formate:oxygen oxidoreductase supercomplex of *Escherichia coli* aerobic respiratory chain. *FEBS Lett.* 2559-64

46. Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L. F. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial *tat* pathway. *J. Biol. Chem.* 274, 13223-13228

-
47. Sargent, F. (2007) The twin-arginine transport system: moving folded proteins across membranes. *Biochem. Soc. Trans.* 35, 835-847
48. Stanley, N. R., Sargent, F., Buchanan, G., Shi, J. R., Stewart, V., Palmer, T., and Berks, B. C. (2002) Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol. Microbiol.* 43, 1005-1021
49. Berks, B. C. (1996) A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* 22, 393-404
50. Cristobal, S., de Gier, J. W., Nielsen, H., and von Heijne, G. (1999) Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J.* 18, 2982-2990
51. Bertero, M. G., Rothery, R. A., Palak, M., Hou, C., Lim, D., Blasco, F., Weiner, J. H., and Strynadka, N. C. J. (2003) Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A. *Nat. Struct. Biol.* 10, 681-687
52. Mitchell, P. (1979) Keilin's respiratory chain concept and its chemiosmotic consequences. *Science* 206, 1148-1159
53. Richardson, D., and Sawers, G. (2002) Structural biology. PMF through the redox loop. *Science* 295, 1842-1843
54. Jormakka, M., Richardson, D., Byrne, B., and Iwata, S. (2004) Architecture of NarGH reveals a structural classification of Mo-bisMGD enzymes. *Structure* 12, 95-104
55. Uden, G., and Bongaerts, J. (1997) Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochim. Biophys. Acta* 1320, 217-234
56. Schwarz, K., Bieri, J. G., Briggs, G. M., and Scott, M. L. (1957) Prevention of exudative diathesis in chicks by factor 3 and selenium. *Proc. Soc. Exp. Biol. Med.* 95, 621-625
57. Cone, J. E., Del Rio, R. M., Davis, J. N., and Stadtman, T. C. (1976) Chemical characterization of the selenoprotein component of clostridial glycine reductase:

identification of selenocysteine as the organoselenium moiety. *Proc. Natl. Acad. Sci. USA* 73, 2659-2663

58. Flohé, L. (1989) The selenoprotein glutathione peroxidase. In: *Glutathione: Chemical, Biochemical and Medical Aspects, Part A.* (D. Dolphin, R. Poulson and O. Avramovic, eds.) pp. 643-731 John Wiley & Sons, New York

59. Arkowitz, R. A., and Abeles, R. H. (1989) Identification of acetyl phosphate as the product of clostridial glycine reductase: Evidence for an acyl enzyme intermediate. *Biochemistry* 28, 4639-4644

60. Axley, M. J., Bock, A., and Stadtman, T. C. (1991) Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proc. Natl. Acad. Sci. USA* 88, 8450-8454

61. Zinoni, F., Birkmann, A., Leinfelder, W., and Bock, A. (1987) Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. *Proc. Natl. Acad. Sci. USA* 84, 3156-3160

62. Zinoni, F., Birkmann, A., Stadtman, T. C., and Bock, A. (1986) Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 83, 4650-4654

63. Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., and Harrison, P. R. (1986) The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. *EMBO J.* 5, 1221-1227

64. Yoshizawa, S., and Bock, A. (2009) The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* 1790, 1404-1414

65. Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot, M. A., and Bock, A. (1988) *Escherichia coli* genes whose products are involved in selenium metabolism. *J. Bacteriol.* 170, 540-546

-
66. Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M. A., and Bock, A. (1988) Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* 331, 723-725
67. Forchhammer, K., Leinfelder, W., and Bock, A. (1989) Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature* 342, 453-456
68. Stadtman, T. C. (1996) Selenocysteine. *Annu. Rev. Biochem.* 65, 83-100
69. Cowie, D. B., and Cohen, G. N. (1957) Biosynthesis by *Escherichia coli* of active altered proteins containing selenium instead of sulfur. *Biochim. Biophys. Acta* 26, 252-261
70. Muller, S., Senn, H., Gsell, B., Vetter, W., Baron, C., and Bock, A. (1994) The formation of diselenide bridges in proteins by incorporation of selenocysteine residues: biosynthesis and characterization of (Se)₂-thioredoxin. *Biochemistry* 33, 3404-3412
71. Muller, S., Heider, J., and Bock, A. (1997) The path of unspecific incorporation of selenium in *Escherichia coli*. *Arch. Microbiol.* 168, 421-427
72. Fimmel, A. L., and Loughlin, R. E. (1977) Isolation and characterization of *cysK* mutants of *Escherichia coli* K12. *J. Gen. Microbiol.* 103, 37-43
73. Young, P. A., and Kaiser, II (1975) Aminoacylation of *Escherichia coli* cysteine tRNA by selenocysteine. *Arch. Biochem. Biophys.* 171, 483-489
74. Stadtman, T. C., Davis, J. N., Zehelein, E., and Bock, A. (1989) Biochemical and genetic analysis of *Salmonella typhimurium* and *Escherichia coli* mutants defective in specific incorporation of selenium into formate dehydrogenase and tRNAs. *BioFactors* 2, 35-44
75. Lehmann, W. D. (1996) Massenspektrometrie in der Bioanalytik. *Heidelberg: Spektrum Akademischer Verlag*
76. Gross, J. H. (2010) Mass Spectrometry – A textbook, 2nd ed. Berlin/Heidelberg: *Springer Verlag*

-
77. Horning, E. C., Horning, M. G., Carroll, D. I., Dzidic, I., and Stillwell, R. N. (1973) New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure. *Anal. Chem.* 45, 936-943
78. Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., and Alice, M. B. (1968) Molecular beams of macroions. *J. Chem. Phys.* 49, 2240-2249
79. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64-71
80. Yamashita, M., and Fenn, J. B. (1984) Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* 88, 4451-4459
81. Dole, R. B. (1997) Electrospray Ionization Mass Spectrometry - Fundamentals, instrumentation and applications. *John Wiley & Sons: Chichester*
82. Fenn, J. B., Mann, M., Meng, C. K., and Wong, S. F. (1990) ESI - principles and practice. *Mass. Spectrom. Rev.* 9, 37-70
83. Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., and Udseth, H. R. (1990) New developments in biochemical mass spectrometry: electrospray ionization. *Anal. Chem.* 62, 882-899
84. Fenn, J. B. (2003) Electrospray wings for molecular elephants (nobel lecture). *Angewandte Chemie* 42, 3871-3894
85. Zeleny, J. (1917) Instability of electrified liquid surfaces. *Phys. Rev.* 10, 1-7
86. Taylor, G. I. (1964) Disintegration of water drops in an electric field. *Proc. R. Soc. Lond. A* 280, 383-397
87. Wilm, M. S., and Mann, M. (1994) Electrospray and taylor-cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion. Proc.* 136, 167-180
88. Raleigh, L. (1882) On the equilibrium of liquid conducting masses charged with electricity. *London, Edinburgh, Dublin Phil. Mag. J. Sci.* 14, 184-186

89. Gomez, A., and Tang, K. (1994) Charge and fission of droplets in electrostatic sprays. *Phys. Fluids* 6, 404-414
90. Duft, D., Achtzehn, T., Muller, R., Huber, B. A., and Leisner, T. (2003) Coulomb fission: Rayleigh jets from levitated microdroplets. *Nature* 421, 128
91. Konermann, L. (2007) A minimalist model for exploring conformational effects on the electrospray charge state distribution of proteins. *J. Phys. Chem. B* 111, 6534-6543
92. Konermann, L. (2009) A simple model for the disintegration of highly charged solvent droplets during electrospray ionization. *J. Am. Soc. Mass Spectrom.* 20, 496-506
93. Cole, R. B. (2000) Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass Spectrom.* 35, 763-772
94. Kebarle, P., and Tang, L. (1993) From ions in solution to ions in the gasphase - the mechanism of ESI-MS. *Anal. Chem.* 65
95. *Photograph by Robert Lamberts at The New Zealand Institute for Plant and Food Research Ltd, (2008)*
96. Felitsyn, N., Peschke, M., and Kebarle, P. (2002) Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *J. Mass Spectrom.* 219, 39-62
97. Mack, L. L., Kralik, P., Rheude, A., and Dole, M. (1970) Molecular beams of macroions. II. *J. Chem. Phys.* 52, 4977-4986
98. Loo, J. A., Edmonds, C. G., Udseth, H. R., and Smith, R. D. (1990) Effect of reducing disulfide-containing proteins on electrospray ionization mass spectra. *Anal. Chem.* 62, 693-698
99. Thomson, B. A., and Iribarne, J. V. (1979) Field-induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* 71, 4451-4463

-
100. Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) Probing conformational changes in proteins by MS. *J. Am. Chem. Soc.* 112, 9012-9013
101. Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* 68, 1-8
102. Schwartz, J. C., Senko, M. W., and Syka, J. E. (2002) A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 13, 659-669
103. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* 78, 2113-2120
104. Douglas, D. J., Frank, A. J., and Mao, D. (2005) Linear ion traps in mass spectrometry. *Mass Spectrom. Rev.* 24, 1-29
105. Schwartz, J. C., Zhou, X. G., and Bier, E. U.S. Patent 5,572,022
106. Makarov, A. (1999) U.S. Patent 5, 5,886,346
107. Makarov, A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* 72, 1156-1162
108. Kingdon, K. H. (1923) A Method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Phys. Rev.* 21, 408-418
109. Scigelova, M., and Makarov, A. (2006) Orbitrap mass analyzer-overview and applications in proteomics. *Proteomics* 6 Suppl 2, 16-21
110. Perry, R. H., Cooks, R. G., and Noll, R. J. (2008) Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom. Rev.* 27, 661-699
111. Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., and Graham Cooks, R. (2005) The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 40, 430-443
112. Baldwin, M. A. (2004) Protein identification by mass spectrometry: issues to be considered. *Mol. Cell. Proteomics* 3, 1-9

113. Aebersold, R., and Goodlett, D. R. (2001) Mass spectrometry in proteomics. *Chem. Rev.* 101, 269-295
114. Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* 11, 601
115. Biemann, K. (1992) Mass spectrometry of peptides and proteins. *Annu. Rev. Biochem.* 61, 977-1010
116. de Hoog, C. L., and Mann, M. (2004) Proteomics. *Annu. Rev. Genomics Hum. Genet.* 5, 267-293
117. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255-261
118. Virshup, D. M., Eide, E. J., Forger, D. B., Gallego, M., and Harnish, E. V. (2007) Reversible protein phosphorylation regulates circadian rhythms. *Cold Spring Harbor Laboratory: Symposia on Quantitative Biology* 72, 413-420
119. Yang, X. J. (2004) Lysine acetylation and the bromodomain: a new partnership for signaling. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 26, 1076-1087
120. Gill, G. (2004) SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev.* 18, 2046-2059
121. Zhi, Y., and Sandri-Goldin, R. M. (1999) Analysis of the phosphorylation sites of herpes simplex virus type 1 regulatory protein ICP27. *J. Virol.* 73, 3246-3257
122. Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y., Tran, H., Ross, S. E., Mostoslavsky, R., Cohen, H. Y., Hu, L. S., Cheng, H. L., Jedrychowski, M. P., Gygi, S. P., Sinclair, D. A., Alt, F. W., and Greenberg, M. E. (2004) Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science* 303, 2011-2015
123. Yuan, Z. L., Guan, Y. J., Chatterjee, D., and Chin, Y. E. (2005) Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science* 307, 269-273

124. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648
125. Elortza, F., Nuhse, T. S., Foster, L. J., Stensballe, A., Peck, S. C., and Jensen, O. N. (2003) Proteomic analysis of glycosylphosphatidylinositol-anchored membrane proteins. *Mol. Cell. Proteomics* 2, 1261-1270
126. Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., Grishin, N. V., White, M., Yang, X. J., and Zhao, Y. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* 23, 607-618
127. Denison, C., Rudner, A. D., Gerber, S. A., Bakalarski, C. E., Moazed, D., and Gygi, S. P. (2005) A proteomic strategy for gaining insights into protein sumoylation in yeast. *Mol. Cell. Proteomics* 4, 246-254
128. Martinez-Ruiz, A., and Lamas, S. (2007) Proteomic identification of S-nitrosylated proteins in endothelial cells. *Meth. Mol. Biol.* 357, 215-223
129. Zhao, Y., and Jensen, O. N. (2009) Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 9, 4632-4641
130. Mirza, U. A., Chait, B. T., and Lander, H. M. (1995) Monitoring reactions of nitric oxide with peptides and proteins by electrospray ionization-mass spectrometry. *J. Biol. Chem.* 270, 17185-17188
131. Ma, S., Caprioli, R. M., Hill, K. E., and Burk, R. F. (2003) Loss of selenium from selenoproteins: conversion of selenocysteine to dehydroalanine *in vitro*. *J. Am. Soc. Mass Spectrom.* 14, 593-600
132. Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037-1053

133. Herbert, B. R., Harry, J. L., Packer, N. H., Gooley, A. A., Pedersen, S. K., and Williams, K. L. (2001) What place for polyacrylamide in proteomics? *Trends Biotechnol.* 19, 3-9
134. Patton, W. F., and Beechem, J. M. (2002) Rainbow's end: the quest for multiplexed fluorescence quantitative analysis in proteomics. *Curr. Opin. Chem. Biol.* 6, 63-69
135. Zhou, G., Li, H., DeCamp, D., Chen, S., Shu, H., Gong, Y., Flaig, M., Gillespie, J. W., Hu, N., Taylor, P. R., Emmert-Buck, M. R., Liotta, L. A., Petricoin, E. F., 3rd, and Zhao, Y. (2002) 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol. Cell. Proteomics* 1, 117-124
136. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* 389, 1017-1031
137. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007-4021
138. Klose, J., and Kobalz, U. (1995) Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034-1059
139. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207
140. Gygi, S. P., Rist, B., and Aebersold, R. (2000) Measuring gene expression by quantitative proteome analysis. *Curr. Opin. Biotechnol.* 11, 396-401
141. Heck, A. J., and Krijgsveld, J. (2004) Mass spectrometry-based quantitative proteomics. *Expert Rev. Proteomics* 1, 317-326
142. Ong, S. E., Foster, L. J., and Mann, M. (2003) Mass spectrometric-based approaches in quantitative proteomics. *Methods* 29, 124-130

-
143. Righetti, P. G., Campostrini, N., Pascali, J., Hamdan, M., and Astner, H. (2004) Quantitative proteomics: a review of different methodologies. *Eur. J. Mass Spectrom. (Chichester, Eng)* 10, 335-348
144. Sechi, S., and Oda, Y. (2003) Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* 7, 70-77
145. Tao, W. A., and Aebersold, R. (2003) Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr. Opin. Biotechnol.* 14, 110-118
146. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252-262
147. Lahm, H. W., and Langen, H. (2000) Mass spectrometry: a tool for the identification of proteins separated by gels. *Electrophoresis* 21, 2105-2114
148. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA* 96, 6591-6596
149. 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. *Mol. Cell. Proteomics* 1, 376-386
150. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* 73, 2836-2842
151. Reynolds, K. J., Yao, X., and Fenselau, C. (2002) Proteolytic ^{18}O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J. Proteome Res.* 1, 27-33
152. Rose, K., Simona, M. G., Offord, R. E., Prior, C. P., Otto, B., and Thatcher, D. R. (1983) A new mass-spectrometric C-terminal sequencing technique finds a similarity between gamma-interferon and alpha 2-interferon and identifies a

proteolytically clipped gamma-interferon that retains full antiviral activity. *Biochem. J.* 215, 273-277

153. 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. *Nat. Biotechnol.* 17, 994-999

154. Schmidt, A., Kellermann, J., and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4-15

155. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlett-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154-1169

156. Hsu, J. L., Huang, S. Y., and Chen, S. H. (2006) Dimethyl multiplexed labeling combined with microcolumn separation and MS analysis for time course study in proteomics. *Electrophoresis* 27, 3652-3660

157. Ji, J., Chakraborty, A., Geng, M., Zhang, X., Amini, A., Bina, M., and Regnier, F. (2000) Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr., B. Biomed. Sci. Appl.* 745, 197-210

158. Che, F. Y., and Fricker, L. D. (2002) Quantitation of neuropeptides in Cpe(fat)/Cpe(fat) mice using differential isotopic tags and mass spectrometry. *Anal. Chem.* 74, 3190-3198

159. Zhang, X., Jin, Q. K., Carr, S. A., and Annan, R. S. (2002) N-Terminal peptide labeling strategy for incorporation of isotopic tags: a method for the determination of site-specific absolute phosphorylation stoichiometry. *Rapid Commun. Mass Spectrom.* 16, 2325-2332

160. Glocker, M. O., Borchers, C., Fiedler, W., Suckau, D., and Przybylski, M. (1994) Molecular characterization of surface topology in protein tertiary structures by amino-acylation and mass spectrometric peptide mapping. *Bioconjugate Chem.* 5, 583-590

-
161. Bondarenko, P. V., Chelius, D., and Shaler, T. A. (2002) Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal. Chem.* 74, 4741-4749
162. Chelius, D., and Bondarenko, P. V. (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J. Proteome Res.* 1, 317-323
163. Wang, W., Zhou, H., Lin, H., Roy, S., Shaler, T. A., Hill, L. R., Norton, S., Kumar, P., Anderle, M., and Becker, C. H. (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* 75, 4818-4826
164. Wiener, M. C., Sachs, J. R., Deyanova, E. G., and Yates, N. A. (2004) Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal. Chem.* 76, 6085-6096
165. Higgs, R. E., Knierman, M. D., Gelfanova, V., Butler, J. P., and Hale, J. E. (2005) Comprehensive label-free method for the relative quantification of proteins from biological samples. *J. Proteome Res.* 4, 1442-1450
166. Wang, G., Wu, W. W., Zeng, W., Chou, C. L., and Shen, R. F. (2006) Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: Reproducibility, linearity, and application with complex proteomes. *J. Proteome Res.* 5, 1214-1223
167. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19, 242-247
168. Liu, H., Sadygov, R. G., and Yates, J. R., 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 76, 4193-4201
169. Gilchrist, A., Au, C. E., Hiding, J., Bell, A. W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S. J., Hallett, M., Paiement, J., Kearney,

- R. E., Nilsson, T., and Bergeron, J. J. (2006) Quantitative proteomics analysis of the secretory pathway. *Cell* 127, 1265-1281
170. Braun, P., and Gingras, A. C. (2012) History of protein-protein interactions: from egg-white to complex networks. *Proteomics* 12, 1478-1498
171. Yanagida, M. (2002) Functional proteomics; current achievements. *J. chromatography. B: Anal. Technol. Biomed. Life Sci.* 771, 89-106
172. Berggard, T., Linse, S., and James, P. (2007) Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7, 2833-2842
173. Ofran, Y., and Rost, B. (2003) Analysing six types of protein-protein interfaces. *J. Mol. Biol.* 325, 377-387
174. Bashor, C. J., Horwitz, A. A., Peisajovich, S. G., and Lim, W. A. (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu. Rev. Biophys.* 39, 515-537
175. McNally, F. J., and Vale, R. D. (1993) Identification of katanin, an ATPase that severs and disassembles stable microtubules. *Cell* 75, 419-429
176. Dutcher, S. K. (2001) The tubulin fraternity: alpha to eta. *Curr. Opin. Cell Biol.* 13, 49-54
177. Hemmerich, P., Schmiedeberg, L., and Diekmann, S. (2011) Dynamic as well as stable protein interactions contribute to genome function and maintenance. *Chromosome Res.* 19, 131-151
178. Schreiber, G., Haran, G., and Zhou, H. X. (2009) Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* 109, 839-860
179. Gavin, A. C., Maeda, K., and Kuhner, S. (2011) Recent advances in charting protein-protein interaction: mass spectrometry-based approaches. *Curr. Opin. Biotechnol.* 22, 42-49
180. Sprinzak, E., Sattath, S., and Margalit, H. (2003) How reliable are experimental protein-protein interaction data? *J. Mol. Biol.* 327, 919-923

-
181. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? *Nat. Rev. Drug Discov.* 5, 993-996
182. Rao, V. S., Srinivas, K., Sujini, G. N., and Kumar, G. N. (2014) Protein-Protein interaction detection: methods and analysis. *Int. J. Proteomics* 2014, 147648
183. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. USA* 97, 5802-5806
184. Hermanson, G. T. (1996) *Bioconjugate techniques* Academic Press, San Diego, CA
185. Sinz, A. (2003) Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass Spectrom.* 38, 1225-1237
186. Sinz, A. (2005) Chemical cross-linking and FTICR mass spectrometry for protein structure characterization. *Anal. Bioanal. Chem.* 381, 44-47
187. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* 25, 663-682
188. Sinz, A. (2005) Chemical cross-linking and FTICR mass spectrometry for protein structure characterization. *Anal. Bioanal. Chem.* 381, 44-47
189. Gingras, A. C., Gstaiger, M., Raught, B., and Aebersold, R. (2007) Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* 8, 645-654
190. Back, J. W., de Jong, L., Muijsers, A. O., and de Koster, C. G. (2003) Chemical cross-linking and mass spectrometry for protein structural modeling. *J. Mol. Biol.* 331, 303-313
191. Rappsilber, J. (2011) The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* 173, 530-540

-
192. Sinz, A. (2010) Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.* 397, 3433-3440
193. Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J. C., Nilges, M., Cramer, P., and Rappsilber, J. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* 29, 717-726
194. Kalisman, N., Adams, C. M., and Levitt, M. (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl. Acad. Sci. USA* 109, 2884-2889
195. Kalkhof, S., and Sinz, A. (2008) Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Anal. Bioanal. Chem.* 392, 305-312
196. Madler, S., Bich, C., Touboul, D., and Zenobi, R. (2009) Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* 44, 694-706
197. Krauth, F., Ihling, C. H., Ruttinger, H. H., and Sinz, A. (2009) Heterobifunctional isotope-labeled amine-reactive photo-cross-linker for structural investigation of proteins by matrix-assisted laser desorption/ionization tandem time-of-flight and electrospray ionization LTQ-Orbitrap mass spectrometry. *Rapid Commun. Mass Spectrom.* 23, 2811-2818
198. Platz, M. S., Huang, H. Y., and Ford, F. (1997) Photochemical rearrangements of diazirines and thermal rearrangements of carbenes. *Pure Appl. Chem.* 69, 803-807
199. Clegg, C., and Hayes, D. (1974) Identification of neighbouring proteins in the ribosomes of *Escherichia coli*. A topographical study with the cross-linking reagent dimethyl suberimidate. *Eur. J. Biochem. / FEBS* 42, 21-28
200. Schilling, B., Row, R. H., Gibson, B. W., Guo, X., and Young, M. M. (2003) MS²Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.* 14, 834-850

201. Fritzsche, R., Ihling, C. H., Gotze, M., and Sinz, A. (2012) Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Spectrom.* 26, 653-658
202. Alley, S. C., Ishmael, F. T., and Jones, A. D. (2000) Mapping protein-protein interactions in the bacteriophage T4 DNA polymerase holoenzyme using a novel trifunctional photo-cross-linking and affinity reagent. *J. Am. Chem. Soc.* 122, 6126-6127
203. Muller, D. R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., and Steinmetz, M. O. (2001) Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* 73, 1927-1934
204. Muller, M. Q., Dreiocker, F., Ihling, C. H., Schafer, M., and Sinz, A. (2010) Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS. *Anal. Chem.* 82, 6958-6968
205. Peri, S., Steen, H., and Pandey, A. (2001) GPMAW - a software tool for analyzing proteins and peptides. *Trends Biochem. Sci.* 26, 687-689
206. de Koning, L. J., Kasper, P. T., Back, J. W., Nessen, M. A., Vanrobaeys, F., Van Beeumen, J., Gherardi, E., de Koster, C. G., and de Jong, L. (2006) Computer-assisted mass spectrometric analysis of naturally occurring and artificially introduced cross-links in proteins and protein complexes. *FEBS J.* 273, 281-291
207. Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L. N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008) Identification of cross-linked peptides from large sequence databases. *Nat. Meth.* 5, 315-318
208. Gotze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C. H., Krauth, F., Fritzsche, R., Kuhn, U., and Sinz, A. (2012) StavroX - a software for analyzing crosslinked products in protein interaction studies. *J. Am. Soc. Mass Spectrom.* 23, 76-87
209. Tuve, T., and Williams, H. H. (1961) Metabolism of selenium by *Escherichia coli*: biosynthesis of selenomethionine. *J. Biol. Chem.* 236, 597-601

210. Luke, I., Butland, G., Moore, K., Buchanan, G., Lyall, V., Fairhurst, S., Greenblatt, J., Emili, A., Palmer, T., and Sargent, F. (2008) Biosynthesis of the respiratory formate dehydrogenases from *Escherichia coli*: characterization of the FdhE protein. *Arch. Microbiol.* 190, 685-696
211. Manske, M. (2006) GENTle, a free multi-purpose molecular biology tool. Dissertation. *Universität zu Köln*
212. Chen, V. B., Davis, I. W., and Richardson, D. C. (2009) KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. *Protein Sci.* 18, 2403-2409
213. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12-21
214. Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551-3567
215. www.expasy.org
216. Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G. M., Carnemolla, B., Orecchia, P., Zardi, L., and Righetti, P. G. (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25, 1327-1333
217. Casadaban, M. J., and Cohen, S. N. (1979) Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: *in vivo* probe for transcriptional control sequences. *Proc. Natl. Acad. Sci. USA* 76, 4530-4533
218. Sambrook, J., and Russell, D. (2001) Molecular cloning: a laboratory manual. *Cold Spring Harbor N.Y., Cold Spring Harbor Laboratory Press*

219. Begg, Y. A., Whyte, J. N., and Haddock, B. A. (1977) The Identification of Mutants of *Escherichia coli* deficient in formate dehydrogenase and nitrate reductase activities using dye indicator plates. *FEMS Microbiol. Lett.* 2, 47-50
220. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685
221. Consortium, U. (2011) Reorganizing the protein space at the universal protein resource (UniProt). *Nucleic acids Res.* 40, D71-D75
222. Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P. D., Smith, C. A., Sheffler, W., Davis, I. W., Cooper, S., Treuille, A., Mandell, D. J., Richter, F., Ban, Y. E., Fleishman, S. J., Corn, J. E., Kim, D. E., Lyskov, S., Berrondo, M., Mentzer, S., Popovic, Z., Havranek, J. J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J. J., Kuhlman, B., Baker, D., and Bradley, P. (2011) ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* 487, 545-574
223. Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., and Zardecki, C. (2002) The protein data bank. *Acta Crystallogr. D, Biol. Crystallogr.* 58, 899-907
224. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112, 535-542
225. Fleishman, S. J., Leaver-Fay, A., Corn, J. E., Strauch, E. M., Khare, S. D., Koga, N., Ashworth, J., Murphy, P., Richter, F., Lemmon, G., Meiler, J., and Baker, D. (2011) RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoSone* 6, e20161
226. Viklund, H., and Elofsson, A. (2008) OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics* 24, 1662-1668

227. Qian, B., Raman, S., Das, R., Bradley, P., McCoy, A. J., Read, R. J., and Baker, D. (2007) High-resolution structure prediction and the crystallographic phase problem. *Nature* 450, 259-264
228. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003) Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* 331, 281-299
229. Alexander, N., Woetzel, N., and Meiler, J. (2011) Bcl::Cluster : A method for clustering biological molecules coupled with visualization in the Pymol Molecular Graphics System. In *Computational Advances in Bio and Medical Sciences (ICCABS), 2011 IEEE 1st International Conference on (February 2011)*, pp. 13-18, doi:10.1109/iccabs.2011.5729867
230. Sezonov, G., Joseleau-Petit, D., and D'Ari, R. (2007) *Escherichia coli* physiology in Luria-Bertani broth. *J. Bacteriol.* 189, 8746-8749
231. Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F. U., Ban, N., Malmstrom, L., and Aebersold, R. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337, 1348-1352
232. Quan, S., Hiniker, A., Collet, J. F., and Bardwell, J. C. (2013) Isolation of bacteria envelope proteins. *Methods Mol. Biol.* 966, 359-366
233. Dimova, K., Kalkhof, S., Pottratz, I., Ihling, C., Rodriguez-Castaneda, F., Liepold, T., Griesinger, C., Brose, N., Sinz, A., and Jahn, O. (2009) Structural insights into the calmodulin-Munc13 interaction obtained by cross-linking and mass spectrometry. *Biochemistry* 48, 5908-5921
234. Jones-Mortimer, M. C., Wheldrake, J. F., and Pasternak, C. A. (1968) The control of sulphate reduction in *Escherichia coli* by O-acetyl-L-serine. *Biochem. J.* 107, 51-53
235. Sharp, P. M., Cowe, E., Higgins, D. G., Shields, D. C., Wolfe, K. H., and Wright, F. (1988) Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster*

and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Res.* 16, 8207-8211

236. Ngounou Wetie, A. G., Sokolowska, I., Woods, A. G., Roy, U., Deinhardt, K., and Darie, C. C. (2014) Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. *Cell. Mol. Life Sci.* 71, 205-228

237. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* 9, 1634-1649

238. Zhang, H., Tang, X., Munske, G. R., Tolic, N., Anderson, G. A., and Bruce, J. E. (2009) Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. *Mol. Cell. Proteomics* 8, 409-420

239. Jennebach, S., Herzog, F., Aebersold, R., and Cramer, P. (2012) Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. *Nucleic acids Res.* 40, 5591-5601

240. Lasker, K., Forster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. USA* 109, 1380-1387

241. Lauber, M. A., and Reilly, J. P. (2010) Novel amidinating cross-linker for facilitating analyses of protein structures and interactions. *Anal. Chem.* 82, 7736-7743

242. Hirst, S. J., Alexander, N., McHaourab, H. S., and Meiler, J. (2011) RosettaEPR: an integrated tool for protein structure determination from sparse EPR data. *J. Struct. Biol.* 173, 506-514

243. Chen, R., and Henning, U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol. Microbiol.* 19, 1287-1294

244. Burmann, B. M., Wang, C. W., and Hiller, S. (2013) Conformation and dynamics of the periplasmic membrane-protein- chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* 20, 1265-U1186
245. Walton, T. A., and Sousa, M. C. (2004) Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell* 15, 367-374
246. Korndorfer, I. P., Dommel, M. K., and Skerra, A. (2004) Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat. Struct. Mol. Biol.* 11, 1015-1020
247. Qu, J., Mayer, C., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2007) The trimeric periplasmic chaperone Skp of *Escherichia coli* forms 1 : 1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions. *J. Mol. Biol.* 374, 91-105
248. Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. USA* 106, 1772-1777
249. Missiakas, D., Betton, J. M., and Raina, S. (1996) New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* 21, 871-884
250. Schafer, U., Beck, K., and Muller, M. (1999) Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J. Biol. Chem.* 274, 24567-24574
251. Entzminger, K. C., Chang, C., Myhre, R. O., McCallum, K. C., and Maynard, J. A. (2012) The Skp chaperone helps fold soluble proteins *in vitro* by inhibiting aggregation. *Biochemistry* 51, 4822-4834
252. Jarchow, S., Luck, C., Gorg, A., and Skerra, A. (2008) Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. *Proteomics* 8, 4987-4994

253. Goemans, C., Denoncin, K., and Collet, J. F. (2013) Folding mechanisms of periplasmic proteins. *Biochim. Biophys. Acta* 1843, 1517-28
254. de Gier, J. W., and Luirink, J. (2001) Biogenesis of inner membrane proteins in *Escherichia coli*. *Mol. Microbiol.* 40, 314-322
255. Hatzixanthis, K., Richardson, D. J., and Sargent, F. (2005) Chaperones involved in assembly and export of N-oxide reductases. *Biochem. Soc. Trans.* 33, 124-126
256. Sargent, F., Berks, B. C., and Palmer, T. (2002) Assembly of membrane-bound respiratory complexes by the Tat protein-transport system. *Arch. Microbiol.* 178, 77-84
257. Olmsted, I. R., Xiao, Y., Cho, M., Csordas, A. T., Sheehan, J. H., Meiler, J., Soh, H. T., and Bornhop, D. J. (2011) Measurement of aptamer-protein interactions with back-scattering interferometry. *Anal. Chem.* 83, 8867-8870
258. Pearce, S. R., Mimmack, M. L., Gallagher, M. P., Gileadi, U., Hyde, S. C., and Higgins, C. F. (1992) Membrane topology of the integral membrane-components, OppB and OppC, of the oligopeptide permease of *Salmonella typhimurium*. *Mol. Microbiol.* 6, 47-57
259. Hiles, I. D., Gallagher, M. P., Jamieson, D. J., and Higgins, C. F. (1987) Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. *J. Mol. Biol.* 195, 125-142
260. Goodell, E. W., and Higgins, C. F. (1987) Uptake of cell-wall peptides by *Salmonella typhimurium* and *Escherichia coli*. *J. Bacteriol.* 169, 3861-3865
261. Tame, J. R. H., Murshudov, G. N., Dodson, E. J., Neil, T. K., Dodson, G. G., Higgins, C. F., and Wilkinson, A. J. (1994) The structural basis of sequence-independent peptide binding by OppA protein. *Science* 264, 1578-1581
262. Glover, I. D., Denny, R. C., Nguti, N. D., Mcsweeney, S. M., Kinder, S. H., Thompson, A. W., Dodson, E. J., Wilkinson, A. J., and Tame, J. R. H. (1995) Structure determination of OppA at 2.3-Ångstrom resolution using multiple-

wavelength anomalous-dispersion methods. *Acta Crystallogr. D Biol. Crystallogr.* 51, 39-47

263. Vollmer, W. (2008) Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiol. Rev.* 32, 287-306

264. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* 2, a000414 doi: 10.1101/cshperspect.a000414

265. Sanders, A. N., and Pavelka, M. S. (2013) Phenotypic analysis of *Escherichia coli* mutants lacking L,D-transpeptidases. *Microbiology* 159, 1842-1852

266. Magnet, S., Dubost, L., Marie, A., Arthur, M., and Gutmann, L. (2008) Identification of the L,D-transpeptidases for peptidoglycan cross-linking in *Escherichia coli*. *J. Bacteriol.* 190, 4782-4785

267. Derouiche, R., Benedetti, H., Lazzaroni, J. C., Lazdunski, C., and Lloubes, R. (1995) Protein complex within *Escherichia coli* inner membrane - TolA N-terminal domain interacts with TolQ and TolR proteins. *J. Biol. Chem.* 270, 11078-11084

268. Lazzaroni, J. C., Vianney, A., Popot, J. L., Benedetti, H., Samatey, F., Lazdunski, C., Portalier, R., and Geli, V. (1995) Transmembrane alpha-helix interactions are required for the functional assembly of the *Escherichia coli* Tol complex. *J. Mol. Biol.* 246, 1-7

269. Germon, P., Clavel, T., Vianney, A., Portalier, R., and Lazzaroni, J. C. (1998) Mutational analysis of the *Escherichia coli* K-12 TolA N-terminal region and characterization of its TolQ-interacting domain by genetic suppression. *J. Bacteriol.* 180, 6433-6439

270. Bouveret, E., Derouiche, R., Rigal, A., Lloubes, R., Lazdunski, C., and Benedetti, H. (1995) Peptidoglycan-associated lipoprotein-Tolb interaction. *J. Biol. Chem.* 270, 11071-11077

271. Lazzaroni, J. C., Germon, P., Ray, M. C., and Vianney, A. (1999) The Tol proteins of *Escherichia coli* and their involvement in the uptake of biomolecules and outer membrane stability. *FEMS Microbiol. Lett.* 177, 191-197

272. Walburger, A., Lazdunski, C., and Corda, Y. (2002) The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB. *Mol. Microbiol.* 44, 695-708
273. Clavel, T., Germon, P., Vianney, A., Portalier, R., and Lazzaroni, J. C. (1998) TolB protein of *Escherichia coli* K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. *Mol. Microbiol.* 29, 359-367
274. Papadakos, G., Housden, N. G., Lilly, K. J., Kaminska, R., and Kleanthous, C. (2012) Kinetic basis for the competitive recruitment of TolB by the intrinsically disordered translocation domain of colicin E9. *J. Mol. Biol.* 418, 269-280
275. Mullineaux, C. W., Nenninger, A., Ray, N., and Robinson, C. (2006) Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. *J. Bacteriol.* 188, 3442-3448
276. Mogensen, J. E., and Otzen, D. E. (2005) Interactions between folding factors and bacterial outer membrane proteins. *Mol. Microbiol.* 57, 326-346
277. Beveridge, T. J. (1995) The periplasmic space and the concept of the periplasm in gram-positive and gram-negative bacteria. *Am. Soc. Microbiol. News* 61, 125-130
278. Beveridge, T. J., and Graham, L. L. (1991) Surface layers of bacteria. *Microbiol. Rev.* 55, 684-705
279. Kessler, D., Herth, W., and Knappe, J. (1992) Ultrastructure and pyruvate formate-lyase radical quenching property of the multienzymic AdhE protein of *Escherichia coli*. *J. Biol. Chem.* 267, 18073-18079
280. Kulzer, R., Pils, T., Kappl, R., Huttermann, J., and Knappe, J. (1998) Reconstitution and characterization of the polynuclear iron-sulfur cluster in pyruvate formate-lyase-activating enzyme. Molecular properties of the holoenzyme form. *J. Biol. Chem.* 273, 4897-4903
281. Beyer, L., Doberenz, C., Falke, D., Hunger, D., Suppmann, B., and Sawers, R. G. (2013) Coordination of FocA and pyruvate formate-lyase synthesis in *Escherichia*

coli demonstrates preferential translocation of formate over other mixed-acid fermentation products. *J. Bacteriol.* 195, 1428-1435

282. Lu, W., Du, J., Schwarzer, N. J., Gerbig-Smentek, E., Einsle, O., and Andrade, S. L. (2012) The formate channel FocA exports the products of mixed-acid fermentation. *Proc. Natl. Acad. Sci. USA* 109, 13254-13259

283. Lu, W., Du, J., Schwarzer, N. J., Wacker, T., Andrade, S. L., and Einsle, O. (2013) The formate/nitrite transporter family of anion channels. *Biol. Chem.* 394, 715-727

284. Green, N. S., Reisler, E., and Houk, K. N. (2001) Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.* 10, 1293-1304

285. Lv, X., Liu, H., Ke, M., and Gong, H. (2013) Exploring the pH-dependent substrate transport mechanism of FocA using molecular dynamics simulation. *Biophys. J.* 105, 2714-2723

Danksagung

Mein innigster Dank gilt Frau Prof. Dr. Andrea Sinz für die Überlassung des interessanten Themas, das Engagement und Interesse an dem Fortgang der Arbeit, die stetige Diskussionsbereitschaft und die Hilfsbereitschaft bei fachlichen Problemen oder Fragen, sowie für die Übernahme des Erstgutachtens.

Ein besonderer Dank gilt Herrn Prof. Dr. R. Gary Sawers für die fachlich kompetente Unterstützung während der gesamten Arbeit, die uneingeschränkte Hilfsbereitschaft, die sehr herzliche Betreuung, das Vertrauen in meine Arbeit sowie der Übernahme des Zweitgutachtens.

Prof. Dr. Jens Meiler danke ich für die großartige Möglichkeit die bioinformatisch gestützte Charakterisierung von Protein-Protein-Interaktionen mit Hilfe des Software-Pakets Rosetta zu erlernen. Desweiteren danke ich ihm für seine tolle Unterstützung während des Forschungsaufenthaltes an der Vanderbilt University (Nashville, Tennessee, USA) sowie der Übernahme des Drittgutachtens.

Dr. David Nannemann danke ich für die herausragende wissenschaftliche Unterstützung zur Modellierung des FocA/PfIB-Komplexes.

Ich danke Stephanie DeLuca und Elizabeth Dong für die Optimierung von Rosetta für die Strukturvorhersage von Membranproteinen.

Danken möchte ich auch Dr. Christian Ihling für die Instandhaltung der LC/MS-Geräte sowie der fachkundigen Beantwortung aller massenspektrometrischen Fragestellungen. Auch Dr. Knut Kölbl und Dirk Tänzler danke ich für die Hilfe bei wissenschaftlichen Fragestellungen und der Unterstützung im Labor.

Weiterhin danke ich PD Dr. Ralph Golbik für die kompetente Unterstützung zur Produktion und Reingung der Formiatdehydrogenasen.

Ich möchte mich bei allen Kollegen des Instituts für Pharmazie für die angenehme Zusammenarbeit bedanken.

Den Mitgliedern des Graduiertenkollegs 1026 danke ich für die Kooperationen und zahlreiche Hilfestellungen verschiedenster Art sowie der Möglichkeit mich in der Präsentation meiner Daten stetig zu verbessern.

Besonders Danken möchte ich Björn, Christine und Doreen für ihr Interesse und die Unterstützung bei fachlichen und nicht-fachlichen Angelegenheiten ;-)) sowie dem Korrekturlesen der Arbeit.

Mein größter Dank gilt meiner großartigen Familie, insbesondere meinen Eltern, meiner Freundin Sophie sowie meinem Freundeskreis für die stetige und bedingungslose Unterstützung.

Publikationen

Michael Zorn, Christian H. Ihling, Ralph Golbik, R. Gary Sawers and Andrea Sinz, Selective *se/C*-independent Selenocysteine Incorporation into Formate Dehydrogenases. *PLoS ONE* 2013, 8 (4) : e61913.

Claudia Doberenz, Michael Zorn, Dörte Falke, David Nannemann, Doreen Hunger, Lydia Beyer, Christian H. Ihling, Jens Meiler, Andrea Sinz and R. Gary Sawers, Pyruvate Formate-Lyase Interacts Directly with the Formate Channel FocA to Regulate Formate Translocation, *J. Mol. Biol.* (2014) Jul 29;426(15):2827-39. doi: 10.1016/j.jmb.2014.05.023. Epub 2014 Jun 2.

Zukünftige Publikationen

Michael Zorn, Christian H. Ihling, Ralph Golbik, R. Gary Sawers, and Andrea Sinz, Mapping Cell Envelope and Periplasm Protein Interactions of *Escherichia coli* Respiratory Formate Dehydrogenases by Chemical Cross-linking and Mass Spectrometry.

Manuskript eingereicht

Poster

Michael Zorn, Christian H. Ihling, Ralph Golbik, R. Gary Sawers, and Andrea Sinz, Purification of Native Formate Dehydrogenase-N from *Escherichia coli*. 2nd International Meeting GRK 1026, Halle 2011.

Michael Zorn, Christian H. Ihling, Ralph Golbik, R. Gary Sawers, and Andrea Sinz, Purification of Native Formate Dehydrogenase-N from *Escherichia coli*. 44. Jahrestagung der Deutschen Gesellschaft für Massenspektrometrie, Dortmund 2011.

Michael Zorn, Christian H. Ihling, Ralph Golbik, R. Gary Sawers, and Andrea Sinz, Identification of Interaction Partners from Formate Dehydrogenase-N from *Escherichia coli*. Fall Meeting GRK 1026, Spergau 2012.

Michael Zorn, Claudia Doberenz, Dörte Falke, David Nannemann, Doreen Hunger, Lydia Beyer, Christian H. Ihling, Jens Meiler, Andrea Sinz and R. Gary Sawers, Pyruvate Formate-Lyase Interacts Directly with the Formate Channel FocA to Regulate Formate Translocation. 47. Jahrestagung der Deutschen Gesellschaft für Massenspektrometrie, Frankfurt (Main) 2014.

Vorträge

Identification of Interaction Partners of Formate Dehydrogenases from *Escherichia coli*. Fall Meeting GRK 1026, Blankenburg 2010.

Identification of Interaction Partners of Formate Dehydrogenases from *Escherichia coli*. Fall Meeting GRK 1026, Leuna 2011.

Selective *sefC*-independent Selenocysteine Incorporation into Formate Dehydrogenases. Fall Meeting GRK 1026, Naumburg 2012.

Identification of Interaction Partners of Formate Dehydrogenases from *Escherichia coli*. Spring Meeting GRK 1026, Meißen 2013.

Selected Selenocysteine Incorporation in Formate Dehydrogenases. 46. Jahrestagung der Deutschen Gesellschaft für Massenspektrometrie, Berlin 2013.

Structural and Functional Insights into FocA/Pyruvate Formate-Lyase Complex. 3rd International Meeting GRK 1026, Halle 2013.

Lebenslauf

Persönliche Daten

geboren am 02.02.1983 in Halle
ledig, keine Kinder
Staatsangehörigkeit: deutsch

Schulausbildung

09/1989 – 07/1993 Grundschule Thomas Mann
09/1993 – 07/2002 Adolf Reichwein-Gymnasium
Abschluss: Abitur

Hochschulausbildung

10/2004 – 04/2010 Studium Biologie/Diplom an der
Martin-Luther-Universität Halle-Wittenberg
Vertiefungsrichtung: Mikrobiologie
08/2009 – 04/2010 Anfertigung der Diplomarbeit (Arbeitsgruppe Prof. Gray R.
Sawers) zum Thema: Die Rolle von HypF bei der Reifung
der Ni-Fe-Hydrogenasen von *E. coli*
04/2010 Abschluss: Diplom Biologe (Prädikat *sehr gut*)

Promotion

05/2010 – 04/2014 Wissenschaftlicher Mitarbeiter am Institut für Pharmazie
der Martin-Luther-Universität Halle-Wittenberg
(Arbeitsgruppe Prof. Andrea Sinz)
Thema: Identifizierung und Charakterisierung
von Protein-Protein-Interaktionen des
Formiatstoffwechsels von *E. coli* mittels Massen-
spektrometrie-basierter Methoden
01/2012 – 04/2012 Wissenschaftlicher Aufenthalt an der Vanderbilt University
(Department of Chemistry; Computational Chemical and
Structural Biology), Nashville Tennessee, USA

Berufliche Erfahrung

seit 07/2014 Wissenschaftlicher Mitarbeiter in der Immundiagnostik AG
(Bensheim)

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Dissertation selbständig und ohne unerlaubte fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen genutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet. Bei der Auswahl und Auswertung des Materials, bei der Herstellung des Manuskripts, sowie bei der geistigen Herstellung der vorgelegten Arbeit waren keine anderen Personen beteiligt. Insbesondere wurde weder die Hilfe eines Promotionsberaters in Anspruch genommen, noch haben Dritte von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die vorgelegte Arbeit ist weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt worden. Ich habe keine früheren erfolglosen Promotionsversuche unternommen.

Mühltal, den 01.10.2014

Michael Zorn