

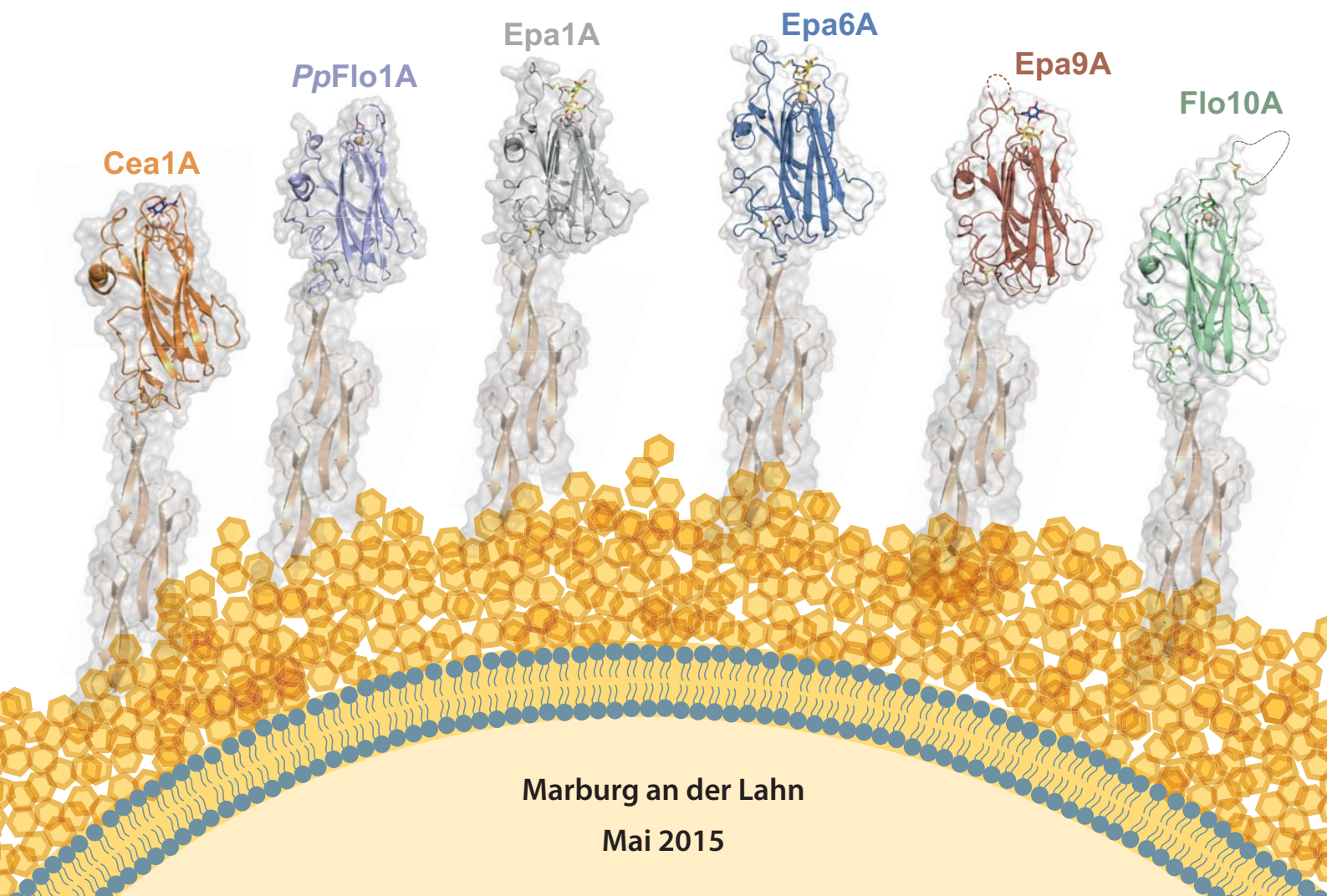
Vergleichende Struktur-/Funktionsanalyse von Adhäsionsdomänen pathogener und nicht-pathogener Hefen

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Chemie
der Philipps-Universität Marburg
vorgelegt von

Michael Alexander Kock



Vergleichende Struktur-/Funktionsanalyse von Adhäsionsdomänen pathogener und nicht-pathogener Hefen

D i s s e r t a t i o n

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

dem

Fachbereich Chemie

der Philipps-Universität Marburg

Hochschulkenziffer 1180

vorgelegt von

Michael Alexander Kock, M.Sc. Chemie

aus Marburg an der Lahn

Erstgutachter: Prof. Dr. Lars-Oliver Essen (Fachbereich Chemie)

Zweitgutachter: Prof. Dr. Hans-Ulrich Mösch (Fachbereich Biologie)

Einreichungstermin: _____

Prüfungstermin: _____

Marburg 2015



Die Untersuchungen zur vorliegenden Dissertation wurden in der Zeit vom Februar 2012 bis zum Februar 2015 unter der Leitung von Herrn Prof. Dr. Lars-Oliver Essen an der Philipps-Universität Marburg durchgeführt.

Erklärung

Ich erkläre, dass eine Promotion noch an keiner anderen Hochschule als der Philipps-Universität Marburg, Fachbereich Chemie, versucht wurde.

Ich versichere, dass meine vorgelegte Dissertation

„Vergleichende Struktur-/Funktionsanalyse von Adhäsionsdomänen pathogener und nicht-pathogener Hefen“

selbst und ohne fremde Hilfe verfasst, nicht andere als die in ihr angegebenen Quellen oder Hilfsmittel benutzt, alle vollständig oder sinngemäß übernommenen Zitate als solche gekennzeichnet sowie die Dissertation in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen in- oder ausländischen Hochschule anlässlich eines Promotionsgesuchs oder zu anderen Prüfungszwecken eingereicht habe.

Ort/Datum

Unterschrift

Publikationen

Kock M, Brückner S, Wozniak N, Veelders M, Schlereth J, Mösch H-U, Essen L-O, *High-affinity recognition of non-reducing chitinous ends by the yeast adhesin Cea1*. In Vorbereitung.

Diderrich R, **Kock M**, Maestre-Reyna M, Keller P, Steuber H, Rupp S, Essen L-O, Mösch H-U, *Structural Hotspots Determine Functional Diversity of the Candida glabrata Epithelial Adhesin Family*. In Revision.

Wozniak N, Steuber H, **Kock M**, Essen L-O, Mösch H-U, *An evolutionary novel subdomain confers functional specificity within the PA14-type adhesin family in Saccharomyces cerevisiae*. In Vorbereitung.

Patente

Kock M, Brückner S, Essen L-O, Mösch H-U, Schlereth J, *Recombinant adhesin protein and peptide thereof and methods and uses thereof*. Anmeldenummer: EP15160171.3, Anmeldedatum: 20.03.2015, Anmelder: Philipps-Universität Marburg

Konferenzbeiträge

Federation of European Biochemical Societies Congress: Mechanisms in Biology, 2013

Kock M, Brückner S, Veelders M, Schlereth J, Mösch H-U, Essen L-O, *Adhesins from Pichia pastoris – A structural basis for a symbiotic lifestyle?*

XI International Fungal Biology Conference, 2013

Kock M, Brückner S, Veelders M, Mösch H-U, Essen L-O, *Adhesins from Pichia pastoris – A structural basis for a symbiotic lifestyle?*

Molecular Life Sciences, 2013

Essen L-O, Kalugin V, **Kock M**, Kraushaar T, Veelders M, Maestre-Reyna M, *Green beards, beer & candidiasis: Fungal adhesion and its molecular base*.

In der wwPDB hinterlegte Proteinstrukturen

Kock M, Maestre-Reyna M, Diderrich R, Mösch H-U & Essen L-O

4COU Crystal Structure of Epithelial Adhesin 6 A Domain (Epa6A) from *Candida glabrata* in complex with Lactose

4COV	Crystal Structure of Epithelial Adhesin 6 A Domain (Epa6A) from <i>Candida glabrata</i> in complex with α -Galactobiose
4COW	Crystal Structure of Epithelial Adhesin 6 A Domain (Epa6A) from <i>Candida glabrata</i> in complex with the T-Antigen
4COY	Crystal Structure of Epithelial Adhesin 6 A Domain (Epa6A) from <i>Candida glabrata</i> in complex with <i>N</i> -Acetyllactosamine
4COZ	Crystal Structure of Epithelial Adhesin 6 A Domain (Epa6A) from <i>Candida glabrata</i> in complex with Lacto- <i>N</i> -Biose
4CP0	Crystal Structure of Epithelial Adhesin 9 A Domain (Epa9A) from <i>Candida glabrata</i> in complex with Lactose
4CP1	Crystal Structure of Epithelial Adhesin 9 A Domain (EpaA) from <i>Candida glabrata</i> in complex with Lacto- <i>N</i> -Biose
4CP2	Crystal Structure of Epithelial Adhesin 9 A Domain (EpaA) from <i>Candida glabrata</i> in complex with <i>N</i> -Acetyllactosamine
4D3W	Crystal Structure of Epithelial Adhesin 1 A Domain (Epa1A) from <i>Candida glabrata</i> in complex with the T-antigen (Galb1-3GalNAc)

Wissenschaftlicher Werdegang

Diese Seite enthält persönliche Daten und ist daher nicht Teil der Online-Veröffentlichung.

*„Der Geist ist nicht wie ein Gefäß, das gefüllt werden soll,
sondern wie Holz, das lediglich entzündet werden will“*

Moralia I, Über das Zuhören, Kapitel 18
— Plutarch (-45 - -125) —



Inhaltsverzeichnis

1	Einleitung	17
1.1	Hefen – Einzellige Arten des Reiches Fungi	17
1.1.1	<i>Saccharomyces cerevisiae</i> – Ein vielfältiger Modelorganismus	18
1.1.2	<i>Candida glabrata</i> - Hefen als Pathogene	19
1.1.3	<i>Pichia pastoris</i> – Eine methylothrophe Hefe	20
1.2	Die Hefezellwand als physische Barriere und Interaktionsfläche mit der Umwelt	22
1.2.1	Chitin - Ein weit verbreitetes Biopolymer vermittelt Rigidität	22
1.2.2	Vernetzte β -Glucane sorgen für strukturelle Integrität	23
1.2.3	Mannoproteine sind die Basis für wichtige zelluläre Funktionen	23
1.3	Pilzliche Zellwandproteine zur Adhäsion von Hefen	24
1.3.1	Das protektive Antigen von <i>Bacillus anthracis</i> als strukturelle Basis	26
1.3.2	Flokkuline aus <i>S. cerevisiae</i>	27
1.3.3	Epitheliale Adhäsine aus <i>C. glabrata</i>	29
1.3.4	PA14-Domänen in <i>P. pastoris</i> und weiteren Hefen	31
1.4	Zielsetzung dieser Arbeit	32
2	Material	33
2.1	Geräte	33
2.2	Verbrauchsmaterialien	34
2.3	Chemikalien	35
2.4	Saccharide und Oligosaccharide	36
2.5	Enzyme	36
2.6	Primer zur Genamplifikation	37
2.7	Primer zur Herstellung von Hefe-Expressionskonstrukten	37
2.8	Primer für die qRT-PCR	37
2.9	Vektoren	38
2.9.1	pET-28a(+)-Vektor	38
2.9.2	YCplac33-Vektor	38
2.10	Verwendete Plasmide zur Expression in <i>E. coli</i>	38
2.11	Verwendete Plasmide zur Expression in <i>S. cerevisiae</i>	39

2.12 Bakterienstämme	39
2.12.1 <i>E. coli</i> DH5	39
2.12.2 <i>E. coli</i> SHuffle® T7 Express	39
2.13 Hefestämme	40
2.13.1 <i>P. pastoris</i> DSMZ70382	40
2.13.2 <i>P. pastoris</i> GS115	40
2.13.3 <i>S. cerevisiae</i> RH2520	40
2.14 Insekten	40
2.14.1 <i>Shelfordella lateralis</i>	40
3 Methoden	42
3.1 Bioinformatische Methoden	42
3.1.1 <i>Genome-mining</i> nach Adhäsionsproteinen des PA14-Typs innerhalb der Ascomyceten	42
3.1.2 Multiples Sequenzalignment und phylogenetische Analyse	43
3.2 Molekularbiologische Methoden	45
3.2.1 Kultivierung von <i>E. coli</i>	45
3.2.2 Kultivierung von <i>S. cerevisiae</i>	45
3.2.3 Herstellung chemisch kompetenter <i>E. coli</i> -Zellen	46
3.2.4 Transformation chemisch kompetenter <i>E. coli</i> -Zellen	47
3.2.5 Stammhaltung von <i>E. coli</i>	47
3.2.6 Transformation von <i>S. cerevisiae</i> nach CHEN <i>et al.</i>	47
3.2.7 Präparation von Plasmid-DNA aus <i>E. coli</i>	48
3.2.8 Präparation von RNA aus Hefen	48
3.2.9 Konzentrationsbestimmung von DNA & RNA	49
3.2.10 Reverse Transkription mit anschließender q-PCR (qRT-PCR)	49
3.3 Proteinchemische Methoden	51
3.3.1 Heterologe Produktion von Proteinen im präparativen Maßstab	51
3.3.2 Zellaufschluss im präparativen Maßstab	52
3.3.3 Ni-NTA-Affinitätschromatografie	52
3.3.4 Größenausschlusschromatografie	53
3.3.5 Konzentration von Proteinen durch Ultrafiltration	53
3.3.6 Pufferwechsel	54
3.4 Analytische Methoden	54
3.4.1 SDS-PAGE	54
3.4.2 Konzentrationsbestimmung von Proteinlösungen	55
3.4.3 Fluoreszenzmarkierung mit AlexaFluor-488-TFP	56
3.4.4 <i>Glycan Array</i> Analyse durch das CFG	56
3.4.5 Isothermale Titrationskalorimetrie - ITC	57
3.5 Bestimmung der dreidimensionalen Struktur von Proteinen	59
3.5.1 Kristallisationsscreening	61

3.5.2	Optimierung von Kristallisationsbedingungen	62
3.5.3	<i>Streak-Seeding</i> mit Kristallkeimen	63
3.5.4	<i>Soaking</i> von Kristallen mit Sacchariden	63
3.5.5	Vorbereiten von Kristallen zur Röntgenbeugung	64
3.5.6	Röntgenbeugungsexperimente	64
3.5.7	Prozessierung und Reduktion der Daten	65
3.5.8	Das Phasenproblem – Strukturlösung mittels molekularen Ersatzes	66
3.5.9	Verfeinerung von initialen Modellen & Generierung von Ligandenbibliotheken	69
3.5.10	Strukturvisualisierung und Analyse von Proteinstrukturen	70
3.6	Methoden zur Funktionsuntersuchung von Adhäsinen <i>in vivo</i>	70
3.6.1	Immunfluoreszenz-Mikroskopie	70
3.6.2	Bindung von Adhäsin-präsentierenden Zellen an Chitin-Beads	72
3.6.3	Fluoreszenzmikroskopie mit AlexaFluor 488 markiertem Protein	72
3.6.4	Interaktionsuntersuchung zwischen Hefen und Schaben	73
4	Ergebnisse	75
4.1	Phylogenetische Analyse der PA14-Domänen enthaltenden GPI-CWP's	76
4.2	Adhäsin aus <i>P. pastoris</i> – Charakterisierung eines neuen Subtyps	80
4.2.1	Expression, Isolation und Aufreinigung von Cea1A	80
4.2.2	Untersuchungen zur Glycanbindung von Cea1A	80
4.2.3	Kristallisation von Cea1A und Strukturlösung mittels molekularen Ersatzes	84
4.2.4	Die Struktur von Cea1A	87
4.2.5	Die Cea1A-vermittelte Chitin-Bindung von <i>S. cerevisiae</i>	92
4.2.6	Produktion, Isolation und Aufreinigung von PpFlo1A	94
4.2.7	Studien zur Funktion von PpFlo1A	95
4.2.8	Die Struktur von PpFlo1A	96
4.2.9	Untersuchungen zur Vergesellschaftung von Hefen und qPCR	101
4.3	Epitheliale Adhäsin aus <i>C. glabrata</i> – Struktur und Funktion	104
4.3.1	Die Kokristallstruktur von Epa1A	104
4.3.2	Lösung weiterer Komplexstrukturen von Epa1A	107
4.3.3	Die Struktur und Bindungseigenschaften von Epa6A	112
4.3.4	Die Struktur und Komplexstrukturen von Epa9A	117
4.4	Strukturelle Untersuchungen bezüglich Flokkuline – Flo10A	120
4.4.1	Strukturlösung des Flo10A-Wildtyps	120
5	Diskussion	124
5.1	Globale Betrachtung der pilzlichen PA14-Domänen	125
5.1.1	Interne Gemeinsamkeiten definiert die einzelnen Subtypen	125
5.1.2	Diversität innerhalb der Subtypen wird durch Anpassung erzeugt	127
5.2	Adhäsin aus <i>Pichia pastoris</i> bilden einen eigenständigen Subtyp	129

5.2.1	Der Domänenaufbau von Cea1A	130
5.2.2	Cea1A zeigt strukturelle Alleinstellungsmerkmale	131
5.2.3	Energetische Betrachtung des DcisD-Motivs	134
5.2.4	Der Einfluss der hypervariablen Bereiche auf die Ligandenbindung	136
5.2.5	Die biologische Funktion von Cea1	138
5.2.6	Potenzielle Anwendung von Cea1 als Reporter in der Medizin	140
5.2.7	<i>PpFlo1A</i> besitzt eine PA14-Domäne mit unbekannter Funktion	141
5.2.8	<i>PpFlo1A</i> besitzt möglicherweise keinen GPI-Anker	142
5.2.9	Der Übergang zur B-Region von Cea1A und <i>PpFlo1A</i> im Vergleich	142
5.3	Die epithelialen Adhäsine aus <i>C. glabrata</i>	145
5.3.1	Phylogenetische Nähe impliziert keine funktionelle Ähnlichkeit	145
5.3.2	Die Struktur/Funktions-Beziehung der Epas ist sehr komplex	147
5.4	Die Flokkulin-spezifische Subdomäne	148
5.5	Ausblick	150
6	Literaturverzeichnis	153
7	Anhang	179
7.1	Alignment der phylogenetischen Analyse	179
7.2	Kristallographische Statistiken	186
7.2.1	Datensammlungs- und Verfeinerungsstatistiken - Epa6A-Komplexe	186
7.2.1	Datensammlungs- und Verfeinerungsstatistiken - Epa9A-Komplexe	187
7.3	Produzierte und analysierte Proteine	188
7.3.1	Cea1A	188
7.3.2	<i>PpFlo1A</i>	188
7.3.3	Epa1A	188
7.3.4	Epa6A	189
7.3.5	Epa9A	189
7.3.6	Flo10A	189
7.4	Abkürzungen	190
8	Danksagung	191

Zusammenfassung

Hefen sind in der Lage sich an unterschiedlichste Umweltbedingungen anzupassen und dadurch eine große Anzahl von Habitaten zu besiedeln. Die erste Kontaktfläche mit ihrer Umwelt bilden Hefen mit ihrer einzigartigen Zellwand und den darin integrierten Zellwandproteinen. Ein großer Teil dieser Proteine, darunter die GPI-verankerten Zellwandproteine (GPI-CWP), dient zur Zell/Zell- oder Zell/Substrat-Adhäsion und spielt damit eine essentielle Rolle in der Lebensweise der Hefen. Einige dieser Proteine sind in der Lage Glycanstrukturen zu binden und zeigen einen gemeinsamen Domänenaufbau. Einer C-Typ-Lektin-artigen A-Domäne folgt eine mittlere, hoch repetitive und O-glycosylierte B-Region. Die Fixierung dieser Zellwandproteine gelingt durch die Transglycosylierung eines in der C-Domäne befindlichen Glycosylphosphatidylinositolankers. Die Glycanbindedomäne basiert strukturell auf der PA14-Domäne des protektiven Antigens von *Bacillus anthracis*. Sie enthält ein ungewöhnliches Motiv aus zwei konsekutiven, mit einer *cis*-Peptidbindung verbundenen Aspartaten, welche ein Ca^{2+} -Ion komplexieren. Zusammen mit Aminosäureresten aus verschiedenen Schleifenbereichen wird hiermit die Affinität und Spezifität der Glycanbindung erzeugt.

In dieser Arbeit gelang es, mittels einer sorgfältigen bioinformatischen und phylogenetischen Studie, neue Subtypen mit konservierten Eigenschaften dieser Adhäsinfamilie in Ascomyceten zu identifizieren. Cea1A, ein charakteristisches Mitglied des *Pichia*-Subtyps, konnte hierbei funktionell sowie strukturell charakterisiert werden. Dieses Adhäsion bindet an N-Acetylglucosamin und vermittelt damit die spezifische und hochaffine Erkennung nicht-reduzierender Enden nicht-kristalliner chitinöser Polymere. Die Struktur von Cea1A zeigt, neben einer neuartigen, stielartigen Subdomäne, einen eigenständigen Bindungsmodus, welcher eine kompakte Bindungstasche und ladungsgetriebene Interaktionen verbindet. Cea1A könnte hiermit die molekulare Grundlage der Vergesellschaftung von Hefen mit Insekten darstellen. PpFlo1A, ein weiteres Adhäsion dieser neuartigen Gruppe, zeigt strukturelle Ähnlichkeiten und dient möglicherweise zur Erkennung einer Subpopulation von *Pichia*-Hefen.

Im Rahmen dieser Arbeit konnten weiterhin drei epitheliale Adhäsine aus dem Humanpathogen *Candida glabrata*, Epa1A, Epa6A und Epa9A, strukturell und funktionell weitergehend untersucht werden. Dieser Subtyp der pilzlichen Adhäsine dient, unter Bindung von Galactosiden auf Epithelzellen, zur Etablierung der Pathogenität von *C. glabrata*. Durch die Lösung von Komplexstrukturen konnte gezeigt werden, dass eine funktionelle Klassifizierung nicht zwingend mit phylogenetischer Verwandtschaft einhergeht.

Mit der strukturellen Untersuchung von Flo10A, einem Adhäsion des Flokkulin-Subtyps, welcher zur vegetativen Zellaggregation in *Saccharomyces cerevisiae* dient, konnte schließlich gezeigt werden, dass die Insertion von Subdomänen im Bereich der flexiblen Loops einen direkten Einfluss auf die Konformation der Bindungstasche und damit die Bindungseigenschaften hat.



Summary

Yeasts are able to adapt to a broad variety of different ecological conditions and to therefore colonize a vast number of habitats. The first interface for interaction with the surrounding medium is the unique fungal cell wall with its embedded integral cell wall proteins. A huge number of these proteins serve in cell/cell- and cell/substrate-adhesion. They play a crucial role in many processes of fungal lifestyles. An important family of fungal adhesins uses a common domain architecture for binding of glycan structures. The A-domain, which works in a C-type-lectin like manner, is followed by a repetitive and highly *O*-glycosylated intermediate B-region. The fixation to the cell wall is provided by transglycosylation of a glycosylphosphatidylinositol-anchor (GPI-anchor), which is part of the terminal C-domain. The glycan recognizing A-domain is structurally related to the PA14 domain of the protective antigen of *Bacillus anthracis*. It harbours a unique motif of two consecutive aspartate-residues connected via a *cis*-peptide bond, which coordinate a Ca^{2+} -ion. In combination with different loop-regions the affinity and specificity for specific glycans is generated.

Using a global bioinformatical and phylogenetic analysis, new subtypes of these adhesins with conserved features could be identified in several ascomycetes. Cea1A, a member of a new subtype from *Pichia pastoris*, was investigated using X-ray crystallography and biophysical as well as *in vivo* studies. This adhesin binds to *N*-acetylglucosamine and mediates the highly specific and affine recognition of non-reducing ends of non-crystalline chitinous polymers. The structure of Cea1A shows, beside a unique stalk-shaped subdomain, a subtype-specific binding mode combining a compact binding pocket with ionic interactions. This adhesin may represent the molecular base of yeast/insect-symbiosis or pathogenesis. PpFlo1A is a further adhesin of this subtype, which is showing structural similarities, and which might be important in the recognition of *Pichia*-subpopulations.

Three epithelial adhesins, Epa1A, Epa6A and Epa9A were structurally and functionally further characterized. This adhesin-subtype is used by the pathogenic yeast *Candida glabrata* for adhesion to epithelial cells via their recognition of surface galactosides. Different complex structures of these adhesins revealed that a functional classification is not strictly linked to phylogenetic dependencies.

Flo10A belongs to the flocculin-subtype, which is applied by the budding yeast *Saccharomyces cerevisiae* in vegetative aggregation and floc formation. Structural investigations using this flocculin showed that the insertion of subdomains in the region of the flexible loops has a major effect on the conformation of the binding pocket and ligand recognition.

The connection of intra-subtype investigations with a global analysis led to the picture of a complex yet highly adaptive kind of adhesion-proteins with conserved features.



1 Einleitung

1.1 Hefen – Einzellige Arten des Reiches Fungi

Hefen sind meist einzellige Mikroorganismen des Reichs Fungi aus der Domäne der Eukaryoten. Sie stellen *per se* keine eigene taxonomische Einheit dar; die hierunter fallenden Mikroben teilen sich auf die beiden Phyla Ascomycota (Schlauchpilze) und Basidiomycota (Ständerpilze) des Unterreiches der Dikarya auf. Pilze und ebenfalls Hefen sind erdgeschichtlich sehr alte Lebewesen; erste potentielle Fossilien stammen bereits aus der Zeit des Präkambriums¹. Diese aquatischen Pilze haben vermutlich im Kambrium, bereits vergesellschaftet in Form von Mykorrhiza, das Land besiedelt^{2,3}.

Heute sind Pilze und Hefen ubiquitär zu finden, sie kommen häufig sogar mit extremen Bedingungen, wie Trockenheit (*Hortaea werneckii*)⁴ oder radioaktiver Strahlung (*Cryptococcus neoformans*)⁵ zurecht. Neben Bakterien sind sie hauptverantwortlich für Recycling und Kompostierung von biologischem Material. Sie sind chemoorganotroph und nutzen organische Moleküle, meist Hexosen wie Glucose und Fructose aber auch komplexere Oligosaccharide, organische Säuren und Alkohole, zur Energiegewinnung⁶. Damit sind Hefen und Pilze in der Lage, ansonsten meist unverdaubare biologische Makromoleküle zu verwerten und erneut dem Ökosystem zugänglich zu machen. Im Gegensatz zu Bakterien sind die Diversität und der Einfluss von Hefen auf ihr umgebendes Ökosystem nur begrenzt bekannt⁷. Eine Studie von HERRERA *et al.* kam bspw. zu dem Schluss, dass einige Hefen durch Metabolisierung von Zuckern aus Nektar die Mikroumgebung in der Blume um bis zu 6 °C erwärmen können⁸. Häufig sind Hefen in Vergesellschaftung zu finden. Sie bilden mit Algen Flechten oder sie kommen als Symbionten und Pathogene von Insekten, Pflanzen und Tieren vor⁹⁻¹². Obwohl sich Hefen sexuell reproduzieren können, findet man häufig die asexuelle Reproduktion mittels Knospung. Hierbei wird eine neue Zelle direkt durch die Mutterzelle gebildet und das duplizierte Genom in die Tochterzelle überführt. Nach weiterem Wachstum folgt die Abspaltung der Tochterzelle.¹³

Die Kombination all dieser Eigenschaften liefert das Bild einer vielfältigen Art von eukaryotischen Einzellern. Aufgrund dieser Vielfalt fanden Hefen bereits früh in der Menschheitsgeschichte Anwendung und finden diese auch heute noch, bspw. beim Brauen von Bier, der Herstellung biotechnologischer Produkte oder der Gewinnung von Wirkstoffen wie Antibiotika¹⁴⁻¹⁶

Obwohl zurzeit bereits über 1500 Hefespezies bekannt sind, wird davon ausgegangen, dass dies nur etwa 1% der zu erwartenden Arten darstellt^{17,18}.

1.1.1 *Saccharomyces cerevisiae* – Ein vielfältiger Modelorganismus

Saccharomyces cerevisiae ist eine Hefe aus der Familie Saccharomycetaceae und der bekannteste Vertreter dieser eukaryotischen Einzeller. Diese, auch als Bäckerhefe bekannte Art, wird bereits seit der Urgeschichte vom Menschen zur Produktion von Brot und Bier mittels Fermentation und Gärung genutzt¹⁴. Dies ist möglich, da Hefe als fakultativ anaerob gilt. Hierbei werden als primäre Nahrungsquelle Zucker wie Glucose, Fructose oder Maltose genutzt. Die hauptsächlich entstehenden Produkte bei deren Metabolisierung sind CO₂ und Ethanol, welches nach Verbrauch des Zuckers unter Oxidation wiederum als Nahrungsquelle dienen kann. Diese Ausgangsstoffe sind in der Natur in reifen Früchte vorhanden, welche daher das natürliche Habitat von *S. cerevisiae* darstellen; die Domestizierung erfolgte vermutlich zufällig¹⁹. Hefen nutzen Wespen als Vektoren, um sich neue Habitate zu erschließen und zu überwintern²⁰. Die einfache Zugänglichkeit und Kultivierbarkeit veranlasste den Menschen zum detaillierten Studium von *S. cerevisiae*, welcher als eukaryotischer Modelorganismus dient. Dies führte zur frühen Entwicklung molekularbiologischer Methoden, zur ersten Sequenzierung eines eukaryotischen Genoms sowie der Etablierung der heterologen Produktion von Proteinen und Metaboliten in der Biotechnologie^{21,22}. Ein weiterer Schritt war die kürzlich erfolgte Vollsynthese eines der 16 Chromosomen²³.

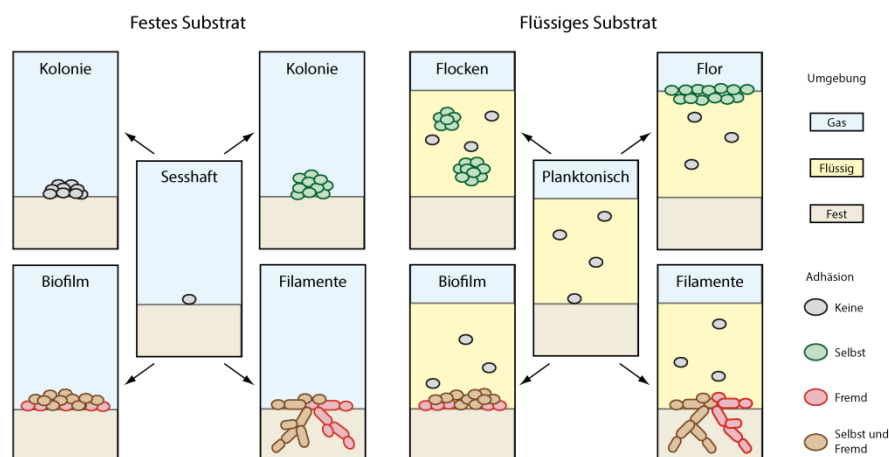


Abbildung 1.1 – Wachstumsformen von *S. cerevisiae*. In Abhängigkeit des verwendeten Mediums ist *S. cerevisiae* in der Lage unterschiedliche Wachstumsformen einzunehmen. Hierbei ist ebenfalls Zell/Zell- oder Zell/Substrat-Erkennung zu unterscheiden. Die Eigenschaften des umgebenden Mediums sowie der einzelnen Zellen sind farbig hervorgehoben. Mit freundlicher Genehmigung von STEFAN BRÜCKNER²⁴.

Der hohe Konservierungsgrad der grundlegenden eukaryotischen Zellmerkmale ermöglicht die Erforschung der zugrundeliegenden molekularen Prinzipien. Darüber hinaus können Homologe von medizinisch relevanten Genen in Hefe gut untersucht werden^{25,26}. Hefen, darunter *S. cerevisiae*, sind in der Lage, durch Nutzung unterschiedlicher Wachstumsformen (pseudohyphales- oder adhäsives Wachstum) und damit eingehendes Zell/Zell- oder Zell/Substrat-Adhäsion, neue Nährstoffquellen zu erschließen und sich ihrer Umwelt anzupassen^{27,28}. Hierbei bilden sie in flüssigem Medium und in Abhängigkeit der chemischen Umgebung bspw. Flokken²⁹ oder einen schwimmenden Flor³⁰ aus (Abbildung 1.1). Auf

Festmedium können ebenfalls Zellverbände in Form von Biofilmen³¹ oder Pseudohyphen²⁷ erzeugt werden.

1.1.2 *Candida glabrata* - Hefen als Pathogene

Da Hefen unter einer Vielzahl von Umweltbedingungen leben können, sind sie ebenfalls in der Lage bspw. die menschliche Haut oder auch den Verdauungstrakt von Säugetieren zu besiedeln. Infektionen kommen hierbei selten vor; sie sind eher als Kommensalen zu betrachten³². Als opportunistische Pathogene bei immunkompromittierten Menschen allerdings spielen Hefen des Phylums *Candida* in den letzten Jahren eine immer größere Rolle³³. Sie stellen hier nicht nur die dritthäufigste Ursache für systemische Erkrankungsbilder auf Intensivstationen dar, sondern bilden generell die Ursache von bis zu 10% der nosokomialen Infektionen. Kommt es zu einer systemischen Infektion beträgt die Mortalitätsrate 47%. Hauptsächlich sind für diese Krankheitsbilder die beiden Spezies *Candida albicans* und *Candida glabrata* verantwortlich. Beide gehören, wie die Bäckerhefe, zur Gattung Saccharomycetaceae. Interessanterweise zeigt *C. glabrata* hierbei auf Genomebene und im Phänotyp eine größere Ähnlichkeit zu *S. cerevisiae* als zu *C. albicans*³⁴. Hauptsächlich wird dies durch eine große Ähnlichkeit bei orthologen Genen (65%) sowie einem vergleichbare niedrigen GC-Gehalt (38%) deutlich³⁵.

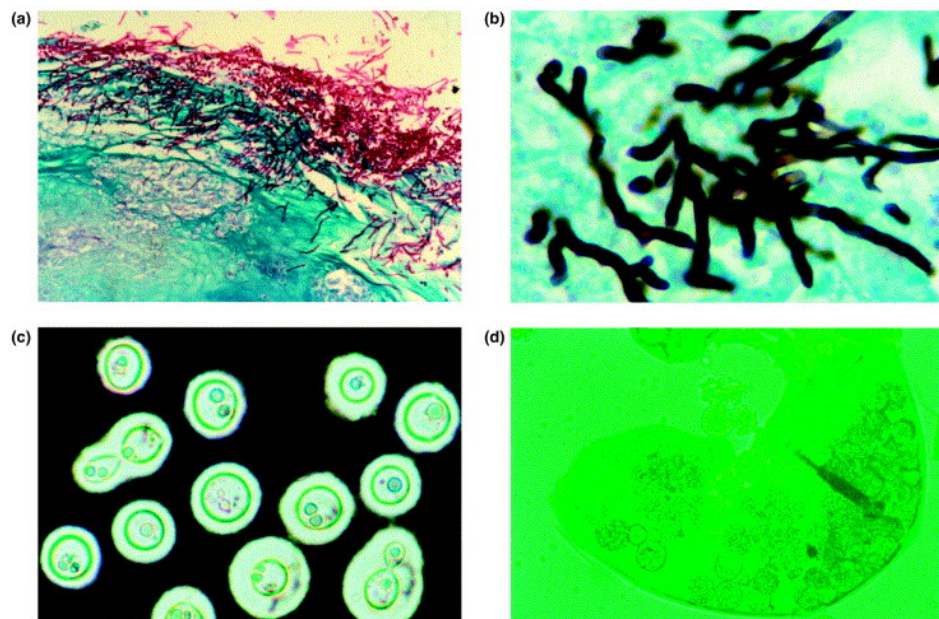


Abbildung 1.2 – Morphogenese pilzlicher Pathogene. (A) Hyphen von *C. albicans* bei der Penetration von Epithelzellen. (B) Hyphales Wachstum von *Aspergillus fumigatus* in Lungengewebe. (C) *Cryptococcus neoformans* als gekapselte Hefe aus der Abteilung der Basidiomycota. (D) Ruptur einer Sphäre von *Coccidioides immitis* und Freisetzung einzelliger Individuen als pathogene Spezies³⁶.

Dadurch, dass *C. glabrata* kommensalistisch mit Säugetieren vergesellschaftet ist, ergeben sich jedoch auch etwaige Unterschiede zu *S. cerevisiae*. Aufgrund der Verfügbarkeit wichtiger Metabolite im Säugetiergewebe ist die Stoffwechsellkapazität von *C. glabrata* gegenüber

S. cerevisiae stark reduziert³⁴. In Kombination deutet dies darauf hin, dass beide Spezies des Phylums *Candida* unabhängig voneinander ihre Pathogenität entwickelt haben. Die verwandte Gattung *Nakaseomyces* stellt möglicherweise ein Bindeglied zwischen der mit Früchten und Insekten verbundenen Lebensweise und der Entwicklung der Pathogenität dar³⁷. Unterschiedliche Wachstumsformen und Adhäsion spielen bei der Pathogenität ebenfalls eine große Rolle (Abbildung 1.2); so korreliert die Fähigkeit zur Ausbildung von Filamenten oft mit der Qualität der Pathogenität^{38,39}.

Auch ohne direkt virulent zu agieren, können immunologische Sensibilisierungen gegenüber Hefen auftreten. So findet man bei bis zu 70% der *Morbus Crohn*-Patienten Antikörper gegen *S. cerevisiae*, während diese nur bei etwa 8% der gesunden Bevölkerung auftreten⁴⁰. Für die Krankheit selbst ist die Hefe allerdings nicht direkt verantwortlich zu machen⁴¹.

1.1.3 *Pichia pastoris* – Eine methylothrophe Hefe

Die Diversität der Hefen spiegelt sich auch in den von ihnen verwertbaren Nahrungsquellen wieder. Während *S. cerevisiae* Ethanol metabolisieren kann, ist die ebenfalls zu den Saccharomycetaceae gehörende Hefe *Pichia pastoris* in der Lage, Methanol umzusetzen⁴² (Abbildung 1.3). Diese sogenannte Methylothrophie, also die Verwendung von Methanol als einzige Energiequelle, ist eine biotechnologisch nutzbare Eigenschaft. Methanol ist im Vergleich zu anderen Nährmedien sehr kostengünstig zu erhalten; die Herstellung erfolgt hauptsächlich katalytisch aus CO und H₂⁴³. Dies ist auch der Grund, warum zu Beginn der 70er Jahre des 20. Jahrhunderts versucht wurde, *Pichia pastoris* als industrielles Nahrungsmittel für die Tierhaltung zu etablieren^{44,45}. Neben der günstigen Nahrungsquelle lässt sich *P. pastoris* zu hohen Zelldichten kultivieren. Aufgrund des steigenden Ölpreises 1973 verteuerte sich allerdings ebenfalls die Methanolproduktion, womit diese Idee verworfen wurde.

Die meisten methylothropen Hefen sind phylogenetisch einem großen Cluster zuzuordnen; mit *P. pastoris* als Ausnahme⁴⁶. Diese ist anhand der Studien von YAMADA *et al.* eher einer eigenen Subgruppe zuzuordnen, welche den Namen *Komagataella* trägt⁴⁷. Da diese Neueinteilung nicht final geregelt ist⁴⁸, wird im Rahmen dieser Arbeit die allgemein bekanntere Bezeichnung *Pichia* beibehalten. Der Referenzstamm DSMZ70382, welcher als erstes Isolat gilt, wurde in den Exsudaten von Rosskastanien und Eichen gefunden. Der Stamm GS115, welcher aus dem Wildtypisolat NRRL Y-11430 u.a. durch Einführung einer Histidin-Auxotrophie generiert wurde, dient hauptsächlich der heterologen Produktion von Proteinen in der Biotechnologie und der Wissenschaft⁴⁹. Die Genome beider Stämme wurden im Jahr 2009 sequenziert^{50,51}. Ein dritter Stamm, CBS7435, folgte im Jahr 2011⁵². Alle drei Genome sind mit Hilfe eines Suchprogramms (<http://www.pichiagenome.org/>) zugänglich⁵³.

Die oben erwähnte Erstbeschreibung durch GUILLIERMOND im Jahr 1919 gelang durch Untersuchung von Baumexsudaten⁵⁴. Bei Exsudaten handelt es sich um über die Pflanzenoberfläche abgesonderte Ausscheidungen, welche vorwiegend aus Polysacchariden, aber auch aus

organischen Säuren, Aminosäuren und Phenolen in Wasser bestehen⁵⁵. Dieser viskose Schleim (*mucilage*) wird bspw. von vielen Mikroorganismen oder Invertebraten als Nahrungsquelle genutzt. Die sich hierbei ansiedelnden Mikroorganismen können dabei in Wechselwirkung mit der absondernden Pflanze treten⁵⁶.

Diese Exsudate stellen ein Habitat der Fruchtfliege *Drosophila melanogaster* dar. Diese ernährt sich von den hier lebenden Mikroorganismen und ist der Lage mit diesen in Wechselwirkung zu treten^{57,58}. Darüber hinaus ist die Vergesellschaftung von *Pichia*-Spezies mit weiteren Insekten bekannt, jedoch nicht genau studiert⁵⁹⁻⁶².

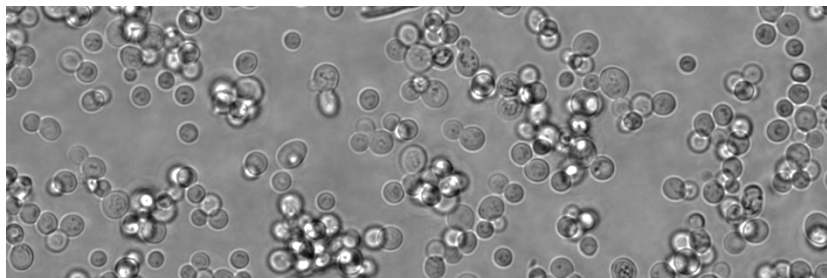


Abbildung 1.3 – Mikroskopische Aufnahme von Zellen des *Pichia pastoris* Stamms GS115. Vergrößerung 40x.

Biotechnologisch ist *Pichia pastoris*, vor allem der Stamm GS115, ein weit verbreiteter Organismus zur heterologen Proteinproduktion. Gerade die Fähigkeit der posttranslationalen Modifikation durch Glycosylierung ist ein wichtiges Kriterium für die Untersuchung und Herstellung vieler eukaryotischer Proteine⁶³. Die Anpassung des Glycosylierungsmusters an humane Verhältnisse ermöglicht zudem die Produktion von „humanisierten“ Proteinen zur medizinischen Anwendung⁶⁴⁻⁶⁶. Die *Pichia*-eigene Glycosylierung kann aufgrund der bei Hefen häufig vorkommenden Hypermannosylierung zu einer Immunantwort führen⁶⁷. Zur Proteinproduktion macht man sich ebenfalls die Methylophorie dieser Hefeart zu Nutze. Die zwei Alkoholoxidasen AOX1 und AOX2 ermöglichen die Verwertung von Methanol im Gegensatz zu Ethanol. Die Promotoren beider Gene sind Methanol-induzierbar und mit Glucose reprimierbar. Dementsprechend kann die Expression eines Zielgens, welches unter Kontrolle des AOX-Promotors steht, durch Zugabe von Methanol induziert werden^{68,69}.

Die Kenntnis des Genoms und die breite Anwendung von *Pichia pastoris* führten jedoch, im Gegensatz zu *S. cerevisiae*, nicht zur Untersuchung der weiteren Biologie und Ökologie dieser Hefe. Diese Aspekte sind so gut wie unbekannt und bieten Raum für weitere Forschung.

1.2 Die Hefezellwand als physische Barriere und Interaktionsfläche mit der Umwelt

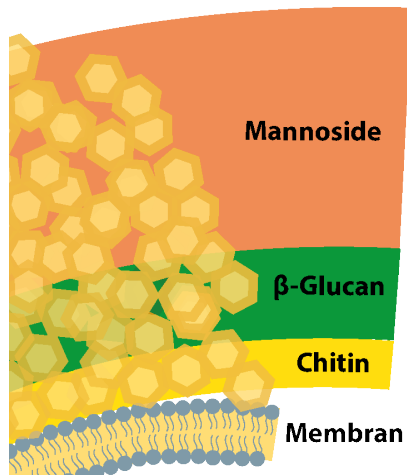


Abbildung 1.4 – Schematischer Aufbau der pilzlichen Zellwand. Der Lipiddoppelschicht als Zellmembran folgt eine dünne Schicht aus polymerem *N*-Acetylglucosamin, Chitin. Hierdurch und durch die folgende Schicht aus β -Glucanen erhält die Zellwand ihre Rigidität. Diese Polysaccharide sind stark quervernetzt und bilden ein dichtes Netzwerk. Eine dicke, bürstenförmige Schicht aus Mannose, welche von mannosylierten Zellwandproteinen herrührt, bildet den Glycokalyx-artigen Abschluss.

Hefen besitzen eine einzigartige Zellwand, welche etwa 15-30% ihrer Trockenmasse ausmacht und der einzelnen Zelle Schutz vor äußeren physikalischen und osmotischen Einflüssen bietet⁷⁰. Bis zu 25% der Gene von *S. cerevisiae* stehen mit der Synthese und der Integrität der Zellwand in Verbindung, was die außerordentliche Bedeutsamkeit dieser Struktur unterstreicht⁷¹. Die mechanische Stabilität wird hauptsächlich durch ihren schichtweisen Aufbau erzeugt, welcher im Mittel eine Dicke von 70-100 nm aufweist^{72,73}. Die äußere Schicht, welche eine bürstenartige Oberfläche zeigt, ist mit Proteasen entfernbar^{73,74}. Dies deutet bereits den hohen Gehalt an hoch-mannosylierten Proteinen (Mannoproteinen) an. Der nachfolgende Teil kann hingegen nur mit β -Glucanasen entfernt werden, er besteht aus Oligo-/Polysacchariden⁷⁵. Die strukturelle und chemische Einteilung ist jedoch komplexer. Darüber hinaus variieren die Zusammensetzung und die Lokalisation der Zellwandkomponenten mit äußeren Einflüssen, dem Zellzyklus und Nahrungsbedingungen⁷¹. Eine Einteilung unter chemischen Gesichtspunkten liefert einen etwa 3-lagigen Aufbau (Abbildung 1.4).

1.2.1 Chitin - Ein weit verbreitetes Biopolymer vermittelt Rigidität

Die erste Schicht, welche der Lipiddoppelschicht der Zellmembran folgt, besteht hauptsächlich aus Chitin, einem linearen Polymer aus *N*-Acetylglucosamin (GlcNAc)⁷⁶. Der Anteil an der Gesamttrockenmasse der Zellwand beträgt nur etwa 1-2%⁷⁷. In Hefen, sowie bei den meisten Arthropoden, kommt Chitin in der α -Konformation vor, welche eine α -helikale Spiral-/Fibrillenform darstellt⁷⁸. β -Chitin ist viel seltener zu finden; so bspw. in den monokristallinen Stacheln der Kieselalge *Thalassiosira fluviatilis*⁷⁹. Da Chitin eher starr ist, dient es zur Ausbildung der nötigen Rigidität der pilzlichen Zellwand. Nur in Ausnahmen ist auch nicht-kristallines Chitin in Hefen zu finden⁸⁰. Die durchschnittliche Kettenlänge beträgt bei *S. cerevisiae* etwa 190 GlcNAc-Einheiten⁸¹. Chitin tritt hauptsächlich in Septa bei der Knospung, an Sprossnarben sowie in geringerem Umfang in der lateralen Zellwand selbst auf⁷¹. α - und β -Chitin unterscheiden sich anhand der zugrunde liegenden Kristallpackung und den damit einhergehenden internen Wechselwirkungen (Abbildung 1.5).

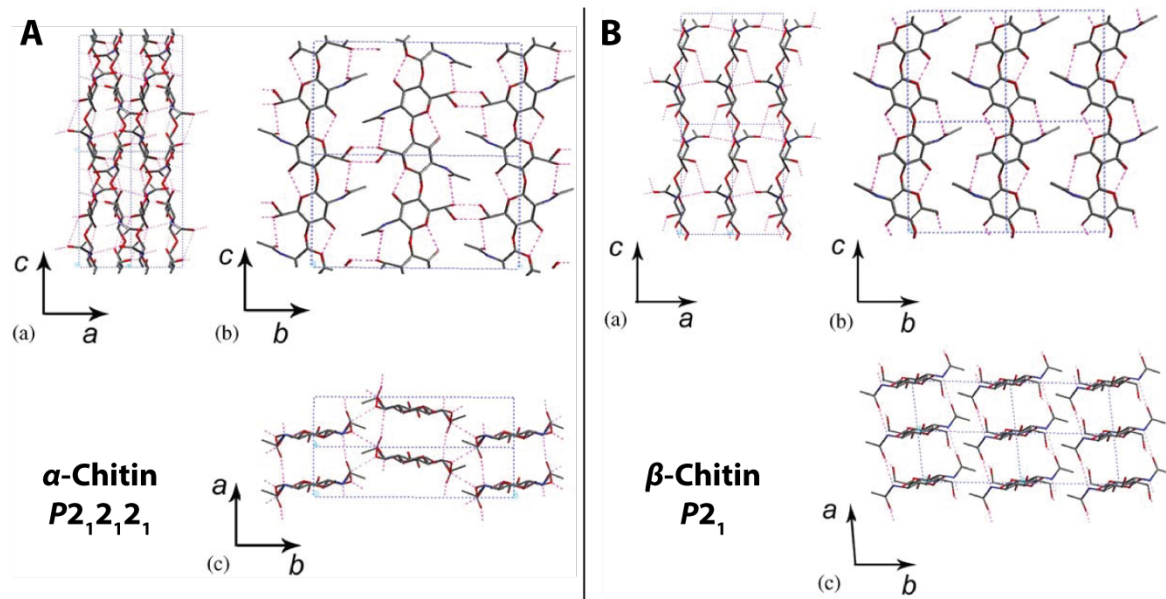


Abbildung 1.5 – α - und β -Chitin unterscheiden sich in ihrer Kristallpackung und den zugrunde liegenden Wechselwirkungen. A: α -Chitin ist die häufigste Variante dieses Biopolymers und zeichnet sich durch Kristallisation in der Raumgruppe $P2_12_12_1$ sowie durch das Vorhandensein zweier Konformationen der $-\text{CH}_2\text{OH}$ Gruppen aus. B: Das seltenere β -Chitin hingegen kristallisiert in der Raumgruppe $P2_1$ und zeigt eine homogene Orientierung der funktionellen Gruppen (Nach RINAUDO)⁷⁸.

In beiden Strukturen sind die Chitinketten in Strängen angeordnet, welche interne C-O...NH Wasserstoffbrücken ausbilden. Die Ketten nehmen dadurch einen Abstand von etwa 0.47 nm ein⁷⁸. Beide Formen kristallisieren in verschiedenen Raumgruppen; $P2_12_12_1$ für α -Chitin und $P2_1$ für β -Chitin. Im Vergleich zeigt ersteres zusätzliche Wasserstoffbrückenbindungen zwischen den einzelnen Strängen in b -Richtung des Gitters. Diese fehlen bei β -Chitin und verursachen vermutlich die Quellung dieses Materials in Verbindung mit Wasser⁸². Die Konversion von β -Chitin zu α -Chitin ist irreversibel, was die größere thermodynamische Stabilität und das häufigere Auftreten letzterer Konformation erklärt⁸³.

1.2.2 Vernetzte β -Glucane sorgen für strukturelle Integrität

Auf Chitin folgt eine stark quervernetzte Schicht aus β 1,3 und β 1,6-Glucanen, wobei letzteres eine höhere Verknüpfungsrates zeigt. β -Glucane stellen 30-60% der Trockenmasse der Zellwand⁸⁴. In lateralen Zellwandbereichen kann, in geringem Umfang, auch die glycosidische Verbindung der Chitinketten mit beiden β -Glucanen beobachtet werden^{85,86}. Der große Anteil an der Gesamtmasse der Zellwand und die Vernetzung tragen zur mechanischen Stabilität der Zellwand bei⁸⁷. Hierdurch wird bspw. ebenfalls deren Flexibilität unter osmotischem Stress erzeugt^{30,88}. Ketten von β 1,3-Glucanen sind moderat vernetzt und zeigen eine Länge von bis zu etwa 1500 Glucose-Resten⁸⁹.

1.2.3 Mannoproteine sind die Basis für wichtige zelluläre Funktionen

Als letztes ist eine für Makromoleküle fast unpassierbare Mannoseschicht vorhanden⁷⁴. Diese rührt von der starken O -Mannosylierung (Man) Serin-/Threonin-reicher *Repeat*-Sequenzen

der in der Zellwand verankerten Proteine her. Die Kettenlänge variiert, zeigt im Mittel aber Längen von 50 und mehr Mannosiden vor einer Verzweigung in α -1,2-Man-Bäume⁷¹. Es wird davon ausgegangen, dass die *O*-Mannosylierung essentiell für die Funktion einiger Zellwandproteine ist oder eine Rolle bei der Zell-/Zell-Kommunikation spielt.⁹⁰ Von diesen Proteinen existieren zwei Gruppen. Die erste Gruppe ist mit Hilfe eines Glycosylphosphatidylinositol-Ankers (GPI-CWP, *glycosylphosphatidylinositol anchored cell wall proteins*) fixiert. Hierbei wird der Anker prozessiert, indem eine Trennung zwischen Glucosamin (GlcN) und Mannose stattfindet⁹¹. Über letzteren Rest wird das Protein an das nicht-reduzierende Ende eines β 1,6-Glucanen transglycosyliert^{86,92}. Die zweite Gruppe innerhalb der Mannoproteine besitzt keinen solchen Anker. Diese PIR-Proteine (*proteins with internal repeats*) sind über Aminosäuren der internen *Repeats* direkt mit β 1,3-Glucanen verestert⁹³. Zur Sekretion besitzen die zellwandverankerten Proteine eine *N*-terminale Signalsequenz⁹⁴. Mit der Zellwand assoziierte Proteine lassen sich funktionell in drei Gruppen einteilen. Proteine der ersten Gruppe sind an Konstruktion und Remodellierung der Zellwand beteiligt, wie Glucosidasen und Hydrolasen. Bei der zweiten Gruppe handelt es sich um Mechanosensoren, welche membranständig sind und in den extrazellulären Bereich hinaus ragen. In die dritte Gruppe lassen sich nicht-enzymatische Proteine wie Adhäsine und Agglutinine einordnen, welche eine wichtige Rolle an der Grenzfläche Zelle/Zelle oder Zelle/Substrat spielen und die Basis für die Interaktion mit dem umgebenden Medium darstellen⁷¹. Adhäsine dieser Gruppe sind zentrales Objekt dieser Arbeit.

Während β -Glucane vornehmlich in der pflanzlichen Zellwand zu finden sind, wird Chitin hauptsächlich in Arthropoden gefunden. Die Kombination beider Materialien im Aufbau der Zellwand ist nur bei Pilzen zu finden⁹⁵. Dies erlaubt eine große Kombinationsmöglichkeit und Flexibilität. Der Aufbau der Zellwand aus Chitin, β -Glucanen und Mannoproteinen stellt hierbei die Regel dar, allerdings finden sich auch weitere Zucker, wie bspw. Phosphomannose⁹⁶. Eine Studie von KOPECKÁ konnte darüber hinaus zeigen, dass *in vitro* eine Mischung linearer β 1,3-Glucane, verzweigter β 1,6-Glucane, Oligomannosiden (Mannane) und Glycogen spontan zur Bildung Zellwand-ähnlicher fibrillärer und amorpher Strukturen führt⁹⁷.

Die Organisation und Synthese der Hefezellwand ist detailliert von ORLEAN und KLIS beschrieben^{71,81}

1.3 Pilzliche Zellwandproteine zur Adhäsion von Hefen

In der Gruppe der nicht-enzymatischen pilzlichen GPI-CWPs bilden Adhäsionsproteine eine wichtige Subgruppe mit vielfältigen Funktionen. Prinzipiell unterscheidet man Adhäsion zur geschlechtlichen Vermehrung und vegetativen Aggregation. Eine heterotypische Protein/Protein-Interaktion sichert hier bspw. die Paarung zweier Zellen unterschiedlichen Geschlechtes (α und α)⁹⁸. Für die nicht-sexuelle, also vegetative Adhäsion, nutzen Hefen ein

Spektrum unterschiedlicher Adhäsionsproteine⁹⁹. Diese zeigen einen ähnlichen Aufbau, wobei drei Domänen differenziert werden können¹⁰⁰. Dem Sekretionssignal folgt die A-Domäne, welche die adhäsive Funktion ausübt. Diese wird von der B-Region flankiert, welche stark *O*-glycosyliert ist und repetitive, Serin-/Threonin-reiche Einheiten enthält. Die C-Domäne dient der Fixierung in der Zellwand über Transglycosylierung ihres GPI-Ankers (Abbildung 1.6).

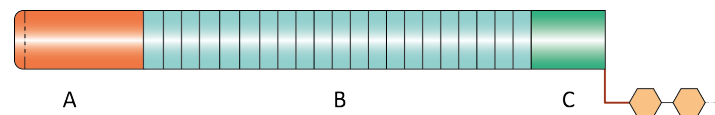


Abbildung 1.6 – Die pilzlichen Adhäsine haben den gleichen Domänenaufbau. Der *N*-terminalen Signalsequenz (gestrichelte Linie) folgt die funktionelle A-Domäne (rot). Es schließt sich eine hoch repetitive B-Region an, welche Serin-/Threonin-reich und stark *O*-glycosyliert ist (blau). Die Fixierung in der Zellwand gelingt über den GPI-Anker der C-Domäne (grün).

Die Funktion der repetitiven B-Region ist nicht abschließend geklärt. Einerseits wird davon ausgegangen, dass unter Bindung von Ca^{2+} -Ionen mit Hilfe der starken *O*-Glycosylierung an Serinen und Threoninen eine stäbchenförmige Konformation eintritt. Diese präsentiert die bindende A-Domäne außerhalb der Zellwand¹⁰¹. Andererseits nimmt man an, dass amyloidbildende Bereiche existieren welche zur Erhöhung der Avidität führt, indem mehrere Adhäsine gebündelt werden¹⁰²⁻¹⁰⁴. Hierdurch soll bspw. auch *cis*-Bindung, also die Bindung von oberflächenexponierten Liganden der eigenen Zelle unterdrückt werden. Das Vorhandensein der starken *O*-Glycosylierung scheint offenbar auch eine Rolle in der Funktion mancher Adhäsine zu spielen¹⁰⁵.

Die A-Domäne enthält in der Superfamilie der pilzlichen Adhäsine die aktive Gruppe, welche die Interaktion mit biotischen oder abiotischen Substraten und Matrices über verschiedene Mechanismen erlaubt^{100,106}. Gut untersucht sind die Adhäsine des Flokkulin-Subtyps aus *S. cerevisiae* oder die epithelialen Adhäsine (Epas) aus *C. glabrata*^{107,108}. Diese C-Typ-Lektin artigen Adhäsine, welche calciumabhängig an Saccharide binden, stehen im Gegensatz zu Als-Adhäsinen (*Agglutinin-like*) aus *C. albicans* oder Adhäsinen des Flo11-Typs aus *S. cerevisiae*^{109,110}. Erstere sind in der Lage mit Hilfe ihrer A-Domäne bestimmte Proteine der Zellwand des Wirts zu erkennen. Letztere verursachen eine homotypische Interaktion, welche zu vegetativer

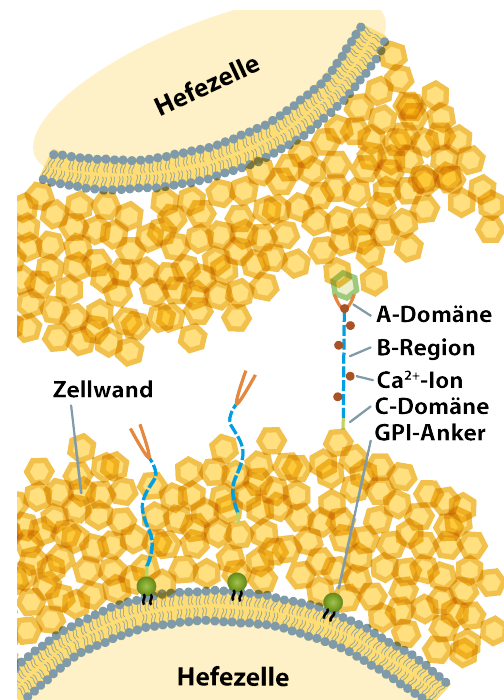


Abbildung 1.7 – Schematische Darstellung der Funktionsweise von Adhäsinen mit PA14-Domäne in Hefen am Beispiel der Zell/Zell-Interaktion. Von links nach rechts: Nach dem Transport durch die Zellmembran wird der GPI-Anker abgespalten und das Adhäsinn mittels Transglycosylierung in der Zellwand fixiert. Unter Bindung von Ca^{2+} -Ionen wird die A-Domäne aktiviert und die B-Region nimmt ihre stabförmige Konformation an.

Aggregation führt und zeigen eine Ig-artige Struktur¹¹⁰. Die lektinartige Wechselwirkung wird durch eine Domäne erzeugt, welche auf Struktur- und Sequenzebene Ähnlichkeit mit der PA14-Domäne des *protective antigens* (PA) von *Bacillus anthracis* aufweist¹¹¹. Hierbei führt, meist unter Bindung von Ca²⁺-Ionen, die Erkennung von Sacchariden in der Umgebung zur Adhäsion (Abbildung 1.7). Die subtelomere Lokalisation der pilzlichen Adhäsine macht diese Proteinfamilie sehr vielseitig, einerseits durch die ständige Entstehung von chimären Adhäsinen¹¹² durch Rekombination, andererseits durch die mit der genomischen Lage einhergehende epigenetische Regulation, bspw. durch *silencing*^{100,113,114}. Gerade die B-Region kann durch *slippage* der DNA-Polymerase unterschiedliche Längen annehmen, was Einfluss auf die Effizienz der Bindung hat¹⁰⁰. Somit kann bereits eine kleine Population von Hefen eine große Anzahl unterschiedlicher Adhäsine tragen und sich optimal an die ökologischen Gegebenheiten anpassen¹¹⁵.

1.3.1 Das protektive Antigen von *Bacillus anthracis* als strukturelle Basis

Die PA14-Domäne ist Bestandteil des Anthrax-Toxin Propeptids und wird während dessen weiterer Reifung proteolytisch abgespalten. Das Molekulargewicht dieses abgespaltenen Teils beträgt 14 kDa, was dem Namensbestandteil der Domäne entspricht. Proteine mit einer PA14-Domäne wurden bioinformatisch bisher in allen drei Domänen des Lebens nachgewiesen, darunter ebenfalls im Menschen. Zumeist ist die Domäne im Zusammenhang mit der Bindung und Prozessierung von Kohlenhydraten zu finden ohne zwingendermaßen in der direkten Bindung involviert zu sein^{116,117}. Die Struktur der PA14-Domäne selbst konnte im Jahr 1997 gelöst werden und zeigt ein dicht gepacktes β -Sandwich Motiv welches aus zwei antiparallelen Faltblättern, eines mit 4 und eines mit 6 Strängen, aufgebaut ist¹¹⁸.

Dieses Faltungsmotiv scheint sich evolutionär bewährt zu haben, da es nicht nur in einer großen Anzahl unterschiedlicher Spezies (720 Spezies nach *Pfam* Datenbank*, auch in *Homo sapiens*), sondern auch in Proteinen mit unterschiedlicher Funktion zu finden ist. Dazu zählen, neben wei-

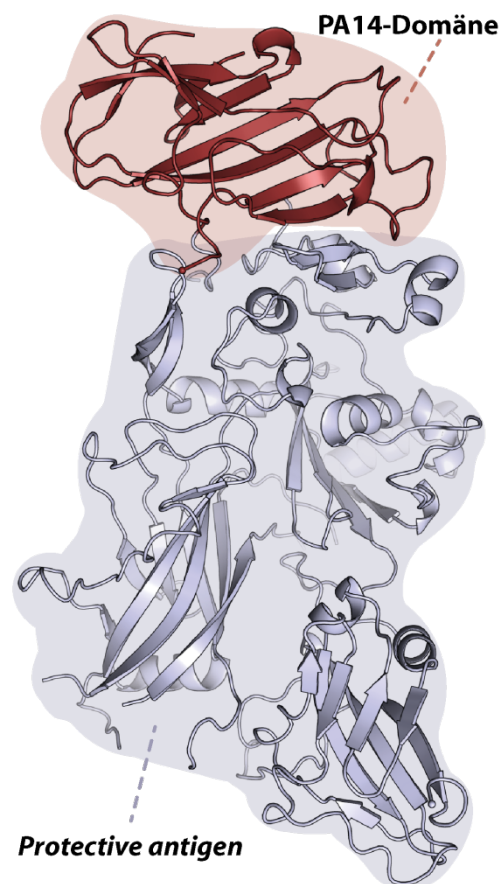


Abbildung 1.8 – Die Struktur des *protective antigens* aus *B. anthracis* (PDB-ID: 1ACC). Das größere, reife Antigen (blau) wird unter proteolytischer Abspaltung der PA14-Untereinheit (rot) erhalten. Diese stellt die strukturelle Grundlage für einige Saccharid-bindende pilzliche Adhäsine dar. Es besteht aus einem β -Sandwich welches, aus je einem 4 strängigem und einem 6 strängigem Faltblatt gebildet wird.

* Stand 12.04.2015

teren bakteriellen Toxinen, vor Allem β -Glucosidasen, Glycosidasen, Glycosyltransferasen oder Adhäsine¹¹⁹. In der letzteren Gruppe von Proteinen wird darüber hinaus zwischen der PA14-Domäne (*Pfam*-ID: PF07691) und der PA14_2-Domäne (*Pfam*-ID: PF10528), welche auch GLEYA-Domäne[†] genannt wird, unterschieden. Während erstere vorwiegend am *N*-Terminus von pilzlichen Adhäsinen gefunden wird, kann letztere, eng verwandte Domäne, am *C*-Terminus einer Familie von Proteinen mit putativer Agglutinations-Funktion bspw. in der Spaltheefe *Schizosaccharomyces pombe* gefunden werden¹²⁰. Darüber hinaus konnte in diesen Proteinen bisher kein GPI-Anker nachgewiesen werden. Die automatische Annotation der Datenbank *Pfam*¹²¹ unterscheidet jedoch aufgrund der großen Ähnlichkeit nicht zwischen beiden Gruppen, so dass das Flokkulin Flo5 bspw. beiden Gruppen zugeordnet wird.

In den bekannten pilzlichen Adhäsinen ist die PA14-Domäne in eine C-Typ-Lektin artige Bindung von terminalen Sacchariden involviert. Die PA14-Domäne aus *B. anthracis* selbst zeigt hingegen keine Bindung an Glycoside.

1.3.2 Flokkuline aus *S. cerevisiae*

Unter Flokkulation ist bei Hefen der Prozess der Agglutination und der damit einhergehenden Ausbildung von Flocken und nachfolgender Sedimentation zu verstehen^{122,123}. Dieser Phänotyp wird durch sog. Flokkuline verursacht. Hierbei handelt es sich um einen Subtyp pilzlicher Adhäsionsproteine welcher die vegetative Zelladhäsion über deren A-Domäne vermittelt¹²⁴⁻¹²⁶. Der Nutzen dieses Phänomens liegt vor Allem im Schutz der Zellen im inneren solcher Flocken. Diese sind vor äußeren Stressbedingungen geschützt¹²⁷. Flokkuline werden den sog. *Green Beard*-Genen zugeordnet. Diese, auf RICHARD DAWKINS zurückgehende Theorie, geht von einem altruistischen Prinzip aus¹²⁸. Auf die Bäckerhefe bezogen bedeutet dies: Zellen, welche mit einem entsprechenden Flokkulin ausgestattet sind aber sich im genetischen Hintergrund unterscheiden können, nehmen am Prozess der Flockenbildung teil und sind dadurch, so sie sich im inneren der Flocken befinden, vor äußeren Einflüssen geschützt^{127,129,130}. Zellen, welchen das Genprodukt fehlt, haben keinen Anteil an diesem Schutzmechanismus, da sie entweder nicht in einer Flocke, oder nur an deren äußersten Rand zu finden sind. Die Flockenbildung wird biotechnologisch und bei der Produktion von Nahrungsmitteln, vor allem Bier, ausgenutzt. Die hierdurch ausgelöste Sedimentation der Zellen, bspw. am Ende des Gärungsprozesses, vereinfacht die Abtrennung der Hefe¹³¹.

Bei den Flokkulinen aus *S. cerevisiae* und verwandten Arten wird zwischen zwei Mechanismen der Interaktion unterschieden^{106,110}. Neben der C-Typ-Lektin artigen Erkennung von Mannosiden der Hefezellwand durch die Flokkuline Flo1, Flo5, Flo9, Flo10 und Lg-Flo1 aus *S. pastorianus* ist die homotypische Interaktion von Flo11 Saccharid- und Calcium-unabhängig. Die Flo11-induzierte Flokkulation ist dementsprechend nicht mit Zuckerzugabe inhibierbar. In der ersten, Saccharid-bindenden Gruppe, lässt sich darüber hinaus weiter zwi-

[†] Abgeleitet von einem konservierten Sequenzmotiv GLEYA.

schen Flo1-Typ (Flo1, Flo5 und Flo9) und *NewFlo*-Typ (Lg-Flo1) differenzieren¹³². Die Interaktion des Flo1-Typs ist hierbei durch Mannose-Zugabe inhibierbar, die des *NewFlo*-Typs zusätzlich durch Maltose, Glucose und Saccharose^{113,132,133}. Die Bildung von Flocken wird bei beiden Typen durch Bindung von terminalen α -Mannosiden der pilzlichen Zellwand erreicht.

Untersuchungen von VEELDERS *et al.* konnten detaillierte Einblicke in die Struktur und Funktion der A-Domäne eines dominanten Flokkulins des calciumabhängigen Flo1-Typs, Flo5, liefern (PDB-ID: 2XJS)¹⁰⁷. Dessen Struktur dient als Modell für die Saccharidbindenden Adhäsine vom PA14-Typ. Es besteht, wie vorhergesagt, aus einer dem *protective antigen* ähnlichen β -Sandwich mit einzigartigen Eigenschaften (Abbildung 1.9). Das Hauptmerkmal der Bindungstasche ist ein Ca^{2+} -Ion, welches über ein Motiv aus zwei konsekutiven, über eine *cis*-Peptidbindung verknüpften Aspartaten komplexiert wird (*DcisD*-Motiv). Der diese Reste enthaltene Loopbereich wird daher *Calcium-binding Loop 1* genannt (CBL1). Anstelle dieses Motives ist im *protective antigen* von Anthrax ein *DtransN*-Motiv zu finden, welches kein Ca^{2+} -Ion koordinieren kann. Dies erklärt die fehlende Saccharidbindung. Ein zweiter, darunter liegender Loop (CBL2) bildet den unteren Teil der Bindungstasche und interagiert direkt mit einem gebundenen Mannoserest sowie über ein Asparagin mit dem Ca^{2+} -Ion. Die 4 Reste, welche in direktem Kontakt mit dem Liganden stehen, werden als Position I-IV bezeichnet. Neben dem auf Struktur- und Sequenzebene im Flokkulin-Subtyp konservierten Kernbereich der PA14-Domäne zeigen drei Loops (L1-L3), welche ebenfalls die den äußeren Teil der Bindungstasche formen und mit dem gebundenen Ligand interagieren, eine große Variabilität.

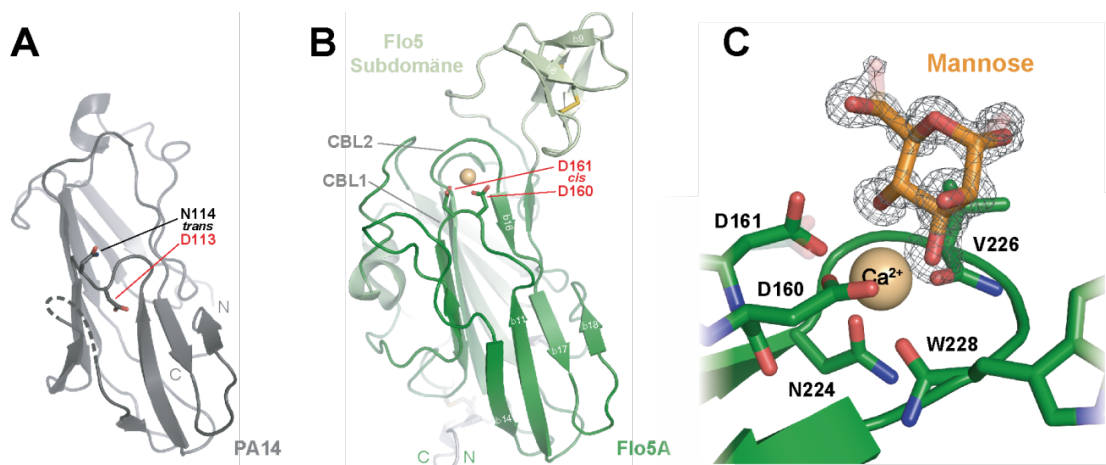


Abbildung 1.9 – Die Struktur der A-Domäne von Flo5 (grün), ein Modell für pilzliche Adhäsine mit PA14-Domäne. A: Die PA14-Domäne des *protective antigens* weist ein *DtransN*-Motiv auf (grau). B: An der äquivalenten Position in der PA14-Domäne von Flo5A findet sich ein einzigartiges *DcisD*-Motiv, welches eine intrinsische Eigenschaft dieser Adhäsine darstellt. Es befindet sich in einem von zwei Loops, CBL1 und CBL2, welche ein Ca^{2+} -Ion koordinieren, um die C-Typ-Lektin artige Bindung zu vermitteln. C: In der inneren Bindungstasche, welche aus den gleichen Resten gebildet wird, liegt eine Mannose (orange) als Ligand vor. Abbildung nach VEELDERS *et al.*¹⁰⁷.

Der Loop L2 ist in den Flokkulinen häufig durch spezifische Subdomänen mit unterschiedlichen Funktionen ersetzt. So trägt die Subdomäne von Flo5 zur Bildung der äußeren Bindungstasche bei und ist in die Bindung der Mannose involviert. In Flo10 hingegen, welches ein eng verwandtes Ortholog darstellt, wird davon ausgegangen, dass die Subdomäne für weitere Prozesse, wie die Bindung an Agar, verantwortlich ist¹⁰⁰. Die beiden Termini der A-Domäne werden über ein Motiv aus zwei konsekutiven Cystein-Resten in räumliche Nähe gebracht. Dass die Flokkulation durch Bindung von α -Mannosiden in der Zellwand der Hefe ausgelöst wird, konnte hiermit molekular gezeigt werden. Bindungsstudien deuteten nur auf eine Dissoziationskonstante im millimolaren Bereich hin. Die Präsentation einer großen Anzahl von Flokkulin-Molekülen (bis zu 10^6 Moleküle pro Zelle) auf der Zelloberfläche und die hohe Dichte von Liganden erzeugt jedoch eine hohe Avidität^{99,134}. Die Strukturlösung von Lg-Flo1A (PDB-ID: 4GQ7) durch SIM *et al.* und von Flo1A durch IELASIS *et al.* bestätigte den C-Typ-Lektin artigen Mechanismus und machte die starke strukturelle Konservierung des PA14-Motivs deutlich^{135,136}. Darüber hinaus konnte gezeigt werden, dass Lg-Flo1A in der Lage ist, an Phosphomannoside mit einer Dissoziationskonstante im hohen mikromolaren Bereich zu binden.

1.3.3 Epitheliale Adhäsine aus *C. glabrata*

Um einen Wirt kolonisieren zu können, macht sich *Candida glabrata* ebenfalls Adhäsine des PA14-Typs zu Nutze, die epithelialen Adhäsine (Epas)^{111,137}. *In vivo* ermöglichen diese die Bindung an humane Epithelzellen, welche die gesamte Körperoberfläche, aber auch Hohlor-gane, wie den Verdauungs- und Urogenitaltrakt, auskleiden¹³⁸. Dies gelingt durch Erkennung von auf der Zelloberfläche präsentierten Sacchariden, bspw. Galactosiden. Aufgrund der essentiellen Funktion, welche diese Adhäsine ausüben, tragen *Candida* Stämme häufig ein ganzes Spektrum von Epa-Orthologen. Das Genom des Stamms CBS138 kodiert bspw. für 17 dieser Proteine. Insgesamt sind 23 verschiedene Epas bekannt^{34,35}. Die Regulation der Expression sowie die subtelomere Lage der Gene dieser Adhäsine zeigen starke Parallelen zu den Flokkulinen^{139,140}. Epa1 ist unter Laborbedingung für etwa 95% des adhäsiven Phänotyps an HEp-2 Epithelzellen verantwortlich¹³⁷. Eine Deletionsmutante zeigt *in vivo* jedoch keinerlei Beeinträchtigung bei der Infektion von Mäusen. Dies wird aufgrund einer Komplementierung der Funktion durch weitere Epas mit ähnlicher Bindungsspezifität möglich. Werden dagegen EPA1, EPA6 und EPA7 deletiert, zeigt sich der Verlust der Adhäsion^{139,141}. Es wird davon ausgegangen, dass die Kolonisation verschiedener Gewebetypen auf die Bindung durch ein spezielles Epa-Protein erfolgt¹⁴²⁻¹⁴⁴. Mit Hilfe heterologer Expression und Präsentation dieser drei Adhäsine auf *S. cerevisiae*-Zellen konnte deren Bindungsspezifität untersucht werden¹³⁸. Diese Experimente zeigten eine selektive Adhäsion an Galactoside, jedoch mit jeweils unterschiedlichen Restriktionen was das sekundäre Saccharid oder die glycosidische Bindung angeht. Dies machte deutlich, dass die reine Zusammensetzung von CBL2, welche für Epa1 und Epa7 identisch ist, nicht ausreichend ist, um die Bindungsspezifität der

Adhäsine mit PA14-Domäne zu erklären. Die Strukturlösung und eingehende biochemische Analyse zur A-Domäne von Epa1 (PDB-ID: 4ASL) verdeutlichte diesen Zusammenhang¹⁰⁸. Trotz einer verhältnismäßig geringen Sequenzidentität von rund 25% zu Flo5A ist global eine ähnliche Faltung zu beobachten. Neben dem strukturell konservierten β -Sandwich der PA14-Domäne und dem *DcisD*-Motiv sind jedoch einige Unterschiede erkennbar (Abbildung 1.10).

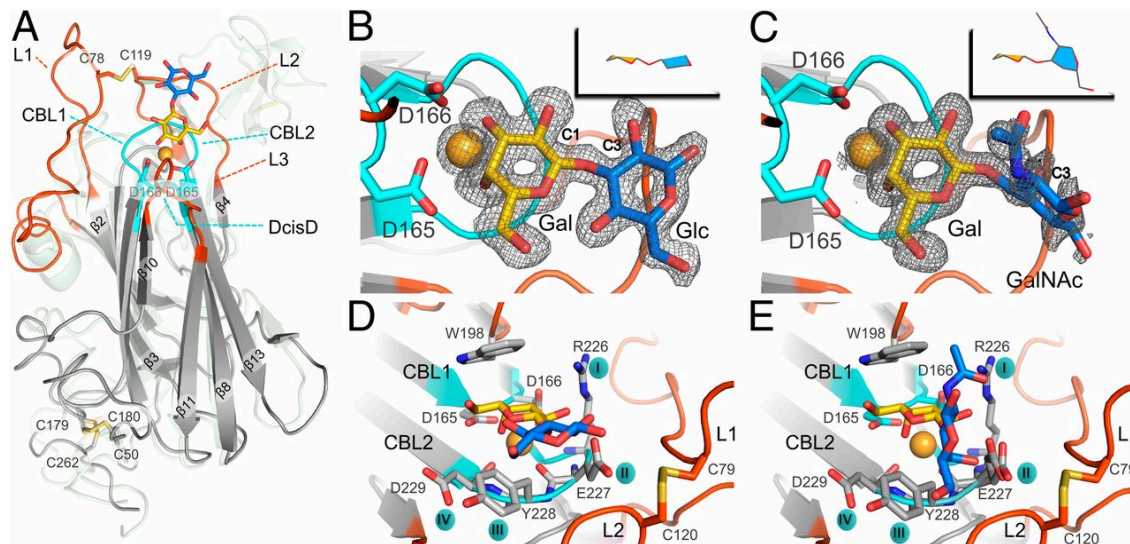


Abbildung 1.10 – Die Struktur der A-Domäne von Epa1. A: Die Struktur von Epa1A (grau) im Komplex mit Galβ1-3Glc (Gal: gelb; Glc: blau). CBL1 und CBL2 (cyan) komplexieren ein Ca²⁺-Ion (orange). Die flexiblen Loops L1-L3 gehören zur Bindungstasche und schützen den gebundenen Liganden vor Einflüssen durch das umgebende Solvens. Darüber hinaus vermitteln diese zusammen mit CBL2 die Ligandenspezifität. Zum Vergleich ist die Struktur der Flo5A-Domäne gezeigt (hellgrün). B-E: Die Epa1A-Domäne als Komplex mit unterschiedlichen Liganden (B+D: Galβ1-3Glc; C+E: Galβ1-3GalNAc/T-Antigen)¹⁰⁸.

So besitzt Epa1A keine Subdomäne, sondern einen verkürzten Loop in L2. Dieser ist darüber hinaus mit einer Disulfidbrücke an Loop L1 fixiert, welcher im Vergleich zu Flo5A verlängert ist. Während L3 über der Bindungstasche in Flo5A kaum in Wechselwirkung mit der gebundenen Mannose tritt, steht dieser in Epa1A mittels eines Tryptophans in direktem Kontakt mit der terminalen Galactose. Dies erklärt die stärkere Adhäsion von Epa1A mit einer Dissoziationskonstante im niedrigen mikromolaren Bereich. Solch eine starke Bindung, gekoppelt mit entsprechender Avidität, unterstreicht die Wirkung als Pathogenitätsfaktor. Auch in Epa1A ist die innere Bindungstasche durch CBL1 und CBL2, der äußere Teil durch L1-L3 geformt. Die Orientierung der beiden Saccharide (Mannose bei Flo5A, Galactose bei Epa1A) unterscheidet sich durch eine Rotation um 180° um die glycosidische Bindung. Die vicinalen Hydroxylgruppen tauschen dabei ihre Position. Dies gelingt durch die Abschirmung unterschiedlicher Bereiche in beiden Adhäsinen sowie durch die unterschiedlichen Reste in CBL2. So ist A225 in Position I von Flo5A in Epa1A gegen ein raumforderndes R226 ausgetauscht. Diese Änderung der Form der Bindungstasche ist wichtig, um die Bindung β -verknüpfter Galactoside zuzulassen. Die Spezies *Candida* besitzt neben den epithelialen Adhäsinen noch weitere Adhäsionsdomänen, mit und ohne PA14-Domäne. Die 7 Prote-

ine des *PA14-containing wall proteins*-Subtyps (Pwp) aus *C. glabrata* enthalten, laut bioinformatischen Vorhersagen, ebenfalls eine PA14-Domäne^{111,145}. Entscheidend ist hierbei das Fehlen des in den Epa essentiellen Tryptophans aus L3. Für Pwp7 konnte *in vitro* bereits die Adhäsion an Epithelzellen gezeigt werden¹⁴⁵. Darüber hinaus existieren im gleichen Organismus etwa 42 weitere putative Adhäsine mit keiner Ähnlichkeit zum Epa- oder PWP-Subtyp. *C. albicans* trägt bspw. keines der Epa-Gene, die Adhäsion wird hier hauptsächlich durch Als-Adhäsine vermittelt, welche Peptid-Termini binden können^{109,146}. Der Aufbau aus 3 Domänen bleibt erhalten, der Typ der A-Domäne unterscheidet sich aber mitunter deutlich von der PA14-Domäne. Obwohl die Etablierung der Adhäsion speziesabhängig durch unterschiedliche Mechanismen erfüllt werden kann, konnten GABALDON *et al.* anhand des Epa-Subtyps zeigen, dass die Evolution dieses Subtyps durchaus mit der Pathogenität in Zusammenhang steht³⁷. Nicht oder minder pathogene Verwandte von *C. glabrata* aus dem Phylum *Nakaseomyces* zeigen entweder eine stark verminderte Anzahl von Epa-Proteinen (1-2 Orthologe) oder eher Flokkulin-ähnliche Homologe.

1.3.4 PA14-Domänen in *P. pastoris* und weiteren Hefen

Das Vorhandensein von Adhäsinen mit PA14-Domänen in weiteren pilzlichen Spezies wird aufgrund der beobachtbaren, Zucker-inhibierbaren Flokkulation und Sequenzvergleichen vermutet^{147,148}. So wurde bereits von 15 flokkulierenden Hefen, darunter *Kluyveromyces* und *Pichia*-Stämme, berichtet. Die Varianz innerhalb eines Phylums kann hierbei sehr groß sein. Die Flokkulation von *K. marxianus* ist offenbar nur mit Mannose inhibierbar, während in *K. bulgaricus* mindestens 6 Saccharide, darunter Fucose, Galactose oder GlcNAc, in Frage kommen¹⁴⁹⁻¹⁵¹. Dies kann einerseits auf eine große Spezifität vorhandener Adhäsine zurückzuführen sein oder andererseits auf ein großes Spektrum von Orthologen mit unterschiedlicher Funktion. Hierbei ist zu beachten, dass es sich nur um den Phänotyp der Flokkulation handelt und weitere, möglicherweise unbekannte Funktionalitäten nicht untersucht sind. Untersuchungen zur Flokkulation der Spaltheefe *Schizosaccharomyces pombe* deuten darüber hinaus auf das Vorhandensein einer zweiten Gruppe Ca²⁺-sensitiver, Saccharid-inhibierbarer Flokkuline hin¹⁵². In diesem speziellen Fall wird die Flokkulation durch das putative Adhäsिन Gsf2 (*galactose specific flocculation*) vermittelt, welches keinerlei Homologie zu den bekannten Flokkulinen und Adhäsinen aufweist.

Gerade die Adhäsionseigenschaften von *Pichia pastoris* sind von großem Interesse, denn es handelt sich hierbei um einen Modellorganismus mit unbekannter Ökologie aber weit verbreiteter Anwendung. Dies eröffnet die Chance zur Entdeckung unbekannter Funktionalitäten und Phänotypen. In früheren Arbeiten konnten erste putative PA14-Domänen in *Pichia pastoris* gefunden werden¹⁵³⁻¹⁵⁵. Diese zeigten in vorläufigen bioinformatischen Studien auf Sequenzebene starke Ähnlichkeit zu den PA14-Domänen von Flo5A und Epa1A bei nur begrenzter Sequenzidentität (~25%). Darüber hinaus konnte das konservierte DD-Motiv sowie zwei konsekutive Cysteine identifiziert werden. Das Fehlen einer entsprechenden Subdomäne

wurde ebenfalls vorhergesagt. Erste Experimente deuteten auf eine Bindung von terminalen GlcNAc-Resten durch eines dieser putativen Adhäsine hin.

1.4 Zielsetzung dieser Arbeit

Aufgrund der Applikation von Adhäsinen mit PA14-Domänen in für Hefen essentiellen Prozessen, ist davon auszugehen, dass dieses Prinzip ein weit verbreitetes Konzept darstellt. Da allerdings die globalen phylogenetischen Zusammenhänge zwischen den unterschiedlichen Adhäsinen so gut wie unbekannt sind, lassen sich kaum Rückschlüsse auf deren Entwicklung und Funktion in der Interaktion mit der Umwelt der Hefen treffen. Die moderne Genomanalyse erlaubt das Auffinden neuer Adhäsine mit unbekannter Funktion in weiteren Organismen. Durch Untersuchung neuartiger PA14-Domänen in pilzlichen Adhäsinen ergibt sich dadurch nicht nur die Möglichkeit, die molekulare Grundlage der Spezifitätserzeugung besser zu verstehen, sondern auch, neue Funktionalitäten zu entdecken. Da diese Adhäsine in Hefen weit verbreitet sind, lässt sich durch das Studium dieser Proteinfamilie ebenfalls etwas über die Biologie und Ökologie dieser vielfältigen Organismen erfahren.

Im Speziellen sollte durch die detaillierte bioinformatische, biochemische und röntgenkristallographische Untersuchung der nur schlecht charakterisierten putativen Adhäsine aus *Pichia pastoris* ein tieferes Verständnis für die molekularen Grundlagen der pilzlichen Adhäsion geschaffen werden. Neu entdeckte Funktionen könnten dabei auch Anwendung in Biotechnologie und Medizin finden. Phylogenetische Studien zu der gesamten Adhäsinfamilie mit PA14-Domäne sollten darüber hinaus einen Einblick in die globalen Zusammenhänge liefern und zeigen, ob sich im Laufe der Evolution spezielle Subtypen ausgebildet haben.

Um einen Einblick in die Ausbildung von Spezifität und struktureller Integrität eines einzelnen Subtyps zu erlangen, sollten Bindungsstudien und röntgenkristallographische Untersuchungen an verschiedenen Mitgliedern der epithelialen Adhäsine durchgeführt werden. Gerade die Betrachtung dieses medizinisch relevanten Subtyps könnte einen Beitrag zur Entwicklung von Inhibitoren der Adhäsion liefern. Darüber hinaus sollten, anhand strukturbioologischer Untersuchungen von Flo10A, Informationen zum Einfluss der in L2 vorhandenen Subdomäne gewonnen werden.

2 Material

2.1 Geräte

Gerät	Bezeichnung (Hersteller)
Äkta Prime-System	(<i>GE Healthcare</i>)
Autoklav	T-Line (<i>Fedegari</i>), HV 110-L (<i>HMC</i>)
Bidest. Wasser-Anlage	Seralpur Pro90CN (<i>Seral</i>)
Brutschrank	BFED-53 (<i>Binder</i>)
Chromatografiesäulen	XK 16/60-Säulenkörper (<i>GE Healthcare</i>), Superdex 200/75-Material (<i>GE Healthcare</i>), His-Trap Ni-NTA (<i>Macherey Nagel</i>)
Digitalkamera	D70s (<i>Nikon</i>)
Dokumentationssystem für Agarosegele	Computer E.A.S.Y. (<i>UVP</i>), Kamera (<i>UVP</i>), UV-Schirm (<i>Herolab</i>)
Dokumentationssystem für Kristallisationsansätze	Rockimager 1000 (<i>Formulatrix</i>)
Elektrophoresekammer für Agarosegele	Eigenbau (<i>Feinmechanische Werkstatt, FB Chemie, Philipps-Universität Marburg</i>)
Elektrophoresekammer für SDS-PAGE	Mighty Small II SE250 / SE260 (<i>Hoefler</i>)
Fluoreszenzmikroskop	Axiovert 200M (<i>Carl-Zeiss AG</i>), Plan Apochromat Öl-Objektiv NA 1,4 (<i>Zeiss</i>), Orca ER (<i>Hamamatsu</i>), GFP- / Rhodamin- / DAPI-Filter Set (<i>AHF Analysetechnik AG</i>), Axiophot (<i>Zeiss</i>)
FPLC-System	Gradient Programmer GP-250 Plus, Pump P-500, Single Path Monitor UV-1, Fraktionskollektor Frac-100, Schreiber Rec112 (<i>GE Healthcare</i>)
Glasperlenmühle	FastPrep-24 (<i>MP Biomedicals</i>)
Hochdruckhomogenisator	French Press (<i>Aminco</i>), Hochdruckzelle (<i>Feinmechanische Werkstatt, FB Biologie, Philipps-Universität Marburg</i>)
ITC-System	ITC ₂₀₀ (<i>MicroCal</i>)
Kristallisationsroboter	Cartesian MicroSys 4000-XL (<i>Genomic Solutions</i>), Cartesian HoneyBee 963 (<i>Genomic Solutions</i>)
Mikroskop	B601 (<i>Olympus</i>), MZ 8 (<i>Leica</i>)
NGC-System	NGC Chromatography System (<i>Biorad</i>)
Peristaltikpumpe	P1 (<i>GE Healthcare</i>)
pH-Meter	Basic Meter PB-11 (<i>Sartorius</i>)
Photometer	ND-1000 (<i>Nanodrop</i>), Ultrospec 3100pro (<i>GE Healthcare</i>), OD600 (<i>Implen</i>)
Pipetten	Research (<i>Eppendorf</i>)
Real-Time PCR-Cycler	Rotor-Gene Q (<i>Qiagen</i>)

Röntgendiffraktometer	Rotating Anode AXS FR591 (<i>Bruker</i>), MAR345dtb detector (<i>MAR research</i>), Oxford cryosystem (<i>Oxford</i>)
Schüttelinkubatoren	Certomat IS (<i>Sartorius</i>), Controlled Environment Incubator Shaker (<i>New Brunswick Scientific</i>), Multitron (<i>InforsHT</i>)
Spannungsquelle	EPS 300 / 301 (<i>GE Healthcare</i>)
Synchrotron-Beamlines	Beamline BL14.1 & BL14.2 BESSY-II Berlin (<i>Helmholtz-Gesellschaft</i>), Beamline P13 & P14 PETRA III Hamburg (<i>DESY</i>)
Waagen	PC2200 (<i>Mettler</i>), LabStyle 54 (<i>MettlerToledo</i>)
Wasserbad	NK22 (<i>Haake</i>)
Zentrifugen	Centrifuge 5810 R (<i>Eppendorf</i>), J2-HS Zentrifuge mit JA10- / JA20-Rotor (<i>Beckmann</i>), Biofuge 15 (<i>Heraeus</i>), Fresco21 (<i>Thermo Scientific</i>)

2.2 Verbrauchsmaterialien

Verbrauchsmaterial	Bezeichnung (Hersteller)
Einmalküvetten	67.724 (<i>Sarstedt</i>)
Entsalzungssäulen	PD-10 Desalting Column (<i>GE Healthcare</i>), PD MiniTrap G25 (<i>GE Healthcare</i>)
Kit für die qRT-PCR	qScript One-Step SYBR Green Kit (<i>Quanta Bioscience</i>)
Kit für DNA-Reinigung	QIAquick PCR Purification Kit (<i>Qiagen</i>)
Kit für Gelextraktion	QIAquick Gel Extraction Kit (<i>Qiagen</i>)
Kit für Plasmidpräparation	QIAprep Spin Miniprep Kit (<i>Qiagen</i>)
Kit zum Aufschluss von Hefezellen	QIAshredder (<i>Qiagen</i>)
Kit zur Präparation und Reinigung von mRNA	RNeasy Mini Kit (<i>Qiagen</i>)
Konzentratoren	Amicon Ultra-15 / Amicon Ultra-5 Centrifugal Filter Devices, 10000 MWCO (<i>Merck Millipore</i>)
Kristallisationszubehör	96-well Innovaplate (<i>Innovadyne</i>), VIEWseal (<i>Greiner BIOone</i>), SilverSeal (<i>Greiner BIOone</i>), 24-well Crystallization Tool mit Crystal Support (<i>Qiagen</i>), Micromounts (<i>MiTeGen</i>), CrystalCap HAT™ (<i>Hampton research</i>), Cryotong / Crystal Wand™ / Micro-Tools (<i>Hampton Research</i>)
Spritzenvorsatzfilter	Filtropur S 0.2 mit 0.2 µm Porengröße (<i>Sarstedt</i>), Filtropur S 0.45 mit 0.45 µm Porengröße (<i>Sarstedt</i>)
Sterilfilter	Ultrafree-MC (<i>Merck Millipore</i>), Flaschenaufsatzfilter (<i>Merck Millipore</i>)

2.3 Chemikalien

Chemikalie	Hersteller
Acrylamid (30%)	<i>Roth</i>
Agar	<i>Roth</i>
Agarose	<i>Invitrogen</i>
Alexa-Fluor® 488 (A30005)	<i>Life Technologies</i>
APS	<i>Fluka</i>
Bacto-Hefeextrakt	<i>Difco</i>
Bacto-Trypton	<i>Difco</i>
Bradford-Reagenz	<i>Biorad</i>
BSA	<i>Roth</i>
Calciumchlorid	<i>Riedel-de-Haën</i>
Chitin-Beads	<i>New England Biolabs</i>
Coomassie-Brilliant-Blue	<i>Serva</i>
D-Glucose	<i>Sigma</i>
Decon-90	<i>Decon Laboratories Limited</i>
DiOC ₆	<i>Santa Cruz Biotechnology</i>
MassRuler DNA Ladder Mix	<i>Fermentas</i>
Ethanol	<i>VWR</i>
Ethidiumbromid	<i>Roth</i>
Glycerin	<i>Fluka</i>
Imidazol	<i>Merck</i>
IPTG	<i>Duchefa</i>
Kanamycin	<i>AppliChem</i>
Kristallisationsscreens	<i>Qiagen NeXtal</i>
2-Mercaptoethanol	<i>Roth</i>
Na ₂ EDTA	<i>Merck</i>
Ni-NTA Superflow	<i>Qiagen</i>
peQGOLD Protein-Marker	<i>peqlab</i>
Peroxyessigsäure	<i>Sigma</i>
PMSF	<i>Fluka</i>
SDS	<i>Serva</i>
Spectinomycin	<i>Appllichem</i>
Streptomycin	<i>Sigma</i>
TEMED	<i>Roth</i>
Tris	<i>Roth</i>
Tween-20	<i>Roth</i>
Virkon	<i>Du Pont</i>
Weizenkleie	<i>Spielberger GmbH</i>
Yeast Nitrogen Base	<i>Roth</i>

Alle nicht gesondert aufgeführten Chemikalien wurden in p.a.-Qualität von den Firmen *Sigma-Aldrich*, *Fluka*, *Merck* oder *Roth* bezogen.

2.4 Saccharide und Oligosaccharide

Saccharid	Abkürzung	Hersteller
Chitobiose	GlcNAc β 1-4GlcNAc	<i>Dextra Laboratories Ltd</i>
Galactose-6-O-sulfat	Gal-6S	<i>Dextra Laboratories Ltd</i>
Galactose- β -(1,3)-6-Carboxy-Glucose	Gal β 1-3Glc-6COOH	<i>H. Lin, Academia Sinica</i>
Galactose- β -(1,3)-6-Sulfo-Glucose	Gal β 1-3Glc-6S	<i>H. Lin, Academia Sinica</i>
Galactose- β -(1,3)-Glucose	Gal β 1-3Glc	<i>H. Lin, Academia Sinica</i>
Glucosamin	GlcN	<i>Alfa Aesar</i>
Glucose	Glc	<i>Fluka</i>
Lacto-N-Biose	Gal β 1-3GlcNAc	<i>Dextra Laboratories Ltd</i>
Lactose	Gal β 1-4Glc	<i>Roth</i>
Mannose	Man	<i>Fluka</i>
N-Acetyl-D -lactosamin	Gal β 1-4GlcNAc	<i>Dextra Laboratories Ltd</i>
N-Acetylglucosamin	GlcNAc	<i>Sigma</i>
N-Acetylneuraminsäure	Neu5Ac	<i>TCI Europe</i>
T-Antigen	Gal β 1-3GalNAc	<i>Dextra Laboratories Ltd</i>
α -Galactobiose	Gal α 1-3Gal	<i>Dextra Laboratories Ltd</i>

Soweit nicht anders erwähnt handelt es sich bei den in dieser Arbeit verwendeten Sacchariden um D-Saccharide. Die synthetischen Disaccharide, welche von Prof. LIN zur Verfügung gestellt wurden, besitzen zusätzlich einen Azido-n-Hexanol Linker an Position 1 der Glucose.

2.5 Enzyme

Enzym	Hersteller
BSA	<i>New England Biolabs / Fluka</i>
Chitinase aus <i>Trichoderma viride</i> (Endo- & Exo-Chitinase)	<i>Sigma Aldrich</i>
DNaseI	<i>Boehringer</i>
Lysozym (50 mg/mL)	<i>Boehringer</i>
<i>Nde</i> I (20 U/ μ L)	<i>New England Biolabs</i>
Phusion HiFi-DNA-Polymerase (2 U/ μ L)	<i>New England Biolabs</i>
T4-DNA-Ligase (400 U/ μ L)	<i>New England Biolabs</i>
<i>Xho</i> I (20 U/ μ L)	<i>New England Biolabs</i>

2.6 Primer zur Genamplifikation

Im Rahmen dieser Dissertation wurden pET-28-basierte Plasmidkonstrukte genutzt, welche bereits in früheren Arbeiten kloniert wurden. Der Vollständigkeit halber sind die hierfür verwendeten Primer im Folgenden aufgelistet. Die für die Klonierung in pET-28a(+)-Vektoren genutzten Restriktionsschnittstellen *Nde*I und *Xho*I sind fett gedruckt.

Primer	Sequenz in 5' ->3'	Quelle
Cea1A-fwd	CAGTCGAC ATATGG ATGACAGTGGAAATGG	155-157
Cea1A-rev	GCTCGAG TTATTCATGGCAGGAGTTCTC	155-157
Epa1A-fwd	CCATATG ACATCTTCCAATGATATCAG	108,158
Epa1A-rev	ACTCGAG TTAAGAAGAATCGTAGCTG	108,158
<i>PpFlo</i> 1A-fwd	CAAGTTAC ATATGC AGGAAAGTGGT GATGG	155,156
<i>PpFlo</i> 1A-rev	GCTCGAG TTATGCTTGGCATTGTTCTTC	155,156

2.7 Primer zur Herstellung von Hefe-Expressionskonstrukten

Zur Erstellung von Hefe-Expressionsplasmiden für *in vivo* Experimente im Rahmen dieser Arbeit wurden folgende Primer von NINA WOZNAK verwendet¹⁵⁹. Als Restriktionsenzyme dienten hierbei *Sac*I und *Sac*II.

Primer	Sequenz in 5' ->3'	Quelle
Cea1A-fwd	AAAA ACCGCGG GATGACAGTGGAAATGGCG	157,159
Cea1A-rev	AAAA AAGAGCTC TTTCATGGCAGGAGTTCTCATC	157,159
<i>PpFlo</i> 1A-fwd	AAAA ACCGCGG CAGGAAAGTGGT GATGGT GAC	159
<i>PpFlo</i> 1A-rev	AAAA AAGAGCTC TGCTTGGCATTGTTCTTCGTCGAGTGCACC	159

2.8 Primer für die qRT-PCR

Primer für qRT-PCR wurden mit Hilfe des Programms *Primer3Plus* erstellt, welches das Auffinden geeigneter Sequenzbereiche und Amplifikate für diese Methode erleichtert¹⁶⁰.

Primer	Sequenz in 5' ->3'	Quelle
ACTI-fwd	TCTGTTTGGGTCTGGAAGC	Diese Arbeit
ACTI-rev	TGCATACGCTCAGCAATACC	Diese Arbeit
Cea1A-fwd	TGGAATCCAAGTTCCAGGAC	Diese Arbeit
Cea1A-rev	CTCCTGAAACGTCAGCTTTG	Diese Arbeit
<i>PpFlo</i> 1A-fwd	TTGATGCCACACGTTACGAC	Diese Arbeit
<i>PpFlo</i> 1A-rev	TGGATCCCGTGTTAACCTTC	Diese Arbeit

2.9 Vektoren

2.9.1 pET-28a(+)-Vektor

Der Vektor pET-28a(+) (*Novagen*) wurde zur heterologen Produktion von rekombinanten A-Domänen in *Escherichia coli* eingesetzt. In Verbindung mit kompatiblen Stämmen kann über eine hierin genomisch kodierte T7-RNA-Polymerase die Transkription erfolgen¹⁶¹. Sowohl die Expression dieser Polymerase als auch die Expression des auf dem Vektor kodierten rekombinanten Gens stehen unter Kontrolle des *lacI*-Repressors. Dieser bindet an die Sequenz des *lac*-Operators im Bereich des *lac*- und des T7-Promotors. Das für den Repressor kodierende Gen liegt ebenfalls auf dem Vektor vor. Durch Zugabe von Isopropyl- β -D-thiogalactopyranosid (IPTG) lässt sich der Repressor lösen und die Expression induzieren¹⁶². Der Sequenzabschnitt der *multiple cloning site* (MCS) enthält eine große Anzahl an Schnittstellen für Restriktionsenzyme und ermöglicht dadurch das Einbringen eines entsprechenden Zielgens. Diese Region stellt außerdem die Fusion mit einem N- oder C-terminalen Hexahistidin-Tag zur Verfügung. Die im Rahmen dieser Arbeit verwendeten Fusionsproteine wurden mit einem N-terminalen Tag hergestellt.

2.9.2 YCplac33-Vektor

Der Shuttle-Vektor YCplac33 diente als Grundlage zur Erstellung von Expressionsplasmiden für *S. cerevisiae*. Zur Klonierung in *E. coli* enthält der Vektor neben einer Ampicillin-Resistenz zur Selektion einen entsprechenden *Origin of Replication* (ORI). Um eine Funktion und Replikation in *S. cerevisiae* zu gewährleisten ist auf dem Vektor eine *Autonomously Replicating Sequence* (ARS), ein Hefe-Centromer (CEN) sowie das URA3-Gen zur Synthese von Uracil als Selektionsmarker zu finden, wodurch bei transformierten Zellen eine entsprechende Komplementation erfolgt. Neben einer hohen Transformationseffizienz zeichnet er sich durch seine mitotische Stabilität sowie durch seine einfache Kopienzahl aus.¹⁶³

2.10 Verwendete Plasmide zur Expression in *E. coli*

Folgende Tabelle enthält die Plasmide, mit welchen rekombinante Proteine im Rahmen dieser Arbeit produziert worden sind. Darüber hinaus beinhaltet diese Arbeit auch die Strukturlösung von weiteren Adhäsionsdomänen. Diese wurden von Kooperationspartnern produziert und biochemisch untersucht. Der Vollständigkeit halber sind diese in der Tabelle ebenfalls aufgeführt.

Name	Konstrukt	Quelle
pET28a-Cea1A	CEA1A ²³⁻²⁴¹ in pET28a(+)	155
pET28a-Epa1A	EPA1A ³¹⁻²⁷¹ in pET28a(+)	158
pET28a-Epa6A	EPA6A ²⁶⁻²⁷¹ in pET28a(+)	164,165
pET28a-Epa9A	EPA9A ³⁹⁻³⁰⁵ in pET28a(+)	164,165
pET28a-Flo10A	FLO10A ²³⁻²⁹³ in pET28a(+)	159
pET28a-PpFlo1A	PPFLO1A ²⁶⁻²⁴⁶ in pET28a(+)	155

2.11 Verwendete Plasmide zur Expression in *S. cerevisiae*

Die in dieser Arbeit verwendeten Plasmide auf Basis von YCplac33 zur Expression in *S. cerevisiae* sind in untenstehender Tabelle aufgelistet. Die Expression erfolgt unter Kontrolle des konstitutiven PGK1-Promotors (Phosphoglyceratkinase 1). Der Flo11 Signalsequenz, welche den extrazellulären Transport sicherstellt, folgt nach einer Spacersequenz (SP) und einem 3HA-Epitop, die entsprechende A-Domäne sowie B- und C-Region von Flo11. Das Gen endet mit dem Flo11-Terminator.

Name	Konstrukt	Quelle
BHUM2200	P _{PGK1} -FLO11 ¹⁻³⁰ -FLO11 ²¹⁴⁻¹³⁶⁰ -T _{FLO11}	165
BHUM2296	P _{PGK1} -FLO11 ¹⁻³⁰ -SP-3HA-PPFLO1 ²⁶⁻²⁴⁶ -FLO11 ²¹⁴⁻¹³⁶⁰ -T _{FLO11}	157,159
BHUM2297	P _{PGK1} -FLO11 ¹⁻³⁰ -SP-3HA-CEA1 ²³⁻²⁴¹ -FLO11 ²¹⁴⁻¹³⁶⁰ -T _{FLO11}	157,159

2.12 Bakterienstämme

2.12.1 *E. coli* DH5

Genotyp: F- Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 *recA1 endA1 hsdR17* (*r_K⁻, m_m⁺*) *phoA supE44* λ - *thi-1 gyrA96 relA1*

E. coli Zellen dieses Stammes sind besonders geeignet zur Plasmidgewinnung¹⁶⁶. Dafür besitzen sie neben einer hohen Plasmidreplikationsrate bspw. eine Defizienz in Endonuclease A1 (*endA1*).

2.12.2 *E. coli* SHuffle® T7 Express

Genotyp: F' *lac, pro, lacI^Q* / Δ (*ara-leu*)7697 *araD139 fhuA2 lacZ::T7 gene1* Δ (*phoA*)*PvuII phoR ahpC* galE* (or U) *galK* λ *att::pNEB3-r1-cDsbC* (*Spec^R, lacI^q*) Δ *trxB rpsL150(Str^R)* Δ *gor* Δ (*malF*)3 (K12 Stammhintergrund)

Der SHuffle® T7 Express Stamm der Firma NEB wurde zur heterologen Proteinproduktion genutzt. Der Vorteil dieses Stammes liegt zum einen in seinem oxidativen Zellmilieu, was durch eine Mutation in der Glutathion-Reduktase und der Thioredoxin-Reduktase sichergestellt wird. Zum anderen wird unter Kontrolle eines konstitutiven Promotors die Disulfidbrücken-Isomerase *DsbC* exprimiert. Weiterhin ist diese ein Chaperon und unterstützt damit nicht nur die Knüpfung nativer Disulfidbrücken, sondern unterstützt den Faltungsprozess im Ganzen¹⁶⁷.

2.13 Hefestämme

2.13.1 *P. pastoris* DSMZ70382

Genotyp: Wildtyp-Isolat

Dieser Wildtyp-Stamm von *Pichia pastoris* wurde von der *Deutschen Sammlung von Mikroorganismen und Zellkulturen* (DSMZ) kommerziell erworben und dient weiterhin als Referenzstamm. Hierbei handelt es sich um ein Isolat aus dem Exsudat eines Kastanien-Baums aus der Gegend um Lyon. Da das Genom dieses Stammes komplett sequenziert vorliegt, stellte er eine gute Grundlage für die Studien an unbekanntem ascomycetalen Adhäsionsdomänen dar. Genomische DNA von *P. pastoris* DSMZ70382 wurde für die Genamplifikation genutzt. Zellkulturen des Stammes wurden zu *in vivo*-Experimenten genutzt⁵⁰.

2.13.2 *P. pastoris* GS115

Genotyp: *his4* (NRRL Y-11430 Stammhintergrund)

Der Stamm GS115 wurde aus dem Wildtyp-Isolat NRRL Y-11430 abgeleitet, indem u.a. eine Histidin-Auxotrophie eingeführt wurde⁴⁹. In Kombination mit entsprechenden Vektoren lässt sich diese Eigenschaft als Selektionsmarker nutzen. Das Genom dieses Stammes ist ebenfalls sequenziert und er bildete die Grundlage für die Weiterentwicklung von *P. pastoris* zu einem weit verbreiteten System zur heterologen Proteinproduktion. Dieser Stamm wurde ebenfalls für bioinformatische, biochemische sowie *in vivo* Studien eingesetzt.

2.13.3 *S. cerevisiae* RH2520

Genotyp: *MATa ura3-52* (S288c Stammhintergrund)

Dieser Stamm der Bäckerhefe *Saccharomyces cerevisiae* basiert auf dem Stammhintergrund S288c. Er besitzt den *Mating type a* und eine Uracil-Auxotrophie. Letztere kann als Selektionsmarker genutzt werden. Dieser Stamm wurde zur Expression und Oberflächenpräsentation von Fusionskonstrukten aus unterschiedlichen Adhäsion-A-Domänen und Flo11 B- und C-Domänen genutzt. Diese Zellen wurden für Experimente *in vivo* verwendet. Durch eine Mutation im Transkriptionsfaktor FLO8 ist der Stammhintergrund S288c Flokkulationsdefizient und eignet sich dadurch für Experimente an heterologen Adhäsionsdomänen^{168,169}.

2.14 Insekten

2.14.1 *Shelfordella lateralis*

Die Schabe *Shelfordella lateralis* (dt. Trivialname: Schokoschabe) gehört zur Familie der Blattellidae und ist ein enger Verwandter der amerikanischen Großschabe *Periplaneta americana*. Die Schaben wurden bei einem Züchter erworben (*J. Bernhard*, Helbigsdorf) und dienen als Modellorganismus zur Untersuchung von putativen Adhäsion-vermittelten Wechselwirkungen

zwischen Hefen und Insekten. *S. lateralis* wurde in früheren Arbeiten bereits zur Charakterisierung der bakteriellen Darmflora von Schaben genutzt¹⁷⁰. Hierbei konnte gezeigt werden, dass die drei markanten Bereiche des Darms (*Crop*, *Midgut* & *Hindgut*) unterschiedliche biochemische und physikochemische Eigenschaften besitzen.

3 Methoden

3.1 Bioinformatische Methoden

3.1.1 *Genome-mining* nach Adhäsionsproteinen des PA14-Typs innerhalb der Ascomyceten

Um ein tieferes Verständnis über die Evolution und die Diversität der ascomycetalen PA14-Domänen enthaltenden Adhäsine zu erhalten, wurden sequenzierte Genome unterschiedlicher Hefen nach putativen Adhäsinen durchsucht. Für diesen *Genome-mining* ähnlichen Ansatz wurde hauptsächlich das Online *BLAST* (*B*asic *L*ocal *A*lignment *S*earch *T*ool) des *National Institute of Health* (NIH) genutzt. Der Algorithmus dieses Tools erlaubt das Durchsuchen einer Datenbank von Nukleotid- oder Proteinsequenzen mit Hilfe eines vorgegebenen Templates nach verwandten/ähnlichen Sequenzen¹⁷¹. Hierbei wurde im Speziellen auf die Submethoden *nucleotide blast*, *protein blast*, *tblastn* sowie *PSI-BLAST*, eine Variante des *protein blasts*, zurückgegriffen. Die Methoden unterscheiden sich hierbei in dem Format des verwendeten Suchtemplates sowie in der durchsuchten Datenbank.

BLAST Programm	Templat	Datenbank
<i>nucleotide blast</i>	Nukleotidsequenz	Nukleotide
<i>protein blast</i>	Proteinsequenz	Proteine
<i>Tblastn</i>	Proteinsequenz	Translatierte Nukleotide

PSI-BLAST erlaubt hierbei das Auffinden von weiter entfernt verwandten Sequenzen indem aufbauend auf einen ersten Suchschritt eine benutzerdefinierte *scoring matrix*[‡] erstellt wird, mit welcher die weiteren Schritte erfolgen^{172,173}.

Standardmäßig wird mit dem *BLAST* eine Datenbank bestehend aus annotierten Genen und Proteinen durchsucht. Davon abweichend erlaubt das Tool auch die Suche in annotierten & nicht-annotierten Genomen (72 655 Sequenzen[§]) sowie in Genom-Fragmenten welche Resultate von *shotgun*-Sequenzierungen sind (166 270 376 Sequenzen)^{174,175}. Als Grundlage für die Suche nach unbekanntem, PA14-Domänen enthaltenden Homologen wurden die Sequenzen der A-Domäne der biochemisch und strukturell bereits gut untersuchten Adhäsine Flo5 und Epa1 genutzt. Hierdurch konnten vor allem Homologe in den Genomen dreier *Pichia- / Komogataella*-Stämme identifiziert werden⁵⁰⁻⁵².

Als weitere Quelle für putative Adhäsine des PA14-Typs in Ascomyceten dienten die Proteinsequenzdatenbank der Online-Plattform *InterPro* sowie die Domänendatenbank *pfam*^{121,176,177}. *InterPro* ist ein Tool des *European Bioinformatics Institutes* zur funktionellen Analyse von Proteinsequenzen. Dieses wird unter anderem genutzt um Proteinsequenzen der

[‡] Standardmäßig wird in *BLAST* die BLOSUM62 Matrix verwendet.

[§] Stand 09.12.2014

Datenbank *UniProt* zu annotieren¹⁷⁸. Über Sequenzvergleiche werden die Primärsequenzen automatisch Domänen oder Familien zugeordnet; diese Automatisierung bildet andererseits auch die Grundlage für fehlerhafte Annotationen. *Pfam*, eine Domänendatenbank des *Sanger Institutes*, liefert zudem eine maschinelle Mustererkennung in Proteinsequenzen und lässt, nach Zuordnung von Sequenzen zu Domänen und funktionellen Familien, einen Rückschluss auf die Funktion der Proteine zu. Die pilzlichen Adhäsine des PA14-Typs sind sowohl hierin als auch in *InterPro* weitgehend zwei Domänen-Familien zugeordnet: Der PA14-Domäne (Pfam-ID: PF07691) und der PA14_2- oder Gleya-Domäne (Pfam-ID: PF10528).

3.1.2 Multiples Sequenzalignment und phylogenetische Analyse

Die somit identifizierten putativen Homologe von Flo5 und Epa1 wurden zur weiteren Analyse einem multiplen Sequenzalignment (MSA) unterzogen. Da darüber hinaus die dreidimensionalen Strukturen beider Adhäsionsdomänen zur Verfügung standen, wurde auf einen strukturbasierten Algorithmus des Programmes *TCoffee* zurückgegriffen, *3D-Coffee*¹⁷⁹. Hierbei wird das Alignment auf struktureller Ebene mit *TMalign* auf Basis der bekannten Struktur-Template gebildet. Zuerst wurde ein robustes Alignment auf Grundlage der Sequenzen bekannter Adhäsionsdomänen generiert: Epitheliale Adhäsine, Flokkuline, PWP's & Adhäsine aus *Nakaseomyces*³⁷. Um die zuvor identifizierten, putativen Homologe in das Alignment einzubringen wurden Algorithmen des Tools *mafft--add*, welches vom Alignment Programm MAFFT7 abgeleitet ist, verwendet. Dieses erlaubt das Einfügen neuer Sequenzen in ein bestehendes Alignment¹⁸⁰. Dafür wird die phylogenetische Beziehung zwischen den Sequenzen im bestehenden Alignment genutzt, um die Verwandtschaft der hinzuzufügenden Sequenzen daraus abzuleiten¹⁸¹.

Dieses primäre Alignment (303 Sequenzen) diente, neben der visuellen Inspektion und dem Vergleich mit Flo5 und Epa1, zur Definition der wahrscheinlichsten Grenzen der Adhäsionsdomänen**. Des Weiteren wurden 19 redundante Sequenzen und 177 Sequenzen, welche kein putatives *DcisD*-Motiv sowie das entsprechende Asparagin von CBL1 im Alignment zeigten, gelöscht.

Um den verwandtschaftlichen und evolutionären Zusammenhang zwischen bekannten sowie unbekanntem Adhäsionsproteinen des PA14-Typs zu untersuchen, wurde eine phylogenetische Studie auf Basis der *Maximum-Likelihood* Methode (*ML*) durchgeführt. Eine phylogenetische Analyse leitet mit Hilfe rechnergestützter Methoden und Algorithmen einen evolutionären Zusammenhang zwischen unterschiedlichen biologischen Entitäten (bspw. Spezies/Taxa/Proteine) her. Als Grundlage hierfür dienten historisch gesehen morphologische oder phänotypische Merkmale, während im heutigen Zeitalter molekulare Determinanten wie Gen- & Proteinsequenzen Verwendung finden. Neben den *Distance-Matrix*-Methoden,

** Für das Alignment sowie die nachfolgende phylogenetische Untersuchung wurden ausschließlich die Sequenzen der A-Domäne benutzt, um den funktionellen Bezug in den Vordergrund zu stellen.

welche einen phylogenetischen Zusammenhang auf alleiniger Grundlage der „genetischen Distanz“ zwischen Sequenzen herleiten, werden auch weitere Methoden/Algorithmen wie *ML*, *Maximum Parsimony (MP)* oder *Bayesian* genutzt.

Bei der hier gebrauchten *ML*-Methodik werden statistische Techniken verwendet um einer größeren Anzahl generierter Bäume eine Wahrscheinlichkeitsverteilung zuzuordnen und damit den geeignetsten phylogenetischen Baum identifizieren zu können. Hierzu wird ein Substitutionsmodell verwendet, welches die Wahrscheinlichkeit eines Austausches zweier bestimmter Nukleotide bzw. Aminosäuren beinhaltet.¹⁸²

Für das finale multiple Sequenzalignment sowie die darauf basierende phylogenetische Analyse wurden folgende 107 Sequenzen genutzt, wobei nur Adhäsine der hervorgehobenen Gruppen bisher genauer biologisch und biochemisch charakterisiert wurden:

<i>Proteinfamilie</i>	Anzahl der Sequenzen
<i>Epitheliale Adhäsine</i>	19
<i>Flokkuline</i>	5
<i>Pichia pastoris Adhäsine</i> aus <i>genome mining</i>	5
<i>Putative Homologe</i>	50
<i>PWPs</i>	6
<i>Nakaseomyces Adhäsine</i>	22

Für die Konstruktion des phylogenetischen Baumes wurde die Software *MEGA6* verwendet¹⁸³ und zu Beginn das am besten geeignete Modell zur Generierung des Baumes ermittelt. Hierzu werden alle in *MEGA6* möglichen Kombinationen aus Substitutionsmethoden der Aminosäurereste, Rate der Substitution, Behandlung von Lücken im Alignment und der heuristischen Methode zum Ableiten des entsprechenden Baumes verwendet. Jede Kombination wird für einige Zyklen berechnet und die Statistiken der vorläufigen rekonstruierten Phylogenien verglichen. Die Kombination, welche die besten statistischen Werte liefert (*Bayesian Information Criterion & Log Likelihood Gain*) wird für die Berechnung der phylogenetischen Abstände zwischen den Sequenzen sowie für die Konstruktion des Baumes verwendet.¹⁸⁴

Für diese Arbeit folgte daraus die Verwendung eines *LG Models*^{††} mit Gamma Verteilung als Substitutionsmatrix¹⁸⁵ unter teilweiser Beachtung der Lücken und der heuristischen *nearest-neighbor-interchange* (NNI) Methode, welche die Konnektivität der Äste eines vorgegebenen Baumes untereinander austauscht um die optimale Struktur mit den kürzesten Abständen zwischen den Gruppen zu errechnen¹⁸⁶. Die Beachtung der Lücken im Alignment spielt gerade bei Proteinen wie den Adhäsinen vom PA14-Typ eine wichtige Rolle, da hierdurch der Variabilität der Loop-Bereiche sowie vorhandener Subdomänen Rechnung getragen wird. Der initiale phylogenetische Baum wurde hierbei automatisch durch *MEGA6* auf Basis der

†† Nach den Autoren LE & GASCUEL

Neighbor-joining Methode erstellt.¹⁸⁷ Der finale Baum wurde mit der Software *FigTree* 1.4 dargestellt und für die weitere Aufarbeitung in einem Grafikprogramm vorbereitet.

Weiterführende Literatur zur Theorie und Praxis der phylogenetischen Analyse bieten SALEMI *et al.*, die praktische Durchführung in MEGA6 wurde von HALL zusammengefasst.^{188,189}

3.2 Molekularbiologische Methoden

Ein Teil der molekularbiologischen Grundlagen dieser Arbeit wurde bereits in vorhergehenden Arbeiten gelegt. Daher wird hier nur auf die für diese Arbeit relevanten Methoden näher eingegangen.

3.2.1 Kultivierung von *E. coli*

Kulturen von *E. coli*-Stämmen wurden mit *lysogeny broth*-Vollmedium¹⁹⁰ (LB) oder *Terrific-Broth*-Medium¹⁹¹ (TB) hergestellt. Festmedium wurde unter Zugabe von 2% Agar (w/v) erhalten.

Lysogeny broth		Terrific broth	
Bacto-Trypton	10 g/L	Bacto-Tryton	12 g/L
Bacto-Hefe-Extrakt	5 g/L	Bacto-Hefe-Extrakt	24 g/L
NaCl	10 g/L	Glycerin (99%)	4 mL
NaOH-Lösung (10 M)	0.4 mL/L	Phosphat-Puffer	1:10
Dest. H ₂ O	ad 1 L	Dest. H ₂ O	ad 1 L
Phosphat-Puffer (10x)			
		KH ₂ PO ₄	0.17 M
		K ₂ HPO ₄ x 3 H ₂ O	0.72 M

Die Medien wurden bei 121 °C und 2 bar 30 min lang sterilisiert, wobei der Phosphat-Puffer getrennt autoklaviert und nachfolgend dem Medium zugegeben wurde. Die Zugabe entsprechender Selektionsmarker erfolgt nach Abkühlung des Mediums auf etwa 55 °C. Das Wachstum angeimpfter Kulturen wurde durch Messung der optischen Dichte bei 595 nm verfolgt¹⁹². Als Selektionsmarker wurden Kanamycin (35 µg/mL; pET-28a(+)), Streptomycin (50 µg/mL; *E. coli* SHuffle T7) und Spectinomycin (12.5 µg/mL; *E. coli* SHuffle T7) verwendet.

3.2.2 Kultivierung von *S. cerevisiae*

S. cerevisiae-Stämme wurden in YEPD-Medium oder zur Selektion in SC-Medium kultiviert. YNB-If (*Yeast-Nitrogen-Base-low fluorescence*) wurde für Immunfluoreszenzassays benutzt. Festmedien wurden auch hier mit 2% (w/v) Agar hergestellt, die Sterilisation erfolgt ebenfalls durch Autoklavieren.

YEPD-Medium		SC-Medium	
Pepton	2% (w/v)	<i>Yeast Nitrogen Base</i> ohne Aminosäuren und (NH ₄) ₂ SO ₄	0.15% (w/v)
Hefe-Extrakt	1% (w/v)	(NH ₄) ₂ SO ₄	0.5% (w/v)
Glucose	2% (w/v)	<i>myo</i> -Inositol	0.2 mM
		Aminosäure-Mix	0.2% (w/v)
		<i>p</i> -Aminobenzoat	0.02% (w/v)

10x Salz-Lösung		Spurenelemente-Lösung	
KH ₂ PO ₄	10 g/L	H ₃ BO ₄	50 mg/100 mL
MgSO ₄	5 g/L	CuSO ₄	4 mg/100 mL
NaCl	1 g/L	KI	10 mg/100mL
CaCl ₂	1 g/L	FeCl ₃	20 mg/100 mL
(NH ₄) ₂ SO ₄	50 g/L	MnSO ₄	40 mg/100 mL
Dest. H ₂ O	ad 1 L	Na ₂ MoO ₄	20 mg/100 mL
		ZnSO ₄	40 mg/100 mL
		Dest. H ₂ O	ad 100 mL

Vitamin-Lösung		YNB-If	
Biotin	0.2 mg/100 mL	10x Salz-Lösung	100 mL/L
Calciumpantothenat	40 mg/100 mL	Spurenelemente-Lösung	0.1 mL/L
Inositol	200 mg/100mL	Vitamin-Lösung	0.1 mL/L
Niacin	40 mg/100 mL	Glucose (20% w/v)	100 mL/L
<i>p</i> -Aminobenzoensäure	20 mg/100 mL	Aminosäure-Mix	0.2% (w/v)
Pyridoxin-HCl	40 mg/100 mL		
Thiamin-HCl	40 mg/100 mL		
Dest. H ₂ O	ad 100 mL		

Der Aminosäuremix enthielt alle Standard-L-Aminosäuren abzüglich Histidin, Leucin und Tryptophan. Nach Inokulation wurden die Kulturen schüttelnd oder rotierend bei 30 °C inkubiert. Der Fortschritt des Wachstums wurde durch Bestimmung der optischen Dichte verfolgt. Hierbei entspricht eine OD₅₉₅=1 einer Anzahl von etwa 1.5·10⁷ Zellen/mL¹⁹².

3.2.3 Herstellung chemisch kompetenter *E. coli*-Zellen

Mit einer Übernachtskultur (ÜNK) des zu präparierenden *E. coli*-Stamms, welche den entsprechenden Selektionsmarker enthielt, wurden 4x 50 mL LB-Medium, ebenfalls mit Selektionsmarker versetzt, im Verhältnis 1:100 inokuliert. Diese Kulturen wurden bei 225 rpm und 37 °C bis zu einer OD₅₉₅ von 0.6 inkubiert. Die Zellen wurden mittels Zentrifugation bei 1500 g und 4 °C für 20 min geerntet. Nach Verwerfen des Überstandes wird das Zellpellet in 60 mL TFBII-Puffer resuspendiert und 2 h auf Eis inkubiert. Nach erneuter Zentrifugation über 10 min bei 4000 rpm und 4 °C wurde der Überstand verworfen und das Pellet in 4 mL TFBII-Puffer resuspendiert und die Suspension zu 50 µL aliquotiert und in flüssigem Stickstoff eingefroren. Die Lagerung der Zellen erfolgte nach MORRISON bei -80 °C¹⁹³.

TFBI-Puffer		TFBII-Puffer	
RbCl	100 mM	RbCl	10 mM
MnCl ₂	50 mM	MOPS	10 mM
KOAc	30 mM	CaCl ₂	75 mM
CaCl ₂	10 mM	Glycerin	15% (v/v)
Glycerin	15% (v/v)		

3.2.4 Transformation chemisch kompetenter *E. coli*-Zellen

Die Transformation der zuvor behandelten *E. coli*-Zellen erfolgte nach einem abgewandelten Hitzeschock-Protokoll nach MANDEL *et al.*¹⁹⁴ Hierfür wurde ein tiefgefrorenes Aliquot (50 µL) mit 1 µL Plasmidlösung (50-250 ng) versetzt und 5 bis 20 min auf Eis inkubiert. Die Zellen werden einem Hitzeschock von 40 s bei 42 °C im Wasserbad unterzogen. Die Probe wurde auf Eis mit 1 mL LB-Medium versetzt und 40 min bei 37 °C und 125 rpm schüttelnd inkubiert. Anschließend wurde 1 min bei 5000 rpm zentrifugiert und ca. $\frac{3}{4}$ des Überstandes verworfen. Die Zellen wurden im übrigen Medium resuspendiert, auf einer LB-Agar Platte mit entsprechendem Selektionsmarker ausgestrichen und über Nacht bei 37 °C im Brutschrank inkubiert.

3.2.5 Stammhaltung von *E. coli*

Mit unterschiedlichen Konstrukten transformierte *E. coli*-Stämme wurden zur Stammhaltung als Glycerinkulturen in Anlehnung an das Protokoll von MORRISON *et al.* präpariert¹⁹³. Hierzu wurden 500 µL einer Übernachtskultur mit 1 mL Glycerin (87%) versetzt und in flüssigem Stickstoff tiefgefroren. Die Lagerung erfolgte bei -80 °C.

3.2.6 Transformation von *S. cerevisiae* nach CHEN *et al.*

Zur Transformation von *S. cerevisiae* Zellen wurde ein verkürztes *One-Step* Protokoll nach CHEN *et al.* verwendet¹⁹⁵. Hierfür wurde eine ÜNK, welche sich in der stationären Phase befand, pelletiert und in 100 µL *One-step* Puffer resuspendiert. Es folgte die Zugabe von 0.1-1.0 µg Plasmid-DNA sowie 5 µL frisch denaturierte Lachsspermien-DNA (10 mg/mL). Nach einem Hitzeschock von 42 °C für ca. 25 min wurden 0.9 mL YEPD Medium hinzugefügt und die Zellen bei 3000 g für 1 min pelletiert. Etwa $\frac{3}{4}$ des Überstandes wurden verworfen, die Zellen im übrigen Medium resuspendiert und auf Agar-Platten mit entsprechendem Selektionsmarker ausgestrichen.

<i>One-Step</i> Puffer	
PEG 4000	40% (w/v)
DTT	100 mM
Lithiumacetat	200 mM
pH	5.0

3.2.7 Präparation von Plasmid-DNA aus *E. coli*

Um Plasmide aus *E. coli* Zellen zu präparieren wurde auf ein kommerziell erhältliches Kit der Firma *Qiagen* zurückgegriffen (*QIAprep Spin Miniprep Kit*). Dieses basiert auf der alkalischen Lyse von Zellen nach BIRNBOIM und DOLY sowie der Adsorption von Plasmid-DNA an Silika-Matrices nach VOGELSTEIN und GILLESPIE^{196,197}. 4 mL einer ÜNK wurden sequentiell in einem 2 mL Reaktionsgefäß pelletiert (1 min, 4000 rpm) und der Überstand verworfen. Das Zellpellet wurde in 250 µL Puffer P1 resuspendiert und nach Zugabe von 250 µL Puffer P2 4-6x invertiert. Die Suspension wurde mit 350 µL Puffer N3 versetzt und erneut invertiert. Nach Zentrifugation von 10 min bei 13000 rpm wurde der Überstand auf eine *QIAprep Spin* Säule gegeben, für weitere 60 s bei 13000 rpm zentrifugiert und der Durchfluss verworfen. Es folgte ein Waschschrift mit 750 µL PB-Puffer, erneute Zentrifugation bei 13000 rpm für 60 s sowie das Verwerfen des Durchflusses. Die Zentrifugation wurde zur Trocknung der Säule wiederholt. Als letzter Schritt erfolgte die Elution mit 35 µL EB Puffer oder bidest. H₂O bei 13000 rpm für 60 s. Die Plasmide wurden bei -20 °C gelagert. Das Protokoll ist ebenfalls dem Handbuch des Kits zu entnehmen¹⁹⁸.

3.2.8 Präparation von RNA aus Hefen

Zur Untersuchung der Expression von Adhäsinen in *Pichia pastoris* wurde auf die Methode der quantitativen Polymerasekettenreaktion zurückgegriffen. Hierfür musste die Gesamt-RNA aus Hefezellen präpariert werden. Bei der Arbeit mit RNA ist große Vorsicht und Sorgfalt geboten, da ubiquitär vorhandene RNAsen schnell zum Abbau der zu präparierenden RNA führen¹⁹⁹. Die Präparation wurde mit einem entsprechenden Kit der Firma *Qiagen* durchgeführt und basierte auf der Adhäsion von RNA >200 bp an Silika-Matrices. Die Extraktion und Isolation der RNA wurde auf Grundlage der Guanidiniumthiocyanat-Extraktion nach CHOMCZYNSKI *et al.* durchgeführt²⁰⁰. Um eine maximale Ausbeute an gereinigter RNA zu erhalten wurden $2 \cdot 10^6$ bis $5 \cdot 10^7$ Zellen verwendet (siehe 1.2.2.). Die Kultur, welche sich in der logarithmischen Wachstumsphase befand, wurde für 5 min bei 1000 g und 4 °C pelletiert. Der Überstand wurde sorgfältig abgesaugt und in 600 µL Guanidiniumthiocyanat-haltigem RLT-Puffer resuspendiert. Die Zellen wurden daraufhin mit Hilfe einer Glasmühle (Fast-prep 24, *MP Biomedicals*) aufgeschlossen. Dafür wurden etwa 600 µL säuregewaschene Glaskugeln mit einem Durchmesser von 0.25-0.5 mm zu den Zellen gegeben und der Aufschluss in 2x 30 s bei einer Geschwindigkeit von 6.0 m/s unter zwischenzeitlichem Kühlen durchgeführt. Das Lysat (ca. 350 µL) wurde in ein neues Reaktionsgefäß überführt und 2 min bei 13000 rpm zentrifugiert. Nach Überführung des Überstandes in ein weiteres Reaktionsgefäß wurde das gleiche Volumen an 70% Ethanol zum homogenisierten Lysat gegeben und mittels auf und ab pipettieren gemischt. Die Probe wurde auf eine *RNeasy spin* Säule gegeben und 15 s bei 13000 rpm zentrifugiert. Nachdem 700 µL Puffer RW1 auf die Säule gegeben wurden, erfolgte die Zentrifugation bei 13000 rpm für 15 s. Dieser Schritt

wurde 2x mit 500 μL RPE Puffer wiederholt. Die abschließende Elution erfolgte in 30-50 μL RNase-freies Wasser indem für 1 min bei 13000 rpm zentrifugiert wurde.

3.2.9 Konzentrationsbestimmung von DNA & RNA

Die Konzentration von DNA und RNA wurde photometrisch bestimmt²⁰¹. Hierbei wird die Absorption von DNA/RNA bei 260 nm genutzt. Weiterhin lässt sich durch Vergleich mit der Absorption bei 280 nm die Verunreinigung mit Proteinen bestimmen. Die Messungen erfolgten mit einem ND-1000 (*Nanodrop*). Mit folgenden Zusammenhängen lässt sich die Konzentration der Nukleinsäuren bestimmen:

$$c_{DNA} = A_{260} \cdot 50 \frac{ng}{\mu L}$$

$$c_{RNA} = A_{260} \cdot 40 \frac{ng}{\mu L}$$

Bei den Konzentrationsmessungen kann nicht zwischen der Absorption von DNA oder RNA unterscheiden werden, somit sind Mischlösungen nicht zu identifizieren.

3.2.10 Reverse Transkription mit anschließender q-PCR (qRT-PCR)

Die Stärke der Genexpression bestimmt, in welchem Maße der Genotyp einer Zelle bzw. eines Organismus phänotypisch in Erscheinung tritt. Die Analyse der Expression von Genen spielt somit in vielen Bereichen der Biowissenschaften eine wichtige Rolle; lassen sich hierbei doch Informationen entwicklungsbiologischen, zellbiologischen oder molekularbiologischen Inhalts gewinnen^{202,203}. Anwendung findet die Analyse der Genexpression bspw. in der Untersuchung geänderter Expressionsstärken bei unterschiedlichen Krankheitsbildern (Onkogen-Expression bei Krebs)²⁰⁴ und zur Identifikation von Virusinfektionen sowie der Kontrolle der Viruslast²⁰⁵. Das direkte Abbild der Expressionsstärke eines Gens stellt die relative Menge an mRNA-Transkripten dar. Der wichtigste Schritt in Richtung der Entwicklung neuer, günstiger, sowie hochdurchsatz-fähiger Techniken zur Untersuchung der Expressionsstärke von Genen gelang mit der Entwicklung der Polymerasekettenreaktion (eng. *polymerase chain reaction*, PCR)²⁰⁶. Für die Anwendung der PCR auf RNA-Template ist es nötig, diese zuvor in cDNA zu transkribieren. Hierfür werden reverse Transkriptasen verwendet^{207,208}. Der Einsatz von thermostabilen DNA-Polymerasen erlaubte erstmals die Durchführung einer sog. *One-Step-RT-PCR*, der reversen Transkription mit anschließender PCR in einem Ansatz²⁰⁹. Um eine Quantifizierung durchführen zu können, muss ein geeignetes Messsignal generiert werden. Dies gelingt bspw. mit interkalierenden Farbstoffen²¹⁰ wie SYBR® Green. Während eine absolute Quantifizierung eine numerische Anzahl an RNA-Transkripten pro Gen liefert, wird bei der relativen Quantifizierung die Expressionsstärke des Zielgens mit der Expression eines sog. Haushalts-Gens in Relation gesetzt^{211,212}. Bei diesem Haushalts-Gen, welches damit als interne Kontrolle genutzt wird, ist darauf zu achten, dass die Expression dieses Gens unabhängig von Zellzyklus oder exogenen Einflüssen ist. Hierfür eignet sich beispielsweise die

Gene von β -Actin, Tubulin oder der Glycerinaldehyd-3-phosphat-Dehydrogenase (GAPDH)²¹³. Die reverse Transkription sowie die quantitative PCR werden mit Zielgen und Haushaltsgen durchgeführt und der *crossing threshold* (C_T), welcher den Zyklus angibt, bei welchem die Fluoreszenz einen festgelegten Wert übersteigt, miteinander verglichen. Dieser soll in einem möglichst frühen Stadium der PCR-Reaktion festgelegt werden (exponentieller Bereich), da hier die Effizienz der Reaktion am Maximum liegt und die Substratkonzentration und damit probenabhängige Determinanten keine limitierenden Faktoren darstellen²¹². Der beispielhafte Verlauf einer qPCR-Reaktion mit eingezeichnetem C_T -Wert ist in Abbildung 3.1 dargestellt.

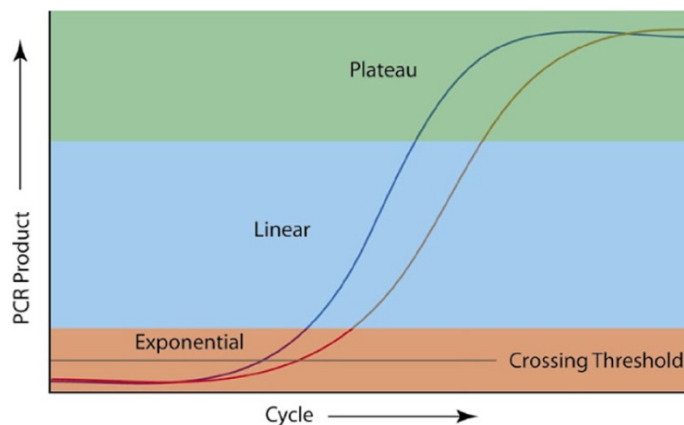


Abbildung 3.1 - Darstellung der unterschiedlichen Phasen einer PCR-Reaktion mit eingezeichnetem C_T -Wert.²¹⁴

Da im Rahmen dieser Arbeit eine *One-Step-qRT-PCR* durchgeführt wurde, also die reverse Transkription mit anschließender quantitativer PCR in einem Reaktionsansatz, musste besonderes Augenmerk auf die Generierung von Primern gelegt werden, da diese den Anforderungen beider Reaktionen gerecht werden mussten. Beispielsweise wird empfohlen für Zielgen und Haushaltsgen ähnliche Amplifikatgrößen zu nutzen, da die Stärke der entstehenden Fluoreszenz hiervon abhängt. Weitere Kriterien sind einer Zusammenfassung von THORNTON und BASU zu entnehmen²¹⁵.

In dieser Arbeit wurde die *One-Step-qRT-PCR* zur Untersuchung der Expression von Adhäsion-codierenden Genen in *P. pastoris* eingesetzt. Hierfür wurde das *qScriptTM One-Step SYBR Green qRT-PCR-Kit* der Firma Quanta Bioscience verwendet, welches eine verbesserte Variante der reversen Transkriptase des murinen Leukämievirus nutzt (MMLV). Der Vorteil dieses Kits besteht darin, dass die DNA-Polymerase zu Beginn durch einen Antikörper inaktiviert vorliegt und damit keine Störung der reversen Transkriptase erfolgt. Durch einen Hitzeshock von 95 °C, welcher in das PCR-Programm integriert wurde, löst sich der Antikörper von der Taq-Polymerase. Das Design der Primer wurde mit Hilfe des Programms *Primer3Plus* durchgeführt, welches entsprechende Optionen zur Erstellung von Primern für die qPCR enthält¹⁶⁰. Als Haushaltsgen im Stamm DSMZ70382 wurde β -Actin verwendet. Dieses wurde zuvor mittels des *BLAST* und dem Homolog aus GS115 identifiziert. Auf Eis wur-

de folgender Reaktionsansatz zusammen gegeben (Die Zugabe des RNA Templats aus 1.2.8. erfolgt zum Ende):

Komponente	Volumen	Endkonzentration
One-Step SYBR Green Master Mix	12.50 μL	1x
Forward Primer	0.62 μL	250 nM
Reverse Primer	0.62 μL	250 nM
RNA Templat	0.35 μL	280 ng/ μL
Reverse Transkriptase	0.50 μL	1x
Nuklease-freies Wasser	10.41 μL	
Endvolumen	25.00 μL	

Pro Gen wurden je drei Ansätze von jeweils drei unabhängigen RNA-Präparationen verwendet. Als Kontrolle wurde bei je einer Probe statt der reversen Transkriptase Wasser verwendet um Falschergebnisse aufgrund vorliegender DNA-Kontaminationen zu identifizieren. Zum Durchführen der Reaktion wurde ein *Rotorgene Q (Qiagen)*-Thermocycler verwendet und im Kanal SYBR-Green bei automatischer Wahl der Sensitivität aufgezeichnet. Um falschpositive Fluoreszenz durch Primer-Dimere oder unspezifische Produkte zu identifizieren, wurde am Ende der Reaktion eine Schmelzkurve der Amplifikate aufgenommen. Folgender Reaktionsverlauf wurde gewählt:

Schritt	Zeit	Temperatur
cDNA-Synthese (reverse Transkription)	10 min	50 °C
Taq-Aktivierung	5 min	95 °C
qPCR (45 Zyklen)	10 s	95 °C
	30 s	60 °C

3.3 Proteinchemische Methoden

3.3.1 Heterologe Produktion von Proteinen im präparativen Maßstab

Zur Produktion von Proteinen im präparativen Maßstab wurde der *E. coli*-Stamm SHuffle® T7 Express genutzt. Die Produktion von *PpFlo1A* und *Cea1A* erfolgte in TB-Medium, während *Epa1A* in LB Medium produziert wurde. Hierfür wurden aus der jeweiligen Glycerinkultur oder von einer Agar-Platte mit frischen Transformanten 3x 50 mL LB-Medium mit entsprechendem Selektionsmarker als Übernachtskultur angeimpft. Am nächsten Tag wurden die Kulturen vereinigt und je 2 L vorgewärmtes (37 °C) Medium mit Selektionsmarker in Schikanekolben mit 40 mL dieser Kultur inokuliert (Verhältnis 1:200). Die Kolben wurden bei 37 °C schüttelnd bei 150 rpm bis zu einer OD₆₀₀ von 0.2 inkubiert und anschließend die Temperatur auf 12 °C gesenkt. Nach 2 h und bei einer OD₆₀₀ von etwa 0.6 wurden die Kulturen mit 0.01 mM Isopropyl- β -D-thiogalactopyranosid (IPTG) versetzt und somit die Expression induziert. Die Kulturen wurden bei 12 °C für 72 h schüttelnd inkubiert und anschließend mittels Zentrifugation geerntet (JA 10, 6000 rpm, 4 °C, 15 min). Die Zellpellets

von 12 L wurden in etwa 25 mL AM-Puffer (Aufschlussmedium) resuspendiert, in flüssigem Stickstoff tiefgefroren und bei $-80\text{ }^{\circ}\text{C}$ gelagert. Die Produktion und Aufreinigung von Flo10A erfolgte durch NINA WOZNIAK¹⁵⁹, von Epa6A und Epa9A durch RIKE DIDERICH¹⁶⁵.

AM-Puffer	
Tris/HCl	100 mM
NaCl	200 mM
pH (HCl)	8.0

3.3.2 Zellaufschluss im präparativen Maßstab

Der Aufschluss der geernteten Zellen erfolgte mit Hilfe eines Hochdruckhomogenisators (*French Press*). Die Zellen werden hierbei einem hohen Druck ausgesetzt und schlagartig entspannt. Die dabei entstehenden Scherkräfte führen zur Lyse der Zellen. Die tiefgefrorene Zellsuspension wurde dazu auf Eis aufgetaut und mit 20 μL Lysozym-Lösung (50 mg/ml), 5 μL EDTA-Lösung (100 mM), 5 μL PMSF-Lösung (200 mM) sowie einer Spatelspitze Desoxyribonuklease I (DNase) versetzt. Die Hochdruckzelle ($4\text{ }^{\circ}\text{C}$) wurde nach etwa 15 min mit der Suspension befüllt und die Zellen in 2-3 Durchläufen bei einem Druck von 900 psi lytiert. Während des gesamten Vorgangs wurden die Zellen auf Eis gekühlt. Mittels Zentrifugation wurden löslicher und unlöslicher Teil voneinander getrennt (JA 20, 18000 rpm, $4\text{ }^{\circ}\text{C}$, 40 min). Der Überstand wurde in ein frisches Gefäß überführt und mit Hilfe eines Spritzenvorsatzfilters (0.45 μm) filtriert.

3.3.3 Ni-NTA-Affinitätschromatografie

Wie bereits in Abschnitt 2.9.1 beschrieben wurden mittels des Vektors pET-28a(+) Fusionsproteine mit einem *N*-terminalen Hexa-Histidin-*Tag* hergestellt. Diese Technik wurde bereits 1975 von PORATH verwendet um Proteine mittels Chelatisierung an Metallionen zu isolieren. Die Fusionsproteine können so in einer Säule aus geeignetem Material an immobilisierte Metallionen binden während andere Proteine die Säule ungehindert passieren²¹⁶. Die Elution des gebundenen Proteins erfolgt durch Zugabe des Histidin-Strukturanalogons Imidazol in erhöhter Konzentration, welches das Fusionsprotein von der Säule verdrängt. Im Rahmen dieser Arbeit wurde die Komplexbildung zwischen Nickel-Ionen und dem Hexa-Histidin-*Tag* ausgenutzt um produzierte Proteine zu isolieren und aufzureinigen. Die verwendeten His-Trap Ni-NTA-Säulen (1 mL oder 5 mL Säulenvolumen, *Macherey Nagel*) enthalten Ni^{2+} -Ionen, welche als Komplex an Nitrilotriessigsäure (NTA) gebunden sind. Diese ist wiederum an einer Gelmatrix aus Sepharose 6 Fast Flow Material verankert.

Der zuvor steril filtrierte, lösliche Teil aus 3.3.2 wurde bei $4\text{ }^{\circ}\text{C}$ über eine Peristaltikpumpe auf die bereits mit AM-Puffer äquilibrierte Säule aufgetragen und diese an ein FPLC-System angeschlossen (ebenfalls $4\text{ }^{\circ}\text{C}$). Bei einer Flussrate von 0.5 mL/min wurde mit einer Konzent-

ration von 15 mM Imidazol in AM-Puffer (AMI-Puffer) unspezifisch gebundene Proteine von der Säule gewaschen um den Reinigungseffekt zu erhöhen. Die Elution des Fusionsproteins erfolgte bei Imidazolkonzentrationen von 30-100 mM in AM-Puffer. Die Fraktionen wurden in ihrem Volumen entsprechenden Gefäßen aufgefangen und die Säule bei 500 mM Imidazol gewaschen. Nach Spülen mit dest. Wasser und 20% Ethanol wurde die Säule bei 4 °C gelagert. Um die Säule nach häufigem Gebrauch zu regenerieren wurde auf ein Protokoll der Firma *Qiagen* zurückgegriffen²¹⁷.

3.3.4 Größenausschlusschromatografie

Weit vor der Verwendung von Isolationsverfahren auf Basis der Wechselwirkung von Fusionsproteinen mit Metallionen publizierte PORATH eine Methode zur Reinigung und Isolation von Proteinen auf Grundlage ihrer apparenten Größe²¹⁸. Hierbei steht Proteinen mit einem kleineren hydrodynamischen Radius beim Passieren eines porösen Säulenmaterials ein größeres Diffusionsvolumen zur Verfügung als Voluminöseren. Durch das größere Diffusionsvolumen ist die Retentionszeit kleinerer Moleküle im Vergleich zu Proteinen, deren Größe über dem Ausschlussvolumen der Poren liegen, erhöht.

Als stationäre Phase wurde in dieser Arbeit eine Säule mit 120 mL Volumen aus Superdex 200 Material verwendet (*GE Healthcare*), welches ein Kompositmaterial aus verzweigter Agarose und Dextran ist. Vor Benutzung wurde die Säule mit 3 Säulenvolumina SEC-Puffer (*Size-Exclusion*) für Cea1A oder AME-Puffer (Aufschlussmedium mit EDTA) für Epa1A äquilibriert und Proteinlösung mit einer Konzentration von 30-60 mg/mL aufgetragen. Die Gelfiltrationsläufe wurden bei einem Durchfluss von 1 mL/min durchgeführt, Fraktionen in 1 mL aufgefangen und deren Reinheit mittels SDS-PAGE überprüft.

SEC-Puffer		AME-Puffer	
Tris/HCl	20 mM	Tris/HCl	100 mM
NaCl	200 mM	NaCl	200 mM
pH (HCl)	8.0	pH (HCl)	8.0
		EDTA	10 mM

3.3.5 Konzentration von Proteinen durch Ultrafiltration

Um die Konzentration von Proteinlösungen zu erhöhen, wurden diese in Ultra-Konzentratoren (*MerckMillipore*) zentrifugiert. Da die Adhäsion-A-Domänen eine Masse von etwa 25-30 kDa besitzen, wurden Filter mit einer Ausschlussgröße von 10 kDa verwendet. Um die Filtermembran zu äquilibrieren wurde der Filter mit dem entsprechenden Puffer gefüllt und zentrifugiert (4000 rpm, 4 °C, 10 min). Anschließend wurde die Proteinprobe in den Konzentrator gegeben und sequentiell mit zwischenzeitlichem auf und ab pipettieren zentrifugiert, (4000 rpm, 4 °C, 2-4 min) bis das gewünschte Volumen erreicht war.

3.3.6 Pufferwechsel

Die Verwendung des Proteins in mehreren analytischen Methoden setzte einen Wechsel des Puffersystems voraus. Hierfür wurde eine spezielle Größenausschlusschromatografie genutzt, bei welcher Säulen mit einem sehr kleinen Ausschlussvolumen (5 kDa) verwendet wurden. Dies führt dazu dass alle untersuchten Proteine im Ausschlussvolumen in der neuen mobilen Phase eluieren. Neben PD-10 Säulen wurden kleinere PD MiniTrap G25-Säulen eingesetzt (*GE Healthcare*), welche beide mittels Gravitationsfluss betrieben wurden. Nach Äquilibration mit 3 Säulenvolumina des Zielpuffers wurden 2.5 mL (PD-10 Säule) bzw. 0.5 mL (G-25 Säule) Proteinlösung aufgetragen und mit 3.5 mL respektive 1 mL eluiert.

3.4 Analytische Methoden

3.4.1 SDS-PAGE

Bei der SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE) nach LAEMMLI werden Proteine anhand ihrer apparenten Masse im elektrischen Feld mit Hilfe eines Gels aus Polyacrylamid aufgetrennt²¹⁹. Proteine werden unter Zugabe des Detergenz Natriumdodecylsulfat (SDS), β -Mercaptoethanol zum reduzieren vorhandener Disulfidbrücken sowie hoher Temperatur (95 °C) denaturiert. Ihre Eigenladung wird hierbei durch Anlagerung des SDS an das somit linearisierte Protein überdeckt und es entsteht ein nahezu konstantes Masse/Ladungs-Verhältnis. In einem Netz aus polymerisiertem Acrylamid, welches mit Bisacrylamid quervernetzt ist, werden die Proteine bei Anlegen eines elektrischen Feldes anhand ihres Ladungs-/Gyrationsradius-Verhältnisses aufgetrennt und mittels Anfärben im Gel sichtbar gemacht. Der Gehalt an Acrylamid im Trenngel bestimmt die Trennleistung, da mit höherem Acrylamidgehalt die Maschen des Polymers kleiner werden (15% für 10-60 kDa, 10% für 30-120 kDa, 8% für 50-200 kDa).

Für die diskontinuierliche Gelelektrophorese wurden Mighty Small II-Kammern (*Hoefler*) genutzt; 11 Gele mit einem Acrylamidgehalt von 4% im Sammelgel und 12% im Trenngel konnten parallel in einem Caster gegossen werden. Die zu untersuchenden Proteinproben wurden mit 2x SDS-Probenpuffer versetzt und bei 95 °C für 5 min erhitzt. Die Elektrophorese erfolgte bei einer Stromstärke von 20-25 mA pro Gel und wurde beendet sobald die Lauffront das Ende des Gels erreichte. Als Massenstandard wurde eine Mischung von Proteinstandards aufgetragen (*PeQLab*). Nach Trennung von Sammel- und Trenngel wurde letzteres durch Schwenken in erhitzter Färbelösung gefärbt und anschließend der Hintergrund durch sukzessives Erhitzen und Schwenken in Entfärbelösung entfärbt²²⁰. Folgendes Rezept bezieht sich auf die Herstellung von 11 Gelen in einem Caster.

SDS-Gel 12%	Trenngel	Sammelgel
Acrylamid (30%)	24 mL	3.33 mL
SDS-Lösung (10% w/v)	600 µL	250 µL
Trenngelpuffer	20 mL	
Sammelgelpuffer		5 mL
TMEDA	60 µL	250 µL
APS (10% w/v)	300 µL	75 µL
Dest. H ₂ O	15 mL	16 mL

Sammelgelpuffer		Trenngelpuffer	
Tris/HCl, pH 6.8	625 mM	Tris/HCl, pH 8.8	1.125 M
		Saccharose	30% (w/v)

10x SDS-Laufpuffer		2x SDS-Probenpuffer	
Tris/HCl, pH 8.8	1.125 M	Tris/HCl, pH 6.8	62.5 mM
Saccharose	30% (w/v)	Glycerin	15% (v/v)
		SDS	4% (w/v)
		Bromphenolblau	Spatelspitze
		β-Mercaptoethanol	4% (v/v)

3.4.2 Konzentrationsbestimmung von Proteinlösungen

Die Konzentration von Proteinen wurde photometrisch bestimmt. Diese Methode wurde bereits 1941 von WARBURG und CHRISTIAN verwendet und stetig weiterentwickelt²²¹. Sie beruht auf der Absorption von elektromagnetischer Strahlung mit einer Wellenlänge von 280 nm. Diese wird in Proteinen hauptsächlich durch aromatische Aminosäuren wie Tyrosin und Tryptophan sowie durch in Disulfidbrücken gebundenen Cysteinen verursacht. Diese Eigenschaft erlaubt die Konzentrationsbestimmung von Proteinen in Lösung mit Hilfe des LAMBERT-BEER'schen Gesetzes²²². Der Server *ProtParam* diente zur Berechnung der theoretischen molaren Extinktionskoeffizienten²²³.

$$E = \varepsilon \cdot c \cdot d$$

$$E = \frac{\varepsilon \cdot m}{(V \cdot M) \cdot d}$$

$$c_M = \frac{E \cdot M}{\varepsilon \cdot d}$$

E: Extinktion; ε : molarer Extinktionskoeffizient; m: Masse; c: Stoffmengenkonzentration; d: Schichtdicke; V: Volumen; M: Molekulargewicht; c_M : Massenkonzentration

Protein	$\epsilon / \text{M}^{-1}\cdot\text{cm}^{-1}$	M / g·M ⁻¹
Cea1A	28100	26681.5
Epa1A	41175	29401.6
Epa6A	48165	29909.0
Epa9A	54125	33203.5
<i>PpFlo1A</i>	28100	27245.8

3.4.3 Fluoreszenzmarkierung mit AlexaFluor-488-TFP

Sowohl für fluoreszenzmikroskopische Untersuchungen als auch für *Glycan-Array*-Studien wurden Adhäsion-A-Domänen fluoreszenzmarkiert. Hierfür wurde ein Tetrafluorphenylester des Fluoreszenzfarbstoffes AlexaFluor 488 verwendet (*Invitrogen*). Farbstoffe der AlexaFluor^{##}-Familie sind in der biotechnologischen Forschung weit verbreitet und werden durch Sulfonierung von bspw. Cumarin, Fluorescein und weiteren Farbstoffen gewonnen²²⁴. AlexaFluor 488 hat eine molekulare Masse von 643 g/mol und das Absorptionsmaximum liegt bei 495 nm. Die maximale Emissionsintensität wird bei 519 nm erreicht²²⁵. Das hier verwendete Derivat bildet mit primären Aminen, wie sie in Proteinen am *N*-Terminus und in unterschiedlichen Aminosäuren vorkommen, ein Amid.

Da die Proteine in Tris-haltigem Puffer vorlagen und Tris selbst ein primäres Amin enthält, wurde die Probe zuvor über eine G-25-Säule umgepuffert. Dazu wurde ein Puffer gewählt, welcher weiterhin sicherstellen sollte das bspw. ein großer Teil der ϵ -Amine von Lysinen deprotoniert vorlagen ($\text{pK}_a \sim 10$) um die Konjugation mit AlexaFluor 488 auch an diesen durchführen zu können. Hierfür wurde ein NaHCO_3 -Puffer mit pH 8.3 verwendet.

Die Reaktion wurde mit 0.5 mL einer Lösung des zu markierenden Proteins (5 mg/mL) in einem Thermomixer bei 200 rpm und 20 °C unter Zugabe von 25 μL Farbstoff (10 $\mu\text{g}/\mu\text{L}$ in DMF) durchgeführt. Nach 1 h erfolgte die Abtrennung des unreaktierten Farbstoffs über eine G-25-Säule unter gleichzeitigem Pufferwechsel in *SEC*-Puffer. Der Erfolg der Reaktion ließ sich mittels SDS-PAGE und anschließender Betrachtung des Gels auf einem UV-Schirm überprüfen.

3.4.4 Glycan Array Analyse durch das CFG

Zur Untersuchung der Spezifität der PA14-Domänen enthaltenden Adhäsine wurden *Glycan-Array-Screens* in Kooperation mit dem *Consortium for Functional Glycomics* (CFG) durchgeführt. Hierbei wird fluoreszenzmarkiertes Protein auf einen gedruckten Glas-Chip gegeben, auf welchem 611 unterschiedliche, synthetische Glycane gebunden sind (*Mammalian Screen V 5.1*)²²⁶. Diese unterscheiden sich in ihrer Länge, Verknüpfung oder schlicht in der Art der Kohlenhydrate (Glucose, Mannose etc.). Mit einem Fluoreszenzfarbstoff konjugiertes Protein kann somit auf diesen Chip aufgetragen werden und nach Waschen des Chips

^{##} Nach ALEX HAUGLAND, dem Sohn der Gründer von *Molecular Probes*, welche die Farbstoffe entwickelten.

die resultierende Fluoreszenz gemessen werden. Die Quantifizierung der Fluoreszenz von 6 Messwerten pro Kohlenhydratstruktur führt zur Ableitung eines spezifischen Erkennungsmusters des untersuchten Proteins.

Das nach 3.4.3 markierte Protein wurde in einer Konzentration von ca. 2 mg/mL eisgekühlt zum CFG geschickt, wo dieses mit einem Laufpuffer auf 200 µg/mL und 20 µg/mL verdünnt wird. Die Lösungen wurden 2 h auf dem Chip inkubiert, der Chip gewaschen und die resultierende Fluoreszenz aufgezeichnet und quantifiziert (Request #2555). Als Positivkontrolle wurde Streptavidin verwendet. Das Vorgehen wurde außerdem detailliert von HEIMBURG-MOLINARO *et al.* beschrieben²²⁷.

Glycan-Array-Laufpuffer	
Tris/HCl	20 mM
NaCl	150 mM
CaCl ₂	2 mM
MgCl ₂	2 mM
Tween 20	0.05% (v/v)
BSA	1% (m/v)
pH (HCl)	7.4

3.4.5 Isothermale Titrationskalorimetrie - ITC

Die Kalorimetrie wird bereits seit langer Zeit verwendet, um Informationen zur Thermodynamik chemischer Reaktionen zu gewinnen und ist eine der ältesten physikochemischen Methoden²²⁸. Hierbei wird die Wärmeenergie gemessen, welche während einer Reaktion frei oder konsumiert wird, um diese in direkten Zusammenhang mit der Enthalpie ΔH und damit der freien Energie ΔG zu bringen. Die stetige Optimierung dieser Methode führte zur Entwicklung von drei Gruppen von Instrumenten: Batch-Kalorimeter, Durchflusskalorimeter und Titrationskalorimeter. Zur Untersuchung von Protein/Ligand-Wechselwirkung dienen heutzutage hauptsächlich Geräte letzteren Typs, welche auf Grundlage der Arbeiten von BRANDTS *et al.* und WISEMAN *et al.*^{229,230} entwickelt wurden. Diese basieren auf der direkten Messung der Wärme-Menge beim Mischen zweier Komponenten mittels Titration.

Typischerweise werden etwa 1-10 µL eines Liganden zu 200 µL-1000 µL des Bindungspartners gegeben. Die stattfindende Reaktion, bspw. die Bindung eines Proteins an ein Kleinmolekül, erzeugt oder konsumiert Wärme (q) proportional zur Menge des in der jeweiligen Injektion zugegebenen Kleinmoleküls ($v \cdot \Delta L$) und zur Enthalpie der Bindung (ΔH). Es folgt:

$$q = v \cdot \Delta H \cdot \Delta L$$

Durch einen Feedback-Mechanismus nach dem Wärme-Kompensations-Prinzip hält das Gerät die Temperaturdifferenz zwischen Mess- und Referenzzelle konstant. Das Messsignal entspricht somit der Energie, welche aufgewandt werden muss, um diese Differenz sicherzustellen.

len (Abbildung 3.2). Mit steigender Ligandenkonzentration nimmt das Messsignal aufgrund der Absättigung der Bindung ab. Nach Sättigung entstehen nur noch kleine Injektionspeaks, welche der Verdünnungswärme des Liganden entsprechen. Diese ist von der Messung zu subtrahieren. Die Wärme der einzelnen Injektionen in Form einer Bindungsisotherme erhält man durch Integration der Kurve, welche aus dem zeitlichen Verlauf der Injektionen und dem Messsignal besteht.²³¹

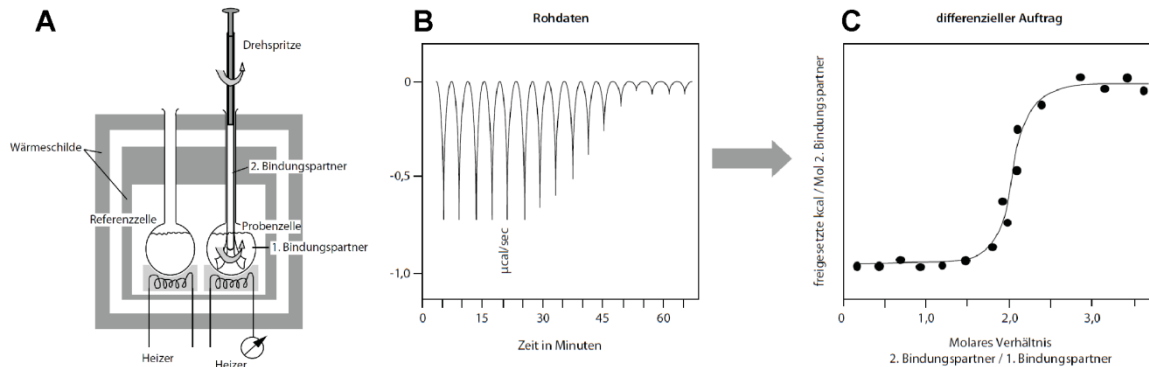


Abbildung 3.2 – A: Schematische Darstellung eines ITC-Geräts. B: Rohdaten eines typischen ITC-Experiments, ein Peak entspricht der freigesetzten Wärmeenergie der jeweiligen Injektion. C: Die Integration der Messung aus B liefert die sigmoide Bindungsisotherme. Abbildung nach REHM und LETZEL²³².

Folgende Gleichung gilt für ein Modell in welchem ein Protein mit einer Bindungstasche angenommen wird²³¹:

$$q_i = v \cdot \Delta H \cdot [P] \cdot \left(\frac{K_a [L]_i}{1 + K_a [L]_i} - \frac{K_a [L]_{i-1}}{1 + K_a [L]_{i-1}} \right)$$

q_i : Wärmemenge der Injektion i ; ΔH : Enthalpie; $[P]$: Konzentration des Proteins; K_a : Bindungskonstante; $[L]_i$: Konzentration des Liganden nach Injektion i

Diese Gleichung wird als nicht-linearer Fit numerisch an die Bindungsisotherme angepasst. WISEMAN *et al.* zeigten 1989, dass um eine Bindungsreaktion direkt messen zu können K_a einen Bereich von etwa 10^8 - 10^9 M^{-1} nicht überschreiten darf²³⁰. Weiterhin stößt die Methode bei sehr geringen Affinitäten an ihre Grenzen, da im mM (K_d) Bereich die Titrationskurve ebenfalls ihre charakteristische Form verliert. Diese Limitationen können durch Konkurrenzexperimente umgangen werden. Hierbei werden ein Protein und ein schwach gebundener Liganden mit einem sehr affinen Liganden titriert und umgekehrt^{233,234}.

Die aus obiger Gleichung erhaltenen Werte für die Enthalpie und die Bindungskonstante erlauben nun die Berechnung der freien Energie (ΔG):

$$\Delta G = -RT \cdot \ln K_a$$

ΔG : freie Enthalpie; R : Universelle Gaskonstante $8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$; T : Temperatur; K_a : Affinitätskonstante

Dies erlaubt wiederum die Kalkulation der Entropie über:

$$\Delta G = \Delta H - T\Delta S$$

ΔG : freie Energie; ΔH : Enthalpie; T: Temperatur; ΔS : Entropie

Die so erhaltenen thermodynamischen Daten geben Aufschluss über die Art der Wechselwirkung zwischen Protein und Ligand und damit über die chemische und strukturelle Umgebung der Bindungstasche.

Im Rahmen dieser Arbeit wurde ein ITC₂₀₀ der Firma *MicroCal* genutzt. Dieses ist mit 200 μL fassenden Zellen (Mess- & Referenzzelle) sowie einer Titrationsspritze mit 40 μL Volumen ausgestattet. Vor Beginn der Messung wurden beide Zellen gründlich durch Spülen mit 2% SDS-Lösung sowie dest. H_2O gereinigt. Eine Funktionskontrolle des Geräts erfolgte mittels der Funktion *y-axis calibration check*, einer Titration von $\text{H}_2\text{O}/\text{H}_2\text{O}$ sowie einer Titration von CaCl_2 (5 mM) mit EDTA (0.4 mM) in MES Puffer (10 mM MES, pH 5.6).

Für die Messung wurden 200 μL H_2O in der Referenzzelle, Proteinlösungen im mittleren mikromolaren Bereich ($\sim 450 \mu\text{M}$ in SEC-Puffer) in der Probenzelle, sowie Ligand im niedrigen millimolaren Bereich ($\sim 5 \text{ mM}$) verwendet. Wichtig ist hierbei der Parameter c , welcher die Konzentration des zu untersuchenden Proteins [P] in der Messzelle in Relation zur Assoziationskonstante K_a setzt ($c \equiv K_a \times [P]$). Idealerweise soll dieser einen Wert von 1000 nicht überschreiten²³¹. Außerdem ist die möglichst genaue Kenntnis der Konzentrationswerte für die Berechnung der Ausgleichskurve der Bindungsisotherme nötig, da hieraus direkt die thermodynamischen Parameter gewonnen werden. ITC-Messungen wurden für Cea1A und Epa6A mit unterschiedlichen Liganden durchgeführt. Hierfür wurden hauptsächlich folgende Parameter gewählt:

Parameter für ITC-Experimente	
# Injektionen	20
Temperatur	10 °C
Referenzenergie	10 $\mu\text{cal/s}$
Startverzögerung	60 s
Volumen 1. Injektion	0.4 μL
Dauer 1. Injektion	0.8 s
Zeit zwischen Inj.	360 s
Filter	5 s
Volumen übrige Inj.	Je 2 μL
Dauer übrige Inj.	Je 4 s

3.5 Bestimmung der dreidimensionalen Struktur von Proteinen

Die Möglichkeit, durch Strukturbestimmung von Proteinen und Nukleinsäuren, einen detaillierten Einblick in die molekularen Vorgänge biochemischer und molekularbiologischer

Prozesse zu erlangen, hat die Wissenschaft weitreichend beeinflusst. Zurzeit stehen mehrere Techniken zur Verfügung, um strukturelle Informationen über biologische Makromoleküle zu erhalten. Neben der Röntgenstrukturanalyse, welche sich die Beugung von harter Röntgenstrahlung an Kristallen zu Nutze macht, wie dies bspw. zuerst mit den ersten Diffraktionsbildern eines Enzyms (Pepsin) von BERNAL und CROWFOOT²³⁵ und der Struktur von Myoglobin durch KENDREW gezeigt wurde²³⁶, steht zur Strukturlösung mit atomarer Auflösung noch die Kernresonanzspektroskopie (NMR) zur Verfügung. Diese wurde hauptsächlich von RICHARD ERNST und KURT WÜTHRICH an Proteinen etabliert²³⁷. Weitere Methoden um strukturelle Informationen mit geringerer Auflösung zu erhalten sind die Kleinwinkelstreuung an Proteinen in Lösung (SAXS, *small angle X-ray scattering*)²³⁸ oder die Cryoelektronenmikroskopie²³⁹.

Während letztere Methoden hauptsächlich der Untersuchung der Quartärstruktur großer Proteine/Komplexe sowie von Protein/Protein-Interaktionen dienen, werden die Kristallstrukturanalyse und NMR genutzt um Informationen auf atomarer Ebene zu erhalten. Bei der NMR werden Proteine Spin-gelabelt (¹⁵N- und ¹³C-Isotope) und mehrdimensionale Spektren (hauptsächlich COSY und NOESY) aufgenommen. Die Kopplung der Kerne in Kombination mit entsprechenden *restraints* (bspw. Bindungslängen) liefert dadurch ein Ensemble an Strukturen. Hierdurch wird auch der Mobilität von Strukturbereichen oder Domänen Rechnung getragen. Der limitierende Faktor der Methode besteht in der Größenrestriktion auf etwa 20-30 kDa, da bei größeren Proteinen oder Komplexen die Zuordnung von Kopplungen (*i*) schwieriger wird und (*ii*) sich Peaks zu überlappen beginnen²⁴⁰. Neben der Möglichkeit Strukturen in Lösung untersuchen zu können bietet die NMR im Vergleich zur Kristallstrukturanalyse eine einfachere Automatisierung²⁴¹.

Da für das detaillierte mechanistische Verständnis biochemischer Prozesse ein Einblick in die atomare Umgebung bspw. einer Bindungstasche von großem Interesse ist, stellt die Methode der Kristallstrukturanalyse von Proteinen auch heutzutage, fast 50 Jahre nach KENDREWs Myoglobin-Struktur, noch den „Goldstandard“ der Strukturbestimmung dar²⁴². Dies wird ebenfalls deutlich, wenn man die Anzahl der in der *Worldwide Protein Data Bank* (*wwPDB*) hinterlegten Strukturen vergleicht²⁴³. So stehen 93738 Strukturen der Kristallstrukturanalyse 10721 Strukturen aus NMR-Experimenten sowie 711 Strukturen der Elektronenmikroskopie gegenüber. Die kürzlich vom *European Molecular Biology Laboratory* (EMBL) etablierte Datenbank für Strukturen aus SAXS-Experimenten, SASDB, enthält 178 Modelle^{244§§}. Neben der hohen Auflösung, bei welcher bisher die von SCHMIDT *et al.* erhaltene als höchst aufgelöste Struktur (0.48 Å) gilt²⁴⁵, ist ein weiterer Vorteil, das keine Markierung des Proteins erforderlich ist.

§§ Alle Stand 09.01.2015

Häufig wird angeführt, dass Proteine *in crystallo* nicht in nativer Umgebung vorlägen, doch konnte gezeigt werden, dass durch den hohen Lösungsmittelgehalt von etwa 50% nur eine verhältnismäßige geringe Kontaktfläche zwischen den Proteinen vorliegt²⁴⁶. Nichtsdestotrotz ist die Flexibilität der Proteine im Kristall eingeschränkt, so dass eine Beeinflussung bspw. der Erkennung von Substraten nicht ausgeschlossen werden kann^{247,248}.

Den Vorteilen der Kristallstrukturanalyse steht die Erzeugung von Einkristallen mit genügender Streuqualität gegenüber. Da die chemischen Bedingungen, welche für die Kristallisation von Proteinen nötig sind nur unzureichend rational abgeleitet werden können, stellt die Gewinnung geeigneter Kristalle den kritischsten Schritt dieser Methode dar^{249,250}.

3.5.1 Kristallisationsscreening

Wird die Löslichkeit von Proteinen bei Zugabe von Präzipitantien langsam herabgesetzt, können diese Kristalle bilden. Das Ziel solcher Experimente ist es, etwa 40 Proteinmoleküle dazu zu bringen sich zu einem Nukleus zu formieren, welcher dann zu einem makroskopischen Kristall mit etwa 10^{15} Molekülen heranwachsen kann²⁴². Da es nicht möglich ist, Kristallisationsbedingungen für ein Protein rational aus dessen Primärsequenz abzuleiten, wird in einem ersten Schritt eine große Anzahl von Bedingungen, welche bereits häufig zur Bildung von Kristallen führten, getestet. Diese empirischen, sog. *sparse-matrix-screens*, folgen der TAGUCHI-Methodik. Ein Experiment, welches von k Faktoren abhängt und dieser Faktor n mal variiert werden soll, benötigt $E=n^k$ individuelle Experimente. Nach TAGUCHI lässt sich diese enorme Anzahl ohne signifikanten Verlust an Informationen auf $E=2k+1$ Experimente reduzieren²⁵¹. Das große Interesse an strukturellen Informationen führte zur Einrichtung strukturgénomischer Konsortien, was die weitere Verbesserung vorhandener *sparse-matrix-screens* zur Folge hatte²⁵²⁻²⁵⁴.

Im Rahmen dieser Arbeit wurden kommerzielle Screens der Firma Qiagen genutzt, welche zum Teil in Zusammenarbeit mit dem *Joint Center for Structural Genomics* (JCSG) entwickelt wurden:

Screen	Informationen
<i>JCSG Core I-IV</i>	Optimierte Literaturbedingungen
<i>JCSG+</i>	Vorgänger der <i>Core I-IV</i> Screens
<i>PACT</i>	Systematischer Screen: pH-, Anionen-, Kationen-Variationen
<i>Classics, Classics light</i>	Literaturbekannte Lösungen
<i>Cryo</i>	Literaturbedingungen mit Kryo-Protektion
<i>Anions, Cations</i>	Systematische Ionen-Screens: pH-, Konzentrationsvariationen
<i>PEGs</i>	Verschiedene Polyethylenglycole als Präzipitants
<i>AmSO4</i>	Systematischer Ammoniumsulfat-Screen
<i>MbClass, MbClass II</i>	Literaturbedingungen vorzugsweise für Membranproteine

Die Kristallisationsexperimente mit diesen Screens wurden in versiegelbaren *96well*-Platten im *sitting drop*-Dampffusionsverfahren durchgeführt. Hierbei wird eine geringe Menge kon-

zentrierter Proteinlösung 1:1 mit der Kristallisationsbedingung vermischt und die Platte luftdicht verschlossen. Während der Inkubationszeit steht so der Proteintropfen mit einem Reservoir der Kristallisationsbedingung über die Dampfphase in Kontakt. Aufgrund des Konzentrationsgradienten zwischen beiden Lösungen kommt es zur Dampfdiffusion des Lösungsmittels vom Proteintropfen zum Reservoir. Hierbei erhöht sich die Konzentration des Proteins im Idealfall soweit, dass ein labiler übersättigter Bereich erreicht wird, in welchem sich Kristallnuklei bilden. Der Übergang in einen Bereich mit geringerer Proteinkonzentration in Lösung, den metastabilen Übersättigungsbereich, führt dann ausschließlich zum Wachstum der Kristalle, neue Nuklei werden nicht gebildet^{255,256}.

Für diese Arbeit wurden mit Hilfe eines Pipettierrobotors (Cartesian Microsys SQ 4004 oder Honeybee 963, beide *Genomic Solutions*) 300 nL einer Proteinlösung mit 20-60 mg/mL mit 300 nL der Präzipitanslösung in 96well Innovaplates (*Innovadyne*) gemischt. Da jedes well mit zwei Vertiefungen versehen war, wurde zusätzlich entweder Puffer als Negativkontrolle oder einer Proteinlösung mit halber Konzentration genutzt. Das Reservoir wurde mit 80 µL der Präzipitanslösung gefüllt, die Platte mit einer durchsichtigen Kunststoffolie (VIEWseal, *Greiner BIOone*) verschlossen und diese bei 18 °C und/oder 4 °C gelagert. Die Dokumentation erfolgte mit Hilfe eines Lichtmikroskops bei entsprechender Temperatur oder mit einem Rockimager-Dokumentationssystem (*Formulatrix*). Die Kokristallisationsexperimente mit Kohlenhydratliganden wurde durch Zugabe von 5 mM des jeweiligen Liganden und 5 mM CaCl₂ durchgeführt.

3.5.2 Optimierung von Kristallisationsbedingungen

Die Qualität der Kristalle, welche durch *sparse-matrix-screening* erhalten werden, ist meist nicht ausreichend um sie direkt zur Strukturlösung einzusetzen. Hierbei spielen vor allem die meist zu geringe Größe, zu große Fehlordnung im Kristall (Mosaizität) aber auch die Morphologie der Kristalle eine wichtige Rolle. Diese liegen hier häufig noch nicht als Einkristalle vor^{257,258}. In Einzelfällen konnte aber bereits aus dem *sparse-matrix-screen* Kristalle mit ausreichender Qualität gewonnen werden.

Wurde in einer Bedingung aus dem vorhergehenden Screening Kristallwachstum beobachtet, konnte eine Optimierung dieser erfolgen. Dazu wurden die chemischen Determinanten, wie der pH-Wert (Schritte von 0.1-0.2 Stufen), die Konzentration der Präzipitanzen und von Additiven wie Salzen oder der Puffersubstanz systematisch variiert. Hierfür wurden sowohl 96well-Platten (*Innovadyne*) im *sitting-drop*-Verfahren als auch 24well-Platten (*Qiagen*) im *sitting*- und *hanging-drop*-Verfahren genutzt. Im letzteren Fall konnten somit auch unterschiedliche Zusammensetzungen des Kristallisationstropfens aus Protein- und Präzipitanslösung getestet werden (1:1, 1:2, 1:3, 2:3). Diese wurden in Volumina von 1-2 µL auf einem Schraubdeckel gemischt und nach Verschrauben über dem 1 mL fassenden Reservoir inkubiert. Die 24well-Platte für das *sitting-drop*-Verfahren bot nur Platz für einen Tropfen (1-

2 μL), welcher ebenfalls in Kontakt mit einem Reservoir von 1 mL Präzipitanslösung stand. Die Dokumentation erfolgte für die 96well-Platten durch das *Rockimager* Dokumentationssystem (*Formulatrix*), für die restlichen Platten mit dem Lichtmikroskop.

3.5.3 *Streak-Seeding* mit Kristallkeimen

Um die morphologische Qualität und die Größe von Proteinkristallen weiter zu verbessern, macht man sich zu Nutze, dass die Bildung eines Kristallkeims und das Wachstum der Kristalle unabhängige Prozesse sind. Wie oben beschrieben ist die Bildung von kritischen Nuklei in der labilen übersättigten Phase möglich, während das Wachstum von Kristallen erst bei niedrigerer Übersättigung erfolgt. Um das Wachstum von wenigen, dafür größeren und wohlgeordneteren Kristallen zu begünstigen können bereits in einem frühen Stadium Kristallisationskeime in Kristallisationstropfen manuell eingebracht werden²⁵⁹. Bei diesem, *Mikro-Seeding* (vom eng. *to seed*, sähen) genannten Prozess, können neben Nuklei des zu kristallisierenden Proteins auch exogene Fremdkörper, wie Glasstaub, Fasern oder Nanopartikel genutzt werden. Der Einsatz von Nanopartikeln dient auch bei schwer zu kristallisierenden Proteinen dazu, die Bildung von Nuklei erst zu ermöglichen^{260,261}. Das Seeding kann darüber hinaus angewandt werden, um die Bildung einer bestimmten Kristallform zu steuern. Dazu werden Keime des zu kristallisierenden Proteins nur in der gewünschten Kristallform eingebracht. Hiervon zu unterscheiden ist das *Makro seeding*, bei welchem ganze Kristalle in einen neuen Kristallisationstropfen überführt werden, um deren Größe zu beeinflussen²⁶².

Für diese Arbeit wurde die Qualität von Cea1A-Kristallen durch *Seeding* verbessert. Hierfür wurden zwei Varianten der *Microseeding*-Methode verwendet^{263,264}. Bei der ersten Ausführung wurden Kristallkeime durch vortexieren (15 s) eines Proteinkristalls mit einer Stahlkugel in einem Teil der Mutterlösung erzeugt und mit einem Katzenschnurrbarthaar in Form einer Verdünnungsreihe nacheinander durch mehrere neue Kristallisationstropfen gezogen. Dieses sog. *streak-seeding* wurde ebenfalls in der zweiten Variante genutzt. Dazu wurde die auf nanoskaliger Ebene raue Oberfläche des Haares direkt zu Nutze gemacht, um einen Proteinkristall zu berühren, randomisiert mikroskopische Bruchstücke herauszuberechnen und diese ebenfalls als Verdünnungsreihe in neue Tropfen einzubringen. Neben Katzenschnurrbarthaaren können ebenfalls Pferdeschweifhaar oder Hasenschnurrbarthaare verwendet werden²⁶³.

3.5.4 *Soaking* von Kristallen mit Sacchariden

Es gibt mehrere Möglichkeiten, um Komplexstrukturen von Proteinen mit Liganden zu erhalten. Neben der gemeinsamen Kristallisation eines Proteins mit einem Liganden kann ebenso ein bestehender Proteinkristall in einer Lösung des Liganden *gesoakt* werden. Außerdem können Liganden bereits während der Aufreinigung zugesetzt werden oder, solange es sich um einen Kofaktor handelt, koexprimiert werden. Einen guten Überblick über Methoden zum Erzeugen von Kokristallen liefern HASSELL *et al.*²⁶⁵.

Neben der Kokristallisation von Cea1A, Epa1A, Epa6A, Epa9A und Flo10A mit unterschiedlichen Sacchariden wurden auch erhaltene Kokristalle mit anderen Liganden *gesoakt*. Hierbei wurde die erfolgreiche Kristallisation von bspw. Epa6A und Epa9A mit Lactose ausgenutzt, um weitere Komplexstrukturen zu erhalten. Beim Einsatz dieses sogenannten *replacement-soaking* muss beachtet werden, dass die neuen Liganden möglicherweise eine geringere Affinität zum Protein aufweisen als derjenige, mit dem kokristallisiert wurde²⁶⁶. Um dieses Problem zu umgehen, wurde in einem ersten Schritt der Kristall mit EDTA-haltigem Puffer gewaschen um das Ca²⁺-Ion der Bindungstasche inklusive dem gebundenem Ligand zu entfernen. Kokristalle von Epa6A und Epa9A wurden von RIKE DIDERRICH hergestellt¹⁶⁵.

Das *replacement-soaking* von Epa1A und Cea1A wurden mit Hilfe folgenden Protokolls durchgeführt: Kristalle wurden dem Kristallisationstropfen entnommen und wiederholt in der entsprechenden, reproduzierten Kristallisationslösung mit 2 mM EDTA *gesoakt* (2 h, 1 h, 30 min, 20 min, 10 min). Diese Lösung wurde ebenfalls im Reservoir vorgelegt (1 mL, 24well Platte). Die Kristalle wurden in einen Tropfen mit 5 mM CaCl₂ und 5 mM Saccharid überführt, wiederholt über unterschiedliche Zeiträume *gesoakt* (1 h, 2 h, 6 h, über Nacht) und der Lösung entnommen.

3.5.5 Vorbereiten von Kristallen zur Röntgenbeugung

Kristalle, welche für Röntgenbeugungsexperimente genutzt werden sollten, wurden mit CryoLoops (*Hampton Research*) oder MicroMounts (*MiTeGen*) mit Hilfe eines Stereomikroskops aus der Lösung aufgenommen. Strahlenschäden, wie Radikalbildung und Destruktion von Disulfidbrücken, welche bei einem Beugungsexperiment entstehen und zu vermehrter Unordnung im Kristall führen, lassen sich durch Reduktion der Temperatur auf 100 K stark verringern²⁶⁷. Um zu verhindern, dass bei dieser geringen Temperatur das im Kristall enthaltene Lösungsmittel kristallin gefriert, nutzt man unterschiedliche organische Substanzen wie Glycerin, MPD oder auch Kohlenhydrate als Kryoprotektanzen²⁶⁸. In dieser Arbeit wurden hierzu die Kristalle in reproduzierter Kristallisationslösung mit 15-20% Glycerin etwa 30 s *gesoakt* und danach in flüssigem Stickstoff schockgefroren, gelagert und in einem Transportdewar zur Röntgenquelle transportiert.

3.5.6 Röntgenbeugungsexperimente

Die Röntgenstrukturanalyse basiert auf der Beugung von Röntgenstrahlung mit einer Wellenlänge welche etwa den Gitterparametern des Proteinkristalls entspricht. Die Beugung wird durch die LAUE-Bedingung bzw. das BRAGG'sche Gesetz beschrieben. Die Intensität und Position der durch die konstruktive Interferenz entstehenden Beugungsmaxima sind abhängig von der Art und des Inhalts der Einheitszelle des Kristalls. Sie werden mit Hilfe geeigneter Detektoren aufgenommen und weiter verarbeitet. Die Theorie des Röntgenbeugungsexperiments ist in der Fachliteratur ausführlich beschrieben^{269,270}.

Kristalle wurden zum Teil mit Hilfe einer $I\mu^{\text{TM}}$ Mikrofokusröhre (*Bruker*) und einem MAR345 Detektor (*marresearch*) vorcharakterisiert, bevor diese an einem Messplatz des Deutschen Elektronensynchrotrons (DESY, PETRA-III, Hamburg) oder des BESSY-II Synchrotrons (HZB, Berlin-Adlershof) auf einem Goniometerkopf montiert wurden und bei einer Temperatur von 100 K im Röntgenstrahl ausgerichtet wurden. Die detaillierte Ausstattung der Messplätze kann für das BESSY der Publikation von MÜLLER *et al.* und für PETRA-III der Webseite des EMBL entnommen werden^{271,272}.

Die Menge an Daten, die gesammelt werden muss, hängt stark von der Symmetrie und den Zellparametern der Kristalle ab. Daher wurden zu Beginn 2 Testbilder bei einem Oszillationswinkel von 1° in einem Winkelabstand von 90° zueinander aufgenommen. Diese gaben, neben einem ersten Eindruck der Streuqualität des Kristalls, auch die Möglichkeit, mit Hilfe der Software *Imosflm*^{273,274}, die Zellparameter und damit eine Strategie zur Sammlung eines kompletten Datensatzes bei der höchstmöglichen Auflösung und Multiplizität zu berechnen. Die anzupassenden Parameter waren (i) die Anzahl und (ii) die Dauer der einzelnen Expositionen des Kristalls (je Bild 0.1 s-20 min), (iii) der Oszillationswinkel um welchen während der Exposition gedreht wurde (0.1° - 1°) sowie (iv) der Abstand des Detektors zum Kristall. Nach der Messung wurde der Kristall erneut in flüssigem Stickstoff gefroren und ggf. zum *Seeding* verwendet.

3.5.7 Prozessierung und Reduktion der Daten

Der Zusammenhang zwischen der gebeugten Welle und der Elektronendichte im Kristall ist durch folgende Fourier-Transformation gegeben:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{2\pi i(hx+ky+lz)+i\Phi(hkl)}$$

$\rho(x, y, z)$: Elektronendichte am Ort x, y, z ; V : Volumen der Einheitszelle, hkl : Miller'sche Indizes; $|F_{hkl}|$: Strukturfaktoramplitude; Φ : Phasenwinkel

Aus der Gleichung wird ersichtlich, dass die Berechnung der Elektronendichte an jeder Stelle der Einheitszelle $\rho(x, y, z)$ möglich ist, wenn man die Strukturfaktoren F_{hkl} mit den zugehörigen Phasenwinkeln Φ kennt. Unter Berücksichtigung eines Korrekturfaktors lassen sich die Strukturfaktoramplituden durch die Intensitäten der Reflexe errechnen²⁷⁰:

$$I(hkl) \propto |F(hkl)|^2$$

Die Intensitäten der gemessenen Reflexe können dazu mit Hilfe geeigneter Software integriert und skaliert werden. Die Integration erfolgte mit den Programmen *XDS*^{275,276} oder *imosflm*^{273,274}. Hierbei wird jeder Reflex hkl erfasst, seine Intensität gemessen, Korrekturen für die Strahlposition, den Detektorabstand und die Mosaizität einberechnet sowie die genauen Zellparameter berücksichtigt. Qualitätskriterien stellen an dieser Stelle etwa die Vollständigkeit, welche den Anteil der gemessenen an den theoretisch möglichen Reflexen wiedergibt,

der Intensitätsunterschied zwischen symmetrieäquivalenten Reflexen (R_{merge}) und der Rauschabstand $I/\sigma(I)$ dar.

Die Skalierung wurde mit der Software *XSCALE*²⁷⁷ oder *SCALA*²⁷⁸ des CCP4-Pakets^{279,280} durchgeführt. Hier werden die symmetrieäquivalenten Reflexe zusammengeführt (*merging*) und skaliert. Neben der Zusammenfassung der Reflexe wurde in diesem Schritt die real erreichbare Auflösung bestimmt. Hierbei wurde darauf geachtet, dass der Wilson-Plot zu hohen Auflösungen hin möglichst linear verläuft, der Rauschabstand einen Wert von 2 nicht unterschreitet sowie R_{merge} bzw. R_{meas} , welcher im Gegensatz zu R_{merge} die Multiplizität berücksichtigt, nicht größer als 70% war. Häufig werden an dieser Stelle in der Literatur geringere Werte als Ausschlusskriterium für die Strukturlösung angegeben, doch konnte gezeigt werden, dass ein fester Wert für die Festlegung der Auflösung experimentellen Daten nicht ausreichend gerecht wird^{281,282***}.

3.5.8 Das Phasenproblem – Strukturlösung mittels molekularen Ersatzes

Nachdem nun die Intensitäten der Reflexe bekannt sind, fehlt noch der Phasenwinkel zur Berechnung der Elektronendichte und zur Lösung der Struktur. Wie zuvor angedeutet, korrespondiert die Amplitude der gebeugten elektromagnetischen Welle mit den Intensitäten der Reflexe. Die Phase der Röntgenstrahlung geht allerdings während des Experiments verloren, sie lässt sich nicht aus den experimentellen Daten direkt errechnen. Problematisch ist hierbei, dass der Informationsgehalt der Phasen größer ist als der der Intensitäten (Abbildung 3.3).

Dennoch lassen sich die Phasen über Umwege berechnen²⁸³. In dieser Arbeit wurde zur Lösung der Strukturen der Adhäsionsdomänen die Methode des molekularen Ersatzes (MR, *molecular replacement*) genutzt. Diese wurde in den 1960er Jahren zuerst von ROSSMANN und BLOW sowie unabhängig von ROBERT HUBER^{†††} angewandt^{284,285}. Hierbei macht man sich zu Nutze, das verwandte Proteine häufig ein ähnliches Faltungsmotiv zeigen²⁸⁶. Die Phasen lassen sich damit durch Korrelation der experimentellen Daten mit einer bekannten Struktur erhalten bzw. annähern. PATTERSON entwickelte 1935 eine Fourier-Transformation der Intensitäten statt der Strukturfaktoren, indem er die Phasenwinkel auf 0 setzte²⁸⁷.

$$P(x, y, z) = \sum_h \sum_k \sum_l |F_{hkl}|^2 e^{-2\pi i(hx+ky+lz)}$$

$P(x, y, z)$: Pattersonfunktion am Ort x, y, z ; hkl : Miller'sche Indizes; $|F_{hkl}|^2$: Betragsquadrat der Strukturfaktoren

*** „The traditional R factors measuring internal consistency, R_{merge} or better R_{meas} , are not suitable measures for setting a resolution cutoff despite their apparent popularity with referees.“ (EVANS & MURSHUDOV, 2013)

††† Hier als „Faltenmolekülmethode“ bezeichnet.

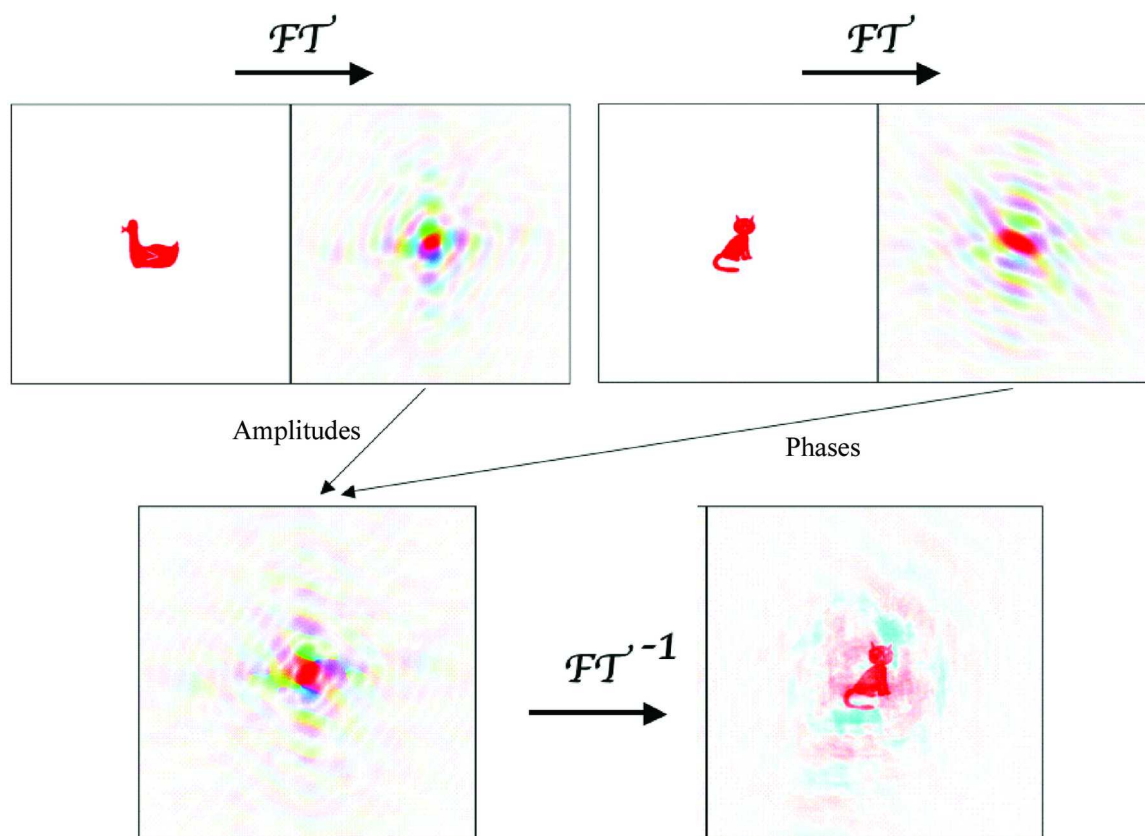


Abbildung 3.3 – Die Bedeutsamkeit der Phasenwinkel. In der oberen Reihe ist sowohl die Fourier-Transformation einer Ente als auch die einer Katze gezeigt. Werden die Amplituden der Ente mit den Phasenwinkeln der Katze kombiniert und das ursprüngliche Bild errechnet erhält man eine Katze²⁸³.

Die sog. PATTERSON-Funktion lässt sich direkt über die experimentellen Daten berechnen und liefert Informationen über die interatomaren Abstände in der Einheitszelle. Wird die PATTERSON-Funktion nun für die unbekannte Struktur und ein Strukturhomolog erstellt, kann versucht werden, diese über *maximum-likelihood*-basierte Algorithmen zu korrelieren. Die Methode ist häufig bereits bei einer Sequenzidentität von 20%-30% erfolgreich²⁸⁸. Diese 6-dimensionale Suche (3 Freiheitsgrade für die Rotation, 3 Freiheitsgrade für die Translation) wird in zwei Schritten durchgeführt. Zuerst wird die berechnete PATTERSON-Karte der bekannten Struktur einer Rotation um alle drei Raumachsen unterzogen und mit der PATTERSON-Karte der experimentellen Daten verglichen. Hierbei wird die größte Korrelation erzielt, wenn beide Strukturen sich in ähnlicher Orientierung befinden. Das so orientierte Modell wird nun durch Translationsfunktionen versucht in der Einheitszelle erfolgreich zu platzieren. Gelingt dieser Vorgang, lassen sich über das bekannte Modell initiale Phasen berechnen, welche ausreichend sind, um eine erste Elektronendichtekarte zu erhalten. Da, wie oben beschrieben, die Phasen einen höheren Informationsgehalt haben als die Intensitäten, muss hierbei darauf geachtet werden, dass der Einfluss des Suchmodells nicht zu groß wird und dies die Elektronendichte beeinflusst²⁷⁰. Das Qualitätsmerkmal für Struktururlösungen mittels MR stellt der Z-Score dar, welcher sich aus dem logarithmischen *likelihood gain* der errechneten Lösungen (LLG_L) abzüglich eines logarithmischen *likelihood gain* einer zufälligen

Lösung (LLG_R) zusammensetzt²⁸⁹. Dieser Zusammenhang wird durch die Standardabweichung einer Auswahl zufälliger LLG-Werten vom Mittel geteilt.

$$Z = \frac{LLG_L - LLG_R}{\sigma_{LLG}}$$

Z: Qualitätskriterium; **LLG_L:** log. *likelihood gain* der Lösung; **LLG_R:** log. *likelihood gain* einer zufälligen Lösung; **σ_{LLG} :** Standardabweichung der LLG vom arithmetischen Mittel

Für den molekularen Ersatz wurde hauptsächlich das Programm *Phase*²⁹⁰ verwendet, implementiert in das CCP4-Paket²⁸⁰ und die *Phenix Suite*²⁹¹. Epa6A und Epa9A konnten mit Hilfe von Epa1A als Suchmodell gelöst werden, während die Lösung der Strukturen von PpFlo1A und Cea1A erst nach mehrfacher Modellierung und Bearbeitung von Homologiemodellen, welche mit *modeler9v7*²⁹² auf Basis von Epa1A erstellt wurden, gelang. Die Lösung der Struktur von Flo10A gelang mit Hilfe eines zuvor aufgeklärten Strukturmodells eines Fusionsproteins aus Flo10A mit der Subdomäne von Flo5A. Wurden erfolgreich Phasen erhalten, konnte mit ARP/wARP²⁹³ ein erstes Modell der unbekanntenen Struktur automatisiert gebaut werden.

Es existieren, neben der reinen Kalkulation der Phasenwinkel, auch die Möglichkeit diese experimentell zu bestimmen^{270,283}. Hier ist zum einen der isomorphe Ersatz zu nennen, welcher die Abhängigkeit eines jeden Reflexes von den Eigenschaften jedes einzelnen Atoms der Einheitszelle ausnutzt. Durch Einbringen von elektronenreichen Atomen in den Kristall wird der Einheitszelle ein starker Streuer zugefügt; dies gelingt meist durch *Soaking* eines Kristalls in der Lösung eines Schwermetalls. Die Komplexität der Substruktur der Schwermetallatome ist im Vergleich zum Proteinkristall sehr gering, weshalb sich die Phasen dieser Substruktur mittels direkter Methoden²⁸³ oder der PATTERSON-Funktion²⁸⁷ lösen lassen. Über die nachfolgende HARKER-Konstruktion²⁹⁴ lassen sich mit den Intensitätsanteilen der Proteinatome und der Schwermetallatome jeweils zwei Phasenwinkel berechnen, wenn ein nativer sowie ein isomorpher Schweratom-Datensatz zur Verfügung stehen (*single isomorphous replacement*, SIR). Sind zwei oder mehrere Schweratomderivate vorhanden, kann eine eindeutige Phase ermittelt werden (*multiple isomorphous replacement*, MIR).

Eine dritte Methode beruht auf der Absorption von Röntgenstrahlung durch elektronenreiche Atome in der Nähe ihrer Absorptionskante. Hier ist die Gültigkeit des FRIEDEL'schen Gesetzes aufgehoben, da die Schweratome die Röntgenstrahlung absorbieren und reemittieren, was die Phase und die Intensität aller Reflexe beeinflusst. Da dies allein durch die Position der Schweratome gegeben ist, lassen sich auch hier die Phasenwinkel mittels der PATTERSON-Funktion und der HARKER-Konstruktion errechnen. Für die SAD-Methode (*single wavelength anomalous dispersion*) wird ein nativer Datensatz sowie ein anomaler Datensatz benötigt, während für MAD (*multi wavelength anomalous dispersion*), welche eindeutige Phasen liefert, zwei anomale Datensätze bei unterschiedlichen Wellenlängen benötigt werden.

Unter der Voraussetzung, dass ein Kristall bis zu einer Auflösung von $<1.2 \text{ \AA}$ streut und die Größe des Proteins etwa 1000 Atome nicht übersteigt^{†††}, gibt es die Möglichkeit, die Phasenwinkel direkt zu bestimmen²⁹⁵. Hierbei werden die Form der Atome, welche durch ihre positive Elektronendichte definiert wird, sowie statistische Zusammenhänge zwischen den Strukturfaktoren ausgenutzt. Die Kenntnis der Phasenwinkel weniger Reflexe, welche informativ mittels Exhaustionsmethode erhalten werden, kann hierbei über *Bootstrapping*-Methoden zu den Phasenwinkeln aller Reflexe führen²⁸³.

3.5.9 Verfeinerung von initialen Modellen & Generierung von Ligandenbibliotheken

Die Anpassung der erhaltenen Modelle an die Elektronendichte, die sog. Verfeinerung, erfolgte alternierend mit der automatischen Verfeinerung durch *refmac5*²⁹⁶ und manueller Inspektion sowie Modellbau mittels *oot*²⁹⁷. Die Qualität der *maximum-likelihood* Verfeinerung und damit des Modells wurde mit dem R-Faktor (*reliability-factor*) quantifiziert, welcher die Güte der Korrelation zwischen den aus dem Modell berechneten Strukturfaktoramplituden und den experimentell beobachteten darstellt. Hierbei wird zwischen dem R_{work} und dem R_{free} unterschieden. Der R_{work} gibt den Wert für die während der Verfeinerung direkt genutzten Reflexe an, während R_{free} den Wert für Testreflexe, welche im Schritt der Datenreduktion ausgewählt wurden ($\sim 5\%$, max 2000 Reflexe), beschreibt.

$$R = \frac{\sum | |F_{\text{obs}}| - |F_{\text{calc}}| |}{\sum |F_{\text{obs}}|}$$

R: Zuverlässigkeitsfaktor; $|F_{\text{obs}}|$: Experimentelle Strukturfaktoramplituden; F_{calc} : Berechnete Strukturfaktoramplituden

Die Testreflexe wurden bei Strukturen mit einem Molekül in der asymmetrischen Einheit (mol/asu) linear über alle Auflösungsschalen hinweg zufällig ausgewählt. Bei Cea1A und PpFlo1A hingegen, welche aufgrund nichtkristallografischer Symmetrie (NCS) 4 respektive 2 mol/asu besitzen, wurden die Testreflexe in dünnen Auflösungsschalen mit dem Programm *DataMan* festgelegt²⁹⁸. Dies war nötig, da gezeigt wurde, dass die R-Werte mit zunehmender NCS beeinflusst sein können, wenn die Testreflexe in linearer Form festgelegt werden²⁹⁹. Die NCS wurde außerdem teilweise als *restraint* während der Verfeinerung genutzt; dieses kann gerade zu Beginn der Verfeinerung eine Verbesserung des Modells bewirken³⁰⁰.

Bei der Verfeinerung wird versucht die kristallographische Zielfunktion E_{total} zu minimieren um experimentelle Daten und chemische Vorgaben (*restraints*) wie Bindungslängen, Bindungswinkel *et cetera* in Einklang zu bringen³⁰⁰.

^{†††} „SHELDRICK's Regel“

$$E_{total} = E_{chem} + WA \cdot E_{exp}$$

E_{total} : kristallographische Zielfunktion; E_{chem} : Abweichung von den chemischen Vorgaben; WA: Gewichtungsfaktor der experimentellen Daten; E_{exp} : Abweichung von den experimentellen Daten.

Alle zur Verfeinerung benutzen Programme benötigen somit Bibliotheken welche die *restraints* und damit die Beschreibung der im Model vorkommenden Moleküle beinhalten. Die proteinogenen Aminosäuren sowie viele Kleinmoleküle, wie Puffersubstanzen und etwaige Liganden, sind bereits in internen Bibliotheken der Programme enthalten. Des Weiteren können weitere Bibliotheken über in der PDB hinterlegte Strukturen bezogen werden. Im Falle dieser Arbeit wurden allerdings synthetische Kohlenhydrate verwendet, welche weder einen korrespondierenden Eintrag in einer Bibliothek noch in der PDB haben. Um eine eigene Bibliothek für diese Liganden zu erstellen, wurde einerseits das Programm *eLBOW*³⁰¹ der *PhenixSuite* genutzt, welches mit Hilfe von sog *SMILES*-Strukturcodes³⁰² (*Simplified Molecular Input Line Entry Specification*) eine energieminierte PDB-Datei des Liganden sowie eine Bibliothek mit entsprechenden *restraints* erstellt. Für komplexere Fälle andererseits wurde der Server *PRODRG*³⁰³ verwendet, welcher das Zeichnen eines Liganden mit anschließender Verarbeitung erlaubte. Die so generierten Bibliotheken konnten daraufhin in *refmac5* und *coot* verwendet werden.

3.5.10 Strukturvisualisierung und Analyse von Proteinstrukturen

Die Visualisierung von Strukturmodellen und das Anfertigen von Abbildungen erfolgten mit *Pymol*³⁰⁴ sowie *Chimera*³⁰⁵. Die Berechnung elektrostatischer Oberflächen wurde mit dem Plugin *APBS*³⁰⁶ durchgeführt und die Strukturen mit dem *PDBePISA*³⁰⁷-Server auf Interaktionsflächen untersucht.

3.6 Methoden zur Funktionsuntersuchung von Adhäsinen *in vivo*

Die im Folgenden beschriebenen Untersuchungen der Funktion von Adhäsionsproteinen aus *Pichia pastoris in vivo* wurde in enger Zusammenarbeit mit Dr. STEFAN BRÜCKNER durchgeführt. Die Interaktionsstudien mit Schaben erfolgten mit Unterstützung von DOROTHEE TEGTMEIER.

3.6.1 Immunfluoreszenz-Mikroskopie

Um Experimente mit Adhäsin-präsentierenden *S. cerevisiae*-Zellen durchführen zu können, muss zuvor sichergestellt werden, dass die Zellen das Fusionsprotein produzieren und auf der Zelloberfläche tragen. Dies erfolgte mittels Fluoreszenzmikroskopie nach Antikörpermarkierung. Dazu wurden zunächst *S. cerevisiae*-Zellen vom Stamm RH2520 mit den Konstrukten BHUM2200 (Δ A-Domäne), BHUM2296 (*PpFlo1A*) und BHUM2297 (*Cea1A*) transformiert, siehe 3.2.6. Diese tragen, neben der entsprechenden A-Domäne auf Basis der Flo11 B- und C-Domäne, ein 3HA-Epitop (3x Hämagglutinin), welches die Markierung mit Antikörpern und Fluoreszenzfarbstoff-konjugierten Sekundärantikörpern erlaubt um die erfolgreiche

Expression sowie die Lokalisation der Proteine in der Zellwand zu bestätigen (Abbildung 3.4).

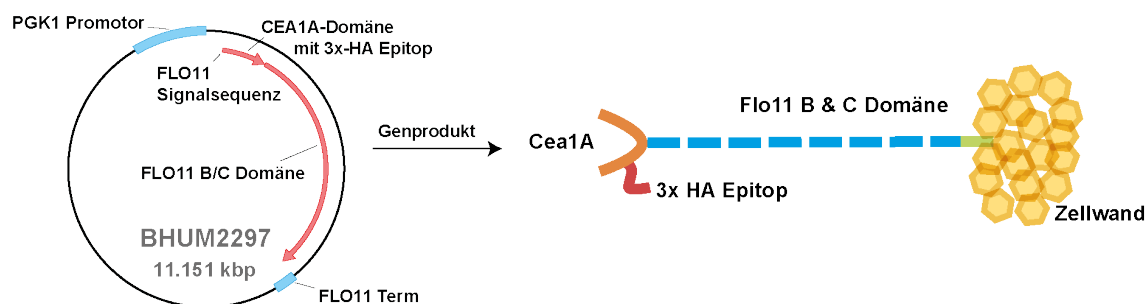


Abbildung 3.4 – Schematische Darstellung des Konstruktes BHUM2297 und des codierten Fusionsproteins. Die Expression des Gens steht unter Kontrolle des PGK1-Promotors; die Flo11 Signalsequenz stellt die Lokalisation auf der Zelloberfläche sicher. Das Genprodukt besteht aus einem 3x HA-Epitop (rot), gefolgt von der Cea1A-Domäne (orange) und der Flo11 B & C Domäne (blau, grün).

Kulturen der Konstrukt-tragenden Stämme wurden in YNB-If Medium (*Yeast-Nitrogen-Base, low fluorescence*) hergestellt und bis zu einer $OD_{595}=1$ wachsen gelassen. 3 mL der Kultur wurden pelletiert (3 min, 2000 rpm) und dreimal mit 1 mL PBS (+1% BSA) gewaschen. Der Überstand wurde verworfen und die Zellen in PBS(+1% BSA) mit monoklonalem Maus-Anti-HA-Antikörper (1:1000; H3663, *Sigma Aldrich*) resuspendiert und 30 min bei Raumtemperatur inkubiert. Nach drei weiteren Waschschrritten wurden die Zellen in PBS(+1% BSA) mit Cy3-konjugiertem sekundären Ziege-Anti-Maus-Antikörper (1:10000; C2181, *Sigma Aldrich*) aufgenommen und 20 min bei Raumtemperatur im Dunklen inkubiert. Nach erneutem dreimaligem Waschen wurden 2-5 μL der Zellen auf ein Bett aus 1% Agarose gegeben und mit einem inversen *Zeiss* Axiovert 200 M Fluoreszenzmikroskop mit CCD-Kamera untersucht. Zuerst wurden die immunochemisch markierten Zellen mittels Differentialinterferenzkontrast-Mikroskopie visualisiert (DIC) und anschließend die Fluoreszenz des Sekundärantikörpers mit einem Rhodamin-Filter-Set (Anregung 551 nm, Emission 573 nm) bei einer Belichtungszeit von 600 ms analysiert. Die Aufnahme der Bilder mit einer Abmessung von 10.062 Pixeln/ μm wurde ohne softwareseitige Veränderung mit *Velocity* (*PerkinElmer*) durchgeführt und die Bearbeitung erfolgte anschließend mit dem Bildbearbeitungsprogramm *ImageJ* (*Improvision*)³⁰⁸.

PBS-Puffer	
KH_2PO_4	2 mM
Na_2HPO_4	10 mM
NaCl	137 mM
KCl	2.7 mM

3.6.2 Bindung von Adhäsion-präsentierenden Zellen an Chitin-Beads

Um zu untersuchen ob die in vorigen Experimenten identifizierte und charakterisierte Bindung von Cea1A an GlcNAc, welches den Baustein von Chitin darstellt, *in vivo* funktional ist und in welchem Maße diese Funktion phänotypisch beobachtbar ist, wurden CeaA-präsentierende *S. cerevisiae* Zellen mit Chitin-Beads gemischt und das Resultat fluoreszenzmikroskopisch analysiert.

Dazu wurden Zellen des Stammes *S. cerevisiae* RH2520 mit den Konstrukten BHUM2200 (Δ A-Domäne), BHUM2296 (*PpFLo1A*) und BHUM2297 (Cea1A) transformiert, siehe 3.2.6. Eine Kultur in SC-4 Medium (exponentielle Phase) wurde auf eine $OD_{595}=10$ durch Pelletieren konzentriert und 1 mL mit 50 μ L DiOC₆ (3,3'-Dihexyloxycarbocyanin-iodid; *Santa Cruz Biotechnology*; 0.3 mg/mL in 1:9 PBS/EtOH), einem Lebendfarbstoff welcher das endoplasmatische Retikulum anfärbt, versetzt³⁰⁹. Nach Vortexieren wurden die Zellen geerntet (3 min, 2000 rpm) und in 200 μ L SC-4 mit CaCl₂ (10 mM) resuspendiert. Die Zellsuspension wurde im Verhältnis 3:1 mit Chitin-Beads (*New England Biolabs*) gemischt.

Die Beads wurden zuvor mit einer Mischung aus einer Endo- und einer Exochitinase aus *Trichoderma viride* (*Sigma Aldrich*) behandelt, um das Vorhandensein terminaler GlcNAc-Enden sicherzustellen. Dafür wurden 1.5 mL aufgeschlämmte Beads pelletiert (3 g, 1 min) und 5x mit 500 μ L Phosphat-Puffer (100 mM, pH 6.1) gewaschen und der Überstand verworfen. Die Beads wurden in 500 μ L Phosphat-Puffer mit 0.5 mg Chitinase resuspendiert und bei 25 °C 3 h bei 500 rpm inkubiert. Die Inaktivierung der Chitinase erfolgte durch 3x waschen mit 800 μ L 70% Ethanol. Nach 5x waschen mit Phosphat-Puffer wurden die Beads in demselben Puffer resuspendiert und bei 4 °C gelagert.

Zu 200 μ L der *S. cerevisiae*-Zell/Bead-Mischung wurde 1 mL SC-4 (+10 mM CaCl₂) gegeben, vorsichtig durch Invertieren gemischt und 2 min bei Raumtemperatur inkubiert. Die Suspension wurde erneut invertiert und nach 10 s der Überstand abgenommen um ungebundene Zellen zu entfernen. Dieser Schritt wurde erneut durchgeführt und die sedimentierten Beads auf 1% Agarose-Betten mikroskopiert (Axiovert 200 M, *Zeiss*). Anstatt eines Rhodamin-Filter Sets wurde diesmal ein GFP-Filter-Set für die Detektion der DiOC₆-Fluoreszenz (Anregung 485 nm, Emission 501 nm) verwendet. Eine Quantifizierung erfolgte, indem freie und gebundene Zellen in 10 vergleichbaren Bildern gezählt wurden (*ImageJ*)³⁰⁸.

3.6.3 Fluoreszenzmikroskopie mit AlexaFluor 488 markiertem Protein

Neben Experimenten mit Adhäsion-präsentierenden Zellen wurden ebenfalls *S. cerevisiae*-Zellen mit heterolog produziertem Protein versetzt, welches zuvor mit AlexaFluor 488 markiert wurde; siehe 3.4.3.

Übernachtskulturen unterschiedlicher *S. cerevisiae* und *P. pastoris* Stämme in YNB-If wurden auf eine $OD_{595}=2$ gebracht und Objektträger mit 6 Wells mit 30 μ L 0.1% Polylysin beschichtet. Nach 3-5 min wurde der Überstand von den Wells gesaugt und 3x mit 30 μ L YNB-If gespült. 30 μ L frisch vortexierter Zellsuspension wurden auf ein Well gegeben, 20 min abgedeckt bei Raumtemperatur inkubiert und der Überstand vorsichtig entfernt. Zur Absättigung des Polylysins wurden 30 μ L PBS (+1% Triton X100) pro Well für 5 min inkubiert und entfernt. Mit je 30 μ L dieser Lösung wurde weitere 3x gespült und nach dem letzten Schritt 20 min inkubiert. Der Überstand wurde entfernt und 5 μ L des gelabelten Proteins (1 mg/mL) + 25 μ L PBS (+1% Triton X100) auf die Zellen gegeben und 5 min inkubiert. Nachdem 3x mit 30 μ L PBS (+1% Triton X100) gewaschen wurde, wurden 25 μ L des Überstandes entfernt und die Zellen an einem Fluoreszenzmikroskop untersucht (Axiovert 200 M, Zeiss). Hierbei wurde erneut auf das GFP-Filterset zurückgegriffen (Anregung 495 nm, Emission 519 nm) und die Bilder anschließend mit *ImageJ* bearbeitet.

3.6.4 Interaktionsuntersuchung zwischen Hefen und Schaben

Da bereits in früheren Arbeiten von Interaktionen zwischen Hefen und Insekten berichtet wurde (siehe Einleitung) und bekannt ist, dass Insekten eine komplexe Darmflora besitzen, sollte überprüft werden inwiefern die untersuchten Adhäsine eine mögliche Wechselwirkung vermitteln können. Weiterhin ist bisher nicht bekannt, ob eine gezielte Etablierung von Hefen im Darm von Insekten möglich ist. Als Model wurde auf die Schabe *Shelfordella lateralis* zurückgegriffen, da diese einfach zu kultivieren ist, die biochemischen Eigenschaften der drei distinkten Bereiche des Darms, wie pH-Wert und RedOx-Potential, (*Crop*, *midgut* & *hindgut*) gut bekannt sind und die Schabe für seine enorme Darmflora bekannt ist¹⁷⁰. Um mögliche symbiotische/pathogene Interaktionen zu untersuchen, wurden die Adhäsinspräsentierenden Zellen aus 3.6.1 an frisch kultivierte, sterile Schaben verfüttert.

Folgender Abschnitt ist bildlich in Abbildung 3.5 dokumentiert. Zunächst wurden Ootheken von in Plastikkästen gehaltenen Schaben gesammelt. Die Schaben werden darin bei 25 °C, 50% Luftfeuchtigkeit und einem 12 h Hell/Dunkel Rhythmus gehalten. Nachdem die Ootheken mit 2% Peressigsäure sterilisiert (max. 5 min) und in sterile Falcon-Tubes (50 mL) gelegt wurden, folgte eine 2-3 wöchige Inkubation bis zum Schlüpfen der sterilen Larven. Die Sterilität wurde durch Ausplattieren einer Larve auf Vollmedium kontrolliert. Die Fütterung mit den entsprechenden Hefe-Stämmen erfolgte durch Aufsaugen von ca. 1 mL einer vortexierten Übernachtskultur ($OD_{595}=5-8$) in ein autoklaviertes Filterpapier, welches in das Falcon mit den Larven gegeben wurde. Es folgte eine Inkubation von einem Tag bei 25 °C, 50% Luftfeuchtigkeit und einem 12 h Hell/Dunkel Rhythmus. Die Schaben wurden auf einer Clean-Bench in ein neues Gefäß überführt und ab diesem Zeitpunkt mit Stärke (*Spielberger GmbH*: 36% Ballaststoffe, 15% Proteine) oder Hühnerfutter (Gold Plus, *Versele-Laga*) sowie Wasser ernährt, welches *ad libitum* zugeführt wurde. Die Sektion der Schaben erfolgte nach 4 Tagen, einer Woche und zwei Wochen *post inoculation*. Hierzu wurden die zu unter-

suchenden Schaben isoliert und in einem neuen Gefäß 5 min auf Eis immobilisiert. Mit sterilen Pinzetten wurde zuerst der Kopf entfernt und der Darm durch die entstandene Öffnung präpariert. *Crop*, *Midgut* und *Hindgut*, welche morphologisch gut voneinander zu unterscheiden sind, wurden getrennt und auf einem Bett aus 1% Agarose mittels Differentialinterferenzkontrast mikroskopiert (Axiophot; Zeiss). Nach der Mikroskopie wurden die Darmproben mit Glasskugeln mit einem Durchmesser von 0.25-0.5 mm unter vortexieren homogenisiert und auf Platten aus SC-4 Festmedium ausgestrichen um eine halbquantitative Aussage über die Besiedlung treffen zu können.

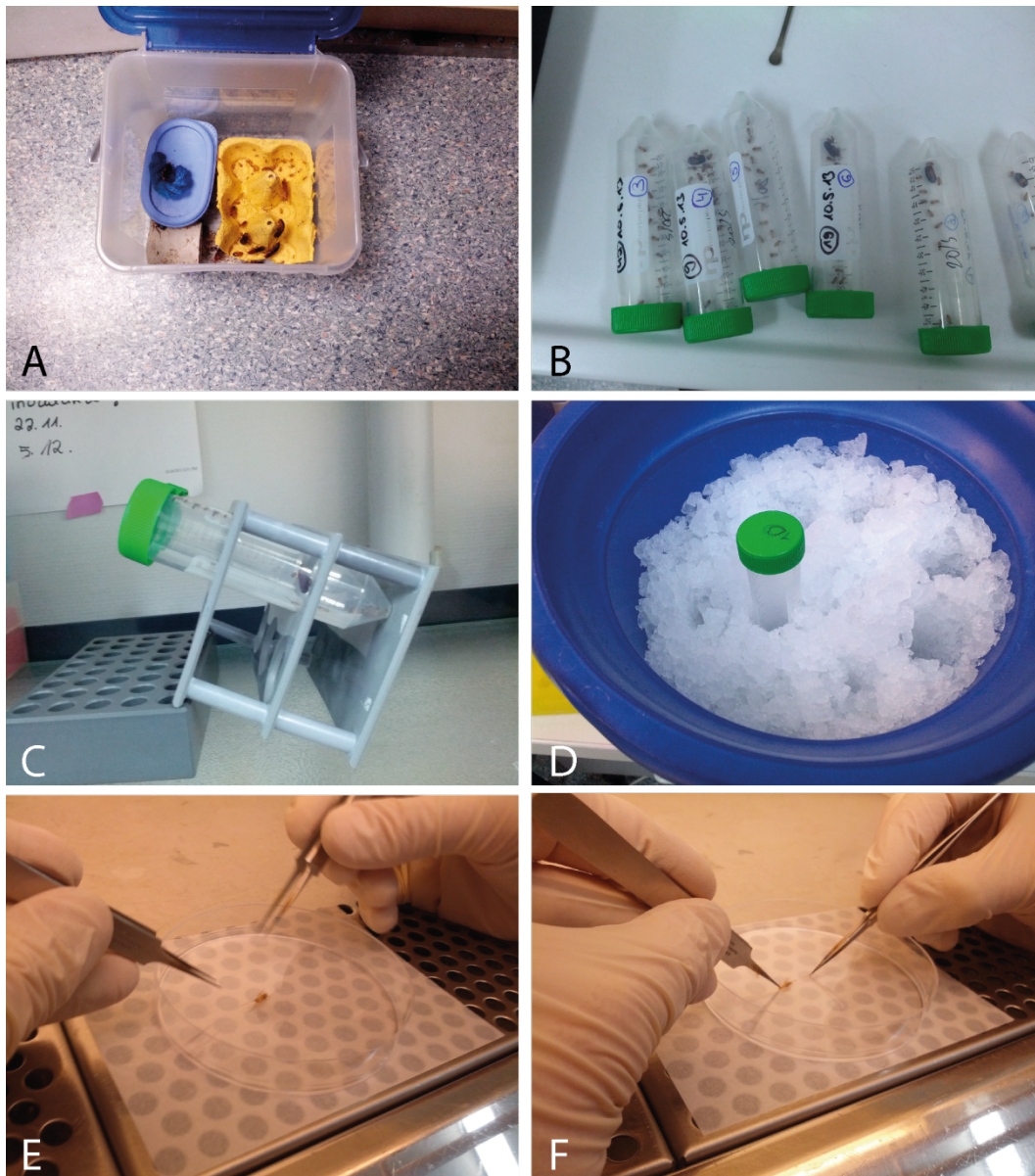


Abbildung 3.5 – Interaktionsuntersuchungen an Schaben. A: Schaben werden in Plastikkästen gehalten. B: Ootheken werden mit 2% Peressigsäure sterilisiert und 2-3 Wochen bis zum Schlüpfen der sterilen Larven inkubiert. C: Sterilisiertes, mit Hefe-Übernachtskultur getränktes Filterpapier wird 1 d mit den Schaben zum Füttern inkubiert. D-F: Schaben werden nach Immobilisation auf Eis seziiert und der Darm präpariert.

4 Ergebnisse

Die in dieser Arbeit untersuchten GPI-verankerten Hefezellwandproteine sind in Abbildung 4.1 zusammengefasst. Der Fokus lag hierbei auf deren A-Domäne, welche die Glycanbindung vermittelt. Die molekulare Grundlage der Interaktion von Hefen mit ihrer Umwelt stellt, unter anderem, die Ca^{2+} -vermittelte Bindung dieser Zellwandproteine an Glycane dar. Während die Genome von *Candida glabrata* CBS138³⁵ und *Saccharomyces cerevisiae* S288c²¹ vollständig sequenziert und größtenteils annotiert sind, stehen für *Pichia pastoris* DSMZ70382 nur einzelne Abschnitte aus *shotgun*-Sequenzierungen zur Verfügung⁵⁰. Die Gene von *Cea1* und *PpFlo1* sind daher nur teilweise bekannt. Während die A-Domäne vollständig verfügbar ist, endet die bekannte Sequenz am Beginn der B-Region.

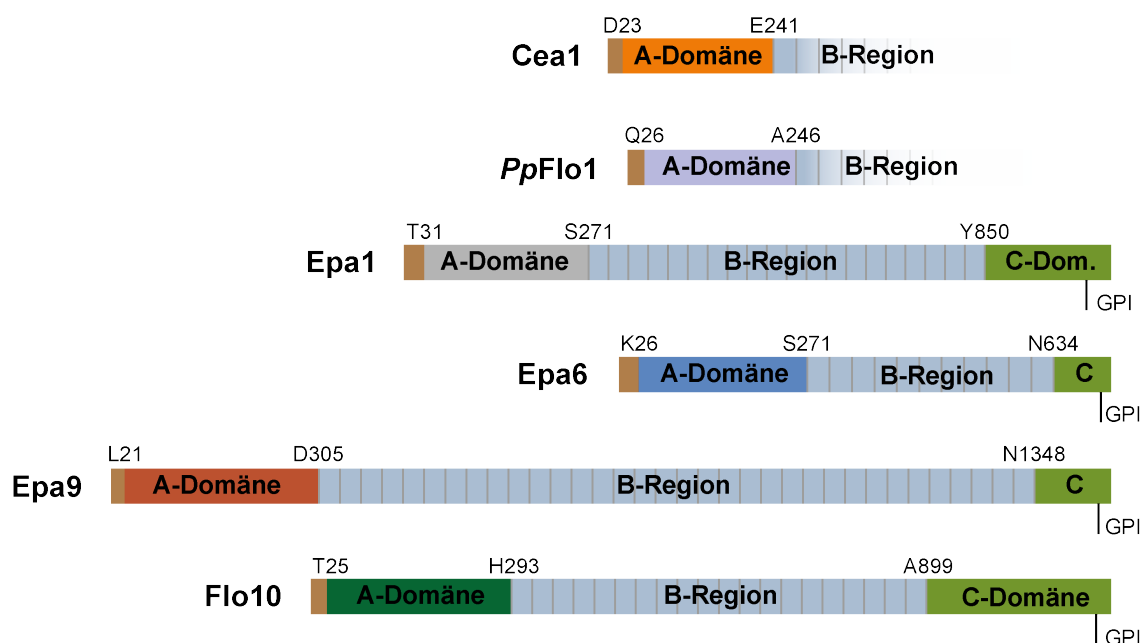


Abbildung 4.1 – Übersicht über die untersuchten Zellwandproteine. Der A-Domäne ist jeweils eine Signalsequenz vorgelagert (braun) um die Sekretion sicherzustellen. Gefolgt wird die A-Domäne von der Serin-/Threonin-reichen, repetitiven B-Region (blau). Die C-Domäne (grün) bildet die Basis der Verankerung des Proteins in der Zellwand durch Transglycosylierung über den GPI-Anker.

In nachfolgender Tabelle 4.1 sind die untersuchten A-Domänen aufgelistet und beschrieben.

Tabelle 4.1 – Informationen zu den untersuchten Adhäsinen und den zugehörigen Fusionsproteinen. Die Datenbank-ID bezieht sich bei *Cea1A* und *PpFlo1A* auf den *Pichia* Genome Browser, bei dem Rest auf UniProt.^{53,178}

Adhäsion	Organismus	Datenbank-ID	Aminosäuren	Mol. Masse mit His-Tag/ Da
<i>Cea1A</i>	<i>P. pastoris</i> DSMZ70382	PIPA03174	D23-E241	26681.5
<i>PpFlo1A</i>	<i>P. pastoris</i> DSMZ70382	PIPA05092	Q26-A246	27245.8
<i>Epa1A</i>	<i>C. glabrata</i> CBS138	Q6VBJ0	T31-S271	29401.6
<i>Epa6A</i>	<i>C. glabrata</i> CBS138	Q6DTM7	K26-S271	29909.0
<i>Epa9A</i>	<i>C. glabrata</i> CBS138	B4UMX2	L21-D305	33203.5
<i>Flo10A</i>	<i>S. cerevisiae</i> S288c	P36170	T25-H293	31599.7

4.1 Phylogenetische Analyse der PA14-Domänen enthaltenden GPI-CWP's

Die Hefezellwand ist sehr dynamisch, denn sie ist die Grenzfläche, an der die molekulare Interaktion zwischen Zelle und Umwelt stattfindet. Da zu erwarten ist, dass sich Hefen durch diese Wechselwirkung an ihre Lebensbedingungen anpassen, sollte eine phylogenetische Analyse einer großen Anzahl PA14-Domänen enthaltender GPI-verankerter Zellwandproteine Informationen darüber liefern, inwiefern sich dieser Prozess auch auf der Ebene der Adhäsion widerspiegelt.

Hierzu wurde, wie in Abschnitt 3.1.1 detailliert beschrieben, neben unterschiedlichen Datenbanken auch beispielhaft die Genome⁵⁰⁻⁵² dreier Stämme aus dem Genus *Pichia* /*Komagataella* nach putativen PA14-Domänen enthaltenden Adhäsinen durchsucht und diese anschließend einem multiplen Sequenzalignment sowie einer phylogenetischen Analyse nach der *ML*-Methodik unterzogen. Als Grundlage für diese Analyse dienten die bereits bekannten GPI-CWPs mit PA14-Domäne: Flokkuline, epitheliale Adhäsine (Epa), *PA14-containing wall proteins* (PWP's) sowie Epa-Homologe aus *Nakaseomyces*-Stämmen. Hierunter waren Epa1A und Flo5A biochemisch und strukturell bereits gut charakterisiert.

Der phylogenetische Baum in Abbildung 4.2 zeigt, dass der Großteil der Sequenzen in gut abgegrenzten Gruppen vorliegt. Diese Subtypen unterscheiden sich stark in ihrer Diversität. Während die Flokkuline (grün) sehr nah beieinander liegen, ist der Subtyp der epithelialen Adhäsine und Homologe weit aufgespalten (blau). Die Datenbanksuche führte weiterhin zur Identifikation von engen Verwandten der Flokkuline aus *S. cerevisiae* in weiteren Hefen. Der Subtyp der Epas hingegen zeigt neben verwandten Sequenzen aus dem Genus *Nakaseomyces* keine neuen putativen Adhäsine. Interessant ist an dieser Stelle, dass in der Spezies *C. braccarensis* dieses Genus sowohl Flokkulin- als auch Epa-ähnliche Genprodukte gefunden wurden. Dies ist im Einklang mit den Beobachtungen von GABALDON *et al.*, welche nicht-pathogene Stämme dieses Genus ohne oder mit nur wenigen EPA-Genen identifizierten und stattdessen unter anderem Flokkulin-ähnliche Sequenzen entdeckten³⁷.

Die Subtypen der PWPs (gelb) und der Adhäsine aus *Pichia*-Stämmen (orange) sind hingegen sehr klein und die Sequenzen scheinen weniger variabel. Ein weiterer distinkter Subtyp putativer Adhäsine besteht hauptsächlich aus Sequenzen welche zu dem Ascomycet *Kluyveromyces lactis* und Verwandten gehören (grau). Ein großer, weitgefaster Subtyp enthält neben Sequenzen aus einer großen Anzahl unterschiedlichster Hefen auch Sequenzen aus Insekten- und Pflanzenpathogenen (braun). Entfernte Verwandte, welche eine größere Ähnlichkeit zur PA14-Domäne selbst zeigen, bilden eine weiter distanzierte, kleine Gruppe (rot).

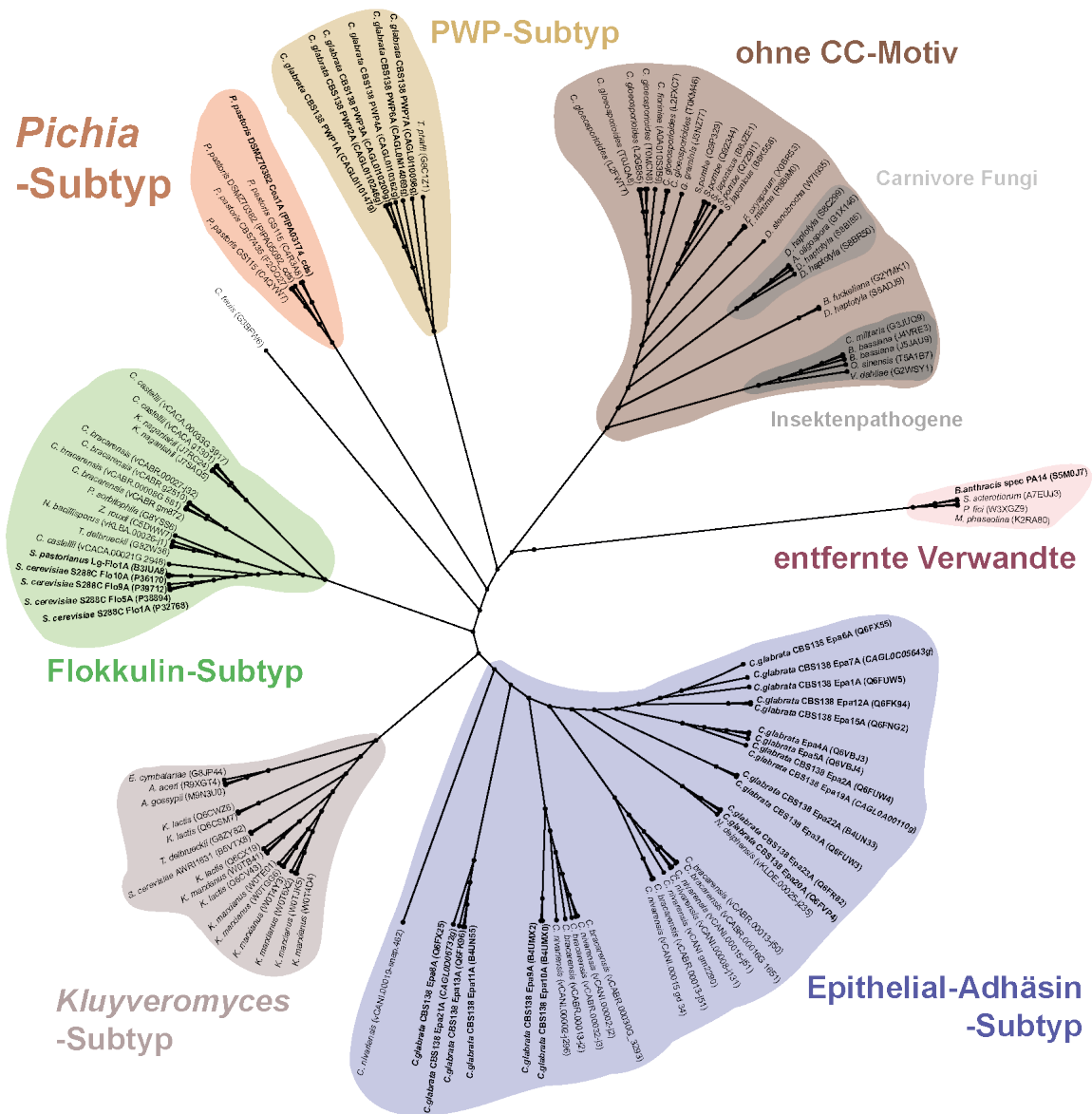


Abbildung 4.2 –Phylogenetische Analyse der ascomycetalen Adhäsine/putativen Adhäsine mit PA14 oder Gleya-Domänen. Die Sequenzen bilden definierte Cluster, welche farblich gekennzeichnet sind: *Pichia*-Klasse (orange), Flokkulin-Klasse (grün), Epa-Klasse (blau), *K. marxianus*-Klasse (grau), PWP-Klasse (gelb), Insekten- / Pflanzenpathogene (braun) & entfernte Verwandte (Rot). Einige Gruppen wie die Epa-Klasse zeigen eine starke Diversität, während bspw. die Gruppe der putativen *Pichia*-Adhäsine einen vergleichsweise kleinen Cluster bilden. Dies spricht unter anderem für eine geringere Variabilität auf Gen-Ebene.

Die pilzlichen Adhäsionsdomänen des PA14-Typs scheinen sich damit neben einer funktionellen Klassifizierung in bspw. Flokkuline und epitheliale Adhäsine auch phylogenetisch in Subtypen aufzuteilen. Ein Alignment (Abbildung 4.3 A) von jeweils einer repräsentativen Sequenz pro Subtyp gibt einen Überblick über die Unterschiede zwischen den einzelnen Subtypen.

Im Vergleich fällt auf, dass große Subdomänen an der Position des flexiblen Loops L2, wie die Flo5-Subdomäne, hauptsächlich im Subtyp der Flokkuline auftreten. Größere Einschübe in L1 treten vorwiegend bei Epa-ähnlichen Adhäsinen sowie bei putativen Adhäsinen aus *K. marxianus* und in PWP's auf. Cea1A aus *Pichia pastoris*, welches seinen Namen aufgrund der im Rahmen dieser Arbeit charakterisierten Erkennung von terminalen Chitin-Enden trägt (*Chitin End Adhesin 1*), zeigt keine Insertionen in diesen Bereichen.

Die mittleren Bereiche mit auffallend großer Ähnlichkeit stellen Bestandteile des strukturell konservierten Faltungsmotivs der Adhäsine dar, eines β -Sandwichs vom PA14/Flo5-Typ. Während alle Sequenzen, außer die der PA14-Domäne selbst, das prominente DD-Motiv (blau hinterlegt) in CBL1 tragen, sind größere Unterschiede in CBL2 (rot hinterlegt) zu erkennen. Während an Position I oft ein Arginin zu finden ist, außer bei Flo5 und Cea1, welche ein Alanin zeigen, sind die weiteren Positionen auch innerhalb der Subtypen sehr variabel. Augenscheinlich tragen Mitglieder des *K. marxianus* Subtyps sehr voluminöse Reste in CBL2, wie zwei konsekutive Tyrosine. Das putative Adhäsine aus dem Subtyp der Insekten- / Pflanzenpathogene hingegen zeigt sehr kleine Reste (G283, G284, P285) an dieser Position.

Vergleicht man die Cysteine, welche in den Subtypen der Flokkuline und der epithelialen Adhäsine Disulfidbrücken zeigen, fällt auf, dass das erste und letzte Cystein stark konserviert ist. Allen Mitgliedern des Insekten- / Pflanzenpathogene-Subtyps fehlt das prominente CC-Motiv hingegen. In den PWP's auf der anderen Seite ist dieses Motiv auf ein Cystein reduziert. Allerdings findet sich bei diesen ein zusätzliches Cystein in der Nähe des C-Terminus der Domäne. Hiermit zeigt sich, dass die Subtyp-interne Konservierung der Cysteine und damit der putativen Disulfidbrücken ein Charakteristikum des jeweiligen Subtyps darstellt.

In Kombination mit den prozentualen Identitäten und Ähnlichkeiten (Abbildung 4.3 B), lässt sich feststellen, dass die im Alignment abgebildeten Sequenzen zwar stark konservierte Bereiche zeigen, allerdings die Zuordnung zu Subtypen nicht willkürlich, sondern auf variablen Bereichen basiert, welche diesen putativen Adhäsinen möglicherweise spezifische Funktionen verleihen können.

Um ein tieferes Verständnis dieser diversen Hefezellwandproteinfamilie zu erhalten und mögliche neuartige Selektivitäten und Bindungsmodi zu entdecken, wurden im Rahmen dieser Arbeit mehrere Mitglieder strukturell und funktionell untersucht.

4.2 Adhäsiene aus *P. pastoris* – Charakterisierung eines neuen Subtyps

Zur Untersuchung einer mittels phylogenetischer Untersuchung identifizierten neuartigen Subtyps pilzlicher Adhäsiene wurden exemplarisch putative Adhäsiene aus *Pichia pastoris* gewählt (Abbildung 4.2, orange). Diese Spezies ist gut bekannt und ein weit verbreitetes Expressionssystem, doch stehen zu Ökologie und Adhäsionsverhalten keine detaillierten Informationen zur Verfügung (siehe Einleitung).

Die Klonierung geeigneter Plasmide zur rekombinanten Expression von Cea1A und *PpFlo1A* in *E. coli* wurden bereits in früheren Arbeiten durchgeführt^{155,156}.

4.2.1 Expression, Isolation und Aufreinigung von Cea1A

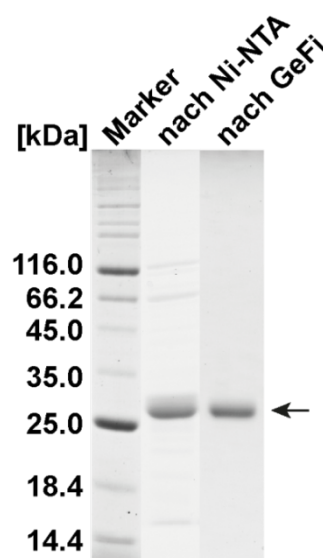


Abbildung 4.4 – SDS-PAGE von Cea1A (26.7 kDa): Neben dem Größenstandard ist eine Probe nach Ni-NTA- sowie nach nachfolgender Größenausschlusschromatografie gezeigt

Die Produktion von Cea1A wurde mit Hilfe des T7-Expressionssystems in *E. coli* Shuffle T7 Express-Zellen bei 12 °C über einen Zeitraum von 72 h durchgeführt. Nach dem Aufschluss der Zellen mittels French-Press oder Glaskugelmühle und Klärung des Lysats wurde die Isolation und Aufreinigung von Cea1A mit Ni-NTA-Affinitätschromatographie und nachfolgender Größenausschlusschromatografie vollzogen. Wie in früheren Arbeiten konnte bei letzterem Schritt eine starke Retention von Cea1A auf der Säule beobachtet werden. Eine Bindung an die Glycanstruktur der Säulenmatrix wurde hierfür als Auslöser angeführt¹⁵⁵. Der Erfolg der Isolation und Aufreinigung kann anhand der SDS-PAGE in Abbildung 4.4 abgelesen werden. Es konnten nach beiden Chromatografien etwa 80 mg reines, monomeres Cea1A pro Liter Expressionskultur erhalten werden. Eine Lagerung des Proteins erfolgte über mehrere Wochen bei 4 °C.

4.2.2 Untersuchungen zur Glycanbindung von Cea1A

Obwohl die Bindung von Cea1A an die Säule der Größenausschlusschromatografie bereits erste Hinweise auf die Bindung an Glycanen lieferte, wurde die Spezifität und die Qualität der Bindung mit Hilfe einer chipbasierten *Glycan-Array*-Analyse untersucht (Abschnitt 3.4.4). Fluoreszenzmarkiertes Cea1A wurde hierfür in Konzentrationen von 20 µg/mL und 200 µg/mL auf einen mit diversen Glycanen besetzten Glas-Chip gegeben.

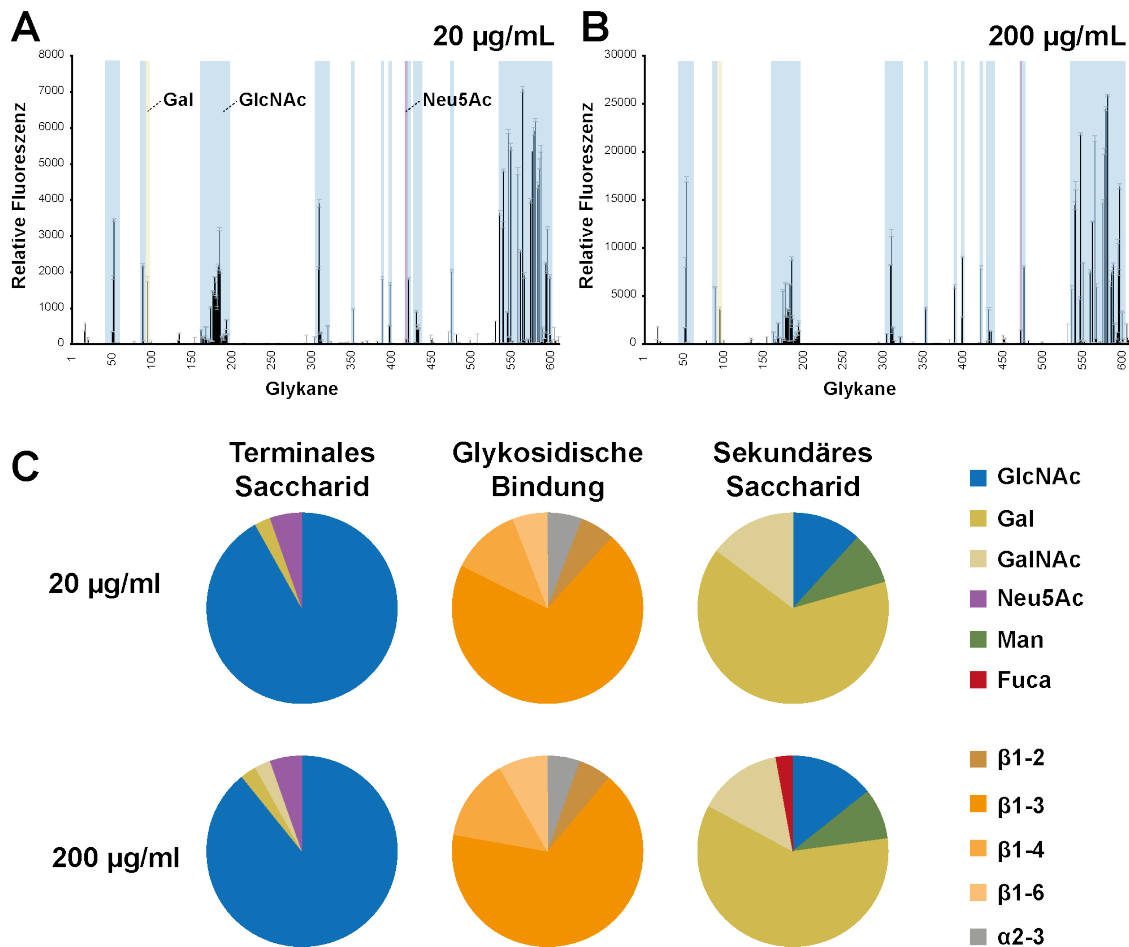


Abbildung 4.5 – *Glycan-Array* Analyse in Kooperation mit dem *Consortium for Functional Glycomics*. A: Auslesung der relativen Fluoreszenz bei einer Cea1A-Konzentration von 20 µg/mL. B: Auslesung der relativen Fluoreszenz bei einer Cea1A-Konzentration von 200 µg/mL. C: Quantifizierung der gebundenen, eindeutigen Glycane.

Unter den 611 vorhandenen Glycan-Strukturen zeigte Cea1A eine eindeutige Präferenz für die Bindung von endständigem *N*-Acetylglucosamin (GlcNAc), welches den Baustein weit verbreiteter biologischer Polymere wie Chitin darstellt (Abbildung 4.5 A & B). Kein anderer endständiger Zuckerrest war in der Lage, eine vergleichbar starke relative Fluoreszenz zu liefern.

Zur Quantifizierung der Spezifität wurde der Anteil eindeutiger Glycanstrukturen^{sss} gezählt, welche mindestens 4% der maximal beobachteten relativen Fluoreszenz erreichten; diese untere Schwelle entspricht einem einzelnen GlcNAc-Rest welcher über einen Linker direkt an den Chip gebunden ist (Abbildung 4.5 C). Während hierbei die Spezifität für den terminalen Rest sehr stark ist, scheint die Auswahl der folgenden glycosidischen Bindung sowie des zweiten Zuckers selbst weniger determiniert. Es ist zu beobachten dass in den *Glycan-Array* Studien ebenfalls *N*-Acetylneuraminsäure (Neu5Ac, eine Sialinsäure), ein Saccharid der Glycokalyx, eine Bindung an Cea1A zeigte.

^{sss} Unverzweigte Glycane und solche, welche am Terminus von Verzweigungen den gleichen Zucker tragen.

Isothermale titrationskalorimetrische Messungen mit GlcNAc (orange) und dem Dimer *N,N'*-Diacetylchitobiose (grün) (5 mM in SEC-Puffer, Abbildung 4.6 A) zeigten, dass die Bindung an GlcNAc nicht nur durch eine hohe Spezifität ausgezeichnet ist, sondern ebenfalls durch eine relativ hohe Affinität im niedrigen μM -Bereich (Abbildung 4.6 B). Diese ist mit der Bindung von Epa1A an galactosidische Liganden vergleichbar, welche hierbei für eine starke, pathogene Interaktion sorgt. Zum weiteren Vergleich dienen Flo5A und Lg-Flo1A, deren Bindungsaffinität an Mannose respektive Mannosephosphat im mittleren bis hohen mM Bereich liegt^{107,108,135}.

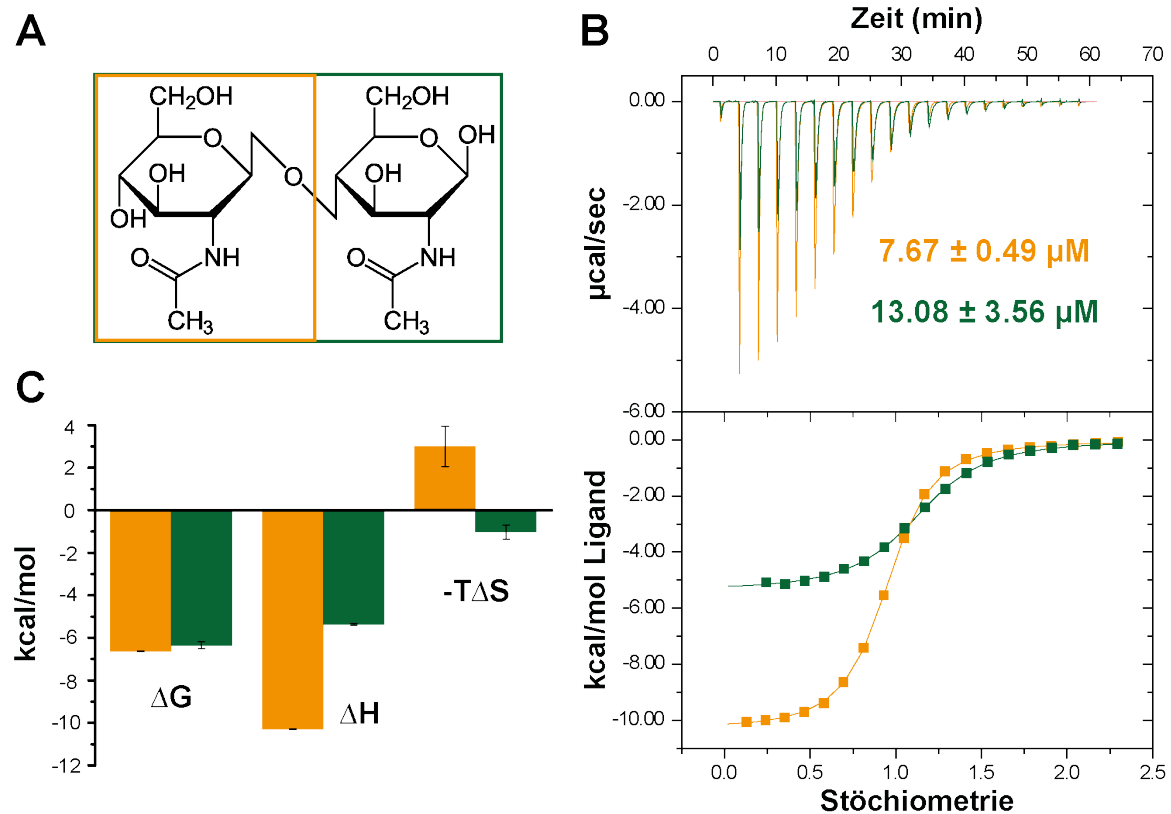


Abbildung 4.6 – ITC Experimente mit GlcNAc (orange) und Chitobiose (grün). A: Strukturen beider Liganden. B: Vergleich der ITC-Experimente beider Liganden, oben Rohdaten, unten Bindungsisotherme. C: Thermodynamische Auswertung.

Beide Liganden binden in einem exothermen Prozess an Cea1A, allerdings zeigt die Bindung von Chitobiose einen kleineren enthalpischen Beitrag (ΔH) als die von GlcNAc. Da die auftretende freie Energie (ΔG) allerdings in der gleichen Größenordnung liegt, ergibt sich eine Differenz im entropischen Beitrag (ΔS) der Bindungs-Thermodynamik (Abbildung 4.6 C & Tabelle 4.2). Die Bindung des Monomers scheint entropisch ungünstig, während bei der Bindung an Chitobiose die Entropie geringfügig steigt. Der Grund könnte die Freisetzung einer größeren Anzahl von Wasser-Molekülen bei Bindung des Dimers sein³¹⁰. In früheren Arbeiten konnte bereits gezeigt werden, dass die Bindung durch Zugabe von 5 mM EDTA nahezu vollständig zum Erliegen kommt, was eine Bestätigung der C-Typ Lektin-artigen Bindung darstellt¹⁵⁵.

Tabelle 4.2 – Thermodynamische Parameter der Bindung von Cea1A an GlcNAc und Chitobiose.

Thermodynamischer Parameter	Cea1A•GlcNAc	Cea1A•Chitobiose
K_D (μM)	7.67 \pm 0.49	13.08 \pm 3.56
ΔH ($\text{kJ}\cdot\text{mol}^{-1}$)	-43.03 \pm 0.06	-22.46 \pm 0.20
$-T\Delta S$ ($\text{kJ}\cdot\text{mol}^{-1}$)	12.55 \pm 3.94	-4.33 \pm 1.39
ΔG ($\text{kJ}\cdot\text{mol}^{-1}$)	-27.76 \pm 0.11	-26.57 \pm 0.67
N (Bindungsstellen)	0.919 \pm 0.001	1.046 \pm 0.007

Die Titration von Cea1A mit Glucose (blau) oder Glucosamin (rot) (jeweils 5 mM in SEC-Puffer) hingegen lieferte keine signifikante Bindung (Abbildung 4.7), was den Einfluss der Acetylgruppe (orange) auf die Bindung unterstreicht. Ebenfalls die vom Glycan-Array suggerierte Bindung von Cea1A an Neu5Ac (schwarz) zeigte kein Signal (5 mM in SEC-Puffer). Hierbei könnte eine unvollständige Synthese bzw. Kettenabbrüche der synthetischen Glycane auf dem Chip für die falsch-positiven Ergebnisse verantwortlich sein; jede dieser Ketten zeigt an den Positionen 3-5 ebenfalls ein GlcNAc.

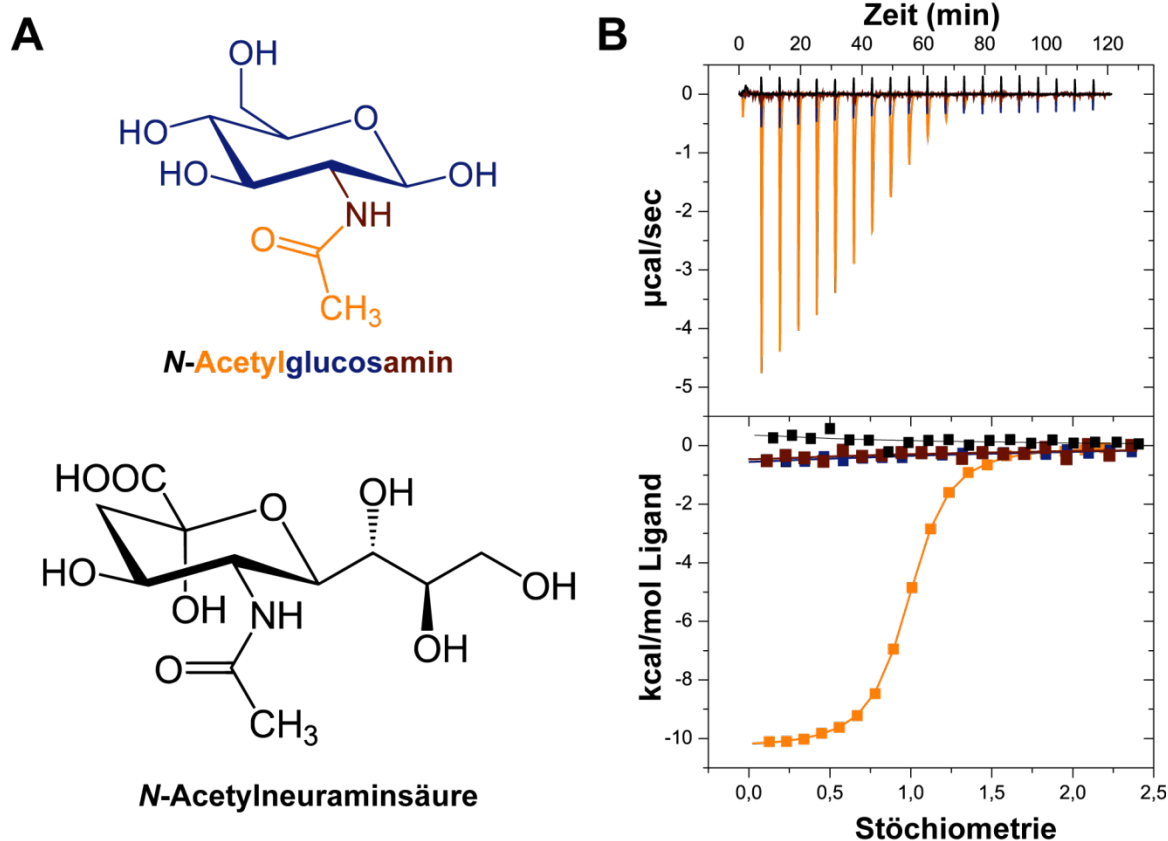


Abbildung 4.7 – ITC-Experimente mit weiteren Sacchariden. A: Um den Einfluss der funktionellen Gruppen von GlcNAc auf die Bindung zu zeigen wurde Glucose (blau), Glucosamin (rot), und N-Acetylneuraminsäure (schwarz) zu Cea1A gegeben. B: Die Bindungswärmen (oben) sowie -isothermen (unten) zeigen deutlich den Einfluss der N-Acetylgruppe (orange) auf die Bindung.

4.2.3 Kristallisation von Cea1A und Strukturlösung mittels molekularen Ersatzes

Cea1A konnte bereits zuvor erfolgreich als Apoprotein kristallisiert werden. Das Streuvermögen der Kristalle war allerdings nicht ausreichend um die Struktur lösen zu können. Hierbei konnte bei einer Auflösung von ~ 4 Å eine Vollständigkeit von lediglich knapp 25% erzielt werden¹⁵⁵.

Die Zugabe von 5 mM der in *Glycan-Array*- und ITC-Analysen als passende Liganden identifizierten Saccharide GlcNAc bzw. Chitobiose in Kombination mit 5 mM CaCl_2 führte daraufhin zur Bildung dreidimensionaler, orthorhombischer Kristalle in den Bedingungen 60 & 87 des kommerziellen Screens MBC II bei 4 °C. Zusammen mit Chitobiose konnten weiterhin nadelförmige Kristalle in Bedingung 82 dieses Screens erhalten werden (Tabelle 4.3). Bei allen Kristallisationsexperimenten wurde Cea1A in einer Konzentration von 15-30 mg/mL in SEC-Puffer verwendet (Abbildung 4.8).

Tabelle 4.3 – Kristallisationsbedingungen von Cea1A.

MBC #60		MBC #82		MBC #87	
MgCl_2	100 mM	NaOAc	200 mM	Li_2SO_4	500 mM
NaCl	100 mM	MES	100 mM	PEG 8000	10%
Na_3citrat	100 mM	pH	6.5	-	-
pH	3.5	PEG 8000	10%	-	-
PEG 4000	12%	-	-	-	-

Mikroseeding mit Kristallen aus Bedingung 60 sowie Reproduktion & Verfeinerung der Bedingung im 24 well / *hanging drop* Maßstab führte zu dreidimensionalen, orthorhombischen Einkristallen mit einer Größe von bis zu etwa 200 μm .

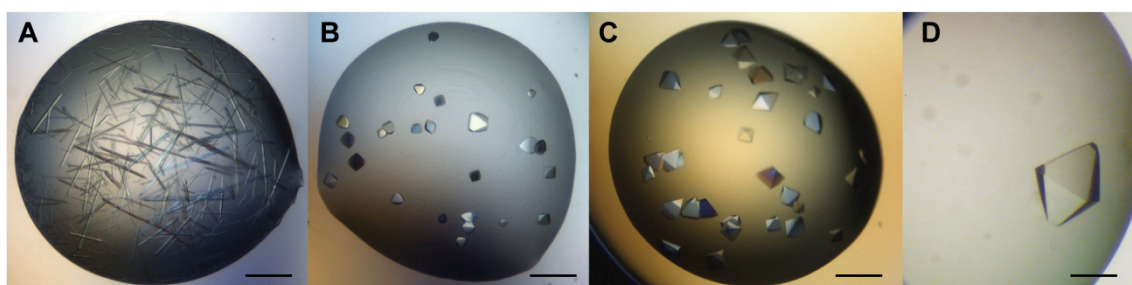


Abbildung 4.8 – Kristalle von Cea1A. A: Nadelförmige Kristalle aus Bedingung 82. B: Kleinere orthorhombische Kristalle aus Bedingung 87. C: Große dreidimensionale, orthorhombische Kristalle aus Bedingung 60. D: Kristalle nach Optimierung und Mikroseeding. Größenmaßstab A-C: 100 μm ; D: 200 μm .

Mit diesen Kokristallen von Cea1A mit beiden Zuckern wurden nach initialer Charakterisierung an einer *inhouse* Röntgenquelle Datensätze im Auflösungsbereich bis ~ 1.7 Å am BESSY II-Synchrotron aufgenommen (siehe Tabelle 4.4). Die Integration und Datenreduktion wurde mit *XDS* sowie *XSCALE* durchgeführt und Testreflexe aufgrund der vorhandenen nicht-

kristallographischen Symmetrie in dünnen Auflösungsschalen mit *DATAMAN* ausgewählt^{277,298}.

Tabelle 4.4 – Datensammlungsstatistik von Cea1A-Kokristallen.

	Cea1A•GlcNAc	Cea1A•Chitobiose
Datensatz	MK113	MK120
Datum	25.05.2012	16.11.2012
Röntgenquelle	BESSY II BL14.1	BESSY II BL14.3
Detektor	MARmosaic 255 mm	
Temperatur (K)	100 K	
Wellenlänge (Å)	0.91841	0.89120
Raumgruppe	<i>P</i> ₂ ₁ ₂ ₁	<i>P</i> ₂ ₁ ₂ ₁
Zellmaße (Å)		
a	102.34	101.70
b	106.21	105.35
c	107.60	106.47
α, β, γ	90°, 90°, 90°	90°, 90°, 90°
Moleküle / a.s.u.	4	4
Auflösung (Å)	19.84-1.66 (1.75-1.66)	19.98-1.75 (1.84-1.75)
Beobachtete Reflexe	797571	423782
Multiplizität	5.8 (5.8)	4.0 (4.0)
Einzigartige Reflexe	137801	106663
R_{merge} (%)	9.9 (59.0)	9.9 (52.3)
Vollständigkeit (%)	99.6 (100.0)	92.7 (95.3)
Mittleres $I/\sigma(I)$	14.9 (3.3)	12.4 (2.8)
Mosaizität (°)	0.12	0.09
Wilson B-Faktor (Å²)	10.7	8.9

Aufgrund der aus der phylogenetischen Analyse zu erwartenden strukturellen Ähnlichkeit zu den bekannten Adhäsins-Strukturen von Flo5A und Epa1A wurde eine Lösung des Phasenproblems mittels molekularem Ersatzes in *phaser* versucht²⁸⁹. Hierzu wurde auf Epa1A (PDB-ID: 4ASL) als Suchmodell zurückgegriffen, da wie in Abbildung 4.2 A&B zu erkennen ist, Cea1A eine höhere Ähnlichkeit auf Primärsequenzebene zu Epa1A als zu Flo5A aufweist (43% vs. 40%). Des Weiteren konnte aufgrund des Alignments erwartet werden, dass, wenn keine Subdomäne in Cea1A vorhanden ist, die Faltung vergleichbar mit Epa1A sein sollte.

Die Verwendung des vollständigen Modells von Epa1A^{****} und der Primärsequenz von Cea1A in *phaser*²⁸⁹ führte auch mit unterschiedlichen Suchparametern und Auflösungsgrenzen nicht zum gewünschten Erfolg (Abbildung 4.9 A). Hauptsächlich führte das Suchmodell zu häufigen Zusammenstößen in der Kristallpackung, weshalb keine geeignete Lösung erhalten werden konnte. Um dies zu umgehen, wurden alle flexiblen Bereiche, wie Loops L1-L3, sowie

**** Die Vollständigkeit ist hier nur auf die Polypeptidkette bezogen; um den Einfluss der Phasen des genutzten Modells so gering wie möglich zu halten, wurden alle anderen Atome entfernt.

der Bereich der Termini mit dem konservierten CC-Motiv entfernt (Abbildung 4.9 B). Das somit nur noch aus dem β -Sandwich bestehenden Modell führte ebenfalls zu keiner Lösung. Das primäre Problem in diesem Fall war, dass das Suchmodell nicht korrekt platziert werden konnte; offenbar beschreibt dieses Kernmotiv die Struktur nicht ausreichend um eine eindeutige Anordnung zu erhalten. Somit wurde die Verbindung der Vorzüge beider Modelle, möglichst ähnliche globuläre Form bei gegebener Freiheit auf Ebene der flexiblen Bereiche, getestet. Dies sollte quasi *a minori ad maius* von einer möglichst erfolgreichen Platzierung der 4 Moleküle in der asymmetrischen Einheit ebenfalls zu wenigen bzw. keinen Zusammenstößen in der Packung führen. Dazu wurde mittels *modeler9v7²⁹²* ein Homologiemodell von Cea1A auf Basis von Epa1A erstellt und die flexiblen Bereiche bis auf die Fixierung der Termini entfernt (Abbildung 4.9 C).

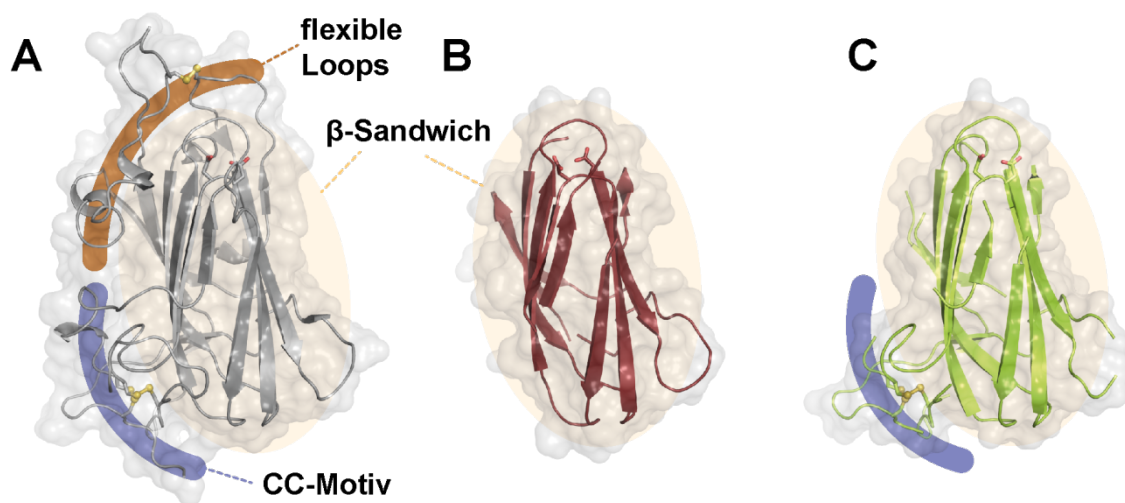


Abbildung 4.9 – Modelle für den molekularen Ersatz. A: Zu Beginn wurde auf ein vollständiges Modell von Epa1A zurückgegriffen (PDB-ID: 4ASL). Dies führte zu einer großen Anzahl von Zusammenstößen in der Kristallpackung. B: Stark bearbeitetes Modell, bei welchem folgende Bestandteile entfernt wurden: flexible Loops, Termini & CC-Motiv zur Fixierung der Termini. Auch dies führte nicht zur Lösung des Phasenproblems C: Die Verwendung eines Homologiemodells von Cea1A auf Basis von Epa1A mit entfernten flexiblen Bereichen aber erhaltenem CC-Motiv führte zum Erfolg.

Mit einer maximal genutzten Auflösung von 3.5 Å konnte damit eine einzige Lösung in der Raumgruppe $P2_12_12_1$ erhalten werden. Zwar wurden keine Zusammenstöße in der Kristallpackung entdeckt, die erhaltenen Translations-Z-Werte (TFZ) für die 4 Moleküle waren allerdings eher gering. Werte zwischen 5-6 weisen generell darauf hin dass eine korrekte Lösung eher noch unwahrscheinlich ist, hier wurden Werte von 5.5, 4.9, 5.4 & 5.1 für die einzelnen Moleküle erhalten. Eine Inspektion der erhaltenen Elektronendichtekarte zeigte dennoch eine gute Korrelation mit dem von *phaser* generierten Modell. Dessen Korrektheit konnte durch erfolgreiche Verfeinerung ($R_{\text{work}}=17.1\%$; $R_{\text{free}}=20.2\%$; Tabelle 4.5) und vollständige Beschreibung der Elektronendichte bestätigt werden. Die Ursache der geringen statistischen TFZ-Werte dieser „borderline“-Lösung wurde bei detaillierter Untersuchung der endgültigen Struktur ersichtlich.

Tabelle 4.5 – Verfeinerungsstatistik für Cea1A-Komplexe.

	Cea1A•GlcNAc	Cea1A•Chitobiose
Auflösung (Å)	19.84-1.66	19.99-1.75
R _{works} , R _{free} (%)	17.07, 20.21	18.26, 21.14
Reflexe (Arbeits-, Test-)	135716, 2084	105044, 1619
Vollständigkeit (%)	99.46	92.20
r.m.s.d.		
Bindungslängen(Å)	0.011	0.012
Bindungswinkel (°)	1.57	1.63
Anzahl Atome	7996	8035
Mittlerer B-Wert (Å ²)	16.28	14.98

4.2.4 Die Struktur von Cea1A

Der molekulare Ersatz lieferte eine Lösung mit 4 Molekülen in der asymmetrischen Einheit. Diese sind mit einer pseudo-222-Punktsymmetrie antiparallel zueinander angeordnet (Abbildung 4.10) und unterscheiden sich aufgrund individueller Kristallkontakte mit einem maximalen r.m.s.d. von 0.22 Å nur sehr geringfügig voneinander. Diese Differenzen beschränken sich zudem bis auf die Seitenkettenkonformation von Arginin 206 in der Bindungstasche nur auf den Bereich der Termini. Daher erfolgt die weitere Betrachtung anhand des Referenzmoleküls A.

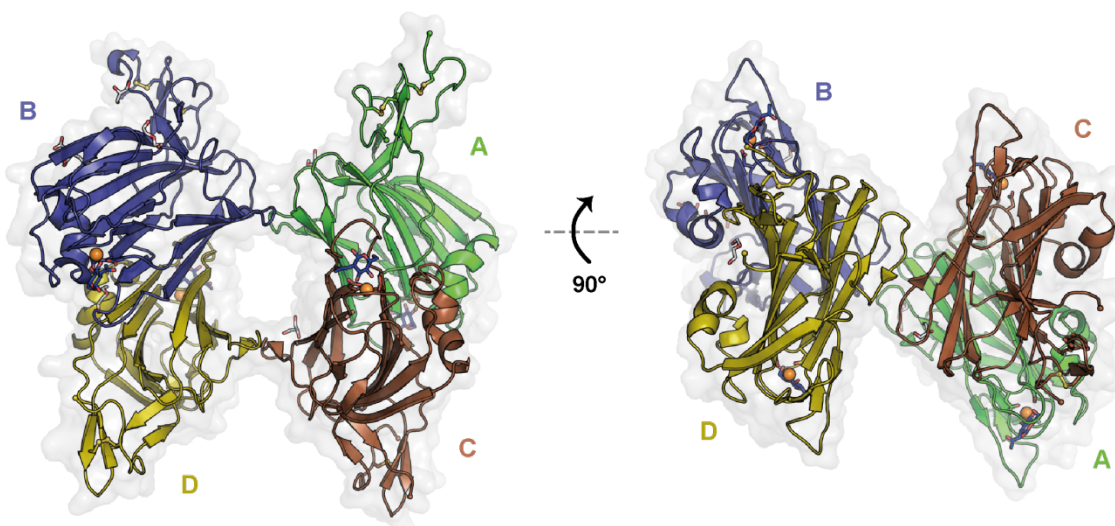


Abbildung 4.10 – Orientierung der 4 Moleküle der asymmetrischen Einheit. Die Cea1A-Moleküle bilden *in crystallo* ein pseudo-222-punktsymmetrisches Tetramer. Die Moleküle unterscheiden sich aufgrund individueller Kristallkontakte nur geringfügig voneinander.

Als Mitglied der PA14-Domänenfamilie zeigt Cea1A ein ausgeprägtes β -Sandwich. Dieses strukturell konservierte Motiv besteht aus 10 β -Strängen, wobei ein 6-strängiges β -Faltblatt ($\beta_3\beta_2\beta_5\beta_{12}\beta_7\beta_{10}$) gegen ein kleineres 4-strängiges β -Faltblatt ($\beta_4/14\beta_{13}\beta_6\beta_{11}$) gepackt ist (Abbildung 4.11). Neben dem D_{cis}D-Motiv in CBL1 (Inlay, hellbraun), welches ein Ca²⁺-Ion in Kombination mit einem Asparagin aus CBL2 koordiniert (Inlay, dunkelbraun),

formt letzterer zusammen mit den flexiblen Loops L1-L3 die Bindungstasche. Die in diesen Bereichen exponierten Aminosäuren sorgen für die Spezifität und in Kombination mit dem durch das *DcisD*-Motiv komplexierte Ca^{2+} -Ion für die Affinität der C-Typ Lektin-artigen Bindung an *N*-Acetylglucosamin (GlcNAc).

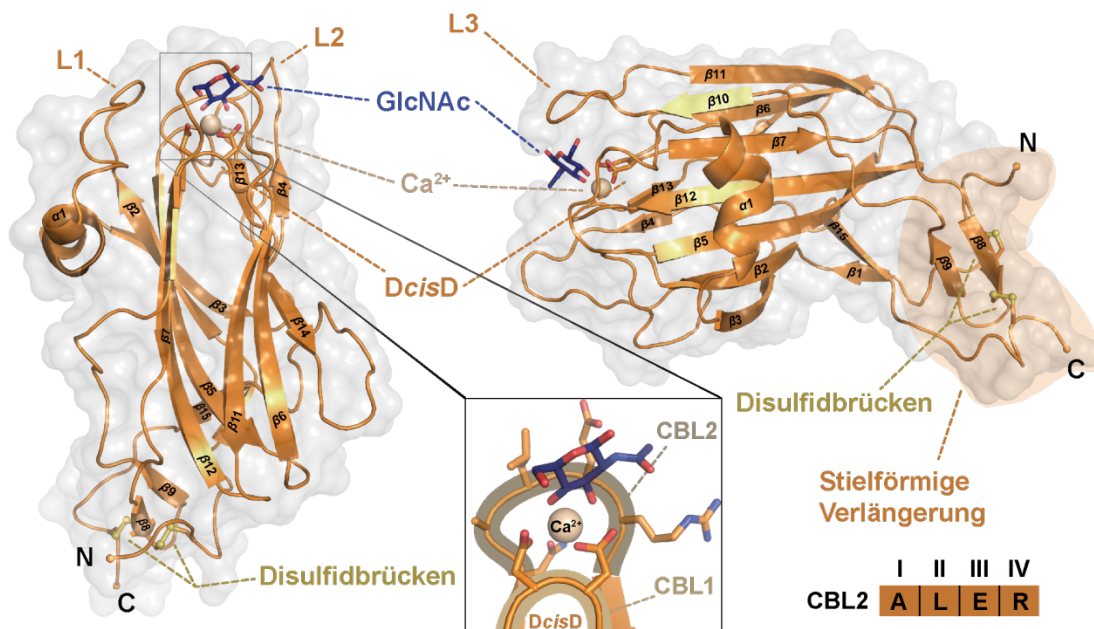


Abbildung 4.11 – Die Struktur von Cea1A. Neben dem konservierten β -Sandwich und den Disulfidbrücken zeigt Cea1A eine spezifische stielförmige Verlängerung (orange hinterlegt). Der Loop CBL1 (hellbraun) komplexiert mit Hilfe des *DcisD*-Motivs ein Ca^{2+} -Ion (ocker) über welches der Ligand GlcNAc (blau) gebunden wird. Hauptsächlich die Reste aus CBL2 (dunkelbraun) vermitteln die Selektivität und die Hauptinteraktionen mit dem terminalen Zucker.

Der hypervariable Loop L1 zeigt, wie in Flo5A und Lg-Flo1A, keine Disulfidbrücke zur Fixierung von L1 an L2, wie diese in Epa1A vorkommt. Die Mobilität von L1 wird in Cea1A jedoch durch einen hydrophoben Bereich zwischen diesem und der PA14-Domäne selbst eingeschränkt (Abbildung 4.12 A). Die Disulfidbrücke zur Fixierung dieses Loops L1 in Epa1A ist aufgrund seiner Länge offenbar nötig um dessen Beitrag zur Form der Bindungstasche zu erhalten. Der zweite flexible Loop (L2) enthält keine Flokkulin-artige Subdomäne, er ist im Vergleich der kürzeste. Ein weiterer Unterscheid kann in L3 festgestellt werden. Dieser weist eine deckelförmige Konformation auf und umschließt das terminale Saccharid und die Bindungstasche. In Flo5A und Lg-Flo1A hingegen ist dieser Loop vom Liganden weggeklappt und zeigt nur einen untergeordneten Beitrag zur Erkennung des Glycans. Epa1A trägt an dieser Stelle einen in der Epa-Familie konservierten Tryptophan-Rest, welcher direkt in die Interaktion mit Galactosiden eingebunden ist.

Für Flo5A wurde eine Stabilisierung des *DcisD*-Motivs durch eine Wasserstoffbrücke zwischen einem Tyrosin im β -Sandwich und dem Peptidrückgrat des Motivs beobachtet¹⁰⁷. In Cea1A ist dieses Y222 gegen ein F200 ausgetauscht, welches den Einfluss dieser Wasserstoffbrückenbindung in Frage stellt (Abbildung 4.12 B).

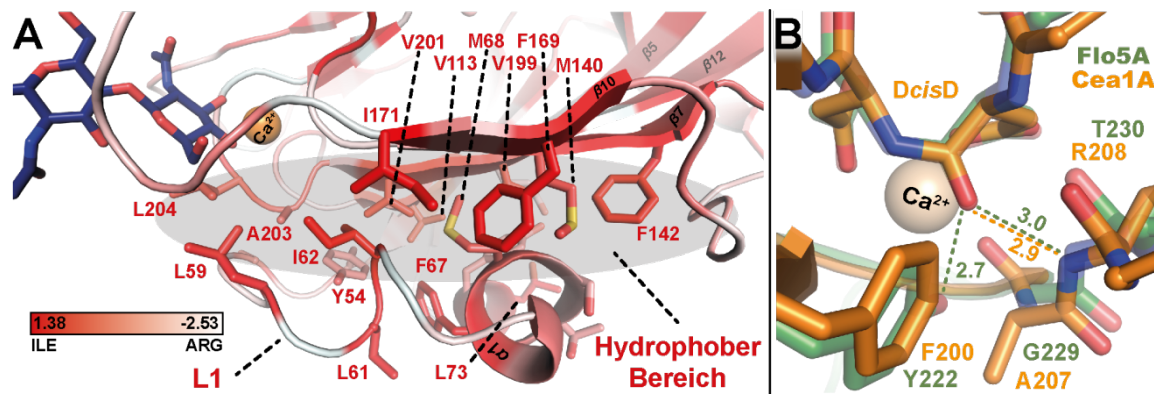


Abbildung 4.12 - Strukturelle Details von Cea1A. A: Hydrophober Bereich zur Fixierung von L1. Die Struktur ist anhand der EISENBERG'schen Hydrophobizitätsskala eingefärbt³¹¹. B: Stabilisierung des DcisD-Motivs im Vergleich zu Flo5A.

Ein einzigartiger struktureller Bestandteil stellt die stielartige Verlängerung des Bereiches um die beiden Termini dar. Obwohl die Cysteine, welche die beiden konsekutiven Disulfidbrücken bilden, hochkonserviert sind, zeigt sich eine alternative Faltung in Cea1A im Vergleich mit anderen bekannten Adhäsinstrukturen mit PA14-Domäne. Statt eines helikalen/loop-förmigen Bereiches wird der Übergang zur repetitiven B-Region in diesem Fall ausgehend von $\beta 7$ über $\beta 8$ - $\beta 9$ als β -Schleife geformt (Abbildung 4.13 A & B). In Epa1A sowie in den weiteren bekannten Strukturen ist diese Region in der Regel als kurze α -Helix oder als Loop-Bereich direkt am β -Sandwich fixiert. Als Resultat ergäbe sich ein Winkel von etwa 55° zwischen A-Domäne und B-Region in Cea1A. Diese unvorhersehbare Konformation bzw. die Ausbildung dieser kleinen Subdomäne stellt vermutlich den Grund für den großen Anteil der suboptimalen Statistiken bei der Lösung des Phasenproblems mittels molekularen Ersatzes dar.

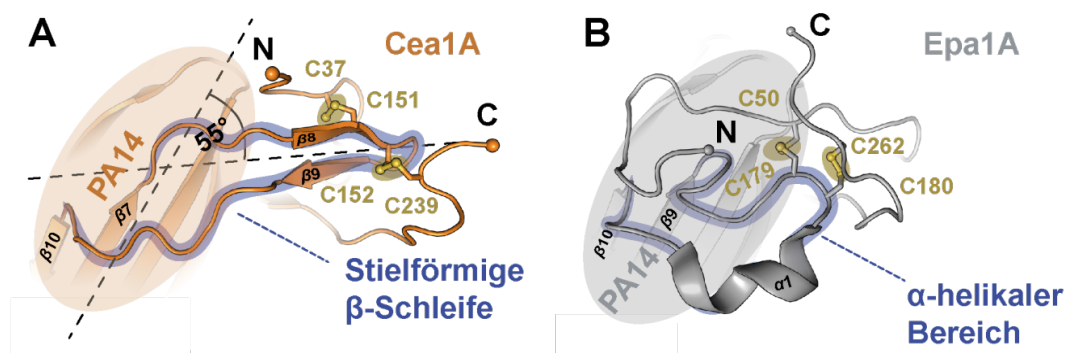


Abbildung 4.13 – Die Bereiche der Termini von Cea1A und Epa1A. Im Vergleich fällt der unterschiedliche Verlauf im Bereich der Fixierung über das CC-Motiv auf. Während Cea1A eine β -Schleife zeigt, welche einen Winkel zwischen der A-Domäne und der B-Region von etwa 55° erzeugen würde, ist in Epa1A sowie in den anderen bekannten Strukturen von Adhäsinen mit PA14-Domänen ein α -helikaler Bereich zu finden, welcher die Termini in eine andere Orientierung bringt.

Nachdem bereits die globale Analyse der Struktur Alleinstellungsmerkmale zeigt, verdeutlicht eine eingehende Betrachtung des Bindungsmodus von Cea1A an GlcNAc/Chitobiose die Eigenständigkeit des neu entdeckten Adhäsinsubtyps. Die Bindung des Liganden wird durch ein komplexes Netzwerk aus Wasserstoffbrückenbindungen und elektrostatischen

Wechselwirkungen herbeigeführt (Abbildung 4.14 A & B). Neben dem Ca^{2+} -Ion (ocker), welches durch das *DcisD*-Motiv sowie N202 aus CBL2 fixiert ist, wird die Bindung des terminalen GlcNAc hauptsächlich durch die Reste in CBL2 vermittelt. In Position I-IV von CBL2 sind dies die Aminosäuren A203, L204, E205 & R206.

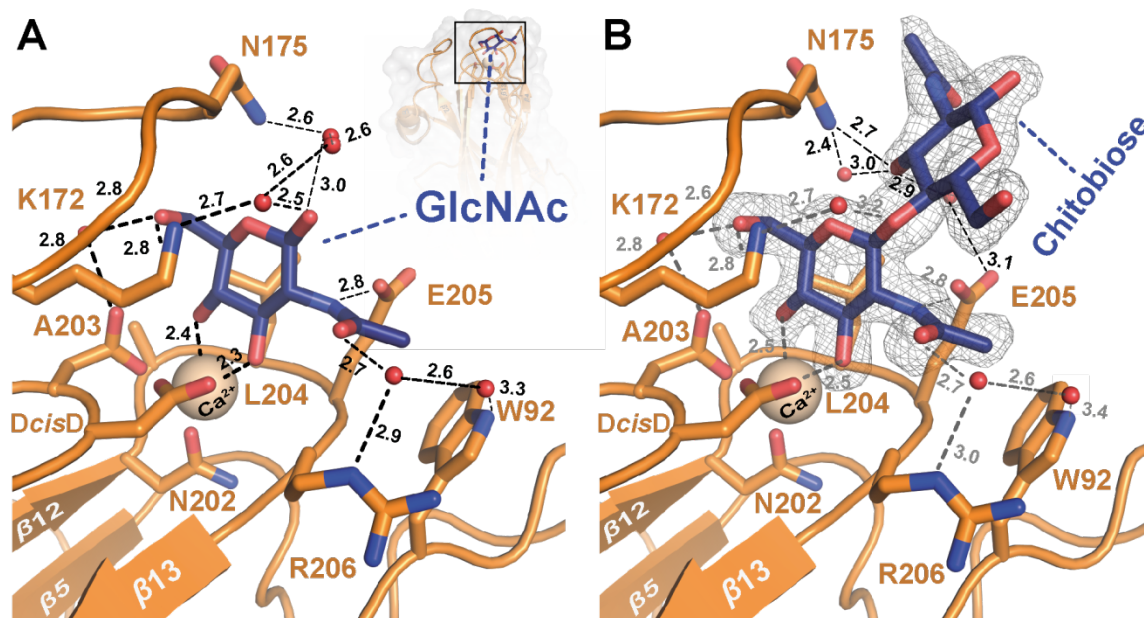


Abbildung 4.14 – Blick in die Bindungstasche von Cea1A. A: Die Bindung von GlcNAc führt zur Ausbildung eines komplexen Netzwerkes aus Wasserstoffbrückenbindungen sowie ladungsgesteuerten und hydrophoben Interaktionen. B: Durch die Bindung der dimeren Chitobiose wird der sekundäre Zucker ebenfalls in die Wechselwirkungen integriert, obwohl die Anzahl der direkten Kontakte mit dem zweiten Zucker untergeordneter ist. Die Chitobiose ist mit SigmaA-gewichteter $2F_{\text{obs}}-F_{\text{calc}}$ -Elektronendichte gezeigt (Konturierungsgrad 1.2σ).

Während A203 und L204 einen hydrophoben Bereich der Bindungstasche bilden um mit dem Kohlenstoffatom in Position 6 des GlcNAc zu interagieren, zeigt die restliche Tasche

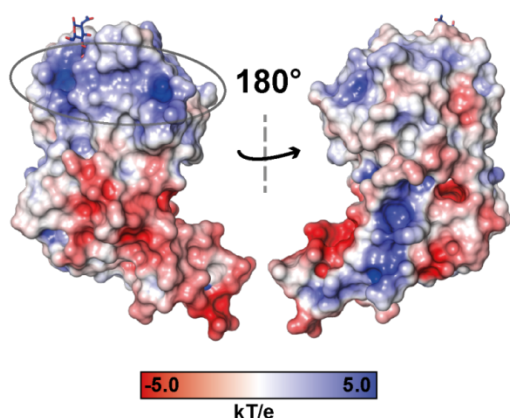


Abbildung 4.15 – Elektrostatistische Oberfläche von Cea1A. Die mit APBS bei 0.15 M NaCl berechnete Oberfläche zeigt dass die Bindungstasche von einem Bereich positiver Ladung umgeben ist.

einen eher polaren Charakter. Eine Rückschau auf die globalen Eigenschaften von Cea1A zeigt darüber hinaus eine Häufung positiver Ladung in der Nähe der Bindungstasche wie diese bei keinem der anderen Adhäsine beobachtet wurde, obwohl der berechnete pI 5.23 beträgt²²³ (Abbildung 4.15). Im Vergleich zu Flo5A ist in der Nähe der Bindungstasche eher eine Häufung negativer Ladungen zu beobachten. Dies verstärkt den Eindruck einer primär polaren Wechselwirkung. Hierbei interagiert E205 direkt mit dem Stickstoffatom des acetylierten Amids während R206 gemeinsam mit W92 aus L2 eine

Wasserstoffbrückenbindung zur Acetylgruppe beisteuert. Die Indolgruppe von W92 bildet eine kleine, hydrophobe Interaktionsfläche welche direkt in Richtung des Acetyl-Kohlenstoffs

orientiert ist. Der flexible Loop L3 spielt ebenfalls eine wichtige Rolle bei der Bindung des Glycans. Er zeigt eine Konformation, welche die Bindungstasche von oben wie ein Deckel gegen Lösungsmittel abschirmt. K172 interagiert zum Einen direkt mit der 6-OH-Gruppe und trägt zum Anderen zu einem komplexen Wasserstoffbrückennetzwerk über und um den monomeren Liganden herum bei.

Der zweite GlcNAc-Rest in Chitobiose führt zur Substitution einiger in Wasserstoffbrücken eingebundener Wassermoleküle. Hierbei wird allerdings auch deutlich, dass die Interaktionsfläche zwischen dem sekundären Zucker und Cea1A, abgesehen von einer Interaktion mit N175, eher klein ist. Dies ist im Einklang mit den Ergebnissen des *Glycan-Array* und zeigt damit, dass die Selektivität diesen zweiten Rest betreffend eher untergeordnet ist. Die in Abbildung 4.14 B dargestellte Elektronendichte der Chitobiose zeigt darüber hinaus deutlich, dass neben dem terminalen GlcNAc auch das sekundäre Saccharid sehr geordnet und räumlich fixiert in der Bindungstasche vorliegt.

Im Vergleich des Bindungsmodus von Cea1A mit der Situation in Flo5A und Epa1A fällt auf, dass die Orientierung des GlcNAc der Position der Mannose in Flo5A sehr stark ähnelt (r.m.s.d. des Hexosenrings 0.3 Å, Abbildung 4.16 A).

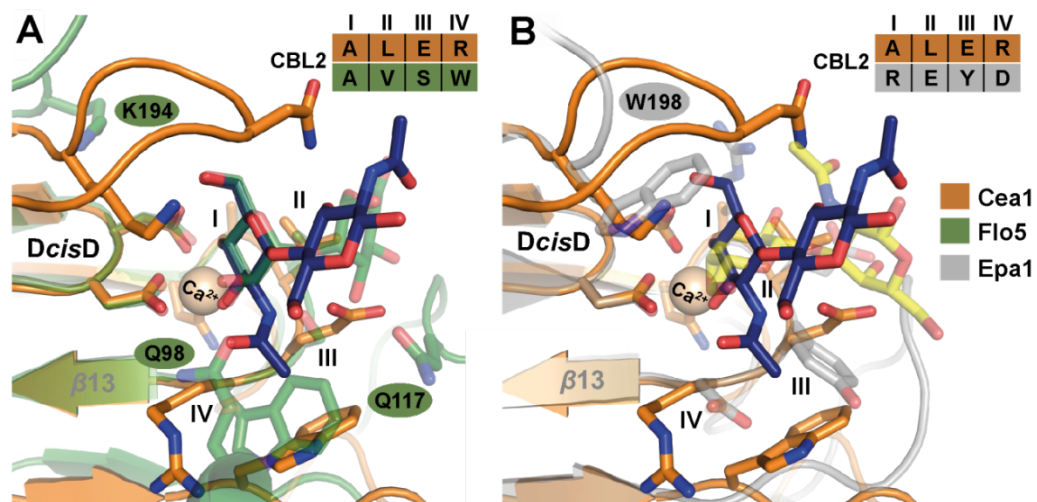


Abbildung 4.16 – Der Bindungsmodus von Cea1A im Vergleich. A: Die Orientierung des primären Zuckers in Flo5A (grün) und Cea1A (orange) ist sehr ähnlich, während die Orientierung des zweiten Rests sich aufgrund der unterschiedlichen glycosidischen Bindung (α 1,2 in mannobiose und β 1,4 in Chitobiose) stark unterscheidet. B: Zwar ist die Anordnung des gebundenen Saccharids in Epa1A (grau) deutlich abweichend von der Position von GlcNAc in Cea1A, doch ist der flexible Loop L3 in beiden Adhäsinen in eine direkte Interaktion mit dem Ligand eingebunden.

Position I und II in CBL2 sind ähnlich, sie bilden in beiden Adhäsinen einen hydrophoben Bereich zur Bindung des Hexosen-C-6-Kohlenstoffs. Der größte Unterschied ist in den beiden nächsten Positionen zu finden. Während Flo5 in Position III ein kleines Serin zur Bindung der α -Mannobiose trägt, ist in Cea1A an dieser Stelle ein größerer Aspartatrest zu finden, welcher die Bindung dieses Disaccharids sterisch verhindern würde. Ebenfalls fehlt ein Pendant zu Q117 aus der Flo5-Subdomäne, welches ebenfalls mit der glycosidischen Bin-

dung interagiert. Die Situation in Position IV ist *vice versa*. In Flo5A findet man an dieser Stelle ein voluminöses Tryptophan, welches in Kombination mit Q98 aus seiner Subdomäne diesen Bereich der Tasche stark abschirmt und damit die Bindung an GlcNAc oder ManNAc sterisch hindern würde. Cea1A hingegen besitzt an dieser Stelle einen flexiblen Arginin-Rest in einer Konformation welche die spezifische Interaktion mit der Acetylgruppe des Liganden erlaubt. Das bereits zuvor angesprochene, konservierte Lysin im flexiblen Loop L3 nimmt im Gegensatz zu Flo5A in Cea1A direkt an der Bindung teil. Die Konformation des Loops in Flo5A verhindert eine direkte Interaktion mit gebundener Mannose und exponiert die Bindungstasche viel stärker gegenüber dem Lösungsmittel.

Im Vergleich mit der zweiten strukturell literaturbekannten Familie der epithelialen Adhäsine¹⁰⁸ zeigt sich eine um 180° entlang der glycosidischen Bindung rotierte Position der Hexose (Abbildung 4.16 B). In diese Orientierung wird die Galactose durch die Interaktion mit W198 aus L3 gezwungen. Weiterhin konnte gezeigt werden, dass dieser Rest essentiell für die korrekte Ligandenerkennung und Bindung ist. Interessanterweise nimmt dieses Tryptophan die gleiche räumliche Position ein wie K172 in Cea1A. Umgekehrt unterscheidet sich die Zusammensetzung von CBL2 sehr stark zwischen Epa1A und Cea1A. So trägt Epa1A an Position I anstatt eines hydrophoben Restes ein Arginin, welches damit die Bindung von GlcNAc in der gegebenen Orientierung ausschließen würde. Kein Pendant in Flo5A und Epa1A hat hingegen W92 aus L2, welches zur Bindung der Acetylgruppe beiträgt.

Es bleibt festzuhalten, das Cea1A Charakteristika beider Adhäsine-Subtypen zeigt, allerdings einen komplett eigenständigen Bindungsmodus entwickelt hat. Die phylogenetische Einordnung in einen neuen Subtyp wird hierdurch funktionell sowie strukturell untermauert.

4.2.5 Die Cea1A-vermittelte Chitin-Bindung von *S. cerevisiae*

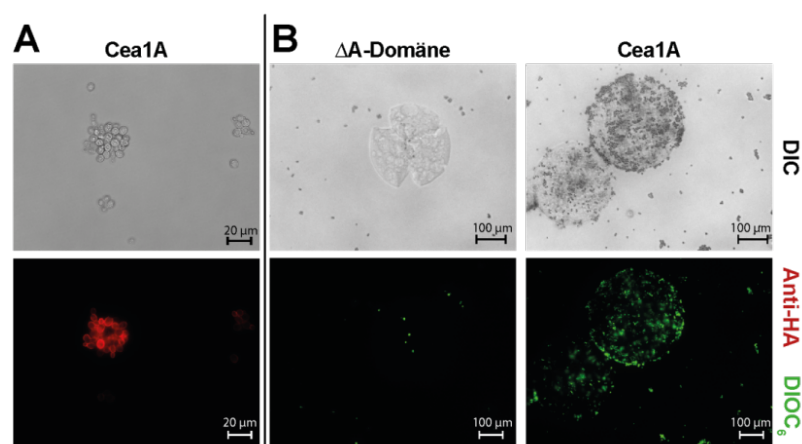


Abbildung 4.17 – Fluoreszenzmikroskopische Untersuchungen zu Cea1A. A: Immunfluoreszenz-Anfärbung zur Lokalisation des chimären Adhäsins. B: Cea1A vermittelte Adhäsion an Chitin-Beads.

Zur Untersuchung der Effektivität der Erkennung von *N*-Acetylglucosamin und damit chitinösen Polymeren durch Cea1 *in vivo* wurde dessen A-Domäne als Fusionsprotein mit der B- und C-Domäne von Flo11 auf *S. cerevisiae* Zellen des Stamms RH2520 präsentiert (BHUM2297; siehe Abschnitt 3.6.4). Die korrekte Lokalisation dieses chimären GPI-CWP in der Zellwand wurde mit Hilfe von immunfluoreszenter Anfärbung eines amino-terminal codierten 3xHA-Tags und konsekutiver Mikroskopie bestätigt (Abbildung 4.17 A). Ein Expressionsplasmid, welches lediglich zur Präsentation der B- und C-Domäne führte (ΔA) diente in den nachfolgenden Experimenten als Negativkontrolle. Die mit der Cea1A/Flo11BC-Chimäre modifizierten Zellen adhärten nach Inkubation und multiplen Waschschritten stark an mit Chitinase behandelten *Chitin Beads* (Abbildung 4.17 B). Die Variante mit Deletion der A-Domänen hingegen zeigte weder eine initiale Adhäsion an die *Chitin Beads* noch eine größere Anzahl an Zellen nach dem Waschen der *Beads*.

Die Cea1A-präsentierenden Zellen waren nicht in der Lage an unbehandelte *Beads* zu binden (Abbildung 4.18 A). Dies untermauert zum Einen, dass die beobachtete Adhäsion durch die C-Typ-Lektinartige Bindung von Cea1A vermittelt wird und keine unspezifische Erkennung der Struktur des polymeren, nicht-kristallinen Chitins auftritt, zum Anderen allerdings auch, dass nicht-reduzierende GlcNAc-Termini für die Bindung benötigt werden. Die Beobachtung das die Zellen nach Inkubation, dreimaligem pelletieren und waschen noch in der Lage sind an den *Beads* gebunden zu bleiben, deckt sich mit den Ergebnissen der ITC-Messungen, welche eine Affinität im mikromolaren Bereich anzeigte. Die Auszählung der adhärerenden und freien Zellen quantifiziert den Einfluss der Cea1A Domäne zudem. Es zeigte sich, dass pro 0.54 mm² großem Bildausschnitt etwa durchschnittlich 1100 statt 100 Zellen an *Chitin Beads* binden, wenn sie die Cea1A-Domäne auf der Oberfläche tragen (Abbildung 4.18 B). Darüber hinaus ist bei diesen Zellen die Bildung von mikroskopisch beobachtbaren Zellaggregaten zu beobachten; im Gegensatz zu Zellen mit fehlender A-Domäne (Abbildung 4.18 C). Die Zellwand von Hefen besteht aus einer basalen Schicht Chitin welche auf dem Großteil der Zelloberfläche durch β 1-3-verknüpfte Glucane und stark mannosylierte Proteine verdeckt ist^{312,313}. Da diese unterste Schicht an begrenzten Flächen

Cea1A & unverdaute Beads

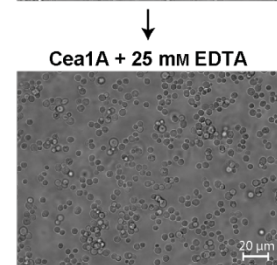
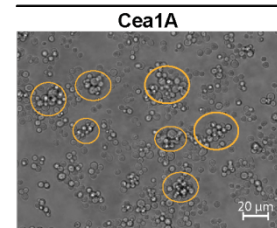
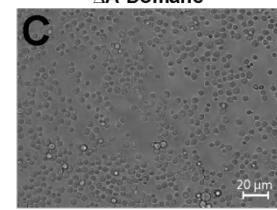
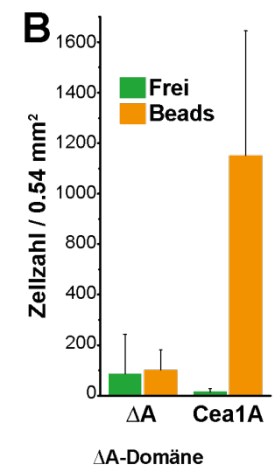
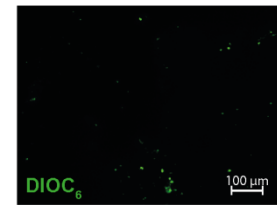
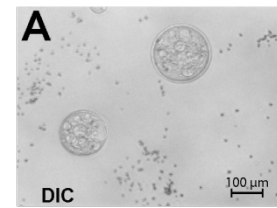


Abbildung 4.18 – Weitere mikroskopische Untersuchungen. A: Unbehandelte *Beads* bieten keine Bindungsstellen für Cea1A. B: Quantifizierung adhärerender Zellen. C: Bildung mikroskopischer Aggregate.

wie Sproßnarben und der Geburtsnarbe von Zellen zu Tage tritt, kann durch Cea1A-vermittelte Adhäsion die Bildung mikroskopisch kleiner Zellaggregate möglicherweise ausgelöst werden³¹⁴. Die Aggregate könnten ebenfalls das Resultat von Cea1A-induzierter unvollständiger Ablösung von Tochterzellen sein.

4.2.6 Produktion, Isolation und Aufreinigung von PpFlo1A

PpFlo1A, ein zweiter Vertreter des *Pichia*-Subtyps wurde zur weiteren Charakterisierung dieses phylogenetisch eigenständigen Subtyps untersucht.

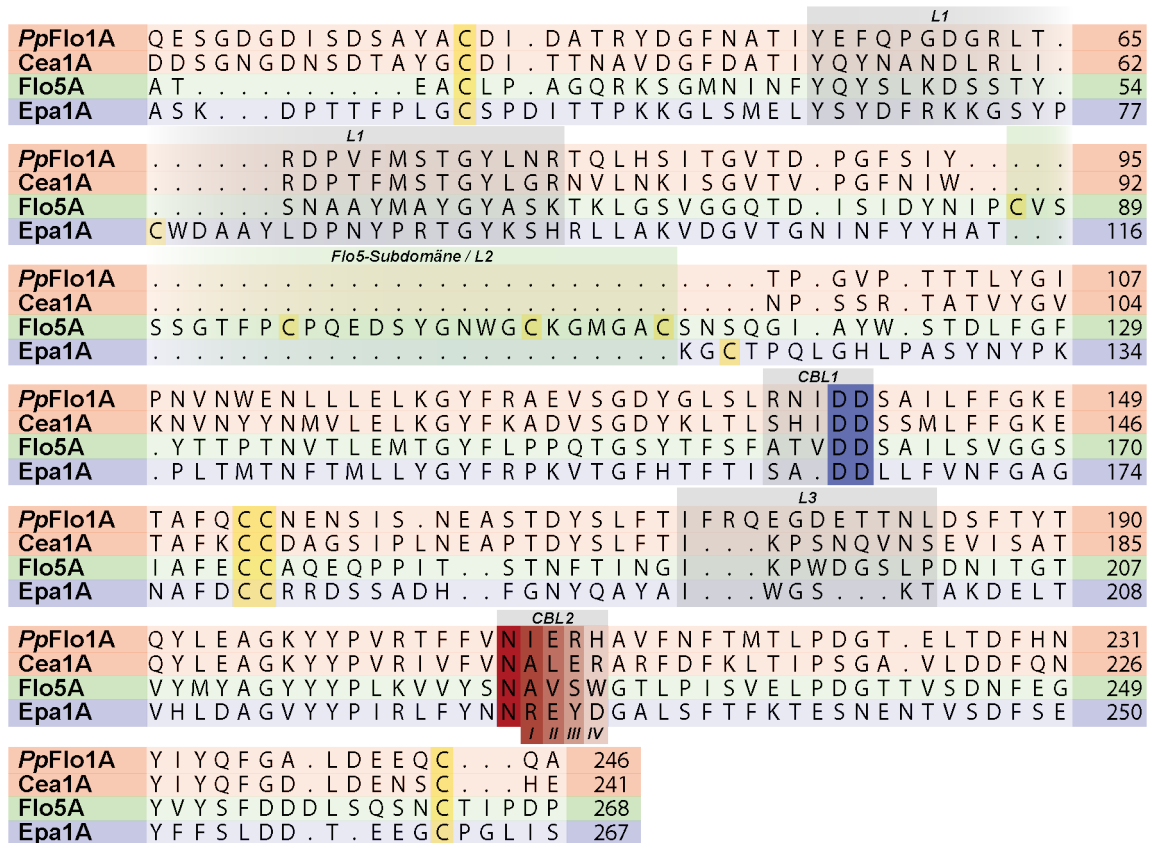
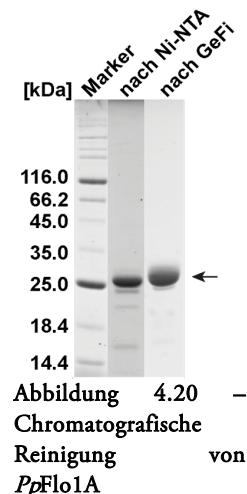


Abbildung 4.19 – Multiples Sequenzalignment zwischen PpFlo1A, Cea1A, Flo5A und Epa1A

Trotz der erhöhten Sequenzidentität von 59% zwischen PpFlo1A und Cea1A zeigt ein multiples Sequenzalignment von PpFlo1A mit weiteren Adhäsinen Differenzen innerhalb des *Pichia*-Subtyps sowie zwischen den Subtypen auf (Abbildung 4.19). Neben der unterschiedlichen Zusammensetzung von CBL1 und CBL2 sind auch die Loops, welche die Bindungstasche formend, variabel in Sequenz und Länge. Dies wird vor allem an Loop L3, welcher über der Bindungstasche liegt, deutlich. Dieser erscheint in PpFlo1A am längsten. Darüber hinaus scheint sich PpFlo1A durch dasselbe Disulfidbrückenmuster wie Cea1A auszuzeichnen und es trägt keine Flokkulin-artige Subdomäne. Die hohe Sequenzidentität spiegelt sich daher eher in Bereichen wieder, welche nicht direkt in an der Ligandenerkennung beteiligt sind.

Die heterologe Produktion von *PpFlo1A* erfolgte nach dem für *Cea1A* etablierten Protokoll. Das putative Adhäsins konnte hierbei ebenfalls über Ni-NTA-Affinitätschromatografie und anschließende Größenausschlusschromatografie gereinigt werden (Abbildung 4.20). Wie zuvor beschrieben, konnte bei *PpFlo1A*, im Gegensatz zu *Cea1A*, keine Bindung an die Säulenmatrix während beiden chromatografischen Schritten beobachtet werden¹⁵⁵. Die Produktion, Isolation und Reinigung des Adhäsins konnte erfolgreich optimiert werden, sodass nach Größenausschlusschromatografie bis zu 100 mg monomeres, homogenes *PpFlo1A* pro Liter Expressionskultur erhalten werden konnten.



4.2.7 Studien zur Funktion von *PpFlo1A*

Da zu Beginn dieser Arbeit dieses Adhäsins, ebenso wie *Cea1A*, weitgehend unbekannt war, konnte auch kein Rückschluss auf die vermutliche Funktion gezogen werden. Da das multiple Sequenzalignment auf eine Verwandtschaft mit den zuvor untersuchten PA14-Domänen GPI-CWP's hinwies, wurde mittels Glycan-Array Analyse versucht einen kompatiblen Liganden für *PpFlo1A* zu finden. Trotz der zweimaligen unabhängigen Durchführung von Sextett-Messungen für jedes Glycan konnte keinerlei messbare, nicht fehlerbehaftete Fluoreszenz gemessen werden (Abbildung 4.21). Dies wurde beispielsweise an Glycan 401 deutlich; hier lag die gemessene Fluoreszenz bei 273 RFU (*Relative Fluorescence Units*), der Fehler allerdings bei 492 RFU.

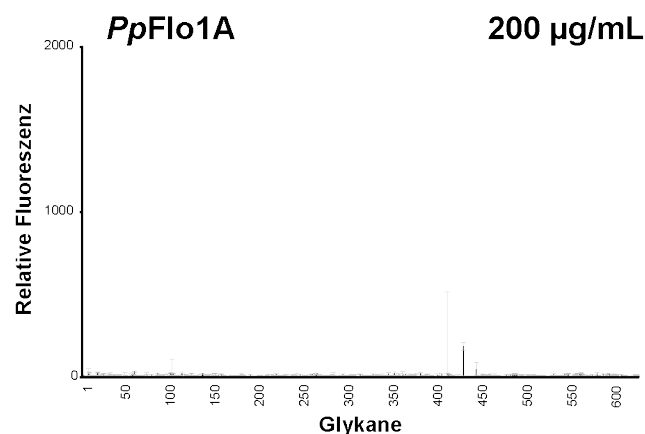


Abbildung 4.21 – *Glycan Array* Analyse von *PpFlo1A*. Es konnte hierbei keine Bindung an einen Zucker des Chips gemessen werden.

Hierbei kommen mehrere Dinge in Frage, welche für dieses Ergebnis verantwortlich sein könnten. Neben dem Fall, dass es sich bei *PpFlo1A* (*i*) nicht um ein lektinartiges Adhäsins handeln könnte, kam in Frage, dass (*ii*) die Konzentration des Liganden eine ausschlaggebende Rolle spielt und in diesem Fall die Bindung nicht beobachtet wurde. Weiterhin bestand die Möglichkeit das *PpFlo1A* (*iii*) nicht aktiv war, sei dies aufgrund fehlendem Ca^{2+}

oder einem unzureichenden pH-Bereich oder (*iv*) der passende Ligand nicht auf dem Chip vorlag.

Dieselbe Charge fluoreszenzmarkierten *PpFlo1A*s wurde in Interaktionsstudien mit *S. cerevisiae* und *P. pastoris* Zellen verwendet um eine mögliche Funktion von *PpFlo1A* weiter eingrenzen zu können. Hierbei wurde das markierte Protein auf immobilisierte Zellen beider Spezies gegeben, nach Inkubation gewaschen und die Proben fluoreszenzmikroskopisch untersucht (Abschnitt 3.6.3).



Abbildung 4.22 – Fluoreszenzmikroskopische Untersuchungen zu *PpFlo1A*. Nach der Inkubation fluoreszenzmarkierten *PpFlo1A* mit *S. cerevisiae* und *P. pastoris* Zellen zeigt sich bei etwa 5-10% der letzteren Fluoreszenz, während diese bei *S. cerevisiae* nicht beobachtet werden konnte.

Während etwa 5-10% der Zellen beider *P. pastoris* Stämme (GS115, DSMZ70382) Fluoreszenz zeigten, konnte diese bei *S. cerevisiae* Stämmen nicht beobachtet werden (Abbildung 4.22). Die Aufnahmen könnten darauf hindeuten, dass *PpFlo1A* zur Wechselwirkung zwischen Zellen einer bestimmten Subpopulation dient, welche das korrekte Substrat auf der Oberfläche tragen. So ist bekannt, dass sich die Zusammensetzung der Zellwand bei Hefen nicht nur mit der Spezies, sondern ebenfalls mit ihrem Zellstadium, ihrer Nahrung und äußeren Faktoren ändern kann^{71,315-317}. SIM *et al.*, welche die Struktur von Lg-Flo1A studierten, konnten zeigen, dass dieses Flokkulin offenbar eine Selektivität für phosphorylierte Mannose zeigt¹³⁵, welche ebenfalls in der Zellwand von Hefen beobachtet wird^{67,318,319}, aber nicht auf dem Chip der *Glycan Array* Analyse vorhanden ist.

4.2.8 Die Struktur von *PpFlo1A*

Um zu klären, ob es sich bei *PpFlo1A* wirklich um ein Adhäsion mit PA14-Domäne handelt und wenn ja, welche strukturellen Eigenschaften es hat, wurden Versuche zur Kristallisation und Röntgenstrukturanalyse durchgeführt. Wie für Cea1A konnten bereits in früheren Arbeiten erste Kristalle erhalten werden¹⁵⁵. Da diese nur untergeordnete Streuqualität hauptsächlich aufgrund sehr starker anisotroper Streuung zeigten, wurde ein erneutes Screening mit anschließender Optimierung durchgeführt. Da kein zugehöriger Ligand für *PpFlo1A* bekannt war, wurde die Kristallisation des Apoproteins versucht. Dies stellte sich als recht aufwendig heraus, da eine große Zahl Kristalle unterschiedlicher Kristallformen getestet und optimiert werden musste um einen ausreichend streufähigen Kristall zu erhalten. Dies mag darauf zurück zu führen sein, das durch das Fehlen eines Liganden *PpFlo1A* nicht stark genug

in seiner Flexibilität eingeschränkt war um zur Ausbildung von gut streuenden Kristallen zu neigen.

Ein monokliner Kristall mit ausreichender Qualität konnte durch Reproduktion von Bedingung 42 des Screens JCSG Core I bei 18 °C und einer Konzentration von 30 mg/mL in SEC-Puffer erhalten werden (Tabelle 4.6).

Tabelle 4.6 – Kristallisationsbedingung von *PpFlo1A*.

JCSG Core I #42	
MgCl ₂	200 mM
NaCaco	200 mM
pH	6.5
PEG 200	50%

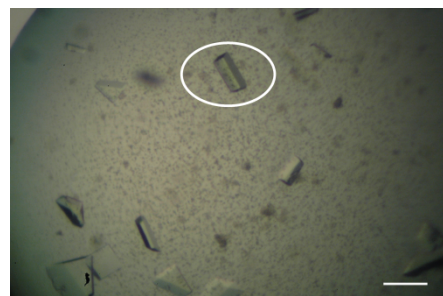


Abbildung 4.23 – Kristallisation von *PpFlo1A*. Der markierte Kristall MK102 wurde zur Strukturbestimmung verwandt. Größenmaßstab: 100 µm.

Der etwa 100 µm große, dreidimensionale Einkristall (Markierung in Abbildung 4.23) wurde mit 20% Glycerin als Kryoschutz in flüssigem Stickstoff zum BESSY-II Synchrotron gebracht und einem Röntgenbeugungsexperiment unterzogen.

Hier konnte ein vollständiger Datensatz des Kristalls an Beamline 14.1 gesammelt werden (siehe Tabelle 4.7) und die Daten mit Hilfe von *imosflm* integriert und mit *SCALA* reduziert werden^{274,278}. Aufgrund der vorhandenen nichtkristallographischen Symmetrie wurden die Testreflexe auch hier in dünnen Auflösungsschalen mit *DATAMAN* gewählt²⁹⁸.

Da ebenfalls wie im Fall von *Cea1A* eine enge strukturelle Verwandtschaft mit den phylogenetischen Subtypen der Flokkuline und epithelialen Adhäsine vorausgesetzt wurde, wurde eine Lösung des Phasenproblems mittels molekularen Ersatzes versucht. Hierzu wurde ein Model von *PpFlo1A* auf Basis der Struktur von *Epa1A* (PDB-ID: 4ASL) mit der Software *chainsaw* erstellt³²⁰. Die flexiblen Bereiche des Modells, wie L1-L3, wurden in angemessener Weise gekürzt und dieses in *phaser* bei einer maximalen Auflösung von 2.5 Å verwandt²⁹⁰. Hierbei konnte eine einzige Lösung in Raumgruppe C121 mit den TFZ-Werten 4.0 und 8.6 erhalten werden, welche in *autoBuild* zur Generierung eines ersten Modells genutzt wurde³²¹. Die Verfeinerung mit 5 alternierenden Runden *Coot* und *REFMAC5* lieferte bisher einen $R_{\text{work}}=21.3\%$ und einen $R_{\text{free}}=25.25\%$. Weitere vorläufige Statistiken sind in Tabelle 4.8 gelistet.

Tabelle 4.7 - Datensammlungsstatistik des *PpFlo1A* – Kristalls.

<i>PpFlo1A</i>	
Datensatz	MK102
Datum	25.05.2012
Röntgenquelle	BESSY II BL14.1
Detektor	MARmosaic 255 mm
Temperatur (K)	100 K
Wellenlänge (Å)	0.91841
Raumgruppe	C121
Zellmaße (Å)	
a	79.94
b	103.19
c	72.10
α, β, γ	90°, 113.51°, 90°
Moleküle / a.s.u.	2
Auflösung (Å)	40.67-2.15 (2.27-2.15)
Beobachtete Reflexe	107631
Multiplizität	3.7 (3.7)
Einzigartige Reflexe	29204
R_{merge} (%)	7.1 (56.1)
Vollständigkeit (%)	100.0 (100.0)
Mittleres $I/\sigma(I)$	11.7 (2.4)
Mosaizität (°)	0.19
Wilson B-Faktor (Å ²)	32.6

Tabelle 4.8 – Verfeinerungsstatistik für *PpFlo1A*.

<i>PpFlo1A</i>	
Auflösung (Å)	39.41-2.15
$R_{\text{work}}, R_{\text{free}}$ (%)	21.33, 25.25
Reflexe (Arbeits-, Test-)	28187, 1005
Vollständigkeit (%)	99.96
r.m.s.d.	
Bindungslängen(Å)	0.012
Bindungswinkel (°)	1.33
Anzahl Atome	3524
Mittlerer B-Wert (Å ²)	34.3

Die Anordnung der beiden Moleküle in der asymmetrischen Einheit entspricht einem antiparallelen Dimer (Abbildung 4.24). Beide sind sich bei einem durchschnittlichen r.m.s.d. von 0.085 Å über 177 C_{α} -Atome sehr ähnlich. Die Kontakte zwischen den Molekülen werden im Kristall durch einen zusätzlichen β -Strang erzeugt, welcher keinem der beiden Moleküle definitiv zugeordnet werden kann, da die Elektronendichte an den Enden des Stranges abbricht. *PpFlo1A* besteht ebenfalls aus einem kompakten β -Sandwich, welches

typisch für die pilzlichen Adhäsine dieser Familie ist. Darüberhinaus ist ebenfalls in *PpFlo1A* das konservierte *DcisD*-Motiv zu finden.

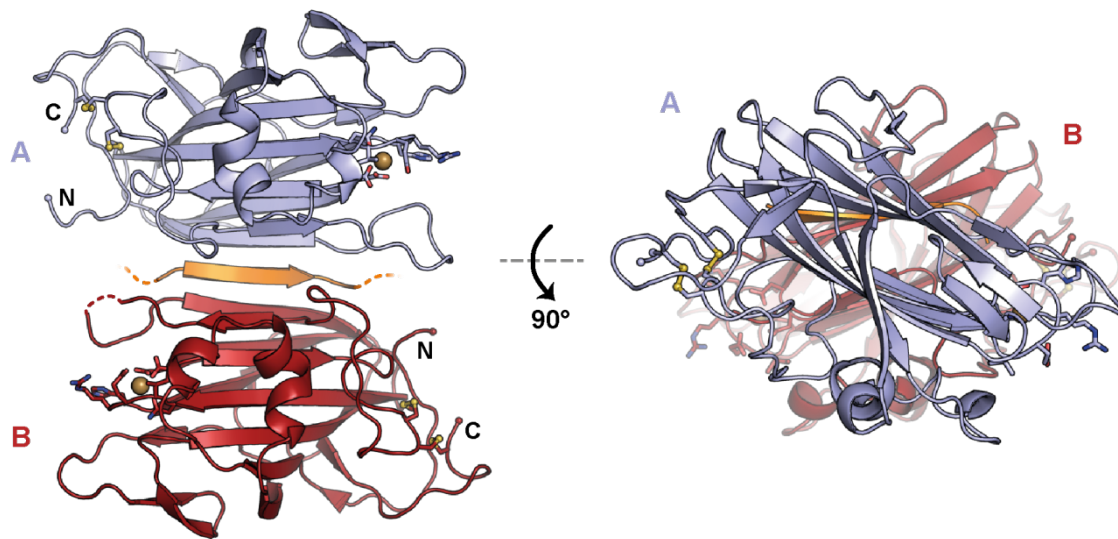


Abbildung 4.24 – Orientierung beider Moleküle der asymmetrischen Einheit. Die antiparallele Anordnung wird durch die Wechselwirkung beider *PpFlo1A*-Moleküle mit einem einzelnen, nicht zuzuordnenden β -Strang stabilisiert.

Neben dem *DcisD*-Motiv sind auch die flexiblen Loops L1-L3 zu erkennen, sowie CBL2 und das Ca^{2+} -Ion der Bindungstasche. Die Struktur einerseits beweist damit, dass *PpFlo1A* sehr wohl ein Mitglied der PA14-Domänen enthaltenden Lektin-artigen GPI-CWP's ist, während die Anwesenheit des Calcium-Ions andererseits die Aktivität des Proteins sicherstellen sollte. Darüber hinaus konnte von SIM *et al.* für Lg-Flo1A eine Dissoziationskonstante von etwa 14 nM für das Ca^{2+} -Ion gezeigt werden¹³⁵. Dies widerlegt bereits zwei Annahmen der negativen *Glycan-Array* Analyse.

Während die flexiblen Loops L1 und L2 gut durch die Elektronendichte definiert sind, ist die Position von L3 nicht so eindeutig. Dessen Rückgrat ließ sich in Molekül A noch darstellen, in Molekül B ist die Position von L3 allerdings nicht gut abgegrenzt (gestrichelte Linie Abbildung 4.24).

Interessant ist im Fall von *PpFlo1A* die Orientierung der beiden Disulfidbrücken und die Region der Termini (Abbildung 4.25). Diese entspricht hier der in Flo5A und Epa1A gefundenen Helix-/Loop-Region. Dies bestätigt die Komplikationen der Lösung des Phasenproblems für Cea1A, da bei *PpFlo1A* die Verwendung des Epa1A-Modells mit vorhandenem CC-Motiv zum Erfolg des molekularen Ersatzes führte.

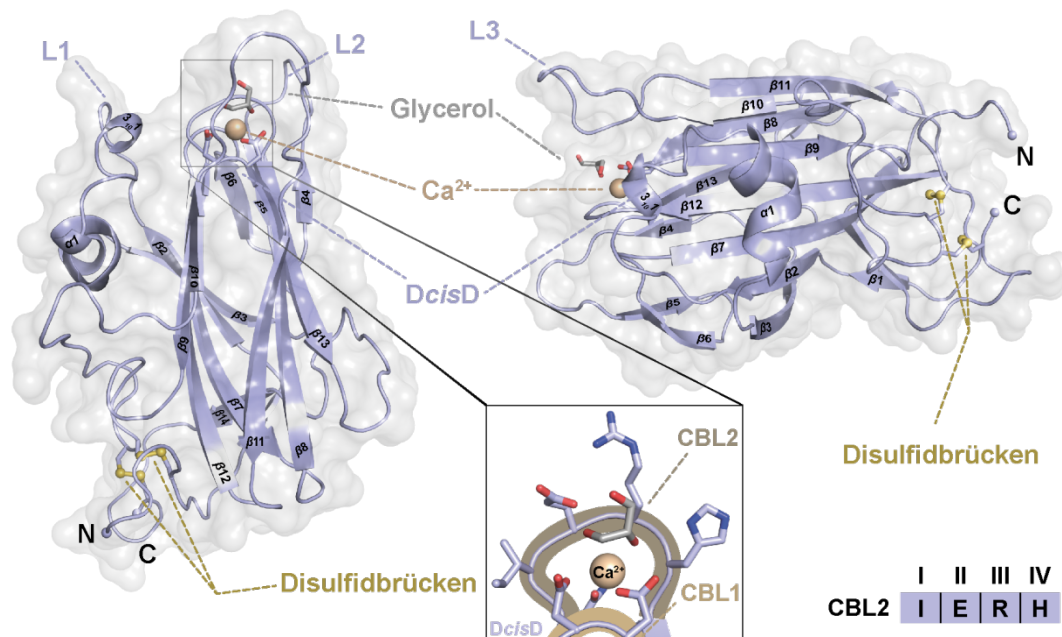


Abbildung 4.25 - Die Struktur von *PpFlo1A*. Neben dem konservierten β -Sandwich und den Disulfidbrücken trägt *PpFlo1A* einen stark verlängerten Loop L3. Die Region beider Termini bildet hier den aus den Flokkulinen und den Epas gewohnten Verlauf. CBL1 (hellbraun) komplexiert mit Hilfe des *DcisD*-Motivs ein Ca^{2+} -Ion (ocker) über welches hier in der apo-Form des Adhäsins ein Glycerol gebunden vorliegt.

Der β -Strang, welcher den Kontakt zwischen den Molekülen A und B sicherstellt und keinem der beiden eindeutig zuzuordnen ist, lässt sich durch Modellierung eines Histidin-Hexapeptidst gut beschreiben; offensichtlich handelt es sich hierbei um den *N*-terminalen Hexa-Histidin-*Tag*. Die Interaktion über ein solch regelmäßiges Element wird verständlicher wenn berücksichtigt wird, dass der Kontakt von Molekül A und B jeweils mit dem gleichen β -Strang der PA14-Domäne ($\beta 11$) und damit anhand der gleichen Kristallkontakte über den Hexa-His Strang erfolgt (Abbildung 4.26).

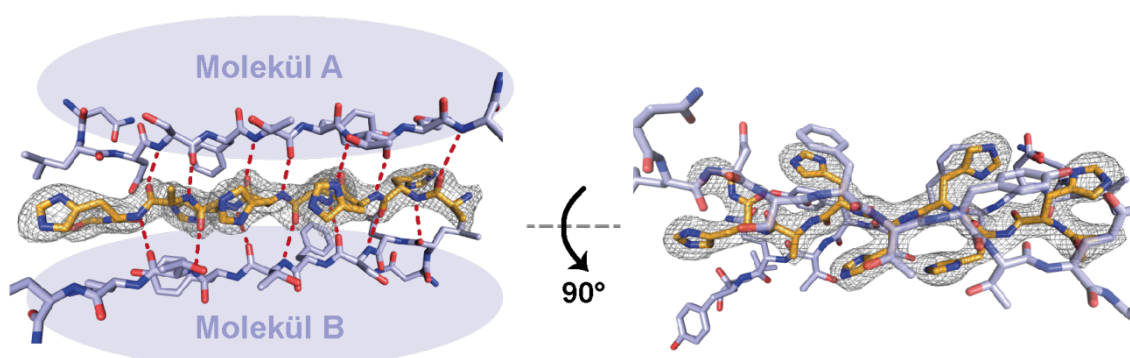


Abbildung 4.26 – Der einzelne β -Strang der asymmetrischen Einheit. Die SigmaA-gewichtete $2F_{obs}-F_{calc}$ -Elektronendichte bei einem Konturierungsgrad von $\sigma = 1.0$ zwischen den Molekülen A & B deutet auf einen Teil des Hexa-Histidin-*Tags* hin, welcher sich aufgrund der Symmetrie keinem der beiden Moleküle eindeutig zuordnen lässt.

Die Bindungstasche zeigt deutliche Unterschiede zwischen *PpFlo1A*, *Cea1A* und den weiteren Adhäsinen. Der flexible Loop L3 ist nicht nur der Längste in den bekannten Strukturen, er besteht darüber hinaus aus einem noch unbekanntem Sequenzmotiv. Tragen Epas einen

konservierten Tryptophan-Rest und die meisten Flokkuline und Cea1A ein Lysin in L3, so zeigt *PpFlo1A* in dieser Position das ungeladene F174. L1 und L2 sind, wie in Cea1A, sehr kurz im Vergleich zu beiden anderen Familien. Komplementär zu W92 in Cea1A trägt *PpFlo1A* an dieser räumlichen Position Tyrosin 95.

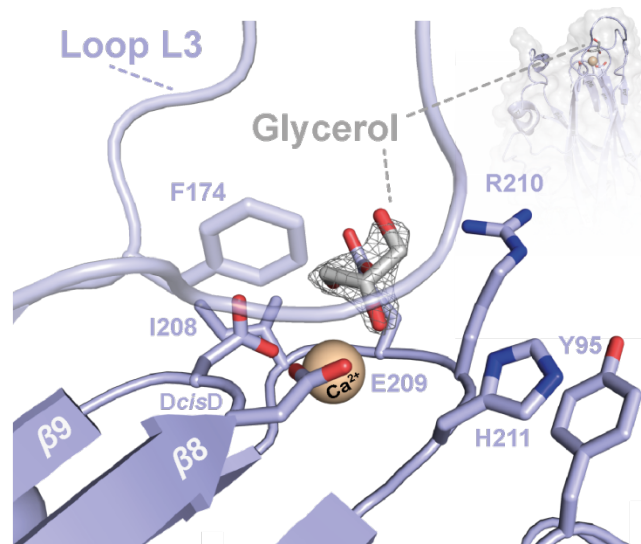


Abbildung 4.27 – Die Bindungstasche von *PpFlo1A*. Neben einem hydrophoben Bereich, welcher durch I208 und F174 gebildet wird, ist die Bindungstasche durch die Präsenz voluminöser Aminosäuren sterisch sehr anspruchsvoll. Darüber hinaus bieten sich sowohl Arginin, Histidin und Tyrosin aus Loop L2 durch ihre Ladungen und exponierte Lage als Wechselwirkungspunkte für einen möglichen Glycanliganden an. Das in der Tasche vorhandene Glycerol ist mit SigmaA-gewichteter $2F_{obs}-F_{calc}$ -Elektronendichte bei einem Konturierungsgrad von $\sigma = 1.0$ gezeigt.

CBL2 ist durch große Reste sehr raumfordernd. So besteht er in Position I aus I208, welches ähnlich wie in den Flokkulinen eine hydrophobe Fläche in der Tasche bildet und darauf folgend aus Q209, R210 sowie H211. Letztere formen einen positiv geladenen Bereich. In der Summe scheint die Bindungstasche sterisch am anspruchsvollsten zu sein. Da *PpFlo1A* ohne Glycan kristallisiert wurde, findet sich in der Bindungstasche ein Glycerin, dessen vicinale Diole an das Ca^{2+} -Ion analog zu einem Zucker koordiniert sind (Abbildung 4.27). Dessen Anwesenheit spricht für die vorhandene Aktivität des Adhäsins da bspw. Flo5A im apo-Zustand ebenfalls ein Glycerin bindet.

4.2.9 Untersuchungen zur Vergesellschaftung von Hefen und qPCR

Um weitere Informationen zum biologischen Kontext des phylogenetischen *Pichia*-Subtyps der Adhäsine zu erlangen wurden in Kooperation mit DOROTHEE TEGTMEIER erste Experimente mit Schaben durchgeführt um den Einfluss der Erkennung von terminalen GlcNAc-Resten durch Cea1A auf die Interaktion zwischen Hefen und Schaben, welche Chitin bzw. chitinöse Polymere in einem Großteil ihres Körperaufbaus enthalten, zu untersuchen. Da Wechselwirkungen zwischen Hefen und Insekten literaturbekannt sind (siehe Abschnitt 1.1) und die Bindungseigenschaften von Cea1A auf dessen Verwicklung in einen solchen Prozess hinweisen könnten, wurde untersucht, ob sich Adhäsins-präsentierende Hefen selektiv in Insekten etablieren können. Da darüber hinaus die Funktion von *PpFlo1A* unbekannt war,

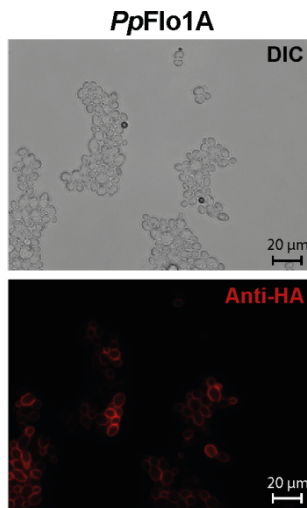


Abbildung 4.28 – Präsentation von *PpFlo1A* auf der Zelloberfläche von *S. cerevisiae*.

wurde ein Plasmid auf Basis von BHUM2297 erzeugt, welches die Präsentation von *PpFlo1A* auf der Zelloberfläche von *S. cerevisiae* ermöglicht. Für dieses Fusionsprotein konnte ebenfalls die Lokalisation auf der Zelloberfläche nachgewiesen werden (Abbildung 4.28).

Wie in Abschnitt 3.6.4 beschrieben, wurden Schaben der Gattung *Shelfordella lateralis* mit *S. cerevisiae* Zellen, welche die Cea1A-, die *PpFlo1A*-, die Flo11A- sowie die Flo11BC-Domäne trugen, gefüttert. Während letztere Probe als Negativkontrolle diente, wurden *P. pastoris* DSMZ70382 und *S. cerevisiae* S288c als Wildtypstämme eingesetzt. Flo11A wurde als Adhäsion ohne PA14-Domäne zur Kontrolle verwendet. Nach 4 Tagen sowie einer und zwei Wochen wurden Schaben seziiert um auf die Anwesenheit

und Verteilung von Hefen in den unterschiedlichen Darmbereichen der Schaben zu überprüfen. Bereits nach 4 Tagen konnten bei allen verwendeten Hefestämmen, so auch bei der Negativkontrolle, Zellen in jedem Bereich des Darms nachgewiesen werden (Abbildung 4.29). Dieser Trend setzte sich für die folgenden Sektionen *post inoculation* fort; in jedem Teil des Darms konnte Besiedlung nachgewiesen werden.

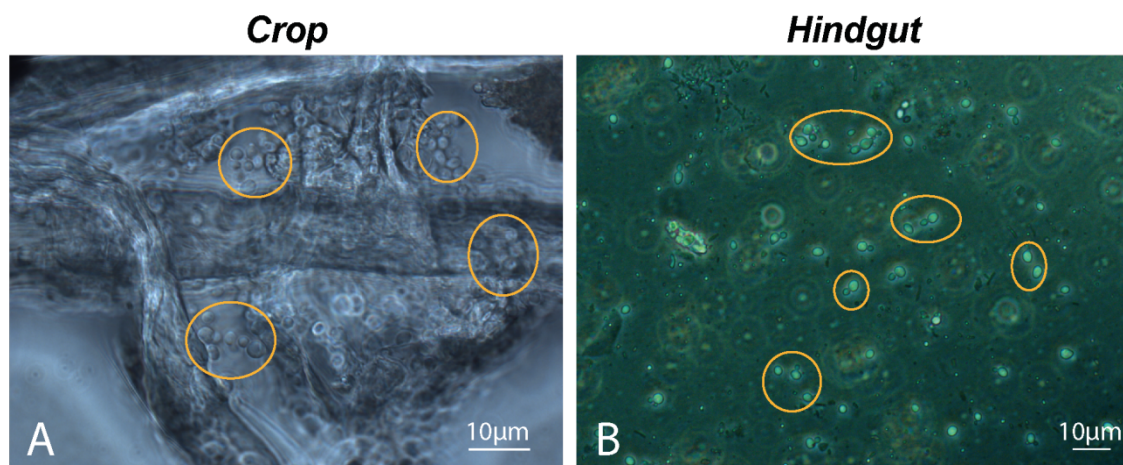


Abbildung 4.29 – Mikroskopische Aufnahmen von *S. cerevisiae* Zellen in zwei unterschiedlichen Darmbereichen von *Shelfordella lateralis* 4 Tage *post inoculation*. A: Flo11-präsentierende Zellen im Crop-Bereich; B: *P. pastoris* DSMZ70382 Zellen im Hindgut-Bereich. Hefezellen sind durch orange Umrandungen hervorgehoben.

Um eine halbquantitative Aussage über den Gehalt an Hefen zu erhalten, wurden Schabendärme nach 1.5 Wochen homogenisiert und auf SC4-Medium ausgestrichen. Dazu wurde ebenfalls eine Probe des Futters homogenisiert und auf Bewuchs getestet. Hierbei wurde deutlich, dass Hefen, hier sogar Laborstämme, offenbar in der Lage sind, nicht nur in *Shelfordella lateralis* zu überleben, sondern diese dauerhaft zu besiedeln. Es zeigten sich zwar größere Unterschiede zwischen dem *P. pastoris*- und dem *S. cerevisiae*-Wildtypstamm, doch war der Bewuchs offenbar unabhängig vom verwendeten Stamm, da ebenfalls Zellen ohne präsentier-

te A-Domäne Wachstum zeigten. Die Betrachtung der Futterprobe deutet allerdings darauf hin, dass eine Besiedlung des jeweils neu hinzugegebenen Futters durch auf der Schabenoberfläche vorhandene Hefen möglicherweise zu einem *perpetuum mobile* von Reinfektionen führte. Obwohl es sich hierbei um vorläufige Ergebnisse erster Experimente handelt zeigen diese dennoch eine erfolgreiche Besiedlung von Schaben durch Hefen unter Laborbedingungen und bilden die Grundlage für weitergehende Untersuchungen.

Weiterhin wurden erste Analysen zur Expressionsstärke von Adhäsion-Genen in *Pichia pastoris* mittels qPCR durchgeführt. Diese dienten im Rahmen dieser Arbeit primär der Etablierung dieser Methode als Basis für die strukturierte Untersuchung des Expressionsverhaltens der Adhäsion-Gene in Hefen. Nachdem Gesamt-RNA aus *P. pastoris* DSMZ70382 präpariert wurde, konnte mit entsprechenden Primern die relative Menge an Transkripten durch Bestimmung des C_T -Werts ermittelt werden.

Um einen relativen Vergleich durchführen zu können, wurde Actin als Haushaltsgen verwendet, da dieses einen stabilen Expressionsverlauf zeigt²¹³. In Abbildung 4.30 ist ein qPCR-Experiment mit *PpFlo1A* (blau) und *Cea1A* (orange) als Zielgene und Actin (schwarz) als Haushaltsgen gezeigt. Vergleicht man die C_T -Werte der Reaktionen wird deutlich, dass diese sich in einem ähnlichen Bereich bewegen. Um falsch-positive Ergebnisse zu identifizieren, wurde eine Probe mit Actin-Primern mit Wasser anstatt der reversen Transkriptase versetzt (rot).

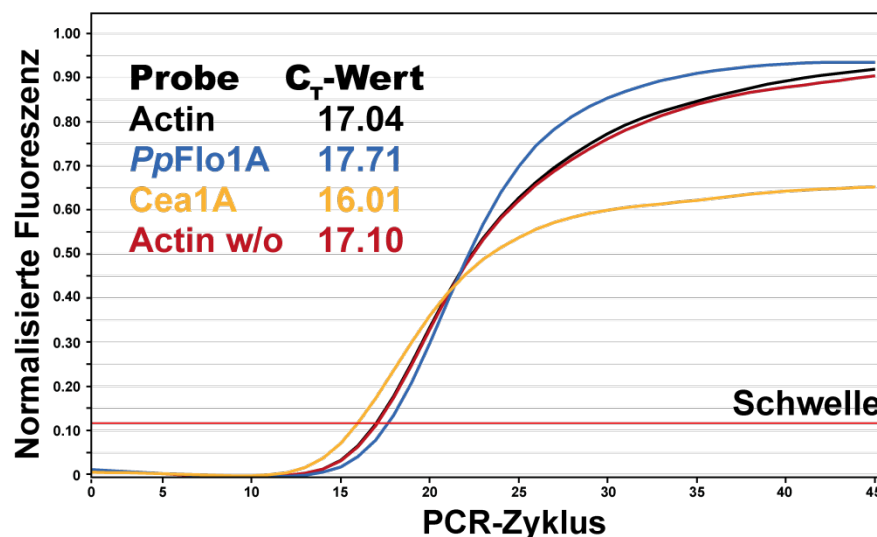


Abbildung 4.30 – Etablierung der qPCR zur Analyse des Expressionsverhaltens von Adhäsion-Genen in Hefen. Der C_T -Wert der Kontrolle ohne reverse Transkriptase (rot) liegt im gleichen Bereich wie die Werte der Proben. Dies deutet auf eine Kontamination mit genomischer DNA hin. Dennoch konnte die prinzipielle Funktionalität der Primer und der PCR-Protokolls bewiesen werden.

Da der C_T -Wert dieser Probe 17.10 beträgt und dieser damit im Bereich der anderen Proben liegt, kann davon ausgegangen werden, dass bei der RNA-Isolation eine Kontamination mit genomischer DNA erfolgt ist. Diese führt zu einer Verfälschung der C_T -Werte und damit der Quantifizierung. Die Verwendung einer DNA-spezifischen Nuklease im Prozess der RNA-

Extraktion sollte hier zur Vermeidung dieser Kreuzkontamination führen. Trotz des Optimierungsbedarfs können diese Experimente *ex nunc* als Grundlage für Genexpressionsanalysen der Adhäsion-Gene in Hefen dienen.

4.3 Epitheliale Adhäsine aus *C. glabrata* – Struktur und Funktion

Der phylogenetisch abgegrenzte, aber dennoch sehr breite Subtyp der epithelialen Adhäsine (Epas) ist bekannt für ihre Beteiligung an der Pathogenität von *Candida glabrata*. Die Epas stellen hier wichtige Pathogenitätsfaktoren dieser Hefe dar, da sie zur Adhäsion an Epithelzellen und damit zur Etablierung einer Infektion benötigt werden. Im Rahmen dieser Arbeit und in Zusammenarbeit mit RIKE DIDERRICH wurden strukturelle Untersuchungen in Kombination mit funktionellen Betrachtungen durchgeführt, um einen vertieften Einblick in die Bindungseigenschaften dieser physiologisch relevanten Familie der PA14-Domänen GPI-CWP's zu erhalten. Hierbei wurde besonderes Augenmerk auf die Korrelation zwischen Struktur und Funktion einerseits und der phylogenetischen Beziehung der einzelnen Epas andererseits gelegt.

4.3.1 Die Kokristallstruktur von Epa1A

Das beststudierte epitheliale Adhäsion stellt Epa1A dar (siehe Einleitung). Hier ist neben der physiologischen Relevanz auch die Struktur und der Mechanismus der molekularen Erkennung von Galactosiden bekannt. Die Komplexstrukturen von Epa1A konnten zuvor hauptsächlich nur durch *replacement-soaking* von Lactose-Kokristallen erhalten werden. Hierzu wurde durch Inkubation des Kristalls in EDTA-haltige Kristallisationspuffer der Ligand mit dem Ca^{2+} -Ion aus der Bindungstasche gewaschen und ein neuer Ligand in den Kristall *gesoakt*^{108,158}. Die Loops L1-L3 weisen eine gewisse Flexibilität auf, was strukturell bspw. bei *PpFlo1A*, *Lg-Flo1A* (PDB-ID: 4GQ7) sowie *Flo1A* (PDB-ID: 4LHL), welche ohne Ligand kristallisiert werden konnten, durch die nur schwach definierte Elektronendichten und hohe Temperatur-Faktoren in den Loop-Bereichen deutlich wird (Abbildung 4.31).

Da hierdurch nicht auszuschließen ist, dass die Fixierung dieser Loop-Regionen durch alternative Wechselwirkungen bei Bindung verschiedener Liganden voneinander abweicht, wurde die Kokristallisation von Epa1A mit einem zweiten Liganden neben Lactose durchgeführt. Neben der Möglichkeit, die Flexibilität der Loop-Bereiche studieren zu können, würde solch eine Struktur bei unveränderten Interaktionen mit dem Liganden das etablierte *soaking*-Protokoll bestätigen.

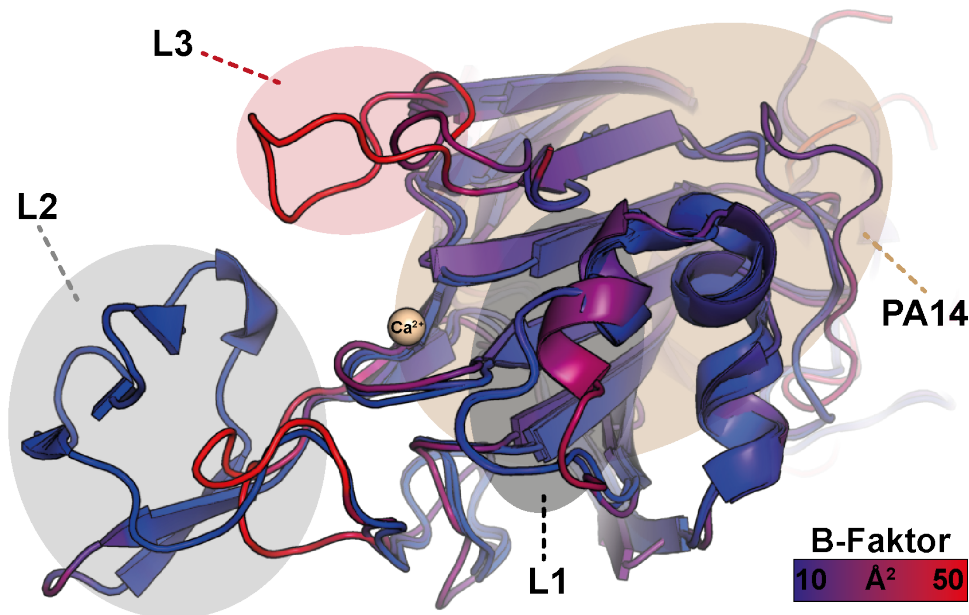


Abbildung 4.31 – Überlagerung der apo-Strukturen von *PpFlo1A*, *Flo1A* (PDB-ID: 4LHL) und *Lg-Flo1A* (PDB-ID: 4GQ7). Die Einfärbung nach Wert des B-Faktors von 10 \AA^2 - 50 \AA^2 zeigt das die Loops L1-L3 eine gewisse Flexibilität aufweisen, wenn kein Ligand gebunden ist.

Da *Epa1A* eine hohe Affinität zu $\text{Gal}\beta 1\text{-3GalNAc}$ (T-Antigen) aufweist, wurde dieser Ligand für die Kokristallisation ausgewählt. Durch ein Screening mit *Epa1A* (15 mg/mL) und 5 mM T-Antigen sowie 5 mM CaCl_2 konnten große dreidimensionale Kristalle in Bedingung #34 des Screens JCSG Core I bei $18 \text{ }^\circ\text{C}$ erhalten werden (Tabelle 4.9, Abbildung 4.32).

Tabelle 4.9 – Kokristallisationsbedingung für *Epa1A* mit 5 mM T-Antigen

JCSG Core I #34	
HEPES pH 6.5	100 mM
PEG 6000	20%
Finaler pH	7.0

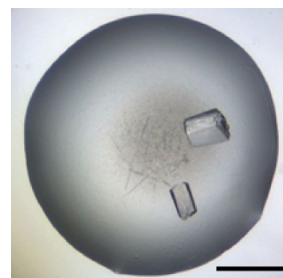


Abbildung 4.32 – Fotografie der Kokristalle von *Epa1A* mit T-Antigen. Maßstab: $100 \text{ }\mu\text{m}$

Bereits bei der Erstcharakterisierung des Kristalls deutete sich an, dass der erhaltene Kristall eine zweite Kristallform für *Epa1A* darstellt. Statt in der Raumgruppe $C222_1$ kristallisierte der *Epa1A*•T-Antigen Komplex in der Raumgruppe $P4_12_12$. Erhält man verschiedene Kristallformen eines Proteins, kann durch Studium neuartiger Kristallkontakte und damit einhergehender Änderungen der Proteinkonformation die Zuverlässigkeit der molekularen Struktur erhöht werden.

Der molekulare Ersatz mittels *Phaser*²⁸⁹ und der *Epa1A*-Struktur (PDB-ID: 4ASL) in Raumgruppe $C222_1$ lieferte bei einer Auflösung von 2.5 \AA eine einzige Lösung mit einem Translation-Z-Wert von 56. Die Verfeinerung erfolgte alternierend mit *REFMAC5*³⁰⁰ und *Coot*²⁹⁷.

Tabelle 4.10 - Datensammlungsstatistik des Epa1A•T-Antigen-Komplexes.

Epa1A•T-Antigen	
Datensatz	MK119
Datum	25.05.2012
Röntgenquelle	BESSY II BL14.1
Detektor	MARmosaic 255 mm
Temperatur (K)	100 K
Wellenlänge (Å)	0.91841
Raumgruppe	$P4_12_12$
Zellmaße (Å)	
a	55.01
b	55.01
c	139.75
α, β, γ	90°, 90°, 90°
Moleküle / a.s.u.	1
Auflösung (Å)	24.60-1.50 (1.58-1.50)
Beobachtete Reflexe	384090
Multiplizität	10.9 (10.6)
Einzigartige Reflexe	35383
R_{merge} (%)	10.1 (50.5)
Vollständigkeit (%)	100.0 (100.0)
Mittleres $I/\sigma(I)$	13.4 (4.2)
Mosaizität (°)	0.48
Wilson B-Faktor (Å ²)	9.1

Tabelle 4.11 – Verfeinerungsstatistik des Epa1A•T-Antigen-Komplexes.

Epa1A•T-Antigen	
Auflösung (Å)	24.60-1.50
$R_{\text{work}}, R_{\text{free}}$ (%)	15.79, 18.30
Reflexe (Arbeits-, Test-)	31.761, 1767
Vollständigkeit (%)	99.93
r.m.s.d.	
Bindungslängen(Å)	0.012
Bindungswinkel (°)	1.61
Anzahl Atome	2183
Mittlerer B-Wert (Å ²)	15.6

Die Kokristallstruktur von Epa1A mit T-Antigen in Raumgruppe $P4_12_12$ zeigt im Vergleich mit der zuvor bekannten Struktur nur sehr marginale Abweichungen bei einem *r.m.s.d.* von 0.31 Å über 172 C_{α} -Atome. In Abbildung 4.33 ist ein Überblick über die Kokristallstruktur von Epa1A gezeigt, welches neben dem strukturell dominierenden β -Sandwich das DcisD-Motiv sowie die bereits in der Einleitung beschriebenen Charakteristika aufzeigt. Darüber hinaus ist eine Überlagerung der *soaking*-Struktur und der Kokristallstruktur zu sehen, welche anhand der auf das jeweilige C_{α} -Atom bezogenen *r.m.s.d.* eingefärbt ist.

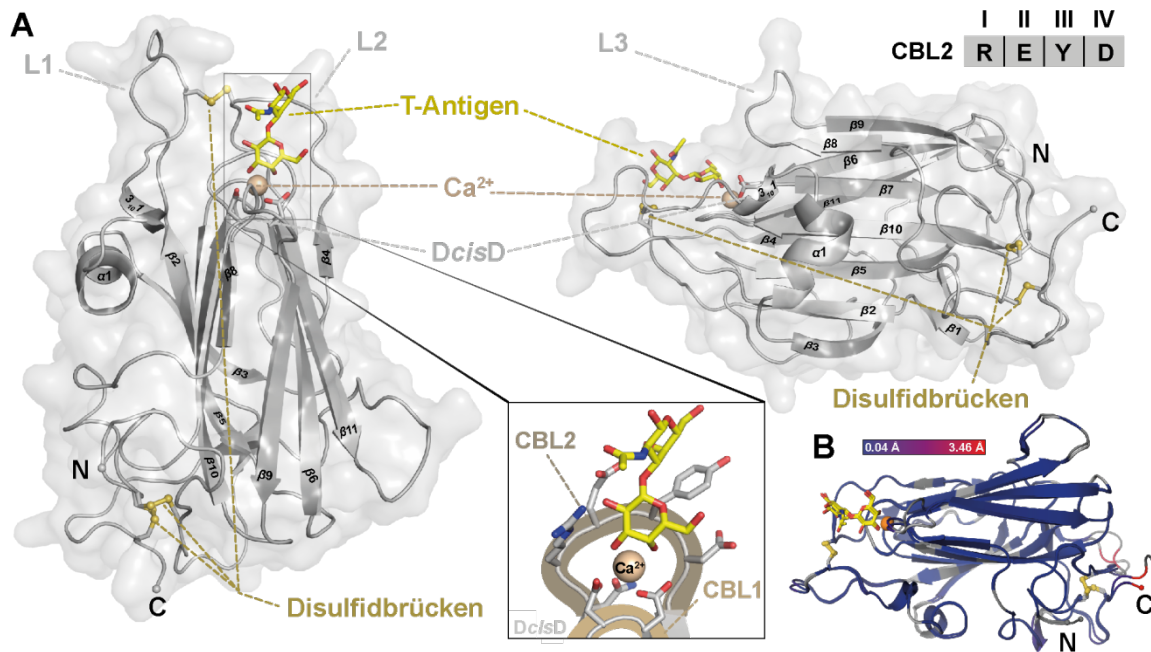


Abbildung 4.33 – Die Kokristallstruktur von Epa1A. A: Neben dem konservierten β -Sandwich zeigt Epa1A das für die epithelialen Adhäsine typische Disulfidbrückenmuster (2x im terminalen Bereich; 1x zur Fixierung von L1 & L2). B: Der Vergleich zwischen der Kokristallstruktur und der Struktur, welche mittels *soakings* erhalten wurde zeigt nur geringe Abweichungen im Bereich der Termini. Die Strukturen sind nach Größe des *r.m.s.d.* gefärbt (rot: hohe Abweichung; blau: geringe Abweichung).

Diese Ergebnisse deuten darauf hin, dass (i) das β -Sandwich von Epa1A sehr rigide ist und (ii) die flexiblen Loops L1-L3 unter Bindung eines Liganden ebenfalls eine vorgegebene Konformation, welche global unabhängig vom gebundenen Liganden ist, einnehmen. Darüber hinaus (iii) stellt die geringe Abweichung der *soaking*-Struktur von der Kokristallstruktur eine Validierung der vorgenannten *replacement-soaking*-Methodik dar.

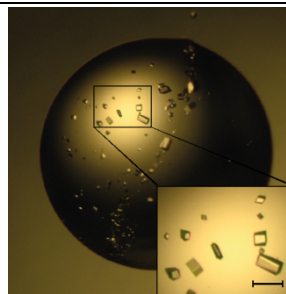
4.3.2 Lösung weiterer Komplexstrukturen von Epa1A

Nach Bestätigung der Zuverlässigkeit des Liganden-*soaking*, wurden die Bindungseigenschaften von Epa1A strukturell detaillierter untersucht. Bereits in Arbeiten von MANUEL MAESTRE-REYNA und RIKE DIDERRICH konnte gezeigt werden, dass epitheliale Adhäsine in der Lage sind sulfatierte Glycane, meist Galactoside, zu binden^{158,165}. Diese kommen unter anderem in Sulfomucinen in der humanen Mucosa vor und könnten hier als Liganden für Epas dienen³²². Um im Speziellen die Interaktion von Epa1A mit sulfatierten Sacchariden zu untersuchen, wurden zuerst Epa1A Kristalle mit Lactose^{††††} im *24well*-Format reproduziert und mittels *streak-seeding* optimiert (Tabelle 4.12 & Abbildung 4.34).

†††† Der gebundene Ligand ist hier Gal β 1-3Glc. Dieser stammt als Verunreinigung aus der kommerziell erworbenen Lactose.

Tabelle 4.12 – Kristallisationsbedingung für Epa1A

Epa1A X-Tal Bedingung	
MES pH 6	100 mM
(NH ₄) ₂ SO ₄	200 mM
PEG 5000 MME	25%

Abbildung 4.34 – Epa1A-Kristalle nach Reproduktion und *Seeding*. Maßstab 100 µm

Mehrere Kristalle wurden daraufhin mittels *replacement-soaking* nach dem Protokoll aus Abschnitt 3.5.4 behandelt und mit 5 mM 6-*O*-Sulfogalactose versetzt. Kristalle, welche nur wenige Stunden in Glycanlösung inkubierten, zeigten häufig ein verschmiertes Diffraktionsbild bei der Röntgenbeugung. Kristalle, welche hingegen über Nacht im Tropfen verblieben, zeigten häufig klar begrenzte Intensitätsspot und ein störungsfreies Diffraktionsbild. Möglicherweise sorgte die längere Inkubationszeit für eine Relaxation von Packungsstörungen, welche durch den großen Konzentrationsgradienten zu Beginn des *Soakings* induziert worden waren.

Daraus folgend wurde ein vollständiger Datensatz eines Kristalls aufgenommen, welcher über Nacht in der *soaking*-Lösung inkubierte. Die Prozessierung erfolgte mit *XDS* und *XSCALE* während die Phasen mit *MolRep*³²³ aus dem bekannten Epa1A-Modell (PDB-ID: 4ASL) erhalten werden konnten (*CONTRAST*-Wert von 15.01). Die Testreflexe wurden vom ursprünglichen Epa1A-Datensatz MV310 übernommen¹⁵⁸ und die Verfeinerung mit *phenix refine*, *REFMAC5* sowie *Coot* durchgeführt. Die zugehörigen Statistiken sind Tabelle 4.13 und Tabelle 4.14 zu entnehmen.

Tabelle 4.13 - Datensammlungsstatistik des Epa1A•Sulfogalactose-Komplexes.

Epa1A•Sulfogalactose	
Datensatz	MK140
Datum	26.05.2012
Röntgenquelle	BESSY II BL14.1
Detektor	MARmosaic 255 mm
Temperatur (K)	100 K
Wellenlänge (Å)	0.91841
Raumgruppe	C222 ₁
Zellmaße (Å)	
a	75.21
b	104.75
c	69.91
α, β, γ	90°, 90°, 90°
Moleküle / a.s.u.	1
Auflösung (Å)	46.00-1.41 (1.49-1.41)
Beobachtete Reflexe	385473
Multiplizität	7.2 (7.1)
Einzigartige Reflexe	53563
R_{merge} (%)	4.2 (57.6)
Vollständigkeit (%)	99.9 (99.6)
Mittleres $I/\sigma(I)$	21.9 (3.4)
Mosaizität (°)	0.11
Wilson B-Faktor (Å ²)	16.8

Tabelle 4.14 – Verfeinerungsstatistik des Epa1A•Sulfogalactose-Komplexes.

Epa1A•Sulfogalactose	
Auflösung (Å)	41.78-1.41
$R_{\text{work}}, R_{\text{free}}$ (%)	18.21, 20.32
Reflexe (Arbeits-, Test-)	53530, 1866
Vollständigkeit (%)	99.17
r.m.s.d.	
Bindungslängen(Å)	0.016
Bindungswinkel (°)	1.78
Anzahl Atome	1948
Mittlerer B-Wert (Å ²)	15.5

Die Komplexstruktur von Epa1A zeigt, außer der gebundenen 6-O-Sulfogalactose, keinen Unterschied zur Kokristallstruktur mit Gal β 1-3Glc (PDB-ID: 4AF9). Die flexiblen Loops zeigen die gleiche Orientierung, ebenso ist der Großteil der Kontakte zwischen den Liganden und Epa1A konserviert (Abbildung 4.35 A). Die Elektronendichte der Sulfogalactose zeigt, dass diese zwar gebunden ist, die Besetzung der Sulfatgruppe aber nur ~60% beträgt. Die restliche Galactose ist zu 100% vorhanden. Hierbei handelt es sich um ein Artefakt aus dem *replacement soaking*. Im Vergleich mit der Gal β 1-3Glc Komplexstruktur ist die Orientierung

der Hexose unverändert, ihre Position ist nur leicht verkippt (Abbildung 4.35 B). Direkte Interaktionen, welche unter Beteiligung der Sulfat-Gruppe erfolgen, gibt es aufgrund der großen Abstände keine. Hingegen werden durch Präsenz dieser Gruppe drei Wassermoleküle, welche in der Gal β 1-3Glc-Struktur gebunden vorliegen und in komplexe Wasserstoffbrückenbindungen eingebunden sind, ersetzt.

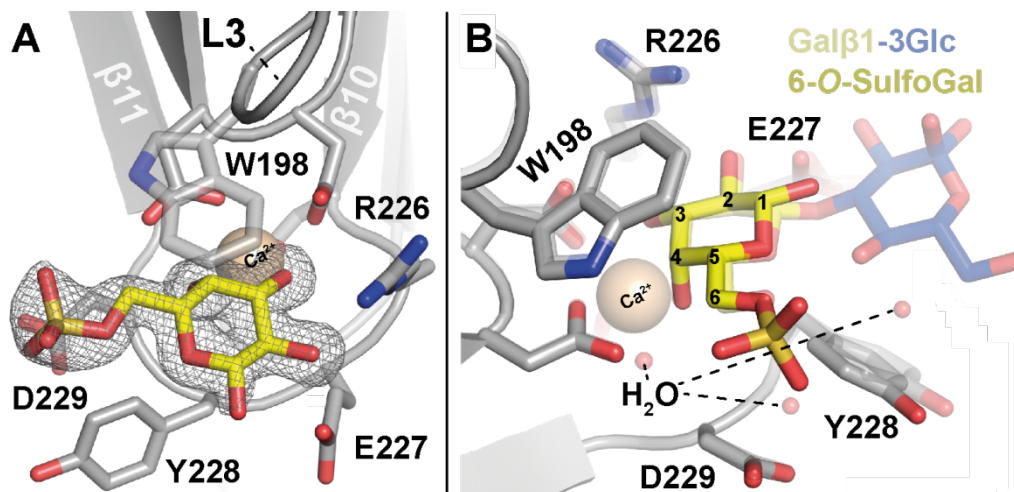


Abbildung 4.35 – Strukturelle Details der Sulfogalactose-Bindung von Epa1A. A: Die 6-O-Sulfogalactose (gelb) zeigt eine fixierte Position in der Bindungstasche, welche durch die SigmaA-gewichtete $2F_{obs}-F_{calc}$ -Elektronendichte bei einem Konturierungsgrad von $\sigma = 1.0$ deutlich wird. Hierbei zeigt die Sulfat-Gruppe aus der Bindungstasche heraus. B: Die Orientierung des Hexoserings der Galactose entspricht der für Gal β 1-3Glc (transparent; PDB-ID: 4AF9) beobachteten. Direkte Wechselwirkungen zwischen dem Sulfat und Epa1A treten nicht auf, allerdings sind drei Wassermoleküle, welche in der Epa1A•Gal β 1-3Glc Struktur vorhanden sind, durch das Sulfat ersetzt.

Obwohl die Sulfat-Gruppe als Elektronen- und damit donor-/akzeptorreiche Gruppe gebunden vorliegt, ist die Anzahl der Interaktionen zwischen diesem Liganden und Epa1A begrenzt. Offenbar sind die konservierten Bereiche (C1-C5) des Glycans die primären Interaktionspartner bei der Erkennung galactosidischer Liganden. Möglicherweise dient dieser Bindungsmodus dazu, Galactoside bei Gewährung einer Variabilität der Derivatisierung, speziell an der C6-Position der Galactose, zu erkennen.

Um weitergehend zu untersuchen, inwiefern sich die Derivatisierung der zweiten, sekundären Hexose auf die molekulare Erkennung der Glycane durch Epa1A auswirkt, wurden ähnliche *replacement-soaking* Experimente mit synthetischen Galactosiden durchgeführt. Diese wurden von Prof. LIN von der *Academia Sinica* in Kooperation mit MANUEL MAESTRE-REYNA zur Verfügung gestellt.

Die zuvor reproduzierten Epa1A-Kristalle wurden für diese Experimente mit drei Disacchariden versetzt, welche aus einer Galactose gefolgt von einer β 1-3-verknüpften derivatisierten Glucose bestanden. Dies entspricht dem Liganden, welcher als primärer Bindungspartner in Lactose-Kokristallisationsexperimenten gefunden wurde¹⁵⁸. Neben ihrer Grundform, in welcher die Glucose an der C6-Position eine (i) OH-Gruppe trägt, wurde diese durch eine (ii) Sulfat-Gruppe oder eine (iii) Carboxy-Gruppe modifiziert. Von Kristallen, welche jeweils mit einem der Sacchride *gesoakt* wurden, konnten am BESSY-II Synchrotron vollständige Da-

tensätze aufgenommen werden. Die Prozessierung erfolgte in diesem Fall mittels *imosflm*²⁷⁴ und *SCALA*²⁷⁸, zwei Runden der Verfeinerung erfolgten wie zuvor für den 6-O-Sulfogalactose-Komplex. Die zugehörigen Statistiken sind folgenden Tabellen zu entnehmen.

Tabelle 4.15 - Datensammlungsstatistik der Epa1A•Disaccharid-Komplexe.

	Epa1A•Galβ1-3Glc	Epa1A•Galβ1-3Glc-6COOH	Epa1A•Galβ1-3Glc-6S
Datensatz	MK251	MK250	MK252
Datum	13.02.2013	13.02.2013	13.02.2013
Röntgenquelle	BESSY II BL14.1	BESSY II BL14.1	BESSY II BL14.1
Detektor	PILATUS 6M	PILATUS 6M	PILATUS 6M
Temperatur (K)	100 K	100 K	100 K
Wellenlänge (Å)	0.91841	0.91841	0.91841
Raumgruppe	C222 ₁	C222 ₁	C222 ₁
Zellmaße (Å)			
a	74.84	74.37	75.02
b	103.78	102.75	104.40
c	69.49	68.98	69.79
α,β,γ	90°, 90°, 90°	90°, 90°, 90°	90°, 90°, 90°
Moleküle / a.s.u.	1	1	1
Auflösung (Å)	37.42-1.35 (1.42-1.35)	41.20-1.25 (1.32-1.25)	37.50-1.25 (1.32-1.25)
Beobachtete Reflexe	361183	468781	497773
Multiplizität	6.1 (5.7)	6.4 (6.5)	6.6 (6.7)
Einzigartige Reflexe	58995	73154	75797
R_{merge} (%)	8.2 (58.1)	5.2 (45.8)	5.1 (56.2)
Vollständigkeit (%)	99.1 (98.4)	100.0 (100.0)	99.9 (100.0)
Mittleres I/σ(I)	10.5 (2.8)	15.5 (3.8)	15.2 (3.0)
Mosaizität (°)	0.16	0.12	0.11
Wilson B-Faktor (Å²)	14.6	12.1	13.4

Tabelle 4.16 - Verfeinerungsstatistik der Epa1A•Disaccharid-Komplexe.

	Epa1A•Galβ1-3Glc	Epa1A•Galβ1-3Glc-6COOH	Epa1A•Galβ1-3Glc-6S
Auflösung (Å)	31.40-1.35	41.20-1.25	33.04-1.25
R_{work}, R_{free} (%)	15.33, 18.65	14.96, 17.53	17.17, 19.28
Reflexe (Arbeits-, Test-)	58981, 2043	73126, 2526	75768, 2608
Vollständigkeit (%)	98.96	99.97	99.92
r.m.s.d.			
Bindungslängen(Å)	0.015	0.015	0.014
Bindungswinkel (°)	1.56	1.55	1.52
Anzahl Atome	2243	2212	2132
Mittlerer B-Wert (Å²)	14.2	12.2	13.5

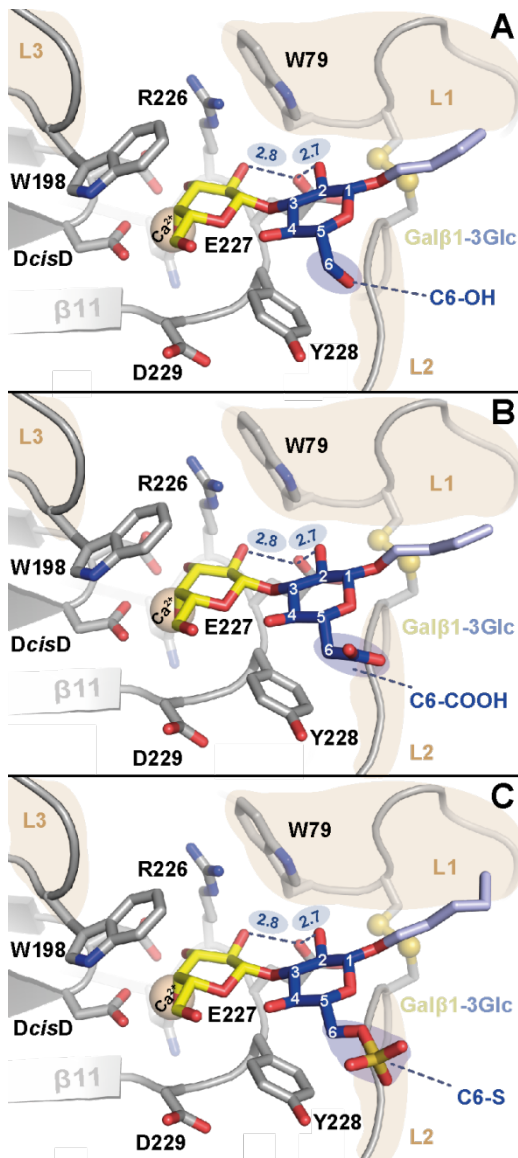


Abbildung 4.36 – Bindung von Liganden mit sekundären, derivatisierten Glucosiden. Der Bindungsmodus aller drei Liganden ist äquivalent. A: Epa1A•Galβ1-3Glc. B: Epa1A•Galβ1-3Glc-6COOH. C: Epa1A•Galβ1-3Glc-6S

Die Bindungsmodi der drei Liganden unterscheiden sich nur minimal. Die Interaktionen von Epa1A mit dem Disaccharid sind jeweils konserviert. Die Derivatisierung an C6 der Glucose zeigt jeweils vom Protein weg und ist damit in keine direkten Wechselwirkungen involviert (Abbildung 4.36). Die Glucose enthält in jedem Disaccharid an C1 einen von der Synthese stammenden Azido-n-Hexanol-Linker, von welchem in der Elektronendichte bis zu 5 CH₂-Gruppen zu erkennen sind.

Möglicherweise sorgt W79 aus Loop L1 für eine sterische Hinderung der Rotation der derivatisierten Glucose und zwingt diese in ihre beobachtete Orientierung. Dieser Effekt wird durch eine Wasserstoffbrücke von W79 über ein Wassermolekül zur C2-OH-Gruppe verstärkt. In der Komplexstruktur Epa1•T-Antigen ist dieser Effekt ebenfalls zu beobachten. Hier sorgt dieses Tryptophan offenbar durch eine Kombination aus sterischer Hinderung und direkter Wechselwirkung mit der C2-Acetylgruppe des *N*-Acetylglucosamins für eine Rotation des GlcNAc um etwa 90°. Weiterhin ist das C2-OH des Glucosederivats in eine Wasserstoffbrücke zu E227 aus CBL2 involviert.

4.3.3 Die Struktur und Bindungseigenschaften von Epa6A

Epa6A als weiterer Vertreter der Familie der epithelialen Adhäsine zeigt ein im Vergleich zu Epa1A verändertes Bindungsmuster. Anstatt, wie Epa1A, hauptsächlich die Bindung von β-glycosidisch verbrückten Galactosiden zu ermöglichen, zeigt Epa6A die Fähigkeit, ebenfalls α-verknüpfte Glycane zu erkennen¹⁶⁵. Weiterhin zeigt Epa6A eine generell stark erhöhte Promiskuität im Vergleich zum spezifischen Epa1A und lässt sich trotz der hohen Sequenzidentität von 79% einer anderen funktionellen Klasse zuordnen¹⁶⁵. Dies mag

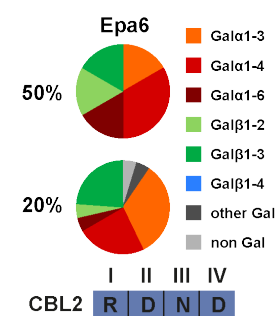


Abbildung 4.37 – Spezifität von Epa6A. Ergebnisse der Glycan Array Studien. Nach DIDERRICH.

vor allem in der unterschiedlichen Zusammensetzung des für die Bindung wichtigen Loops CBL2 begründet liegen, doch zeigt sich die Ligandenerkennung aufgrund der hypervariablen Loops L1-L3 weitaus komplexer. Daraufhin wurden die strukturellen Grundlagen für dieses Verhalten untersucht, um die Bindungseigenschaften der epithelialen Adhäsine auf molekularer Ebene besser verstehen zu können. Da Epa6A eine breite Spezifität aufweist, wurde eine Kollektion von Galactosiden, welche sich in ihrer Verknüpfung und in der sekundären Hexose unterscheiden, eingesetzt, um Komplexstrukturen zu erhalten.

Tabelle 4.17 – Zur Studie von Epa6A-Komplexstrukturen verwendete Disaccharide.

Name	Zusammensetzung
Lactose	Gal β 1-4Glc
T-Antigen	Gal β 1-3GalNAc
Lacto- <i>N</i> -Biose	Gal β 1-3GlcNAc
<i>N</i> -Acetyl-D-lactosamin	Gal β 1-4GlcNAc
α -Galactobiose	Gal α 1-3Gal

Die Bindung von Epa6A an diese Disaccharide wurde zuvor mittels isothermaler Titrationskalorimetrie und durch RIKE DIDERRICH mittels Fluoreszenztitrationsanalysen quantifiziert. Hierbei konnte die Promiskuität von Epa6A ebenfalls beobachtet werden, denn alle Liganden außer Lacto-*N*-Biose binden mit vergleichbarer Stärke an Epa6A. Dieser Trend konnte mit beiden Methoden beobachtet werden. Mit Hilfe der Fluoreszenztitrationsanalysen konnte weiterhin gezeigt werden, dass die Bindung von Epa6A an α 1-3 Galactobiose im Vergleich zu Epa1A und Epa7A am stärksten ist¹⁶⁵.

Die Kristallisation sowie das *replacement-soaking* von Epa6A wurden von RIKE DIDERRICH und HOLGER STEUBER durchgeführt¹⁶⁵. Vollständige Datensätze konnten daraufhin für jeden Komplex gesammelt werden (Statistiken der Datensammlung im Anhang in Tabelle 7.1). Die Prozessierung und Datenreduktion erfolgte mit Hilfe von *XDS*²⁷⁶, *XSCALE*²⁷⁷ und *SCALA*²⁷⁸.

Aufgrund der hohen Sequenzidentität zwischen Epa1A und Epa6A wurde auch in diesem Fall eine Lösung des Phasenproblems mittels molekularem Ersatzes getestet. Hierzu wurde ein Modell auf Basis von Epa1A mit verkürzten Loop-

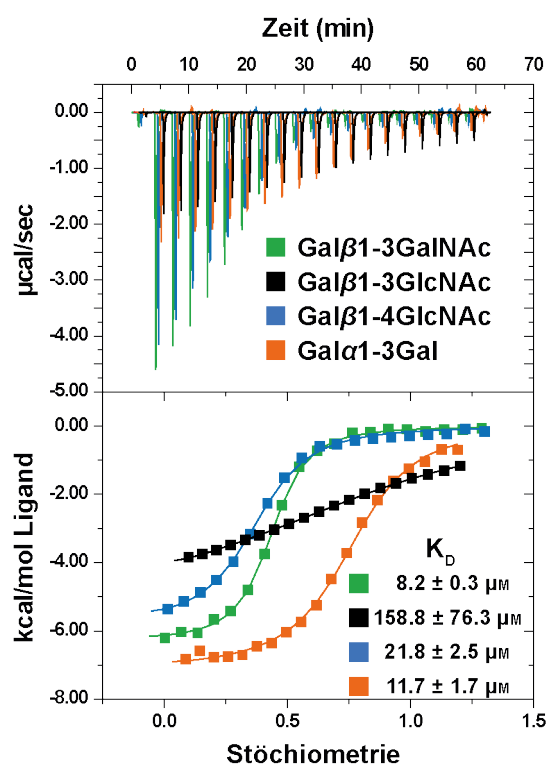


Abbildung 4.38 – ITC Experimente mit Epa6A. Pro Disaccharid ist eine repräsentative Messung gezeigt. Die Bindungskonstante ist für alle Disaccharide, abgesehen von Lacto-*N*-Biose (schwarz), vergleichbar. Ebenfalls konnte hierbei die Bindung von Epa6A an α -verknüpfte Galactoside verfolgt werden (orange). Der Trend in den Affinitäten konnte mittels Fluoreszenztitrationsanalysen bestätigt werden.

Bereichen mittels *modeller9v7* erstellt²⁹². In Zusammenarbeit mit MANUEL MAESTRE-REYNA konnten in *phaser* erfolgreich erste Phasen mit einem TFZ-Wert von 30.3 erhalten werden. Die Verfeinerung der Komplexstrukturen erfolgte mit *phenix refine*²⁹¹, *REFMAC5*³⁰⁰ und *Coot*²⁹⁷ (Verfeinerungsstatistiken im Anhang in Tabelle 7.1).

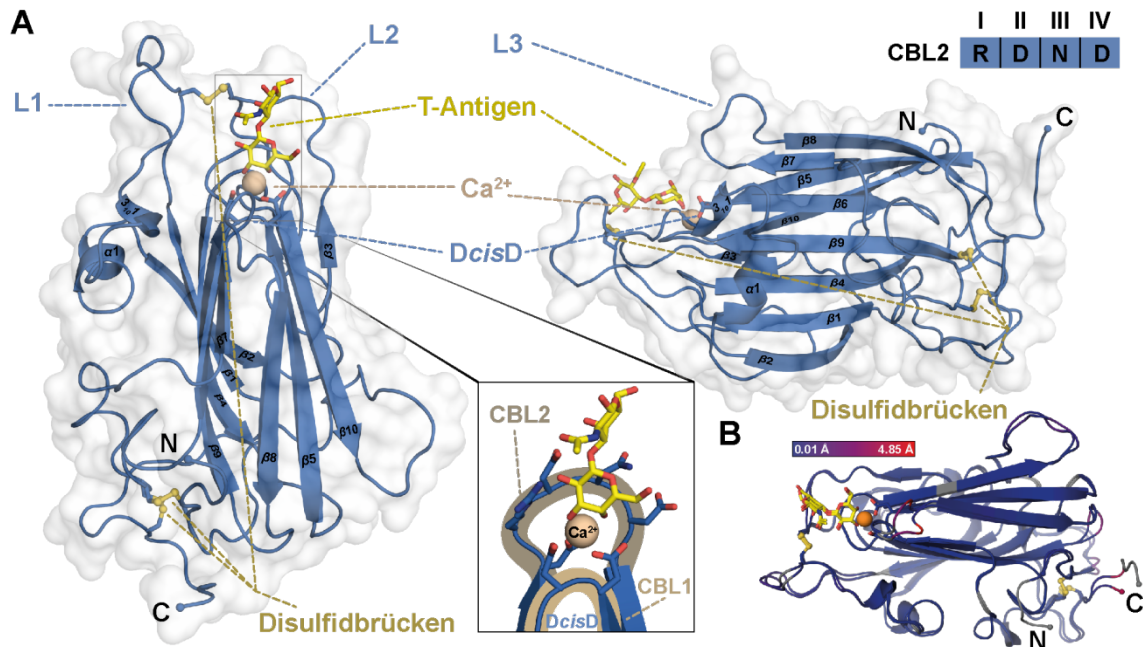


Abbildung 4.39 - Die Kokristallstruktur Epa6A•T-Antigen. A: Neben dem konservierten β -Sandwich zeigt Epa6A das für die epithelialen Adhäsine typische Disulfidbrückenmuster (2x im terminalen Bereich; 1x zur Fixierung von L1 & L2). B: Der Vergleich zwischen Epa6A•T-Antigen und Epa1•T-Antigen zeigt nur geringe Abweichungen im Bereich der Termini und in L3. Die Strukturen sind nach Größe des *r.m.s.d.* gefärbt (rot: hohe Abweichung; blau: geringe Abweichung).

Die globale Faltung von Epa6A ist der von Epa1A mit einem *r.m.s.d.* von 0.37 Å über 177 C_{α} -Positionen sehr ähnlich (Abbildung 4.39). Es besteht ebenfalls aus dem strukturell konservierten β -Sandwich-Motiv der PA14-Domäne, der typischen Bindungstasche aus den zwei Calciumbindungs-Loops CBL1 & 2 sowie den hypervariablen Loops L1-L3, wobei L1 und L2 durch eine Disulfidbrücke miteinander verknüpft sind und L3 das für die Epa-Familie typische Tryptophan enthält. Die Verknüpfung beider Loops verengt die Bindungstasche und schützt diese vor dem umgebenden Lösungsmittel. Unterschiede werden allerdings im Detail deutlich. Neben den beiden konservierten, über eine *cis*-Peptidbindung verknüpften Aspartaten aus CBL1 besitzt CBL2 eine andere Zusammensetzung als in Epa1A. Der Einfluss dieser geänderten Region zeigt sich bei Vergleich der Bindung des T-Antigens (Abbildung 4.40).

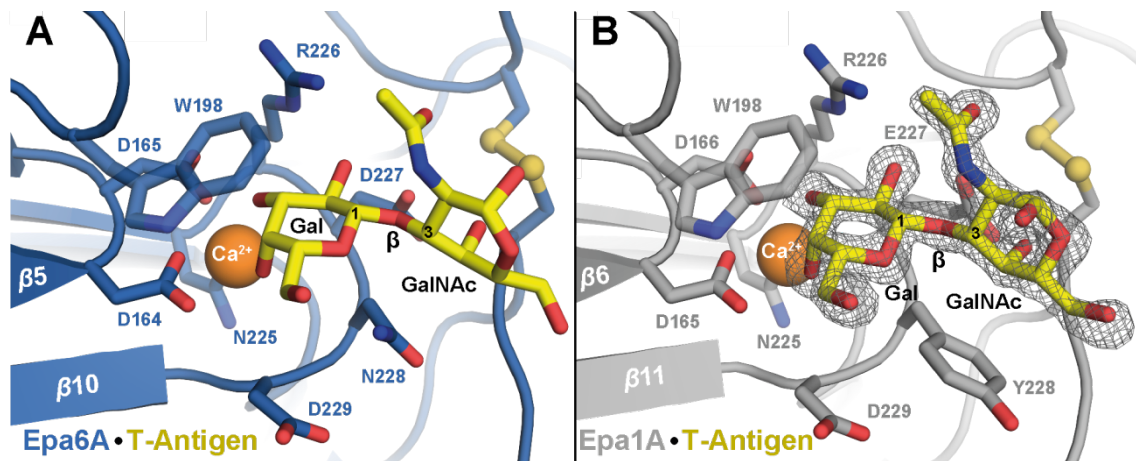


Abbildung 4.40 – Vergleich der Bindungsmodi von Epa6A•T-Antigen und Epa1A•T-Antigen. Obwohl die Orientierung des Liganden in beiden Komplexstrukturen sehr ähnlich ist, zeigen sich verschiedene Interaktionen. Vor allem Position II und III sind direkt in die Erkennung des Liganden und der glycosidischen Bindung involviert.

Obwohl die Orientierung des Liganden vergleichbar ist, unterscheidet sich das Interaktionsmuster. Im Fall von Epa6A sind D227 und N228 an der Ligandenbindung beteiligt, während die korrespondierenden Reste in Epa1A E227 und Y228 sind. Vergleicht man Epa6A• α -Galactobiose mit einem Modell von Epa1A• α -Galactobiose, so wird deutlich, dass die Zusammensetzung von CBL2 einen Einfluss auf die Effizienz der Diskriminierung zwischen α - und β -verknüpften Galactosiden hat (Abbildung 4.41). Denn während Epa6A an Position II von CBL2 das kürzere D227 trägt, findet sich an dieser Stelle in Epa1A das längere E227. Dies deutet darauf hin, dass die Verkürzung dieses Restes eine effizientere Erkennung von α -verknüpften Galactosiden durch eine geringere sterische Hemmung ermöglicht.

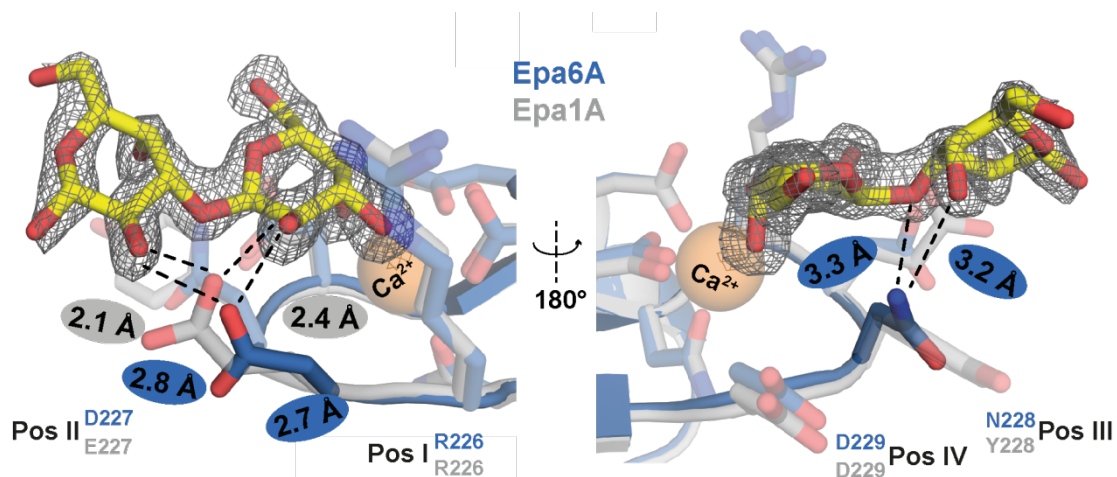


Abbildung 4.41 – Die Bindung α -glycosidisch verknüpfter Galactoside durch Epa6A. Die Überlagerung der Struktur von Epa1A (grau) mit Epa6A• α -Galactobiose (blau) erklärt die größere Selektivität von Epa6A für α -verknüpfte Galactoside. Ein voluminöseres Glutamat in Position II des CBL2 von Epa1A sorgt für eine sterische Hinderung bei Bindung solcher Liganden. Epa6A hingegen trägt an dieser Stelle ein kürzeres Aspartat.

Weiterhin kommt es in Epa6A• α -Galactobiose zu einer in Epa1A abwesenden Wasserstoffbrückenbindung zwischen Position III (N228) und dem Liganden. Diese Beobachtung deckt sich mit Experimenten von MAESTRE-REYNA *et al.*, in welchen die CBL2 Region von Epa1A

gegen die von Epa6A ausgetauscht wurde und dadurch eine verringerte Diskriminierung zwischen beiden Bindungstypen gefunden wurde¹⁰⁸.

Die Bindung von Liganden mit Unterschieden in Verknüpfung und zweiter Hexose führt trotz konservierter Orientierung der primären Galactose zur Torsion des zweiten Restes entlang der glycosidischen Bindung im äußeren Bereich der Bindungstasche (Abbildung 4.42 A-D). Dies wird vor allem durch alternative H-Brücken zwischen D227 und der 4-OH Gruppe des GlcNAc in Gal β 1-3GlcNAc, des 3-OH des GlcNAc in Gal β 1-4GlcNAc und der 6-OH der Glucose in Lactose erreicht.

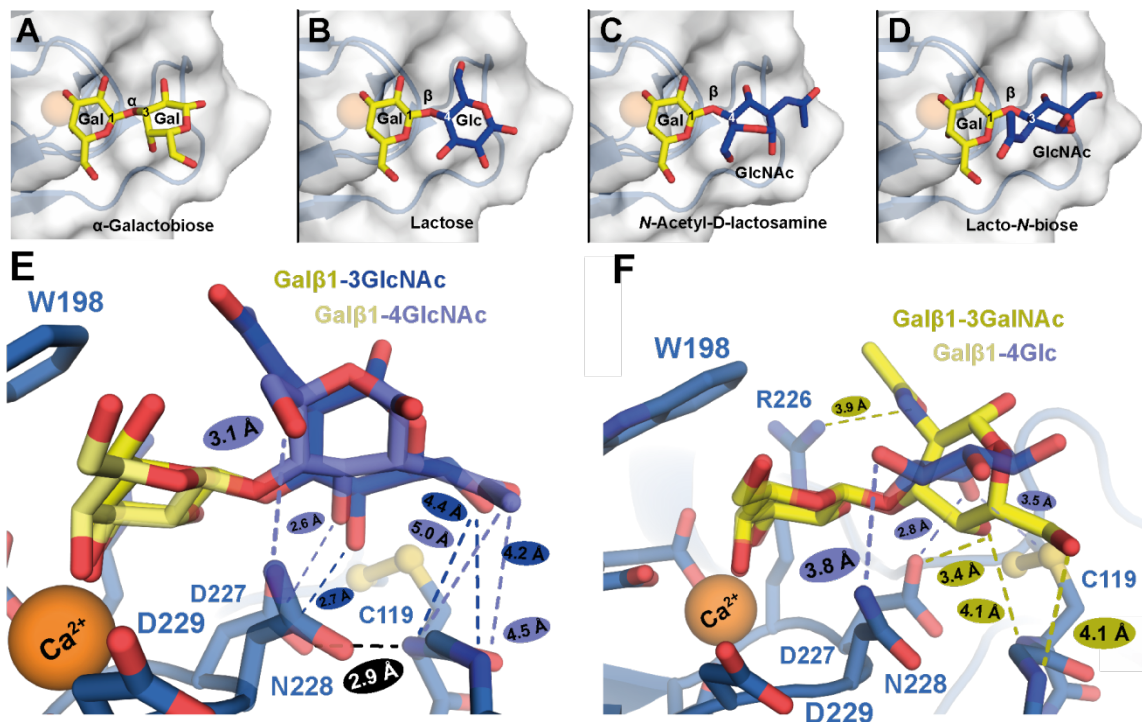


Abbildung 4.42 – Bindungsmodi unterschiedlicher Disaccharide in Epa6A. A-D: In Abhängigkeit der glycosidischen Bindung und des sekundären Saccharids ist die Orientierung des zweiten Zuckers bei Bindung an Epa6A sehr divers. E: Die Änderung der glycosidischen Bindung von β 1-3 zu β 1-4 sorgt für ein alternatives Interaktionsmuster und verursacht damit eine Rotation des GlcNAc um etwa 180° . F: Die Überlagerung der Komplexe von Epa6A mit Gal β 1-3GalNAc und Gal β 1-4Glc unterstreicht die Komplexität der Erkennung des sekundären Zuckers und deutet darüber hinaus auf einen im Vergleich zur Galactose verminderten energetischen Beitrag zur Gesamtbindung hin.

Korrespondierend zu Epa1A wird die Konformation der Seitenkette von D227 in Epa6A durch eine H-Brücke zur Peptid-Gruppe von C119 in L2 stabilisiert.

Wird bei der Erkennung des Disaccharides Gal-GlcNAc die glycosidische Bindung von β 1,4 zu β 1,3 geändert, führt dies zur Rotation des GlcNAc um nahezu 180° und einem Austausch der Interaktionspartner (Abbildung 4.42 E). Während die 6-OH-Gruppe des β 1,3-verknüpften Saccharids mit C119 wechselwirkt, zeigt bei β 1,4-Verknüpfung die *N*-Acetyl-Gruppe in diese Richtung. Bei letzterem Liganden wird darüber hinaus eine Salzbrücke zwischen dessen 6-OH und N228 gebildet. Dies unterstreicht die Komplexität der von pilzlichen Adhäsinen mit PA14-Domänen entwickelten Bindungsmodi; obwohl die gleichen Reste

der Bindungstasche in die Ligandenerkennung involviert sind, ändert sich das Interaktionsmuster deutlich. Dieser Effekt lässt sich ebenfalls bei Vergleich von Epa6A•Galβ1-3GalNAc mit Epa6A•Galβ1-4Glc beobachten; hier ist die zweite Hexose um etwa 90° gedreht, was wiederum zu alternierenden Wechselwirkungen führt (Abbildung 4.42 F).

Diese Differenzen in der Orientierung des zweiten Zuckers deuten darauf hin, dass dessen energetischer Beitrag zur Gesamtbindung des Galactosids unabhängig von seiner Position und Orientierung in der äußeren Bindungstasche ist.

4.3.4 Die Struktur und Komplexstrukturen von Epa9A

Die Kristallisation von Epa9A durch RIKE DIDERRICH ermöglichte, durch die Lösung dessen Struktur im Rahmen dieser Arbeit, neben Epa1A ein zweites epitheliales Adhäsins mit ähnlicher Spezifität (Abbildung 4.43, Funktionsklasse I nach DIDERRICH *et al.*¹⁶⁴) zu untersuchen. Darüber hinaus ist das Studium von Epa9A interessant, da es trotz der Einordnung in dieselbe Funktionsklasse nur eine Sequenzidentität von 42% zu Epa1A aufweist. Dies ist unter anderem damit begründet, dass Epa9A, laut multiplem Sequenzalignment, einen stark verlängerten hypervariablen Loop L1 besitzt¹⁶⁵.

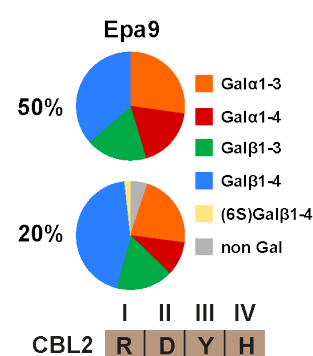


Abbildung 4.43 - Spezifität von Epa9A. Ergebnisse der *Glycan Array* Studien. Nach DIDERRICH.

Für die Analyse von Komplexstrukturen wurden ebenfalls die in Tabelle 4.17 aufgelisteten Disaccharide verwendet. Die Lösung des Phasenproblems sowie die Verfeinerung der Strukturen wurden, unter Vorarbeit von HOLGER STEUBER, in Anlehnung an das Vorgehen für Epa6A durchgeführt. Für den molekularen Ersatz wurde ein Modell auf Basis von Epa1A verwendet und es konnte hierbei ein TFZ-Wert von 21% erhalten werden. Die Statistiken zur Datensammlung und Verfeinerung der Modelle sind im Anhang in Tabelle 7.2 zu finden.

Im Gegensatz zu Epa6A konnten für Epa9A, abgesehen von *Glycan Array Analysen*¹⁶⁵, keine Bindungsstudien *in vitro* durchgeführt werden. Zum Einen konnten mittels isothermaler Titrationskalorimetrie keine eindeutigen Bindungsisothermen erhalten werden. Zum Anderen zeigte Epa9A einen zu geringen Fluoreszenz-*quench* um Fluoreszenztitrationen mit aussagekräftigen Ergebnissen durchführen zu können¹⁶⁵. Die strukturelle Betrachtung der Wechselwirkungen zwischen Epa9A und seinen Liganden liefert allerdings weitere Einblicke in das Bindungsverhalten dieses Adhäsins.

Der strukturelle Kern von Epa9A ist sehr stark vergleichbar mit Epa1A und Epa6A (Abbildung 4.44). So beträgt der *r.m.s.d.* zwischen Epa9A und Epa1A nur 0.64 Å über 144 C_α-Atome. Leichte Abweichungen zeigen hier die Bereiche der Loops L1 und L2. Interessanterweise ist die Verlängerung von L1 nicht in der Elektronendichte definiert. Möglicherweise ist dieser sehr flexibel und daher dessen Orientierung im Einkristall nicht homogen.

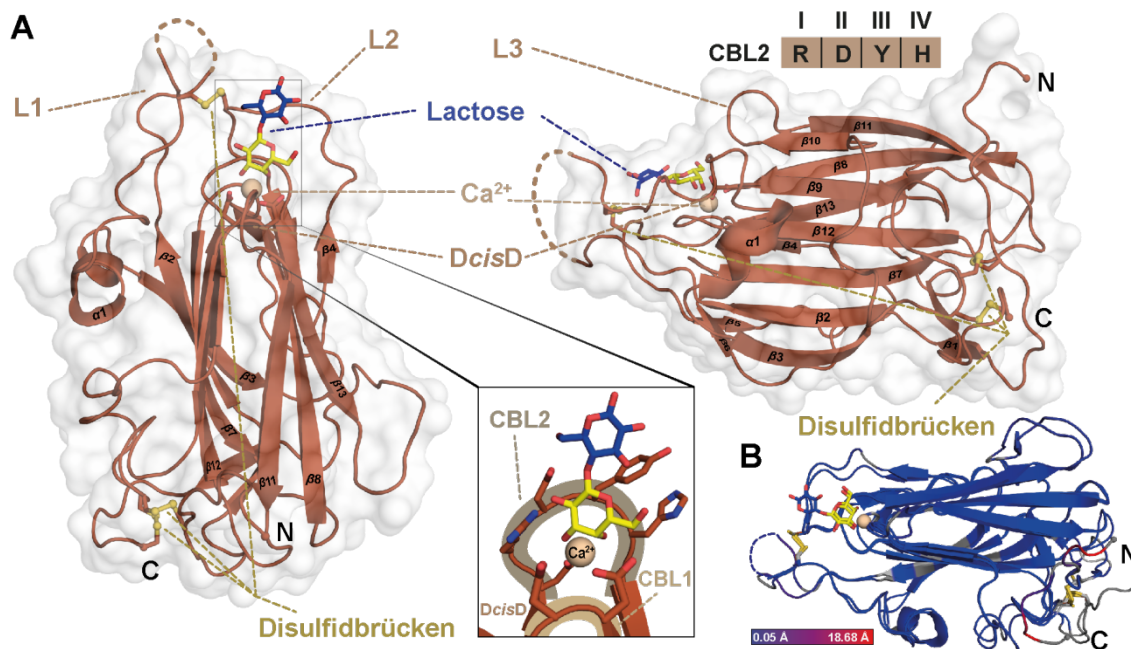


Abbildung 4.44 - Die Kokristallstruktur Epa9A•Lactose. A: Neben dem konservierten β -Sandwich zeigt Epa9A das für die epithelialen Adhäsine typische Disulfidbrückenmuster (2x im terminalen Bereich; 1x zur Fixierung von L1 & L2). B: Der Vergleich zwischen Epa9A•Lactose und Epa1•T-Antigen zeigt Abweichungen im Bereich der Termini (etwa 18,7 Å), in L3 und in L1, welcher in Epa9A stark verlängert, aber nicht durch die Elektronendichte abgebildet ist. Die Strukturen sind nach Größe des *r.m.s.d.* gefärbt (rot: hohe Abweichung; blau: geringe Abweichung).

Die Orientierung der Lactose in den entsprechenden Komplexstrukturen von Epa6A und Epa9A ist vergleichbar, doch sind aufgrund der unterschiedlichen Reste in CBL2 die Interaktionen modifiziert. In diesem Loop der epithelialen Adhäsine ist Position I absolut (N) und Position II sehr stark (R) konserviert. An Position III und IV befinden sich in Epa9A Y259 gefolgt von H260.

Durch diese unterschiedliche Sequenz haben Epa6A und Epa9A verschiedene Mechanismen zur Erkennung von Glycanen, hier speziell von Galactosiden, entwickelt. Dies wird beim Vergleich der Orientierung der Disaccharide in der Bindungstasche deutlich. Während in Epa9A•Lacto-*N*-Biose (Gal β 1-3GlcNAc) der GlcNAc Rest im Vergleich zu Epa6A•Lacto-*N*-Biose um etwa 90° verdreht ist, beträgt der Unterschied in der Rotation des GlcNAc-Rests bei β 1-4-Verknüpfung in der *N*-Acetyl-D-lactosamin-Komplexstruktur fast 180° (Abbildung 4.45).

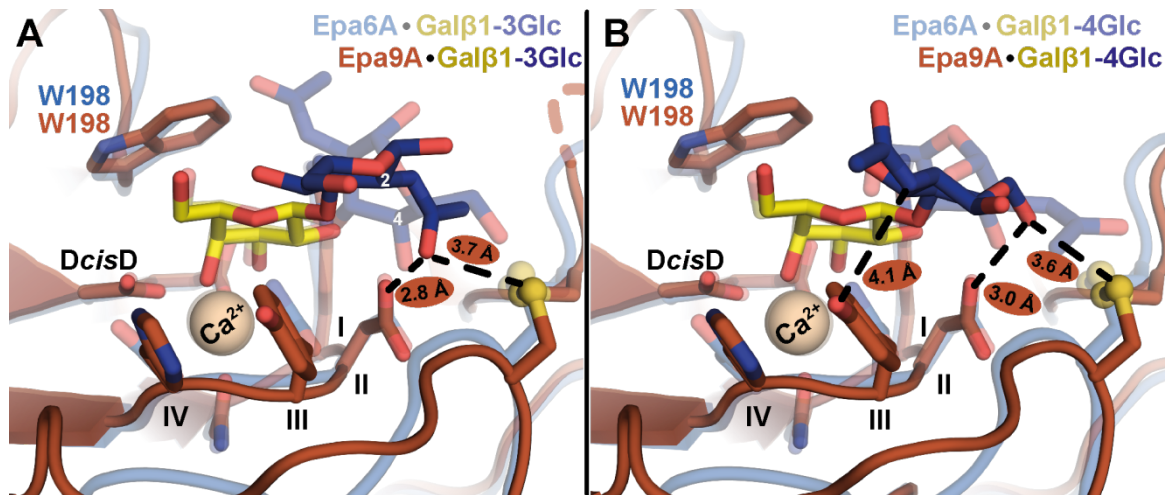


Abbildung 4.45 – Die äußere Bindungstasche kontrolliert die Orientierung der zweiten Hexose. Wie bereits bei Epa6A zeigen auch die Komplexstrukturen von Epa9A ein komplexes Interaktionsmuster in der äußeren Bindungstasche. A: Die veränderte Form der Bindungstasche in Epa9A (braun) führt zur Torsion der Glucose von Galβ1-3Glc um etwa 90° im Vergleich zum Komplex mit Epa6A (blau; transparent) und sorgt für alternative Wechselwirkungen. B: Dieser Effekt sorgt bei β1-4-verknüpfung für eine Rotation um etwa 180° und damit für einen Austausch der Bindungspartner.

Im Epa9A•Lacto-*N*-Biose Komplex ist eine senkrechte Orientierung des GlcNAc zur Galactose, wie dies in Epa6A der Fall ist, sterisch nicht möglich. Eine direkte Interaktion zwischen dem zweiten Zucker und Epa9A wird nur durch D258 vermittelt. Hier nimmt in Epa9A•Lacto-*N*-Biose die *N*-Acetylgruppe des GlcNAc die Position der C4-OH Gruppe des Epa6A-Komplexes ein (Abbildung 4.45 A). Diese verminderte Interaktion, aufgrund sterischer Kontrolle, korrespondiert mit den Ergebnissen der *Glycan-Array* Experimente, in denen Epa6A eine bessere Erkennung von β1-3-verknüpften Galactosiden zeigt.

Die Rotation von 180° im Fall der β1-4-Verknüpfung lässt sich ebenfalls durch die veränderte Anzahl und Art der Interaktionsmöglichkeiten sowie durch den differenten sterischen Anspruch der Bindungstasche erklären. Während Epa6A mittels N228 eine Interaktion mit der 6-OH-Gruppe des GlcNAc ermöglicht, ist Y259 in Epa9A weiter von der *N*-Acetylgruppe, welche die Position der 6-OH Gruppe einnimmt, entfernt (4.1 Å). Stattdessen deutet die 6-OH Gruppe hier in Richtung der Disulfidbrücke und D258 (3.0 Å) (Abbildung 4.45 B).

Die in Epa6A vorhandenen Interaktionen zwingen den GlcNAc Rest in die vorgefundene, verzerrte Orientierung, während in Epa9A beide Hexosen fast koplanar vorliegen. Der sterische Anspruch durch zwei voluminöse Reste in CBL2 lässt offenbar nur eine koplanare Anordnung der beiden Hexosen zu, während die im Vergleich kleineren Reste in Epa6A eine größere Variation in der Orientierung des zweiten Restes ermöglichen. Dies erklärt ebenfalls die bessere Erkennung von sterisch anspruchsvolleren α-verknüpften Galactosiden durch Epa6A.

4.4 Strukturelle Untersuchungen bezüglich Flokkuline – Flo10A

Die Lösung der Struktur von Flo5A legte den Grundstein für die weitergehende strukturelle und funktionelle Untersuchung der pilzlichen Adhäsine mit PA14-Domäne³²⁴. Wie die phylogenetische Studie aus Abschnitt 4.1 darlegte, zeichnen sich die Subtypen unter anderem durch Insertion an den Positionen der hypervariablen Loops L1-L3 aus. Während die epithelialen Adhäsine oft einen verlängerten Loop L1 sowie eine Disulfidbrücke zwischen diesem und L2 aufweisen, tragen Flokkuline häufig eine subdomänenartige Insertion in Loop L2. Flo5 und Flo10 sind sich sehr ähnlich (Identität der A-Domänen etwa 70%). Neben einer abweichenden Anzahl an Repeats in der B-Region zeigt Flo10A eine nicht nur verlängerte, sondern auf Sequenzebene auch sehr unterschiedliche Subdomäne (Abbildung 4.46).

Flo5A	MTIAHHCIFLVILAFLLALINVASGATEA	CLPAGQRKSGMN	40
Flo10A	MPVAARYIFLTGLFLLSVANVALGTTEA	CLPAGEKKNGMT	40
	<i>L1</i>		
Flo5A	INFYQYSLKDSSTYSNAAYMAYGYASKTKLGSVGGQTDIS		80
Flo10A	INFYQYSLKDSSTYSNPSYMAAYGYADA EKLGSVSGQTKLS		80
	<i>Flo-Subdomäne / L2</i>		
Flo5A	IDYNIPCVSSSGTFFPCPQEDSYGNWG CKGMGACS		114
Flo10A	IDYSIPCNGASDTCACSDDDATEYSASQVVPVKRQVVKLCS		120
	<i>Flo-Subdomäne / L2</i>		
Flo5A	N SQGIAYWSTD LFGFYTTPTNVT L		138
Flo10A	DNTTLLSSKTEKRENDDCDQGAAYWSSDLFGFYTTPTNVT V		160
	<i>CBL1</i>		
Flo5A	EMTGYFLPPQTGSYTF SFATVDDSAILSVGG SIAFECCAQ		178
Flo10A	EMTGYFLPPKTGT YTFGFATVDDSAILSVGG NVAFECCKQ		200
	<i>L3</i>		
Flo5A	EQPPIITSTNFTINGIKPWDGSLPDNITGTVYMYAGYYYP L		218
Flo10A	EQPPIITSTDFTINGIKPWNADAPTDIKGSTYMYAGYYYP I		240
	<i>CBL2</i>		
Flo5A	KVVYSNAVSWGTLPI SVELPDGTTVSDNFE GYVYSFDDDL		258
Flo10A	KIVYSNAVSWGTL PVS VVLPDGT EVNDDFEGYVFSFDDNA		280
	<i>I II III IV</i>		
Flo5A	SQSNCTIPDP		268
Flo10A	TQAHCSVPNP		290

Abbildung 4.46 – Ein Alignment der Primärsequenzen von Flo5A (UniProt-ID: P38894) und Flo10A (UniProt-ID: P36170) unterstreicht die hohe Ähnlichkeit und Identität, verdeutlicht allerdings auch die Unterschiede, welche hauptsächlich in der unterschiedlich großen Subdomäne sowie den abweichenden Sequenzen des hypervariablen Loops L3 zu finden sind.

Um den Einfluss der Subdomäne auf die Glycanbindung zu untersuchen, wurde die Kristallstruktur von nativem Flo10A in Kooperation mit NINA WOZNIAK untersucht. Obwohl die Kristallisation von Wildtyp-Flo10A zuvor nicht gelang konnte ein chimäres Adhäsins aus Flo10A und der Subdomäne von Flo5, Flo10A^{5SD} erfolgreich kristallisiert werden und dessen Struktur aufgeklärt werden¹⁵⁹.

4.4.1 Strukturlösung des Flo10A-Wildtyps

Weitere Kristallisationsexperimente von NINA WOZNIAK mit Flo10A und Mannose bei 18 °C lieferten einen trigonalen Kristall in Bedingung 94 des JCSG+-Screens (*QIAGEN*), dessen Streuvermögen am BESSY-II Synchrotron sehr gut für eine Strukturlösung geeignet war¹⁵⁹. Die gesammelten Daten wurden mit *XDS*, *XSCALE* und *SCALA* prozessiert und das Phasenproblem mit *MolRep*³²³ und dem bekannten Flo10A^{5SD}-Modell gelöst (*CONTRAST*-Wert von 21.94). Das Modell wurde daraufhin mit *REFMAC5*³⁰⁰ und *Coot*²⁹⁷ verfeinert.

Tabelle 4.18 – Datensammlungsstatistik für Flo10A•Mannose.

Flo10A•Mannose	
Datensatz	10-H10
Datum	16.05.2013
Röntgenquelle	BESSY II BL14.1
Detektor	Pilatus 6M
Temperatur (K)	100 K
Wellenlänge (Å)	0.91841
Raumgruppe	$P3_121$
Zellmaße (Å)	
a	76.63
b	76.63
c	76.23
α, β, γ	90°, 90°, 120°
Moleküle / a.s.u.	1
Auflösung (Å)	34.23-1.65 (1.74-1.65)
Beobachtete Reflexe	311673
Multiplizität	9.9 (9.4)
Einzigartige Reflexe	31603
R_{merge} (%)	6.8 (52.3)
Vollständigkeit (%)	100.0 (99.0)
Mittleres $I/\sigma(I)$	25.6 (4.8)
Mosaizität (°)	0.07
Wilson B-Faktor (Å ²)	13.0

Tabelle 4.19 – Verfeinerungsstatistik für Flo10A•Mannose.

Flo10A•Mannose	
Auflösung (Å)	34.26-1.65
$R_{\text{work}}, R_{\text{free}}$ (%)	15.92, 18.61
Reflexe (Arbeits-, Test-)	30009, 1594
Vollständigkeit (%)	99.98
r.m.s.d.	
Bindungslängen(Å)	0.008
Bindungswinkel (°)	1.31
Anzahl Atome	2219
Mittlerer B-Wert (Å ²)	19.31

Neben dem konservierten β -Sandwich finden wir in Flo10A die bekannten Strukturelemente der Adhäsine mit PA14-Domänen, wie das *DcisD*-Motiv in CBL1 welches gemeinsam mit CBL2 ein Ca^{2+} -Ion komplexiert, die hypervariablen Loops und das konservierte Disulfidbrückenmuster (Abbildung 4.47). Im Gegensatz zum sehr ähnlichen Flo5A (*r.m.s.d.* von 0.28 Å)

und der Flo10A^{5SD}-Chimäre (*r.m.s.d.* von 0.30 Å) ist die Subdomäne an Position des Loops L2 allerdings nicht durch die Elektronendichte beschrieben.

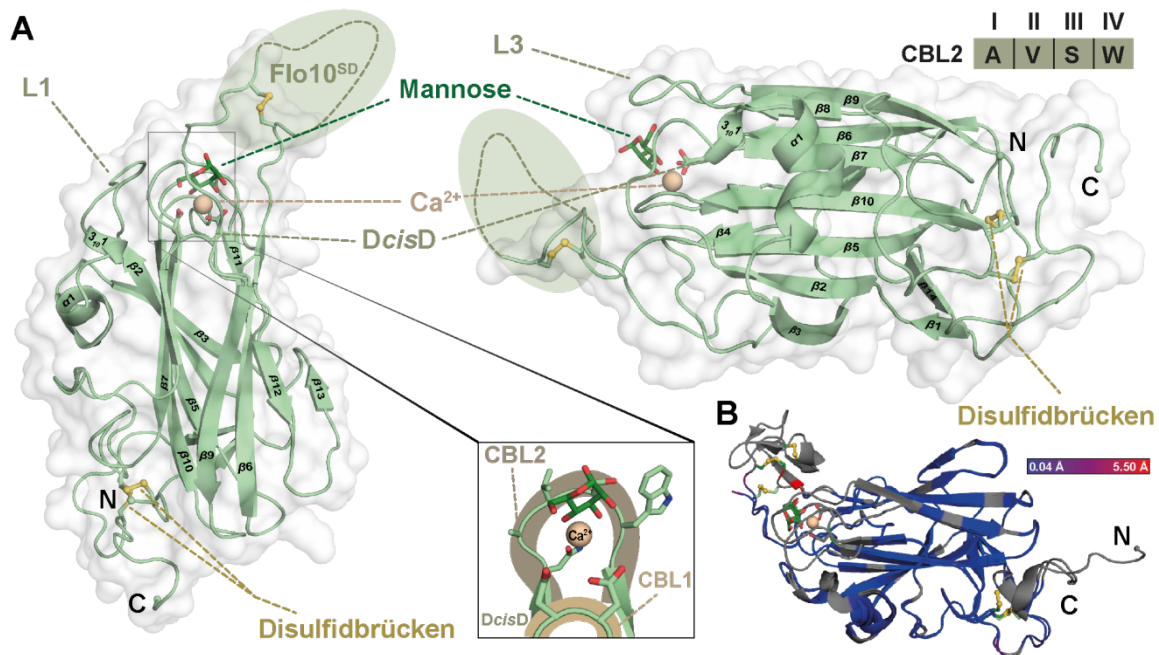


Abbildung 4.47 - Die Struktur von Flo10A. A: Neben dem konservierten β -Sandwich und den Disulfidbrücken trägt Flo10A eine nicht durch die Elektronendichte abgebildete Subdomäne. Die Konformation des hypervariablen Loops L3 ist der deckelförmigen Orientierung in *Pp*Flo1A sehr ähnlich. CBL1 (hellbraun) komplexiert mit Hilfe des DcisD-Motivs ein Ca^{2+} -Ion (ocker) über welches eine Mannose gebunden vorliegt. B: Der Vergleich zwischen Flo10A•Mannose und Flo5A•Mannose zeigt hauptsächlich Abweichungen im Bereich Subdomäne, in L3 und im Bereich der Termini. Die Strukturen sind nach Größe des *r.m.s.d.* gefärbt (rot: hohe Abweichung; blau: geringe Abweichung; grau: keine geeignete Überlagerung)

Das Modell von Flo10A bricht nach der ersten Disulfidbrücke der Subdomäne ab. Möglicherweise ist die Subdomäne von Flo10A flexibler als diejenige von Flo5A, welche in allen bekannten Strukturen eine definierte Orientierung besitzt. Betrachtet man allerdings die Bindungstasche genauer und vergleicht diese mit Flo5A und Flo10A^{5SD}, so fallen weitere Abweichungen auf.

Nicht nur die Orientierung der Mannose ist leicht verändert, sondern vor Allem die Orientierung von L3. Während in Flo10A^{5SD} dieser Loop die Konformation wie in Flo5A einnimmt, zeigt L3 in Flo10A eine geschlossene, deckelförmige Konformation; ähnlich wie in Cea1A. Weiterhin interagiert K216 direkt mit der gebundenen Mannose, was ebenfalls der Bindung des GlcNAc in Cea1A entspricht. Möglicherweise ist die Modifizierung des Bindungsmodus in Flo10^{5SD} durch die Interaktion von Q98 aus der Flo5A-Subdomäne mit der Mannose induziert. Hierdurch wird die Mannose leicht gekippt und Loop L3 ändert seine Konformation. Weiterhin ist ein sterischer Einfluss der rigiden Flo5-Subdomäne möglich.

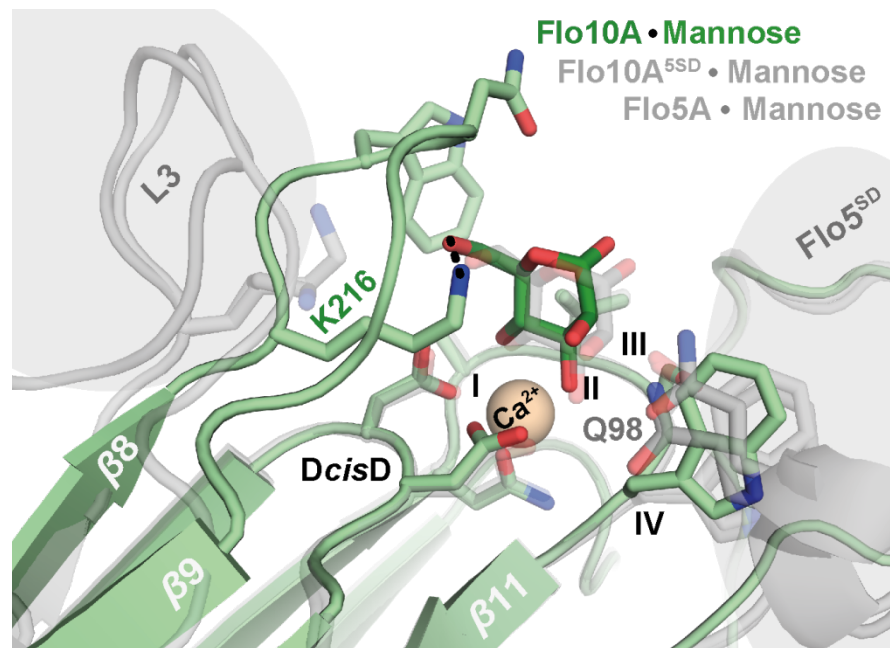


Abbildung 4.48 – Vergleich der Form der Bindungstasche von Flo10A mit Flo5A (PDB-ID: 2XJS) und dem chimären Flo10A^{5SD}. Die Anwesenheit der Flo5A-Subdomäne führt zu einer sterischen Restriktion der Bindungstasche durch Q98. Hierdurch wird die Orientierung der Mannose leicht gekippt und damit L3 in eine veränderte Konformation gezwungen wird. Durch den Austausch der rigiden Flo5A- gegen eine flexiblere Flo10A-eigene Subdomäne erlangt Loop L3 seine native Konformation und interagiert direkt mit der gebundenen Mannose.

Dies macht deutlich dass die Subdomäne einen direkten Einfluss auf die globale Form der Bindungstasche hat. In Flo10A ist durch die flexible Subdomäne der Loop L3 nicht in seiner Orientierung beeinflusst und nimmt seine deckelförmige Orientierung an. Dies korrespondiert darüber hinaus mit der Beobachtung in Cea1A, welches keine Subdomäne besitzt.

Im Sequenzalignment von Flo5A mit Flo10A ist ebenfalls zu erkennen, dass die Zusammensetzung von L3 unterschiedlich ist. NINA WOZNIAK konnte zeigen, dass Flo5A sowie Flo10A ohne Subdomäne nicht funktional sind¹⁵⁹. Möglicherweise ändert hierbei die Bindungstasche zu stark ihre Konformation. Dies erklärt ebenfalls, warum die Chimäre Flo5A^{10SD} in ihren Flokkulationseigenschaften stark eingeschränkt ist. Die Bindungstasche ist degeneriert, da die Flo10A-Subdomäne zu flexibel ist um die Flo5A-Bindungstasche stabil zu halten. Stattdessen zeigt dieses Adhäsins eine Bindung an Agar. Dies lässt in der Summe folgenden Schluss zu: Die Flo10A-Subdomäne ist nicht direkt in die Bindung von Glycanen involviert, stattdessen hat Flo10A die Mannosebindung über einen modifizierten Loop L3 mit direkter Interaktion von K216 erreicht und nutzt die Subdomäne für eine weitere Funktion. Die Flo5A-Subdomäne dagegen ist direkt in die Glycanerkennung involviert, da die Orientierung von L3 in Flo5A ohne die Subdomäne nicht zur Mannosebindung und damit zur Flokkulation ausreicht.

5 Diskussion

Hefen haben sich im Laufe der Evolution an ein großes Spektrum an Umweltbedingungen anpassen können und sind heute fast ubiquitär vorhanden. Zur Interaktion mit Zellen der eigenen Gattung oder mit dem sie umgebenden Ökosystem haben sie eine Palette zellwandverankerter Adhäsionsproteine entwickelt⁹⁹. Ein großer Teil dieser Adhäsine basiert auf der strukturell konservierten PA14-Domäne aus *B. anthracis*. Diese Ca^{2+} -abhängigen GPI-verankerten Zellwandproteine vermitteln die Bindung an diverse Glycanstrukturen und sind damit für einige wichtige Prozesse im Lebenszyklus von Hefen verantwortlich. Neben Flokkulinen, welche in der Bäckerhefe *S. cerevisiae* die vegetative Adhäsion und Verklumpung von Zellen der eigenen Gattung vermitteln, sind als zweite Gruppe die epithelialen Adhäsine aus *C. glabrata*, welche unter Bindung physiologisch relevanter Galactoside als Pathogenitätsfaktoren wirken, gut charakterisiert^{108,324}. Obwohl beide Adhäsingruppen eine hohe strukturelle Ähnlichkeit aufweisen und Glycane binden, finden sie in unabhängigen Prozessen Anwendung.

Hefen sind erdgeschichtlich sehr alt und bilden seit Millionen Jahren einen unverzichtbaren Baustein in der Verwertung und dem „Recycling“ von komplexen biologischen Materialien⁷. Adhäsine mit PA14-Domänen sind an grundlegenden, aber mitunter spezies- bzw. habitatabhängigen Prozessen beteiligt. Dies legt den Schluss nahe, dass deren Anwendung ein in Hefen häufig vorkommendes Prinzip darstellt. Aufgrund der hohen genetischen Varianz, der diese Adhäsine bspw. aufgrund ihrer Lokalisation in subtelomeren Bereichen unterliegen, bilden diese ein breites, größtenteils unbekanntes Spektrum an Funktionalitäten ab³²⁵. Dies erlaubt den Hefen die stetige Anpassung an veränderliche Umweltbedingungen.

Im Rahmen dieser Arbeit wurde ein tieferes Verständnis für die Funktionalitäten, aber auch für die globalen Determinanten der Bindungsspezifität dieser diversen Adhäsine-Gruppe geschaffen. So konnte mit *Cea1A* und *PpFlo1A* ein neuer Subtyp identifiziert und biochemisch charakterisiert werden. Während *Cea1A* die selektive Bindung an *N*-Acetylglucosamin vermittelt, ist für *PpFlo1A* bisher kein Ligand bekannt. Das Studium weiterer Familienmitglieder der beiden bekannten Subtypen der Flokkuline und *Epas* lieferte ein tieferes Verständnis der speziesinternen Variation auf Struktur- und Funktionsebene. Diese bildet die Grundlage für die stetige Entwicklung und Anpassung neuartiger oder verbesserter Varianten in Abhängigkeit der Umweltbedingungen. Aufgrund dieser Eigenschaften verglichen Christiaen *et al.* die pilzlichen Adhäsine mit der Haupthistokompatibilitätskomplex-Familie (MHC) des Immunsystems¹¹².

Aufgrund dieser maßgeblichen Vielfalt an Funktionalitäten bilden die Adhäsine mit PA14-Domänen auch das Potential für die Anwendung in Biotechnologie und Pharmazie bspw. zur Bindung physiologisch relevanter Glycokonjugate.

5.1 Globale Betrachtung der pilzlichen PA14-Domänen – Genetische Varianz führt zum Erfolg

Bei zusammenfassender Betrachtung der globalen Zusammenhänge zwischen den einzelnen Subtypen der pilzlichen Adhäsine mit PA14-Domäne wird eine große Varianz auf Sequenz- und auf Strukturebene deutlich. Wie lässt sich damit eine Klassifizierung in Subtypen in Einklang bringen?

5.1.1 Interne Gemeinsamkeiten definiert die einzelnen Subtypen

Neben der Vielfalt auf Sequenzebene lassen sich Gemeinsamkeiten in den einzelnen Subtypen ausmachen (Abbildung 5.1 A). So ist dies das Vorhandensein von Subdomänen in den Flokkulinen, die Konservierung von Disulfidbrückenmustern innerhalb der Gruppen, aber auch strukturelle Determinanten, welche gruppenspezifisch sind. Hier ist die jeweilige Verlängerung unterschiedlicher flexibler Loops zu nennen. Während in den Epas vorwiegend der Bereich um L1 verlängert ist, ist dies in den Flokkulinen die Subdomäne/L2 und im *Pichia*-Subtyp L3 über der Bindungstasche. Daraus folgend zeigen die Subtypen ähnliche Spezifitäten. Eine tiefgehende Analyse der potentiellen Adhäsine ohne CC-Motiv deutet an, dass diese eine alternative Domänenorganisation besitzen. Hier ist nach der Signalsequenz direkt ein S/T-reicher Bereich zu finden, auf welchen C-Terminal die PA14-Domäne folgt; ein GPI-Anker fehlt. Dies erinnert an die ursprüngliche Definition der GLEYA-Domäne als C-terminale PA14-Domäne. In *S. pombe* sowie *Hirsutella minnesotensis*, einem carnivoren Ascomycet, konnten bereits erste Vertreter gefunden werden^{326,327}. Zwei ebenfalls im phylogenetischen Baum verzeichnete GLEYA-Proteine der Nematoden-fangenden Pilze *Drechlerella stenobrocha* (UniProt-ID: W7I935) und *Arthrobotrys oligospora* (UniProt-ID: G1X146) spielen möglicherweise eine direkte Rolle in der Erkennung und Bindung des Fadenwurms^{328,329}.

Um die Einordnung in Subtypen zu untermauern, wurde ein Sequenz-Ähnlichkeits-Netzwerk (*Sequence-Similarity-Network*, SSN) mit Hilfe des *Enzyme Function Initiative-Enzyme Similarity Tools* (EFI-EST) erstellt und mit Cytoscape visualisiert³³⁰⁻³³². Solch ein Netzwerk gibt die Ähnlichkeit vorgegebener Sequenzen grafisch wieder. Hierfür wurden 5242 in der *Interpro*-Datenbank gelistete Sequenzen mit PA14-Domäne verwendet (Interpro Familie IPR011658). Sequenzen, welche eine Identität von $\geq 95\%$ aufweisen, sind zu einem Knoten zusammengefasst^{***}. Die Knoten fallen hierbei zu wohldefinierten Gruppen zusammen (Abbildung 5.1 B). Neben bakteriellen Enzymen und Toxinen (blau), findet man Gruppen, welche sekretierten, bakteriellen Proteinen (grün) oder der Fibroschleimdomäne (gelb) zuzuordnen sind. Pilzliche Adhäsine (rot), sowie solche ohne CC-Motiv (violett) finden sich vor allem in kleineren aber dennoch gut abgegrenzten Gruppen.

*** Daraus folgen 3086 Knoten, welche 5228 Sequenzen repräsentieren.

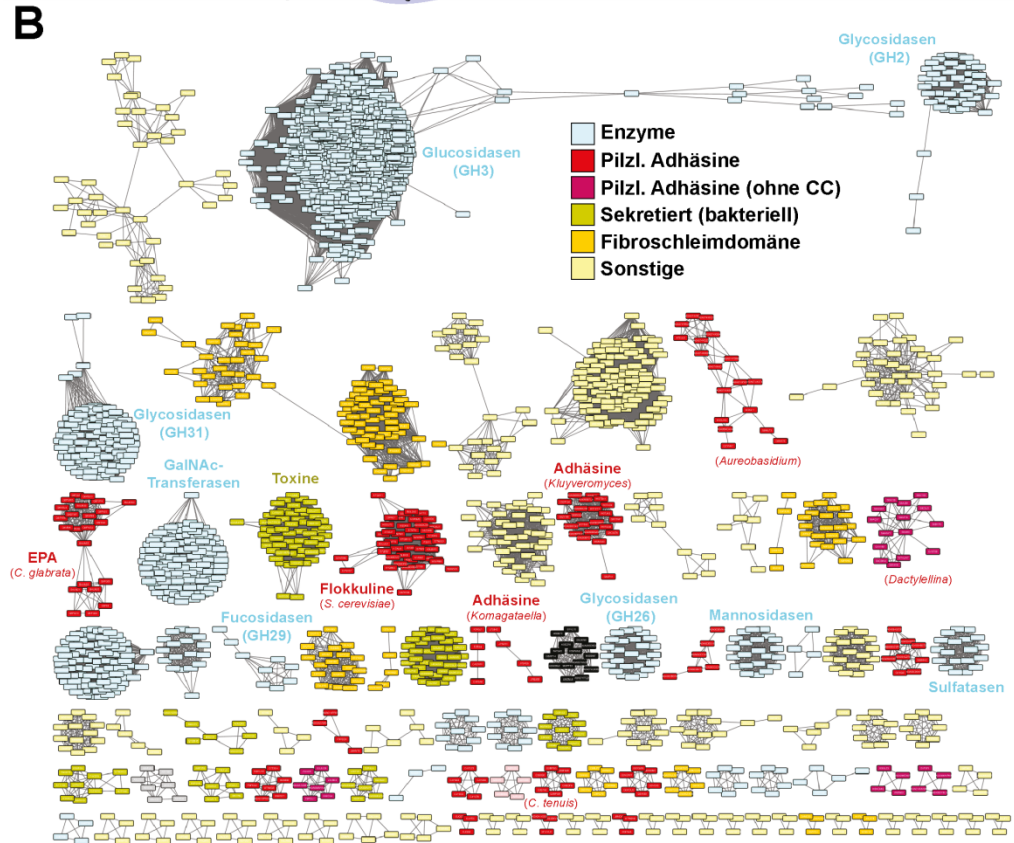
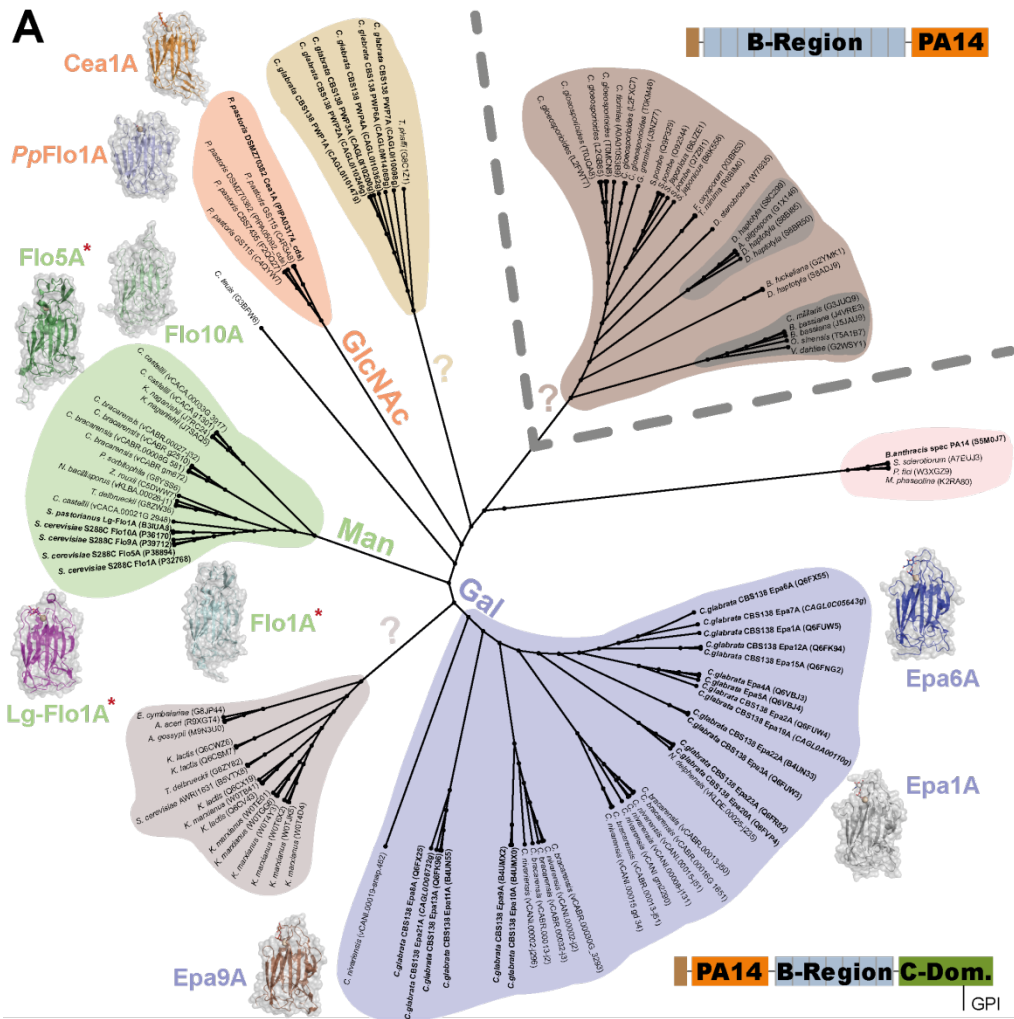


Abbildung 5.1 – Globale Betrachtung der pilzlichen Adhäsine mit PA14-Domäne. A: Der Einordnung der im phylogenetischen Baum abgebildeten Adhäsine in Subtypen liegen funktionelle (Art des erkannten Liganden) und strukturelle Eigenschaften (konservierte Disulfidbrücken, Insertionen in unterschiedliche Schleifenbereiche) zu Grunde. Die Sequenzen zeigen darüber hinaus zwei verschiedene Domänenorganisationen. Während der Großteil der Adhäsine dem bekannten Muster mit *N*-terminaler PA14-Domäne folgt, tragen hauptsächlich die Sequenzen ohne CC-Motiv diese Domäne am *C*-Terminus. Dies entspricht der ursprünglichen GLEYA-Domäne. Im Rahmen dieser Arbeit konnte das funktionelle und strukturelle Bild der *Pichia*-Adhäsine (orange), der Epas (blau) und der Flokkuline (grün) weiter komplettiert werden. Die *Kluyveromyces*-Adhäsine (grau), die PWP's (gelb) sowie die Adhäsine ohne CC-Motiv (violett) sind weitgehend unerforscht. (Adhäsine, welche nicht Gegenstand dieser Arbeit waren, sind mit einem roten Stern markiert.) B: Eine *Sequence-Similarity-Network* Analyse der PA14-Familie aus allen Domänen des Lebens untermauert die Einteilung der pilzlichen Adhäsine in Subtypen. Die dargestellten Cluster sind mit denen der phylogenetischen Analyse in Größe und Diversität vergleichbar.

Die Subtypen, welche im Rahmen dieser Arbeit identifiziert werden konnten, finden sich ebenfalls in dieser Betrachtung wieder. So bilden Flokkuline, epitheliale Adhäsine und *Pichia*-Adhäsine sowie *Kluyveromyces*-Adhäsine und solche ohne CC-Motiv jeweils einen *Cluster* aus. Darüber hinaus zeigen die hier dargestellten Gruppen eine ähnliche interne Streuung wie im phylogenetischen Baum; bspw. ist der Subtyp der Epas breiter als der der Flokkuline.

5.1.2 Diversität innerhalb der Subtypen wird durch Anpassung erzeugt

Die Stärke der Varianz in den einzelnen Subtypen ist möglicherweise mit dem auf den Organismus wirkenden Selektionsdruck zu erklären; vor allem vor dem Hintergrund der essentiellen Funktionen, welche die Adhäsine ausüben. Die epithelialen Adhäsine dienen der Adhäsion an die Epithelzellen eines Wirtsorganismus und sind unverzichtbarer Bestandteil bei der Etablierung der Pathogenität^{137,333}. Da *Candida glabrata* mit der Mikroheterogenität der Glycosylierung, auch im Säugetierkörper, konfrontiert ist³³⁴⁻³³⁶, muss dieser Organismus eine Bandbreite anpassbarer aber effektiver Adhäsionsproteine auf seiner Oberfläche präsentieren, die epithelialen Adhäsine. Dies führt zu einer Verbreiterung der phylogenetischen Gruppe. Im Gegensatz dazu sieht sich *Saccharomyces cerevisiae* einem nur sehr begrenzten Selektionsdruck ausgesetzt, zumindest was die Bildung von multizellulären Aggregaten, d.h. Biofilmen und Flocken, angeht. Es gilt nur eine Art Glycan zu erkennen, Oligomannoside mit Man α 1-2Man-Enden. Dies führt zu der Ähnlichkeit der Flokkulinsequenzen untereinander und damit zu kurzen Ästen im phylogenetischen Baum sowie einer kleineren Gruppe im SSN.

Solche Mechanismen der molekularen und genetischen Anpassung^{337,338} findet man in der Natur häufiger^{335,336}, so bspw. bei Proteinen der Ig-Familie³³⁹, den VSG bei *Trypanosomas*³⁴⁰ oder bei Molekülen des Haupthistokompatibilitätskomplexes (MHC)³⁴¹, mit welchen CHRISTIAENS *et al.* bereits die Subtypen der Flokkuline und epithelialen Adhäsine verglichen¹¹². Die Varianz wird hauptsächlich durch die bereits in der Einleitung erwähnte subtelomere Lage der Adhäsingene und die damit einhergehende Rekombination verursacht. *Slippage* der DNA-Polymerase bei der Transkription führt bspw. ebenfalls zur Variation der Länge der B-Region¹⁰⁰. Dieses „verrutschen“ beim Ablesen des parenteralen Stranges wird durch die hohe Anzahl repetitiver Einheiten verursacht.

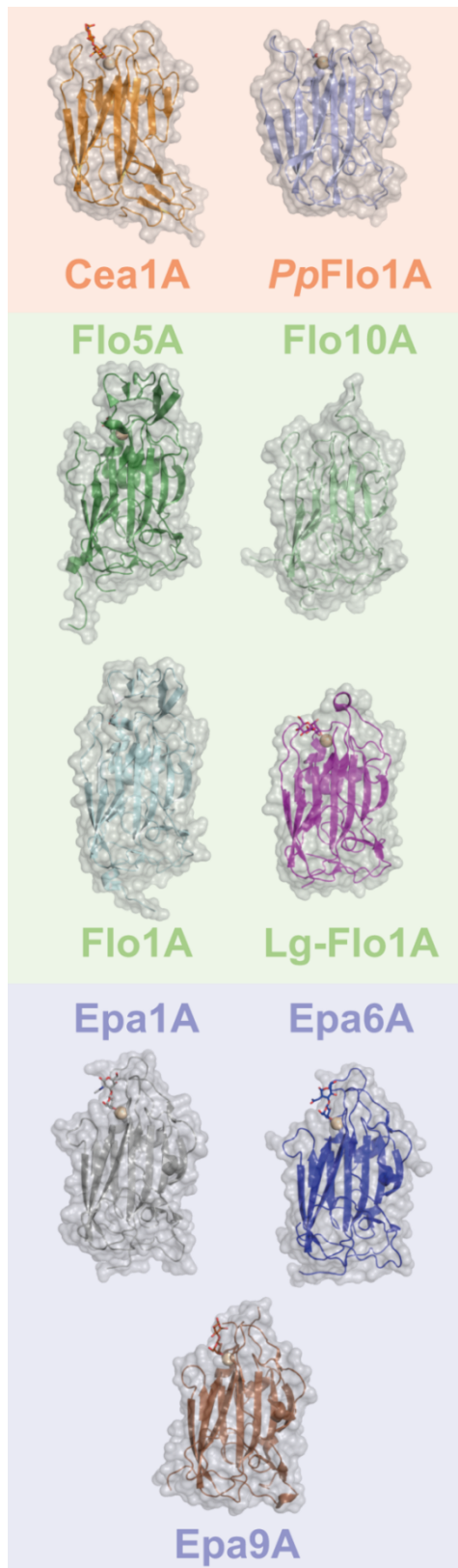


Abbildung 5.2 – Die pilzlichen Adhäsine mit PA14-Domäne sind unabhängig von ihrer Subtyp-Zugehörigkeit morphologisch divers. Orange: *Pichia*-Subtyp, grün: Flokkulin-Subtyp, blau: *Epa*-Subtyp.

Die Rekombination fungiert als eine Art Baukastenprinzip zur ständigen Anpassung an die Umwelt¹⁰⁰. Mutationen und neue Kombinationen erweisen sich als vorteilhaft und werden angewandt oder werden zu funktionslosen Genen oder Genfragmenten, wie dies möglicherweise bei *PpFlo1A* der Fall ist. Das Vorhandensein solcher Genfragmente konnte ebenfalls für die *VSG*'s bei *Trypanosoma* gezeigt werden. Derselbe Effekt ist anzunehmen, wenn sich durch Wechsel des Habitats über mehrere Generationen hinaus ein Adhäsins nicht mehr als positiv für die *fitness* erweist.

Diese Zusammenhänge machen deutlich, dass es sehr schwer ist, die pilzlichen Adhäsine im Ganzen zu betrachten. Die Komplexität und Eigenständigkeit der einzelnen Subtypen verdeutlicht dies bspw. anhand der unterschiedlichen Bindungsmodi; auch innerhalb der einzelnen Subtypen. Diese Vielfalt zeigt sich bereits anhand der diversen strukturellen Morphologien (Abbildung 5.2). Hinzu kommen die dadurch verursachten Unterschiede in den chemischen und physikalischen Eigenschaften der Oberfläche. Auch die Flokkuline, deren Variabilität im Bereich der PA14-Domäne kleiner ist als die der *Epas*, zeigen durch unterschiedliche Subdomänen stark variierende Formen.

Aufgrund dieser Diversität ergibt sich eine Vielzahl von chemischen und strukturellen Determinanten, welche die Funktion der Subtypen und der einzelnen Mitglieder beschreiben. Daher ist es kaum möglich, rational die Funktion eines Adhäsins bloß anhand seiner Primärsequenz vorherzusagen, bzw. ob es sich definitiv um ein Adhäsins mit PA14-Domäne handelt. Dies macht es umso schwieriger, wie im Falle der *Epas* gezeigt werden konnte, Funktion und phylogenetische Nähe in Korrelation zu bringen. Auf der einen Seite kann die Verwandtschaft bzw. die Herkunft der einzelnen Adhäsine nicht von deren

Funktion entkoppelt betrachtet werden, doch zeigt sich ein überaus großes Spektrum von Feinabstimmungen innerhalb der Gruppen. Dies lässt auf der anderen Seite Raum für weitere Entdeckungen, bspw. in den in dieser Arbeit identifizierten Subtypen ohne CC-Motiv oder im Subtyp der *Kluyveromyces*-Adhäsine mit z.T. voluminösen Resten in der vorhergesagten Bindungstasche.

Die Varianz und die Rekombination beleuchten allerdings eine Schwäche der phylogenetischen Analyse. Aufgrund der Rekombination sind die Abstände und Lage der Adhäsine im phylogenetischen Baum, evolutionär betrachtet, verfälscht¹¹². Nichtsdestotrotz gibt der Baum die Ähnlichkeit und Identität der Primärsequenzen wieder und erlaubt somit eine globale Gegenüberstellung. Dies erklärt, warum eine funktionelle Betrachtung mitunter losgelöst von der Verwandtschaft geschehen muss.

Zusammengefasst untermauert und validiert die globale Betrachtung: (i) Die Einteilung in Subtypen sowie (ii) die Existenz eines *Pichia*-Subtyps, (iii) die Konsistenz innerhalb der Subtypen aber auch (iv) Varianz und Anpassungsfähigkeit als generelles Prinzip

Der Erfolg der Mitglieder des Reiches Fungi in der Erdgeschichte ist durchaus mit deren Anpassungsfähigkeit, welche sich auch auf der Ebene der Adhäsionsproteine wiederfindet zu erklären. Dies wird Hefen, neben anderen Mikroorganismen, vermutlich weiterhin zu einem erfolgreichen Teil der eukaryotischen Domäne des Lebens machen^{§§§§}.

5.2 Adhäsine aus *Pichia pastoris* bilden einen eigenständigen Subtyp

Die phylogenetische Analyse einer großen Zahl von etwaigen Adhäsinen mit PA14-Domäne deutete bereits an, dass die untersuchten Hefegattungen eigens an ihr Ökosystem und ihre Lebensbedingungen angepasste Varianten von Adhäsinen entwickelt haben (Abschnitt 4.1). Aufgrund der geringen Sequenzidentitäten im Bereich von max. 20-25% lässt sich allerdings kaum ein Rückschluss auf mögliche Funktionen oder Bindungseigenschaften ziehen. Dies wird durch die Erkenntnisse dieser Arbeit gestärkt. Die Charakteristika der Erkennung von Liganden sind keineswegs nur anhand der Primärsequenz von CBL2 abzulesen, da die Form der Bindungstasche durch eine große Anzahl von Faktoren beeinflusst wird. So spielt bspw. die Länge und Orientierung der hypervariablen Loops eine tragende Rolle. Durch die Beeinflussung dieser Loops untereinander, wie bei Flo10A durch eine alternative Subdomäne gezeigt werden konnte, wird dieser Effekt weiter verstärkt. Die geringe Sequenzidentität führt darüber hinaus dazu, dass selbst die Eigenschaften der einzelnen Loops nur schwer direkt aus der Primärsequenz abzulesen sind, da gerade diese hypervariablen Bereiche unter dem Druck ständiger Anpassung zu stehen scheinen.

^{§§§§} Oder frei nach DAVID MOORE: „Don't turn your back on a fungus!“
(aus MOORE, D. *Slayers, Saviors, Servants, and Sex: An Exposé of Kingdom Fungi*. (Springer New York, 2001))

Um einen Eindruck von der Komplexität und ein tieferes Verständnis der strukturellen Eigenschaften dieser komplexen Proteinfamilie zu erhalten, wurden Cea1A und *PpFlo1A* als erste Beispiele eines weiteren Subtyps untersucht.

5.2.1 Der Domänenaufbau von Cea1A

Die Identifikation von Cea1A durch ein multiples Sequenzalignment¹⁵⁶ und die nachfolgende phylogenetische Analyse lieferte die Primärsequenz der A-Domäne von Cea1A. Die Sequenz der B- und C-Domäne hingegen blieb bisher aufgrund der unvollständigen Sequenzierung unbekannt. In einer kürzlich erschienen Veröffentlichung³⁴² eines *in silico* Screenings nach GPI-verankerten Zellwandproteinen in *P. pastoris* GS115 wurde ein putatives Adhäsins identifiziert, welches sich nach eingehender Betrachtung als Homolog von Cea1A mit einer Sequenzidentität der A-Domäne von etwa 89% herausstellte (UniPort-ID: C4R3A8, Abbildung 5.3). Für dieses Gen ist dagegen die Gesamtsequenz verfügbar. Dem 21 Aminosäuren langen Signalpeptid folgen die A-Domäne mit 219 Resten und die B-Region. Ein Vergleich der vorhandenen Repeats mit RADAR³⁴³ führte zur Identifikation zwei unterschiedlicher Typen (8x32 aa gefolgt von 3x18 aa). Es verbleiben 80 Reste für die C-Domäne und den mittels PredGPI identifizierten GPI-Anker³⁴⁴. Im Vergleich mit den epithelialen Adhäsinen und den Flokkulinen fällt auf, dass die Länge der B-Region stark variiert. In Cea1A beträgt diese etwa 330 aa während Epa6 ~360 aa und Epa9A ~1000 aa in der B-Region zeigen.

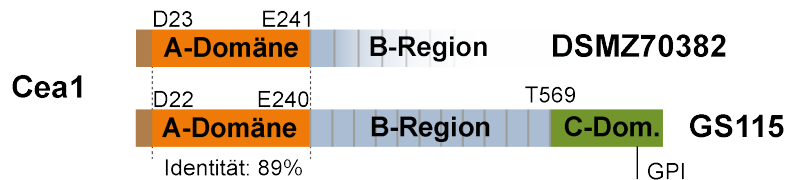


Abbildung 5.3 – Komplettierung des Domänenaufbaus von Cea1A anhand eines homologen Adhäsins aus *P. pastoris* GS115 mit einer Identität von 89% bezogen auf die A-Domäne. Dem Signalpeptid (braun) folgt die A-Domäne (orange), die repetitive B-Region (blau) sowie die zur Fixierung in der Zellwand zuständige C-Domäne (grün).

Der Einfluss der B-Region auf die Bindung wurde in der Literatur bereits häufig diskutiert. Die generelle Auffassung ist hierbei, dass die B-Region unter Ca^{2+} -Bindung durch deren *O*-Glycosylierungen eine stäbchenförmige Konformation einnimmt um die A-Domäne außerhalb der Glycokalyx, der markanten Glycanschicht der Zellen, zu präsentieren¹⁰¹. Auch amyloidbildende Sequenzbereiche zur Verhinderung von *cis*-Bindung und zur Erhöhung der Avidität sind vorgeschlagen worden^{345,346}. Untersuchungen mittels Einzelzellkraftspektroskopie (*single-cell-force-spectroscopy*, SCFS) und *in silico-modelling* am Beispiel vergleichbarer Repeats aus Als-Adhäsinen, welche zu Adhäsinen mit Ig-ähnlicher Faltung zählen¹⁰⁹, konnten erste Einblicke in ihre Struktur vermitteln³⁴⁷. Die Repeats sollen hiernach aus einzelnen, stark glycosylierten, β -Faltblatt-reichen Domänen bestehen, welche sich in der SCFS einzeln entfalten³⁴⁸.

Die Strukturlösung von Cea1A bestätigte die Annahme, dass es sich hierbei um ein strukturell mit Flo5A und Epa1A verwandtes Adhäsins handelt und es das für diese Adhäsins-Familie typische *DcisD*-Motiv besitzt (Abschnitt 4.2.4). Auch hier wird die Bindung von Glycanen über ein Ca^{2+} -Ion vermittelt. Die Selektivität wird durch eine komplexe Anordnung hypervariabler Loops gesteuert.

5.2.2 Cea1A zeigt strukturelle Alleinstellungsmerkmale

Die Eigenständigkeit der Adhäsine aus *P. pastoris* zeigt sich funktionell wie strukturell. Dies wird vor allem am Übergang der A-Domäne zur B-Region deutlich, welcher in Cea1A aus einer stabförmigen Verlängerung gebildet wird und hierbei die B-Region voraussichtlich mit einem Winkel von etwa 55° zur A-Domäne abwinkelt (Abschnitt 4.2.4). Mit Hilfe des *ab initio* Servers QUARK konnte das Modell eines Repeats von Cea1A erstellt werden und mit Hilfe des GLYCAM-Servers *in silico* glycosyliert werden^{****}. Hierbei erkennt der Server potentielle *O*-Glycosylierungen, addiert einen benutzerdefinierten Glycan-Baum und führt eine Energieminimierung über AMBER mit Hilfe des GLYCAM-Kraftfeldes aus³⁴⁹⁻³⁵¹ (Abbildung 5.4 A). In diesem Fall wurde die in Hefen häufig vorkommende Glycosylierung $\text{Man}_8\text{GlcNAc}_2$ an drei Threoninen und einem Serin gewählt. Die voluminöse Struktur und die Komplexität der gesamten B-Region lassen sich hierdurch erahnen. Das Modell ähnelt der von FRANK *et al.* vorgeschlagenen Faltblatt-Struktur, welche bei der Modellierung eines Als-Adhäsins Repeats erhalten wurde³⁴⁷. Verbindet man die kristallografisch erhaltene Struktur der A-Domäne mit dem Repeat und führt eine Energieminimierung mit AMBER durch, erhält man einen Eindruck des Übergangs zur B-Region. Das Modell zeigt hierbei kein Abspreizen von 55° an, doch handelt es sich dabei bestenfalls um eine Annäherung der *in situ* Situation. Hierbei wird eher ein henkeförmiger Übergang zur B-Region erhalten (Abbildung 5.4 B). LIPKE *et al.* & RAMSOOK *et al.* berichteten von der Möglichkeit, dass Adhäsine über ihre B-Region Di- oder Oligomere bilden könnten, welche die Avidität der Wechselwirkung weiter erhöhen, indem die Glycan-bindenden Zentren in räumliche Nähe gebracht werden^{103,104,352}. Zwei der zuvor gebildeten Fusionsproteine aus der A-Domäne und eines Repeats lassen sich beispielsweise zu einem Dimer mit zweizähliger Symmetrie zusammenführen. Die Tatsache, dass in der Kristallstruktur der Kontakt zwischen den 4 mol/asu ebenfalls über eine Verlängerung des β -Sandwichs herbeigeführt wird, bestärkt die Möglichkeit einer solchen Dimer- oder Oligomerisierung (Abbildung 5.4 C).

****Aufgrund der fehlenden Sequenz der B-Region in Cea1A wurde der erste Repeat (Uniprot-ID:C4R3A8, Reste: E277-T309) des Homologs aus *P. pastoris* GS115 verwendet.

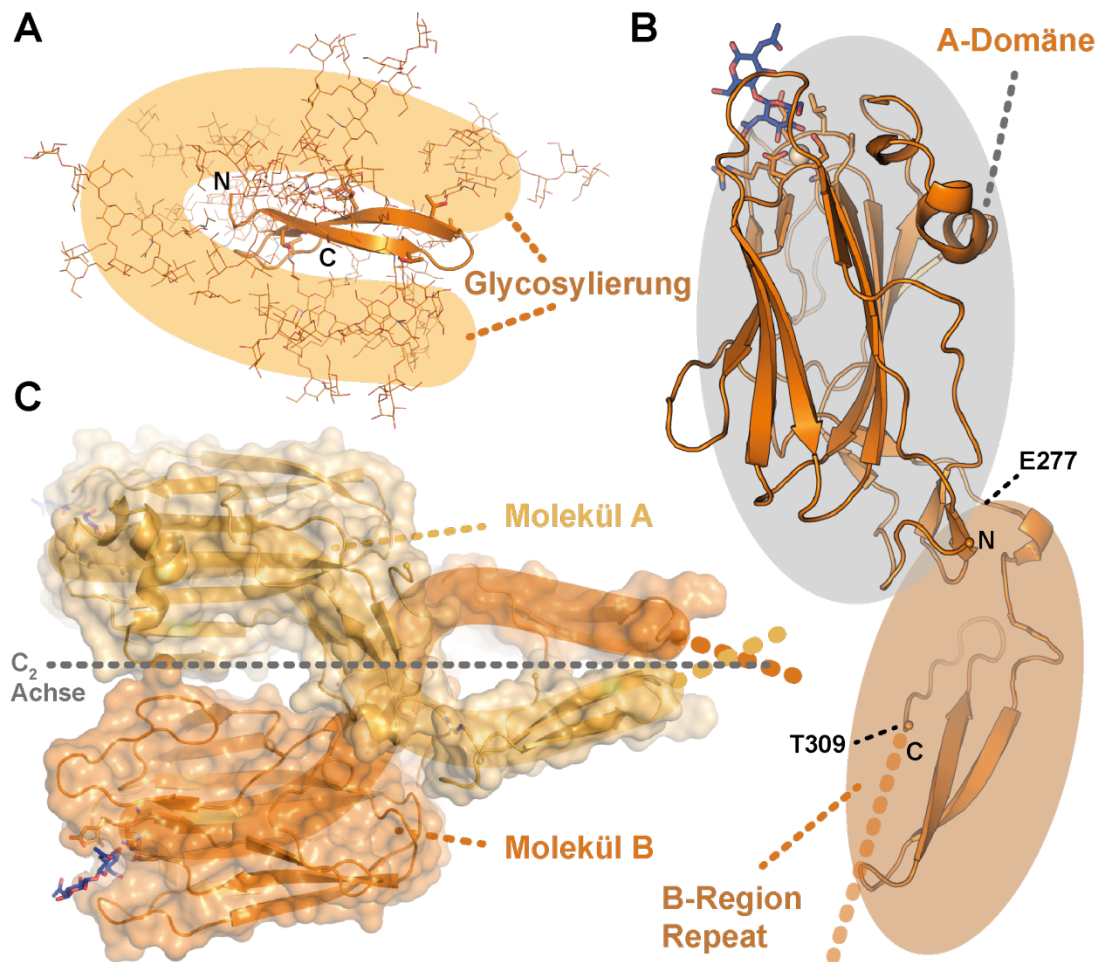


Abbildung 5.4 – Modellierung des Übergangs zur B-Region. A: Das mit Hilfe des QUARK- und des GLYCAM-Servers erstellte Modell eines Repeats aus der B-Region mit $\text{Man}_8\text{GlcNAc}_2$ -Glycosylierung zeigt einen β -Faltblatt-Verlauf, welcher bereits für die B-Region der Als-Adhäsine vorhergesagt wurde. B: Die Fusion des Modells mit der Kristallstruktur der A-Domäne und die Energieminimierung mittels AMBER deuten einen henkelförmigen Übergang zur B-Region an. C: Der henkelförmige Übergang und die Orientierung der B-Region könnte, in Verbindung mit relevanter Literatur, auf eine Oligo- bzw. Dimerisierung von Cea1A bzw. weiterer Adhäsinen auf der Zellwand hindeuten. Hier ist eine zweizählige Symmetrie angedeutet.

Die alternative Faltung im Bereich der Termini ist ein Alleinstellungsmerkmal unter den bekannten Strukturen. Dementsprechend zeigt die Primärsequenz ebenfalls eine starke Abweichung in diesem Bereich (Abbildung 5.5 A). Führt man mit der *N*-terminalen Sequenz des Stielbereichs eine Suche mittels *BLASTp* durch, so findet man ausschließlich die in der phylogenetischen Analyse genutzten putativen Orthologe von Cea1A (*E-value* von etwa 10^{-14})¹⁷². Eine Übereinstimmung mit über 90% ist darüber hinaus nur noch in *bap* (UniProt-ID: Q4ZHU4), einem Biofilm-assoziierten Protein aus *Staphylococcus*, zu finden^{353,354}.

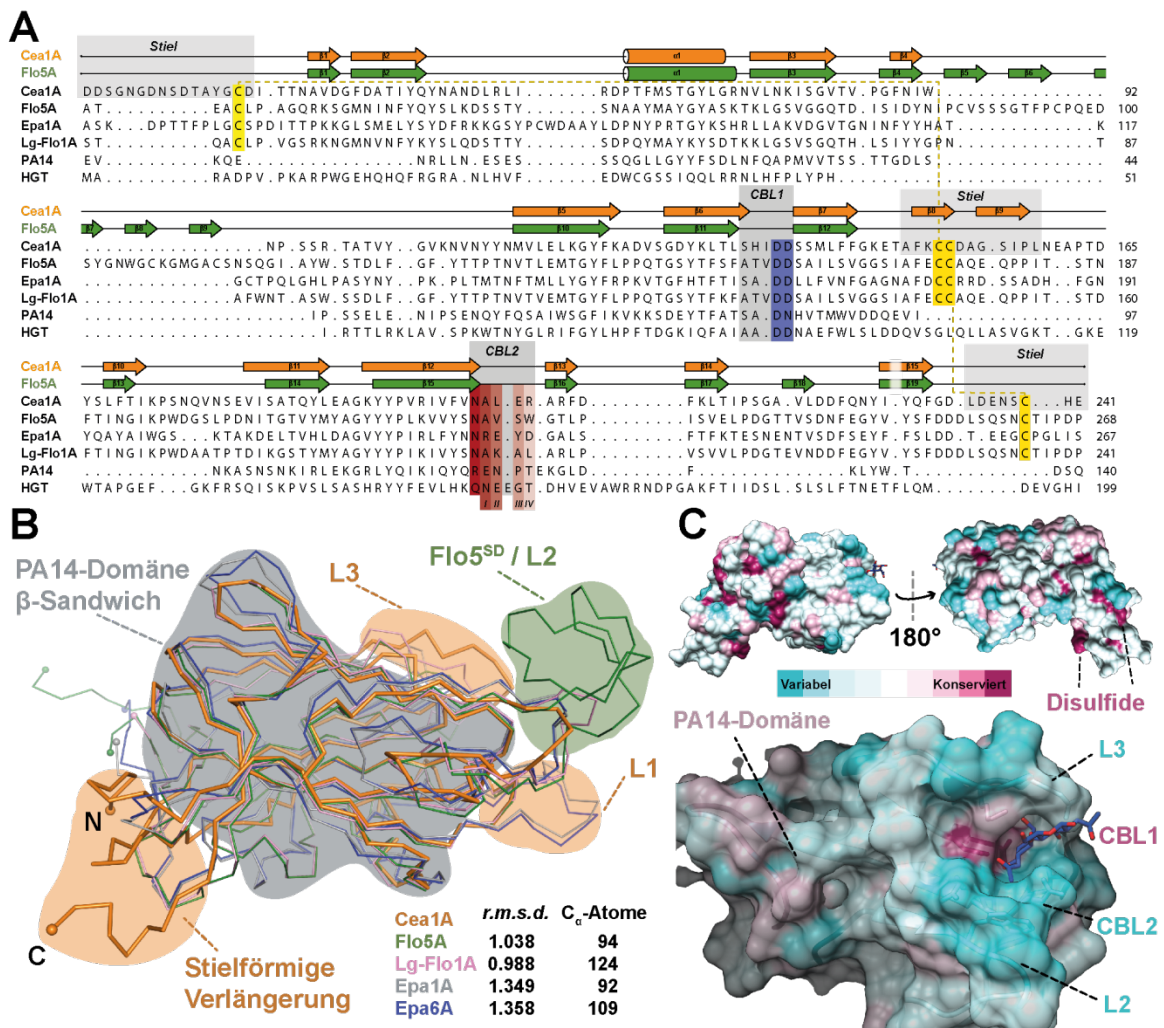


Abbildung 5.5 – Vergleich der strukturellen Eigenschaften repräsentativer Adhäsine mit PA14-Domäne. A: Ein multiples Sequenzalignment mit T-COFFEE beleuchtet Unterschiede auf Primärsequenzebene. Ein Vergleich mit den Sekundärstrukturelementen von Flo5A unterstreicht die alternative Faltung im stielförmigen Bereich, welche in Cea1A hauptsächlich aus β8 und β9 in Verbindung mit den beiden Termini gebildet wird. Das hoch konservierte Dα3D-Motiv (grau/blau) ist ebenso wie der für die Glycanbindung wichtige Bereich des CBL2 (grau/rot) hervorgehoben. (UniProt-IDs: Flo5A, P38894; Epa1A, Q6VBJ0; LgFlo1A, B3IUA8; PA-14, P13423; HGT, Q8N9V0) B: Überlagerung der Rückgrate von PA14-Adhäsinstrukturen. Während die PA14-Domäne als Kern der Strukturen sehr rigide ist (grau hinterlegt), zeigen sich Abweichungen besonders im Bereich der hypervariablen Loops L1-L3 (orange und grün) sowie im Bereich der Termini (orange). C: Die Abbildung des Konservierungsgrades auf Grundlage des T-COFFEE Alignments auf die Strukturfläche von Cea1A zeigt, dass die strukturelle Konservierung nicht unbedingt auf Sequenzebene wiederzuerkennen ist. Über die phylogenetischen Gruppen hinweg sind nur Reste, welche für die strukturelle Integrität und die Glycan- und Calciumbindung essentiell sind, erhalten.

Eine Überlagerung der Rückgrate repräsentativer Strukturen unterstreicht diese Einzigartigkeit (Abbildung 5.5 B). Während der hypervariable Loop L1 eine mit Flo1A, Flo5A und Lg-Flo1A vergleichbare Konformation einnimmt, ist in L2 das Fehlen einer Subdomäne zu erkennen; der Loop von Cea1A ist in allen Strukturen der Kürzeste^{107,108,135}. Am ehesten ist dieser mit dem in *PpFlo1A* beobachteten Loop vergleichbar. Trotz Fehlens der Subdomäne sind beide Loops nicht wie bei den epithelialen Adhäsinen aus *C. glabrata* mit einer zusätzlichen Disulfidbrücke verbunden. Diese wird durch die Verkürzung beider Loops und die Fixierung von L1 mit einem hydrophoben Bereich funktionell ersetzt. Der entsprechend verlängerte

Loop L3, welcher über der Bindungstasche eine Art Deckel bildet, kompensiert die Öffnung der Tasche.

Bildet man den Konservierungsgrad der einzelnen Reste mittels des Alignment auf die 3D-Struktur ab, so zeigt sich eine große Varianz (Abbildung 5.5 C). Die Reste, welche dem Lösungsmittel zugänglich sind, zeigen eine intermediäre Variabilität, während sich die Bindungstasche als hypervariabel erweist. Auch der Bereich, welcher das PA14-Kernmotiv bildet, scheint auf Sequenzebene eher variabel zu sein. Die strukturelle Rigidität und Integrität ohne global konservierte Sequenzbereiche ist ein Kennzeichen weiterer β -Sandwich Domänen wie bspw. Ig-artigen Proteinen³⁵⁵, zu welchen ebenfalls die Adhäsine vom Flo11-Typ zählen¹¹⁰. Während der PA14-Kern eine rigide strukturelle Grundlage zur Glycanbindung bildet, werden die weiteren Bindungseigenschaften durch die intrinsische Variabilität auf Sequenz- und Strukturebene der Loops erreicht, welche die Bindungstasche bilden.

5.2.3 Energetische Betrachtung des DcisD-Motivs

Wie in Abschnitt 4.2.4 beschrieben, wird das DcisD-Motiv in Cea1A, im Vergleich zu Flo5A, nicht durch ein Tyrosin des β -Sandwichs stabilisiert¹⁰⁷. Das an dieser Stelle liegende Phenylalanin ist nicht in der Lage einen Einfluss auf die *cis*-Peptidbindung zu nehmen und entspricht der für Flo5A untersuchten Mutante Y222F, welche ebenfalls zeigte, dass diese Interaktion zumindest nicht essentiell für die Ausbildung der *cis*-Peptidbindung und C-Typ Lektinfunktion ist¹⁰⁷.

Wie ist die Bildung des DcisD-Motivs energetisch zu erklären? *Cis*-Peptidbindungen kommen nur selten nach anderen Aminosäuren als Prolin vor (0.03%)³⁵⁶; zumeist in der Nähe aktiver Zentren von Enzymen, wo Metallionen koordiniert werden³⁵⁷⁻³⁵⁹. Die Aktivierungsenergie, welche für die *cis/trans*-Isomerisierung aufgewendet werden muss, ist mit etwa 20 kcal/mol sehr hoch³⁵⁹⁻³⁶¹. Dies erklärt (*i*) das seltene Auftreten dieser Bindungsform sowie (*ii*) die Isomerisierung als geschwindigkeitsbestimmenden Schritt bei der Faltung von Proteinen mit nicht-Prolin-*cis*-Peptidbindungen aufgrund der kinetischen Barriere bei physiologischer Temperatur^{359,362,363}. Zur Beschleunigung dieses Prozesses haben sich *cis/trans*-Isomerasen entwickelt, welche bisher allerdings nur für Prolyl-*cis*-Peptidbindungen bekannt sind³⁶⁴. Die Energie für die Isomerisierung muss demnach bei der Faltung aufgewandt werden. Möglicherweise lässt sich die Bildung des DcisD-Motivs durch den Energiegewinn erklären, welcher durch Ausbildung des sehr rigiden β -Sandwiches entsteht. Die Energie wird hier (*i*) durch Bildung der beiden großen Faltblätter sowie (*ii*) deren Zusammenlagerung in das Sandwich durch den hydrophoben Effekt gewonnen (Abbildung 5.6 A). Dass die Entstehung des β -Sandwiches einen entscheidenden Einfluss auf die Bildung und Stabilität der *cis*-Peptidbindung hat, wird weiterhin durch Abbildung 5.5 B verdeutlicht. Das bei aller struktureller Variabilität dieser Kernbereich in allen Strukturen nahezu unverändert vorliegt, deutet auf die Unverzichtbarkeit dieser Einheit hin.

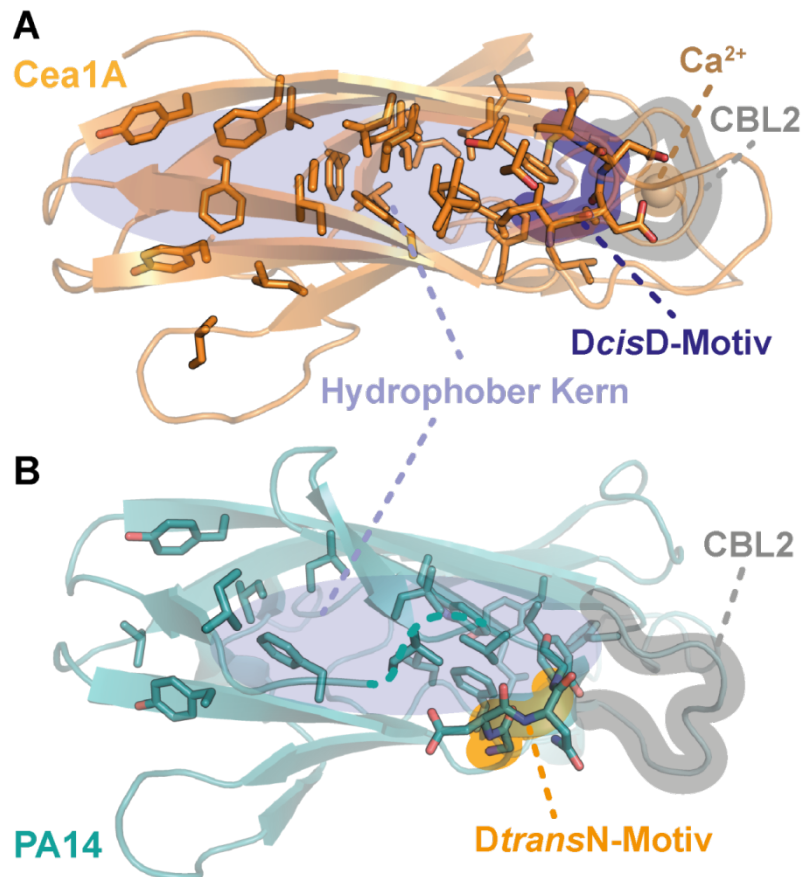


Abbildung 5.6 – Die Bildung des β -Sandwiches trägt maßgeblich zur Ausbildung der *cis*-Peptidbindung bei. A: Im Kern der PA14-Domäne von Cea1A ist eine große Anzahl hydrophober Reste zu erkennen, welche die rigide Verbindung der beiden Faltblätter verursachen. Dadurch wird die Region um das D*cis*D-Motiv (i) konformativ eingeschränkt und (ii) die nötige Energie für die Isomerisierung aufgebracht. Die Bindung des Ca^{2+} -Ions erzeugt weitere Energie sowie eine stabile Umgebung für *cis*-Peptidbindung. B: Auch die PA14-Domäne selbst zeigt einen hydrophoben Kern, welcher allerdings sehr viel weniger stark ausgeprägt ist. Die starke Verlängerung des zu CBL2 analogen Loops und das Fehlen eines rigiden Loops L3 über dem CBL1-Analogen sorgt möglicherweise für verminderte Stabilität und damit für eine vergrößerte Flexibilität in diesem Bereich, was die Bildung der favorisierten *trans*-Peptidbindung zur Folge hat.

Eine Analyse von PAL und CHAKRABARTI kommt ebenfalls zu dem Schluss, dass nicht-prolyl-*cis*-Peptidbindungen eher durch umgebende Sekundärstrukturelemente stabilisiert werden als durch direkte Interaktionen der Peptidbindung mit umliegenden Resten³⁶⁵. Dementsprechend werden diese *cis*-Bindungen fast nur in oder in der Nähe von β -Strängen oder seltener *N*-terminal an Helices gefunden. Das D*cis*D-Motiv ist nur peripher über die Aspartatseitenketten oberflächenexponiert und befindet sich ansonsten eingebettet im Sandwich. Aus diesem Grund ist die faltungsinduzierte Bildung des D*cis*D-Motivs wahrscheinlicher als eine nachfolgende, bspw. Calcium-induzierte Isomerisierung. Letztere Hypothese würde eine aufwendige strukturelle Konformationsänderung durch die sterisch anspruchsvolle Isomerisierung voraussetzen³⁶⁵. Eine konzertierte Verbindung beider Prozesse ist ebenfalls denkbar, da auch die Bindung des Ca^{2+} -Ions einen Energiegewinn bringen sollte, welcher zur Isomerisierung verwendet werden könnte. SIM und Mitarbeiter erhielten für die Komplexierung des Ca^{2+} -Ions an Lg-Flo1A ein K_D -Wert von etwa 14 nM¹³⁵. Aufgrund des hohen Konservie-

ungsgrades des Calcium-Bindungsmotivs ist für Cea1A etc. eine ähnliche Bindungsstärke zu erwarten. Dementsprechend ist davon auszugehen, dass die Ausbildung des *DcisD*-Motivs während des Reifungsprozesses (Faltung, Disulfidbrückenbildung, Ca^{2+} -Bindung, Glycosylierung etc.) beim Transport über das endoplasmatische Retikulum und den Golgi-Apparat zustande kommt. Die detaillierte Untersuchung der Bildung des *DcisD*-Motivs könnte anhand von Flo5A gelingen, da dieses in der CD-spektroskopischen Schmelzpunktbestimmung eine gewisse Tendenz zur Rückfaltung nach Denaturierung zeigte³⁶⁶.

Neben den pilzlichen Adhäsinen sind nur zwei weitere Proteinstrukturen mit einem *DcisD*-Motiv bekannt, eine Zn^{2+} -abhängige Aminopeptidase³⁶⁷ (PDB-ID: 1AMP) und die ATP-Synthase³⁶⁸ (PDB-ID: 1BMF). In beiden Proteinen wird durch dieses Motiv ein Metallion komplexiert, jedoch ist die direkte strukturelle Umgebung der konsekutiven Aspartate hauptsächlich α -helikal und schlecht vergleichbar. In der PA14-Domäne aus *B. anthracis* selbst liegt kein *DcisD*-Motiv vor³⁶⁹. Stattdessen ist hier ein Aspartat mit einem Asparagin über eine *trans*-Peptidbindung verknüpft (Abbildung 5.6 B). Das β -Sandwich besteht hier nicht nur aus kürzeren und weniger Strängen, sondern es ist breiter und nicht so dicht gepackt wie in den Adhäsinen. Dies gewährt dem zu CBL1 analogen Loop eine größere Flexibilität, welche dazu genutzt wird um eine alternative Konformation einzunehmen. Da die Abstände zwischen den beiden Faltblättern vergrößert sind, ist auf eine weitaus geringere Anzahl hydrophober Reste in der Mitte zurückzuführen. Das Vorhandensein der *trans*-Peptidbindung ist weiterhin damit zu begründen, dass Loop L3 über dem DN-Motiv nur vermindert am Sandwich beteiligt ist und darüber hinaus hoch flexibel ist, was zu einer geringeren sterischen Blockade des CBL1-Analogs führt. CBL2, welcher in den Adhäsinen das Ca^{2+} -Ion umschließt, ist in der PA14-Domäne stark verlängert und durch einen großen Loop-Bereich gekennzeichnet, welcher dem Loop mit DN-Motiv weitere Flexibilität ermöglicht. Weiterhin fehlen die in Adhäsinen konservierten Disulfidbrücken, welche hier das Sandwich weiter stabilisieren und „zusammenschnüren“. Zusammenfassend ist anzunehmen, dass die PA14-Domäne (*i*) zu viele flexible Bereiche besitzt, als das sie eine *cis*-Peptidbindung begünstigen könnte und (*ii*) der verlängerte CBL2-Loop keine Bindung eines weiterhin stabilisierend wirkenden Metallions erlaubt.

5.2.4 Der Einfluss der hypervariablen Bereiche auf die Ligandenbindung

In Cea1A wird der Einfluss der hypervariablen Loops auf die Bindung sehr deutlich. Während in Flo5A kaum von einer Bindungstasche, eher von einer breiten, dem Lösungsmittel gut zugänglichen Bindungsfläche, gesprochen werden kann, wird diese im Verlauf über Epa1A nach Cea1A immer enger und abgeschirmter (Abbildung 5.7 A-C). Dies liegt unter anderem an der Konformation der variablen Loops. Während in Flo5A die Subdomäne in L2 eine wichtige Rolle bei der Mannosebindung spielt, wird dies in Epa1A durch L3 über dem Liganden übernommen. L1 und L2, welche über eine Disulfidbrücke miteinander verbunden

sind, bilden den räumlichen Abschluss der Bindungstasche unter dem gebundenen Galactosid.

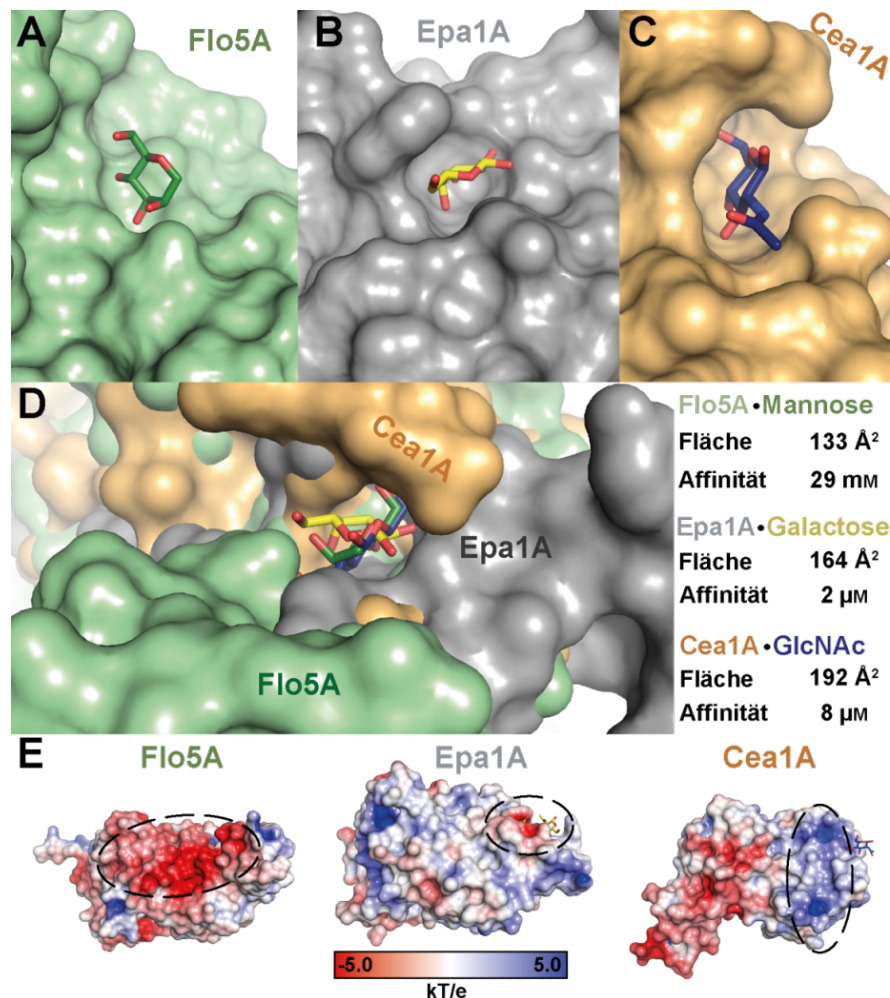


Abbildung 5.7 – Die hypervariablen Loops formen die Bindungstasche. A: Die Bindungstasche in Flo5A (PDB-ID: 2XJS) ist stark solvensexponiert und nur durch die Subdomäne partiell abgeschirmt. B: In Epa1A (PDB-ID: 4ASL) findet eine Abschirmung durch L1 und L3 statt. C: In Cea1A wird der Ligand durch einen stark verlängerten L3 fast umschlossen. D: Ein Vergleich der drei Bindungsmodi macht deutlich, dass durch in Sequenz und Konformation stark abweichende hypervariablen Loops jeweils einen anderen Bereich der Bindungstasche abschirmen. E: Der Vergleich der elektrostatischen Oberflächen von Flo5A, Epa1A und Cea1A zeigt deutliche Unterschiede. Während Flo5A stark negativ geladen ist und Epa1A eine ausgeglichene Ladungsverteilung zeigt, ist in Cea1A eine Häufung positiver Ladungen im Bereich der Bindungstasche zu finden (hervorgehobene Bereiche).

In Cea1A allerdings, welches zwar stark verkürzte L1 und L2 besitzt, liegt der verlängerte Loop L3 deckelartig fast über dem gesamten terminalen GlcNAc und sorgt damit für eine direkte Interaktion mittels K172. Dies erlaubt eine hochaffine Bindung mittels einer Mischung aus ladungsgetriebener und hydrophober Wechselwirkung mit dem Liganden durch Abschirmung vom *bulk*-Lösungsmittel. Die Berechnung der Interaktionsflächen zwischen Glycan und Protein mit Hilfe des PDBePISA-Servers verdeutlichen dies weiter³⁰⁷. Für Flo5A•Mannose erhält man hierbei eine Fläche von 133 Å², für Epa1A•Galactose 164 Å² und für Cea1A•GlcNAc 192 Å². Die Variation der Loops in Sequenz und Konformation führt dazu, dass in jedem der drei Adhäsine jeweils eine andere Seite der Bindungstasche ab-

geschirmt ist (Abbildung 5.7 D). Diese räumliche Abgrenzung der Bindungstasche könnte bei der Steuerung der Selektivität in Bezug auf die glycosidische Bindung und die nachfolgenden Glycane eine Rolle spielen. Darüber hinaus konnte in Cea1A eine einzigartige, zur Bindung der *N*-Acetyl-Gruppe geeignete, positiv geladene Fläche in der Nähe der Bindungstasche gefunden werden. In Flo5A und Epa1A, welche keine geladene Wechselwirkung an dieser Stelle unterstützen, findet sich ein eher leicht negativ geladener Bereich (Abbildung 5.7 E).

Während Flo5A eine schwache, eher durch Avidität getriebene, aber selektive Bindung von Mannose zeigt, vermag Lg-Flo1A Mannose-1-phosphat im hohen mikromolaren Bereich zu erkennen^{107,135}. Nur Epa1A bindet aufgrund einer abgeschirmten Bindungstasche und der hydrophoben Interaktion mittels W198 in vergleichbarer Stärke an Galactoside¹⁰⁸. Cea1A verbindet nun die Merkmale einer abgeschirmten Bindungstasche mit einer ladungsgetriebenen Interaktion.

Cea1A vermittelt die Bindung an terminale *N*-Acetylglucosamine in Glycanketten mit μM Affinität. Hierdurch sind, wie in den Ergebnissen dieser Arbeit gezeigt werden konnte, Cea1A-präsentierende Zellen in der Lage an enzymatisch behandelte Chitin-Beads zu adhären. Es gibt weitere Beispiele Chitin-bindender Proteine in der Literatur; auch außerhalb des Reiches Fungi. Hierzu zählen zellwandassoziierte und cytoplasmatische sowie sekretierte Proteine³⁷⁰. Der Unterschied ist, dass diese, im Gegensatz zu Cea1A, meist eine breite Spezifität zeigen und ausschließlich die fibrilläre Superstruktur des kristallinen α -/ β -Chitins erkennen. Dies ist unter anderem bei vielen Chitinasen der Fall³⁷¹.

Die vergleichende Betrachtung von Cea1A mit den beiden Repräsentanten weiterer Subklassen, Flo5A und Epa1A, verdeutlicht die Komplexität der Glycanerkennung durch diese vielseitigen Zellwandproteine. Flokkuline sind in der Lage Mannoside zu binden, während epitheliale Adhäsiene vordringlich die Erkennung von Galactosiden vermitteln. Die selektive, hochaffine Adhäsion an Glucoside, hier *N*-Acetylglucoside, wurde bei dieser Proteingruppe noch nicht beobachtet. Cea1A schafft es, durch Anpassung der hypervariablen Loops, eine chemisch sowie konformationell einzigartige Bindungstasche zu erzeugen, welche diese Bindung ermöglicht. Weiterhin verdeutlicht das Studium von Cea1A und dessen Einordnung in den Kontext der Adhäsiene mit PA14-Domäne, dass die Erkennung von Glycanen durch diese Proteinfamilie sehr viel komplexer ist, als bislang vermutet.

5.2.5 Die biologische Funktion von Cea1

Obwohl die exakte biologische Rolle von Cea1A nicht eindeutig geklärt ist, lassen die identifizierten Eigenschaften doch eine Einordnung in einen biologischen Kontext zu. Die Bildung mikroskopischer Flocken bei Oberflächenpräsentation von Cea1A auf *S. cerevisiae* ist zwar zu beobachten, doch ist diese weitaus schwächer als die von Flo5A oder Flo11A verursachte. Die Ausbildung dieser Flocken kann durch Bindung von freiem nicht-kristallinem Chitin in der

Zellwand von Hefen verursacht werden. Zwar ist dieses tief in der Zellwand unter Schichten von Mannoproteinen und β -Glycanen verborgen und macht nur 1-2% der Trockenmasse der pilzlichen Zellwand aus, doch ist es bspw. an Sprossnarben oberflächenexponiert^{71,312-314,372}. Die hohe Bindungsaffinität von Cea1A deutet allerdings eher in die Richtung eines adhäsiven, Epa-ähnlichen, anstatt eines flokkulierenden, Flokkulin-artigen, Phänotyps. Die beobachtbare Flokkulation in *P. pastoris* Wildtyp-Zellen ist dadurch erklärbar, dass neben Cea1A weitere putative Adhäsine mit PA14-Domäne in *P. pastoris* vorliegen und diese für diesen Effekt verantwortlich sein können³⁷³.

Bei aktueller Faktenlage ist davon auszugehen, dass das Habitat von Hefen der Genera *Pichia* oder *Komagataella* verrottendes Holz ist³⁷⁴. In der Literatur gibt es wiederum viele Insekten, welche in enger Beziehung zu Hefen stehen. Hierbei sind die Hefen Futter, Symbiont oder Pathogen^{48,61,375,376}. So konnten STEFANINI *et al.* zeigen, dass Wespen von *S. cerevisiae* und weiteren Hefen als Vektoren genutzt werden, um neue Habitate zu erschließen²⁰. Vielmehr wurden auch zu *Pichia* gerechnete Hefen im Darm unterschiedlicher Käfer, wie *Odontotaenius disjunctus*, gefunden³⁷⁷. Als Mycosymbionten dienen Hefen hierbei zur Verdauung komplexer Glycane in speziellen Organellen, den Mycangia. TOKI *et al.* berichteten von sog. „Fungal farming“ in *D. bucculenta*, wo Hefen in abdominalen Mycangien gesammelt und zur Auskleidung des Baues genutzt werden³⁷⁸ (Abbildung 5.8).

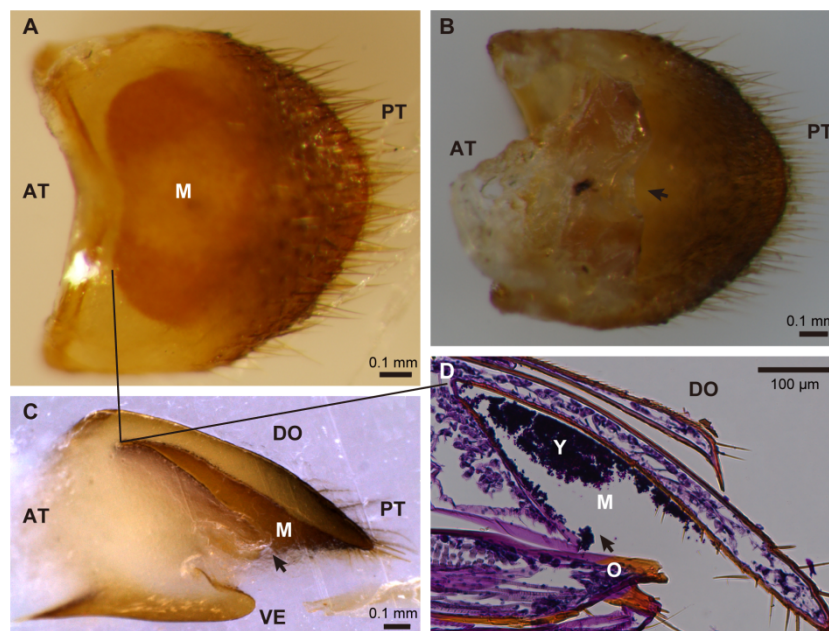


Abbildung 5.8 – Mycangium von *Doubledaya bucculenta* zur gezielten Besiedlung mit symbiotischen Hefen. A: Dorsale Ansicht vom Ende des achten Segmentes. B: Ventrale Ansicht. C: Laterale Ansicht des in Paraffin eingebetteten und longitudinal geschnittenen Segments. D: Longitudinale Sektion des Segments, gefärbt mit SCHIFF's Reagenz und Hämatoxylin. Die Linien deuten den Beginn der Öffnung des Mycangiums an. Abkürzungen: M: Mycangium, O: Ovipositor, Y: Hefe, AT: Anterior, DO: Dorsal, PT: Posterior, VE: Ventral. Entnommen aus TOKI *et al.*³⁷⁸.

In Anlehnung hierzu wurde *Pichia pastoris* zuerst in räumlicher Nähe zur Fruchtfliege *Drosophila melanogaster* isoliert⁵⁴. Alle diese Insekten besitzen nicht nur ein aus Chitin bestehendes

Exoskelett, sondern häufig auch einen aus chitinösen Polymeren aufgebauten Verdauungstrakt, die sog. peritrophe Matrix³⁷⁹. In Verbindung mit sekretierten Chitinasen zum Aufbruch der kristallinen Struktur des Chitins, ist eine Cea1A-vermittelte Adhäsion von Hefen an Insekten denkbar. Solche Chitinasen sind in Hefen durchaus bekannt³⁸⁰. Die starke Besiedlung von *S. lateralis* durch Hefen (sogar durch Laborstämme), welche im Rahmen dieser Arbeit beobachtet werden konnte, sowie die hohe Affinität von Cea1A für nicht-reduzierende Enden nicht-kristallinen Chitins sprechen für solch eine hypothetische Vergesellschaftung. Das Chitin der peritrophen Matrix scheint darüber hinaus von Natur aus mit terminalen, nicht reduzierenden Enden versehen zu sein, welche für die Bindung durch Proteine zugänglich sind³⁸¹.

5.2.6 Potenzielle Anwendung von Cea1 als Reporter in der Medizin

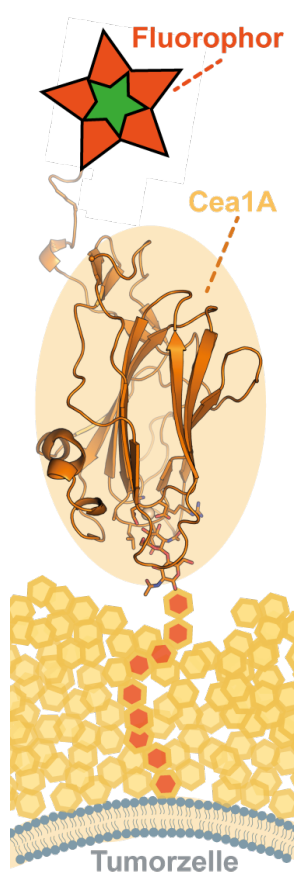


Abbildung 5.9 – Modell einer Anwendung von Cea1A als Reporter. Die Kopplung mit einem Fluorophor könnte die Diagnostik hyper-*O*-GlcNAcylierter Tumorzellen ermöglichen. GlcNAc-Reste sind

(Abbildung 5.9). Aufgrund dieses Potentials wurde Cea1 und dessen A-Domäne sowie daraus resultierende Produkte zur Patentierung angemeldet (Anmeldenummer: EP15160171.3, Anmeldedatum: 20.03.2015)³⁸⁸.

Da Cea1A durch eine hohe Affinität zu *N*-Acetylglucosaminen mit sehr enger Spezifität gekennzeichnet ist, bietet sich dessen Anwendung in der Biomedizin an. Diese Eigenschaften sind wichtige pharmakodynamische und pharmakokinetische Parameter, um Wirkort oder Wirkmechanismus eines Therapeutikums zu steuern. Somit könnten durch Cea1A selektiv Zellen markiert werden, welche terminale GlcNAc-Reste tragen; wie dies bspw. bei einigen Krebsarten der Fall ist. Hier wird durch den Warburg-Effekt die hyper-*O*-GlcNAcylierung an der Zelloberfläche ausgelöst^{382,383}. Die Anwendung GlcNAc-erkennender Proteine für solche Zwecke wurde in der Literatur bereits diskutiert³⁸⁴⁻³⁸⁶. Nachteilig wirken sich bei den hierfür ausgewählten Proteinen allerdings hauptsächlich die schlechtere Selektivität und geringere Affinität aus. Darüber hinaus zeigen so gut wie alle bekannten GlcNAc-bindenden Proteine mehr als eine Bindungstasche oder eine Dimer-/Oligomerisierung³⁸⁷. Dies würde bei therapeutischer oder diagnostischer Anwendung zur unerwünschten Koagulation oder Quervernetzung von Zellen führen. Cea1A hingegen besitzt (*i*) nur eine Bindungstasche, (*ii*) eine durch EDTA-Zugabe lösbare Bindung, (*iii*) ohne B-Region nicht die Neigung zur Oligomerisierung sowie (*iv*) von der Bindungstasche weit entfernte Termini zur Funktionalisierung. Eine Funktionalisierung ist hier einerseits mit Fluoreszenzfarbstoffen zur Detektion von Tumoren denkbar, auf der anderen Seite könnten auch gekoppelte Wirkstoffe an den gewünschten Wirkort gebracht werden

5.2.7 PpFlo1A besitzt eine PA14-Domäne mit unbekannter Funktion

Neben Cea1A konnte im Rahmen dieser Arbeit ein weiteres putatives Adhäsion mit PA14-Domäne aus *P. pastoris* charakterisiert werden. PpFlo1A besteht ebenfalls aus dem strukturell konservierten, rigiden β -Sandwich Motiv (Abschnitt 4.2.8). Die Identität beider A-Domänen beträgt 59%, doch konnte bereits bei den epithelialen Adhäsionen gezeigt werden, dass eine hohe Konservierung auf Sequenzebene nicht automatisch mit einer konservierten Funktion korrelieren muss. Dies ist in PpFlo1A ebenfalls der Fall, wie durch die *Glycan-Array* Experimente gezeigt werden konnte. Trotz der hohen Identität zeigt PpFlo1A keine Bindung an einen Liganden des Mikrochips. Das hier kein Glycan identifiziert werden konnte, kann mehrere Gründe haben. Neben den bereits in Abschnitt 4.2.7 angesprochenen ist denkbar, dass ähnlich wie bei Flo5A, die Affinität von PpFlo1A zu seinem Liganden zu gering ist, um die Wechselwirkung mit Hilfe des *Glycan-Array* nachweisen zu können. Möglicherweise ist der native Ligand von PpFlo1A auch nicht auf dem Chip vorhanden, da dieser primär mit Säuger-spezifischen Oligosacchariden bestückt ist. Die Bindungstasche von PpFlo1A exponiert viele geladene Reste in Richtung des zu bindenden Liganden, besitzt aber, wie alle anderen Adhäsine einen, hier aus F174 und I208 geformten, hydrophoben Bereich für die Bindung eines C6-Kohlenstoffs. Dies eröffnet die Möglichkeit der selektiven Erkennung derivatisierter Glycane, wie sulfatierte oder phosphorylierte Zucker. Für die Aktivität des Proteins spricht, dass in der Kristallstruktur deutlich die Elektronendichte von Glycerol zu erkennen ist, welches, wie ein gebundenes Glycan, mit zwei vicinalen Hydroxylgruppen an das Ca^{2+} -Ion bindet. Darüber hinaus sind phosphorylierte Mannoside in der Zellwand von Hefen literaturbekannt³¹⁸. Die Zusammensetzung der Zellwand variiert mit Zellstadium, Nahrungsaufnahme, chemischer Umgebung und dem Vorhandensein von Stressbedingungen³⁸⁹. Sie unterliegt parallel dem ständigen Umbau durch Glucanasen und Chitinasen^{81,390}. Hierbei ist denkbar, dass nur eine Subpopulation von Hefen gebunden wird. Dies wurde durch die fluoreszenzmikroskopischen Experimente angedeutet. Da bei PpFlo1A präsentierenden Zellen keine ausgeprägte Flokkenbildung beobachtet werden konnte, ist Mannose als primärer Ligand auszuschließen.

Ginge man davon aus, PpFlo1A sei nicht in der Lage, Glycane zu binden, so böten sich in Hefen und dem sie umgebenden Ökosystem noch weitere Liganden mit ähnlich angeordneten funktionellen Gruppen an. Beispielsweise konnte kürzlich gezeigt werden, dass Melanin in der Zellwand und im Cytosol von Hefen vorkommt, dessen Funktion hier aber noch weitgehend unbekannt ist (Abbildung 5.10)³⁹¹. Eine Bindung der beiden Hydroxyle an das Ca^{2+} -Ion wäre hier prinzipiell denkbar.

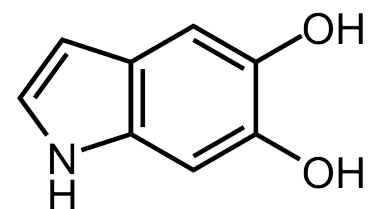


Abbildung 5.10 – Struktur von 5,6-Dihydroxyindol als Monomer von Eumelanin.

5.2.8 PpFlo1A besitzt möglicherweise keinen GPI-Anker

Eine Sonderstellung nimmt *PpFlo1A* möglicherweise auch beim Domänenaufbau ein. Während bei dem hier verwendeten Gen nur die Sequenz der A-Domäne bekannt ist, sind zwei putative Homologe aus *P. pastoris* CBS7435 und GS115 mit jeweils ~74% Sequenzidentität in der A-Domäne vollständig sequenziert (UniProt-ID: F2QQ27 und C4QYW7). Die hier annotierten Genprodukte zeigen nach der A-Domäne nur einen Repeat und brechen danach mit einem Stop-Codon ab. Darüber hinaus sind die Reste in CBL2 nicht identisch mit denen in *PpFlo1A*. In *Cea1A* hingegen ist in den 2 Homologen der Bereich der Bindungstasche konserviert. Weitere annotierte, putative Homologe für *PpFlo1A* sind nicht bekannt. Der vorhergesagte Abbruch des Genproduktes nach dem ersten Repeat kann an einer fehlerhaften Annotation oder Sequenzierung in dem hochvariablen und schlecht zugänglichen subtelomeren Bereich liegen¹⁰⁰. Da es sich hierbei allerdings um Befunde aus zwei unterschiedlichen Stämmen handelt und die Sequenzen in der Mitte eines *shotgun-contigs* liegen, ist davon auszugehen, dass die Sequenzen valide sind. Dabei stellt sich die Frage ob *PpFlo1A* überhaupt eine Funktion hat oder es sich um ein ungenutztes Genfragment handelt, welches durch Rekombinationsereignisse in den subtelomeren Bereichen des Chromosoms entstanden ist. Weitere qPCR-Experimente könnten hier für Klarheit sorgen. *PpFlo1A* besäße demnach auch keinen GPI-Anker und wäre ein sekretiertes PA14-Protein ohne Zellwandverknüpfung, da das Sekretionssignal hingegen vorhanden ist. Dies könnte ebenso gut auch eine neuartige Funktion für PA14-Domänen in Hefen darstellen. So wurde für Flo11A bspw. sogenanntes „shedding“ beobachtet³⁹². Hierbei wird Flo11A, vermutlich durch posttranslationale Reaktion mit einer Protease, vom GPI-Anker getrennt und sekretiert. Dies dient möglicherweise zur Optimierung der Flo11A-induzierten Wechselwirkung und einer Mucus-artigen Funktion, wie diese in höheren Organismen bekannt ist³⁹³. Untersucht man den Bereich, der dem Stop-Codon folgt, ist keine Ähnlichkeit mit der typischen B-Region festzustellen; es handelt sich also auch nicht um eine *nonsense*-Mutation. Der Domänenaufbau könnte ebenso ein *N*-terminales Pendant zu den ursprünglich als GLEYA-Domänen bezeichneten, *C*-terminalen PA14-Domänen ohne GPI-Anker darstellen³²⁷. Hierbei könnte die Fixierung in der Zellwand bspw. über eine Veresterung wie bei *PIR*-Proteinen erfolgen⁹³.

Möglicherweise ist der Genus *Pichia* im Verlauf seiner Entwicklung mit einem Glycan in seiner Umgebung in Kontakt gekommen, für welches sich ein passendes Adhäsion, *PpFlo1*, durch Rekombination und Mutation entwickelte. Entweder war dieser Versuch nicht erfolgreich oder der Selektionsdruck wurde durch Wegfall des Liganden verringert und das Gen bzw. die Funktion verloren. Darüber hinaus ist ebenso ein Verlauf denkbar, bei welchem es sich einfach um eine zufällig generierte Rekombination handelt, welche nie zum Einsatz kam.

5.2.9 Der Übergang zur B-Region von *Cea1A* und *PpFlo1A* im Vergleich

Trotz der erhöhten Sequenzidentität zwischen *Cea1A* und *PpFlo1A* zeigt letzteres den aus den weiteren Strukturen bekannten helikalen/loopförmigen anstatt den in *Cea1A* gefunde-

nen stiel förmigen Verlauf (Abschnitt 4.2.4). Da auch diese Bereiche im *Pichia*-Subtyp auf Sequenzebene gut konserviert sind (Abbildung 5.11 A), lässt sich anhand eines Vergleichs zeigen, warum es zu zur Ausbildung verschiedener Struktur motive kommt.

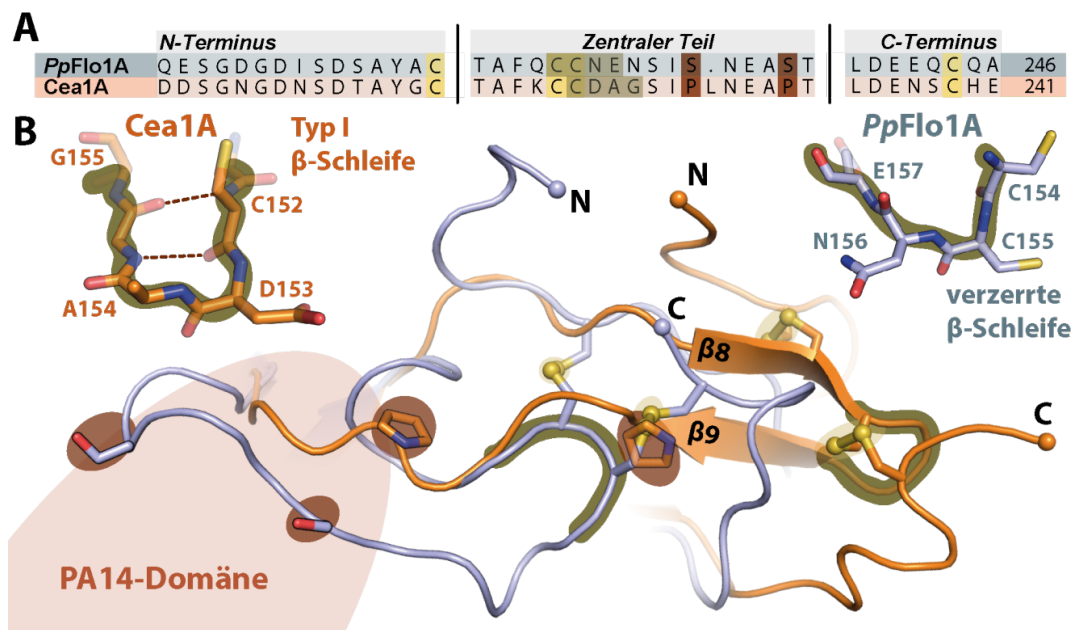


Abbildung 5.11 – Vergleich der Faltungsmotive am CC-Motiv. A: Ein Alignment der Sequenzbereiche die das Faltungsmotiv bilden zeigt nur geringe Abweichungen. B: Der strukturelle Vergleich hingegen macht deutlich, dass bereits die geringen Abweichungen ausreichen das alternative Faltungsmotiv zu erzeugen. Hierzu gehören in *PpFlo1A* (blau) zwei Serine, welche in *Cea1A* (orange) gegen Proline ausgetauscht (rot hinterlegt) sind sowie die Reste, welche direkt in die Bildung der β -Schleife involviert sind. Sterisch weniger anspruchsvolle Reste sorgen in *Cea1A* zur Ausbildung einer Schleife vom Typ I (ocker hinterlegt) in Form einer Haarnadelschleife. In *PpFlo1A* sind diese gegen voluminösere Reste ausgetauscht, was zur Verzerrung dieser Region und der Ausbildung des ungeordneteren Bereichs führt.

Hierbei wird deutlich, dass in *Cea1A* zwei Serine gegen Proline ausgetauscht sind. Dies schränkt die konformationelle Freiheit stark ein und verhindert dadurch die Ausbildung des helikalen Bereichs. Darüber hinaus ist die Sequenz, welche die Schleife bildet, verändert. In *Cea1A* besteht die β -Schleife aus C152, N153 und den sterisch anspruchslosen Resten A154 und G155. Hierbei werden zwei Wasserstoffbrückenbindungen zwischen den Strängen möglich und es kommt zur Ausbildung einer β -Schleife vom Typ I (Abbildung 5.11). Dieser unterscheidet sich von den weiteren Typen durch die Orientierung (Winkel Φ und Ψ) des Peptid Sauerstoffs zwischen Rest $n+1$ (N153) und $n+2$ (A154)³⁹⁴. Typ I ist in Proteinstrukturen am häufigsten zu finden, jedoch wird er in Form einer Haarnadelschleife wie in *Cea1A* nur selten beobachtet^{395,396}. Ebenso selten ist das Vorhandensein eines Cysteins und damit einer Disulfidbrücke an erster Stelle. Zur Ausbildung der β -Haarnadelschleife des Typs I trägt vor allem G155 an Position $n+3$ bei, da dieses sterisch besonders anspruchslos ist und damit sehr häufig in diese Schleifentyp gefunden wird (Über 30% der Fälle)³⁹⁷. Das Sequenzmotiv CDAG wurde laut „*Motivated Proteins*“ Server, welcher eine Datenbank für kleine, durch Wasserstoffbrücken gebildete Struktur motive ist, noch nicht in β -Schleifen gefunden³⁹⁷. So-

mit ist der stielförmige Bereich in *Cea1A* durch das kleine Faltblattmotiv, die β -Schleife sowie zwei Disulfidbrücken stabilisiert und vermutlich sehr rigide.

Der Austausch der an einer Schleife beteiligten Reste in *PpFlo1A* zu CCNE führt aufgrund des sterischen Anspruchs zu der Bildung einer stark verzerrten β -Schleife, welche nicht zur Bildung einer Haarnadelschleife geeignet ist. Offenbar ist der Austausch des Schleifenmotivs und der Serine ausreichend, um in *Cea1A* den stielförmigen Bereich zu erzeugen, während dieser in *PpFlo1A* helikal/loopförmig vorliegt. An den gleichen Positionen finden sich in den Strukturen der übrigen Adhäsine-Subtypen ebenfalls räumlich sehr anspruchsvolle, für solch eine Schleifenregion ungeeignete Reste.

Eine Besonderheit in der Struktur von *PpFlo1A* ist die Ausbildung von Kristallkontakten über den *N*-terminal vorhandenen *His*₆-*Tag* (Abschnitt 4.2.8). Die Etablierung der Aufreinigung von heterolog produzierten Proteinen als Fusionsprotein mit Affinitätstags führte zur Diskussion über den Einfluss solcher *Tags* bei der Untersuchung dieser Proteine. Hierbei wurde dazu tendiert, bei Kristallisation von Proteinen *Tags* abzuschneiden, da deren Mobilität die Kristallisation im schlimmstenfalls verhindern könne^{398,399}. Es konnte allerdings immer häufiger gezeigt werden, dass ein terminaler *His*₅- oder *His*₆-*Tag* in aller Regel keine signifikante Auswirkung auf die Kristallisation von Proteinen hat⁴⁰⁰. Die Ausbildung von direkten Kristallkontakten über einen *His-Tag* blieb hierbei eher die Ausnahme. Bei Betrachtung der Anordnung der beiden *PpFlo1A*-Moleküle in der asymmetrischen Einheit wird deutlich, dass der *His*₆-*Tag* den größten Anteil der Kristallkontakte innerhalb der Einheit stellt. Obwohl, aufgrund der exakt gleichen Orientierung beider Moleküle zum *Tag*, dieser keinem der beiden Moleküle eindeutig zugeordnet werden kann, ist die Validität des Modells auch dadurch bestätigt, dass *N*-terminal zwischen jedem Molekül und dem potentiellen *His-Tag* etwa 20 Aminosäuren nicht durch die Elektronendichte beschrieben sind. Diese Linkerregion wäre ausreichend lang, um eine Verbindung zu dem isolierten Strang herzustellen.

Zusammenfassend lässt sich sagen, dass auch in der Subklasse der *Pichia*-Adhäsine eine hohe Ähnlichkeit auf Sequenzebene nicht auf eine korrelierende Funktion schließen lässt. Obwohl *Cea1A* und *PpFlo1A* phylogenetisch eng verwandt sind, zeigen sich unterschiedliche Bindungstaschen und terminale Bereiche. Ständige Rekombination beeinflusst den phylogenetischen Baum und macht erneut die Komplexität dieser Proteinfamilie bewusst. Nichtsdestotrotz ist die Zuordnung der einzelnen Genprodukte zu den entsprechenden phylogenetischen Subtypen valide, da globale Eigenschaften trotzdem konserviert vorliegen. Im *Pichia*-Subtyp sind dies das Disulfidbrückenmuster, das Fehlen von Subdomänen bzw. die Länge der Loopbereiche und die ähnliche Sequenz im Bereich der PA14-Domäne.

5.3 Die epithelialen Adhäsine aus *C. glabrata* erzeugen funktionelle Vielfalt durch Anpassung hypervariabler Bereiche

Die bioinformatische Untersuchung der epithelialen Adhäsine zeigte die größte Varianz innerhalb eines Subtyps, was im phylogenetischen Baum durch einen breiteren Ast angedeutet wird (Abschnitt 4.1, Abbildung 4.2). Der Vergleich der Sequenzen macht deutlich, dass diese Varianz hauptsächlich in den hypervariablen Loops und CBL2 zu finden ist; also dem für die Glycanerkennung wichtigen Bereich der Struktur. Bereits am Beispiel von *PpFlo1A* und *Cea1A* wurde gezeigt, dass eine phylogenetische Nähe nicht zwingend mit einer funktionellen Konservierung korreliert. Bei den epithelialen Adhäsinen ist dieses Prinzip insofern verändert, als das mit strukturell variablen Adhäsinen ähnliche Liganden erkannt werden. Diese Erkennung von unterschiedlich gebundenen und derivatisierten Galactosiden durch eine Vielzahl der 17 Epa Paralogen in *C. glabrata* dient der Adhäsion an Epithelzellen und damit der Etablierung der Pathogenizität¹⁶⁵. Der hohe Selektionsdruck erfordert eine ständige Adaption an den Wirtsorganismus, was möglicherweise zur Entwicklung einer ganzen Bandbreite an Adhäsinen geführt hat. Ob jedes Mitglied der Epa-Familie bspw. selektiv einen eigenen Epithelzelltyp erkennt oder es sich um eine Batterie an Adhäsinen zur generellen Adhäsion handelt, welche ständiger Adaption unterliegen, ist Gegenstand aktueller Debatten^{108,165,401}. Letztere Hypothese lässt sich komplementär zur Anpassung der Glycosylierung von *Trypanosoma* beschreiben. *T. brucei* besitzt mehr als 1000 Gene, welche für sog. *Variable Surface Glycoproteins* (VSG) codieren³⁴⁰. Durch Veränderung des Expressionsmusters lassen sich so die für den Wirt erkennbaren Oberflächenstrukturen ändern um dem Immunsystem zu entkommen. Diese Gene unterliegen ebenfalls starker Rekombination. Mit den epithelialen Adhäsinen wird nicht versucht, dem Wirt auszuweichen, sondern spezifisch an ihn zu binden. Dieser starke Selektionsdruck resultierte in der Entwicklung hoch-affiner Adhäsine, wie Epa1A, welches im niedrigen μM Bereich an Galactoside bindet. Am Beginn dieser Entwicklung könnten hierbei weniger gut an Galactoside angepasste Adhäsine gestanden haben, wie sie in mit *C. glabrata* verwandten, nicht pathogenen Spezies gefunden wurden. Diese besitzen hingegen keine Epa-artigen Adhäsine, es fehlt bspw. das in L3 vorhandene, essentielle Tryptophan^{37,108}.

Wie RIKE DIDERRICH zeigen konnte, lassen sich die Epas je nach Selektivitätsmuster in funktionelle Gruppen einteilen¹⁶⁵. Hierbei wurde weiterhin deutlich, dass sich die Spezifität der Epas untereinander zwar in Bezug auf glycosidische Verknüpfung oder Derivatisierung unterscheidet, die Bindung von Galactosiden jedoch von dem Großteil der Adhäsine unterstützt wird.

5.3.1 Phylogenetische Nähe impliziert keine funktionelle Ähnlichkeit

Doch wie werden diese unterschiedlichen Selektivitäten erreicht? Eine Zuordnung der Sequenzen der A-Domänen zu ihren Funktionsklassen zeigt, dass die phylogenetische Ver-

wandtschaft nicht zwingend mit ähnlichen Bindungseigenschaften einhergeht (Abbildung 5.12 A). Die A-Domänen funktionell eng verwandter Adhäsine sind über den Ast des phylogenetischen Baumes verteilt. Umgekehrt gehören phylogenetisch verwandte Sequenzen nicht zwingend einer funktionellen Klasse an; die Bindungseigenschaften können sich sogar stark unterscheiden.

Das Studium der strukturellen und funktionellen Diversität des Epa-Subtyps deutet eher darauf hin, dass sich einzelne Epa-Domänen unabhängig voneinander entwickelt haben, was letztere ungerichtete Entwicklungshypothese stützt. Die Konservierung der Reste in den Epa-Subtyp-Mitgliedern variiert stark mit dem betrachteten Strukturbereich (Abbildung 5.12 B). Während alle untersuchten Epas auf einer gemeinsamen PA14-Domäne basieren, zeigen die oberflächenexponierten Reste große Abweichungen.

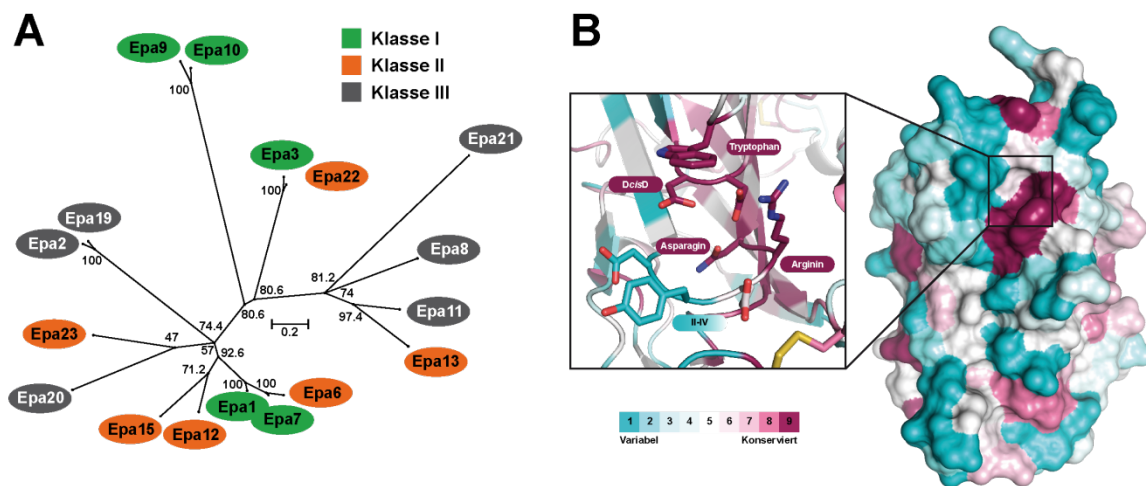


Abbildung 5.12 – Phylogenetische und funktionelle Verwandtschaft sind unabhängig. A: Die Abbildung der Zugehörigkeit zu funktionellen Klassen auf einen phylogenetischen Baum der epithelialen Adhäsine zeigt, dass die phylogenetische Nähe zweier A-Domänen nicht mit funktioneller Vergleichbarkeit korreliert. *Bootstrapping*-Werte (500 Zyklen) zur Validierung sind in Prozent angegeben B: Die Einfärbung der Strukturbereich der epithelialen Adhäsine nach dem Konservierungsgrad unterstreicht die Variabilität der hypervariablen Loops, verdeutlicht allerdings auch die hohe Konservierung einiger Reste in der Bindungstasche sowie der Disulfidbrücken.

Ein detaillierter Blick in die Bindungstasche zeigt allerdings, dass wichtige Bereiche über alle Epa-Subtyp Mitglieder konserviert sind. Neben dem essentiellen *DcisD*-Motiv und dem Ca^{2+} -koordinierenden Asparagin aus CBL2 sind weitere konservierte Reste im vom Glycan aus gesehen rechten Teil der Tasche zu finden (Abbildung 5.12 B). Darüber hinaus ist ein Tryptophan-Rest des Loops L3 und ein Arginin an Position I von CBL2 in jedem epithelialen Adhäsine zu finden. Dieses strukturelle Kernmotiv scheint unabdingbar für eine korrekte Erkennung der zu bindenden Hexose. Diesem unveränderlichen Kern der Bindungstasche steht eine hochvariable Seite gegenüber. Wie die Experimente mit unterschiedlichen Glycanen für Epa1A, Epa6A und Epa9A gezeigt haben, ist eine Variabilität an den Positionen II-IV von CBL2 überaus wichtig zur Differenzierung zwischen bspw. α - und β -verbrückten Galactosiden. Die Komplexstruktur von Epa1A mit Sulfogalactose deutet darauf hin, dass dieser Bereich der Bindungstasche die Erkennung unterschiedlich derivatisierter Galactosen erlaubt.

Der erlaubte Grundbaustein kann hierbei je nach sterischem Anspruch der Reste in den verschiedenen Epas verändert werden. Dies verdeutlicht erneut die zweite Entwicklungsthese; einer funktionierenden Grundstruktur werden variable Bereiche hinzugefügt, um das Spektrum der erkennbaren Liganden bei Restriktion auf physiologisch relevante Galactosederivate breit genug zu halten um eine Etablierung des Organismus im Wirt zu ermöglichen.

5.3.2 Die Struktur/Funktions-Beziehung der Epas ist sehr komplex

Die Komplexstrukturen machen allerdings auch deutlich, dass die Spezifität der Mitglieder des Epa-Subtyps keineswegs nur auf diese wenigen Reste im primären Bereich der Bindungstasche zurückgeführt werden kann und weitaus komplexer ist als bisher angenommen. Wie im Vergleich zwischen Cea1A, Flo5A und Epa1A zu sehen war, schirmen die untersuchten Repräsentanten dreier Subtypen unterschiedliche Bereiche der Tasche ab. Dies gelingt hauptsächlich durch die Länge und Zusammensetzung der hypervariablen Loops L1-L3. Dies lässt sich als Feinabstimmung der hier von Epa1A abgeschirmten Bereiche beschreiben. Es gibt ebenfalls vier Fälle (Epa1A/Ep7A, Epa3A/Epa22A), in denen Epas gleiche Reste in CBL2 aber unterschiedliche Spezifitäten haben. Die Form und die chemischen Eigenschaften der äußeren Bindungstasche hängen sehr stark von den umgebenden Loops ab. Strukturell konnte festgestellt werden, dass die Hauptinteraktionen zwischen Glycan und Protein hauptsächlich im Kern der Bindungstasche etabliert werden. Der äußere Teil vermittelt einen zwar untergeordneteren, aber für die Selektivität wichtigen Beitrag zur Ligandenbindung. Dies trifft gerade in Bezug auf unterschiedliche Verknüpfungen (β 1-3 vs. β 1-4) der Zucker zu, wie die Ergebnisse für Epa6A und Epap9A zeigen konnten. Die Komplexstrukturen von Epa1A mit synthetischen Disacchariden konnten zeigen, dass dieses Adhäsins an der C6-Position des zweiten Zuckers nur wenige Interaktionen eingeht und damit diese Position in der Erkennung komplexer Glycane nur eine untergeordnete Rolle spielt. Hierdurch wird genügend Flexibilität bei der Erkennung von Liganden mit hoher Affinität kombiniert, was die Vorrangstellung von Epa1A als einen der Hauptpathogenitätsfaktoren unter den Epas unterstreicht¹³⁷. Dass die Affinität und die Spezifität zur Bindung des terminalen Galactosids stark von der Form und den chemischen Eigenschaften der primären Bindungstasche abhängig ist, konnte mittels Sättigungsmutation der in CBL2 vorhandenen Reste durch IELASIS *et al.* gezeigt werden. Durch die Mutationen E227A und Y228W konnte hier eine Affinitätssteigerung und Änderung der Selektivität beobachtet werden⁴⁰². Die Beschreibung der globalen Bindungsverhältnisse von Epa1A auf Basis dieses einen Loops ist allerdings unzureichend. Wie zuvor dargelegt werden konnte ist die Beschränkung auf einzelne Strukturbereiche nicht ausreichend um ein aussagekräftiges Bild der Selektivität und Spezifität der Adhäsine zu erhalten. Darüber hinaus konnte bereits durch MAESTRE-REYNA *et al.* gezeigt werden, dass der Austausch der CBL2-Region einen maßgeblichen Effekt auf Spezifität und Affinität hat¹⁰⁸.

Die treibende Kraft hinter der Entwicklung der Epa-Subtyp Adhäsine liefert möglicherweise der hohe Selektionsdruck, dem *C. glabrata* als Pathogen unterliegt. Dies führte zur Etablie-

nung einer großen Zahl hochaffiner Adhäsine, welche durch ihre Komplexität in der Lage sind, eine große Anzahl Galactoside zu erkennen und die selektive Wechselwirkung der Hefe mit ihrem Wirt auszubilden. Untermauert wird dies ebenfalls durch die Untersuchungen von GABALDON *et al.* Hierbei konnte gezeigt werden, dass nicht pathogene, mit *Candida* eng verwandte Spezies aus dem Genus *Nakaseomyces*, (*i*) weniger Adhäsine-Gene tragen und diese (*ii*) eher Flokkulinen oder Adhäsinen ohne essentielles Tryptophan in L3 zuzuordnen sind³⁷.

Eine ausführliche Betrachtung des Epa-Subtyps ist bei RIKE DIDERRICH zu finden¹⁶⁵.

5.4 Die Flokkulin-spezifische Subdomäne trägt zur Integrität der primären Bindungstasche bei

Die vorigen Abschnitte zu den Subtypen der *Pichia*-Adhäsine und Epas hatten bereits gezeigt, dass die Zusammensetzung und Form der primären Bindungstasche den größten Einfluss auf die Spezifität der einzelnen Adhäsine hat. Die Feinabstimmung, welche durch die hypervariablen Loops L1-L3 erfolgt, hat im Subtyp der Flokkuline einen größeren Einfluss als bisher angenommen. Zwar wurde in Arbeiten von VEELDERS und SIM bereits ein wichtiger Einfluss der Subdomäne auf die Bindungseigenschaften erkannt, doch bezog sich dies auf durch die Subdomäne vermittelte Interaktionen mit Glycanen bspw. über Q98^{107,135}. Die Kristallstrukturen von Flo10 und Flo10^{5SD} ermöglichen es, den direkten Einfluss der Subdomäne in Position von L2 zu beleuchten. Dabei wurde vor allem deutlich, dass die Subdomäne von Flo10A im Vergleich zu der von Flo5A sehr flexibel ist. Es konnte gezeigt werden, dass (*i*) sich die Subdomänen stark in ihren strukturellen Eigenschaften wie ihrer Rigidität unterscheiden und (*ii*) dass sie neben der direkten Beteiligung an der Glycanbindung einen großen Einfluss auf die Konformation der Bindungstasche haben. Die Beeinflussung der Konformation von L3 durch die Anwesenheit und die Art der Subdomäne spielt hierbei eine entscheidende Rolle. Die mit Hilfe des PDBePISA-Servers³⁰⁷ berechnete Interaktionsfläche mit gebundener Mannose betragen für Flo10A 180 Å², für Flo1A 166 Å² (PDB-ID: 4LHN) und für Lg-Flo1A 168 Å² (PDB-ID: 4GQ7). Hierbei zeigen Flo1A und Flo10A einen deckelförmigen Verlauf von L3 mit direkter Glycan-Interaktion des darin liegenden Lysins, obwohl Flo1A eine rigide, Flo5-artige Subdomäne trägt (Abbildung 5.13). Ohne gebundene Mannose erscheint L3 in Flo1A, trotz schwer zu interpretierender Elektronendichte, hingegen zurückgeklappt (PDB-ID: 4LHL). Darüber hinaus zeigen Lg-Flo1A, welches keine Subdomäne besitzt, und die Chimäre Flo10A^{5SD} ebenfalls einen zurückgeklappten Loop L3. Wie sind diese Befunde zu erklären?

Bei der C-Typ lektinartigen Flokkulation wird zwischen dem *Flo1*-Typ und dem *NewFlo*-Typ unterschieden¹³². Während die Flokkulation nach dem *Flo1*-Typ nur durch Mannose inhibierbar ist, kann der *NewFlo*-Typ ebenfalls durch Glucose, Maltose und Saccharose inhibiert werden. Hierbei wird dem *Flo1*-Typ die Expression der Paraloge *FLO1*, *FLO5* und *FLO9* zugeordnet, die Expression von *Lg-FLO1* hingegen dem *NewFlo*-Typ^{113,132}. Die breite-

re Spezifität von Lg-Flo1A ist hauptsächlich durch das Fehlen einer entsprechenden Subdomäne zu erklären. Dies verdeutlicht allerdings ebenfalls, warum Flo10A, dessen Subdomäne sehr flexibel ist, in begrenztem Maße Bindung an Galactose zeigt¹⁵⁹.

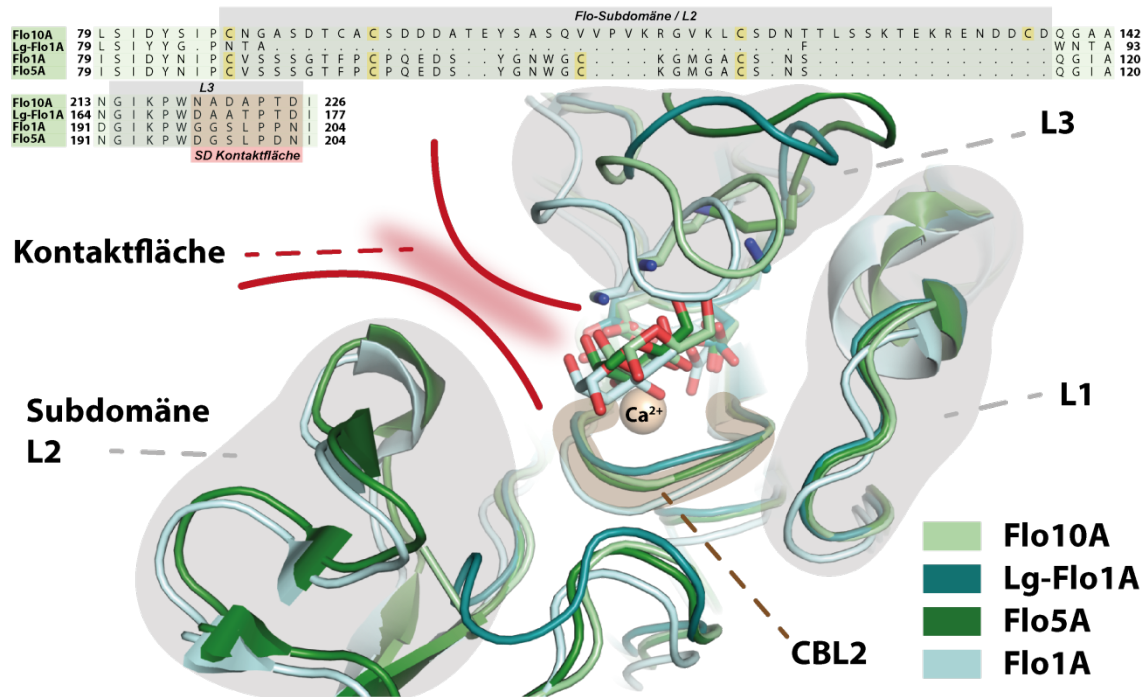


Abbildung 5.13 – Die Interaktionsfläche zwischen der Flokkulin-Subdomäne und dem hypervariablen Loop L3 bestimmt die Form der Bindungstasche. Ein Sequenzalignment der Subdomäne und L3 unterschiedlicher Flokkuline (Flo10A, Lg-Flo1A (PDB-ID:4GQ7), Flo1A (PDB-ID: 4LHN) und Flo5A (PDB-ID: 2XJS)) zeigt, dass diese Bereiche im Vergleich zur fast identischen Sequenz des PA14-Kerns sehr variabel sind. Eine strukturelle Überlagerung deutet an, dass in Abhängigkeit der Kontaktfläche zwischen L3 und der Subdomäne die Form der Bindungstasche variiert und dadurch die Spezifität der Flokkuline verändert wird. Hierdurch kommt es zur Ausbildung der unterschiedlichen Flokkulations-Phänotypen des *Flo1*- und des *NewFlo*-Typs.

Sowohl die Sequenz des hypervariablen Loops L3 als auch der Subdomäne scheinen jeweils aufeinander abgestimmt zu sein. In L3 ist das Motiv GIKPW stark konserviert während der Bereich, welcher in Kontakt mit der jeweiligen Subdomäne tritt, abweicht (Abbildung 5.13). Die Kontaktfläche zwischen L3 und der Subdomäne ist offenbar ausschlaggebend für die Konformation der primären Bindungstasche. Dies zeigt, warum Flo1A und Flo5A, trotz identischer Subdomäne eine abweichende Konformation von L3 zeigen: Die Kontaktflächen sowie die Primärsequenzen von L3 selbst sind leicht unterschiedlich. Dies erklärt weiterhin, warum es zur auffälligen Veränderung dieses Bereiches in der Chimäre Flo10^{SSD} im Vergleich zu Flo10A kommt (Abschnitt 4.4.1).

Darüber hinaus kann durch die Variation der Subdomäne und die damit einhergehende strukturelle Änderung der Bindungstasche die Diskriminierung zwischen *Flo1*-Typ und *NewFlo*-Typ erklärt werden. Die Mutation Q98A, welche von VEELDERS *et al.* für Flo5A beschrieben wurde, zeigt ebenfalls Eigenschaften des *NewFlo*-Typs. Hier ist von einer Ände-

zung der Form der Bindungstasche zugunsten der Bindung von nicht-Mannosiden, möglicherweise unter gleichzeitiger Änderung der Konformation von L3, auszugehen¹⁰⁷.

Zusammenfassend ist festzustellen, dass die Spezifität der Flokkuline neben den Resten an Position I-IV in CBL2 ebenfalls stark von der Orientierung und Länge der hypervariablen Loops abhängt. Ganz besonders deutlich ist dieser Effekt in Bezug auf die Subdomäne, welche eine intrinsische Eigenschaft dieses Subtyps darstellt. Interessanterweise ist die Orientierung von L1 in allen Flokkulinen sehr ähnlich, während dieser den in seiner Länge variabelsten Bereich der Epas darstellt. Bei den Flokkulinen verhält es sich umgekehrt. L2, welcher die Subdomäne enthält ist in Form und Funktion variabel, während dieser in den Epas eine eher untergeordnete Rolle zu spielen scheint. Der PA14-Kern ist innerhalb des Flokkulin-Subtyps konservierter als bspw. bei den Epas. Dies wird an der kleineren Gruppe im phylogenetischen Baum deutlich. Das augenscheinlichste Merkmal der Flokkuline ist damit ihre Subdomäne, welche, wie man an Flo5A und Flo10A sieht, nicht nur unterschiedliche Formen, sondern ebenfalls verschiedenen Funktionen zu haben scheint.

Eine ausführliche Betrachtung von Flo10A wird bei NINA WOZNIAK zu finden sein¹⁵⁹.

5.5 Ausblick

Die bioinformatische Analyse der bekannten und putativen Adhäsine mit PA14- und GLEYA-Domänen aus Ascomyceten liefert die Grundlage für die globale Untersuchung dieser diversen Proteinfamilie. Besonderes Augenmerk sollte hierbei nicht nur auf Bestimmung von Struktur und Funktion von Adhäsinen neuer Subtypen gelegt werden, sondern ebenfalls auf das detaillierte Verständnis innerhalb der Subtypen.

Ausgewählte, putative Adhäsine ohne das konservierte CC-Motiv, welche vorwiegend der C-terminalen GLEYA-Domäne zuzuordnen sind, könnten hierbei heterolog produziert und anschließend funktionell sowie strukturell untersucht werden. Deren Struktur würde ebenfalls Aufschluss darüber geben, in welcher Art und Weise diese Adhäsine die Fixierung des N- und C-Terminus und damit die Rigidität des konservierten β -Sandwich erreichen. Diese Untersuchungen könnten ebenfalls einen Beitrag zum Verständnis der biologischen Funktion dieser ungewöhnlichen Domänenorganisation ohne GPI-Anker liefern, welche nur rudimentär untersucht und beschrieben ist. Darüber hinaus könnte hierdurch die globale Entwicklung der pilzlichen Adhäsine mit PA14-Domäne genauer verstanden werden.

Um das Mosaik zu vervollständigen, sollten ebenfalls Genprodukte des *Kluyveromyces*-Subtyps untersucht werden. Deren putativen Adhäsine tragen voluminöse Aminosäurereste in CBL2 und könnten damit eine neuartige Konformation der Bindungstasche besitzen. Die Betrachtung unterschiedlicher Subtypen sollte helfen, die Anpassung der jeweiligen Adhäsine an eine bestimmte Funktion zu verstehen. Damit können, durch Entdeckung neuer Selektivitäts- und Bindungsmuster, möglicherweise medizinisch oder biotechnologisch relevante

Prozesse untersucht werden. Die Möglichkeit der darüber hinaus gehenden Anwendung potentiell neuartiger Bindungsmodi in Technologie und Wissenschaft ist ein weiteres Argument zur weiteren Untersuchung dieser vielfältigen Proteinfamilie.

Innerhalb der einzelnen Subtypen dagegen lassen sich die molekularen Determinanten, welche Selektivität und Spezifität beeinflussen, detailliert bestimmen. Hierzu könnten vor Allem weitergehende Untersuchungen innerhalb der Flokkuline, der epithelialen Adhäsine sowie der *Pichia*-Adhäsine dienen. In den drei Subtypen ist jeweils eine andere Region der Bindungstasche konformationell modifiziert. Dies scheint ein Kennzeichen des jeweiligen Subtyps zu sein und bietet daher die Grundlage zum Design von chimären Adhäsinen mit alternativen Bindungseigenschaften.

Die detaillierte biologische Funktion von Cea1A könnte ein Fokus von weiteren *in vivo* Experimenten darstellen. Neben dem Ausbau von Experimenten mit Schaben als Modelorganismus zur Untersuchung der Hefen-/Insekten-Vergesellschaftung könnten bspw. anhand von Cryo-EM-Aufnahme die Cea1A-Bindungsstellen innerhalb der *Pichia*-Zellwand bestimmt werden. ITC-Messungen mit unterschiedlichen Glycanstrukturen könnten zur Identifizierung eines geeigneten Bindungspartners von PpFlo1A führen. Da dieses Adhäsins einen dritten Domänenaufbau liefert, eine N-terminale PA14-Domäne ohne geeigneten GPI-Anker, könnte hierdurch eine weitere Funktionsklasse von Adhäsinen weiter untersucht werden.

Anhand des Flokkulin-Subtyps, speziell anhand von Flo1A, Lg-Flo1A, Flo5A und Flo10A, lässt sich der Einfluss der Subdomäne auf die Konformation der Bindungstasche und der vermittelten Funktion, mittels *in vitro* und *in vivo* Tests, weiter untersuchen.

Der Subtyp der epithelialen Adhäsine bietet, aufgrund einer großen Anzahl gut studierter Mitglieder, die Möglichkeit, die Komplexität der Anpassung eines Pathogens an seinen Wirt detailliert zu untersuchen. Durch gezielte Bindungsstudien an unterschiedlichen Gewebetypen ließe sich herbei klären, ob die einzelnen Epas eine Zelltyp-spezifische Adhäsion vermitteln oder ein eher ungerichteter Prozess zu Grunde liegt. Durch die Lösung weiterer Proteinstrukturen dieses Subtyps könnte das Verständnis der Selektivität vertieft und die Entwicklung geeigneter Adhäsions-Inhibitoren erfolgen.

Die korrekte Funktion der Adhäsine hängt stark von der erfolgreichen Präsentation auf der Zelloberfläche ab. Strukturelle Untersuchungen der B-Region könnten hierbei einen Einblick in die zu Grunde liegenden, molekularen Determinanten bieten. Zur Sicherstellung des korrekten O-Glycosylierungsmusters könnte die Proteinproduktion in *Pichi pastoris* durchgeführt werden und strukturelle Untersuchungen mittels Kleinwinkelstreuung (SAXS) durchgeführt werden. Dies würde, neben der realen Faltung der B-Regionen-Repeats, auch die Möglichkeit der Dimer- oder Oligomerisierung überprüfbar machen.

Wie in Abschnitt 5.2.3 besprochen wurde, stellt die Reifung und damit die Ausbildung des *DcisD*-Motivs einen kritischen Schritt in der Etablierung der Funktion dar. Zur Untersuchung dieses Vorgangs könnte möglicherweise ein ITC-Ansatz genutzt werden. Flo5A zeigte die teilweise Rückfaltung nach Hitzedenaturierung. Mittels solche einer, im ITC-Gerät durchgeführten, Reaktion, ließe sich möglicherweise die Kinetik der *trans/cis*-Isomerisierung beobachten. Hierzu könnten Ca^{2+} -Ionen zusammen mit dem hitzedenaturierten Flo5A vorgelegt werden und die bei Abkühlung eintretende Bindung des Metallions anhand der entstehenden Wärmeenergie gemessen werden.

6 Literaturverzeichnis

- 1 Butterfield, N. J. Probable proterozoic fungi. *Paleobiology* **31**, 165-182, (2005).
- 2 Lücking, R., Huhndorf, S., Pfister, D. H., Plata, E. R. & Lumbsch, H. T. Fungi evolved right on track. *Mycologia* **101**, 810-822, (2009).
- 3 Brundrett, M. C. Coevolution of roots and mycorrhizas of land plants. *New phytologist* **154**, 275-304, (2002).
- 4 Vaupotic, T., Veranic, P., Jenoe, P. & Plemenitas, A. Mitochondrial mediation of environmental osmolytes discrimination during osmoadaptation in the extremely halotolerant black yeast *Hortaea werneckii*. *Fungal Genetics and Biology* **45**, 994-1007, (2008).
- 5 Dadachova, E. *et al.* Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. *PloS One* **2**, e457, (2007).
- 6 Kurtzman, C., Fell, J. W. & Boekhout, T. *The yeasts: a taxonomic study*. Vol. 1 (Elsevier, 2011).
- 7 Moore, D. *Slayers, Saviors, Servants, and Sex: An Exposé of Kingdom Fungi*. (Springer New York, 2001).
- 8 Herrera, C. M. & Pozo, M. I. Nectar yeasts warm the flowers of a winter-blooming plant. *Proceedings of the Royal Society B: Biological Sciences* **277**, 1827-1834, (2010).
- 9 Vega, F. E. & Dowd, P. F. in *Insect-Fungal Associations: Ecology and Evolution* 211-243 (2005).
- 10 Botha, A. The importance and ecology of yeasts in soil. *Soil Biology and Biochemistry* **43**, 1-8, (2011).
- 11 Lutzoni, F. & Miadlikowska, J. Lichens. *Current Biology* **19**, R502-R503, (2009).
- 12 Hazen, K. C. New and emerging yeast pathogens. *Clinical Microbiology Reviews* **8**, 462-478, (1995).
- 13 Herskowitz, I. Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiological Reviews* **52**, 536, (1988).
- 14 Cavalieri, D., McGovern, P. E., Hartl, D. L., Mortimer, R. & Polsinelli, M. Evidence for *S. cerevisiae* fermentation in ancient wine. *Journal of Molecular Evolution* **57 Suppl 1**, S226-232, (2003).
- 15 Cereghino, G. P. L. & Cregg, J. M. Applications of yeast in biotechnology: protein production and genetic analysis. *Current Opinion in Biotechnology* **10**, 422-427, (1999).
- 16 McCormack, P., Wildman, H. & Jeffries, P. Production of antibacterial compounds by phylloplane-inhabiting yeasts and yeastlike fungi. *Applied and Environmental Microbiology* **60**, 927-931, (1994).
- 17 Kurtzman, C. P. & Fell, J. W. in *Biodiversity and Ecophysiology of Yeasts* 11-30 (Springer, 2006).

- 18 Sunnerhagen, P., Piskur, J., Kurtzman, C. & Piškur, J. in *Comparative Genomics* Vol. 15 *Topics in Current Genetics* 29-46 (Springer Berlin Heidelberg, 2006).
- 19 Martini, A. Origin and domestication of the wine yeast *Saccharomyces cerevisiae*. *Journal of Wine research* **4**, 165-176, (1993).
- 20 Stefanini, I. *et al.* Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. *Proceedings of the National Academy of Sciences* **109**, 13398-13403, (2012).
- 21 Goffeau, A. *et al.* Life with 6000 Genes. *Science* **274**, 546-567, (1996).
- 22 Hinnen, A., Hicks, J. B. & Fink, G. R. Transformation of yeast. *Proceedings of the National Academy of Sciences* **75**, 1929-1933, (1978).
- 23 Annaluru, N. *et al.* Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55-58, (2014).
- 24 Brückner, S. & Mösche, H.-U. Choosing the right lifestyle: adhesion and development in *Saccharomyces cerevisiae*. *FEMS Microbiology Reviews* **36**, 25-58, (2012).
- 25 Marton, M. J. *et al.* Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nature Medicine* **4**, 1293-1301, (1998).
- 26 Melese, T. & Hieter, P. From genetics and genomics to drug discovery: yeast rises to the challenge. *Trends in Pharmacological Sciences* **23**, 544-547, (2002).
- 27 Gimeno, C. J., Ljungdahl, P. O., Styles, C. A. & Fink, G. R. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* **68**, 1077-1090, (1992).
- 28 Roberts, R. L. & Fink, G. R. Elements of a single MAP kinase cascade in *Saccharomyces cerevisiae* mediate two developmental programs in the same cell type: mating and invasive growth. *Genes and Development* **8**, 2974-2985, (1994).
- 29 Lindquist, W. Cell surface constituents and yeast flocculation. *Nature* **170**, 544-545, (1952).
- 30 Martinez, P., Codon, A., Pérez, L. & Benitez, T. Physiological and molecular characterization of flor yeasts: polymorphism of flor yeast populations. *Yeast* **11**, 1399-1411, (1995).
- 31 Reynolds, T. B. & Fink, G. R. Bakers'yeast, a model for fungal biofilm formation. *Science* **291**, 878-881, (2001).
- 32 Oyeka, C. & Ugwu, L. Fungal flora of human toe webs. *Mycoses* **45**, 488-491, (2002).
- 33 Bolotin-Fukuhara, M. & Fairhead, C. *Candida glabrata*: a deadly companion? *Yeast* **31**, 279-288, (2014).
- 34 Kaur, R., Domergue, R., Zupancic, M. L. & Cormack, B. P. A yeast by any other name: *Candida glabrata* and its interaction with the host. *Current Opinion in Microbiology* **8**, 378-384, (2005).
- 35 Dujon, B. *et al.* Genome evolution in yeasts. *Nature* **430**, 35 - 44, (2004).

- 36 Gow, N. A., Brown, A. J. & Odds, F. C. Fungal morphogenesis and host invasion. *Current Opinion in Microbiology* **5**, 366-371, (2002).
- 37 Gabaldon, T. *et al.* Comparative genomics of emerging pathogens in the *Candida glabrata* clade. *BMC Genomics* **14**, 623, (2013).
- 38 Bölker, M. *Ustilago maydis*—a valuable model system for the study of fungal dimorphism and virulence. *Microbiology* **147**, 1395-1401, (2001).
- 39 Ernst, J. F. & Schmidt, A. *Dimorphism in human pathogenic and apathogenic yeasts*. Vol. 5 (Karger Medical and Scientific Publishers, 2000).
- 40 Walker, L. *et al.* Anti-*Saccharomyces cerevisiae* antibodies (ASCA) in Crohn's disease are associated with disease severity but not NOD2/CARD15 mutations. *Clinical and Experimental Immunology* **135**, 490-496, (2004).
- 41 Main, J. *et al.* Antibody to *Saccharomyces cerevisiae* (bakers' yeast) in Crohn's disease. *BMJ* **297**, 1105-1106, (1988).
- 42 Cletus P, K. in *The Yeasts (Fifth Edition)* 491-495 (Elsevier, 2011).
- 43 Klier, K. Methanol synthesis. *Advances in Catalysis* **31**, 243-313, (1982).
- 44 Wegner, G. H. Emerging applications of the methylotrophic yeasts. *FEMS Microbiology Reviews* **7**, 279-283, (1990).
- 45 Olah, G. A., Goepfert, A. & Prakash, G. K. S. *Beyond oil and gas : the methanol economy*. 2., updated and enl. edn, (Wiley-VCH-Verl., 2009).
- 46 Kurtzman, C. P. & Robnett, C. J. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie van Leeuwenhoek* **73**, 331-371, (1998).
- 47 Yamada, Y., Matsuda, M., Maeda, K. & Mikata, K. The phylogenetic relationships of methanol-assimilating yeasts based on the partial sequences of 18S and 26S ribosomal RNAs: the proposal of *Komagataella* gen. nov. (Saccharomycetaceae). *Bioscience, Biotechnology, and Biochemistry* **59**, 439-444, (1995).
- 48 Kurtzman, C. P. Three new ascomycetous yeasts from insect-associated arboreal habitats. *Canadian Journal of Microbiology* **46**, 50-58, (2000).
- 49 Cregg, J. M. Introduction: distinctions between *Pichia pastoris* and other expression systems. *Methods in molecular biology (Clifton, NJ)* **389**, 1-10, (2006).
- 50 Mattanovich, D. *et al.* Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microbial Cell Factories* **8**, 29, (2009).
- 51 De Schutter, K. *et al.* Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nature Biotechnology* **27**, 561-566, (2009).
- 52 Küberl, A. *et al.* High-quality genome sequence of *Pichia pastoris* CBS7435. *Journal of Biotechnology* **154**, 312-320, (2011).
- 53 Mattanovich, D. *et al.* Open access to sequence: Browsing the *Pichia pastoris* genome. *Microbial Cell Factories* **8**, 53, (2009).

- 54 Phaff, H. J. & Knapp, E. P. The taxonomy of yeasts found in exudates of certain trees and other natural breeding sites of some species of *Drosophila*. *Antonie van Leeuwenhoek* **22**, 117-130, (1956).
- 55 Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology* **57**, 233-266, (2006).
- 56 Bais, H. P., Broeckling, C. D., Vivanco, J. M. & Karlovsky, P. Vol. 14 *Soil Biology* 241-252 (Springer Berlin Heidelberg, 2008).
- 57 Tatum, E. L. Nutritional Requirements of *Drosophila Melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* **25**, 490-497, (1939).
- 58 Baumberg, J. P. A nutritional study of insects, with special reference to microorganisms and their substrata. *Journal of Experimental Zoology* **28**, 1-81, (1919).
- 59 Suh, S.-O., McHugh, J. V., Pollock, D. D. & Blackwell, M. The beetle gut: a hyperdiverse source of novel yeasts. *Mycological Research* **109**, 261-265, (2005).
- 60 Suh, S.-O. & Blackwell, M. Three new beetle-associated yeast species in the *Pichia guilliermondii* clade. *FEMS Yeast Research* **5**, 87-95, (2004).
- 61 Barata, A., Santos, S. C., Malfeito-Ferreira, M. & Loureiro, V. New Insights into the Ecological Interaction Between Grape Berry Microorganisms and *Drosophila* Flies During the Development of Sour Rot. *Microbial Ecology* **64**, 416-430, (2012).
- 62 Gibson, C. M. & Hunter, M. S. Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecology Letters* **13**, 223-234, (2010).
- 63 Cereghino, G., Cereghino, J., Ilgen, C. & Cregg, J. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Current Opinion in Biotechnology* **13**, 329 - 332, (2002).
- 64 Hamilton, S. *et al.* Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**, 1441 - 1443, (2006).
- 65 Hamilton, S. & Gerngross, T. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Current Opinion in Biotechnology* **18**, 387 - 392, (2007).
- 66 Jacobs, P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nature Protocols* **4**, 58 - 70, (2009).
- 67 Ballou, C. E. Isolation, characterization, and properties of *Saccharomyces cerevisiae* mnn mutants with nonconditional protein glycosylation defects. *Methods in Enzymology* **185**, 440-470, (1990).
- 68 Daly, R. & Hearn, M. T. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *Journal of Molecular Recognition* **18**, 119-138, (2005).

- 69 Cregg, J. M. *et al.* Expression in the yeast *Pichia pastoris*. *Methods in Enzymology* **463**, 169-189, (2009).
- 70 Smith, A. E., Zhang, Z., Thomas, C. R., Moxham, K. E. & Middelberg, A. P. The mechanical properties of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* **97**, 9871-9874, (2000).
- 71 Orlean, P. Architecture and Biosynthesis of the *Saccharomyces cerevisiae* Cell Wall. *Genetics* **192**, 775-818, (2012).
- 72 Cappellaro, C., Baldermann, C., Rachel, R. & Tanner, W. Mating type-specific cell-cell recognition of *Saccharomyces cerevisiae*: cell wall attachment and active sites of α - and α -agglutinin. *EMBO Journal* **13**, 4737-4744, (1994).
- 73 Osumi, M. The ultrastructure of yeast: cell wall structure and formation. *Micron* **29**, 207-233, (1998).
- 74 Zlotnik, H., Fernandez, M. P., Bowers, B. & Cabib, E. *Saccharomyces cerevisiae* mannoproteins form an external cell wall layer that determines wall porosity. *Journal of Bacteriology* **159**, 1018-1026, (1984).
- 75 Kopecká, M., Phaff, H. & Fleet, G. Demonstration of a fibrillar component in the cell wall of the yeast *Saccharomyces cerevisiae* and its chemical nature. *The Journal of cell biology* **62**, 66-76, (1974).
- 76 Gabius, H.-J. *The sugar code: fundamentals of glycosciences*. (John Wiley & Sons, 2011).
- 77 Holan, Z. *et al.* The glucan-chitin complex in *Saccharomyces cerevisiae*. *Archives of Microbiology* **130**, 312-318, (1981).
- 78 Rinaudo, M. Chitin and chitosan: properties and applications. *Progress in polymer science* **31**, 603-632, (2006).
- 79 Dweltz, N., Colvin, J. R. & McInnes, A. Studies on chitan (β -(1 \rightarrow 4)-linked 2-acetamido-2-deoxy-D-glucan) fibers of the diatom *Thalassiosira fluviatilis*, Hustedt. III. The structure of chitan from X-ray diffraction and electron microscope observations. *Canadian Journal of Chemistry* **46**, 1513-1521, (1968).
- 80 Vermeulen, C. A. & Wessels, J. G. H. Chitin biosynthesis by a fungal membrane preparation. *European Journal of Biochemistry* **158**, 411-415, (1986).
- 81 Klis, F. M., Mol, P., Hellingwerf, K. & Brul, S. Dynamics of cell wall structure in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev* **26**, 239-256, (2002).
- 82 Rössle, M. *et al.* Fast intracrystalline hydration of β -chitin revealed by combined microdrop generation and on-line synchrotron radiation microdiffraction. *Biomacromolecules* **4**, 981-986, (2003).
- 83 Saito, Y., Putaux, J., Okano, T., Gaill, F. & Chanzy, H. Structural aspects of the swelling of β chitin in HCl and its conversion into α chitin. *Macromolecules* **30**, 3867-3873, (1997).

- 84 Dallies, N., Francois, J. & Paquet, V. A new method for quantitative determination of polysaccharides in the yeast cell wall. Application to the cell wall defective mutants of *Saccharomyces cerevisiae*. *Yeast* **14**, 1297-1306, (1998).
- 85 Kollár, R., Petráková, E., Ashwell, G., Robbins, P. W. & Cabib, E. Architecture of the Yeast Cell Wall - The Linkage between Chitin and β 1,3-Glucan. *Journal of Biological Chemistry* **270**, 1170-1178, (1995).
- 86 Kollár, R. *et al.* Architecture of the Yeast Cell Wall β (1 \rightarrow 6)-glucan interconnects mannoprotein, β (1 \rightarrow 3)-glucan, and chitin. *Journal of Biological Chemistry* **272**, 17762-17775, (1997).
- 87 Young, S.-H. & Jacobs, R. R. Sodium hydroxide-induced conformational change in schizophyllan detected by the fluorescence dye, aniline blue. *Carbohydrate Research* **310**, 91-99, (1998).
- 88 Morris, G., Winters, L., Coulson, G. & Clarke, K. Effect of osmotic stress on the ultrastructure and viability of the yeast *Saccharomyces cerevisiae*. *Journal of General Microbiology* **132**, 2023-2034, (1986).
- 89 Manners, D. J., Masson, A. J. & Patterson, J. C. The structure of a beta-(1 \rightarrow 3)-D-glucan from yeast cell walls. *Biochemical Journal* **135**, 19-30, (1973).
- 90 Gentsch, M. & Tanner, W. The PMT gene family: protein O-glycosylation in *Saccharomyces cerevisiae* is vital. *The EMBO Journal* **15**, 5752, (1996).
- 91 Gonzalez, M., Lipke, P. N. & Ovalle, R. GPI proteins in biogenesis and structure of yeast cell walls. *The Enzymes* **26**, 321-356, (2009).
- 92 Fujita, M. & Kinoshita, T. Structural remodeling of GPI anchors during biosynthesis and after attachment to proteins. *FEBS Letters* **584**, 1670-1677, (2010).
- 93 Ecker, M., Deutzmann, R., Lehle, L., Mrsa, V. & Tanner, W. Pir proteins of *Saccharomyces cerevisiae* are attached to β -1, 3-glucan by a new protein-carbohydrate linkage. *Journal of Biological Chemistry* **281**, 11523-11529, (2006).
- 94 Nuoffer, C., Horvath, A. & Riezman, H. Analysis of the sequence requirements for glycosylphosphatidylinositol anchoring of *Saccharomyces cerevisiae* Gas1 protein. *Journal of Biological Chemistry* **268**, 10558-10563, (1993).
- 95 Alexopoulos, C. J., Mims, C. W. & Blackwell, M. *Introductory mycology*. 4. edn, (Wiley, 1996).
- 96 Thieme, T. R. & Ballou, C. E. Nature of the phosphodiester linkage of the phosphomannan from the yeast *Kloeckera brevis*. *Biochemistry* **10**, 4121-4129, (1971).
- 97 Kopecká, M. Yeast and fungal cell-wall polysaccharides can self-assemble in vitro into an ultrastructure resembling in vivo yeast cell walls. *Microscopy* **62**, 327-339, (2013).
- 98 Lipke, P. & Kurjan, J. Sexual agglutination in budding yeasts: structure, function, and regulation of adhesion glycoproteins. *Microbiological Reviews* **56**, 180-194, (1992).

- 99 Dranginis, A. M., Rauceo, J. M., Coronado, J. E. & Lipke, P. N. A biochemical guide to yeast adhesins: glycoproteins for social and antisocial occasions. *Microbiology and molecular biology reviews : MMBR* **71**, 282-294, (2007).
- 100 Verstrepen, K. J., Reynolds, T. B. & Fink, G. R. Origins of variation in the fungal cell surface. *Nature Reviews Microbiology* **2**, 533-540, (2004).
- 101 Jentoft, N. Why are proteins O-glycosylated? *Trends in Biochemical Sciences* **15**, 291-294, (1990).
- 102 Ramsook, C. B. *et al.* Yeast cell adhesion molecules have functional amyloid-forming sequences. *Eukaryotic cell* **9**, 393-404, (2010).
- 103 Lipke, P. N. *et al.* Between Amyloids and Aggregation Lies a Connection with Strength and Adhesion. *New Journal of Science* **2014**, 815102, (2014).
- 104 Garcia, M. C. *et al.* A role for amyloid in cell aggregation and biofilm formation. *PloS One* **6**, e17632, (2011).
- 105 Meem, M. H. & Cullen, P. J. The impact of protein glycosylation on Flo11-dependent adherence in *Saccharomyces cerevisiae*. *FEMS Yeast Research* **12**, 809-818, (2012).
- 106 Verstrepen, K. J. & Klis, F. M. Flocculation, adhesion and biofilm formation in yeasts. *Molecular Microbiology* **60**, 5-15, (2006).
- 107 Veelders, M. *et al.* Structural basis of flocculin-mediated social behavior in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22511-22516, (2010).
- 108 Maestre-Reyna, M. *et al.* Structural basis for promiscuity and specificity during *Candida glabrata* invasion of host epithelia. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16864-16869, (2012).
- 109 Hoyer, L. L. & Hecht, J. E. The ALS5 gene of *Candida albicans* and analysis of the Als5p N-terminal domain. *Yeast (Chichester, England)* **18**, 49-60, (2001).
- 110 Kraushaar, T. *et al.* Interactions by the fungal Flo11 adhesin depend on a fibronectin type III-like adhesin domain girdled by aromatic bands. *Structure, in press*, (2015).
- 111 de Groot, P. W. J. & Klis, F. M. The conserved PA14 domain of cell wall-associated fungal adhesins governs their glycan-binding specificity. *Molecular Microbiology* **68**, 535-537, (2008).
- 112 Christiaens, J. F. *et al.* Functional divergence of gene duplicates through ectopic recombination. *EMBO Reports* **13**, 1145-1151, (2012).
- 113 Teunissen, A. W. & Steensma, H. Y. The Dominant Flocculation Genes of *Saccharomyces cerevisiae* Constitute a New Subtelomeric Gene Family. *Yeast* **11**, 1001-1013, (1995).
- 114 Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* **63**, 751-762, (1990).

- 115 Halme, A., Bumgarner, S., Styles, C. & Fink, G. R. Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell* **116**, 405-415, (2004).
- 116 Yoshida, E. *et al.* Role of a PA14 domain in determining substrate specificity of a glycoside hydrolase family 3 beta-glucosidase from *Kluyveromyces marxianus*. *Biochemical Journal* **431**, 39-49, (2010).
- 117 Johan, L. *et al.* Structural and enzymatic characterization of a glycoside hydrolase family 31 alpha-xylosidase from *Cellvibrio japonicus* involved in xyloglucan saccharification. *Biochemical Journal* **436**, 567-580, (2011).
- 118 Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H. & Liddington, R. C. Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833-838, (1997).
- 119 Rigden, D. J., Mello, L. V. & Galperin, M. Y. The PA14 domain, a conserved all- β domain in bacterial toxins, enzymes, adhesins and signaling molecules. *Trends in Biochemical Sciences* **29**, 335-339, (2004).
- 120 Linder, T. & Gustafsson, C. M. Molecular phylogenetics of ascomycotal adhesins — A novel family of putative cell-surface adhesive proteins in fission yeasts. *Fungal Genetics and Biology* **45**, 485-497, (2008).
- 121 Finn, R. D. *et al.* The Pfam protein families database. *Nucleic Acids Research* **38**, D211-D222, (2010).
- 122 Stewart, G. & Goring, T. EFFECT OF SOME MONOVALENT AND DIVALENT METAL IONS ON THE FLOCCULATION OF BREWERS YEAST STRAINS*. *Journal of the Institute of Brewing* **82**, 341-342, (1976).
- 123 Stratford, M. Evidence for two mechanisms of flocculation in *Saccharomyces cerevisiae*. *Yeast (Chichester, England)* **5**, S441-445, (1989).
- 124 Watari, J. *et al.* Molecular cloning and analysis of the yeast flocculation gene FLO1. *Yeast* **10**, 211-225, (1994).
- 125 Verstrepen, K. *et al.* Late fermentation expression of FLO1 in *Saccharomyces cerevisiae*. *Journal of the american society of brewing chemists* **59**, 69-76, (2001).
- 126 Govender, P., Domingo, J. L., Bester, M. C., Pretorius, I. S. & Bauer, F. F. Controlled expression of the dominant flocculation genes FLO1, FLO5, and FLO11 in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **74**, 6041-6052, (2008).
- 127 Smukalla, S. *et al.* FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726-737, (2008).
- 128 Dawkins, R. *The Selfish Gene*. (Oxford Univ. Press, 1976).
- 129 Gardner, A. & West, S. A. Greenbeards. *Evolution* **64**, 25-38, (2010).
- 130 West, S. A. & Gardner, A. Altruism, spite, and greenbeards. *Science* **327**, 1341-1344, (2010).

- 131 Verstrepen, K. J., Derdelinckx, G., Verachtert, H. & Delvaux, F. R. Yeast flocculation: what brewers should know. *Applied Microbiology and Biotechnology* **61**, 197-205, (2003).
- 132 Kobayashi, O., Hayashi, N., Kuroki, R. & Sone, H. Region of Flo1 proteins responsible for sugar recognition. *Journal of Bacteriology* **180**, 6503-6510, (1998).
- 133 Guo, B., Styles, C. A., Feng, Q. & Fink, G. R. A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating. *Proceedings of the National Academy of Sciences*, 220420397, (2000).
- 134 Javadekar, V. S., Sivaraman, H., Sainkar, S. R. & Khan, M. I. A mannose-binding protein from the cell surface of flocculent *Saccharomyces cerevisiae* (NCIM 3528): its role in flocculation. *Yeast* **16**, 99-110, (2000).
- 135 Sim, L., Groes, M., Olesen, K. & Henriksen, A. Structural and biochemical characterization of the N-terminal domain of flocculin Lg-Flo1p from *Saccharomyces pastorianus* reveals a unique specificity for phosphorylated mannose. *FEBS Journal* **280**, 1073-1083, (2013).
- 136 Ielasi, F. S., Goyal, P., Sleutel, M., Wohlkonig, A. & Willaert, R. G. The mannose-specific lectin domains of Flo1p from *Saccharomyces cerevisiae* and Lg-Flo1p from *S. pastorianus*: crystallization and preliminary X-ray diffraction analysis of the adhesin-carbohydrate complexes. *Acta Crystallographica. Section F: Structural Biology and Crystallization Communications* **69**, 779-782, (2013).
- 137 Cormack, B. P., Ghori, N. & Falkow, S. An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. *Science* **285**, 578-582, (1999).
- 138 Zupancic, M. L. *et al.* Glycan microarray analysis of *Candida glabrata* adhesin ligand specificity. *Molecular Microbiology* **68**, 547-559, (2008).
- 139 De Las Peñas, A. *et al.* Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. *Genes and Development* **17**, 2245-2258, (2003).
- 140 Castano, I. *et al.* Telomere length control and transcriptional regulation of subtelomeric adhesins in *Candida glabrata*. *Molecular Microbiology* **55**, 1246-1258, (2005).
- 141 Domergue, R. *et al.* Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTI. *Science* **308**, 866-870, (2005).
- 142 Gipson, I. K. *et al.* Mucin genes expressed by human female reproductive tract epithelia. *Biology of Reproduction* **56**, 999-1011, (1997).
- 143 Bartman, A. E. *et al.* The MUC6 secretory mucin gene is expressed in a wide variety of epithelial tissues. *The Journal of pathology* **186**, 398-405, (1998).
- 144 Russo, C. L. *et al.* Mucin gene expression in human male urogenital tract epithelia. *Human Reproduction* **21**, 2783-2793, (2006).

- 145 Desai, C., Mavrianos, J. & Chauhan, N. *Candida glabrata* Pwp7p and Aed1p are required for adherence to human endothelial cells. *FEMS Yeast Research* **11**, 595-601, (2011).
- 146 Hoyer, L., Payne, T. & Hecht, J. Identification of *Candida albicans* ALS2 and ALS4 and Localization of Als Proteins to the Fungal Cell Surface. *Journal of Bacteriology* **180**, 5334-5343, (1998).
- 147 Stratford, M. Lectin-mediated aggregation of yeasts—yeast flocculation. *Biotechnology and Genetic Engineering Reviews* **10**, 283-342, (1992).
- 148 Vallejo, J., Sánchez-Pérez, A., Martínez, J. P. & Villa, T. Cell aggregations in yeasts and their applications. *Applied Microbiology and Biotechnology* **97**, 2305-2318, (2013).
- 149 Teixeira, J., Oliveira, R., Azeredo, J., Sousa, M. & Sil, C. Cell wall surface properties and flocculence of a *Kluyveromyces marxianus* strain. *Colloids and surfaces B: Biointerfaces* **5**, 197-203, (1995).
- 150 Al-Mahmood, S., Colin, S. & Bonaly, R. *Kluyveromyces bulgaricus* yeast lectins. Isolation of two galactose-specific lectin forms from the yeast cell wall. *Journal of Biological Chemistry* **266**, 20882-20887, (1991).
- 151 Hussain, T., Salhi, O., Lematre, J., Charpentier, C. & Bonaly, R. Comparative studies of flocculation and deflocculation of *Saccharomyces uvarum* and *Kluyveromyces bulgaricus*. *Applied Microbiology and Biotechnology* **23**, 269-273, (1986).
- 152 Matsuzawa, T., Morita, T., Tanaka, N., Tohda, H. & Takegawa, K. Identification of a galactose-specific flocculin essential for non-sexual flocculation and filamentous growth in *Schizosaccharomyces pombe*. *Molecular Microbiology* **82**, 1531-1544, (2011).
- 153 Mbawala, A., Al Mahmood, S., Loppinet, V. & Bonaly, R. Acetolysis and ¹H NMR studies on mannans isolated from very flocculent and weakly flocculent cells of *Pichia pastoris* IFP 206. *Journal of General Microbiology* **136**, 1279-1284, (1990).
- 154 Schlereth, J. *Strukturelle und funktionelle Charakterisierung von putativen Adhäsionsdomänen aus Pichia pastoris* Diplom thesis, (2010).
- 155 Kock, M. *Die zuckerbindende Domäne von putativen Adhäsinen aus Pichia pastoris - Eine funktionelle und strukturelle Charakterisierung* M. Sc. thesis, Philipps-Universität Marburg, (2012).
- 156 Schlereth, J. *Strukturelle und funktionelle Charakterisierung von putativen Adhäsionsdomänen aus Pichia pastoris* Diplom thesis, Philipps-Universität, (2010).
- 157 Kock, M. *et al.* A member of a hitherto unknown subfamily of ascomycetal adhesins binds to chitin ends. *in preparation*, (2015).

- 158 Maestre-Reyna, M. *Structural and functional studies of mucin-interacting adhesion domains from Candida glabrata & Helicobacter pylori* Dr. rer. nat. thesis, Philipps-Universität, (2011).
- 159 Wozniak, N. *in preparation* Dr. rer. nat. thesis, Philipps-Universität, (2015).
- 160 Untergasser, A. *et al.* Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* **35**, W71-W74, (2007).
- 161 Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130, (1986).
- 162 Dubendorff, J. W. & Studier, F. W. Controlling Basal Expression in an Inducible T 7 Expression System by Blocking the Target T 7 Promotor with Lac Repressor. *Journal of Molecular Biology* **219**, 45-59, (1991).
- 163 Gietz, R. D. & Akio, S. New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**, 527-534, (1988).
- 164 Diderrich, R. *et al.* Structural and Functional Evolution of the Candida Glabrata Epithelial Adhesin Family. *in preparation*, (2015).
- 165 Diderrich, R. *Strukturelle und funktionelle Charakterisierung der epithelialen Adhäsine aus Candida glabrata* Dr. rer. nat. thesis, Philipps-Universität Marburg, (2014).
- 166 Taylor, R. G., Walker, D. C. & McInnes, R. R. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research* **21**, 1677-1678, (1993).
- 167 Lobstein, J. *et al.* SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories* **11**, 56, (2012).
- 168 Liu, H., Styles, C. A. & Fink, G. R. Saccharomyces cerevisiae S288C Has a Mutation in FLO8, a Gene Required for Filamentous Growth. *Genetics* **144**, 967-978, (1996).
- 169 Grundmann, O., Mösch, H.-U. & Braus, G. H. Repression of GCN4 mRNA Translation by Nitrogen Starvation in Saccharomyces cerevisiae. *Journal of Biological Chemistry* **276**, 25661-25671, (2001).
- 170 Schauer, C., Thompson, C. L. & Brune, A. The Bacterial Community in the Gut of the Cockroach Shelfordella lateralis Reflects the Close Evolutionary Relatedness of Cockroaches and Termites. *Applied and Environmental Microbiology* **78**, 2758-2767, (2012).
- 171 Altschul, S., Gish, W., Miller, W., Myers, E. & Lipman, D. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403 - 410, (1990).
- 172 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402, (1997).

- 173 Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* **89**, 10915-10919, (1992).
- 174 Staden, R. A strategy of DNA sequencing employing computer programs. *Nucleic Acids Research* **6**, 2601-2610, (1979).
- 175 Stranneheim, H. & Lundeberg, J. Stepping stones in DNA sequencing. *Biotechnology Journal* **7**, 1063-1073, (2012).
- 176 Sonnhammer, E. L. L., Eddy, S. R. & Durbin, R. Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins: Structure, Function, and Genetics* **28**, 405-420, (1997).
- 177 Hunter, S. *et al.* InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Research* **40**, D306-D312, (2012).
- 178 Consortium, T. U. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Research* **42**, D191-D198, (2014).
- 179 Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**, 205-217, (2000).
- 180 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780, (2013).
- 181 Katoh, K. & Frith, M. C. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28**, 3144-3146, (2012).
- 182 Liò, P. & Goldman, N. Models of Molecular Evolution and Phylogeny. *Genome Research* **8**, 1233-1244, (1998).
- 183 Tamura, K., Stecher, G., Peterson, D., FilipSKI, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729, (2013).
- 184 Johnson, J. B. & Omland, K. S. Model selection in ecology and evolution. *Trends in Ecology & Evolution* **19**, 101-108, (2004).
- 185 Le, S. Q. & Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution* **25**, 1307-1320, (2008).
- 186 Kubatko, L. S. in *Tutorials in Mathematical Biosciences IV* Vol. 1922 *Lecture Notes in Mathematics* (ed Avner Friedman) Ch. 1, 1-38 (Springer Berlin Heidelberg, 2008).
- 187 Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425, (1987).
- 188 Salemi, M., Lemey, P. & Vandamme, A.-M. *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. (Cambridge University Press, 2009).

- 189 Hall, B. G. Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution* **30**, 1229-1235, (2013).
- 190 Bertani. The Mode of Phage Liberation by Lysogenic *Escherichia coli*. *Journal of Bacteriology* **62**, 293, (1951).
- 191 Tartof, K. D. & Hobbs, C. A. Improved media for growing plasmid and cosmid clones. *Focus* **9**, 12, (1987).
- 192 Amberg, D. C., Burke, D., Strathern, J. N. & Laboratory, C. S. H. *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. (Cold Spring Harbor Laboratory Press, 2005).
- 193 Morrison, D. A. Transformation in *Escherichia coli*: cryogenic preservation of competent cells. *Journal of Bacteriology* **132**, 349-351, (1977).
- 194 Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53**, 159-162, (1970).
- 195 Chen, D.-C., Yang, B.-C. & Kuo, T.-T. One-step transformation of yeast in stationary phase. *Current Genetics* **21**, 83-84, (1992).
- 196 Vogelstein, B. & Gillespie, D. Preparative and Analytical Purification of DNA from Agarose. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 615-619, (1979).
- 197 Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513-1523, (1979).
- 198 QIAGEN. *QIAprep® Miniprep Handbook*. (2012).
- 199 Mülhardt, C. *Der Experimentator: Molekularbiologie/ Genomics*. Vol. 24 (Spektrum Akademischer Verlag, 2009).
- 200 Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**, 156-159, (1987).
- 201 Gallagher, S. R. & Desjardins, P. R. in *Current Protocols in Protein Science* (John Wiley & Sons, Inc., 2001).
- 202 Gomperts, B. D., Kramer, I. M. & Tatham, P. E. R. in *Signal Transduction* (ed Bastien D. Gomperts Ijsbrand M. Kramer Peter E. R. Tatham) 1-17 (Academic Press, 2002).
- 203 Krauss, G. in *Biochemistry of Signal Transduction and Regulation* 89-118 (Wiley-VCH Verlag GmbH, 2001).
- 204 Ståhlberg, A., Zoric, N., Åman, P. & Kubista, M. Quantitative real-time PCR for cancer detection: the lymphoma case. *Expert Review of Molecular Diagnostics* **5**, 221-230, (2005).
- 205 Abbott, M. A. *et al.* Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in vitro. *Journal of Infectious Diseases* **158**, 1158-1169, (1988).

- 206 Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**, 335-350, (1987).
- 207 Temin, H. M. & Mizutani, S. Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. *Nature* **226**, 1211-1213, (1970).
- 208 Baltimore, D. Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* **226**, 1209-1211, (1970).
- 209 Shaffer, A. L., Wojnar, W. & Nelson, W. Amplification, detection, and automated sequencing of gibbon interleukin-2 mRNA by *Thermus aquaticus* DNA polymerase reverse transcription and polymerase chain reaction. *Analytical Biochemistry* **190**, 292-296, (1990).
- 210 Zipper, H., Brunner, H., Bernhagen, J. & Vitzthum, F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Research* **32**, e103, (2004).
- 211 Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**, e45, (2001).
- 212 Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402-408, (2001).
- 213 Thellin, O. *et al.* Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* **75**, 291-295, (1999).
- 214 VanGuilder, H. D., Vrana, K. E. & Freeman, W. M. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques* **44**, 619-626, (2008).
- 215 Thornton, B. & Basu, C. Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education* **39**, 145-154, (2011).
- 216 Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599, (1975).
- 217 Qiagen. *The QIAexpressionist*. (2003).
- 218 Porath, J. Gel filtration: a method for desalting and group separation. *Nature* **183**, 1657-1659, (1959).
- 219 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685, (1970).
- 220 Bennett, J. & Scott, K. J. Quantitative staining of fraction I protein in polyacrylamide gels using Coomassie brilliant blue. *Analytical Biochemistry* **43**, 173-182, (1971).
- 221 Warburg, O. & Christian, W. Isolierung und Kristallisation des Gärungsferments Enolase. *Naturwissenschaften* **29**, 589-590, (1941).

- 222 Gill, S. C. & von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* **182**, 319-326, (1989).
- 223 Gasteiger, E. *et al.* Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*, 571–607, (2005).
- 224 Panchuk-Voloshina, N. *et al.* Alexa Dyes, a Series of New Fluorescent Dyes that Yield Exceptionally Bright, Photostable Conjugates. *Journal of Histochemistry and Cytochemistry* **47**, 1179-1188, (1999).
- 225 Berlier, J. E. *et al.* Quantitative Comparison of Long-wavelength Alexa Fluor Dyes to Cy Dyes: Fluorescence of the Dyes and Their Bioconjugates. *Journal of Histochemistry and Cytochemistry* **51**, 1699-1712, (2003).
- 226 Blixt, O. *et al.* Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17033-17038, (2004).
- 227 Heimbürg-Molinaro, J., Song, X., Smith, D. F. & Cummings, R. D. in *Current Protocols in Protein Science* (John Wiley & Sons, Inc., 2001).
- 228 Breslauer, K. J., Freire, E. & Straume, M. in *Methods in Enzymology* Vol. Volume 211 (ed James E. Dahlberg David M.J. Lilley) 533-567 (Academic Press, 1992).
- 229 Brandts, J. F., Lin, L. N., Wiseman, T., Williston, S. & Yang, C. P. An Instrument for Rapid-Determination of Binding Constants for Biomolecules. *American Laboratory* **22**, 30-&, (1990).
- 230 Wiseman, T., Williston, S., Brandts, J. F. & Lin, L.-N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* **179**, 131-137, (1989).
- 231 Leavitt, S. & Freire, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current Opinion in Structural Biology* **11**, 560-566, (2001).
- 232 Rehm, H. & Letzel, T. *Der Experimentator: Proteinbiochemie, Proteomics*. 6. Aufl.. edn, (Spektrum Akad. Verl., 2010).
- 233 Sigurskjöld, B. W. Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry. *Analytical Biochemistry* **277**, 260-266, (2000).
- 234 Zhang, Y.-L. & Zhang, Z.-Y. Low-Affinity Binding Determined by Titration Calorimetry Using a High-Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B. *Analytical Biochemistry* **261**, 139-148, (1998).
- 235 Bernal, J. D. & Crowfoot, D. X-ray photographs of crystalline pepsin. *Nature* **133**, 794-795, (1934).
- 236 Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-666, (1958).

- 237 Wüthrich, K. The way to NMR structures of proteins. *Nature Structural & Molecular Biology* **8**, 923-925, (2001).
- 238 Lipfert, J. & Doniach, S. Small-Angle X-Ray Scattering from RNA, Proteins, and Protein Complexes. *Annual Review of Biophysics and Biomolecular Structure* **36**, 307-327, (2007).
- 239 Kühlbrandt, W. Cryo-EM enters a new era. *eLife* **3**, (2014).
- 240 Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *Journal of Biological Chemistry* **265**, 22059-22062, (1990).
- 241 Liu, G. *et al.* NMR data collection and analysis protocol for high-throughput protein structure determination. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10487-10492, (2005).
- 242 Buxbaum, E. *Biophysical Chemistry of Proteins*. (Springer, 2011).
- 243 Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology* **10**, 980-980, (2003).
- 244 Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Research*, (2014).
- 245 Schmidt, A., Teeter, M., Weckert, E. & Lamzin, V. S. Crystal structure of small protein crambin at 0.48 Å resolution. *Acta Crystallographica Section F* **67**, 424-428, (2011).
- 246 Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. Extent and nature of contacts between protein molecules in crystal lattices and between subunits of protein oligomers. *Proteins: Structure, Function, and Bioinformatics* **28**, 494-514, (1997).
- 247 Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *Journal of Molecular Biology* **320**, 597-608, (2002).
- 248 Søndergaard, C. R., Garrett, A. E., Carstensen, T., Pollastri, G. & Nielsen, J. E. Structural Artifacts in Protein-Ligand X-ray Structures: Implications for the Development of Docking Scoring Functions. *Journal of Medicinal Chemistry* **52**, 5673-5684, (2009).
- 249 Rupp, B. & Wang, J. Predictive models for protein crystallization. *Methods* **34**, 390-407, (2004).
- 250 Geerlof, A. *et al.* The impact of protein characterization in structural proteomics. *Acta Crystallographica Section D* **62**, 1125-1136, (2006).
- 251 Taguchi, G. *System of experimental design: engineering methods to optimize quality and minimize costs*. (UNIPUB/Kraus International Publications, 1987).
- 252 Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography* **24**, 409-411, (1991).

- 253 Stevens, R. C., Yokoyama, S. & Wilson, I. A. Global Efforts in Structural Genomics. *Science* **294**, 89-92, (2001).
- 254 Blow, N. Structural genomics: inside a protein structure initiative center. *Nat Meth* **5**, 203-207, (2008).
- 255 Davies, D. R. & Segal, D. M. in *Methods in Enzymology* Vol. Volume 22 (ed B. Jakoby William) 266-269 (Academic Press, 1971).
- 256 Saridakis, E. & Chayen, N. E. Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion. *Protein Science* **9**, 755-757, (2000).
- 257 Luft, J. R. *et al.* Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. *Protein Science* **16**, 715-722, (2007).
- 258 Cox, M. J. & Weber, P. C. An Investigation of Protein Crystallization Parameters Using Successive Automated Grid Searches (Sags). *Journal of Crystal Growth* **90**, 318-324, (1988).
- 259 Bergfors, T. Seeds to crystals. *Journal of Structural Biology* **142**, 66-76, (2003).
- 260 Bergfors, T. in *Evolving Methods for Macromolecular Crystallography* Vol. 245 *NATO Science Series* (eds RandyJ Read & Joell Sussman) Ch. 1, 1-10 (Springer Netherlands, 2007).
- 261 D'Arcy, A., Mac Sweeney, A. & Haber, A. Using natural seeding material to generate nucleation in protein crystallization experiments. *Acta Crystallographica Section D* **59**, 1343-1346, (2003).
- 262 Mowbray, S. Macroseeding: a real-life success story. *Protein Crystallization*, 157-162, (1999).
- 263 Stura, E. A. & Wilson, I. A. Applications of the streak seeding technique in protein crystallization. *Journal of Crystal Growth* **110**, 270-282, (1991).
- 264 Luft, J. R. & DeTitta, G. T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D* **55**, 988-993, (1999).
- 265 Hassell, A. M. *et al.* Crystallization of protein-ligand complexes. *Acta Crystallographica Section D* **63**, 72-79, (2007).
- 266 Skarzynski, T. & Thorpe, J. Industrial perspective on X-ray data collection and analysis. *Acta Crystallographica Section D* **62**, 102-107, (2006).
- 267 Henderson, R. *Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction*. Vol. 241 (1990).
- 268 Garman, E. F. & Mitchell, E. P. Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *Journal of Applied Crystallography* **29**, 584-587, (1996).
- 269 Rhodes, G. in *Crystallography Made Crystal Clear (Third Edition)* (ed Gale Rhodes) 49-89 (Academic Press, 2006).
- 270 Drenth, J. *Principles of Protein X-Ray Crystallography*. (Springer, 2007).

- 271 Mueller, U. *et al.* Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *Journal of Synchrotron Radiation* **19**, 442-449, (2012).
- 272 EMBL-Hamburg. *Macromolecular Crystallography*, <http://www.embl-hamburg.de/services/access_infrastructures/mx/index.html> (
- 273 Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography* **67**, 271-281, (2011).
- 274 Powell, H. R., Johnson, O. & Leslie, A. G. W. Autoindexing diffraction images with iMosflm. *Acta Crystallographica Section D* **69**, 1195-1203, (2013).
- 275 Kabsch, W. XDS. *Acta Crystallographica Section D* **66**, 125-132, (2010).
- 276 Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **26**, 795-800, (1993).
- 277 Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D* **66**, 133-144, (2010).
- 278 Evans, P. Scaling and assessment of data quality. *Acta Crystallographica Section D* **62**, 72-82, (2006).
- 279 CCP4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallographica, Section D: Biological Crystallography* **50**, 760-763, (1994).
- 280 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D* **67**, 235-242, (2011).
- 281 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallographica Section D* **69**, 1204-1214, (2013).
- 282 Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS Journal* **280**, 5705-5736, (2013).
- 283 Taylor, G. The phase problem. *Acta Crystallographica Section D* **59**, 1881-1890, (2003).
- 284 Huber, R. Die automatisierte Faltmolekulmethode. *Acta Crystallographica* **19**, 353-356, (1965).
- 285 Rossmann, M. G. & Blow, D. M. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica* **15**, 24-31, (1962).
- 286 Liljas, A. *Textbook of Structural Biology*. (World Scientific, 2009).
- 287 Patterson, A. L. in *Zeitschrift für Kristallographie - Crystalline Materials* Vol. 90 517 (1935).
- 288 DiMaio, F. *et al.* Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* **473**, 540-543, (2011).

- 289 McCoy, A. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D* **63**, 32-41, (2007).
- 290 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674, (2007).
- 291 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213-221, (2010).
- 292 Eswar, N. *et al.* Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* **Chapter 5**, Unit 5 6, (2006).
- 293 Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3**, 1171-1179, (2008).
- 294 Harker, D. in *Advances in Biological and Medical Physics* Vol. Volume 4 (eds Lawrence John H & Tobias Cornelius A) 1-22 (Elsevier, 1956).
- 295 Sheldrick, G. Phase annealing in SHELX-90: direct methods for larger structures. *Acta Crystallographica Section A* **46**, 467-473, (1990).
- 296 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240-255, (1997).
- 297 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126-2132, (2004).
- 298 Kleywegt, G. J. & Jones, T. A. xdlMAPMAN and xdlDATAMAN - Programs for Reformatting, Analysis and Manipulation of Biomacromolecular Electron-Density Maps and Reflection Data Sets. *Acta Crystallographica Section D* **52**, 826-828, (1996).
- 299 Fabiola, F., Korostelev, A. & Chapman, M. S. Bias in cross-validated free R factors: mitigation of the effects of non-crystallographic symmetry. *Acta Crystallographica Section D* **62**, 227-238, (2006).
- 300 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D* **67**, 355-367, (2011).
- 301 Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D* **65**, 1074-1080, (2009).
- 302 Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **28**, 31-36, (1988).
- 303 Schuttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D* **60**, 1355-1363, (2004).

- 304 DeLano, W. L. *The PyMOL molecular graphics system*, 2002).
- 305 Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-1612, (2004).
- 306 Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10037-10041, (2001).
- 307 Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* **372**, 774-797, (2007).
- 308 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675, (2012).
- 309 Sabnis, R. W., Deligeorgiev, T. G., Jachak, M. N. & Dalvi, T. S. DiOC6(3): a useful dye for staining the endoplasmic reticulum. *Biotechnic and Histochemistry* **72**, 253-258, (1997).
- 310 Biela, A. *et al.* Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin. *Angewandte Chemie International Edition* **52**, 1822-1828, (2013).
- 311 Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of Molecular Biology* **179**, 125-142, (1984).
- 312 Latgé, J.-P. The cell wall: a carbohydrate armour for the fungal cell. *Molecular Microbiology* **66**, 279-290, (2007).
- 313 Latgé, J.-P. Tasting the fungal cell wall. *Cellular Microbiology* **12**, 863-872, (2010).
- 314 Cabib, E., Roh, D.-H., Schmidt, M., Crotti, L. B. & Varma, A. The Yeast Cell Wall and Septum as Paradigms of Cell Growth and Morphogenesis. *Journal of Biological Chemistry* **276**, 19679-19682, (2001).
- 315 Coluccio, A. E., Rodriguez, R. K., Kernan, M. J. & Neiman, A. M. The yeast spore wall enables spores to survive passage through the digestive tract of *Drosophila*. *PLoS One* **3**, e2873, (2008).
- 316 Thieme, T. R. & Ballou, C. E. Nature of the phosphodiester linkage of the phosphomannan from the yeast *Kloeckera brevis*. *Biochemistry* **10**, 4121-4129, (1971).
- 317 Griffin, S. R. & Macwilli, I. C. Variation of Cell Wall Content in Flocculent and Non-Flocculent Yeast Strains. *Journal of the Institute of Brewing* **75**, 355-&, (1969).
- 318 Parolis, L. A., Duus, J. O., Parolis, H., Meldal, M. & Bock, K. The extracellular polysaccharide of *Pichia (Hansenula) holstii* NRRL Y-2448: the structure of the phosphomannan backbone. *Carbohydrate Research* **293**, 101-117, (1996).

- 319 Ferro, V., Fewings, K., Palermo, M. C. & Li, C. P. Large-scale preparation of the oligosaccharide phosphate fraction of *Pichia holstii* NRRL Y-2448 phosphomannan for use in the manufacture of PI-88. *Carbohydrate Research* **332**, 183-189, (2001).
- 320 Stein, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography* **41**, 641-643, (2008).
- 321 Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica. Section D: Biological Crystallography* **64**, 61-69, (2008).
- 322 Nieuw Amerongen, A. V., Bolscher, J. G., Bloemena, E. & Veerman, E. C. Sulfomucins in the human body. *Biological Chemistry* **379**, 1-18, (1998).
- 323 Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography* **30**, 1022-1025, (1997).
- 324 Veelders, M. *et al.* Structural basis of flocculin-mediated social behavior in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22511-22516, (2010).
- 325 Verstrepen, K. J. & Klis, F. M. Flocculation, adhesion and biofilm formation in yeasts. *Mol Microbiol* **60**, 5-15, (2006).
- 326 Lai, Y. *et al.* Comparative genomics and transcriptomics analyses reveal divergent lifestyle features of nematode endoparasitic fungus *Hirsutella minnesotensis*. *Genome Biology and Evolution* **6**, 3077-3093, (2014).
- 327 Linder, T. & Gustafsson, C. M. Molecular phylogenetics of ascomycotal adhesins - A novel family of putative cell-surface adhesive proteins in fission yeasts. *Fungal Genetics and Biology* **45**, 485-497, (2008).
- 328 Liu, K. *et al.* *Drechlerella stenobrocha* genome illustrates the mechanism of constricting rings and the origin of nematode predation in fungi. *BMC Genomics* **15**, 114, (2014).
- 329 Yang, J. *et al.* Genomic and proteomic analyses of the fungus *Arthrobotrys oligospora* provide insights into nematode-trap formation. *PLoS Pathogens* **7**, e1002179, (2011).
- 330 Atkinson, H. J., Morris, J. H., Ferrin, T. E. & Babbitt, P. C. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PloS One* **4**, e4345, (2009).
- 331 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**, 2498-2504, (2003).
- 332 Gerlt, J. A. *et al.* Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, (2015).
- 333 De Groot, P. W. *et al.* The cell wall of the human pathogen *Candida glabrata*: differential incorporation of novel adhesin-like wall proteins. *Eukaryotic cell* **7**, 1951-1964, (2008).

- 334 Knezevic, A. *et al.* Variability, heritability and environmental determinants of human plasma N-glycome. *Journal of Proteome Research* **8**, 694-701, (2008).
- 335 Varki, A. Nothing in glycobiology makes sense, except in the light of evolution. *Cell* **126**, 841-845, (2006).
- 336 Gagneux, P. & Varki, A. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**, 747-755, (1999).
- 337 Houle, D. Comparing evolvability and variability of quantitative traits. *Genetics* **130**, 195-204, (1992).
- 338 Brown, R. L. What evolvability really is. *The British Journal for the Philosophy of Science*, axt014, (2013).
- 339 Litman, G. W. *et al.* Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. *Molecular Biology and Evolution* **10**, 60-72, (1993).
- 340 Barry, J. D. & McCulloch, R. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology* **49**, 1-70, (2001).
- 341 The, M. Complete sequence and gene map of a human major histocompatibility complex. *Nature* **401**, 921-923, (1999).
- 342 Zhang, L. *et al.* Screening for Glycosylphosphatidylinositol-Modified Cell Wall Proteins in *Pichia pastoris* and Their Recombinant Expression on the Cell Surface. *Applied and Environmental Microbiology* **79**, 5519-5526, (2013).
- 343 Heger, A. & Holm, L. Rapid automatic detection and alignment of repeats in protein sequences. *Proteins: Structure, Function, and Bioinformatics* **41**, 224-237, (2000).
- 344 Pierleoni, A., Martelli, P. & Casadio, R. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* **9**, 392, (2008).
- 345 Otoo, H. N., Lee, K. G., Qiu, W. & Lipke, P. N. *Candida albicans* Als adhesins have conserved amyloid-forming sequences. *Eukaryotic cell* **7**, 776-782, (2008).
- 346 Ramsook, C. B. *et al.* Yeast cell adhesion molecules have functional amyloid-forming sequences. *Eukaryot Cell* **9**, 393-404, (2010).
- 347 Frank, A. T. *et al.* Structure and Function of Glycosylated Tandem Repeats from *Candida albicans* Als Adhesins. *Eukaryotic Cell* **9**, 405-414, (2010).
- 348 Alsteens, D. *et al.* Unfolding individual als5p adhesion proteins on live cells. *ACS Nano* **3**, 1677-1682, (2009).
- 349 Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **3**, 198-210, (2013).
- 350 Xu, D. & Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715-1735, (2012).

- 351 DeMarco, M. in *Glycoinformatics* Vol. 1273 *Methods in Molecular Biology* (eds Thomas Lütteke & Martin Frank) Ch. 23, 379-390 (Springer New York, 2015).
- 352 Chan, C. X. & Lipke, P. N. Role of force-sensitive amyloid-like interactions in fungal catch bonding and biofilms. *Eukaryot Cell* **13**, 1136-1142, (2014).
- 353 Tormo, M. A., Knecht, E., Gotz, F., Lasa, I. & Penades, J. R. Bap-dependent biofilm formation by pathogenic species of Staphylococcus: evidence of horizontal gene transfer? *Microbiology* **151**, 2465-2475, (2005).
- 354 Cucarella, C. *et al.* Bap, a Staphylococcus aureus surface protein involved in biofilm formation. *Journal of Bacteriology* **183**, 2888-2896, (2001).
- 355 Koide, A., Bailey, C. W., Huang, X. & Koide, S. The fibronectin type III domain as a scaffold for novel binding proteins. *Journal of Molecular Biology* **284**, 1141-1151, (1998).
- 356 Weiss, M. S., Jabs, A. & Hilgenfeld, R. Peptide bonds revisited. *Nature Structural Biology* **5**, 676, (1998).
- 357 Herzberg, O. & Moult, J. Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins* **11**, 223-229, (1991).
- 358 Stoddard, B. L. & Pietrokovski, S. Breaking up is hard to do. *Nature Structural Biology* **5**, 3-5, (1998).
- 359 Jabs, A., Weiss, M. S. & Hilgenfeld, R. Non-proline cis peptide bonds in proteins. *Journal of Molecular Biology* **286**, 291-304, (1999).
- 360 Scherer, G., Kramer, M. L., Schutkowski, M., Reimer, U. & Fischer, G. Barriers to rotation of secondary amide peptide bonds. *Journal of the American Chemical Society* **120**, 5568-5574, (1998).
- 361 Christen.Dh, Kortzebo.Rn & Bak, B. Results of Ab-Initio Calculations on Formamide. *Journal of Chemical Physics* **53**, 3912-&, (1970).
- 362 Brandts, J. F., Halvorson, H. R. & Brennan, M. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953-4963, (1975).
- 363 Odefey, C., Mayr, L. M. & Schmid, F. X. Non-prolyl cis-trans peptide bond isomerization as a rate-determining step in protein unfolding and refolding. *Journal of Molecular Biology* **245**, 69-78, (1995).
- 364 Fischer, G., Bang, H. & Mech, C. Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides. *Biomedica Biochimica Acta* **43**, 1101-1111, (1984).
- 365 Pal, D. & Chakrabarti, P. Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *Journal of Molecular Biology* **294**, 271-288, (1999).
- 366 Veelders, M. *Flokkuline aus Saccharomyces cerevisiae: funktionelle und strukturelle Untersuchungen* Diploma thesis, Philipps-Universität, (2008).

- 367 Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M. & Moras, D. Crystal structure of *Aeromonas proteolytica* aminopeptidase: A prototypical member of the co-catalytic zinc enzyme family. *Structure* **2**, 283, (1994).
- 368 Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* **370**, 621-628, (1994).
- 369 Rigden, D. J., Mello, L. V. & Galperin, M. Y. The PA14 domain, a conserved all-beta domain in bacterial toxins, enzymes, adhesins and signaling molecules. *Trends in Biochemical Sciences* **29**, 335-339, (2004).
- 370 Raikhel, N. V., Lee, H. I. & Broekaert, W. F. Structure and Function of Chitin-Binding Proteins. *Annual Review of Plant Physiology and Plant Molecular Biology* **44**, 591-615, (1993).
- 371 Duo-Chuan, L. Review of fungal chitinases. *Mycopathologia* **161**, 345-360, (2006).
- 372 Cabib, E. & Durán, A. Synthase III-dependent Chitin Is Bound to Different Acceptors Depending on Location on the Cell Wall of Budding Yeast. *Journal of Biological Chemistry* **280**, 9170-9179, (2005).
- 373 Mbawala, A., al Mahmood, S., Loppinet, V. & Bonaly, R. Acetolysis and 1H NMR studies on mannans isolated from very flocculent and weakly flocculent cells of *Pichia pastoris* IFP 206. *J Gen Microbiol* **136**, 1279-1284, (1990).
- 374 Kurtzman, C. P. in *The Yeasts (Fifth Edition)* 491-495 (Elsevier, 2011).
- 375 Zhang, N., Suh, S.-O. & Blackwell, M. Microorganisms in the gut of beetles: evidence from molecular cloning. *Journal of Invertebrate Pathology* **84**, 226-233, (2003).
- 376 Suh, S. O. & Blackwell, M. Three new beetle-associated yeast species in the *Pichia guilliermondii* clade. *FEMS Yeast Res* **5**, 87-95, (2004).
- 377 Suh, S.-O., Marshall, C. J., McHugh, J. V. & Blackwell, M. Wood ingestion by passalid beetles in the presence of xylose-fermenting gut yeasts. *Molecular Ecology* **12**, 3137-3145, (2003).
- 378 Toki, W., Tanahashi, M., Togashi, K. & Fukatsu, T. Fungal farming in a non-social beetle. *PloS One* **7**, e41893, (2012).
- 379 Hegedus, D., Erlandson, M., Gillott, C. & Toprak, U. New Insights into Peritrophic Matrix Synthesis, Architecture, and Function. *Annual Review of Entomology* **54**, 285-302, (2009).
- 380 Hartl, L., Zach, S. & Seidl-Seiboth, V. Fungal chitinases: diversity, mechanistic properties and biotechnological potential. *Applied Microbiology and Biotechnology* **93**, 533-543, (2012).
- 381 Wijffels, G. *et al.* A Novel Family of Chitin-binding Proteins from Insect Type 2 Peritrophic Matrix cDNA SEQUENCES, CHITIN BINDING ACTIVITY, AND

- CELLULAR LOCALIZATION. *Journal of Biological Chemistry* **276**, 15527-15536, (2001).
- 382 Ma, Z. Y. & Vosseller, K. O-GlcNAc in cancer biology. *Amino Acids* **45**, 719-733, (2013).
- 383 Kamigaito, T. *et al.* Overexpression of O-GlcNAc by prostate cancer cells is significantly associated with poor prognosis of patients. *Prostate Cancer and Prostatic Diseases* **17**, 18-22, (2014).
- 384 Nangia-Makker, P., Conklin, J., Hogan, V. & Raz, A. Carbohydrate-binding proteins in cancer, and their ligands as therapeutic agents. *Trends in Molecular Medicine* **8**, 187-192, (2002).
- 385 Oliveira, C., Teixeira, J. A. & Domingues, L. Recombinant lectins: an array of tailor-made glycan-interaction biosynthetic tools. *Critical Reviews in Biotechnology* **33**, 66-80, (2013).
- 386 Shuai, J. *et al.* A novel lectin from *Agrocybe aegerita* shows high binding selectivity for terminal N-acetylglucosamine. *Biochemical Journal* **443**, 369-378, (2012).
- 387 Slifkin, M. & Doyle, R. Lectins and their application to clinical microbiology. *Clinical Microbiology Reviews* **3**, 197-218, (1990).
- 388 Kock, M., Brückner, S., Essen, L.-O., Mösch, H.-U. & Schlereth, J. Recombinant adhesin protein and peptide thereof and methods and uses thereof. (2015).
- 389 Bowman, S. M. & Free, S. J. The structure and synthesis of the fungal cell wall. *Bioessays* **28**, 799-808, (2006).
- 390 Adams, D. Fungal cell wall chitinases and glucanases. *Microbiology* **150**, 2029 - 2035, (2004).
- 391 Morris-Jones, R. *et al.* Synthesis of melanin pigment by *Candida albicans* in vitro and during infection. *Infection and Immunity* **73**, 6147-6150, (2005).
- 392 Karunanithi, S. *et al.* Shedding of the mucin-like flocculin Flo11p reveals a new aspect of fungal adhesion regulation. *Current Biology* **20**, 1389-1395, (2010).
- 393 Agrawal, B., Gendler, S. J. & Longenecker, B. M. The biological role of mucins in cellular interactions and immune regulation: prospects for cancer immunotherapy. *Molecular Medicine Today* **4**, 397-403, (1998).
- 394 Lewis, P. N., Momany, F. A. & Scheraga, H. A. Chain reversals in proteins. *Biochimica et Biophysica Acta (BBA)-Protein Structure* **303**, 211-229, (1973).
- 395 Rose, G. D., Glerasch, L. M. & Smith, J. A. Turns in peptides and proteins. *Advances in Protein Chemistry* **37**, 1-109, (1985).
- 396 Hutchinson, E. G. & Thornton, J. M. A revised set of potentials for β -turn formation in proteins. *Protein Science* **3**, 2207-2216, (1994).
- 397 Leader, D. P. & Milner-White, E. J. Motivated proteins: a web application for studying small three-dimensional protein motifs. *BMC Bioinformatics* **10**, 60, (2009).

- 398 Derewenda, Z. S. The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **34**, 354-363, (2004).
- 399 Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification* **48**, 1-13, (2006).
- 400 Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. His-tag impact on structure. *Acta Crystallographica Section D: Biological Crystallography* **63**, 295-301, (2007).
- 401 Verstrepen, K. J., Jansen, A., Lewitter, F. & Fink, G. R. Intragenic tandem repeats generate functional variability. *Nature Genetics* **37**, 986-990, (2005).
- 402 Ielasi, F. S., Verhaeghe, T., Desmet, T. & Willaert, R. G. Engineering the carbohydrate-binding site of Epa1p from *Candida glabrata*: generation of adhesin mutants with different carbohydrate specificity. *Glycobiology* **24**, 1312-1322, (2014).

Anhang

	100	110	120	130	140	150	160	170	180	190	200
PA14											..-S SQGLLGY
Epa10A_CBS138_CAGLOA01284g					-P	-EGCS			-S	-PP	..-NSVSVGLHML
Epa9A_CBS138_CAGLOA01366g					-P	-EGCS			-S	-PP	..-NAVSVGLHML
Epa11A_CBS138_CAGLOL13299g					-Y	-VGC			-P	-SV	..-G TQLT PGLSLE
Epa13A_CBS138_CAGLOL13332g					-Y	-A	-HGCS		-P	-SI	..-R AKLKPGLSMEI
Epa8A_CBS138_CAGLO000847g					-Y	-L	-LGCS		-P	-LL	..-G TKT PGLSMEI
Epa12A_CBS138_CAGLOM00132g					-Y	-P	-AAC		-P	-KI	..-TNPKGLSMEI
Epa15A_CBS138_CAGLOJ11968g					-F	-P	-LGCS		-P	-EF	..-TTKPKGLSMEI
Epa1A_CBS138_CAGLOE06644g					-F	-P	-LGCS		-P	-DI	..-TTPKPKGLSMEI
Epa7A_CBS138_CAGLO005643g					-F	-P	-LGCS		-P	-NV	..-TTPKPKGLSMEI
Epa6A_CBS138_CAGLO000110g					-F	-P	-LGCS		-P	-NV	..-TTPKPKGLSMEI
Q6DTM7/1-227					-F	-P	-LGCS		-P	-NV	..-TTPKPKGLSMEI
Q6FX5/1-227					-F	-P	-LGCS		-P	-NV	..-TTPKPKGLSMEI
Epa23A_CBS138_CAGLO00220g					-NF	-P	-SGCS		-P	-NH	..-GLYNRGLTMEI
Epa19A_CBS138_CAGLOA00110g					-F	-P	-IGCS		-P	-QV	..-LKTQPLGLAVKL
Epa2A_CBS138_CAGLOE06666g					-F	-P	-MGCS		-P	-QV	..-LKTQPLGLAVKL
Q6VB3/1-220					-F	-P	-IGCS		-P	-QV	..-LKTQPLGLAVQL
Q6VB4/1-220					-F	-P	-IGCS		-P	-QV	..-LKTQPLGLAVQL
Epa22A_CBS138_CAGLOK000170g					-Y	-P	-IGCS		-P	-QS	..-TNV I PGLTMEI
Epa3A_CBS138_CAGLOE06688g					-Y	-P	-IGCS		-P	-QD	..-VA I I PGLTMEI
Epa20A_CBS138_CAGLOE00275g					-Y	-P	-LGCS		-N	-KY	..-SGHLPGLKMEI
Epa21A_CBS138_CAGLOD06732g					-H	-P	-QGCS		-P	-VD	..-SG I PKGLKMDL
Flo10_S288C_P36170							-EAC		-L	-PA	..-GEKKNGMT INF
Lg-Flo1_Spatorianus_B3IU8							-OAC		-L	-PV	..-GSRKNGMNVNF
Flo1A_S288C_P32768							-EAC		-L	-PA	..-GQRKSGMNI NF
Flo5_S288C_P38894							-EAC		-L	-PA	..-GQRKSGMNI NF
Flo9A_S288C_P39712							-EAC		-L	-PA	..-GQRKSGMNI NF
KpFlo1A_G5115_C4QY07					-A	-Y	-YAC		-D	-IG	..-ATPFDGFNATI
P.pastoris_CBS7435/KoFlo1A					-A	-Y	-YAC		-D	-IG	..-ATPFDGFNATI
P.pastoris_DSM270382/PpFlo1A					-A	-Y	-YAC		-D	-ID	..-ATRYDGFNATI
KpFlo2A_G5115_XP_002492223.1					-A	-Y	-YGC		-D	-IT	..-SNAFDGFDATI
P.pastoris_DSM270382/Cea1A					-A	-Y	-YGC		-D	-IT	..-TNAVDFDATI
CAGLOI0098g/PWP7					-P	-NL	-C		-Q	-YPNL	..-DGAAGVFDTEI
CAGLOI01047g/PWP1					-T	-NV	-C		-Q	-FDPN	..-WARTAGFQATT
CAGLOI0246g/PWP2					-N	-NV	-C		-T	-FPND	..-YGRIPGFTVLV
CAGLOI0200g/PWP3					-T	-NV	-C		-Q	-FDPN	..-FSKKGFLATT
CAGLOI0362g/PWP4					-T	-NV	-C		-S	-FKMA	..-WHRTAGFRALT
CAGLOI0406g/PWP6					-T	-NL	-C		-Q	-IPDN	..-VKRTSGFKGEI
G3BFW6/1-198							-GGC		-S	-PT	..-AVEKGLVAYEY
G8Z36/1-235			-A	-Q	-AN	-Y	-VAC		-L	-PQ	..-GERTAGFSARF
CSDW7/1-239							-AC		-L	-PQ	..-NQAQNGFKANF
Q6CSM7/1-156							-DLCP		-S	-P	..-NQAQNGFKANF
Q6CWZ6/1-155							-VACY		-I	-PI	..-NQAQNGFKANF
W0T4D4/1-228							-LGI		-P	-PQ	..-GSLAKGFTVSY
W0TK5/1-228							-LGI		-P	-PQ	..-GSLAKGFTVSY
W0T4Y3/1-226							-LGI		-P	-KT	..-NSLSKGFVSYSY
W0TG6/1-222							-LGI		-P	-KT	..-NSLSKGFVSYSY
W0TES1/1-226							-LGD		-P	-KT	..-NSLSKGFVSYSY
W0T6X2/1-228							-LGI		-P	-KA	..-DSLTKGFSVSY
Q6CV43/1-121											..-DSLTKGFSVSY
Q6CX19/1-225							-LGI		-P	-SK	..-NSPSSGFVEAY
W0TB41/1-220							-GAC		-P	-KK	..-DSPVKGFDISY
B5VTX8/1-229							-LGI		-P	-LK	..-TGSGFNVKY
G8ZY82/1-224							-LGI		-P	-OS	..-VTSGFVSYSY
G8C1Z1/1-211							-HC		-D	-IPAL	..-SSGTAGFRARF
G8YS6/1-209							-SGC		-R	-PT	..-SSARGMNVTF
G8IP4/1-235							-Y		-P	-VSLNPS	..-LVLRRGFDITVY
J75A05/1-261						-GQHD	-M	-GS	-P	-GAC	..-LPL
J7RC2/1-23							-N	-P	-P	-SC	..-GSPKPGFTARF
M9N3U0/1-228							-L	-Q	-ACD		..-P-LRG
R9XGT4/1-228							-L	-Q	-ACD		..-G-ALARAGFDLTL
G2YMK1/1-219							-S	-V	-A	-AT	..-S-TIAP-S
G3JUQ9/1-20							-K	-S	-R	-V	..-I-P
J4VRE3/1-204							-K	-S	-R	-V	..-I-A
J5IAU9/1-202							-K	-S	-R	-V	..-I-P
TS4187/1-181									-A	-S	..-TCGFRGVGYTR
W3XGZ9/1-168									-Y	-V	..-TCGFRGVGYTR
A7EJ3/1-23									-S	-VSA	..-SGNAPSIS
S8B85/1-222									-T	-EA	..-LQ
G2WSY1/1-16									-T	-V	..-V
X0BR53/1-236									-T	-V	..-V
J3NZ77/1-207									-S	-S	..-V
S8BR50/1-193									-P	-I	..-P
S8C299/1-19									-E	-P	..-T
G1X146/1-197									-I	-K	..-P
W7935/1-195									-C	-I	..-L
Q7Z9U/1-234									-D	-V	..-S
Q9P39/1-234									-T	-F	..-V
B6JZE1/1-236									-I	-PGSA	..-T
B6K558/1-233									-I	-SSDG	..-T
Q92344/1-228									-F	-V	..-S
S8ADJ9/1-226									-I	-S	..-T
A0A01053B9/1-150										-T	..-WTIV
L2FXC7/1-201										-G	..-T
TOKM46/1-201										-S	..-C
L2FWT7/1-216										-T	..-PGAEFAL
T0JQ48/1-216										-T	..-SSAST
L2GB85/1-253										-T	..-SSAST
T0MCH8/1-253										-T	..-SSAST
R8BM0/1-18										-T	..-SSAST
K2R480/1-232											..-VGAGYA
vCABR_gm872										-S	..-I
vCABR.00013j2										-T	..-INS
vCABR.00032j3										-W	..-D
vCABR_g2510										-L	..-L
vCABR.00030G_3293										-S	..-D
vCABR.00013j51										-H	..-Q
vCABR.00013j59										-Q	..-G
vCABR.00016C_1651										-Q	..-D
vCABR.00008G_581										-I	..-F
vCABR.00002j32										-T	..-L
vCAN.00015j51										-Q	..-S
vCAN.00019-snap.462										-G	..-N
vCAN.00008j131										-Q	..-D
vCAN.00015_gd_34										-N	..-Q
vCAN.00002j296										-D	..-N
vCAN.00002j2										-S	..-G
vCAN_gm2290										-M	..-L
vKLD.00025j235										-H	..-A
vACA_g1301										-G	..-S
vACA.00021G_2948										-G	..-S
vACA.00033C_3917										-Q	..-Y
vKLB.00026j1										-Q	..-A

	200	210	220	230	240	250	260	270	280	290	300
PA14	Y F F S D	- L									
Epa10A_CBS138_CAGL0A01284g	N L Y N Y	- P Y L Y V		K P S T	- K G	- M F	- T N D T N S		D S D G E T D G D S A G G I E G R A G		Q - C W - N P E Y
Epa9A_CBS138_CAGL0A01366g	D L Y N Y	- P Y L Y V		K N P R	- T G	- F	- T N D T D S		D A D G E T D G D S A G G I E G R A G		Q - C W - N P E Y
Epa11A_CBS138_CAGL0L13299g	E Y Y A Y	- P M D G N		N G							C W - N A T Y
Epa13A_CBS138_CAGL0L13332g	E I Y S Y	- A L R P T		D D							C W - D S A Y
Epa8A_CBS138_CAGL0C06847g	E L Y E Y	- P L D N H		G I							P - H C W - D P A Y
Epa12A_CBS138_CAGL0M00132g	Q L Y S Y	- G H P V P		G							P - C W - D P A Y
Epa15A_CBS138_CAGL0L11968g	E L Y R Y	- D Y L P P		G S							P - C W - D S A Y
Epa1A_CBS138_CAGL0E06644g	E L Y S Y	- D F R K K		G S							P - C W - D A A Y
Epa7A_CBS138_CAGL0C05643g	E L Y S Y	- D L R K N		G S							P - C W - D A A Y
Epa6A_CBS138_CAGL0C00110g	E L Y S Y	- D Y L K S		G S							P - C W - D A A Y
Q6DM7/1-227	E L Y S Y	- D Y L K S		G S							P - C W - D A A Y
Q6FX5/1-227	E L Y S Y	- D Y L K S		G S							P - C W - D A A Y
Epa23A_CBS138_CAGL0I00220g	E L Y N Y	- S Y I H P		Q S		I Q S	- Y	- F N T S T L K K M E	- K G		Q - C W - D P S Y
Epa19A_CBS138_CAGL0A0110g	K L Y S Y	- G H P V P		G							C S - Y I T Y
Epa2Z_CBS138_CAGL0E06666g	K L Y S Y	- G H P V P		G							C A - Y N T Y
Q6IB3/1-220	Q L Y S Y	- G H P V P		G							C A - Y T T Y
Q6VB4/1-220	Q L Y S Y	- G H P V P		G							C A - Y T T Y
Epa22A_CBS138_CAGL0K00170g	E L Y S Y	- P I V S G		T P							N - C Y - D S Y
Epa3A_CBS138_CAGL0E06688g	E L Y S Y	- P I V S G		T P							N - C Y - D S Y
Epa20A_CBS138_CAGL0D00675g	E L Y S Y	- N Y V N P		R H P D	- P Q N P G K L D Q L Y D P N L L						E - C F - D S Y
Epa21A_CBS138_CAGL0D00673g	D L Y R F	- P I T P P		G T G							C W - S W D W L
Flo10_S288C_P36170	N F Y Q Y	- S L K D S		S						Y S N	P
Lg-Flo1_S_pastorianus_B3UA8	N F Y Q Y	- S L Q D S		T						Y S D	P
Flo1A_S288C_P32768	N F Y Q Y	- S L K D S		S						Y S N	A
Flo5_S288C_P38894	N F Y Q Y	- S L K D S		S						Y S N	A
Flo9A_S288C_P39712	N F Y Q Y	- S L R D S		S						Y S N	A
KpFlo1A_GS115_C4QYW7	T I Y Q Y	- Q A S D D		N S						I Q D	P
P.pastoris_CBS7435/KoPflo1A	T I Y Q Y	- Q A S D D		N S						I Q D	P
P.pastoris_DSMZ70382/PpFlo1A	T I Y E F	- Q P G D G		R L						T R D	P
KpFlo2A_GS115_XP_00249223.1	T I Y E Y	- N A N D L		K L						I R D	P
P.pastoris_DSMZ70382/Cea1A	T I Y Q Y	- N A N D L		R L						I R D	P
CAGL010098g/PWP7	E I Y N A	- D P S I L		I						N F	- S L W P - I L
CAGL010147g/PWP1	T T V R L	- G L T I A		S						N	- T Y K E L L
CAGL010246g/PWP2	L V G R L	- G L L I S		A						S	- D Y K E L L
CAGL010200g/PWP3	T T A R L	- G L W I T		G						N	- S Y K E L L
CAGL010362g/PWP4	L T A R L	- G I G I Q		A R L						S	- S Y K E L L
CAGL0M14069g/PWP6	E I R R I	- G L E P P		G						G	- D Y T Y L L
G3BFW6/1-198	E Y I K V	- A N H E W		K E							- G D T
G8ZW36/1-235	R F F P Y	- A L S D L		S						S	- Y K D
CSDWW7/1-239	N F Y Q Y	- D F G D E		V							- Y A D
Q6CSM7/1-156											
Q6CWZ6/1-155											
W0T4D4/1-228	S Y Y H Y	- P M I P N		P N S A							- G C Y L I - N S V F
W0TJK5/1-228	S Y Y H Y	- P M I P N		P N S A							- G C Y L I - N S V F
W0T4Y3/1-226	S Y Y H Y	- P L I P S		G S - S							- G C Y V D D R K T F
W0T6S6/1-22	S Y Y H Y	- P L I P S		G S - S							- G C Y V D D R K T F
W0TE0/1-226	S Y Y H Y	- P L I P S		G S - A							- G C Y V G D S O T F
W0T6X2/1-228	S Y Y H Y	- P M T E N		A T A P							- G C F N L - N N I F
Q6CV43/1-121											
Q6CX19/1-225	A Y Y H Y	- P M V P S		G K							- G C Y N Y - D A T Y
W0TB41/1-220	S Y Y H Y	- P M I P S		K G S N							- K C F T F - D P V Y
BSVXT8/1-229	K Y Y D Y	- P F A P S		G S - N							- P C Y - N T A Y
G8Z8Z2/1-224	K Y Y D Y	- P H A A A		G S - S							- P C Y - N L A Y
G8C12/1-211	R F F P Y	- Y P P E Y S		N						Y S P	- L
G8T556/1-209	T F Y K Y	- T L D F D		S						Y K E	- E
G8P44/1-235	T Y Y N Y	- P W S Y		M F Q I							- N R G V P D R V T Y
J75Q5/1-261	K F F S Y	- P I N V Y		D K							- V L D
J7RC24/1-23	R F F H L	- P I N Q W		V Y							- M R D
M9N3U0/1-228	T I W N Y	- P W T S Y		N H Q T							- N R G V P D R T T F
R9XGT4/1-228	T L W N Y	- P W T S Y		N H Q T							- N R G V P D R T T F
G2YMK1/1-219	A A F A E T P Q I I N S			G							- T Y
G3JUQ9/1-20	A L Y T N	- P F R R D M		S							- P
J4VRE3/1-204	A L Y D N	- P F R R D L		S							- P
J5JAU9/1-202	A L Y D N	- P F R Q D T		S							- P
TSAB18/1-181	T R Y K N	- P F T K D F		S							- S
W3K629/1-168											
A7EUJ3/1-23	A Y F S N	- P T L T T G		V G S							
S8B85/1-222	A M Y N N	- T Y L N S Q		H S							- N
G2W5Y1/1-16											- S F
X0BR53/1-236	E F Y K L	- Q R S P T D		T	- W A						- R
J3NZ77/1-207	A A Y Y F	- D R A S D K		G Q M L W R							- T
S8BR50/1-193	A I Y N Y	- I W S T T		S							- P
S8C299/1-19	A I Y N N	- P Y T E F S		G							- D
G1X146/1-197	A I Y D N	- P Y L T F T		T							- N Y
W7935/1-195	A L Y S N	- P F T F A Y									- P
Q7Z9H/1-234	A Y Y N Y	- D I S S S K									- Q
Q9P329/1-234	A Y Y D Y	- T I S S S M									- E
B6I2E1/1-236	A Y Y N Y	- T L P S K K		K							- N
B6K558/1-233	A I Y N Y	- T L P D W K		E F	- G T	- T G					- H
Q92344/1-228	A Y Y N Y	- D L P S E S		T							- F
S8ADJ9/1-226	A Y Y N I	- E I G R Y		D							- D
A0A0105389/1-150											- D
L2FXC7/1-201	A L H V I	- E E R S T L		C	- R N	- T					- L Q
T0K946/1-201	A L H V I	- E E R S N L		C	- R N	- T					- L Q
L2PWT7/1-216	A L Y S I	- D F N S D L		C	- Q N	- T					- I R
T0JQ8/1-216	A L Y S I	- D R M S D L		C	- Q N	- T					- I R
L2C885/1-253	A I Y T F	- E P R S P E		C	- R G	- T					- V
T0MCM8/1-253	A I Y T F	- E P R S P E		C	- R G	- T					- V
R8BIM0/1-18											- N
K2RA80/1-232	A A Y N N	- P Y A G G G		T A	- S N	- P					- S Y
vCABR_gm872	K V Y S Y	- E L Y D Q		N S							- Y K S
vCABR00013j2	E I W Q Y	- D Y I N P Y Y F C P D K Y D G S Y C M P W R P N E									- Q
vCABR00032j3	E L W E Y	- S Y R F Y		Y R P C I D S P E Q C - W - V N K S							- E
vCABR_g2510	R I Y A Y	- K L F D S		I H							- Y Q Q
vCABR00030G_3293	E L W E Y	- N Y R F Y		K A P C P G S P N N L C - W - V N W Q							- F
vCABR00013j51	E L W E Y	- D Y A T P		Y E L C - P D G Q T L C - P - I							- D G
vCABR00013j50	E L Y K Y	- D Y L N P		L V P C - E S G T - G							- F
vCABR00016G_1651	Q L Y K Y	- D Y L N P		L V P C - E D G I - G							- F
vCABR00008G_581	N V Y E Y	- D L S I G T		E L							- A Y N
vCABR00027j32	R V Y A Y	- T L G D S		F N							- F Q L
vCANI00015j51	E L Y K Y	- D Y L N P		L V P C - E N G I - G							- F
vCANI00019-snap.462	E I Y S Y	- P Y A W T		A A D C - G - N	- E N						- Y D N W H Q H C N D
vCANI00008j131	K L Y K Y	- D Y L N P		L V E C - E D G T - G							- F
vCANI00015_gd_34	E L Y E N	- D Y A N P		Y E L C - P D G Q T L C - P - L							- S G
vCANI00002j296	E I W E Y	- N F L N Y		I V P C P D N P Q E S C - W - A H N T							- D L S
vCANI00002j2	E I W E Y	- R Y R Y T		H E P C A D D P S Q L C - W - V N P L							- E
vCANI_gm2250	D	- G		Q K G							- Y C P Q N G P L V N - D
vHLEDE00025j235	E I Y E Y	- D Y L N P		Y Y K C - P D N Y G Y C - R E F D P A K L							- K
vACA_g1301	R L F S Y	- P L V D E		E							- S
vACA00021G_2948	R L Y D M	- P L N R P		S							- L
vACA00033G_3917	R L F S Y	- P W R N M		E							- L
vKLBAA00026j1	R F Y S Y	- P L W D W R		Q							- F

Anhang

	300	310	320	330	340	350	360	370	380	390	400																															
PA14																																										
Epa10A_CBS138_CAGL0A01284g	Y	QD	PNF	P	RF	GY	K	QY	G	E	FGASD	NVN	G	D	IS	W	DH	NE	FT	EGCKP	V	LA	AL																			
Epa9A_CBS138_CAGL0A01366g	Y	QD	PNF	P	RY	GY	K	KY	G	S	FGSSD	NVN	G	D	IS	W	DH	NE	F	EGCKP	I	MA	RL																			
Epa11A_CBS138_CAGL0L13299g	Y	KD	PEF	P	RS	GY	L	KN	K	L	LGKAD	G	V	G	P	LN	F	LA	K	NK	DSCP	K	VD	TL																		
Epa13A_CBS138_CAGL0L13329g	Y	LD	PMF	P	RT	GY	L	TH	K	L	LGKAE	G	I	G	P	LN	F	LA	K	NP	ACIP	A	VS	TL																		
Epa8A_CBS138_CAGL0C00847g	Y	LN	PEY	P	RT	GF	E	SS	H	L	IGT AQ	NVM	S	LN	F	EH	R	TQ	DCHV	E	FS	NL	TL																			
Epa12A_CBS138_CAGL0M001329g	Y	LD	PNF	P	RV	GY	K	KN	K	L	IARVD	G	V	G	D	IN	F	NF	HP	KR	OCTP	I	AD	YL																		
Epa15A_CBS138_CAGL0J11969g	Y	LN	PSY	P	RT	GY	K	AK	K	L	IATVD	G	V	G	D	IN	F	NF	HP	KR	OCTP	I	AD	YL																		
Epa1A_CBS138_CAGL0E06644g	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVD	G	V	G	N	IN	F	Y	HA	TK	GCTP	Q	LG	HL																		
Epa7A_CBS138_CAGL0C05643g	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVD	G	V	G	N	IN	F	Y	HA	TK	GCTP	E	ID	HL																		
Epa6A_CBS138_CAGL0C00110g	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVE	NVA	G	N	IN	F	Y	HA	PM	GCTS	L	FD	TL																			
Q6DTM7/1-227	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVE	NVA	G	N	IN	F	Y	HA	PM	GCTS	L	FD	TL																			
Q6F55/1-227	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVE	NVA	G	N	IN	F	Y	HA	PM	GCTS	L	FD	TL																			
Epa23A_CBS138_CAGL0E00220g	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Epa19A_CBS138_CAGL0A00110g	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Epa13A_CBS138_CAGL0E06666g	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Q6V83/1-220	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Q6V84/1-220	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Q6V83/1-220	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Q6V84/1-220	Y	QY	PDF	P	RS	GY	K	KE	R	K	LADIT	G	V	G	K	ID	I	DL	KP	ND	PCRV	L	TG	HL																		
Epa22A_CBS138_CAGL0K00170g	Y	VD	PEY	P	RT	GY	L	KR	K	M	IGKST	G	V	G	N	LN	Y	KE	SV	EK	ACMV	K	YG	NL																		
Epa3A_CBS138_CAGL0E06688g	Y	VD	PEY	P	RT	GY	L	KR	K	M	IGKST	G	V	G	N	LN	Y	KE	SV	EK	ACMV	K	YG	NL																		
Epa20A_CBS138_CAGL0E00275g	Y	KD	INY	P	RV	GY	K	SK	T	L	IGRS	G	V	G	I	LN	F	NF	VP	PW	GCMV	Q	KG	QL																		
Epa21A_CBS138_CAGL0D06732g	WL	T	V	T	P	D	F	P	S	G	Y	W	K	W	P	K	I	A	E	V	D	G	V	G	P	L	H	L	D	S	G	H	P	T	C	D	P	L	T	S	V	L
Flo10_S288C_P36170	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVE	NVA	G	N	IN	F	Y	HA	PM	GCTS	L	FD	TL	TL																		
Lg-Flo1_Spasticus_B31U0A8	Y	LD	PNY	P	RT	GY	K	SH	R	L	LAKVE	NVA	G	N	IN	F	Y	HA	PM	GCTS	L	FD	TL	TL																		
Flo1A_S288C_P32768	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Flo5_S288C_P38894	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Flo9A_S288C_P39712	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
KpFlo1A_GS115_C4Q1W7	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Ppactor1_CBS7435_KoFlo1A	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Ppactor1_DSM270382_PpFlo1A	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
KpFlo2A_GS115_XP_002492223.1	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Ppactor1_DSM270382/Cea1A	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL010098g/PWP1	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL010147g/PWP1	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL010246g/PWP2	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL010209g/PWP3	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL010363g/PWP4	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CAGL0M14969g/PWP6	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
G3F7W6/1-198	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
GRZV36/1-235	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
CSDW07/1-239	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Q6CSM7/1-156	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
Q6CWZ6/1-155	Y	LD	VNY	P	RT	GY	K	TH	N	R	FAKVN	G	I	D	G	LD	F	EF	NP	TR	SCVP	S	KD	QL																		
W0T4D4/1-228	F	LD	DEY	Q	HG	GY	E	TLGGG	L	IGRSN	G	V	T	N	LT	Y	YS	PSTCS	PGCKV	Q	VG	NV	NV																			
W0TJ5/1-228	F	LD	DEY	Q	HG	GY	E	TLGGG	L	IGRSN	G	V	T	N	LT	Y	YS	PSTCS	PGCKV	Q	VG	NV	NV																			
W0T4Y3/1-226	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
W0T6G6/1-22	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
W0T6J1/1-226	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
W0T6Z7/1-228	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
Q6CV43/1-121	F	LD	DEY	Q	HG	GY	E	TIDGG	L	IGTSS	G	V	T	N	LT	F	SS	DTTCS	SGCKV	G	TG	KL	KL																			
Q6CX19/1-225	Y	QY	EAY	Q	HG	GY	E	TYGGG	L	IGKSS	G	V	T	N	LT	F	KS	TAA	SNSCKP	V	KA	RL	RL																			
W0T841/1-220	Y	LD	VNY	L	HG	GY	Q	TYGGG	L	LGTS	G	V	T	N	LT	F	SL	GE	RDCN	P	PVNG	KL	KL																			
B5V7X/1-229	Y	TS	NSY	LYEG	GY	E	TLGGG	L	LGESS	G	V	T	N	LT	F	ET	TS	LQIKS	CTV	Y	EG	QL	QL																			
G8Z8Z/1-224	Y	QY	DDY	LYNG	GY	Q	TLGAG	L	LAQSS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM	NM	NM																			
G8C1Z1/1-211	Y	LD	VNY	L	FD	YV	DE	IN	N	E	I	ATTS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM																		
G8Y56/1-209	Y	LD	VNY	L	FD	YV	DE	IN	N	E	I	ATTS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM																		
G8J44/1-235	Y	LD	VNY	L	FD	YV	DE	IN	N	E	I	ATTS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM																		
J75A25/1-261	Y	LD	VNY	L	FD	YV	DE	IN	N	E	I	ATTS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM																		
J7R1C2/1-23	Y	LD	VNY	L	FD	YV	DE	IN	N	E	I	ATTS	G	V	T	N	LT	F	AT	SSNM	SRLF	I	MA	NM																		
M9N3U/1-128	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
R9XG74/1-228	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
G2YMK1/1-219	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
G3JUQ9/1-20	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
J4VRE3/1-204	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
J5JAU9/1-202	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
T5A187/1-181	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
W3XG29/1-168	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
A7EJ3/1-129	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
S8B85/1-222	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
G2W5Y/1-16	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
X0B853/1-236	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
J3N277/1-207	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
S8R50/1-193	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
S8C299/1-19	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
G1X146/1-197	F	LD	DEY	L	HG	GY	A	TYGGG	L	IGSST	G	V	T	N	LT	F	HF	HF	PGGLRP	I	LG	SL	SL																			
W7935/1-195	F																																									

	400	410	420	430	440	450	460	470	480	490	500
PA14											P S S
Epa10A_CBS138_CAGL0A01284g	L										P P G
Epa9A_CBS138_CAGL0A01366g	L										P T A
Epa11A_CBS138_CAGL0L13299g	L										P A G
Epa13A_CBS138_CAGL0L13329g	L										P L A
Epa8A_CBS138_CAGL0C00847g	L										P A S
Epa12A_CBS138_CAGL0M00132g	L										P P N
Epa15A_CBS138_CAGL0J11968g	L										P S N
Epa1A_CBS138_CAGL0E06644g	L										P A S
Epa7A_CBS138_CAGL0C05643g	L										P A S
Epa6A_CBS138_CAGL0C00110g	L										P Q A
QeDTM7/1-227	L										P Q A
QeFX55/1-227	L										P Q A
Epa23A_CBS138_CAGL0O00220g	L										P O N
Epa19A_CBS138_CAGL0A00110g	L										P S N
Epa2A_CBS138_CAGL0E06666g	L										P S N
QeVB33/1-220	L										P P N
QeVB4/1-220	L										P P N
Epa22A_CBS138_CAGL0K00170g	L										P A G
Epa3A_CBS138_CAGL0E06688g	L										P A G
Epa20A_CBS138_CAGL0E00275g	L										P P A
Epa21A_CBS138_CAGL0D06732g	L										P H G
Flo10_S288C_P36170	L	A C S D D D A T E Y	S A S	Q V V P V K R G		V K L C S D N T T L S S K T E K R E N D D C D Q G		A A			Y W S S
Lg-Flo1_Spasterianus_B3IU8	L						F	W N	T A		S W S S
Flo1A_S288C_P32768	L	P C P Q E D S	Y	G N W	G C	K G	M G A C S	N S	Q G	I A	Y W S T
Flo9A_S308C_P39712	L	P C P Q E D S	Y	G N W	G C	K G	M G A C S	N S	Q G	I A	Y W S T
KpFlo1A_GS115_CAGYW7	L	P C P Q E D L	Y	G N W	G C	K G	I G A C S	N N	P I	I A	Y W S T
P.pastoris_CBS7435/KoFlo1A	L							V			A T T
P.pastoris_DSMZ70382/PpFlo1A	L							V			A T T
P.pastoris_DSMZ70382/PpFlo1A	L							V			P T T
KpFlo2A_GS115_XP_002492223.1	L							S			R T A
P.pastoris_DSMZ70382/Cea1A	L							S			R T A
CAGL010098g/PWP7	L							A			S D T
CAGL010147g/PWP1	L							G	F		P F F T
CAGL010246g/PWP2	L							G	F		P F F T
CAGL010209g/PWP3	L							G	T		D L F T
CAGL010362g/PWP4	L							G			L V G
CAGL0M14069g/PWP6	L							G			W F P
G3BFW6/1-198	L							Y C	S K		N T D
G8ZW36/1-235	L	P					Y C		T R	T V	FWG S
CSDWW7/1-239	L							L	W D G R V V V P T E P W N F D Y S K C T N N D Y I Y N Q Q P G		
QeCSM7/1-156	L										P G A
QeCWZ6/1-155	L										P A A
W0T4D4/1-228	V										P P N
W0TJK5/1-228	V										P P N
W0T4Y3/1-226	L										P S N
W0T5G6/1-22	L										P S N
W0TE01/1-226	L										P S N
W0T6X2/1-228	L										P P N
QeCV43/1-121	L										
QeCX19/1-225	L										P S N
W0TB41/1-220	L										P A N
B5VTX8/1-229	L										P S N
G8ZY82/1-224	M										P S N
G8C121/1-211	L						S		A G	T	I A W G
G8Y556/1-209	L								S A	D	V Y Y G
G8Y44/1-235	L										P D N
JFSQ5/1-261	L	T E R C Y H P	F L R G L A C K Y	W C	G S		I H C N	N S	L R	G A	Y W S E
JRC24/1-23	L	T F E C Y T P	F Y E N Y P	R Y	K C	G D	V Q C S	N S	W R	T P	Y W N N
M9N3U0/1-228	L										P P N
R9XGT4/1-228	L										P P N
G2YMK1/1-219	A										V
G3JUQ9/1-20	R										K
J4VRE3/1-204	R										G
J5JAU9/1-202	D										D
TS1A187/1-181	L										
W3KCS29/1-168	L								T	D	G M L L
A7EUJ3/1-23	D										V
S8B85/1-222	G										L
G2WSY1/1-16	P										F
X0BR53/1-236	P										
J3NZ77/1-207	P										F
S8BR50/1-193	N A										H P
S8C299/1-19	V										A
G1X146/1-197	M										N
W7935/1-195	A										
Q7Z91/1-234	P G										P F P
Q9P329/1-234	P P										Y P
B6JZE1/1-236	P S										Y P
B6K558/1-233	P D										Y P
Q92344/1-228	P A										Y P
S8ADJ9/1-226	G										
A0A01053B9/1-150	P										
L2FXC7/1-201	P										
T0KM46/1-201	P										
L2FWT7/1-216	P										
TUJ048/1-216	P										
L3G885/1-253	P										
T0MCM8/1-253	P										
R8BIM0/1-18	P										
K2RA80/1-232	T										
vCABR_gm872	D S										T L
vCABR.00013-j2	L										I N G N E F Y K C D D T S V I C S Y P
vCABR.00032-j3	L										E V
vCABR_g2510	L										T L
vCABR.00030G_3293	L										T D L Y Y C N N D T S V L C Y K P
vCABR.00013-j51	L										S L
vCABR.00013-j50	L										
vCABR.00016G_1651	L										
vCABR.0008G_581	L										
vCABR.00027-j32	L	S K									N Y
vCANI.00015-j51	L	S G									N Y
vCANI.00019-snap.462	L										S I G V K F Y C E G E S A - V C H F
vCANI.00008-j131	L										R D Q Y Y C N N D T S L L C Y T P
vCANI.00015_gd_34	L										D V
vCANI.00002-j296	L										
vCANI.00002-j2	L										
vCANI_gm290	L										
vKLE.00025-j235	L										
vCACA_g1301	L	P L N C W D D T L E V N E G G K F									T C
vCACA.00021G_2948	L										Q G
vCACA.00033G_3917	L										F
vLBA.00026-j1	L										E C R N T
	L										L Y
	L										S
	L										I N Y G
	L										N C K T D K I D L - E P F L C D G W E C - N D
	L										I C C R N D
	L										P G
	L										Q
	L										I S Y G
	L										V Q C S N R
	L										I
	L										S Y G
	L										P Y T S

	600	610	620	630	640	650	660	670	680	690	700
PA14											
Epa10A_CBS138_CAGLO001284g	P N	-	-	-	-	-	-	-	-	-	-
Epa9A_CBS138_CAGLO01366g	P N	-	-	-	-	-	-	-	-	-	-
Epa11A_CBS138_CAGLO11329g	P N	-	-	-	-	-	-	-	-	-	-
Epa13A_CBS138_CAGLO11332g	P G	-	-	-	-	-	-	-	-	-	-
Epa8A_CBS138_CAGLO00847g	AG	-	-	-	-	-	-	-	-	-	-
Epa12A_CBS138_CAGLO011968g	GR	-	-	-	-	-	-	-	-	-	-
Epa7A_CBS138_CAGLO005645g	G K	-	-	-	-	-	-	-	-	-	-
Epa6A_CBS138_CAGLO000110g	G S	-	-	-	-	-	-	-	-	-	-
QeDtm7/1-227	G S	-	-	-	-	-	-	-	-	-	-
QeX55/1-227	G S	-	-	-	-	-	-	-	-	-	-
Epa23A_CBS138_CAGLO00220g	G S	-	-	-	-	-	-	-	-	-	-
Epa19A_CBS138_CAGLO00110g	G T	-	-	-	-	-	-	-	-	-	-
Epa2A_CBS138_CAGLO00666g	G T	-	-	-	-	-	-	-	-	-	-
QeVb13/1-220	G S	-	-	-	-	-	-	-	-	-	-
QeVb14/1-220	G S	-	-	-	-	-	-	-	-	-	-
Epa23A_CBS138_CAGLO00170g	H V	-	-	-	-	-	-	-	-	-	-
Epa3A_CBS138_CAGLO00688g	H V	-	-	-	-	-	-	-	-	-	-
Epa20A_CBS138_CAGLO00275g	G I	-	-	-	-	-	-	-	-	-	-
Epa21A_CBS138_CAGLO006732g	S K	-	-	-	-	-	-	-	-	-	-
Flo10_S288C_P36170	W N	-	-	-	-	-	-	-	-	-	-
Lg-Flo1_Spastorianus_B3IU8A	W D	-	-	-	-	-	-	-	-	-	-
Flo1A_S288C_P32768	W D	-	-	-	-	-	-	-	-	-	-
Flo5_S288C_P38894	W N	-	-	-	-	-	-	-	-	-	-
KpFl61A_G5115_C40Pw7	W N	-	-	-	-	-	-	-	-	-	-
P.pastoris_DSM270382/PpFl61A	G E A	-	-	-	-	-	-	-	-	-	-
KpFl62A_G5115_XP_00249223.1	S N	-	-	-	-	-	-	-	-	-	-
P.pastoris_DSM270382/Cea1A	S N	-	-	-	-	-	-	-	-	-	-
CAGLO10098g/PWP7	Y	-	-	-	-	-	-	-	-	-	-
CAGLO10147g/PWP1	T	-	-	-	-	-	-	-	-	-	-
CAGLO101246g/PWP2	T	-	-	-	-	-	-	-	-	-	-
CAGLO101200g/PWP3	T	-	-	-	-	-	-	-	-	-	-
CAGLO101562g/PWP4	T	-	-	-	-	-	-	-	-	-	-
CAGLO101406g/PWP5	C	-	-	-	-	-	-	-	-	-	-
GBFw6/1-198	C	-	-	-	-	-	-	-	-	-	-
GBZw36/1-235	E	-	-	-	-	-	-	-	-	-	-
CSDWw7/1-239	L D	-	-	-	-	-	-	-	-	-	-
QeC5m7/1-156	N N	-	-	-	-	-	-	-	-	-	-
QeCw26/1-155	D G	-	-	-	-	-	-	-	-	-	-
W074D4/1-228	L G	-	-	-	-	-	-	-	-	-	-
W07KJ5/1-228	P D	-	-	-	-	-	-	-	-	-	-
W074Y3/1-226	P D	-	-	-	-	-	-	-	-	-	-
W071G6/1-22	P S	-	-	-	-	-	-	-	-	-	-
W07E11/1-226	P S	-	-	-	-	-	-	-	-	-	-
W076X2/1-228	R T	-	-	-	-	-	-	-	-	-	-
QeC4X3/1-121	P D	-	-	-	-	-	-	-	-	-	-
QeC6X19/1-225	P Q	-	-	-	-	-	-	-	-	-	-
W07B41/1-220	P K	-	-	-	-	-	-	-	-	-	-
B5V7X8/1-229	Q S	-	-	-	-	-	-	-	-	-	-
GBZ7B21/1-224	A S	-	-	-	-	-	-	-	-	-	-
GBZ556/1-209	K P	-	-	-	-	-	-	-	-	-	-
GBZ12/1-211	Q W	-	-	-	-	-	-	-	-	-	-
GBZ561/1-209	Q W	-	-	-	-	-	-	-	-	-	-
JF54O5/1-261	S R	-	-	-	-	-	-	-	-	-	-
JF7C24/1-23	Y	-	-	-	-	-	-	-	-	-	-
M9N1U0/1-228	A R	-	-	-	-	-	-	-	-	-	-
R9NGT4/1-228	A R	-	-	-	-	-	-	-	-	-	-
GZYMK1/1-219	G	-	-	-	-	-	-	-	-	-	-
G3UJQ9/1-20	G	-	-	-	-	-	-	-	-	-	-
J4VRE3/1-204	G	-	-	-	-	-	-	-	-	-	-
J51AU9/1-202	G	-	-	-	-	-	-	-	-	-	-
TS41B7/1-181	P S	-	-	-	-	-	-	-	-	-	-
W5XG29/1-168	V	-	-	-	-	-	-	-	-	-	-
ATEU13/1-23	D	-	-	-	-	-	-	-	-	-	-
S8B15/1-222	G	-	-	-	-	-	-	-	-	-	-
G2W5Y1/1-16	T	-	-	-	-	-	-	-	-	-	-
OR8R53/1-236	N	-	-	-	-	-	-	-	-	-	-
J3N27/1-207	N	-	-	-	-	-	-	-	-	-	-
S8B29/1-193	A	-	-	-	-	-	-	-	-	-	-
SC299/1-19	S	-	-	-	-	-	-	-	-	-	-
G1X146/1-197	A	-	-	-	-	-	-	-	-	-	-
W7935/1-195	A	-	-	-	-	-	-	-	-	-	-
Q7291/1-234	Y	-	-	-	-	-	-	-	-	-	-
Q9P29/1-234	T	-	-	-	-	-	-	-	-	-	-
B6J2E/1-236	S	-	-	-	-	-	-	-	-	-	-
B6K558/1-233	S	-	-	-	-	-	-	-	-	-	-
Q92344/1-228	G	-	-	-	-	-	-	-	-	-	-
S8ADJ9/1-226	A	-	-	-	-	-	-	-	-	-	-
A0A105389/1-150	R	-	-	-	-	-	-	-	-	-	-
L2FXC7/1-201	R	-	-	-	-	-	-	-	-	-	-
T0RKM6/1-201	R	-	-	-	-	-	-	-	-	-	-
L2FW7/1-216	S	-	-	-	-	-	-	-	-	-	-
T0JC8/1-216	S	-	-	-	-	-	-	-	-	-	-
L2C8B5/1-253	S	-	-	-	-	-	-	-	-	-	-
TMCN8/1-253	T	-	-	-	-	-	-	-	-	-	-
RB8M0/1-18	S	-	-	-	-	-	-	-	-	-	-
K2RA80/1-232	N	-	-	-	-	-	-	-	-	-	-
vCABR_gm872	W N	-	-	-	-	-	-	-	-	-	-
vCABR.00013-j2	G S	-	-	-	-	-	-	-	-	-	-
vCABR.00032-j3	G S	-	-	-	-	-	-	-	-	-	-
vCABR_g2510	G S	-	-	-	-	-	-	-	-	-	-
vCABR.00030g_3293	W F	-	-	-	-	-	-	-	-	-	-
vCABR.00013-j1	D N	-	-	-	-	-	-	-	-	-	-
vCABR.00013-j50	P L	-	-	-	-	-	-	-	-	-	-
vCABR.00016G_1651	P L	-	-	-	-	-	-	-	-	-	-
vCABR.00008G_581	P L	-	-	-	-	-	-	-	-	-	-
vCABR.00027-j32	W S	-	-	-	-	-	-	-	-	-	-
vCANI.00015-j51	P L	-	-	-	-	-	-	-	-	-	-
vCANI.00019-snap.462	S N	-	-	-	-	-	-	-	-	-	-
vCANI.00008-j131	P A	-	-	-	-	-	-	-	-	-	-
vCANI.00015_gd.34	G S	-	-	-	-	-	-	-	-	-	-
vCANI.00002-j2	G S	-	-	-	-	-	-	-	-	-	-
vCANI.00002-j296	G T	-	-	-	-	-	-	-	-	-	-
vCANI_gm2390	P R	-	-	-	-	-	-	-	-	-	-
vKLD.00025-j235	G S	-	-	-	-	-	-	-	-	-	-
vACA_g1301	S I	-	-	-	-	-	-	-	-	-	-
vACA.00021G_2948	A N	-	-	-	-	-	-	-	-	-	-
vACA.00033G_3917	T I	-	-	-	-	-	-	-	-	-	-
vKBA.00026-j1	L	-	-	-	-	-	-	-	-	-	-

7.2 Kristallographische Statistiken

7.2.1 Datensammlungs- und Verfeinerungsstatistiken - Epa6A-Komplexe

Tabelle 7.1 – Datensammlungs- und Verfeinerungsstatistiken der Epa6A-Komplexe.

Ligand	Gal β 1-4Glc	Gal β 1-3GalNAc	Gal β 1-4GlcNAc	Gal β 1-3GlcNAc	Gal α 1-3Gal
Datensatz	Epa6-1_6	T-Anti_03	Epa6-Z4_1	Lactobiose_01	Epa6-Z5
Datum	16.09.2012	21.10.2012	14.02.2013	21.10.2012	14.02.2013
Röntgenquelle	BESSY II BL14.1	PETRA III P14	BESSY II BL14.1	PETRA III P14	BESSY II BL14.1
Detektor	MARmosaic 255 mm	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M
Temperatur (K)	100 K	100 K	100 K	100 K	100 K
Wellenlänge (Å)	0.91841	0.826606	0.918409	0.826606	0.918409
Raumgruppe	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Zellmaße (Å)					
a	41.3	41.4	41.5	41.3	41.5
b	61.1	61.1	63.3	61.3	63.3
c	106.3	106.6	110.3	106.6	110.3
α, β, γ	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Moleküle / a.s.u.	1	1	1	1	1
Auflösung (Å)	28.77 - 1.48 (1.56 - 1.48)	34.24 - 2.00 (2.11 - 2.00)	41.58 - 1.80 (1.90 - 1.80)	53.12 - 2.25 (2.37 - 2.25)	41.58 - 1.50 (1.58 - 1.50)
Beobachtete Reflexe	174129	109348	146305	120749	257032
Multiplizität	3.8 (3.8)	5.8 (5.7)	5.3 (5.0)	9.0 (9.2)	5.4 (5.2)
Einzigartige Reflexe	45573	18879	27537	13420	47367
R_{merge} (%)	4.4 (57.1)	12.3 (60.4)	9.2 (40.4)	17.3 (64.7)	4.6 (52.0)
Vollständigkeit (%)	99.7 (99.9)	99.7 (99.9)	99.7 (99.9)	99.9 (100)	99.9 (99.9)
Mittleres $I/\sigma(I)$	21.1 (2.9)	11.9 (2.7)	17.1 (4.4)	12.1 (3.5)	20.7 (3.0)
Auflösung (Å)	26.50 - 1.48	32.69 - 2.15	41.58 - 1.80	53.17 - 2.30	41.58 - 1.50
R_{work}	14.4	17.7	18.7	17.8	15.0
R_{free} (%)	17.2	23.0	22.3	22.0	17.6
Reflexe	44337	14534	26129	11930	46100
(Arbeits-, Test-)	1173	740	1378	624	1223
r.m.s.d.					
Bdglängen(Å)	0.008	0.010	0.010	0.008	0.007
Bdgsinkel (°)	1.295	1.319	1.384	1.209	1.134
Atome	2109	2056	2197	2013	2219
Mittlerer B-Wert (Å ²)	19.167	25.137	16.266	23.917	20.240

7.2.1 Datensammlungs- und Verfeinerungsstatistiken – Epa9A-Komplexe

Tabelle 7.2 – Datensammlungs- und Verfeinerungsstatistiken der Epa9A-Komplexe.

Ligand	Gal β 1-4Glc	Gal β 1-3GlcNAc	Gal β 1-4GlcNAc
Datensatz	Epa9-1_3	Epa9_LacNBio	Epa9-Z4_1
Datum	16.11.2012	12.08.2012	13.02.2013
Röntgenquelle	BESSY II BL14.1	BESSY II BL14.1	BESSY II BL14.1
Detektor	MARmosaic 255 mm	MARmosaic 255 mm	Pilatus 6M
Temperatur (K)	100 K	100 K	100 K
Wellenlänge (Å)	0.918410	0.918410	0.918409
Raumgruppe	<i>P3₁21</i>	<i>P3₁21</i> ₁	<i>P3₁21</i>
Zellmaße (Å)			
a	66.6	66.6	66.8
b	66.6	66.6	66.8
c	118.0	117,7	116.0
α, β, γ	90, 90, 120	90, 90, 120	90, 90, 120
Moleküle / a.s.u.	1	1	1
Auflösung (Å)	41.23 - 2.15 (2.27 - 2.15)	41.21 - 2.50 (2.64 - 2.50)	40.96 - 2.58 (2.72 - 2.58)
Beobachtete Reflexe	186187	52916	79850
Multiplizität	10.9 (11.1)	4.8 (4.9)	8.1 (8.0)
Einzigartige Reflexe	17016	10922	9889
R_{merge} (%)	9.9 (59.2)	13.7 (50.7)	12.3 (57.8)
Vollständigkeit (%)	99.9 (100)	99.7 (99.9)	99.9 (100)
Mittleres $I/\sigma(I)$	24.9 (4.9)	12.5 (3.4)	15.7 (3.7)
Auflösung (Å)	41.23 – 2.15	33.35 – 2.50	41.00 – 2.60
R_{work}	16.3	17.0	18.1
R_{free} (%)	19.4	23.1	24.2
Reflexe (Arbeits-, Test-)	16140 857	10356 549	9180 492
r.m.s.d.			
Bdglängen(Å)	0.009	0.012	0.009
Bdgsinkel (°)	1.263	1.498	1.330
Atome	2043	2008	1930
Mittlerer B-Wert (Å ²)	27.860	21.827	36.072

7.3 Produzierte und analysierte Proteine

7.3.1 Cea1A

```

10      20      30      40      50      60
MGSSHHHHHH SSSLVPRGSH MDDSGNGDNS DTAYGCDITT NAVDGFDAI YQYNANDLRL

70      80      90     100     110     120
IRDPTFMSTG YLGRNVLNKI SGVTVPGFNI WNPSSRTATV YGVKNVNYYN MVLELKGYFK

130     140     150     160     170     180
ADVSGDYKLT LSHIDDSML FFGKETAFKC CDAGSIPLNE APTDYSLFTI KPSNQVNSEV

190     200     210     220     230     240
ISATQYLEAG KYYPVRIVFV NALERARFDF KLTIPSGAVL DDFQNYIYQF GDLDENSCH

```

7.3.2 PpFlo1A

```

10      20      30      40      50      60
MGSSHHHHHH SSSLVPRGSH MQESGDGDIS DSAYACDIDA TRYDGFNATI YEFQPGDGR

70      80      90     100     110     120
TRDPVFMSTG YLNRTQLHSI TGVTDPGFSI YTPGVPTTTL YGIPNVNWNW LLELKGYFR

130     140     150     160     170     180
AEVSGDYGLS LRNIDDSAIL FFGKETAFQC CNENSISNEA STDYSLFTIF RQEGDETTNL

190     200     210     220     230     240
DSFTYTQYLE AGKYYPVRTF FVNIERHAVF NFTMTLPDGT ELTDFHNYIY QFGALDEEQ

```

QA

7.3.3 Epa1A

```

10      20      30      40      50      60
MGSSHHHHHH SSSLVPRGSH MTSSNDISLA SKDPTTFPLG CSPDITTPKK GLSMELYSYD

70      80      90     100     110     120
FRKKGSYPCW DAAYLDPNYP RTGKSHRLL AKVDGVTGNI NFFYHATKGC TPQLGHLPAS

130     140     150     160     170     180
YNYPKPLTMT NFTMLLYGYF RPKVTGFHTF TISADLLFV NFGAGNAFDC CRRDSSADHF

190     200     210     220     230     240
GNYQAYAIWG SKTAKDELTV HLDAGVYYP I RLFYNNREYD GALSFTFKTE SNENTVSDFS

250     260
EYFFSLDDTE EGCPGLISYD SS

```

7.3.4 Epa6A

10	20	30	40	50	60
MGSSHHHHHH	SSGLVPRGSH	MKDDYSSSLS	NNNLGWDPT	EFPLGCSPNV	TTPKNGLSME
70	80	90	100	110	120
LYSYDYLKSG	SNPCWDAAYL	DPNYPRTYGK	SHRLLAKVEN	VAGNINFYYH	APMGCTSLFD
130	140	150	160	170	180
TLPQAYNYRT	PLTMTNFTML	LYGYFKPKVT	GYHTFTISAD	DLLFVNFAG	NAFDCCKRES
190	200	210	220	230	240
SADDFGNQYA	YAVWGSQTAK	DDLTVHLDAG	LYYPIRIFFN	NRDNDGALSL	TLKTESDPNP
250	260				
VIDFSDYFYS	FDDTKDGCPG	LVSYDTS			

7.3.5 Epa9A

10	20	30	40	50	60
MGSSHHHHHH	SSGLVPRGSH	MASDITPFAH	YPRPEGCSSP	PNAVSVGLHM	DLYNYPYLYV
70	80	90	100	110	120
KNPRTGFTND	TDSADGETD	GDSAGGIEGR	AGQCWNPEYQ	DPNFPRYGYK	KYGSFGSSDH
130	140	150	160	170	180
VNGKISWDHN	EFKEGCKPIM	ARLPTAYNYP	AKITFSNFTM	VLSGYFKPKS	TGLYKFEIHA
190	200	210	220	230	240
DDFILFNFGS	KNAFECNRE	ESIDNFGPYV	AYAMWPNEAD	QELEVYLFED	SYYPIRLFYN
250	260	270	280	290	
NRDYHRSKFMV	GFYPPNTEEI	TYDFDGYLYM	LDDTGNECKD	SIRYKTVCCD	

7.3.6 Flo10A

10	20	30	40	50	60
MGSSHHHHHH	SSGLVPRGSH	MLGTTEACLP	AGEKKNGMTI	NFYQYSLKDS	STYSNPSYMA
70	80	90	100	110	120
YGYADAELKG	SVSQTKLSI	DYSIPCNGAS	DTCACSDDDA	TEYSASQVVP	VKRGVKLCSD
130	140	150	160	170	180
NTTLSSKTEK	RENDDCDQGA	AYWSSDLFGF	YTTPTNVTVE	MTGYFLPPKT	GTYTFGFATV
190	200	210	220	230	240
DDSAILSVGG	NVAFECCKQE	QPPITSTDFT	INGIKPWNAD	APTDIKGSTY	MYAGYYPIK
250	260	270	280	290	
IVYSNAVSWG	TLPVSVVLPD	GTEVNDDFEG	YVFSFDDNAT	QAHCSVPNPA	EH

7.4 Abkürzungen

Å	Ångstrom ($1 \text{ Å} = 100 \text{ pm} = 10^{+10} \text{ m}$)
Bdgslänge	Bindungslänge
Bdgswinkel	Bindungswinkel
C-terminal	Carboxyterminal
Da / kDa	Dalton ($1 \text{ Da} = 1 \text{ u} = 1.660538921 \cdot 10^{-27} \text{ kg}$)
EDTA	Ethylendiamintetraessigsäure
M	Molar (mol/L)
N-terminal	Aminoterminal
PEG	Polyethylenglycol
PMSF	Phenylmethylylsulfonylfluorid
<i>r.m.s.d.</i>	<i>Root mean square deviation</i> (quadratische Mittelwertabweichung)
rpm	<i>Revolutions per minute</i> (Umdrehungen pro Minute)

8 Danksagung

Herrn Prof. Dr. Lars-Oliver Essen danke ich herzlich für die interessante Themenstellung, die damit einhergegangene, hervorragende, wissenschaftliche Betreuung sowie die Freiheit, welche mir bei Planung, Ausgestaltung und Durchführung von Experimenten gewährt wurde.

Bei Herrn Prof. Dr. Hans-Ulrich Mösch bedanke ich mich für die freundliche Übernahme des Zweitgutachtens sowie die stets ergebnisreiche, konstruktive Zusammenarbeit in jeglichen Belangen der pilzlichen Adhäsion.

Herrn Dr. Stefan Brückner danke ich nicht nur für die überaus gelungene Zusammenarbeit auf fachlicher und menschlicher Ebene, sondern auch für die Einführung in die Geheimnisse des Sauerteigs und das (nord-)hessische Esprit.

Bei Dr. Rike Diderrich und Nina Wozniak bedanke ich mich für die gelungene Zusammenarbeit auf den Gebieten der Epas und Flokkuline: Das ist ein weites Feld!

Prof. Dr. Andreas Brune und Dorothee Tegtmeier danke ich für den interessanten Ausflug ins Reich der Insekten.

Dr. Maik Veelders und Dr. Manuel Maestre-Reyna möchte ich dafür danken, dass sie mir gezeigt haben, wie man sich im Adhäsion-Kaninchenbau zurechtfindet und wie man etwas Ordnung in das ganze Durcheinander bringt.

Der gesamten Arbeitsgruppe Essen, deren Teil ich in den letzten Jahren, während Studium und Promotion, sein durfte, gilt meine tiefste Dankbarkeit. Die Mitglieder dieser Arbeitsgruppe, damals wie heute, ob Assistent, Doktorand oder Masterand, sind ein Garant für erfolgreiche Zusammenarbeit, fruchtbare Diskussion, ein offenes Ohr, Spaß außerhalb des Labors und generell großen Zusammenhalt: Ihr seid eine starke Truppe (und das nicht nur beim Kistenschleppen)! Zur Zeit der Abgabe der Arbeit umfasste die Gruppe folgende Personen: Katrin Anders, Ankan Banerjee, Sophie Franz, Sabrina Gies, Petra Gnau, Silke von Horsten, Elisabeth Ignatz, Vitali Kalugin, Timo Kraushaar und Ralf Pöschke. Zusätzlich danke ich Bernard Lutterbach und Oxana Litvinova für ihren Beitrag zu dieser Arbeit.

Ein besonderer Dank gilt hier natürlich den YeastBuddies um $\frac{T/V/M}{K^3}$.

Ganz herzlich möchte ich mich bei meinen Freunden in Nah (Argenstein) und Fern (Brandenburg, Bremen, Ernsthausen, Kufstein, Mainz, Osnabrück etc.⁺⁺⁺⁺) für die schönen Auszeiten und die Unterstützung bedanken. Meinen Kolleginnen und Kollegen der Freiwilligen Feuerwehr danke ich für die Abwechslungen vom Alltag und den guten Zusammenhalt.

⁺⁺⁺⁺ Ihr wisst schon wer gemeint ist

Ein besonderer, nicht nur auf diese Arbeit bezogener, Dank gilt meiner Familie, darunter meinen Eltern sowie meinem Bruder. Wer hätte schon gedacht, dass es so weit kommen würde: Ohne euch wäre es ganz sicher nichts geworden!

Christin Schönfeld, mit deinem Vertrauen, dem Halt den du mir gibst und deiner Liebe machst du mein Leben erst zu dem, was es ist.

große Lage basiert
Zell/Zell Zellwandproteinen Spezifität Enden
Insertion
molekulare Bindungstasche Anzahl
Struktur Aspartaten schließlich
darunter Galactosiden neuartigen darstellen anthracis
weiterhin Subdomänen Zell/Substrat-Adhäsion
Rolle phylogenetischen Adhäsion cis-Peptidbindung binden
neben Vergesellschaftung Subpopulation besiedeln
enthält CealA Hefen neue unterschiedlichste Bindung
bindet mittlere Glycanbindung Eigenschaften konservierten mittels
flexiblen charakterisiert strukturell gemeinsamen konsekutiven
essentielle Erkennung Proteine Grundlage
Epa6A Glycanbindedomäne
Ca²⁺-Ion Etablierung Adhäsine funktionell identifizieren Epa9A
Affinität Lebensweise Pichia-Hefen
A-Domäne GPI-verankerten Flokkulin-Subtyps charakteristisches zwingend
großer einhergeht bioinformatischen phylogenetischer Zellaggregation anzupassen
hochaffine Bindungseigenschaften komplexieren
Teil verbundenen Transglykosylierung einzigartigen kompakte
PpFlo1A strukturellen N-Acetylglucosamin ungewöhnliches vermittelt
chitinöser spezifische Glycosylphosphatidylinositolankers strukturelle folgt
direkten Antigenen Glycanstrukturen Arbeit Interaktionen hoch
untersucht Subtyp Subtypen Bindungsmodus Zusammen gelingt
Epa1A nicht-kristalliner glabrata Ascomyceten C-Domäne
erzeugt Loops Subdomäne funktionelle Umwelt
darin Habitaten integrierten ladunggetriebene Mitglied
gelang B-Region möglicherweise cerevisiae Studie
GPI-CWP Fixierung Zellwandproteine
Motiv protektiven Saccharomyces Untersuchung
Konformation Aminosäureresten Gruppe
sorgfältigen O-glycosylierte Flo10A
repetitive Kontaktfläche