Structural and functional characterization of fungal cell wall proteins involved in adhesion and integrity sensing

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

des Fachbereichs Biologie der Philipps-Universität Marburg

vorgelegt von

Bernard Johannes Lutterbach

aus Herdecke

Marburg 2019

Die vorliegende Dissertation wurde von März 2014 bis Februar 2019 am Fachbereich Biologie der Philipps-Universität Marburg unter der Leitung von Prof. Dr. Hans-Ulrich Mösch angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180)

als Dissertation angenommen am _____

Erstgutachter: Prof. Dr. Hans-Ulrich Mösch

Zweitgutachter: Prof. Dr. Lars-Oliver Essen

Tag der Disputation _____

Erklärung

Ich versichere, dass ich meine Dissertation

"**Structural and functional characterization of fungal cell wall proteins involved in adhesion and integrity sensing**"

selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

_____ _____

(Ort, Datum) (Bernard Lutterbach)

Table of contents

Summary

The human pathogenic yeast *Candida glabrata* harbors a family of seven PA14 domain-containing cell wall proteins (Pwp) with a similar modular structure typically found in fungal adhesins. Fungal adhesins are secreted proteins that usually consist of an N-terminal domain for adhesion (A-domain), a large central segment comprised of a variable number of highly glycosylated, serine- and threonine-rich repeats (B-Domain), and a C-terminal region carrying a GPI (glycosylphosphatidylinositol) anchor required for attachment to the cell wall. Therefore, these proteins are also referred to as GPI-CWP (GPI-anchored cell wall-associated proteins) adhesins. In the first part of this work, the crystal structures of the A-domains of the two paralogs Pwp1A and Pwp5A were elucidated, giving a novel and detailed insights into structural features of PwpA domains. Surprisingly, PwpA domains have an exposed calcium-binding site, rather than a binding pocket for terminal glycan recognition typically found in other GPI-CWP adhesion domains. In addition, the structural rigidity of PwpA domains appears to be significantly lower than that of other PA14 domains, making them more sensitive to environmental stresses, caused by e.g. changes in ionic strength or pH. Also, the Pwp1A/Pwp5A structure-based modeling of all other PwpA domains allowed a detailed structural comparison of the whole Pwp family. Glycan array screening with various fluorescently labeled PwpA domains furthermore allowed to identify glycosaminoglycan as a possible group of ligands and a previously unknown host substrate for GPI-CWPs of *C. glabrata*. In the case of Pwp1A, isothermal titration calorimetry revealed that this adhesion domain is able to bind a synthetic heparin pentasaccharide with low micromolar affinity. These results are relevant, because numerous bacteria, viruses and parasites are known to bind to glycosaminoglycan via a variety of adhesive proteins during host colonization and pathogenesis. As such, these findings represent the first example for heparan-sulfate mediated adhesion by a fungal pathogen.

In the second part of this work, the crystal structure of the cysteine-rich domain (CRD) of the cell surface sensor protein Wsc1 of *S. cerevisiae* could be solved. The CRD of Wsc1 is embedded in the cell wall and has been suggested to detect mechanical stress, leading to Wsc1 sensor clustering and activation of a signaling pathway that confers the execution of a cell wall integrity maintenance program. The structure of the Wsc1-CRD shows that this domain contains four disulfide bonds and therefore is structurally highly rigid. The Wsc1-CRD also harbors three clusters of surface-exposed, aromatic amino acid residues, indicating that these structural motifs could be crucial for mediating

hydrophobic interactions upon sensor activation, which allows the clustering of Wsc1 sensor domains and triggers downstream signaling events.

Zusammenfassung

Die humanpathogene Hefe *Candida glabrata* verfügt über eine Familie von sieben Pwp Proteinen (PA14 domain containing cell wall protein) mit einem modularen Aufbau wie er typischerweise in pilzlichen Adhäsinen zu finden ist. Pilzliche Adhäsine sind sekretierte Proteine, die üblicherweise aus drei Regionen aufgebaut sind: Einer N-terminalen domäne für die Adhäsion (A-domäne), einem zentralen Segment mit einer variablen Anzahl repetetiver, hoch glycosylierter und Serin-/Threonin-reicher Abschnitte, sowie einer C-terminalen Region mit einem GPI-Anker (glycosyl-phosphatidylinositol) für die Befestigung in der Zellwand. Solche Proteine werden auch GPI-CWP Adhäsine genannt (GPI-anchored cell wall-associated proteins). Im ersten Teil dieser Arbeit wurden die Kristallstrukturen der A-Domänen der beiden Paraloge Pwp1A und Pwp5A gelöst, welche einen neuen und detaillierten Einblick in strukturelle Merkmale der PwpA domänen geben. Überraschenderweise haben PwpA Domänen anstatt einer Bindetasche zur Erkennung terminaler Glykane, wie bei anderen GPI-CWP Adhäsionsdomänen üblich, eine exponierte Kalzium-Bindestelle. Außerdem scheint die strukturelle Stabilität von PwpA Domänen signifikant niedriger zu sein als bei anderen PA14 Domänen, wodurch sie empfindlicher gegenüber äußeren Faktoren wie Ionenstärke oder pH sind. Eine Pwp1A/Pwp5A strukturbasierte Modellierung aller PwpA domänen erlaubt zudem einen detaillieten strukturellen Vergleich der gesamten Pwp Familie. Analysen mittels Glykan Arrays mit fluoreszenzmarkierten PwpA Domänen ermöglichten die Identifizierung von Glykosaminoglykan als Gruppe möglicher Liganden und somit ein bisher unbekanntes Substrat für GPI-CWPs von *C. glabrata*. Mithilfe isothermaler Titrationskalorimetrie konnte gezeigt werden, dass Pwp1A ein synthetisches Glykosaminoglykan-Pentasaccharid mit niedriger mikromolarer Affinität bindet. Diese Ergebnisse sind relevant, da zahlreiche Bakterien, Viren und Parasiten verschiedene Adhäsionsproteine nutzen um während der Pathogenese und Wirtskolonisierung an Glykosaminoglykan zu binden.

Im zweiten Teil dieser Arbeit wurde die Kristallstruktur der Cystein-reichen Domäne (CRD) des Sensorproteins Wsc1 von *S. cerevisiae* gelöst. Die CRD von Wsc1 liegt in der Zellwand und erkennt hier vermutlich mechanischen Stress, was zum Sensor-Clustering und der Aktivierung eines Signalwegs führt, der ein Programm zur Instandhaltung der Zellwandintegrität aktiviert. Durch die Kristallstruktur der Wsc1-CRD wird ersichtlich, dass diese Domäne durch vier Disulfidbrücken strukturell sehr rigide ist. Außerdem verfügt die Wsc1-CRD über drei Oberflächen-exponierte Cluster von Resten aromatischer Aminosäuren. Diese Strukturmotive könnten entscheidend für

hydrophobe Wechselwirkungen von Wsc1 Sensoren sein, was wiederum das Clustering von Wsc1 Domänen und die Auslösung von nachgeschalteten Signalen ermöglicht.

1. Introduction

1.1. Yeasts - versatile fungi

Yeasts are unicellular members of the kingdom of fungi, belonging to the domain of eukaryotes. They do not form a single taxonomic group but are distributed among the phyla of ascomycetes and basidiomycetes [1]. Nowadays, there are more than 1500 yeast species recognized, which are assumed to represent about 1 % of all extant species [2].

Yeasts are ubiquitously distributed organisms, which can be found almost everywhere on the planet, including soil and the deep-sea [3,4]. They are most common on sugar-rich material like fruits, grains or plant-exudates, representing natural habitats with high amounts of hexose sugars, their main carbon source. Some members of this diverse group tolerate extreme environmental conditions, including parameters such as temperature, desiccation, radiation, salinity or pH [5]. Moreover, yeasts are frequently found in association with other eukaryotes. Together with photosynthetic algae, they form symbiotic organisms such as lichens. They have even been identified as part of a tertiary composite organism in the form of a beard-like lichen [6]. Yeasts are also found as a constituent of the gut microbiome in insects and mammals [7]. In this context, they appear as commensals and pathogens in association with humans. There are also clinically important species, which can cause serious infections, such as members of the *Candida* genus [8].

Yeasts use organic compounds as energy source and show diverse physiological characteristics, especially concerning carbohydrate metabolism [9]. While glucose is the basic carbohydrate source of cellular energy in eukaryotes within aerobic degradation, many yeast species, including the baker's yeast *Saccharomyces cerevisiae* and close relatives, have evolved the ability to ferment sugars under anaerobic and even aerobic conditions to produce carbon dioxide and ethanol [10,11]. This fermentative lifestyle enables *Saccharomyces* yeasts to outcompete other microorganisms [12].

Although many yeasts can reproduce through sexual cycles, the most common form of vegetative growth, and a main characteristic of yeasts in general, is asexual reproduction by budding [13]. The ability to form haploid spores from diploid cells gives them the great advantage of surviving stressful conditions for extended periods of time and dispersion in a very efficient way [14].

Finally, yeasts are probably one of the earliest domesticated organisms. They have been used for brewing a prototype of modern beer in ancient Sumer over 6.000 years ago [15]. Nowadays, yeasts are employed not only in the production of alcoholic beverages and

bread, they also play an important role as eukaryotic model organisms in science. Moreover, they are implemented in biotechnology and particularly engineered in order to produce pharmaceuticals, technical enzymes, antibiotics and biofuel [16]. The extensive economic, scientific and booze-related applications make yeasts the most important eukaryotic microorganisms known and illustrate our dependence on this diverse fungal group.

1.1.1. The yeast *Candida glabrata* – an opportunistic pathogen

The yeast *Candida glabrata* is part of the *Nakaseomyces* clade, belonging to the *Saccharomycetaceae* family within the phylum of ascomycetes [17,18]. The genome of the reference strain CBS138 has been completely sequenced, revealing that it is phylogenetically more closely related to the baker's yeast *Saccharomyces cerevisiae* than to other *Candida* species such as the extensively studied pathogen *Candida albicans* (Fig. 1) [19]. Hence, the term *Candida* is not a true genus in the phylogenetic sense since it comprises species that belong to different clades within the Saccharomycotina [20,21]. The relation of *C. glabrata* to *S. cerevisiae* is reflected by a number of similarities. For example, the common progenitor of *C. glabrata* and *S. cerevisiae* experienced a whole-genome duplication event and most genes in *S. cerevisiae* have orthologues in *C. glabrata* [19]. Moreover, the chromosomal structure in terms of gene order is largely conserved between the two species [22]. However, despite their common origin, they show clear differences concerning lifestyle and impact on human life (Tab. 1).

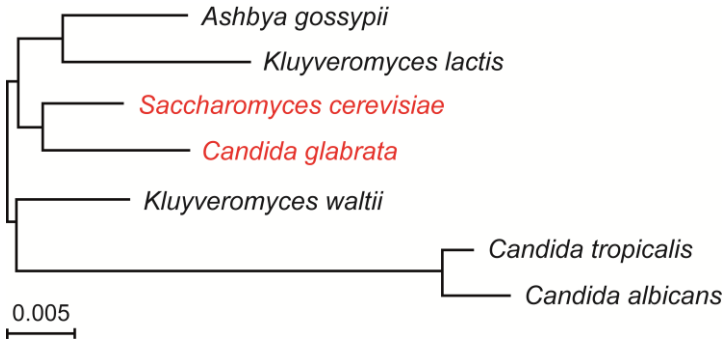

Fig. 1: 18S phylogeny of *Candida* species and other *Saccharomycetes*. *C. glabrata* and *S. cerevisiae* are closely related and distinct from other pathogenic *Candida* species such as *C. tropicalis* or *C. albicans* [23].

C. glabrata, like many other *Candida* species, is a ubiquitous fungus that is found to live in commensal interactions with mammals [24]. Here, it is part of the normal microbiota on mucosal surfaces of the oral cavity as well as the vaginal and gastrointestinal tract [24,25]. Nevertheless, it has also been isolated from a variety of non-mammalian sources [26-28]. In humans, *C. glabrata* is usually found in small numbers and does not cause disease in healthy persons [29]. Due to its commensal lifestyle, *C. glabrata*

has lost some of the genes needed for several metabolic pathways, like for example galactose utilization, niacin-, thiamine- and pyridoxine biosynthesis (Tab. 1) [23]. Such auxotrophies are generally compensated by the host environment [30]. Other adaptations are a short generation time and a growth optimum at 37 °C [31].

Tab. 1 Comparison of features from *C. glabrata*, *C. albicans* and *S. cerevisiae* [23,32,33].

	***C. glabrata* CBS138**	***C. albicans* SC5314**	***S. cerevisiae* S288c**
Ploidy	Haploid	Diploid	Diploid
Genome size	12.3 Mb, 13 chromosomes, 5293 ORFs	15.4 Mb, 8 chromosomes, 6198 ORFs	12.2 Mb, 16 chromosomes, 6049 ORFs
True hyphae	Absent	Present	Absent
Biofilm formation	Present	Present	Present
Major adhesins	Lectins (encoded by *EPA* genes)	Agglutinins (*ALS*), *HWP*	Lectins (encoded by *FLO* genes), agglutinins
Phenotype switching	Present	Present	Absent
Virulence	Opportunistic pathogen	Opportunistic pathogen	Non-pathogenic
Major sites of infection	Oral, vaginal, gastrointestinal, disseminated	Oral, vaginal, gastrointestinal, disseminated	Non-infectious
Auxotrophy	Niacin, thiamine, pyridoxine	None	None
Azole resistance	Innate resistance	Susceptible	Susceptible
Sexual cycle	Unknown	Known (cryptic)	Known

Despite its primarily commensal lifestyle, *C. glabrata*, along with many other *Candida* spp., is considered an opportunistic pathogen. After the widespread pathogen *C. albicans*, *C. glabrata* represents the second most common cause for fungal infections in humans worldwide and a main cause of nosocomial infections [34]. In total, *C. glabrata* accounts for about 20–24 % of all *Candida* infections, increasing every year [35-37]. In immunocompromised patients, *C. glabrata* can become invasive and cause a systemic candidiasis, which may result in a lethal sepsis [24]. The mortality rate after three month of a systemic candidiasis reaches up to 40 % in humans [35]. However, an induced infection of mice with *C. glabrata* does not lead to the development of a systemic candidiasis, whereas it becomes lethal upon immunosuppression [38]. In healthy individuals, *C. glabrata* is restricted by the action of the innate immune system and microbial communities, which counteract dissemination by competition for nutrients and the secretion of toxins [39]. Immunocompromised patients (e.g. cancer, diabetes, HIV,

high age, transplant patients) or persons with disturbed microbiota, e.g. upon antibiotic treatment, are much more susceptible to *Candida* infections [40-42]. However, in contrast to *C. albicans*, some *C. glabrata* strains show an innate resistance to classic azole-based antimycotics [43,44]. Thus, infections with *C. glabrata* are difficult to treat and associated with longer hospital stays and higher costs [45]. Moreover, *C. glabrata* is able to form biofilms, which enables to colonize biotic and abiotic surfaces and to persist in clinical devices, such as catheters [46]. A recent study reports the presence of *C. glabrata* in brain tissue from patients with Alzheimer's disease, indicating a fungal infection as a possible etiology for Alzheimer's disease [47].

The related pathogen *C. albicans* is a diploid, polymorphic yeast that is able to switch between yeast-, hyphal- and pseudohyphal growth, whereas *C. glabrata* is thought to be strictly haploid and restricted to the yeast form (Fig. 2) [23]. Although evidence for sexual reproduction and true hyphae have never been observed for *C. glabrata*, its genome carries homologs of all *S. cerevisiae* genes involved in the mating process and the presence of both mating types has been reported [23,48]. However, *C. glabrata* may develop pseudohyphae under specific *in vitro* conditions [49].

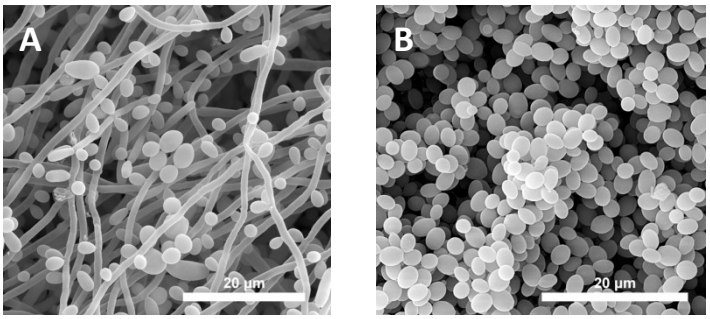

Fig. 2 Cell morphologies of *Candida* spp. in biofilms. A: Filamentous network of *C. albicans* in the form of yeasts, hyphae and pseudohyphae. **B:** *C. glabrata* biofilm, formed smaller spherical yeast cells [50].

Despite their close association with humans, the phylogenetic distance between *C. albicans* and *C. glabrata*, separated by several non-pathogenic yeasts, implies that their ability to infect humans has evolved independently. In fact, *C. albicans* and *C. glabrata* have evolved different strategies for adhesion, tissue invasion, nutrient acquisition and the interaction with immune cells [24]. *C. albicans* reveals an aggressive strategy of infection that includes the active penetration of host epithelium by the formation of true hyphae followed by rapid dissemination into deeper parts of the tissue [51]. In contrast, *C. glabrata* engages in a combination of immune evasion and persistence [19,52,53]. It has been reported that *C. glabrata* misuses the uptake by macrophages to evade other immune cells and to survive and even proliferate inside these cells [54,55]. Thus, *C. glabrata* might utilize phagocytosis as 'trojan horse' mechanism to overcome epithelial barriers and to hide from immune surveillance,

similar to the strategy postulated for *C. neoformans* [53,56]. It has been shown that *C. glabrata* attaches to *C. albicans* hyphae during infection, indicating a kind of piggy-back strategy for host invasion, since *C. glabrata* cannot actively penetrate tissue on its own [57]. *C. glabrata* may also exploit tissue destruction caused by *C. albicans* to gain nutrients. In fact, co-infections by both fungi are a common observation in oral candidiasis [58]. However, knowledge about molecular mechanisms employed by *C. glabrata* to infect humans remains limited and the true lifestyle of this yeast is still subject of discussion since traits considered as virulence factors rather suggest an ancient origin [18].

1.1.2. *Saccharomyces cerevisiae* – a useful yeast

The baker's yeast *Saccharomyces cerevisiae* is a unicellular fungus, belonging to the Saccharomycetaceae family within the phylum of ascomycetes. Due to its facultative anaerobic metabolism, which allows the conversion of sugars like glucose to CO_2 and ethanol under aerobic conditions, it has been used to produce beer, wine and bread for a long time in human history, and is still in use [59]. In nature, *S. cerevisiae* is primarily found on ripe fruits, representing its natural habitat by the abundance of its primary nutrients [60]. Historic human use and ease of cultivation led to an early scientific interest in yeasts; therefore much research has been carried out on *S. cerevisiae*. Due to its basic eukaryotic characteristics, accessibility and short generation time, it has become an important model organism. The genome of *S. cerevisiae* was the first eukaryotic genome that has been completely sequenced [61], which revealed a size of 12 Mb with 6604 open reading frames, distributed over 16 chromosomes (haploid) [33]. It includes more than 23°% of human homologous genes, which contributed to a better understanding of many important eukaryotic cell processes as well as human disease genes [62,63]. Moreover, *S. cerevisiae* has had a considerable economic impact, since it is widely used for a number of commercial applications like the production of biofuels [64] or as cell factories for the production of recombinant proteins or pharmaceutical agents [65]. *S. cerevisiae* can grow in the haploid and diploid form. Vegetatively, it proliferates by budding, which is the mitotic division of mother- and daughter- cells. It has two haploid mating types, *MATa* and *MATα*, which can mate to form diploid cells that can either sporulate or continue to exist as diploid cells. Upon nutritional starvation, the diploid cell is able to build four haploid spores, which can germinate under more favorable nutritional conditions and resume asexual reproduction [66]. Depending on environmental conditions, *S. cerevisiae* can choose between adhesive growth and filamentous growth with pseudohyphae, involving cell-cell and cell-substrate adhesion [67].

1.2. The yeast cell wall

The yeast cell wall provides the cell with shape and mechanical stability. By its robust structure, it acts as a counterbalance to the cellular osmotic pressure and limits permeability, while at once being dynamic enough to allow morphological changes and cell growth. Moreover, it serves as a physical barrier and makes the first contact with the environment. In pathogenic yeasts such as *C. glabrata*, the cell wall also plays a role in host-yeast interactions and harbors a number of proteins that are considered as virulence factors [68-70]. It provides protection against the host defense system and may confer resistance to antifungal drugs [71,72]. The importance of the cell wall is illustrated by the fact that yeasts invest a considerable amount of energy into its biogenesis, which can comprise up to 30 % of the cell dry weight [73,74] and up to 50 % of the cell volume [75]. Moreover, about 25 % of the genes in *S. cerevisiae* are related to cell wall synthesis and integrity [76].

The high stability of the yeast cell wall is given by its layered composition (Fig. 3). The innermost layer, which surrounds the plasma membrane is chitin, a linear polymer of *N*-acetylglucosamine units, which are connected by β-1,4 linkages [77]. Chitin significantly contributes to the cell wall's rigidity [77]. Since chitin contributes only 1-4 % to the cell wall's dry weight, it though represents only a minor part [78]. It is again surrounded by a thick layer of fibrous β-1,3-glucan and branched β-1,6-glucan, while all parts are covalently connected to each other [79]. The β-1,3-glucan chains are primarily responsible for cell wall strength and elasticity, while β-1,6-glucan is important for crosslinking different cell wall components. The β-glucan-chitin complex is the major constituent of the inner cell wall, which has a thickness of 70-100 nm [80]. The branched β-1,6-glucan links components of the inner and outer wall. In the outer wall, there is a huge number of mannoproteins, which are extensively *O*- and *N*-glycosylated and thereby build an almost closed layer that can be up to 200 nm thick [75,81]. The cell wall proteins contribute an amount of 30-50 % to the dry weight of the cell wall [82]. Mannoproteins of the wall are mostly linked to β-1,6-glucan through a processed glycosylphosphatidylinositol (GPI) anchor (GPI *anchored cell wall proteins,* GPI-CWPs). They can also be linked to β-1,3-glucan through an ester bond [83]. The wall-proteins are involved in cell-cell recognition and cell-substrate binding and are therefore often responsible for cell adhesion [84,85]. It is noteworthy that the *C. glabrata* cell wall contains about 50 % more mannoprotein than the cell wall of *S. cerevisiae*, which might account for virulence factors that are absent in *S. cerevisiae* [78].

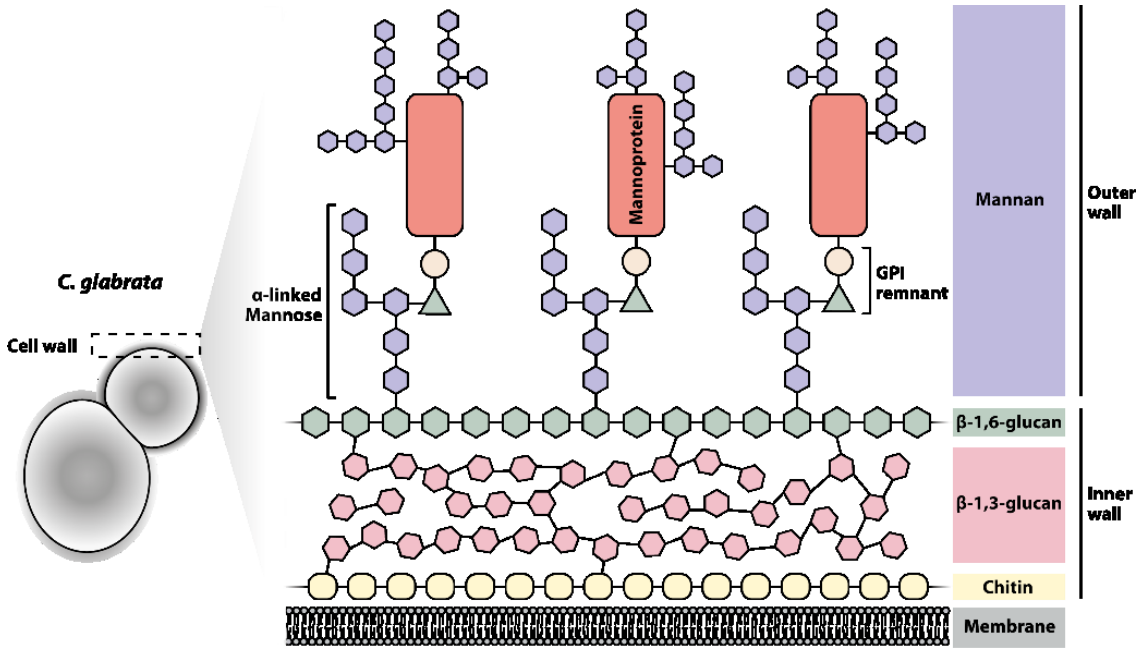

Fig. 3 Overall organization of the *Candida glabrata* cell wall. The cell wall can be distinguished in two layers. The inner wall contains the skeletal polysaccharides chitin and β-glucan, which confer wall strength and cell shape. The outer wall is enriched with *O*- and *N*-linked mannose polymers (mannan) that are covalently associated with GPI anchored glycoproteins.

The yeast cell wall is a highly dynamic structure that is constantly remodeled and structurally supported by the incorporation of a large variety of cell wall proteins [78]. These proteins are required for morphological processes like bud growth and cell division, as well as shmoo formation and cell fusion during mating. Furthermore, fungal cell walls are permanently exposed to various kinds of stresses such as oxidative stress, microbial enzymes, mechanical injuries and, in particular, changing osmolarity. Perturbation of the cell wall integrity or synthesis hence leads to cell lysis. Therefore, stress-sensor proteins in the cell wall detect various environmental influences and transmit signals for the induction of a signaling cascade and cellular response. In *S. cerevisiae*, cell wall stress is detected by the Wsc- and Mid- cell wall integrity sensor families (1.5) [86,87].

1.3. Fungal adhesins

The initial recognition and adhesion of fungi to other cells or surfaces is crucial for a variety of functions, e.g. mating or habitat colonization. In fungi, adhesion is mainly mediated through a large group of surface-exposed adhesion proteins, so-called adhesins. In the pathogenic yeast *C. albicans* for example, adhesion is mainly mediated by the agglutinin-like sequence (Als) protein family [85,88]. The Als-family has been described to contribute to pathogenesis as multifunctional invasins and adhesins that enable adhesion to epithelial- and endothelial cells, as well as components of the extracellular matrix [89-92]. The non-pathogenic yeast *S. cerevisiae* carries flocculins (Flo),

which confer self-recognition and enable the formation of flocs by cell-cell adhesion [93,94]. These flocs, composed of thousands of cells, provide the advantage of protection from outer stress factors for internal cells. The flocs sediment to the bottom or float to the surface during fermentation processes, making it easy to separate them from fermentation products [95]. Another fungal adhesin, Cea1, of the industrially utilized methylotrophic yeast *Komagataella pastoris* has been reported to mediate binding to chitinous polymers [96,97]. The epithelial adhesion protein family (Epa) in *C. glabrata*, mediates adhesion to the glucan matrix of epithelial cells [70,98].

In *C. glabrata*, the majority of cell wall proteins are GPI-anchored. These proteins are mostly present in the outer part of the cell wall and among them, there are several proteins, that govern primary host-pathogen interactions, such as adhesins, of which *Candida glabrata* uses a large variety [24]. A hallmark of fungal adhesins is their common modular domain architecture, that can be separated into three distinct domains, A, B and C (Fig. 4) [99,100].

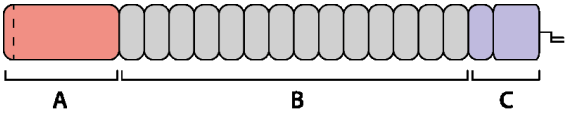

Fig. 4: Domain architecture of fungal adhesins: A secretion signal sequence (dashed line) at the *N*-terminus is followed by an A domain (red), which serves as a functional binding domain and confers adhesion. A central B-region (grey) with serine- and threonine rich repeats builds a highly glycosylated stalk, responsible for presenting the functional A domain outside of the cell wall. A carboxy-terminal C domain (blue) contains a GPI anchor, required for covalent association to the fungal cell wall.

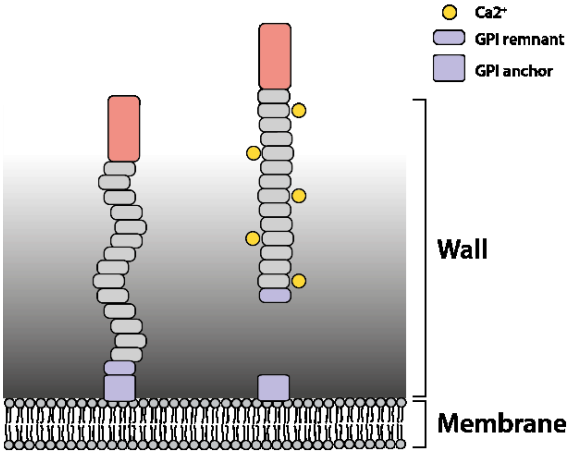

Fig. 5 Localization GPI CWPs in the cell wall. Fungal adhesins are secreted proteins that first attach to the plasma membrane by their GPI anchor. Upon arrival on the membrane, the GPI anchor is cleaved off and the adhesin is linked covalently through a GPI remnant to beta-1,6-glucan via transglycosylation. Interaction of the B region with Ca^{2+} ions putatively promotes a semi-solid structure.

During maturation and transport through the secretory pathway, fungal adhesins undergo extensive post- translational modifications. An *N*-terminal secretion signal sequence enables the translocation through the secretory pathway, where it is removed. The carboxy-terminal C-domain contains a GPI signal peptide, which is replaced by a GPI anchor. After secretion, fungal adhesins usually get linked to the cell wall via a GPI remnant (Fig. 5). A central B-region contains several repetitive sequences, rich in serine- and

threonine, which are highly *N*- and *O*-glycosylated. It is believed that the *O*-linked oligo-saccharide sidechains promote a rod-shaped, semi-solid structure by steric hindrance and stabilization upon binding of calcium ions [100,101]. The postulated primary function of the B-region is to be a passive spacer of low complexity for the presentation of the functional A-domain above the cell wall surface [99]. However, the high glycosylation of the B-region also seems to affect its function [102], and it is considered to be an element that undergoes oligomerization in an amyloid fashion and thereby increases avidity [103]. However, its precise function remains unknown. With respect to the process of establishing contact between the fungal cell and other cells or substrates, the functional A-domain appears to be the most important by mediating specific binding to either glycan ligands, proteins or abiotic surfaces [104]. Calcium is a known cofactor for many fungal adhesins, and C-type lectins form a large subcategory within this class [105]. Major fungal adhesin families often encode several proteins within each species, and different strains can even encode various subsets of adhesins to employ strategies for adhesion.

1.3.1. Adhesin families of *Candida glabrata*

Adhesins of pathogenic fungi have been shown to significantly contribute to pathogenesis [95,106] and one of the most important characteristics of *Candida* spp. concerning virulence is their ability to adhere tightly to different host surfaces, including human skin and mucosal tissues [85]. Therefore, adhesion is considered as a first important step during the initial stages of infection [85]. *Candida* species like *C. albicans* and *C. glabrata* owe their success as pathogens, in parts, to a remarkably large variety of different adhesins, which provide the cell with an array of adhesion properties [78,107]. The genome of *C. glabrata* contains between 67 (CBS138) and 107 (DSY565) sequences respectively, which encode for typical adhesins and which cluster into seven subfamilies [78,85] (Fig. 6). The large repertoire of adhesins is considered to provide *C. glabrata* the opportunity to adhere to and to form biofilms on a wide variety of surfaces and to thrive under a number of environmental conditions [78].

The largest subfamily of adhesins in *C. glabrata* comprises the Epa- (*epithelial adhesin*) family that, depending on the strain, contains between 17 (CBS138) and 23 (BG2) orthologs and is found in cluster I (Fig. 6) [23,78]. Research on the Epa family has shown that it belongs to the most important virulence factors of *C. glabrata* by mediating strong adhesion to mucosal epithelia [24,70]. The *N*-terminal adhesion domains of the Epa family mediate recognition of terminal disaccharide-glycans in the glycocalyx of

mucosal surfaces in a C-type lectin manner [70,108]. Moreover, it has been shown that the Epa proteins mediate adhesion to abiotic surfaces like agar or plastics [109,110].

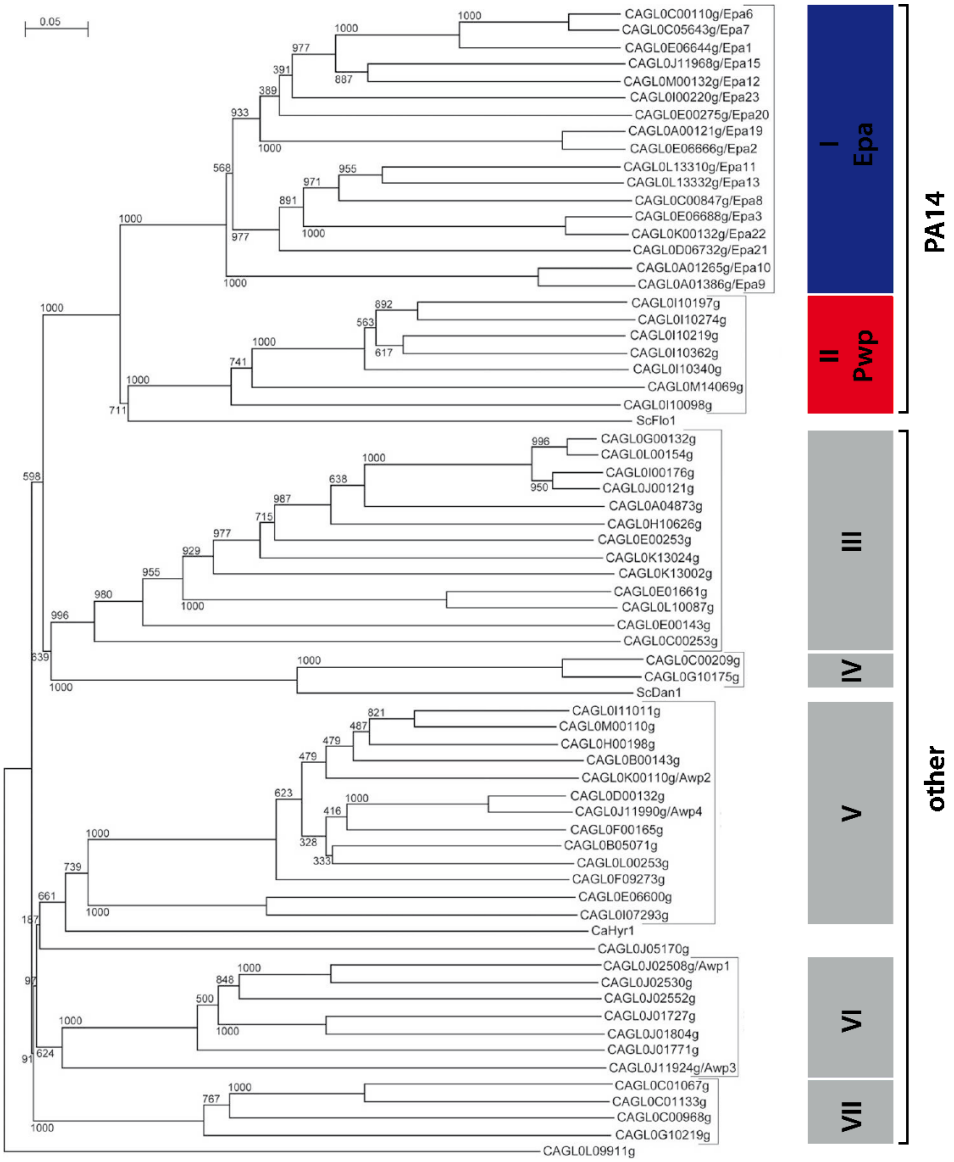

Fig. 6 Subfamilies of adhesin-like wall proteins in *C. glabrata* CBS138 [78]. Phylogenetic neighbor-joining tree based on the putative functional adhesion domains (the 300 *N*-terminal amino acids).

The adhesin-like cluster II in *C. glabrata* represents the Pwp family with seven members, which has not yet been studied in detail. The proteins were termed Pwp1-7, according to *PA14-containing wall protein* [78] since, referring to the *Pfam* database, their putative adhesion domains share a common core structure: a so-called PA14 domain. This structure is also found in adhesion domains of epithelial adhesins from *C. glabrata* and flocculins from *Saccharomyces* and *Komagataella* yeasts [111]. The Pwp family appears to be absent in other *Candida* species, suggesting that *C. glabrata* has possibly evolved a unique strategy for the interaction with its host. Comprising seven orthologs of putative adhesins, the cluster appears relatively small, compared to the Epa family. DE GROOT and KRANEVELD *et al.* identified another family of adhesion-like wall

proteins (Awp) in *C. glabrata* that is involved in the first stage of biofilm development [78,112]. Fourteen Awp domains have been identified, which spread among the clusters III-VII. A gene deletion of *AWP5*, also known as *AED1*, has been shown to reduce adherence to human endothelial cells [113]. Specific knowledge about the functions of the Awp family remains yet mostly unknown. However, the related yeast *C. albicans* carries no genes of the *EPA* or *PWP* family. The adhesion of this pathogenic fungus is mainly mediated by Als adhesins, which bind to peptide-termini [114,115]. While the modular domain architecture remains conserved in the Als family, the structure of their adhesive domains differs from the PA14 domain [107].

1.3.2. The PA14 domain – structural basis for adhesion domains

The PA14 domain was initially characterized as part of the *anthrax protective antigen* PA20 pro-peptide from the pathogen *Bacillus anthracis* and was named after its molecular weight of 14 kDa [116, 117]. The PA14 subdomain is cleaved off during maturation of the PA20 pro-peptide, which enables the final stage of anthrax toxin assembly [118]. The core structure of the PA14 domain is generally characterized by a tightly packed β-'sandwich' motif, composed of two antiparallel β-sheets (Fig. 7) [116]. The initial PA14 of domain *Bacillus anthracis* shows an exposed N*trans*D motif, which reappears as a conserved D*cis*D motif in a variety of surface-exposed fungal adhesins, where it contributes to the cell's glycan binding capabilities [93,108].

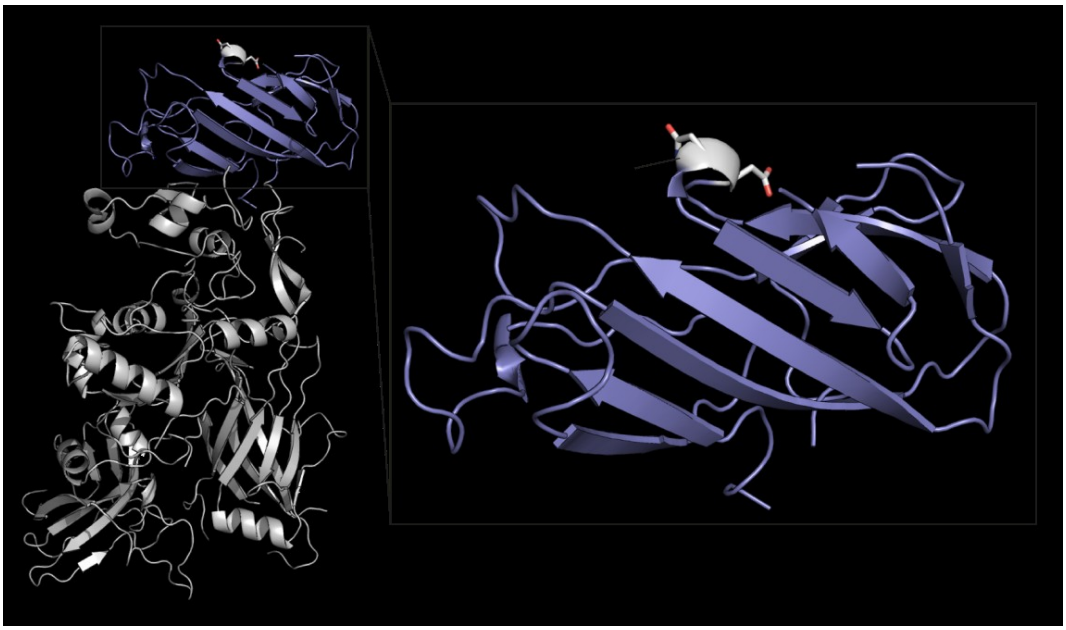

Fig. 7 PA14 domain (blue) in the mature protective antigen (grey) from *B. anthracis* (PDB: 1ACC). The PA14 domain builds a tightly packed β-sandwich that is composed of two opposite sheets comprising six and four β-sheets. It shows an exposed N*trans*D motif, which is found as a conserved D*cis*D motif in a number of fungal adhesion domains where it contributes to glycan binding.

PA14-like domains have been found to be widely distributed among all domains of life with different functions, primarily bacterial toxins, proteases, β-glucosidases, glycoside hydrolases, glycosyltransferases and adhesins [108,117]. Currently, the PA14 domain occurs in 3455 sequences, distributed over 1278 species, referring to the *Pfam* database [111]. Phylogenetic analysis showed that the fungal PA14 GPI-CWP adhesin superfamily clusters into six different subgroups of which the Pwp family forms a unique, small group whereas the *EPA* family clusters into a larger and highly diverse group (Fig. 8).

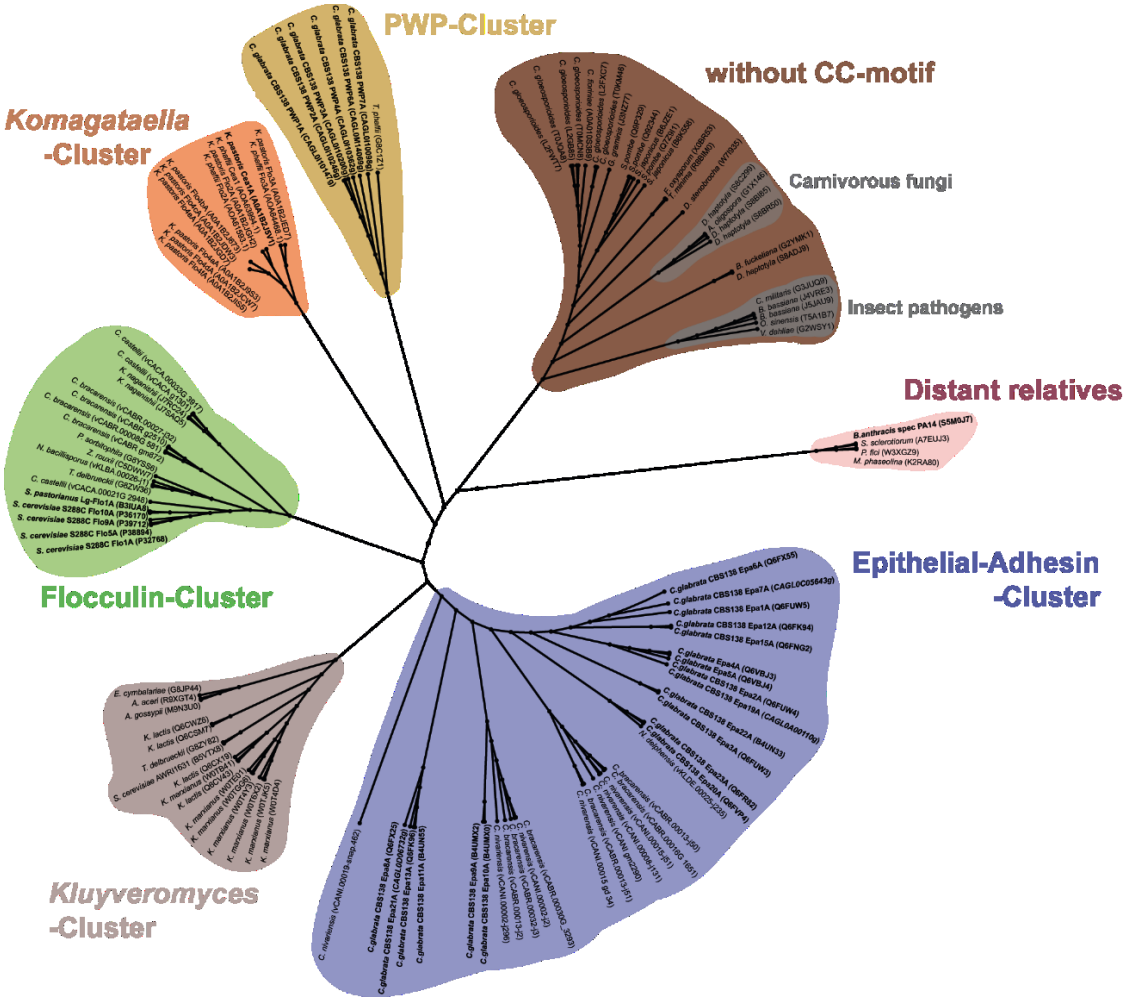

Fig. 8 Phylogenetic analysis of ascomycetous PA14/Flo5-like domain-containing GPI-CWP adhesins [97]**.** The maximum likelihood phylogenetic tree shows the clustering of GPI-CWP PA14 domains from the *Pfam* database of *C. glabrata*, *S. cerevisiae*, *K. pastoris* and further PA14_2 (GLEYA) domain-containing putative adhesins into six different subgroups. The PA14 domain of *B. anthracis* is used as an outliner and clusters together with further distant relatives. A small subgroup is formed by gene products from *Komagataella* species. Flocculins from *S. cerevisiae* cluster in a subgroup (green) and epithelial adhesins (*EPA*) in another, highly diverse subgroup (blue). Other clusters are formed by PWPs from *C. glabrata* (yellow) and putative adhesins mainly from *Kluyveromyces* species (grey). A separate, large group of PA14/Flo5-like proteins lacks a conserved motif of two consecutive cysteines that is present in other clusters (brown).

Fungal PA14 domains in cell wall associated proteins are often involved in glycan binding [119]. For instance, the adhesion domain of the flocculin Flo5 from *S. cerevisiae* mediates mannoside binding and thereby enables cell-cell interactions and

flocculation [93,94]. PA14 domains in epithelial adhesins from *C. glabrata* are known to mediate glycan binding by a C-type lectin function and to confer adhesion to human epithelium *in vivo* [70,108]. This may suggest a C-type lectin function for the Pwp family as well and may, therefore indicate a role in host recognition or biofilm formation. The original PA14 domain from *B. anthracis,* however, shows neither glycan- nor calcium-binding capabilities.

1.3.3. The Epa adhesion domain of *Candida glabrata*

C. glabrata utilizes adhesins of the Epa family for the colonization of host surfaces in the oral-, urogenital- and gastrointestinal tracts. The first member of the Epa family, Epa1, has been described by CORMACK *et al*. in 1999 [98]. Since then, extensive research has been carried out to uncover specific functions of the different Epa proteins. *In vitro* functional analysis of the adhesion domains of the 17 Epa paralogs in the strain CBS138 have shown that the EpaA domains all possess individual ligand binding patterns with different affinities. They together recognize a wide variety of glycosidic ligands containing terminal α- and β-linked galactosides as well as non-galactosidic sugars of glycosidic ligands for conferring epithelial cell adhesion [70]. With respect to their ligand binding profiles, the 17 Epa paralogs have been classified into three different classes. Class I comprises Epa proteins with preference for ligands with terminal β-linked galactose units (Epa1, Epa7, Epa3, Epa10 and Epa9), while in class II, the preferred ligands contain α-linked terminal galactose units or terminal (6S)-galactose units connected by a β-type linkage (Epa6, Epa12, Epa13, Epa15, Epa22, Epa23). The Epas of class III prefer nongalactosides as terminal units [70]. It has been suggested that Epas of class III act as low specificity lectins [70]. However, it is assumed that the different Epas enable *C. glabrata* to adhere to different host surfaces by the complementation of their specificity [70].

In vivo functional studies revealed variable epithelial cell adhesion of the different Epa proteins also dividing them into three different groups. Group I (Epa1, Epa6 and Epa7) confers very efficient adhesion while group II confers medium adhesion (Epa9A, Epa12A, Epa15A, Epa23A, and Epa8A). Other Epa adhesins confer adhesion barely stronger than the control (group III) [70]. The binding of Epa1A to galactose has been shown as strong interaction with a dissociation constant in the lower micromolar range, which points out its medical significance as a virulence factor [108]. However, although it is assumed that Epa1 is responsible for a large part of the cells' adhesion, the deletion of *EPA1* does not affect the adhesive abilities of *C. glabrata in vivo*, whereas the deletion of *EPA1*, *EPA6* and *EPA7* together results in a loss of adhesion [120]. This has also been

explained by a complementary function of the Epa proteins [120]. The regulation of expression of the Epa proteins and their subtelomeric localization shows similarities to yeast flocculins [121,122].

Structural analysis of Epa adhesion domains have shown that their glycan binding capability is conferred by a conserved calcium binding site, which coordinates a Ca^{2+} ion (Fig. 9). This site comprises the side chain of an asparagine, as well as carbonyl groups of peptide bonds on the calcium binding loop 2 (CBL2) and the carboxylic side chains of two consecutive aspartate residues which are connected via an unusual *cis*-peptide bond, a so-called D*cis*D motif on the calcium binding loop 1 (CBL1) [70]. The Ca^{2+} ion directly interacts with the non-reducing sugar moiety over the 3- and 4-hydroxyl groups and is, together with other residues present in the binding site essential for glycan binding [70]. A conserved tryptophan residue on loop 3 (L3) has been shown to significantly contribute to ligand affinity [70,108]. The two outer loops L1 and L2 are linked by a disulfide bond and shield the binding site from the surrounding solvent.

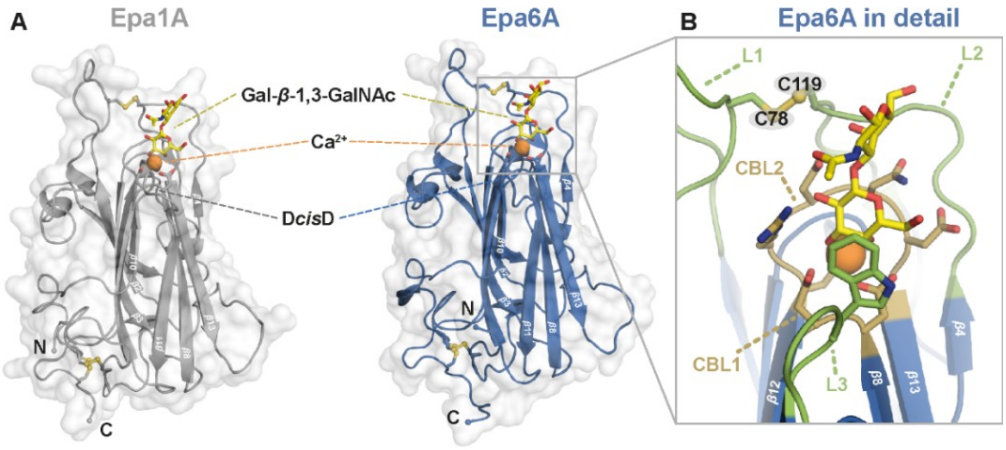

Fig. 9 Structural features of Epa1A and Epa6A [70]. **A**: The overall structure of EpaA domains shows a β-sandwich PA14 backbone. In the binding pocket, T-antigen is bound via a calcium ion that is coordinated by a conserved D*cis*D motif. **B**: Selectivity and affinity of the ligand binding are achieved by two calcium binding loops (CBL1 and CBL2) in combination with three flexible loops (L1-L3). L3 contains a tryptophan residue that covers the binding pocket and is essential for ligand binding.

However, EpaA domains of functionally closely related members are structurally quite diverse. On the other hand, phylogenetically closely related adhesins show distinct ligand binding specificities. Therefore, functionally related Epa variants might have repeatedly developed independently [70].

1.3.4. The adhesin-like Pwp family of *Candida glabrata*

While the majority of *EPA* genes and other adhesin-like clusters in *C. glabrata* CBS183 are localized in subtelomeric regions, the majority of the *PWP* genes is distributed over a centromeric region on chromosome IX (I), except *PWP6*, which is found in proximity to a telomere on chromosome XIII (M) [78].

Tab. 2 Comparison of characteristics of the *PWP* family from *C. glabrata* CBS138 [123,124].

Adhesin	Putative function	Sub-telomeric localization	D*cis*D motif
PWP1	unknown	no	yes
PWP2	unknown	no	yes
PWP3	unknown	no	yes
PWP4	unknown	no	yes
PWP5	unknown	no	no (ND)
PWP6	unknown	yes	yes
PWP7	adhesion to endothelium	no	yes

A subtelomeric localization of fungal adhesins usually allows a constant emergence of chimeric adhesins through ectopic recombination, as well as epigenetic regulation through silencing [125,126]. This is thought to enable adaptations to changing environments and to influence the length of the B-regions, which have been shown to have an effect on the binding efficiency [100]. Surprisingly, Pwp5 has an asparagine substitution on the first position of the typical D*cis*D motif, which comprises the putative calcium binding site; as described in chapter 3.1.1. DESAI *et al*. postulated that Pwp7 is required for adhesion to human endothelial cells, since a deletion mutant of *PWP7* in *C. glabrata* CBS138 resulted in reduced adherence to human umbilical vein endothelial cells [113]. However, this finding was not confirmed by HWANG-WONG in 2016, since overexpression of *PWP7* in a heterologous *S. cerevisiae* expression system did not result in significant adhesion to human endothelial cells [127]. In the same study, the majority of the Pwp family did not confer adhesion to a variety of six different human epithelial-, as well as endothelial cell lines. An exception was *PWP6*, which mediated weak adhesion to epithelial Lec2 cells from chinese hamster ovary tissue [127].

In 2015, LINDE *et al*. analyzed the whole genome transcriptome of *C. glabrata* CBS138 by RNA-sequencing (Fig. 10). Here, gene expression in liquid culture was compared to different stress conditions [128]. *PWP5* showed the highest expression level of the *PWP* family, whereas the overall expression of the *PWP* family was on a comparable level to most other adhesins such as the *EPA*s but on a low level compared to housekeeping genes such as *ACT1* (β-actin). However, the expression of *EPA6* was about ten times larger than that of other *PWP*s or *EPA*s. Hence, the expression of adhesins in

C. glabrata such as *PWP*s and *EPA*s might depend on conditions, which have not yet been studied.

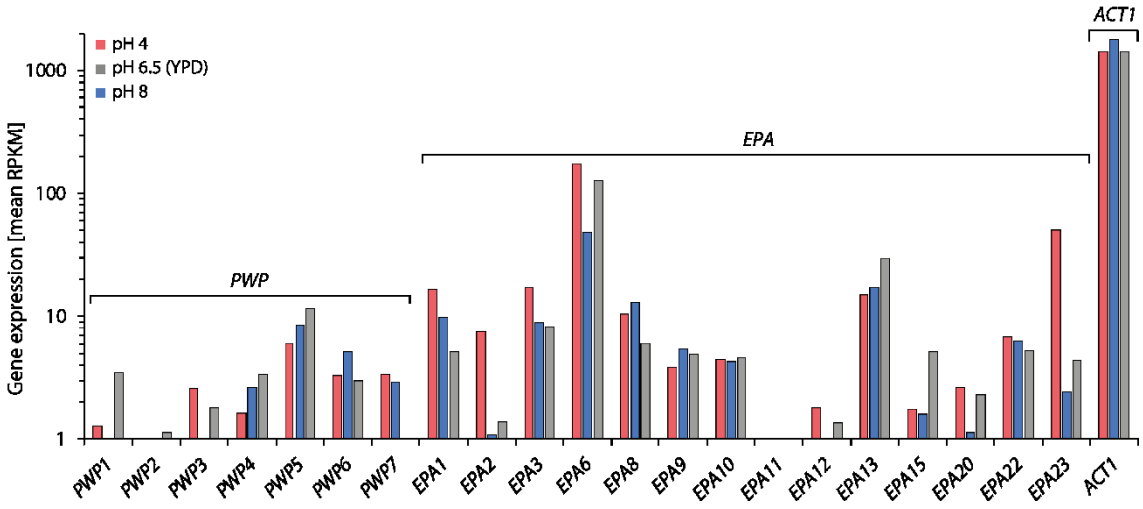

Fig. 10 mRNA expression of the *PWP* family in comparison to the *EPA* family of *C. glabrata* CBS138. Shown is the expression in YPD medium at pH 6.5 (grey) and under pH stress at pH 4 (red) and pH 8 (blue). Overall *PWP* expression levels are comparable to *EPA* expression, except *EPA6*, which expression level is about ten times larger than other *EPA*s. *PWP5* shows the highest expression of the *PWP* family, whereas *PWP2* expression is virtually absent. *PWP7* expression appears only under pH-stress. Household gene *ACT1* expression is a hundred fold higher than *PWP* expression. The figure is based on transcriptomic data by LINDE *et al.* 2015 [128].

PIET DE GROOT *et al*. isolated the covalently bound cell wall proteome of *C. glabrata* strain ATCC90876 followed by liquid chromatography-tandem mass spectrometry [78]. In this study, eighteen GPI-anchored cell wall proteins were identified, of which five were adhesin-like proteins, four were Awps and one was Epa6. However, there was no evidence for members of the Pwp family. The analyzed strain was isolated from blood but was cultivated in YPD medium, which might also affect *PWP* expression since laboratory conditions lack several features of the host such as interaction with the immune system or contact to specific tissues.

1.4. Glycosaminoglycans – sulfated polysaccharides with diverse functions

The Pwp family appears to be unique for *C. glabrata* and might mediate interaction with another set of substrates. A hitherto unstudied set of substrates for *C. glabrata* are the glycosaminoglycans as part of connective tissue and skin. Glycosaminoglycans (GAGs) are linear, heterogeneously sulfated polysaccharides, of which one or several chains are mostly covalently linked to a core protein. Both components together build the so-called proteoglycan. These proteoglycans either span the plasma membrane or are linked to the plasma membrane by a GPI anchor, whereas most proteoglycans are usually secreted [129]. GAG chains are particularly long, resulting in high molecular weight, ranging from 15 kDa to over 100 kDa, which means, that it dominates the chemical properties of the proteoglycan. Glycosaminoglycan chains consist of variable

disaccharide units, which are mostly sulfated, creating individual and complex structures. The disaccharide units always consist of an amino sugar (glucosamine that is *N*-acetylated, *N*-sulfated or *N*-acetylgalactosamine) together with a uronic acid (glucuronic acid or iduronic acid) or galactose [129]. By variations in these subunits, glycosaminoglycans can be classified into five types, comprising heparin/ heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate, and hyaluronan, each type containing unique disaccharide units, chemical linkages and degree of sulfation (Fig. 11).

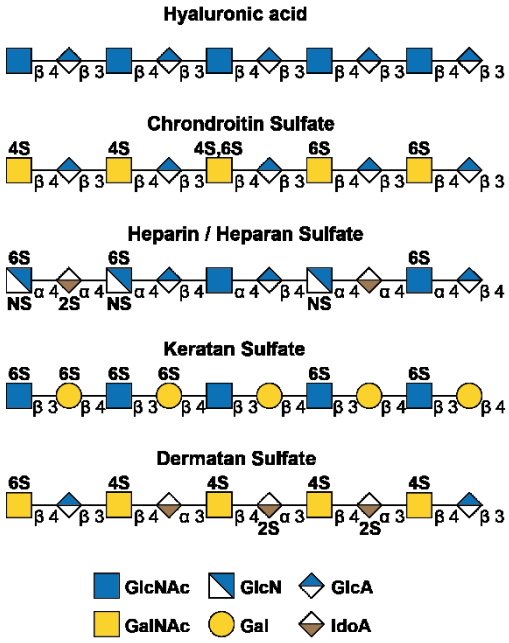

Fig. 11 The five types of glycosaminoglycan. GAGs consist of repeating disaccharide units composed of an *N*-acetylated or *N*-sulfated hexosamine and either a uronic acid (GlcA or IdoA) or galactose**.** Hyaluronan lacks sulfate groups, but other glycosaminoglycans contain sulfates at various positions. Dermatan sulfate is distinguished from chondroitin sulfate by the presence of iduronic acid. Heparan sulfate is the only glycosaminoglycan that contains *N*-sulfated hexosamine. Keratan sulfate lacks a uronic acid and instead consists of galactose and GlcNAc residues. Reducing termini are to the right.

Heparin and heparan sulfate build a group that shares the same polysaccharide backbone but they can be discriminated by their degree of sulfation. In fact, heparin is much more sulfated than heparan sulfate, resulting in the highest negative charge density of all known biomacro-molecules [130]. Hyaluronan, in contrast, is the only glycosaminoglycan that is exclusively non-sulfated and that is not found covalently linked to a core protein [129]. Chondroitin sulfate represents the most prevalent type of glycosaminoglycan in mammals where it makes a main component of cartilage and other connective tissues [130]. Dermatan sulfate is the predominant glycan found in skin, whereas keratan sulfate has its highest concentration in the cornea [131,132]. The type of glycosaminoglycan contributes to biological activities associated with proteoglycans, although GAGs are expressed by virtually all mammalian cells and can be found inside the cell and on the cell surface [129]. Moreover, single cell types have been shown to express multiple types of proteoglycan and glycosaminoglycan [129].

However, GAGs mostly occur in connective tissue as part of the extracellular matrix (ECM), which makes an essential component of metazoa, that provides structural and biochemical support in the intercellular space [129]. In addition, a few bacteria have been

shown to produce extracellular GAGs [133] (4.1.4). In the ECM, glycosaminoglycans help to connect, stabilize and organize fibrillary proteins such as collagens, that provide tensile strength and elasticity [134]. Within this mesh, GAGs provide an intermolecular force between the protein fibers [134]. It has been suggested, that the interaction between GAGs and other ECM components plays an important role in the mechanical properties of blood vessels [135]. Moreover, proteoglycans in the ECM have roles in the innate immunity and regulation of growth factor signaling [129]. They also make an essential part of the thin basement membrane, which lies against epithelial cell layers and surrounds muscle- and fat cells [129].

Beside endogenous functions, GAGs have been demonstrated to function as receptors for a number of pathogens, such as bacteria, viruses and parasites. Pathogens employ host cell-surface GAGs as receptor molecules for adhesion in the initial stage of infection, as well as eukaryotic cell invasion and intercellular migration processes [136-138] (4.1.4). However, the utilization of host GAGs by fungi as receptors for pathogenic adhesion has not yet been demonstrated.

1.5. Maintenance of the cell wall integrity in *S. cerevisiae*

The fungal cell wall is an essential structure that confers cell shape and protects from environmental stress. Thus, yeasts that are growing under cell wall damaging conditions elicit mechanisms to provide cell wall maintenance and cellular integrity. Its proper condition is controlled by the cell wall integrity (CWI) pathway, whose outermost components are membrane-spanning stress-sensors, embedded in the cell wall. This essential function makes fungal stress-sensors to potential targets for the development of newly broad-spectrum antifungals.

1.5.1. The yeast cell wall integrity (CWI) pathway

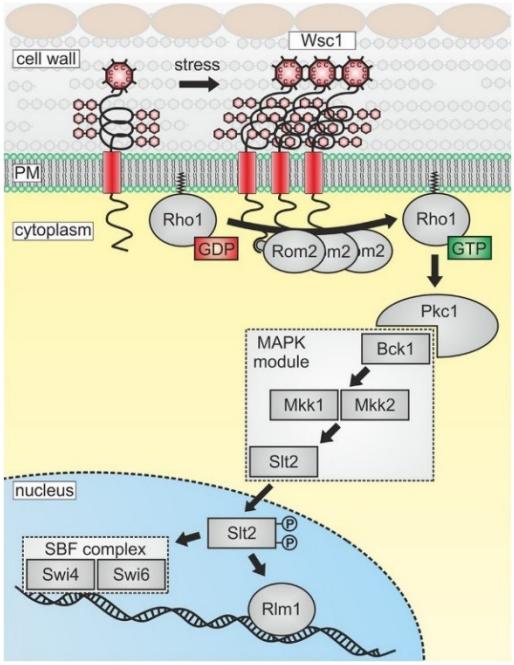

Fig. 12 The CWI pathway and its activation by sensor clustering [139]. The sensor Wsc1 clusters upon cell surface stress and activates the conserved MAPK cascade module, which activates transcription factors for the regulation of cell cycle progression, cell wall synthesis and remodeling

The yeast CWI pathway becomes activated during the requirement of cell wall remodeling, such as vegetative growth or under stress conditions which challenge the cellular integrity. Such stress conditions involve physical damage, temperature shifts, a hypo-osmotic shock or interference with antifungal agents [140]. In theory, cell wall stress is detected by sensors in the cell wall. Upon activation, these sensors bind by their cytoplasmic tail to the GDP/GTP exchange factor Rom2, which in turn activates the small GTPase Rho1 (Fig. 12). Rho1 is considered as a master regulator for the CWI pathway since it is also involved in cell wall synthesis, cell surface- as well as cell cycle signaling [140]. Moreover, Rho1 interacts with the protein kinase C (Pkc1), which activates the MAPK cascade [141]. The MAPK cascade is composed of the MAP kinases Bck1, Mkk1, Mkk2 and Slt2 (also known as Mpk1) and triggered by downstream phosphorylation. Slt2 is mostly present in the nucleus but also occurs in the cytoplasm upon cell wall stress [142]. In the nucleus, Slt2 activates the transcription factor Rlm1 and the heterodimeric SBF complex subunit, composed of Swi4 and Swi6, which are involved in the regulation of cell wall synthesis and cell cycle control [87].

1.5.2. Cell wall integrity sensor proteins

A group of five sensor proteins has been described for the CWI pathway in *S. cerevisiae*, which can be separated into the Wsc-type family with three members (Wsc1, Wsc2 and Wsc3) [86,139,143] and the Mid-type family with two members (Mid2 and Mtl1) [144,145]. A similar protein, Wsc4, is not a CWI sensor, as it resides in the ER membrane and is probably involved in protein translocation [146]. These CWI sensor proteins share a common domain structure (Fig. 13).

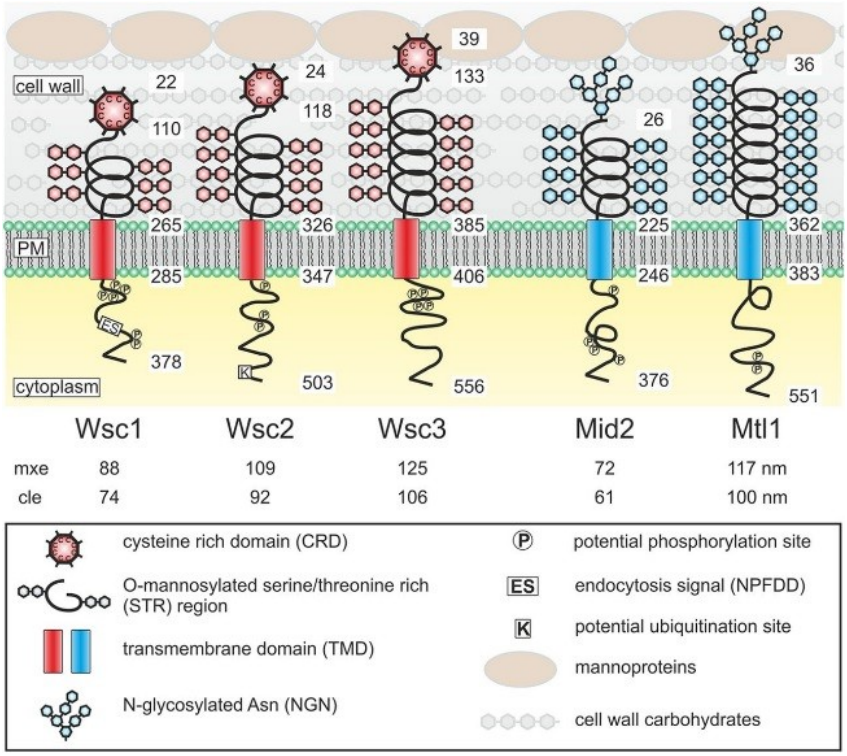

Fig. 13 The CWI sensors Wsc1-3, Mid2, Mtl1 and their proposed structural organization in the *S. cerevisiae* cell wall [139].

An *N*-terminal secretion signal mediates their transport to the cell membrane. In addition, a single transmembrane domain (TMD) anchors them in the cell membrane. These domains possess a cytoplasmic, unstructured tail of various lengths, which enables the interaction with downstream components of the CWI pathway [147]. The tail contains a putative Rom2-binding site, as well as several putative phosphorylation sites. Moreover, an extracellular serine/threonine-rich region (STR), variable in length, builds a spring-like structure, reaching into the cell wall. This region is highly mannosylated, which is thought to stabilize the rod-like shape of the spring and which is necessary for its sensor function [148]. It shows the physical properties of a nanospring, depending on the degree of glycosylation [149]. The main difference between the two CWI sensor families is the presence of an *N*-terminal, conserved and cysteine-rich domain (CRD) in the Wsc-type family, which is absent in the Mid-type family. The CRD contains eight

conserved cysteine-residues by a molecular weight of about 10 kDa, indicating a particularly rigid structure. Instead, the two Mid-type sensors carry an *N*-glycosylated asparagine residue at the head-group position, which has been reported to be required for their sensor function [87,150]. The CRD has been shown to be involved in clustering of the Wsc sensors [151]. It has features of a lectin binding domain and is supposed to be in contact with cell wall β-glycans [152]. In general, CRDs occur in proteins with diverse functions, including dimerization and DNA-interaction [86].

It is believed that the CWI sensors have a mechanoreceptor function: in theory, the CRD anchors within the cell wall and is connected to the TMD via the STR region. When the cell wall is under mechanical stress, the CRD becomes dislocated in relation to the TMD. As a consequence, tension is exerted on the intermediate spring-like STR and it becomes expanded. This might trigger a conformational change in the cytoplasmic tail, which enables interaction with Rom2 and thus induces a stress response [151]. This theory is supported by an atomic force microscopy study, in which Wsc1 showed the dynamic behavior of a linear hookean spring [149]. The different length in the STR might then be explained by the ability of stress detection in different layers of the cell wall, since different stress factors might affect different cell wall layers by varying degrees [87].

The deletion of *WSC1* leads to hypersensitivity to stress conditions including temperature shifts and the addition of cell wall destabilizing compounds, such as calcofluor white or congo red and the defect is enhanced by the additional deletion of *WSC2* and *WSC3* [86,143,153]. In addition, Wsc1-GFP localized at sites of polarized cell growth, like bud-necks and tips of emerging buds, whereas Mid2 has been shown to have a more homogeneous distribution and seems to be related to a general stress response and mating [153]. Mtl1 is supposed to be involved in the stress response upon glucose starvation and oxidative stress [154,155]. Furthermore, the deletion of *MLT1* leads to temperature sensitivity in combination with *mid2*Δ [145]. The sensors Wsc1 and Mid2 have been suggested to be mainly responsible for CWI signaling with maybe partly overlapping functions, although Mid2 has no CRD [152].

1.5.3. The Wsc1 cysteine-rich domain

It has been shown by HEINISCH *et al.*, that Wsc1 builds membrane clusters (Fig. 12) or 'microdomains' *in vivo* with a diameter of about 200 nm, increasing in size under stress conditions [151]. The microdomains of Wsc1 indicate a moderate overlap with other Wsc-type sensor clusters, but not with those of the Mid-type sensors or other established membrane domains [156]. In this context, the CRD indicate an essential function in

clustering and signaling, since the exchange of cysteines to alanine within in the Wsc1-CRD resulted in an even Wsc1-distribution and a sensitivity towards cell wall stress, comparable to that of a *wsc1*Δ strain [151]. In line with this, the addition of dithiothreitol as a reducing agent for disulfides leads to a similar result in cells with the wild type Wsc1 [157]. In general, clustering of membrane sensors can enhance the cytoplasmic signal strength or might play a role in a threshold limiting for signal generation [87]. Since Wsc1 is expected to bind Rom2 by its cytoplasmic part, clustering has been suggested to form a 'Wsc1 sensosome' signaling complex [151]. However, the specific mechanism of the Wsc clustering and its molecular triggers remain unknown.

As mentioned previously, the presence of eight highly conserved cysteines, indicating the formation of four disulfide bridges, is a particularly large amount for a small protein-like the Wsc-CRDs. These cysteines are found in the Wsc family among different yeast species and in homologous domains of other species. Currently, there are 577 proteins with Wsc-domains provided by the *SMART* database [158]. The fungus *Trichoderma harzianum* has a Wsc domain with an exoglucanase (ThCRD2) function, which could indicate a carbohydrate binding function for the Wsc family of *S. cerevisiae*.as well. Moreover, the closely related yeast *Kluyveromyces lactis* has cell wall sensors with Wsc-domains, which are also involved in cell wall integrity [159]. In the more distantly related yeast *Schizosaccharomyces* pombe, Wsc1 homologs have also been identified, but they did not trigger the CWI pathway [160]. The human protein Krm1 has a Wsc-domain as well, which structure has been solved recently [161].

Krm1 is a transmembrane receptor, found in a ternary complex of the Wnt signaling pathway, which is important during embryonic development and tissue homeostasis [162]. The Wsc domain of Krm1 has been shown to be involved in binding of the Wnt antagonist Dkk1 [161]. Yet another Wsc domain is found in the mammalian protein Polycystin-1 (Pdk1) [163]. It has been shown to play an essential role in renal tubular morphogenesis, and malfunction causes cystogenesis in human autosomal-dominant polycystic kidney disease. Moreover, Pdk1 is also considered as an ion-channel regulator or to be involved in protein-protein and protein-carbohydrate interactions [163]. The specific function of the Wsc-domain in Pdk1 remains unknown.

2. Objectives

C. glabrata is the second most common cause of fungal infections worldwide. However, there is still demand for the development of novel antifungal treatments, since *C. glabrata* shows innate resistance to classic antimycotics. Surface exposed adhesins enable *C. glabrata* to form biofilms and the colonization of the human host. For this reason, adhesins represent important virulence factors and are considered as targets for drugs that may inhibit adhesion and hence avoid infections at an early stage. A large variety of adhesin-like proteins and subfamilies in *C. glabrata* imply different ligand specificities and actions of binding, which might be necessary for the colonization of different host surfaces and tissues. While extensive studies have been carried out on the Epa subfamily, knowledge about the structure and function of the Pwp subfamily remains yet poor. Therefore, a central aim of this work was to elucidate the structure and function of the adhesin-like Pwp subfamily from *C. glabrata* and to discover differences to the related Epa family. In order to gain insight into ligand specificity and the structural basis, binding studies and crystallographic analysis should be performed with heterologously produced A-domains of the Pwp subfamily. Subsequently, ligand affinities should be determined and the adhesion to different human tissues and abiotic surfaces should be investigated *in vivo* by the use of a heterologous *S. cerevisiae* expression system. From a prospective view, newly discovered functions of the Pwp family might find applications in medicine and biotechnology, such as novel strategies for inhibiting catheter-associated biofilms.

The five cell wall integrity sensors Wsc1, Wsc2, Wsc3, Mid2 and Mtl1 are part of a signaling cascade, detecting cell surface stress in *S. cerevisiae*. Since the CWI signaling cascade is essentially conserved throughout the fungal kingdom and required for cell survival, it represents a potential target for antifungal drugs. The presence of a cysteine-rich domain (CRD) discriminates the Wsc family from Mid2 and Mtl1, but the structural features and precise biochemical functions of the CRDs of the Wsc family are unknown. Therefore, a second key goal of this work was to characterize the CRDs from Wsc family sensors from a structural and biochemical view, in order to obtain detailed insights into their molecular and biological functions with respect to detecting cell wall stress. The Wsc CRDs should be produced heterologously for crystallographic studies as well as for the identification of potential glycan ligands. Moreover, clustering by homo- and heterotypic protein-protein interactions should be analyzed *in vitro*.

3. Results

3.1. Structural and functional characterization of the Pwp family

3.1.1. *In silico* analysis

For a better characterization of the structure and function of the Pwp family, a comprehensive bioinformatical analysis of the PwpA domains was performed as described in detail in paragraph (5.1). Basis for the analysis were primary protein sequences derived from *C. glabrata* CBS 138 (Tab. 3).

Tab. 3 Database entries of proteins that were analyzed in this study. Sequences were retrieved from the *UniProt Database* [164] and the *Candida Genome Database* [123].

Protein	UniProt-ID	CGD-ID
Pwp1	B4UN20	CAGL0I10147g
Pwp2	B4UN22	CAGL0I10246g
Pwp3	B4UN21	CAGL0I10200g
Pwp4	Q6FQ04	CAGL0I10362g
Pwp5	Q6FQ05	CAGL0I10340g
Pwp6	Q6FII8	CAGL0M14069g
Pwp7	Q6FQ10	CAGL0I10098g
Epa1	Q6VBJ0	CAGL0E06644g

Since the functional A-domains should be subject to following studies, borders of theoretical domain regions had to be determined precisely, excluding parts of the secretion signal and the stalk-like B-region (Fig. 14). This discrimination was performed in order to improve protein solubility, crystallization and to avoid positioning of domain borders within potential secondary structure motifs, which might affect glycan binding ability. The predicted *N*-terminal PA14 domains of the Pwp family comprise 201-203 amino acids in length. It is noteworthy that Pwp4 harbors a secondary short PA14-like domain within the repetitive B-region. This PA14-like domain has 29 % sequence identity to the adhesive domain of Cea1 of *Komagataella pastoris*. Moreover, Pwp1 has a remarkably long B-region comprising 32 tandem repeats, which outreaches the length of other PwpB and EpaB domains by far. A comparison of the Pwp adhesion domain phylogeny and identity highlights a relatively greater distance of Pwp7 to other Pwps, which are overall highly related (Fig. 15, Fig. 16).

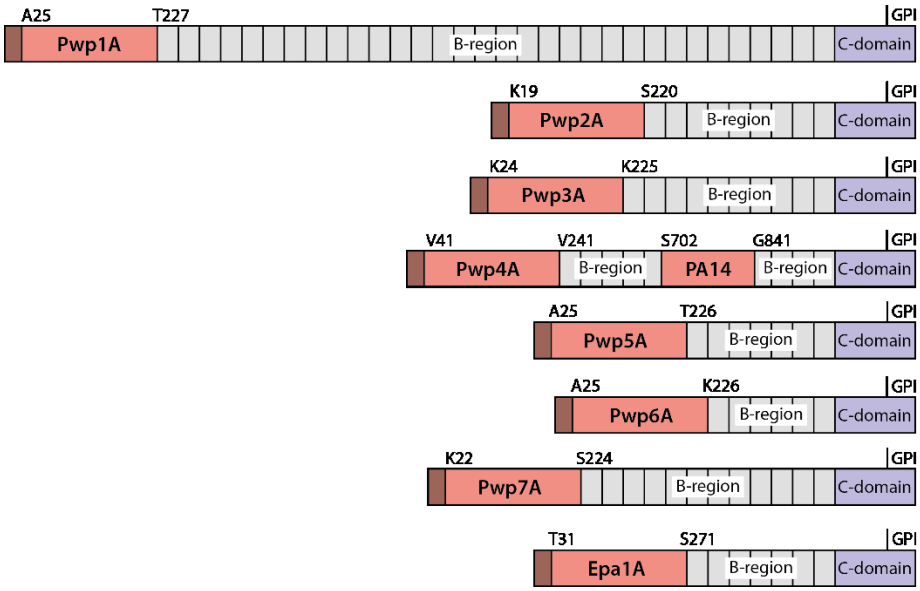

Fig. 14 Theoretical domain organization of the Pwp family from *C. glabratra*. An *N*-terminal secretion signal sequence (brown) is followed by the putative adhesion domain (red), which is followed by a Ser-/Thr- rich repetitive B-region (grey), variable in length. The C-domain (blue) is responsible for membrane anchoring via a GPI anchor. Pwp4 harbors a secondary putative PA14 domain within the B-region.

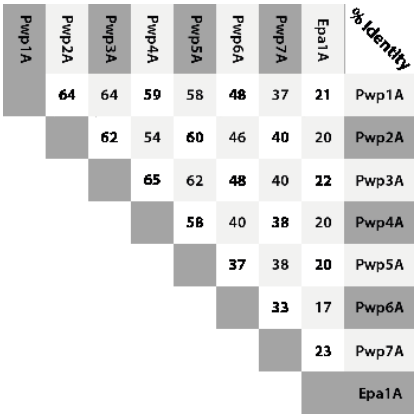

Fig. 15 Sequence identities of PwpA domains and Epa1A on the amino acid level.

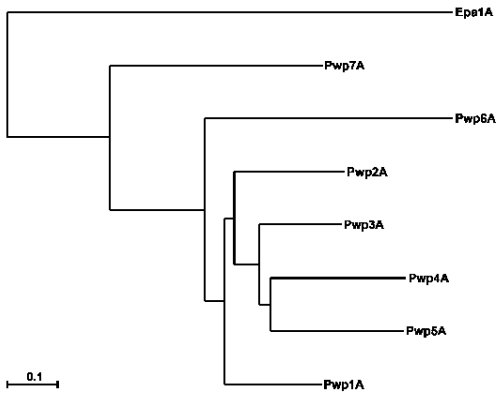

Fig. 16 Structure-based phylogenetic tree of PwpA domains.

The Pwp adhesion domains were further analyzed *in silico* with respect to theoretical information about physicochemical protein parameters as described in 5.1.2 and in order to estimate buffer compatibility, protein stability and protein solubility before heterologous expression and *in vitro* analysis (Tab. 4). This analysis also included a Pwp5A^{N131D} variant that was constructed for heterologous expression in order to analyze putative effects of the unusual asparagine substitution on CBL2 of Pwp5A *in vitro*. Overall, the PwpA domains share a molecular weight of approximately 25 kDa. They are theoretically hydrophilic and soluble, referring to the negative gravy indices based on the algorithm of KYTE and DOOLITTLE [165]. The fused 6×*His*-tag increases the pI of the PwpA domains by about one pH unit. The pI of Pwp3A is therefore at alkaline pH with a fused 6×*His*-tag while the pI of all other domains is in the acidic range.

Tab. 4 Theoretical physicochemical properties of PwpA domains and Epa1A determined with *ProtParam* [166].

Protein	A-domain	Mw [Da]	pI	pI (+ 6×*His*)	Ext. coefficient	Gravy index
Pwp1A	A25-T224	24855.77	5.23	6.69	45630	-0.22
Pwp2A	K19-S220	24927.05	4.5	5.47	36120	-0.14
Pwp3A	K24-K225	24873.99	6.35	7.7	43110	-0.24
Pwp4A	V41-V241	24725.73	5.09	6.17	43110	-0.17
Pwp5A	A25-T226	24390.32	4.66	6.35	36120	-0.12
Pwp5A^{N131D}	A25-T226	24391.3	4.51	6.11	35870	-0.12
Pwp6A	A25-K226	24899.97	5.12	6.28	47120	-0.37
Pwp7A	K22-S224	25039.10	4.87	6.1	45630	-0.3
Epa1A	T31-S271	29401.6	5.46	6.2	41175	-0.5

Also, putative glycosylation sites were determined as described (5.1.2), which suggests overall high glycosylation of the PwpA domains (Tab. 5). However, glycosylation is not assigned during heterologous expression in *E. coli*.

Tab. 5 Theoretical glycosylation sites of PwpA domains used in this study. Glycosylation sites were predicted with *NetNGlyc 1.0* and *NetOGlyc 4.0*.

Protein	*N*-glycosylation sites	*O*-glycosylation sites
Pwp1A	N121, N188, N196	-
Pwp2A	N134	-
Pwp3A	N139, N166	T74, T75, S79, T80, T81, T82
Pwp4A	-	T92, T93, S96, S97, T98, T78, T161
Pwp5A	N67, N140	S73, T75, T76, T79, T80, T81, T82, T83
Pwp6A	-	S70, S73, T74, S76, S80, T82,T83, S88, T89
Pwp7A	N57, N144, N150	T82, S149, S159
Epa1A	N140	-

3.1.2. Protein expression and purification

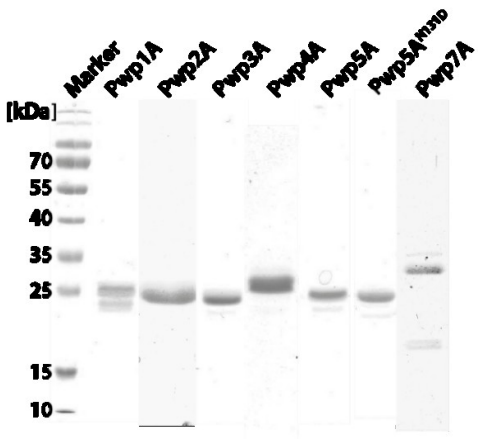

Fig. 17 PwpA domains in SDS-PAGE after size-exclusion chromatography (24.4-25 kDa).

In order to perform structural and functional *in vitro* analysis, the PwpA domains were produced and purified heterologously The construction of the pET-28(a) expression plasmids is described in chapter 5.2.15 and 6.12. The heterologous production was performed in *E. coli SHuffle*® T7 Express cells at 12 °C for 72 h, since these conditions were based on test-expressions, which resulted in highest solubility. Pwp6A, however, remained insoluble. Cell lysis and separation of the soluble fraction by ultra-centrifugation was followed by purification of the target proteins via Ni-NTA affinity- and subsequent size exclusion chromatography (Fig. 17).

Pwp2A showed strong retention on the SEC column, which might be explained through interaction with the *Superdex*® column material. Purified protein was stored sterile-filtered at either 4°C for short term usage or at -80°C for long term storage in 50% (v/v) glycerol. The yield of protein was 2-10 mg per liter of culture (Tab. 6).

Tab. 6 Approximate yields of 6×*His*-tagged PwpA domains after heterologous production in *E.coli*.

	Pwp1A	Pwp2A	Pwp3A	Pwp4A	Pwp5A	Pwp5A^{N131D}	Pwp6A	Pwp7A
[mg l^{-1}]	10	5	10	10	10	10	-	2

3.1.3. General protein analytics

Protein stability largely depends on solvent properties. Since PwpA domains greatly tended to precipitate under standard purification conditions, it was necessary to establish individual buffer systems in order to increase protein stability before further *in vitro* analysis was possible. Therefore, thermal shift assays (5.4.4) were performed under varying buffer conditions and additives with PwpA domains, as well as Epa1A as control (Fig. 18, Fig. 19, Fig. 20, supplements Fig. 21).

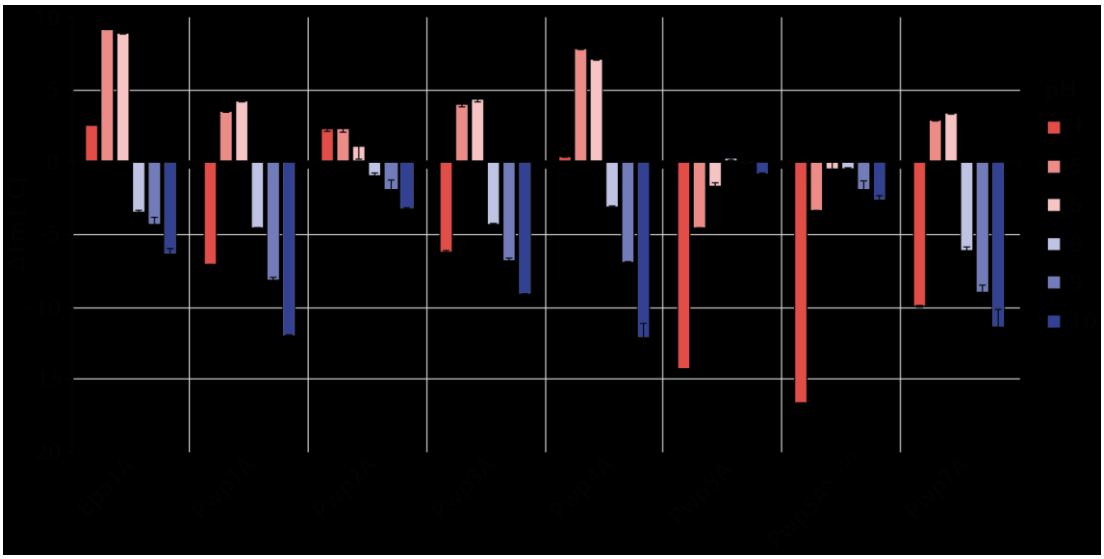

Fig. 18 Shifts in thermal stability of 6×*His*-tagged PwpA domains in a range of pH 4-10. Shown are changes of protein melting temperatures (T_m) in different buffer systems with reference to pH 7 (5.4.4). Samples contain the fluorescent dye *SYPRO Orange* ™ (7.8×) and protein at a concentration of 10 µM.

Changing pH parameters highly influenced the thermal stability of PwpA domains, whereas most PwpA domains and Epa1A showed maximum stability in a slightly acidic environment (pH 5-6) (Fig. 18). In contrast, Pwp5A and Pwp5A^{N131D} showed different stability profiles, since they were most stable at neutral pH and tended to precipitate in an acidic environment. However, the pH-dependent thermal stability profiles do not correlate entirely with the isoelectric points (Tab. 4) since proteins usually become less stable and tend to aggregate at a pH close to the proteins' pI [167].

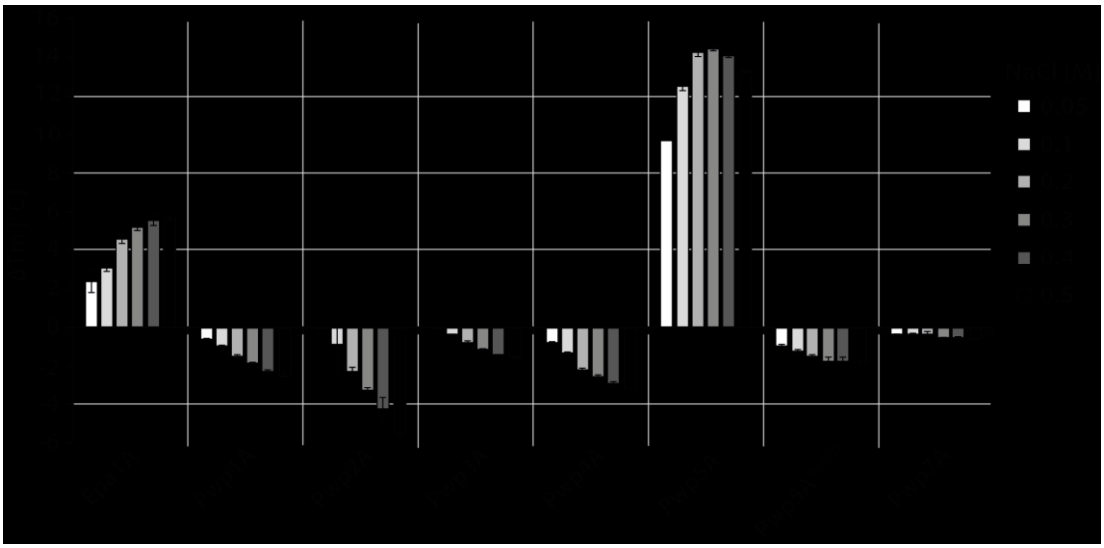

Fig. 19 Shifts in thermal stability of 6×*His*-tagged PwpA domains in a range of 0.05 – 0.5 M NaCl. Shown are changes of protein melting temperatures (T_m) with reference to 0 M NaCl (5.4.4). Samples contain the fluorescent dye *SYPRO Orange* ™ (7.8×) and protein at a concentration of 10 µM and pH 6.0 (Pwp5-samples at pH 7.0).

A striking observation was the high difference in protein stability at a high ionic strength between the natural Pwp5A domain and the $Pwp5A^{N131D}$ variant (Fig. 19). The $Pwp5A^{N131D}$ variant, along with other Pwp adhesion domains, destabilized with increasing NaCl concentrations, whereas the natural Pwp5A domain tolerated high NaCl concentrations up to 500 mM, similar to Epa1A. Since cofactor binding most likely results in increased protein stability and since the Pwp are considered as putative C-type lectins, Ca^{2+} and other divalent cations were screened in thermal shift analysis with PwpA domains (Fig. 20).

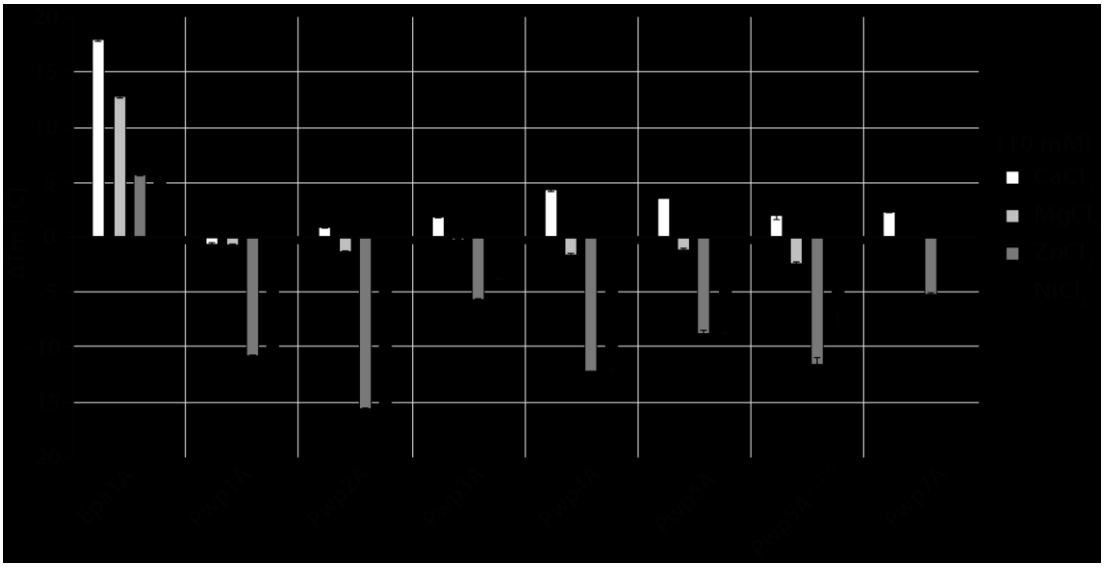

Fig. 20 Shifts in thermal stability of 6×*His*-tagged PwpA domains with different divalent cations. Shown are changes of protein melting temperatures (T_m) with reference to a sample without divalent cations (5.4.4). Samples contain the fluorescent dye *SYPRO Orange* ™ (7.8×) and protein at a concentration of 10 µM after EDTA treatment.

The addition of calcium stabilized the PwpA domains, except Pwp1A. However, calcium binding led to a much stronger intrinsic stabilization of Epa1A. The addition of other divalent cations led to decreasing thermal stability for PwpA domains, whereas Epa1A was stabilized by all of them. The observed destabilization of PwpA domains by divalent cations might be due to an ion strength effect, since the PwpA domains are sensitive to higher NaCl concentrations, although calcium may represent a cofactor. In addition, the tested ionic compounds lead to an ionic strength effect in the solution that is three times larger than an equal amount of NaCl. However, the thermal stability of the Pwp5A^{N131D} variant also decreased under the addition of divalent cations other than calcium although it showed high salt tolerance.

Glycerol is often added to protein buffers since it is known to promote protein stability and to reduce aggregation. Therefore, glycerol was added to PwpA domains in order to analyze the impact on thermal stability. Glycerol increased the melting temperature along with an increasing gradient to up to 1.7 °C for the Pwp family and up to 3 °C for Epa1A, whereas the thermal stability of Pwp7A was not promoted by the addition of glycerol (supplements, Fig. 21). Based on the whole results, final buffer solutions for further analysis were established (Tab. 40 in chapter 6.10).

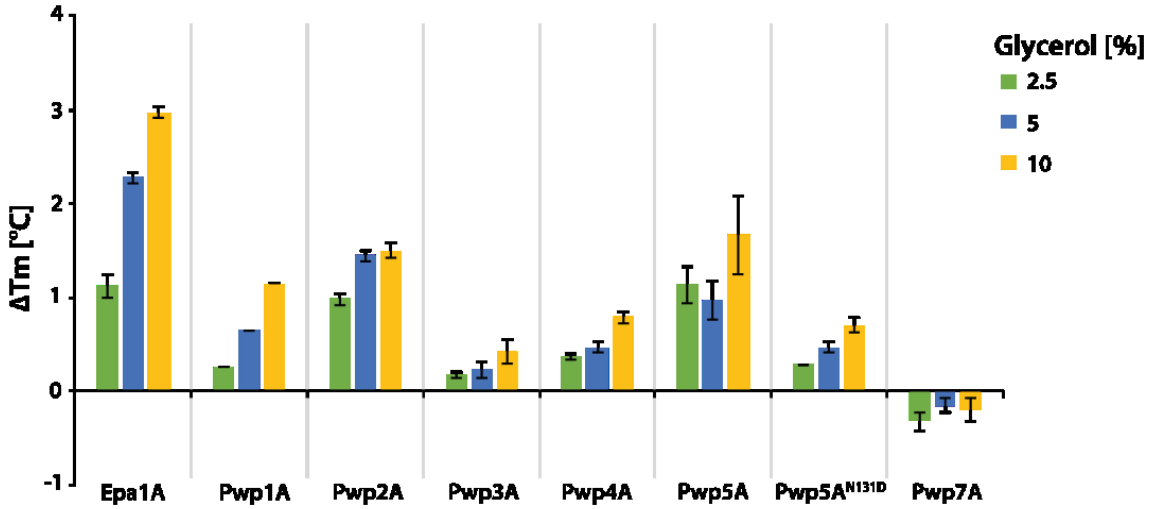

Fig. 21 Thermal stability of PwpA domains under different glycerol concentrations with reference to H_2O.

3.1.4. PwpA domains are PA14/Flo5-like domains with unique structural features

A central aim of this work was to solve the 3D-structures of PwpA domains by X-ray diffraction analysis, in order to gain insight into their specific structural properties and to compare the PwpA structures with other known adhesion domains. Since X-ray diffraction analysis requires the production of well-ordered protein crystals, large-scale crystallization screens were performed (5.5.1) and several conditions resulted in crystalline material. Pwp1A successfully crystallized in conditions #27 and #72 of JCSG Core I at 4 °C and at a protein concentration of 50 mg ml^{-1} (Fig. 22). Pwp5A crystallized in conditions #28 (4 °C), #39 (4 °C) and #54 (18 °C) of JCSG Core I and in #18 (4 °C) of JSCG Core II at a protein concentration of 30 mg ml^{-1} (Fig. 23). Initial crystal growth of both domains was observed after a remarkably long incubation time of approximately one year.

Fig. 22 Crystals of Pwp1A. **A**: Condition #27 of screen JCSG Core I. The framed crystal of this condition was used for structure solution. **B**: Condition #72 of screen JCSG Core I. Scale bar: 100 µm.

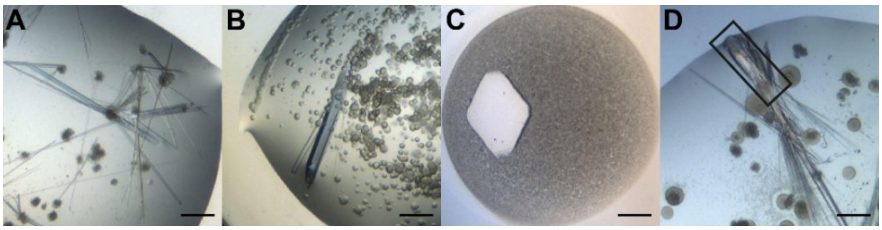

Fig. 23 Crystals of Pwp5A. **A**: Condition #28 of JCSG Core I. **B**: Condition #18 in JCSG Core II. **C**: Condition #54 in JCSG Core I. **D**: Condition #39 in JCSG Core I. The crystal in this condition which was used for structure determination is marked in the black frame. Scale bar: 100 µm.

The crystallization of Pwp3A, Pwp4A and Pwp5A^{N131D} resulted in either fuzzy needle-like clusters or in thin needles with poor diffraction. In the case of Pwp2A and Pwp7A, crystal growth remained completely absent, perhaps as a consequence of missing ligands or poor structural rigidity. However, whenever crystal growth was observed for a PwpA sample, reproduction and optimization, as well as co-crystallization setups were performed in a 24 well/hanging drop scale. A Pwp1A crystal from condition #27 in JCSG Core I and a Pwp5A crystal from condition #39 in JCSG Core I resulted in sufficient diffraction, while other crystals resulted in poor diffraction. Compositions of the respective crystallization conditions for Pwp1A and Pwp5A are shown in Tab. 7.

Tab. 7 Crystallization conditions of Pwp1A and Pwp5A.

Pwp1A - JCSG Core I #27		**Pwp5A - JCSG Core I #39**	
Sodium cacodylate	100 mM	Sodium chloride	200 mM
PEG 8000	5 % (w/v)	PEG 3350	20 % (w/v)
pH	6.5		
2-Methyl-2,4-pentanediol	40 % (v/v)		

Four datasets were recorded for Pwp1A and two datasets were recorded for Pwp5A. After data reduction and scaling, both protein structures were solved by molecular replacement (Tab. 8).

Tab. 8 Data collection statistics of PwpA crystals.

	Pwp1A	**Pwp5A**
Dataset	CC227A_x2	BL3_1
Date	28.02.2018	18.12.2015
X-ray source	ID29, ESRF Grenoble	ID29, ESRF Grenoble
Detector	Pilatus 6M	Pilatus 6M
Wavelength (Å)	0.984	1.03
Space group	$P4_12_12$	$P4_1$
Cell dimensions (Å)		
a	133.96	96.5
b	133.96	96.5
c	71.36	44.34
α, β, γ	90°, 90°, 90°	90°, 90°, 90°
Molecules / a.s.u.	2	2
Resolution (Å)	66.98 – 1.85 (1.88 – 1.85)	96.5 - 1.64 (1.73 – 1.64)
Total reflections	824390	153140
Unique reflections	56123	37340
Multiplicity	14.7 (15.1)	4.1 (3.4)
Completeness (%)	100 (100)	74.4 (25.5)
R_{merge} (%)	0.11 (1.16)	6.6 (48.9)
Mean I/σ(I)	14.1 (2.2)	11.7 (2.5)
Wilson B-factor (Å^2)	35.2	20.2
Mosaicity (°)	0.2	0.2

The structure of Pwp5A was solved on basis of a *Swiss-model* generated, Flo5A-based homology model of Pwp5A. Afterwards, the structure of Pwp1A was solved on basis of the solved Pwp5A structure (Tab. 9).

Tab. 9 Refinement statistics for PwpA molecules.

	Pwp1A	Pwp5A
Resolution (Å)	50.1 – 1.87	68.24 - 1.64
R_{work}	17.29	14.75
R_{free}	20.01	19.77
Completeness (%)	99.99	74.13
r.m.s.d.		
Bond length (Å)	0.02	0.012
Bond angle (°)	1.872	1.534
Total number of atoms	3409	3517
Mean B-value (Å^2)	35.22	25.1

Pwp1A and Pwp5A are PA14 domains with the typical architecture of a β-sheet sandwich, which show a relatively high structural conservation with members of the Epa family, despite a low sequence identity of about 20 % (Fig. 24). The PA14 core is composed of two opposing β-sheets, which build the hydrophobic core of the domain. A five-stranded β-sheet comprises β9β11β7β2β3 (Pwp1A) and β10β8 β12/13β2/β3β4 (Pwp5A) opposite to the seven-stranded β-sheet, which comprises: β10β8 β12/β13β14β4β5β6 (Pwp1A) or β11β9β14β15β5β6β7 (Pwp5A). The flexible loop L1 of both PwpA domains is mostly random coiled but also possesses two short α-helices (α1/α2).

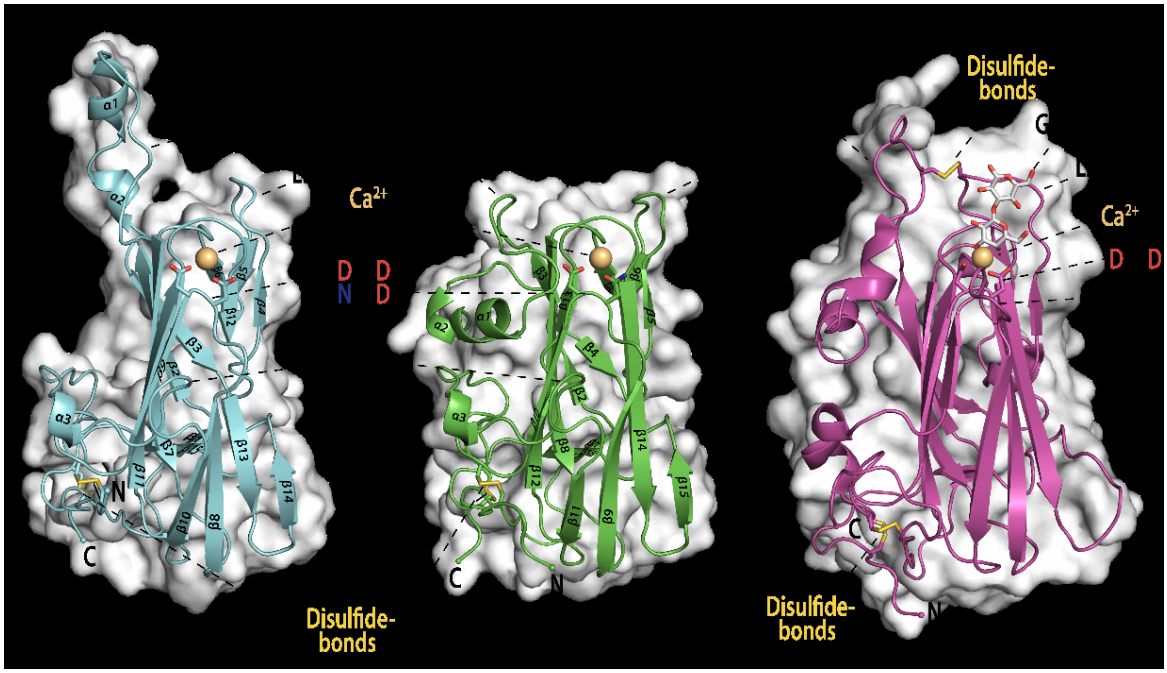

Fig. 24 Structures of Pwp1A (A, cyan) and Pwp5A (B, green) in comparison to Epa1A(C, magenta). The structures of Pwp1A and Pwp5A share a conserved PA14/Flo5-like β-sheet sandwich core. Both domains have two conserved disulfide bonds in the *N*-/*C*-terminal region and a D*cis*D motif (Pwp1A) or N*cis*D motif respectively (Pwp5A) on loop CBL1, which helps to coordinate Ca^{2+}-ion. The loop L1 in Pwp1A shows a turned orientation compared to Pwp5A.

The most striking observation was that loop L1 of Pwp1A is disconnected from the β-sheet core, which discriminates it structurally from Pwp5A and other known fungal

PA14/Flo5-like lectins. L1 of Pwp1A is turned by approximately 90°, leaving an open gap between α3 and the β-sheet, which might function as ligand binding site. The L1 of Pwp5A, in contrast, remains connected to the β-sheet core, such as in EpaA domains. Moreover, the loop L2 is much shorter than loop L1 in both PwpA domains, although L1 is also reduced in length by 8-9 amino acids compared to Epa1A. Another striking observation is that the PwpA domains show a subtype-specific pattern of disulfides and cysteine conservation (Fig. 25, Fig. 26).

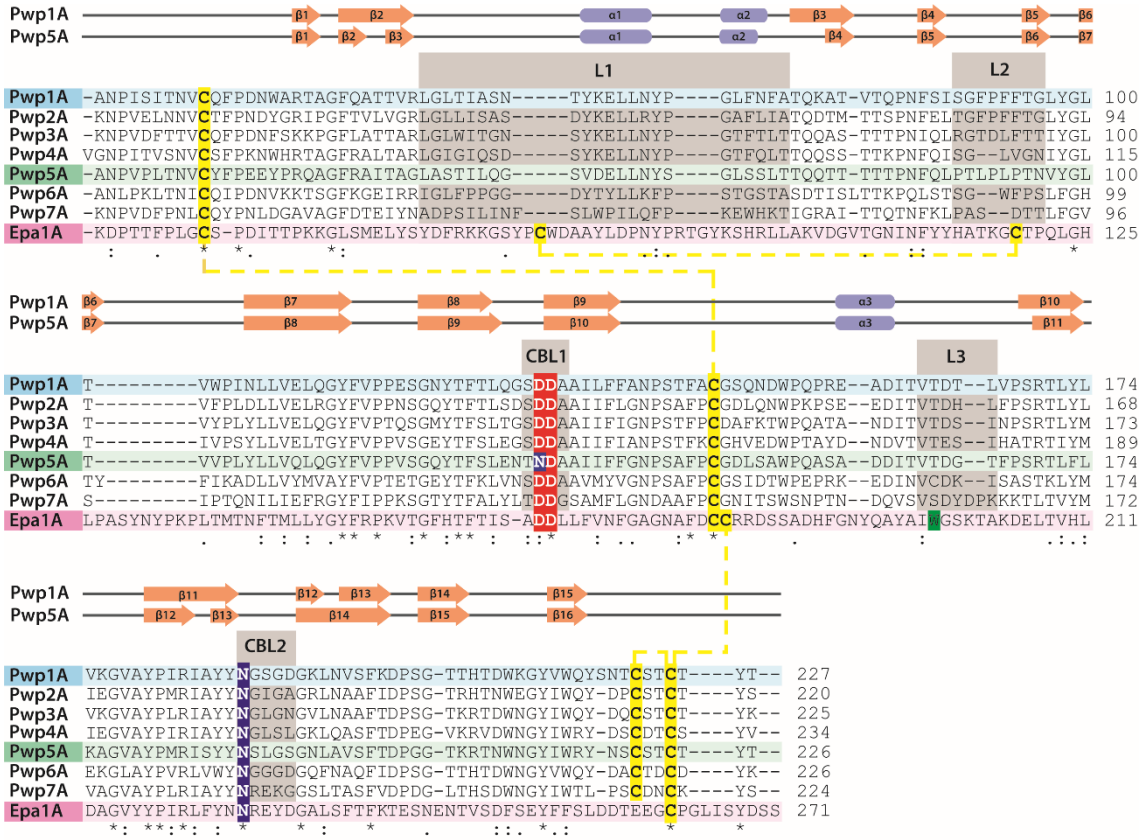

Fig. 25 Structural alignment of PwpA domains in comparison to Epa1A. PwpA domains comprise four conserved cysteines (yellow) that but lack two cysteines, which are conserved in EpaA domains and connect the outer loops L1 and L2 through a disulfide bond. Corresponding loop regions L1, L2 and L3 of the Pwp family are reduced in length by a number of amino acids compared to Epa1A. L3 of the Pwp family lacks aromatic residues as it is found in the Epa family. The PwpA domains lack a CC-motif, which is conserved in the Epa family and other fungal PA14/Flo5-like adhesins. The CBL1 region of the Pwp family shares a conserved motif of two consecutive aspartic acids together with the Epa family (red) and a conserved asparagine residue on CBL2 (blue), which are altogether involved in Ca^{2+}-recognition. In Pwp5A an aspartic acid on CBL1 is substituted by an asparagine. Structural sequence alignment was performed with *TCoffee Expresso*.

The Pwp family lacks two cysteines, which fuse the two front loops L1 and L2 in EpaA domains by the formation of a disulfide bond. Such a disulfide bond generally enhances structural rigidity, restricts the mobility of L1 and shields the ligand binding site from surrounding solvent. However, an L1-L2 disulfide bond is also absent in other fungal PA14/Flo5-like lectins such as *Kp*Cea1A, *Sc*Flo5A and Lg-Flo1A [93,97,168], indicating that a covalent fusion of L1-L2 is not mandatory for effective glycan binding. The first *N*-terminal and the last *C*-terminal cysteines remain highly conserved in the Pwp family,

similar to other PA14/Flo5-like lectins such as the epithelial adhesin-, flocculin-, *Pichia*- and *Kluyveromyces*-subtypes (Fig. 8). In contrast to the Pwp subtype, other subtypes have a conserved CC-motif of two consecutive cysteines in the central part of the PA14 domain (here represented by Epa1A), which build a link between the *N*-terminal part and the *C*-terminal part, providing high structural rigidity. In the Pwp subtype, the CC-motif is reduced to one cysteine that is connected to the *N*-terminal cysteine. The *C*-terminal cysteine, in contrast, builds a disulfide with a subtype-specific cysteine that is separated by only two intermediate amino acids, creating a short loop. As a result, the *C*-terminal region, which connects the adjacent repetitive B-region, differs structurally from other known PA14/Flo5-like domains and shows increased disorder in the *N*-/*C*-terminal region, indicating for a more flexible link.

Apart from the missing CC-motif, the central parts of PwpA domains appear very similar, representing the structurally conserved β-sandwich core of the PA14/Flo5-type. Both PwpA structures show conservation of the features required for Ca^{2+}-binding: an asparagine on CBL2 and the common *cis*-peptide bond on loop CBL1 referred to as D*cis*D (Pwp1) and N*cis*D (Pwp5). The latter has not yet been observed in PA14/Flo5-like adhesins from other sources than *C. glabrata*. Interestingly, the unique asparagine substitution still allows the coordination of a Ca^{2+}-ion together with the side chain of the conserved asparagine and carbonyl groups of peptide bonds on CBL2 (Fig. 26). The Ca^{2+}-ion has been shown to directly participate in glycan ligand binding in the Epa family and other PA14/Flo5-like adhesins [70,96,108].

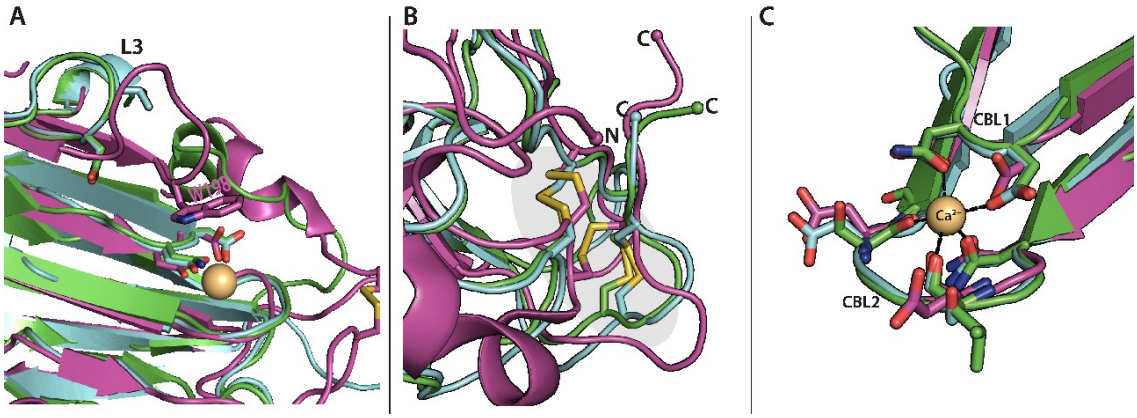

Fig. 26 Superimposition of Pwp1A, Pwp5A and Epa1A. A: L3 of Epa1A (magenta) carries a conserved tryptophan W198, covering the Ca^{2+} binding site. L3 of Pwp1A (cyan) and Pwp5A (green) are withdrawn from the Ca^{2+}-binding site and do not carry aromatic residues. **B**: Different disulfide patterns at the *C*-termini. The conserved CC-motif of EpaA domains and other fungal adhesins is reduced to one in Pwps, which builds a disulfide with an additional Pwp-specific cysteine, creating a short *C*-terminal loop. **C**: Coordination of the Ca^{2+}-ion by the conserved D*cis*D motif on CBL1 (N*cis*D in Pwp5A) together with the side chain of a conserved asparagine and carbonyl groups of peptide bonds on CBL2.

Another striking difference to other PA14/Flo5-like domains can be observed in the top loop L3 of PwpA domains. In Epa1A for instance, the loop L3 reaches the calcium binding site where it presents a tryptophan residue (W198) that is conserved throughout the Epa family. That tryptophan residue covers the binding site and has been shown to enhance the affinity to terminal galactose by forcing a coplanar orientation [108]. In *Kp*Cea1A for instance, the L3 lacks aromatic residues but covers the binding site with an elongated loop, resulting in a different orientation of the terminal hexose ligand, compared to Epa1A [97]. In contrast, aromatic residues are absent on L3 of the PwpA domains and besides, the entire loop L3 is shorter and withdrawn from the calcium binding site. This results in the absence of a shielded binding pocket and exposes the binding-site with the Ca^{2+}-ion to surrounding solvent (Fig. 28), which seems unfavorable for the binding of hexose ligands. A similar exposition of the Ca^{2+}-binding-site can be found in *Sc*Flo5, which has been shown to specifically bind mannose with low affinity in the millimolar range [93].

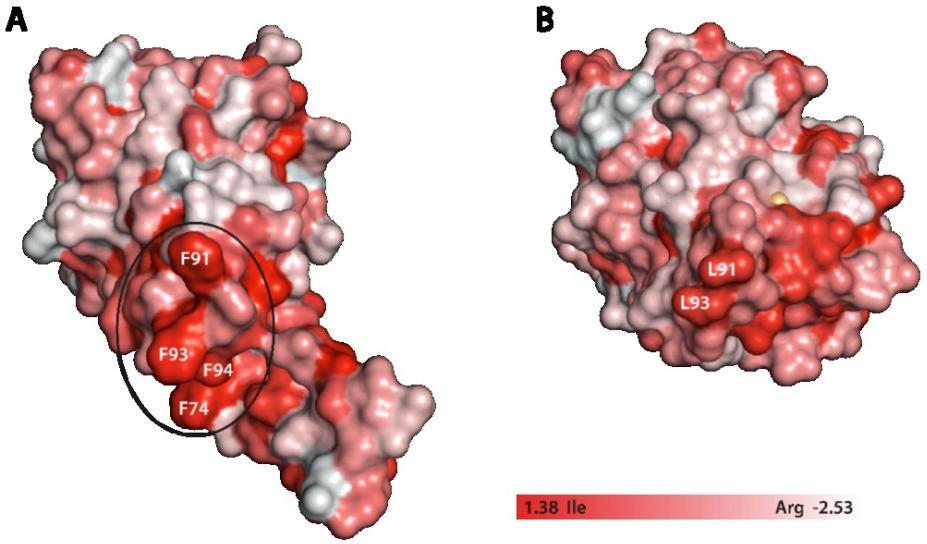

Fig. 27 Surface hydrophobicity of Pwp1A (A) and Pwp5A (B). Shown is the relative hydrophobicity, based on the EISENBERG scale with hydrophobic areas in red and hydrophilic areas in white [169]. The short loop L2 is highly hydrophobic with three phenyalanines in Pwp1A (F91, F93 and F94) and two leucines in Pwp1A (L91 and L93). Loop L1 of Pwp1A contains also an opposing phenylalanine F74, creating a highly hydrophobic region in the gap between L1 and L2.

A comparison of the surface hydrophobicity of Pwp1A and Pwp5A shows that the loops L2 of both domains are much more hydrophobic, compared to the rest of the protein surface (Fig. 27). In particular, the region between loop L2 and L1 of Pwp1A is highly hydrophobic with three surface-exposed phenylalanine residues and an opposing one on loop L1.

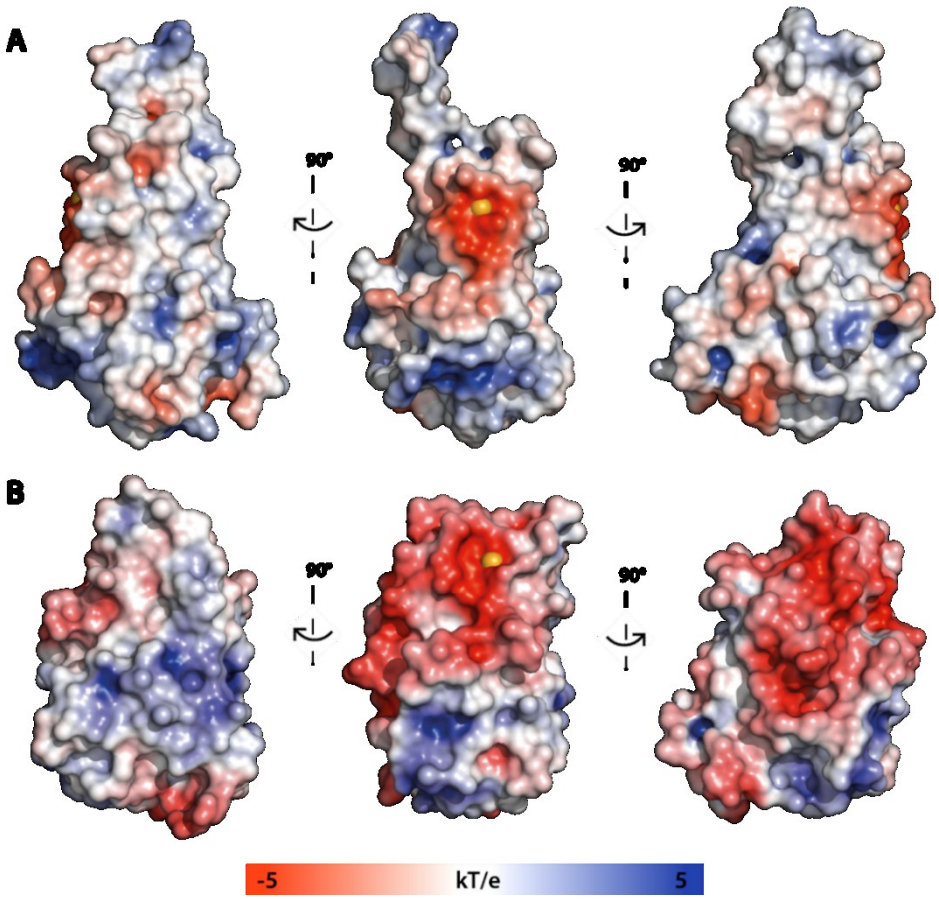

Fig. 28 Electrostatic surface potential of Pwp1A (A) and Pwp5A (B). The calcium binding-site of Pwp5A is surrounded by a negatively charged surface area. Both domains have positively charged areas close to the N- and C-termini. Pwp5A has a much higher positively charged surface area, whereas Pwp1A appears mainly non-polar. Surface potential was determined with *APBS* at pH 6.0 and 0.15 M NaCl.

The electrostatic surface potential of Pwp5A reveals an accumulation of negative charge that largely surrounds the Ca^{2+}-binding site and expands laterally, whereas the opposing lateral side is dominated by positive charges (Fig. 28). In Pwp1A a corresponding area with negative charges is restricted to the direct environment of the Ca^{2+}-ion. However, the overall surface of Pwp1A appears to be mainly non-polar. These differences in surface potential between Pwp1A and Pwp5A correlate with the observed differences in structural stability under changing pH conditions (3.1.3) and differences in the pI (Tab. 4). Moreover, the observed discrepancy in NaCl tolerance between Pwp5A and other PwpA domains correlates with the difference in electrostatic surface potential, probably due to Na^{2+}-interactions. However, both domains show positively charged areas, which are located mainly close to the *N*- and *C*- termini.

3.1.5. *In vitro* glycan binding studies

In a next step, the glycan binding capability of the Pwp family was analyzed in different experiments. First, fluorescently labeled PwpA domains were screened for interaction with a library of 609 mammalian glycan ligands which were printed on a microarray (Fig. 29, Fig. 30). The assay was performed in cooperation with the *Consortium for Functional Glycomics* as described in chapter 5.4.6.

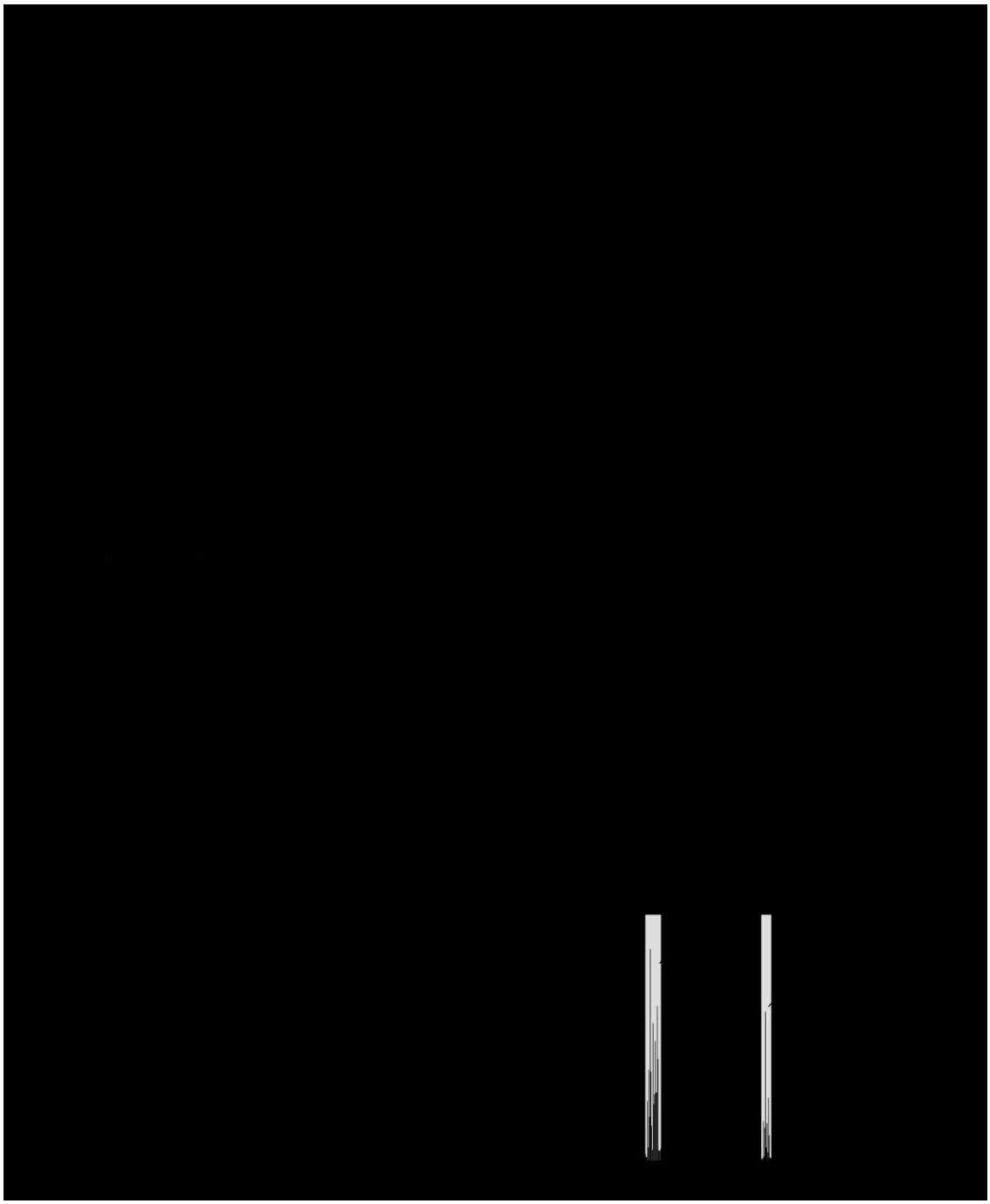

Fig. 29 *CFG* mammalian glycan array with PwpA domains and Epa1A. **A-H:** Quantification of the relative fluorescence of bound protein at a concentration of 8 µM (PwpA) and 70 nM (Epa1A) in CFG#3059 & CFG#3171 (**A**: Pwp1A, **B**: Pwp2A, **C**: Pwp3A, **D**: Pwp4A, **E**: Pwp5A, **F**: Pwp5A^{N131D}, **G**: Pwp7A, **H**: Epa1A). Labeled are terminal saccharides of corresponding glycan compounds. Glycan binding by PwpA domains appears unspecific or absent, whereas Epa1A is highly specific for terminal galactosides.

Quantification of the *CFG* mammalian glycan array showed that glycan binding by PwpA domains was unspecific or absent, whereas the EpaA control showed a clear preference for terminal galactose moieties, as expected. The overall binding profiles of the PwpA domains (Fig. 30) rather correspond to the distribution of terminal glycans on the array than indicating binding specificities, whereas the glycan binding profile of Pwp3A may suggest a preference for terminal galactosides with a majority of 42 % among the best binding glycans. However, potential glycan-binding by PwpA domains might depend on specific buffer conditions that differ from that of Epa adhesion domains. In contrast to the Pwp5A^{N131D} variant, the natural Pwp5A showed a reduced capability for interaction with terminal *N*-acetylglucosamine. Moreover, Pwp5A showed a reduced interaction with glycans in comparison to the Pwp5A^{N131D} variant. Binding of differently sulfated galactose residues was observed consistently among the Pwp family, however with inconsistent sulfation patterns and binding profiles.

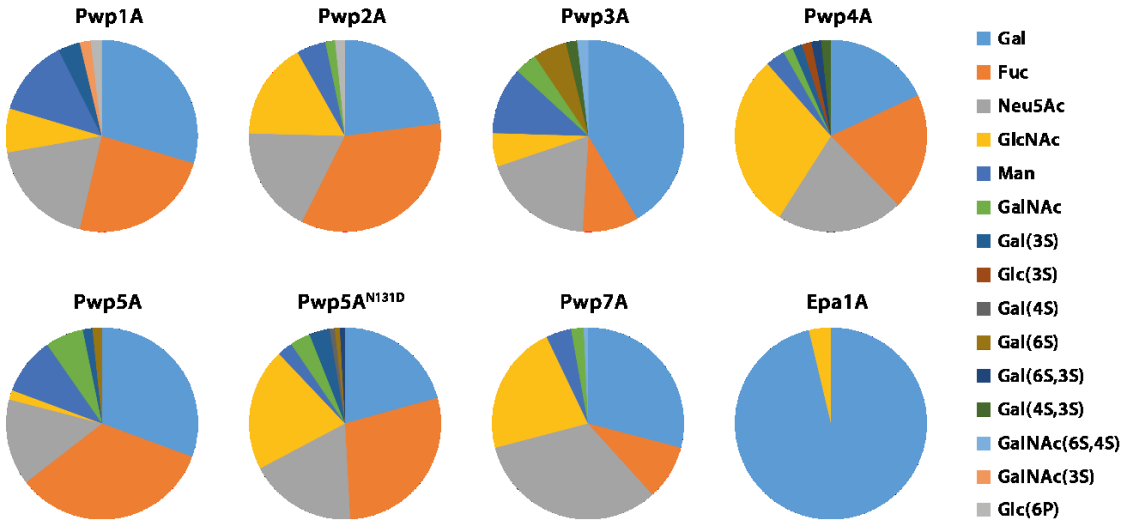

Fig. 30 Quantification of fluorescently labeled PwpA domains on the *CFG* mammalian glycan array in comparison to Epa1A. Incorporated are terminal glycans of unbranched sugars and branched sugars with equal terminal glycans, which achieved a minimum of 80 % of the maximum observed fluorescence intensity.

In a second approach, a pathogen glycan ligand microarray analysis with 140 compounds (Tab. 50), performed in cooperation with coworkers of Prof. Dr. SEEBERGER, revealed binding of Pwp3 and Pwp7A to the synthetic iduronic acid IdoA-2,4-disulfate (Fig. 31), which has been reported to interact with the human chemokine CCL20 with micromolar affinity [170]. Iduronic acid compounds were not printed on the mammalian glycan array. Natural iduronic acid is found either mono-sulfated as 2-*O*-sulpho iduronic acid (IdoA(2S)) or non-sulfated (IdoA) as part of the glycosaminoglycans heparin/heparan sulfate and dermatan sulfate. Pwp3A also bound the disaccharide IdoA(2S)(α1-3)GalNAc(4S), which makes a main component of dermatan sulfate (Fig. 31). Heparin, heparan sulfate and dermatan sulfate have been reported to serve as ligands for microbial and viral attachment as well as

internalization [137]. However, the non-sulfated form of iduronic acid binds to none of the PwpA domains analyzed on this array.

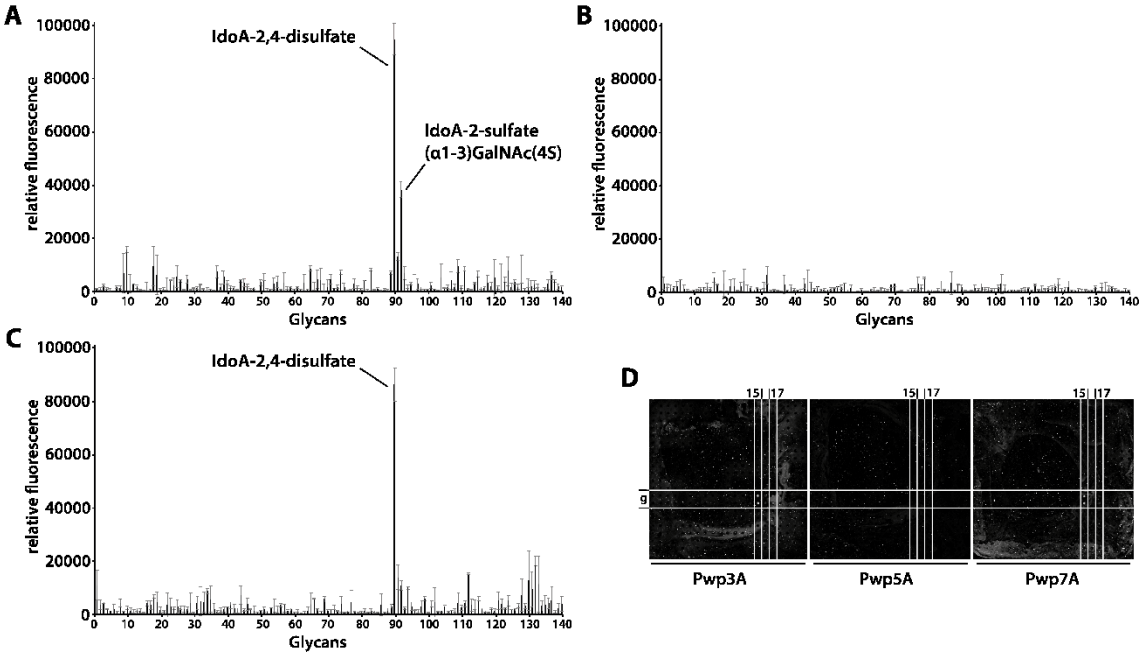

Fig. 31 Binding of fluorescently labeled PwpA domains to pathogen glycan ligands on a microarray. A-C: Quantification of the relative fluorescence of glycan-bound protein at a concentration of 8 µM. Labeled are the best binding glycan compounds. (**A**: Pwp3A, **B**: Pwp5A, **C**: Pwp7A). Pwp3A and Pwp7A bind IdoA-2,4-disulfate (g 15) and Pwp3A additionally binds IdoA-2-sulfate(α1-3)GalNAc(4S) (g17). **D:** Binding of fluorescently labeled PwpA domains to heparin-like oligosaccharides on a microarray with 140 compounds (n = 2).

In a next step, a thermal shift analysis was performed in order to further analyze the binding of PwpA domains to a variety of sulfated and non-sulfated monosaccharides of which several are found as part of glycosaminoglycans (Fig. 32).

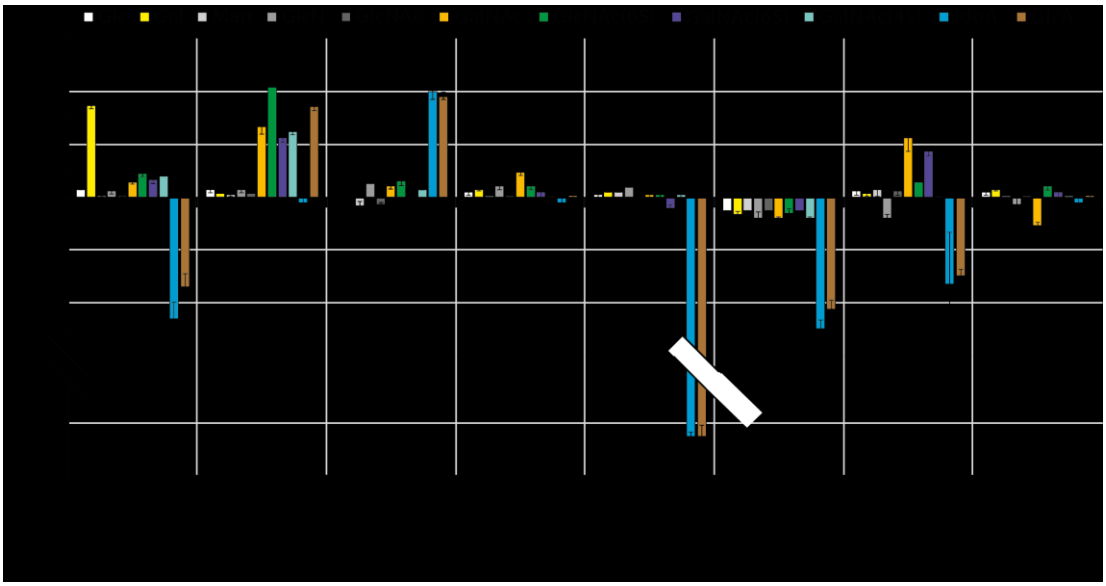

Fig. 32 Thermal shift analysis of PwpA domains with different monosaccharides. Thermal shifts were determined at a protein concentration of 10 µM with the fluorescent dye *SYPRO Orange*®. Melting temperatures were converted to shifts in relation to an H_2O control. Thermal shift reactions were performed in buffer with 20 mM MES at pH 6.0, 10 mM $CaCl_2$ and 50 mM NaCl. Shifts ≥ 1.5 °C are considered as significant. Epa1A served as glycan-binding control in combination with galactose.

Significant positive thermal shifts ≥ 1.5°C were observed for Pwp1A in combination with GlcNAc(6S) and GlcA, indicating interactions on a comparable level to Epa1A with galactose (Fig. 32). It is noteworthy, that although GlcA (β-D-Glucopyranuronic acid) induces positive thermal shifts for Pwp1A, its stereoisomer IdoA (α-L-Idopyranuronic acid) has no effect on Pwp1A. Such discrimination at similar concentrations indicates against a potential pH-related stability effect. However, IdoA stabilizes exclusively Pwp2A, whereas it strongly destabilizes Pwp4A, and moderately destabilizes Pwp5A, Pwp5A^{N131D} and Epa1A. The observation of an IdoA-induced destabilization is always accompanied by a GlcA-induced destabilization. In this context, the binding of glycan ligands may also lead to a negative thermal shift. However, it seems unusual and in theory, such an interaction should result in a positive thermal shift. While GlcA is found as part of all types of glycosaminoglycan, except keratan sulfate, IdoA is exclusively found in heparin/heparan sulfate and dermatan sulfate, both sulfated and non-sulfated. Furthermore, the monosaccharide GlcNAc(6S), which induced shifts for Pwp1A, is found as part of heparin/heparan sulfate and keratan sulfate. In this context, the sulfate group of GlcNAc(6S) appears to be required for Pwp1A binding, since the non-sulfated variant GlcNAc does not induce thermal shifts for this domain.

3.1.6. *In vitro* heparin-binding analysis

A glycan microarray with thirty-six printed heparin-like compounds [171,172] was screened for binding to Pwp3A, Pwp5A and Pwp7A.

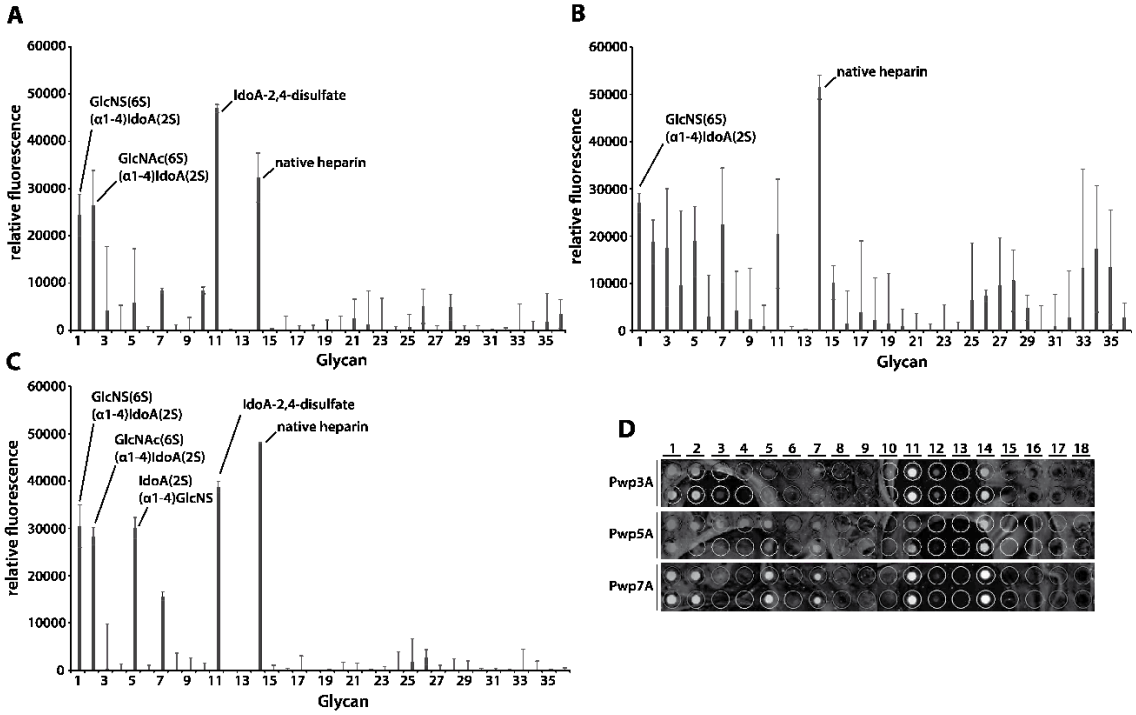

Fig. 33 Binding of fluorescently labeled PwpA domains to heparin-like compounds on a microarray. A-C: Quantification of the relative fluorescence of glycan-bound protein at a concentration of 200 µg/ml (**A**: Pwp3A, **B**: Pwp5A, **C**: Pwp7A). Labels show the terminal saccharides of best binding heparin-like compounds. **D**: Binding of fluorescently labeled PwpA domains to heparin-like oligosaccharides on the microarray ($n = 2$). Corresponding full names of glycans on the array are shown in Tab. 51.

All three PwpA domains showed strong binding to several oligosaccharides, which are naturally occurring in the glycosaminoglycans heparin and heparan sulfate (Fig. 33, Fig. 34).

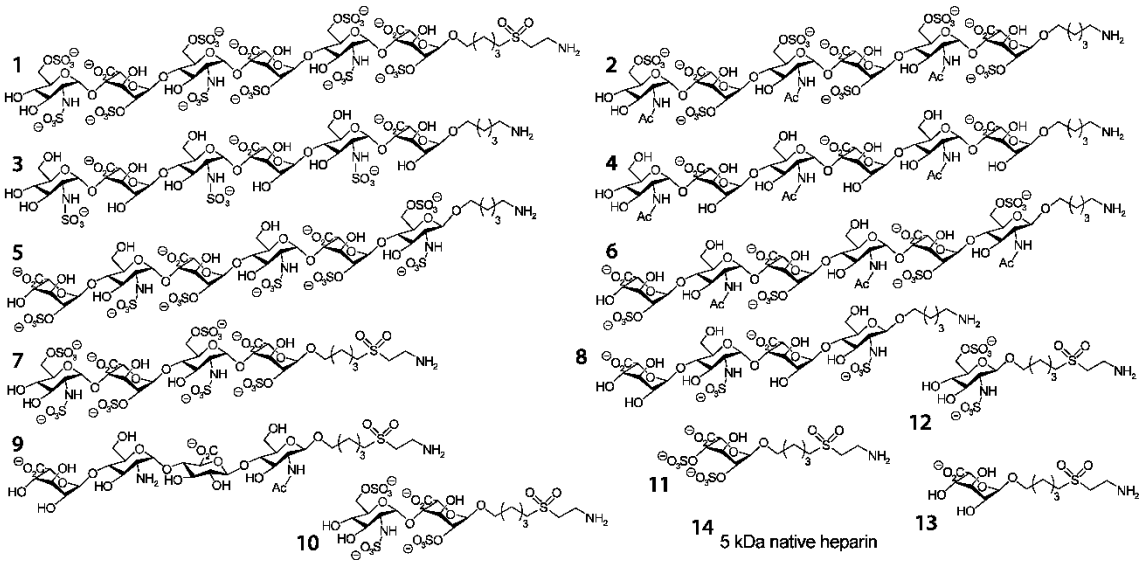

Fig. 34 Heparin-like compounds tested on the heparin microarray. Numbers donate the different glycan structures that are listed in. Tab. 51 and shown in Fig. 33.

Pwp7 additionally showed strong binding to the heparin-like tetrasaccharide #5, whereas this compound bound weakly to Pwp3A and moderately to Pwp5A. The synthetic iduronic acid IdoA-2,4-disulfate again strongly interacts with Pwp3A and Pwp7A but shows weaker signals for Pwp5A. Additionally, the tested PwpA domains strongly bind to native heparin. It is noteworthy, that those glycan compounds on the array, which showed strong interaction with PwpA domains, possess at least one or more iduronic acids that carry a sulfate group on carbon position two. However, the heparin-like hexasaccharide #6, which contains three sulfated iduronic acids did not bind to any of the tested PwpA domains. Moreover, the results indicate that the overall affinity of PwpA domains to heparin-like compounds correlates with increasing sulfation levels.

Since PwpA domains showed binding to heparin-like compounds, thermal shift analysis with PwpA domains were performed in combination with low molecular weight heparin (*Dalteparin*), the synthetic heparin pentasaccharide *Fondaparinux*, unsaturated disaccharide subunits of dermatan sulfate (ΔIdoA(β1-3)GalNAc & ΔIdoA(β1-3) GalNAc(4S)) and DNA as negative control, due to its polyanionic properties (Fig. 35). The most striking observation was that the heparin analog *Fondaparinux* induced significant positive shifts of 2.6 °C for Pwp1A. Pwp3A and Pwp5A^{N131D} were less affected in thermal stability by *Fondaparinux*, whereas other PwpA domains and the control Epa1A were almost not affected.

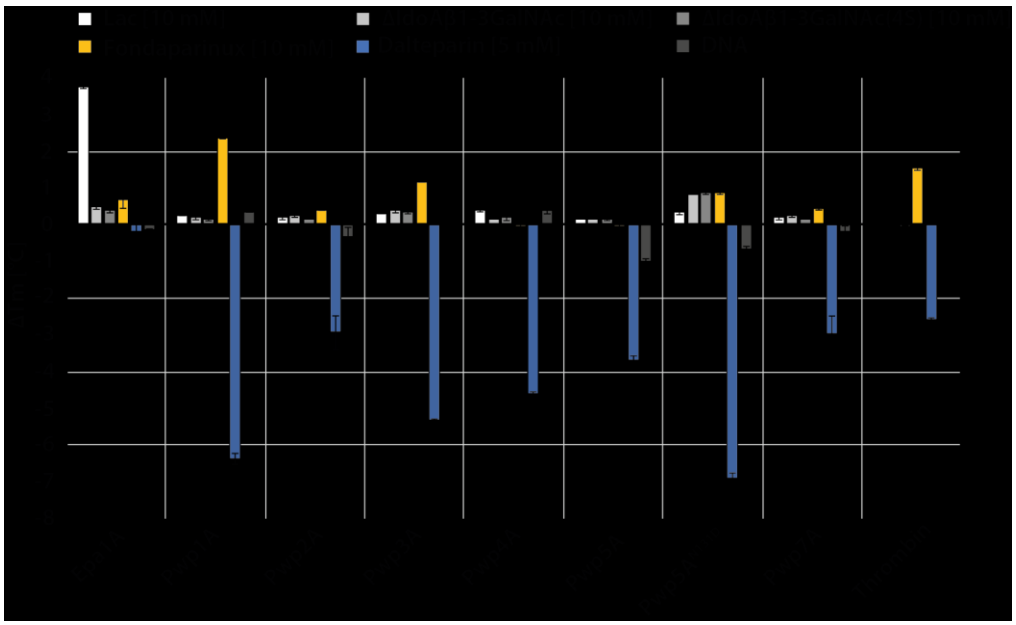

Fig. 35 Thermal shift analysis of PwpA domains with GAG-like compounds. Epa1A in combination with lactose served as a positive control for glycan interaction. Thrombin served as a positive control for heparin/heparan sulfate interaction with *Fondaparinux* and *Dalteparin*. DNA served as a polyanionic negative control. ΔIdoAβ1-3GalNAc and ΔIdoAβ1-3GalNAc represent the main subunits of dermatan sulfate. Thermal shifts were determined at a protein concentration of 10 µM. Melting temperatures were converted to shifts in relation to an H_2O control. Reactions were performed in 20 mM MES pH 6.0, 10 mM $CaCl_2$ and 50 mM NaCl. Shifts ≥ 1.5 °C are considered as significant.

Epa1A in combination with lactose served as positive glycan binding control, which induced a strong positive thermal shift of 3.7 °C. Thrombin served as a positive control since it binds to heparin with micromolar affinity [173]. The heparin pentasaccharide *Fondaparinux* induced a positive shift of 1.5 °C for thrombin, whereas low molecular weight heparin (~ 5kDa) induced a decrease of -2.6 °C. In line with this observed destabilizing effect of polyanionic heparin on thrombin, PwpA domains were also destabilized, indicating for interaction. The synthetic heparin pentasaccharide *Fondaparinux* contains the monosaccharide subunits GlcNS(6S) and GlcA (Fig. 36). GlcNAc(6S) and GlcA also led to positive thermal shifts for Pwp1A (Fig. 32). However, these two monosaccharides are also found in low molecular weight heparin, which led to destabilization. Complexation of a protein-bound calcium ion by the negatively charged sulfate groups in low molecular weight heparin might explain such a decrease in thermal stability. A high negative charge density, such as it is found in heparin might also lead to protein aggregation. The observed negative thermal shifts induced by low molecular weight heparin might also reflect protein aggregation along the polysaccharide chain upon binding. However, polyanionic DNA did not induce significant thermal shifts, such as unsaturated dermatan sulfate disaccharides. Dermatan sulfate contains neither GlcA nor sulfated or non-sulfated GlcNAc.

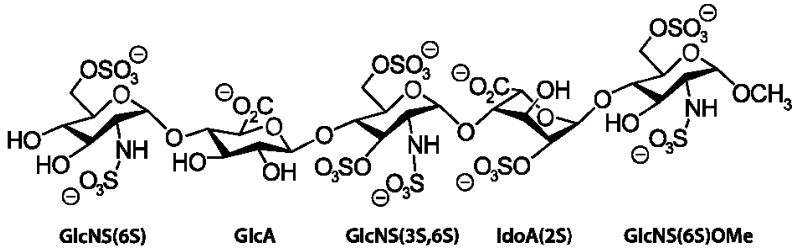

Fig. 36 Chemical structure of the synthetic heparin analog *Fondaparinux*.

The observed binding of PwpA domains to heparin-like compounds was further analyzed in a heparin-column binding assay (Fig. 37) (method described in chapter 5.4.8). BSA was used as a negative control since it does not bind to heparin. I addition, a thrombin sample was used as a positive control, since it is known to bind to heparin with high affinity (6.5 µM) [174], however in an unspecific manner. Furthermore, an Epa1A control was used in order to compare the binding properties of the Pwp family with a representative of the Epa family. In order to analyze if potential interactions with the column material are calcium-dependent, Ca^{2+}-ions were removed by purging the column with 25 mM EDTA before bound protein was eluted with 1 M NaCl.

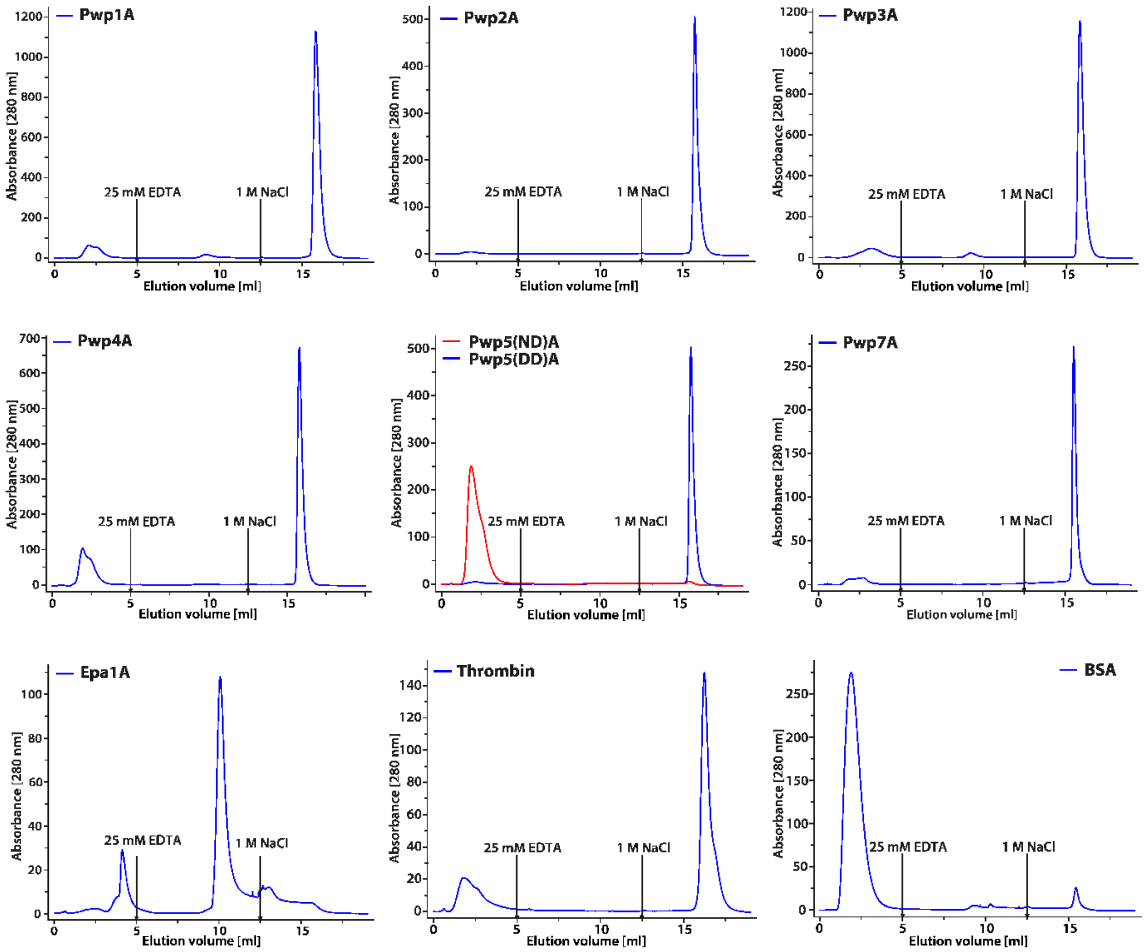

Fig. 37 Elution of PwpA domains from a 1 ml-heparin column. After a flow-through of 5 ml SEC buffer, the column was flushed with 7.5 ml of 25 mM EDTA in SEC buffer, followed by 6.5 ml of 1M NaCl in SEC buffer.

All PwpA domains, except Pwp5A showed binding to the heparin column, comparable to the thrombin control. In contrast, the Pwp5A^{N131D} variant did also to the heparin column, indicating that the aspartate substitution on CBL1 might enable a heparin-interaction and that the natural N^{131} leads to loss of interaction. This finding suggests, that the heparin binding capability of PwpA domains depends on the integrity of the calcium binding capability. In this context, it is also worth to mention, that a high discrepancy in salt tolerance was observed between Pwp5A and the mutant Pwp5A^{N131D} (3.1.3).

While the negative control BSA did not bind to the column, Epa1A showed strong retention but elutes with 25 mM EDTA, which can be explained by Ca^{2+}-mediated interactions with glycan components of the agarose-matrix. In line with this finding, Epa1A retention on agarose-matrix material has also been described by MAESTRE-REYNA, which could be prevented by the addition of EDTA [175]. It is remarkable, that the removal of calcium by 25 mM EDTA did however not result in the elution of column-

bound PwpA domains, indicating that the interaction is not Ca^{2+}-mediated. This observation is contrary to the suggested C-type lectin function for the Pwp family.

In order to further compare heparin affinities of different PwpA domains, the elution from the heparin-column was monitored during linearly increasing NaCl gradients (Fig. 38, Tab. 10). The different PwpA domains revealed different elution behaviors. While Pwp1A shows the highest affinity and elutes in a monodisperse fraction at about 0.9 M NaCl, other PwpA domains elute in two distinction fractions between 0.5 M and 0.95 M NaCl. Pwp7A elutes already at a concentration of about 0.3 M NaCl, whereas binding of Pwp5A was absent. The positive control thrombin, which is known to bind to heparin, elutes in three fractions between 0.3 M and 0.6 M NaCl, although it has been purified in a single fraction by size exclusion chromatography before.

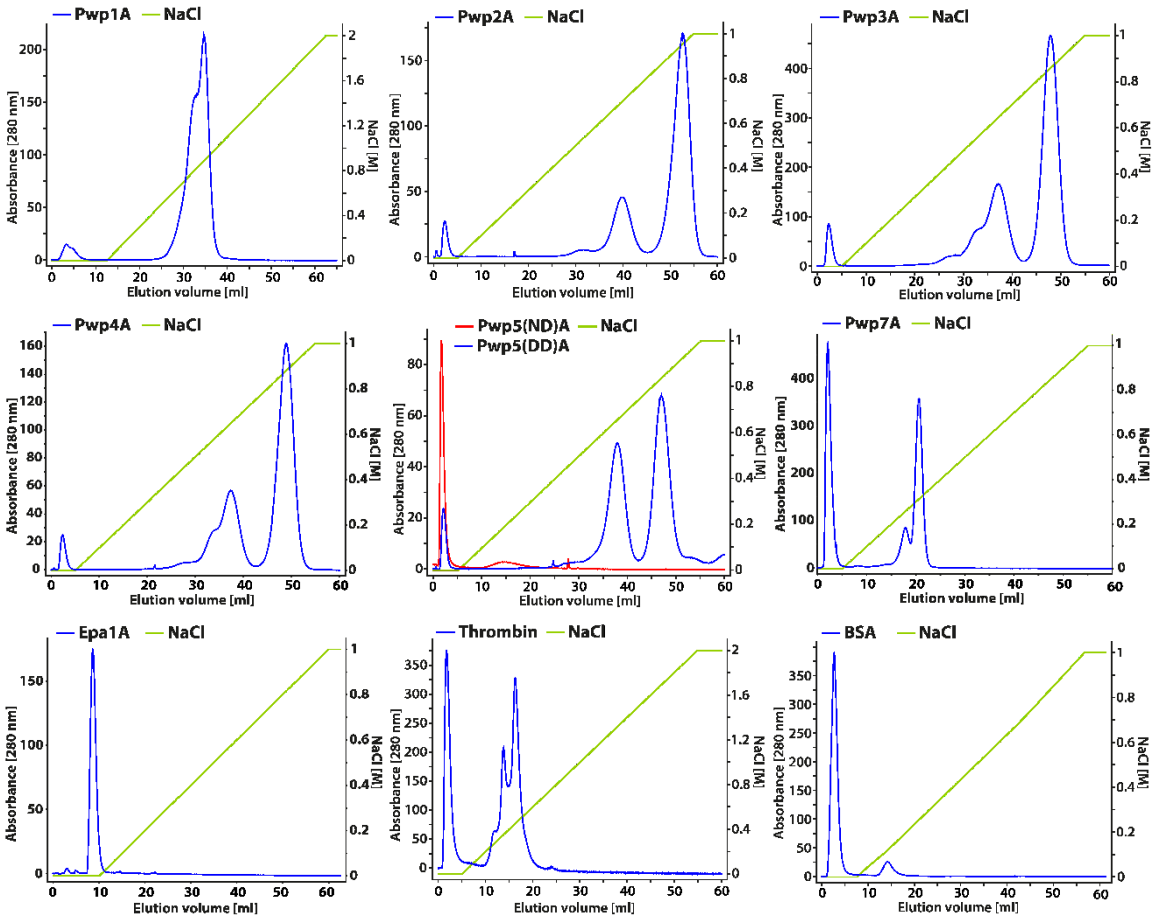

Fig. 38 Elution of PwpA domains from a heparin column: Elution of column bound protein by a linearly increasing NaCl gradient (0.05 M – 1 or 2 M NaCl by 50 min).

Tab. 10 NaCl [M] concentration at the elution of PwpA domains from a 1 ml-heparin column.

Protein	Fraction 1	Fraction 2	Protein	Fraction 1	Fraction 2
Pwp1A	0.9 M	-	Pwp5A	0.65 M	0.85 M
Pwp2A	0.65 M	0.95M	Pwp5A^{N131D}	-	-
Pwp3A	0.65 M	0.85 M	Pwp7A	0.25 M	0.35 M
Pwp4A	0.65 M	0.9 M	Thrombin	0.4 M	0.5 M

In order to validate the observed Pwp-heparin interactions and to exclude potential PwpA-agarose interactions, the PwpA domains were loaded on a 1 ml GST-agarose column and submitted to a linearly increasing NaCl gradient to up to 1M NaCl. In this assay, Pwp-GST-agarose binding was not observed (supplements Fig. 71).

In order to further determine the affinity of PwpA domains to heparin, Pwp1A, Pwp5A and the Pwp5A^{N131D} variant were analyzed in ITC measurements by the titration of the synthetic heparin analog *Fondaparinux*, since natural heparin is subjected to variations in molecular weight. In addition, the Pwp samples were compared to the controls BSA and Epa1A

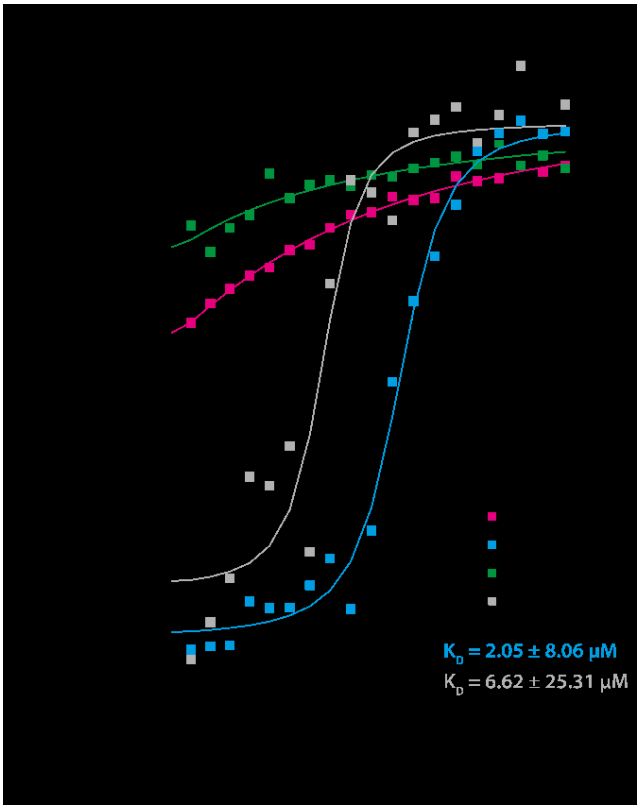

Fig. 39 ITC measurements with *Fondaparinux*. Pwp1A and Pwp5A^{N131D} bind to *Fondaparinux* with micromolar affinity. Shown are titration curves of the fit to a single-class binding-site model using *MicroCal Origin* software after integration of injection peaks with correction for heats of dilution. (Injected: 2 µl of 2 mM *Fondaparinux* into 200 µl of 200 µM Protein at 20° C).

The ITC measurements revealed a high affinity of Pwp1A and Pwp5A^{N131D} to the heparin analog *Fondaparinux* in the lower micromolar range. Such affinities are comparable to the specific interaction of Epa1A with terminal galactose moieties [70,108] (Fig. 39). Interestingly, the natural Pwp5A domain did not show such affinity to *Fondaparinux*, similar to the negative controls Epa1A and BSA, indicating again, that the D^{131} substitution in Pwp5A is necessary for interaction and that binding depends on Ca^{2+}-integrity. The interaction of Pwp1A and Pwp5A^{N131D} with *Fondaparinux* appears to be an exothermic driven reaction (Tab. 11).

Tab. 11 ITC thermodynamic parameters of the binding of PwpA domains to Fondaparinux.

Parameter	Pwp1A	Pwp5A^{N131D}
KD [µM]	2.05 ± 8.06	6.62 ± 25.03
ΔH [kJ mol-1]	-12.24 ± 0.08	-12.88 ± 0.81
-TΔS [kJ mol-1]	-15.83 ± 0.45	-23.85 ± 7.85
ΔG [kJ mol-1]	-7.59 ± 0.12	-5.89 ± 1.49
N [Stoichiometry]	0.97 ± 0.01	056 ± 0.02

The observed binding of heparin and *Fondaparinux* led to the question if certain monosaccharides in the pentasaccharide chain are responsible for the observed interaction. Therefore, ITC measurements of Pwp1A with the monosaccharides GlcA, IdoA and GlcNAc(6S)) were performed. However, the measurements did not result in measurable affinity, although GlcA and GlcNAc(6S) led to thermal shifts for Pwp1A, just as GalNAc and GalNAc(4S)) (Fig. 40). This might be explained by (*i*) the need of 2-*O*-sulfation of iduronic acid, (*ii*) the need of *N*-sulfation of the glucosamine (possibly in combination with 6-*O*-sulfation and/or 3-*O*-sulfation) or (*iii*) a combination of such modifications. Another possibility could be that Pwp1A rather binds to di- or oligosaccharide subunits of heparin/heparan sulfate than to monosaccharide subunits.

However, the affinity to monosaccharide parts might also be too low to be detected by the ITC method. In order to check for an impact of calcium on the affinity of Pwp1A to *Fondaparinux*, the measurement was repeated in absence of calcium (with 5 mM EDTA), which however did not reduce affinity (Fig. 40). The result indicates that the interaction with *Fondaparinux* is not calcium-mediated or that the contribution is only minor.

Fig. 40 ITC profiles for the binding of Pwp1A to (A) GlcA, (B) IdoA, (C) GlcNAc6S, (D) GalNAc, (E) GalNAc4S and (F) *Fondaparinux-EDTA*. The upper panels represent the released heat upon injections of 2 µl 2 mM ligand from the syringe into 200 µl of 200 µM Pwp1A in SEC buffer at 20°C. The reaction with *Fondaparinux* was performed in SEC buffer with 5 mM EDTA. Heats of reactions were determined by the integration of injection peaks with correction for heats of dilution. The lower panels represent the resulting titration curves of the fit to a single-class binding-site model using *MicroCal Origin* software.

3.1.7. PwpA-mediated adhesion is absent *in vivo*

In order to examine the adhesive properties of the Pwp family *in vivo*, *S. cerevisiae* was used as a heterologous expression system for PwpA domains. The used strain BY4741 has an S288c background and is, therefore, non-flocculent, which reduces cell-cell interaction bias and allows single cell sticking via particularly expressed adhesion domains (5.6). The cells were transformed with plasmids (Fig. 41), which were specifically constructed to enable a comparable and strong expression for the presentation of different A-domains on the yeast cell surface and for the interaction with the environment (5.6.1). The presence of corresponding A-domains on the yeast cell surface was verified by immunofluorescence microscopy (5.6.6) using an N-terminal 3× hemagglutinin epitope (Fig. 41) with the result, that the A-domains were found evenly distributed on the cell surface (Fig. 42).

Fig. 41 Expression plasmid for *in vivo* experiments with Pwp adhesion domains in *S. cerevisiae*. The expression is driven by a PGK1 promotor, followed by a signal sequence (SS), a spacer sequence (SP), a 3× hemagglutinin tag (3HA), the adhesion domain (A-domain), a Flo11B/C domain and the terminator (T).

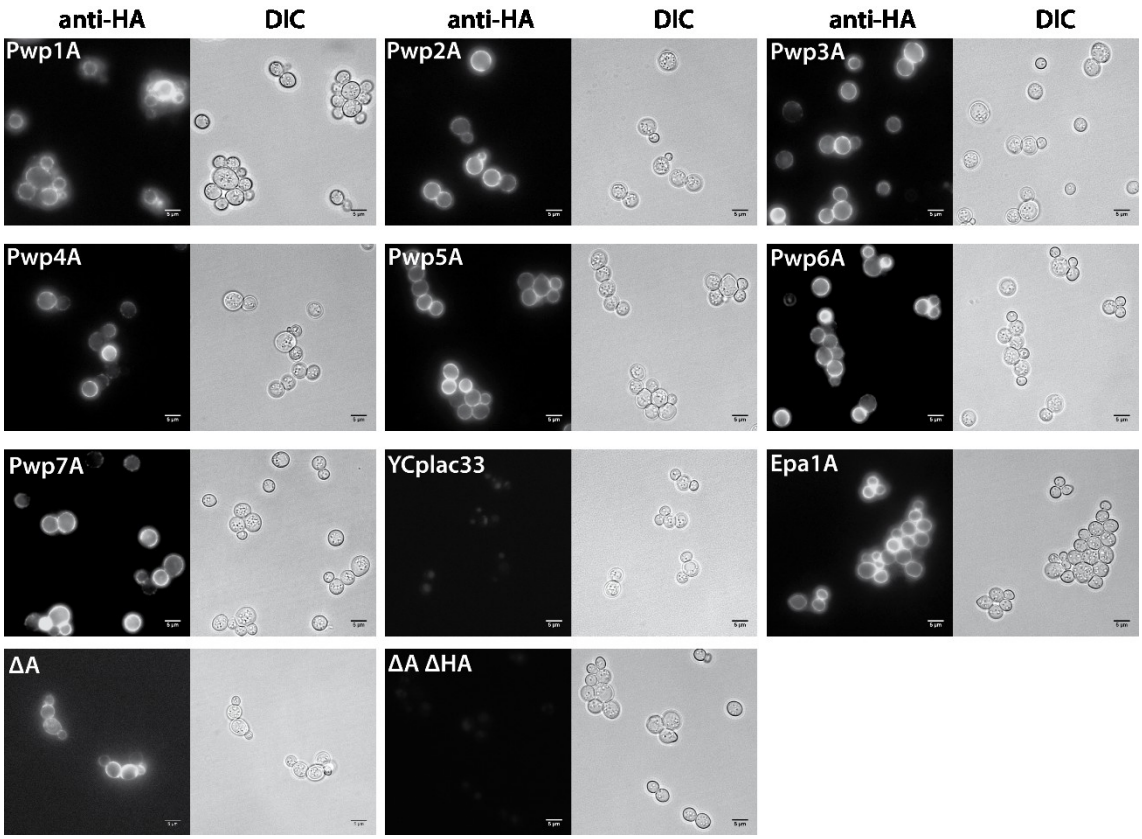

Fig. 42 Detection of PwpA domains on the cell surface of *S. cerevisiae* S288c by immunofluorescence microscopy. Fluorescence microscopy (left) with a Cy3-linked HA-antibody and differential interference contrast microscopy (right) of PwpA-domains and Epa1A as control. YCplac33 represents the empty plasmid backbone. ΔA represents the expression plasmid without any A-domain and ΔA ΔHA carries neither an A-domain nor a 3×HA-tag.

The PwpA-domain presenting yeast cells were next used to perform an agar adhesion-assay as described in chapter 5.6.4. However, these cells did not show significant agar adhesion since they washed off rapidly from the agar surface, similar to the negative control plasmids YCplac33 (empty plasmid backbone) and ΔA (no adhesion domain) (Fig. 43). In contrast, the positive controls Flo11A and Epa1A induced strong adhesion to agar. There is recognizable floc formation in cell suspensions of the controls Flo11A and Epa1A but not in PwpA samples, which appear similar to the negative controls.

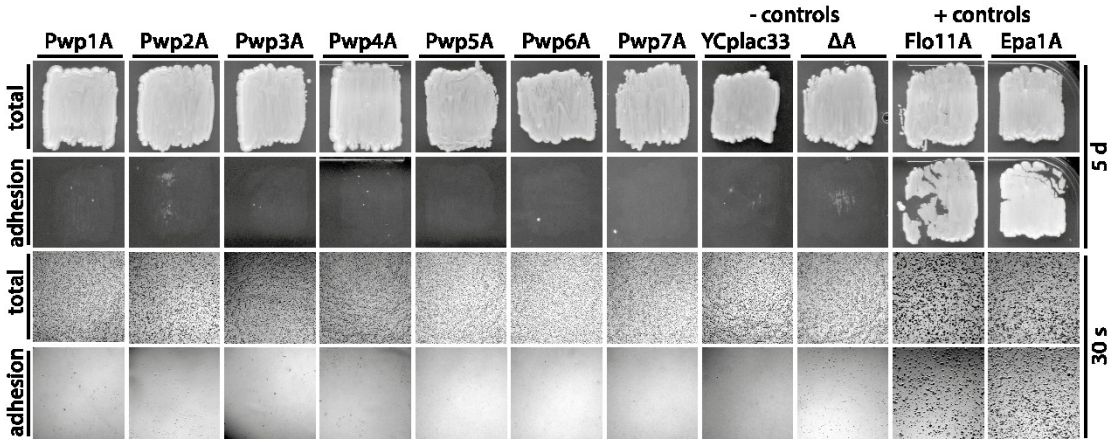

Fig. 43 Adhesion to agar by PwpA domains *in vivo*. Shown are samples before (total) and after (adhesion) wash. Adhesion domains are fused to a B/C-domain of Flo11A from the *S. cerevisiae* strain Σ1278b and possess a 3×hemagglutinin tag for fluorescent detection on the cell surface of the non-flocculent strain *S. cerevisiae* S288c. Biofilms were incubated on solid SC-Ura agar for five days. Single drops of cell suspension were washed off after 30 s of incubation on solid SC-Ura agar.

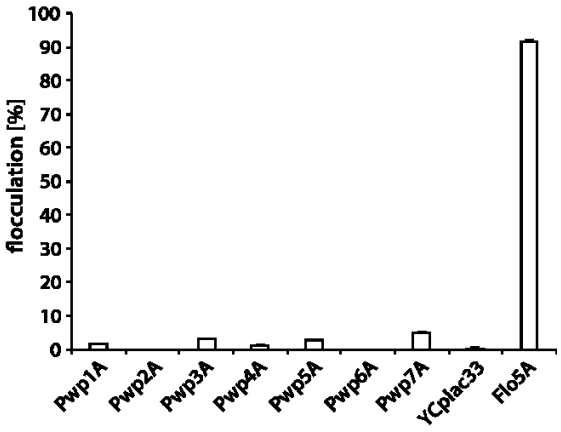

Fig. 44 Relative Ca^{2+}-dependent flocculation by PwpA domains *in vivo*. Flocculation is absent upon expression of PwpA domains as it is on a comparable empty expression plasmid backbone of YCplac33. Expression of the control Flo5A induces strong flocculation.

In a further adhesion test, Ca^{2+}-dependent flocculation was measured with the PwpA-domain presenting cells. The flocculin Flo5A from *S. cerevisiae* was used as a positive control, which induced strong Ca^{2+}-dependent flocculation (Fig. 44). Flo5A is a PA14-like C-type lectin that facilitates cell-cell adhesion or flocculation by the binding of mannose moieties in the yeast cell wall [93]. However, flocculation induced by homotypic or heterotypic PwpA interactions has not been observed.

Furthermore, adhesion to human epithelial cell monolayers was analyzed *using* the *S. cerevisiae* expression strain as described in chapter 5.6.2. However, significant adhesion to different types of human epithelium was not observed for PwpA domains, whereas Epa1A induced strong epithelial adhesion (Fig. 45), as it has been described before [70,109].

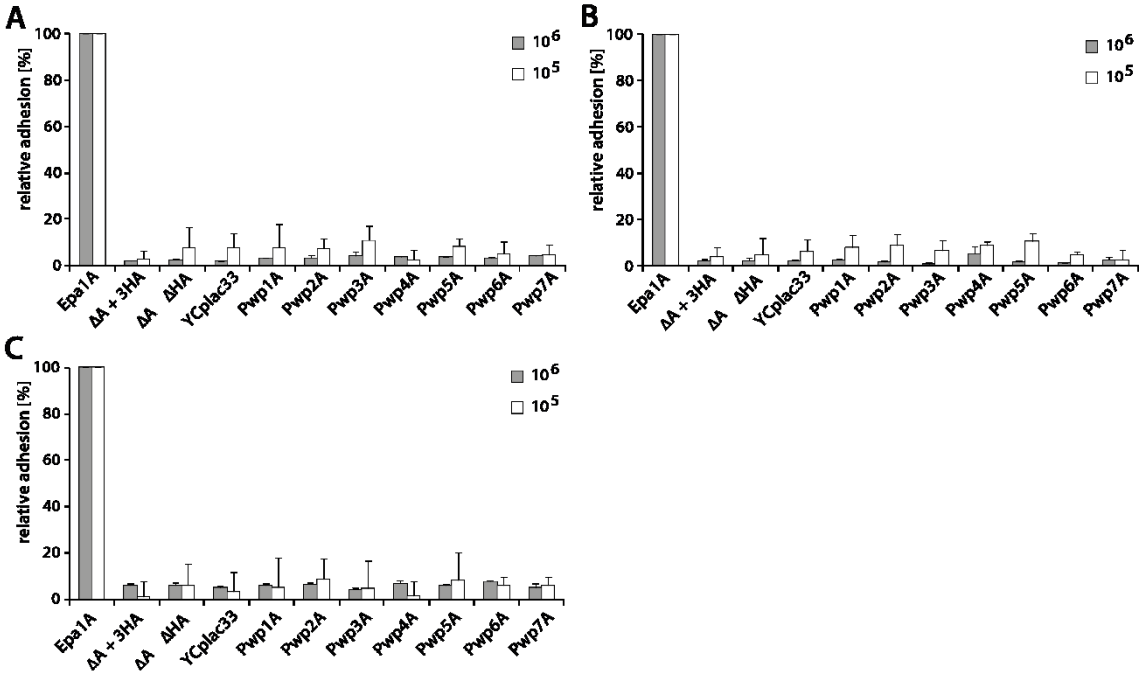

Fig. 45 Adhesion by PwpA domains to different human epithelia *in vivo* and in relation to Epa1A. Adhesion was absent for the Pwp family. **A**: Colon epithelium (Caco2). **B**: Oral epithelium (TR146). **C**: Vaginal epithelium (A431). Either 10^6 or 10^5 yeast cells, were added per well of monolayer epithelium.

Finally, an adhesion assay with human extracellular matrix proteins was performed as described in chapter 5.6.4. PwpA domains induced adhesion to human type IV collagen in a dimension of 50-70% of the positive control Als3 from *C. albicans*. Als3 is an adhesin that has been reported to bind collagen type IV *in vivo* [176,177]. Interestingly collagen, along with other ECM proteins, is found as part of a variety of connective tissues, where it builds a net-like structure with glycosaminoglycans, such as heparan sulfate. Collagen type IV has several sites for glycosylation [178] but the specific structures of linked glycans are yet unknown.

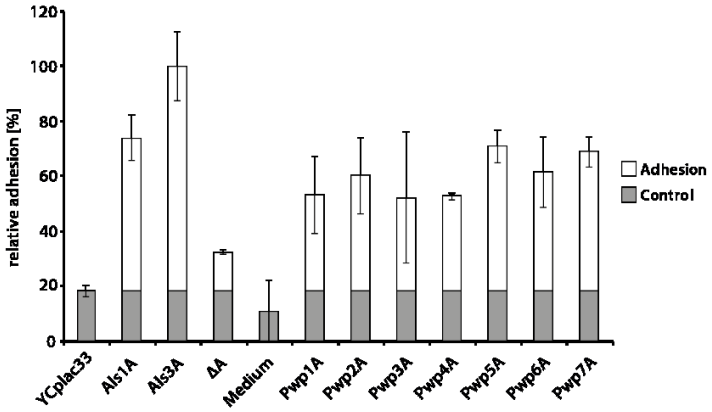

Fig. 46 Adhesion of PwpA domains to type IV collagen *in vivo*. Adhesion is shown in relation to the positive control Als3 from *C. albicans*. 10^6 cells were applied per well of a coated 96-well plate coated with human collagen type IV.

3.2. Structural and functional characterization of Wsc domains

As the uppermost components of the CWI signaling cascade in *S. cerevisiae*, the cell wall integrity sensors Wsc1, Wsc2 and Wsc3 represent potential targets for antifungal drugs. Their cysteine-rich domains appear to be necessary for the function as sensors for mechanical stress. Since these domains required a structural and functional characterization, they were here used for crystallographic studies, as well as for *in vitro* analytics. Parts of the following results were obtained by the guidance of a Master thesis (Philipp Schöppner, 2017)

3.2.1. *In silico* analysis

Structural sequence alignment of the CRDs from *Sc*Wsc1-3 reveals eight highly conserved cysteines (Fig. 47). Moreover, a relatively high amount of conserved or similar aromatic residues, in particular tyrosine and phenylalanine, indicates highly hydrophobic regions in the CRDs of Wsc1-3.

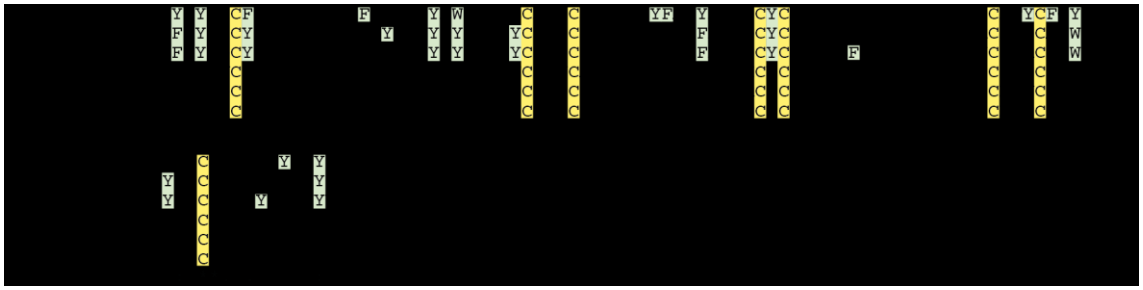

Fig. 47 Structural alignment of the CRDs from *Sc*Wsc1-3 in comparison to Wsc of human kremen1 and an uncharacterized Wsc domain from *C. glabrata*. Eight conserved cysteines (yellow) build the CRD region (cysteine-rich domain). A high amount of aromatic residues (light green) can be found within the CRD region of *Sc*Wsc1-3. The alignment was performed with the *TCoffee Expresso* server.

Homologous CRDs of human Kremen1 and the *C. glabrata*-protein CAGL0F01507g have a relatively low level of sequence identity to the Wsc family from *S. cerevisiae* (Fig. 48). The CRDs of Wsc2 and Wsc3 show the highest degree of conservation with a sequence identity of 63%. The CRD of Wsc1, in contrast, is more distantly related to Wsc2 and Wsc3 with an identity of 35% and 36 %, whereas shares 63% sequence identity with the uncharacterized CRD of *C. glabrata* but merely 27% to the CRD of human Kremen1.

Domain borders for heterologous expression were determined as described in chapter 5.1.1 and the final domain organization is shown in Fig. 48

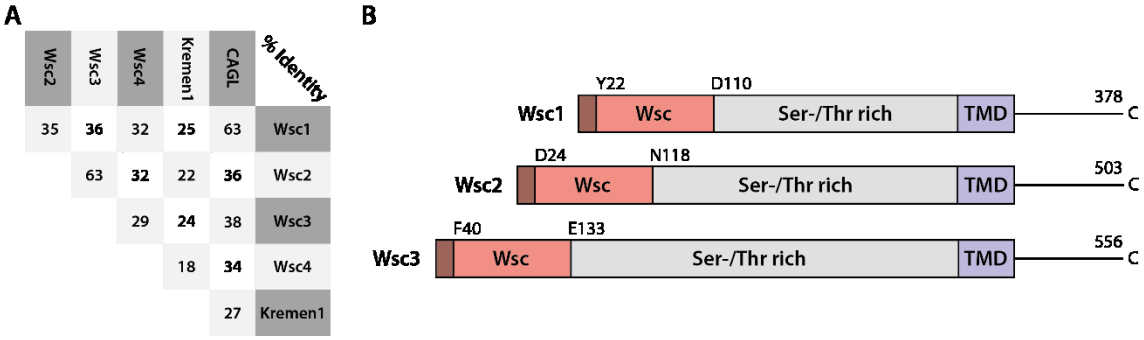

Fig. 48 A: **Sequence identity of Wsc domains.** CRDs of *Sc*Wsc1-3 in comparison to the Wsc domain of human Kremen1 and an uncharacterized Wsc domain of *C. glabrata*. **B**: **Domain organization of *Sc*Wsc1-3**. An *N*-terminal secretion signal sequence (brown) is followed by the cysteine-rich Wsc domain (red). That domain is followed by a Ser-/Thr- rich region (grey), which is variable in length and is thought to be highly glycosylated. A transmembrane domain (blue; TMD) anchors the protein in the cell membrane. An unstructured, C-terminal region is supposed to be responsible for intracellular signal transduction.

Based on the results of the *in silico* analysis, expression plasmids were produced for recombinant expression in *E. coli SHuffle*® T7 Express and for the fusion with an N-terminal 6×-histidine tag. The predicted physicochemical properties of the *His*-tagged CRDs are shown in Tab. 12 and putative glycosylation sites are listed in Tab. 13.

Tab. 12 Theoretical physicochemical properties of *His*-tagged *Sc*Wsc domains used in this study; values were determined with *ProtParam*.

CRD	A-domain	Mw [Da]	pI	Ext. coefficient	Gravy index
Wsc1	Y22-I118	12877.6	5.2	20900	-0.75
Wsc2	D24-N118	12524.88	7.01	19410	-0.34
Wsc3	F40-E133	12428.68	6.34	19410	-0.36

Tab. 13 Theoretical glycosylation sites of *Sc*Wsc domains used in this study. Glycosylation sites were predicted with *NetNGlyc 1.0* and *NetOGlyc 4.0*.

CRD	*N*-glycosylation sites	*O*-glycosylation sites
Wsc1	N65, N75	S81, S83, S85
Wsc2	N70	S85
Wsc3	-	-

The recombinant CRDs of Wsc1-3 from *S. cerevisiae* S288c share a molecular weight between 12-13 kDa and are predicted to be highly soluble, referring to the negative gravy indices based on the algorithm of KYTE and DOOLITTLE [165]. However, the three CRDs of Wsc1-3 show relatively high variations in their pI. The CRDs of Wsc1 and Wsc2 are putatively *N*- and *O*- glycosylated, whereas Wsc3 does not have potential glycosylation sites.

3.2.2. Heterologous expression and purification of Wsc domains

The highest solubility for heterologous production was achieved with the *E. coli* strain *Shuffle T7 Express* at 12°C for 72 h in LB-medium, induced with 0.1 mM IPTG (Tab. 14).

Tab. 14 Yield of His-tagged Wsc CRDs by heterologous production in *E. coli*.

	Wsc1	Wsc2	Wsc3
Yield [mg/ml]	1.75	0.375	0.625

It is striking that during size exclusion chromatography, the *Sc*Wsc CRDs elute in distinctly separated fractions rather than in monodisperse fractions, which is probably caused by dimer- or oligomerization via the *His*-tags, since removal of the *His*-tags by thrombin cleavage results in a monodisperse fraction at the right size (Fig. 49). However, the solubility of the Wsc1 CRD was strongly reduced after *His*-tag removal. The recombinant *Sc*Wsc CRDs also show an unusual separation behavior when subjected to SDS-PAGE, since they appear as 20 kDa proteins rather than their corresponding molecular weight of ~13 kDa (Fig. 49). This property might be caused by incomplete denaturation due to high structural rigidity.

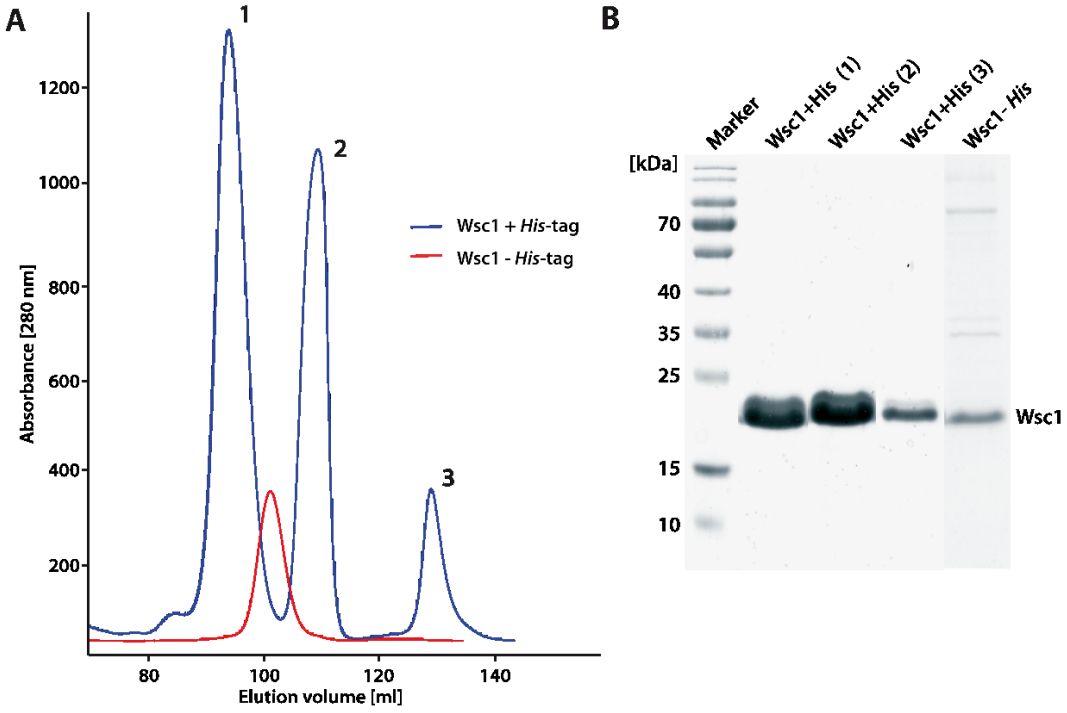

Fig. 49 A: Size-exclusion chromatogram of the *Sc*Wsc1 CRD. The Wsc1 CRD elutes in distinctly separated fractions (blue) and in a monodisperse peak after *His*-tag cleavage. **B: SDS PAGE of *Sc*Wsc1 CRD peaks upon size-exclusion chromatography**. The Wsc CRDs of separated peaks appear as ~20 kDa protein rather than their actual molecular weight of ~13 kDa (+*His*) or ~10 kDa (-*His*) respectively. Other *Sc*Wsc CRDs show similar properties.

The expected high structural stability of the *Sc*Wsc CRDs was confirmed by the fact, that the proteins did not precipitate in thermal shift analysis when subjected to 95°C and pH 3-10. The addition of 20 mM DTT, however, resulted in a strong reduction of the melting temperature, which speaks for the disulfide-mediated exceptionally high stability of the *Sc*Wsc CRDs (supplements Fig. 76).

3.2.3. *In vitro* analysis of Wsc protein-protein interactions

It has been shown, that Wsc1 builds compartments in the membrane, which partly overlap with Wsc2 and that the CRDs are necessary for CWI sensor clustering, which enables stress sensing. Since CRD clustering is supposed to be mediated through protein-protein interactions, the formation of CRD-dimers or -multimers was analyzed by microscale thermophoresis (Fig. 50). The measurement compared interactions in the presence and absence of the hexa-histidine tag (*His*-tag) and under changing pH conditions. Homotypic protein-protein interactions were observed in the presence of a *His*-tag with a K_D of 1.3 µM. After cleaving off the His-tag homotypic interactions were absent under the same conditions as well as in a slightly acidic environment (pH 5.5), indicating that the observed interaction is *His*-tag mediated.

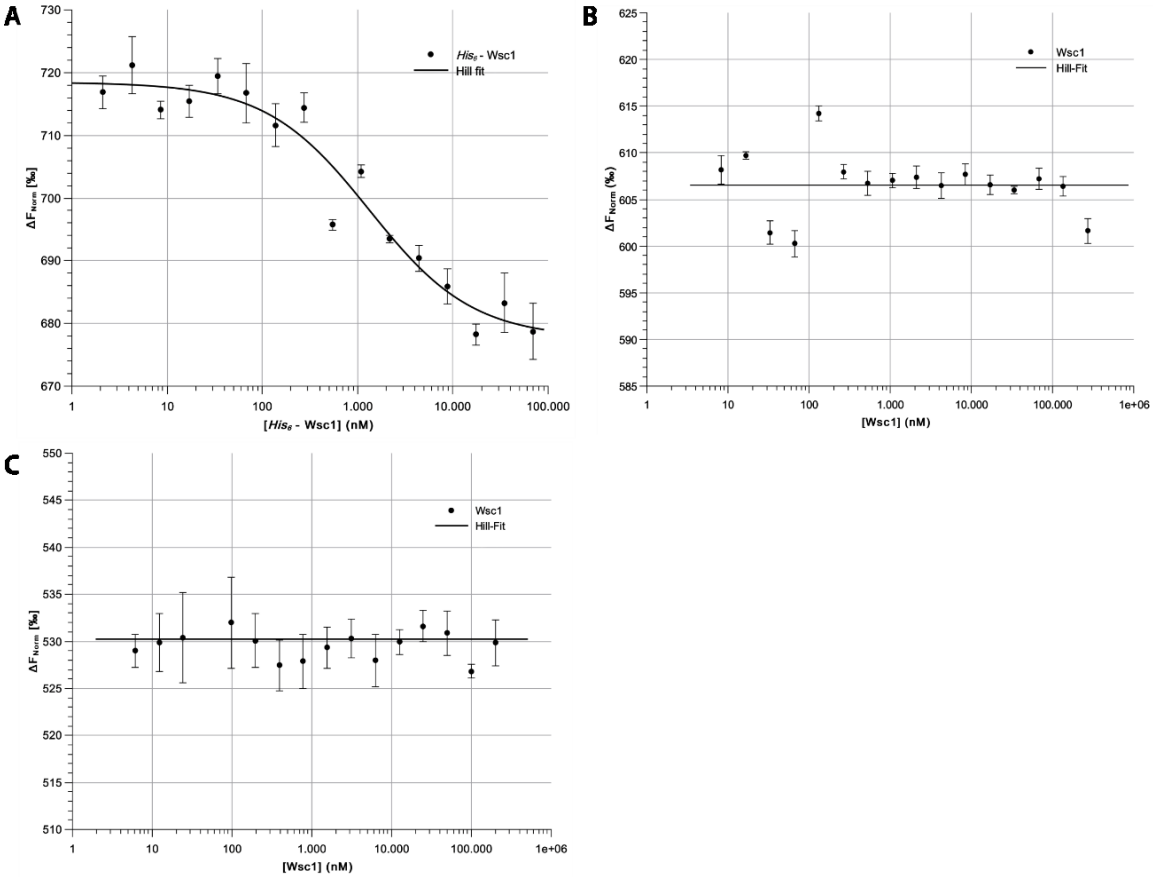

Fig. 50 Homotypic protein-protein interaction analysis of fluorescently labeled *Sc*Wsc1 CRDs. A:6×*His*-tagged Wsc1 at pH 7.5. **B:** Wsc1 without tag at pH 7.5. **C:** Wsc1 without tag at pH 5.5. Homotypic interaction (K_D = 1.3 µM) is observed only in presence of a 6×*His*-tag.

3.2.4. Glycan binding analysis of Wsc domains

In order to check for possible Wsc-glycan interactions, the fluorescently labeled CRDs of *Sc*Wsc1-3 were analyzed on a microarray with 140 printed pathogen glycan ligands (Tab. 50) as described in (5.4.6). In this analysis, glycan binding was not observed, whereas Epa1A showed binding to terminal galactose residues.

3.2.5. Structural analysis of Wsc domains

In order to solve the 3D-structure of the Wsc CRDs from *S. cerevisiae* and to gain insight into structural properties of the Wsc family, X-ray diffraction analysis were performed. Since X-ray diffraction analysis requires the production of well-ordered protein crystals, large-scale crystallization conditions were screened (5.5.1). Therefore, the Wsc CRDs were concentrated to a maximum of 70 mg/ml (Wsc1), 15 mg/ml (Wsc2) and 30 mg/ml (Wsc3). The Wsc1 CRD showed by far the highest solubility and resulted in the highest yield, which correlates with the Gravy index in Tab. 12. The Wsc1 CRD successfully crystallized in condition #77 of JCSG Core II at 4 °C and at a protein concentration of 70 mg ml/ml. Crystal growth was first observed after an incubation time of about three weeks. However, the Wsc2 and Wsc3 CRDs resulted in either proto-crystalline material or crystal growth was absent.

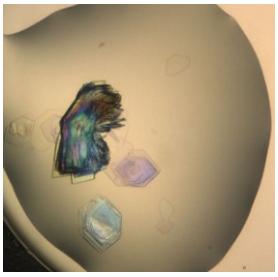

Fig. 51 Crystals of Wsc1. Two-dimensional crystals in JCSG Core II #77 that resulted in sufficient diffraction.

Tab. 15 Crystallization conditions of Wsc1.

Wsc1- JCSG Core II #77	
Sodium acetate	0.08 M
pH	4.6
Ammonium sulfate	1.6M
Glycerol	20% (w/v)

One dataset for the Wsc1 CRD was collected. After data reduction and scaling the crystal structure of the Wsc1 CRD was solved by molecular replacement, based on the crystal structure of the Wsc domain from the human transmembrane sensor Kremen1 (Tab. 16, Tab. 17). Despite a relatively low sequence identity of about 20%, the structure was solved at a resolution of 1.6 Å.

Tab. 16 Data collection statistics for the *Sc*Wsc1 crystal.

	***Sc*Wsc1**
Dataset	Wsc1 1
Date	17.02.2017
X-ray source	ID29, ESRF Grenoble
Detector	Pilatus 6M
Wavelength (Å)	0.979
Space group	$P2_1$
Cell dimensions (Å)	
a	31.87
b	53.72
c	52.07
α, β, γ	90°, 95.596°, 90°
Molecules / a.s.u.	2
Resolution (Å)	28.31 – 1.59 (1.62 – 1.59)
Total reflections	69876
Unique reflections	23057
Multiplicity	3 (2.7)
Completeness (%)	97.4 (94.2)
R_{merge} **(%)**	6 (43)
Mean I/σ(I)	10.5 ()2.3
Wilson B-factor (Å^2)	23.986
Mosaicity (°)	0.2

Tab. 17 Refinement statistics for *Sc*Wsc1.

	***Sc*Wsc1**
Resolution (Å)	51.75 – 1.58
R_{work}	16.66
R_{free}	19.96
Completeness (%)	96.49
r.m.s.d.	
Bond length (Å)	16.66
Bond angle (°)	19.96
Total number of atoms	1602
Mean *B*-value (Å^2)	18.43

Overall, the structure of the Wsc1 CRD is characterized by a central β-sheet, composed of five antiparallel β-strands (β1-5), which build the core of a tightly packed globular domain (Fig. 52).

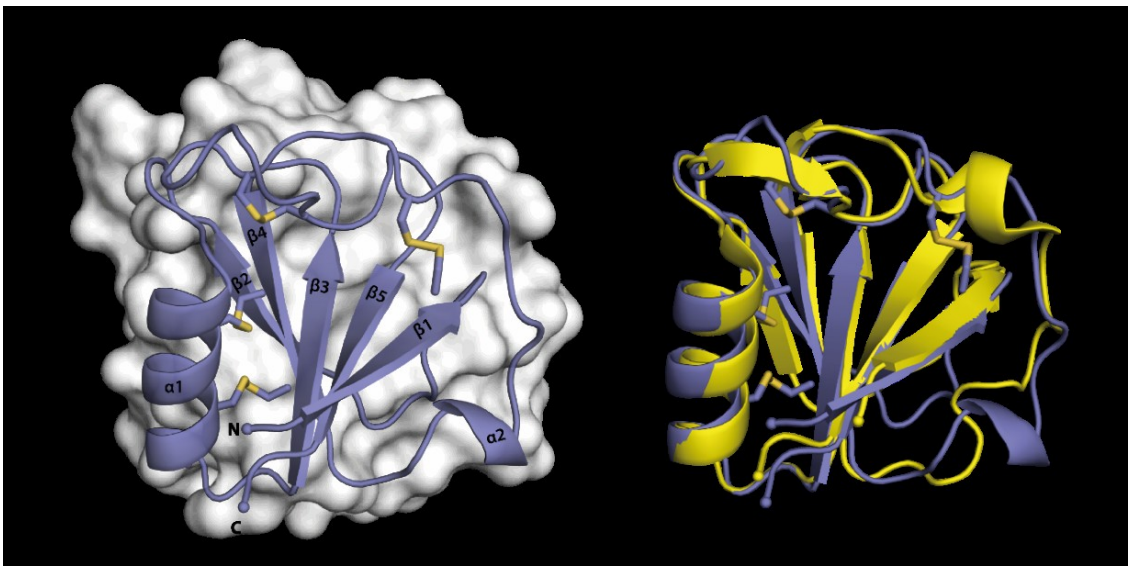

Fig. 52 A: Crystal structure of the CRD from *Sc*Wsc1. A central β-sheet composed of five β-strands builds the core of the globular CRD that is stabilized by the formation of four disulfides. **B: Overlay of the CRD from *Sc*Wsc1 (blue) and the Wsc domain from human Kremen1 (yellow).** Both domains show high similarities in the central core structure organization but differences in the flexible outer areas.

The domain is stabilized by the formation of four disulfide bonds from eight highly conserved cysteines, which allows a highly rigid and dense structure (Fig. 53).

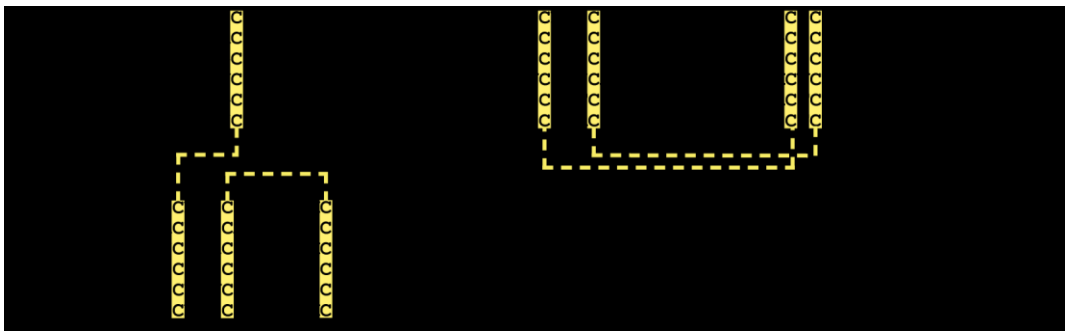

Fig. 53 Formation of disulfide bonds in Wsc-CRDs. Shown is a structural alignment of *Sc*Wsc CRDs with homologous domains (*TCoffe Expresso*). Cysteines are highlighted in yellow and disulfide bonds are indicated by dashed yellow lines.

Moreover, an α-helix (α1) and an α-helical turn (α2) can be found as part of the core-surrounding material, which is mainly composed of unstructured flexible loops. The helix α1 is closely attached to the β-sheet core by two disulfides. The outer flexible and mostly unstructured areas represent the main differences to the Wsc domain from human Kremen1. Moreover, the Wsc1 CRD does not show an apparent binding site or pocket for glycan ligands. Intrinsic metal ions that could be involved in ligand binding are also absent. Interestingly, Wsc1 CRD crystal-contacts in the asymmetric unit are formed by the complexation of a sulfate ion via the hexa-histidine tags of two opposing Wsc1 CRD molecules (Fig. 75). However, an apparent physiological relevant dimerization has not been observed.

An estimation of the electrostatic surface potential of the *Sc*Wsc1 CRD reveals a mainly negative charge on one side of the domain, whereas the opposite side has mainly non-polar characteristics (Fig. 54). The negative surface potential results from eleven surface-exposed residues of aspartic acids and glutamic acids, which points out the highly negative pI of 4.0 (without hexa-histidine tag) of the *Sc*Wsc1 CRD.

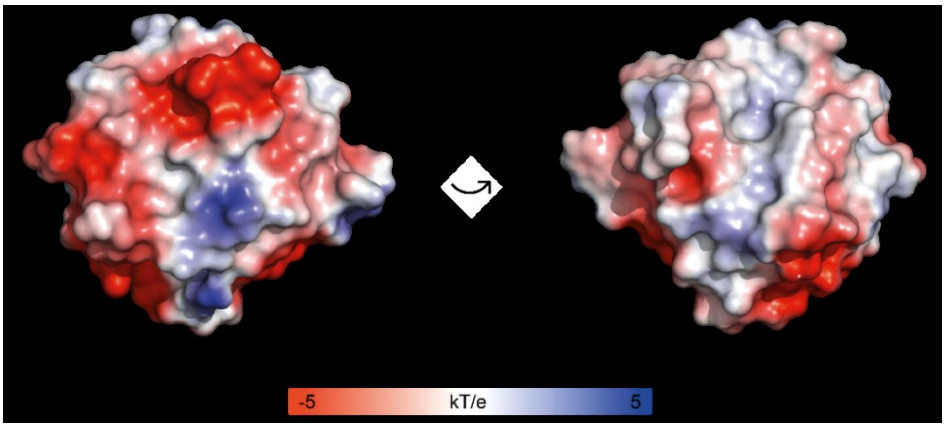

Fig. 54 Electrostatic surface potential of the *Sc*Wsc CRD. Determined with *APBS* at pH 7.0 and 0.15 M NaCl

An estimation of the surface hydrophobicity characteristics does not reveal particularly hydrophobic areas (Fig. 55). However, the Wsc1 CRD has a high amount of 15 aromatic amino acid residues, mainly tyrosine and phenylalanine, of which 12 are surface exposed (Fig. 56).

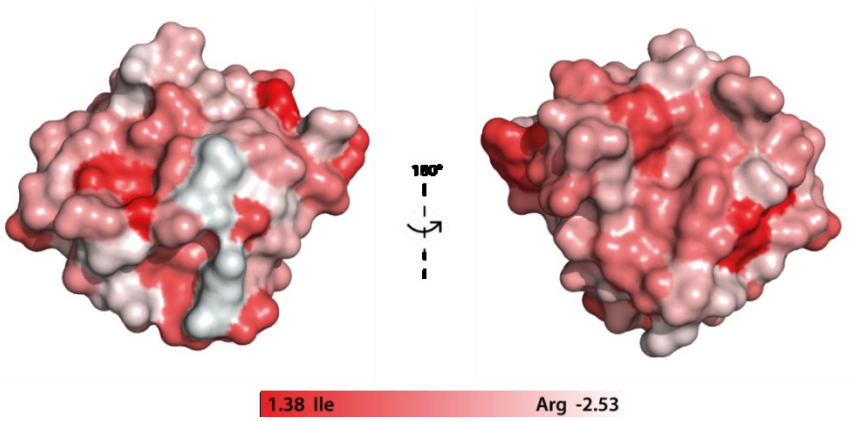

Fig. 55 Surface hydrophobicity of the *Sc*Wsc1 CRD. Shown is the relative hydrophobicity based on the EISENBERG scale [169] with hydrophobic areas in red and hydrophilic areas in white.

These surface-exposed aromatic residues build three distinct clusters (1-3). The largest cluster (cluster 1) comprises the residues of Y41/W43/Y89/F91/Y93, on the apical side of the domain. Two additional clusters comprise the tyrosine residues Y64/Y70/Y104 (cluster 2) and Y22/Y24/Y107 (cluster 3). The conservation and clustering of the aromatic amino acids on the protein surface might indicate a functional significance, such as protein-protein interactions. Therefore clusters of aromatic residues might play a role in sensor-clustering.

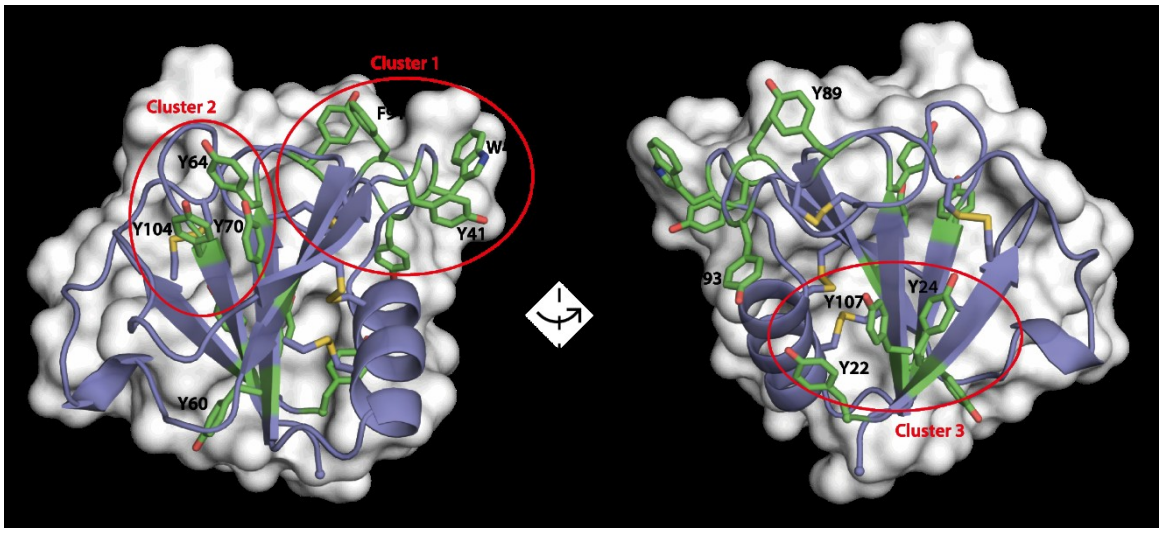

Fig. 56 Aromatic amino acids in the Wsc domain of Wsc1. Aromatic residues are highlighted in green which build three distinct clusters (red).

4. Discussion

Yeasts are ubiquitous fungi that have adapted to different ecosystems and evolved diverse strategies to thrive under changing environmental conditions. The opportunistic pathogen *C. glabrata* colonizes different habitats including the human body. Its natural resistance to classic antimycotics and the ability to cause systemic candidiasis makes it a clinically highly relevant yeast. The non-pathogenic yeast *S. cerevisiae* in contrast primarily occurs on sugar-rich materials. A large variety of cell wall proteins such as adhesins and integrity sensors help yeasts to interact with the environment and to react to changing environmental conditions. This work gives a deeper understanding of two groups of yeast cell wall proteins (*Cg*Pwp and *Sc*Wsc), of which knowledge was hitherto limited.

4.1. The Pwp family of *Candida glabrata*

A large part of fungal adhesins are cell wall-linked C-type lectins, that base on the structurally conserved PA14 domain from *B. anthracis* and confer cell-cell adhesion and cell-substrate adhesion. Despite their structural conservation, different fungal PA14/Flo5-like adhesins bind different ligands and are employed in independent biological processes. Large sets of these adhesins in *C. glabrata* cluster in several subgroups, which is thought to cover a wide range of different substrates and to enable diverse functionalities under changing environmental conditions. The structural basis of the *N*-terminal domains from the *C. glabrata* Pwp subgroup was uncovered by the crystal structures of Pwp1A and Pwp5A, revealing unique structural features. Functional studies indicate that Pwp proteins bind to components of glycosaminoglycan with albeit not yet clear specificity and mode of binding.

4.1.1. Evidence for high structural similarity of PwpA domains

The crystal structures of Pwp1A and Pwp5A were successfully solved, whereas crystallization of other PwpA domains either failed to yield well-diffracting crystals or crystal formation remained absent. Therefore, homology models of the remaining PwpA domains from *C. glabrata* have been created on the basis of the Pwp5A crystal structure. These models indicate that structural differences among the PwpA domains are most likely minor, due to the conserved β-sandwich core and an overall high structural similarity. However, the unusual turned orientation of loop L1 in domain Pwp1A appears to be a unique feature (Fig. 57). The structural model of Pwp7A misses one β-strand in the core backbone that corresponds to β10 in Pwp1A (or β11 in

Pwp5A). As expected, other PwpA domains are most likely Ca^{2+}-coordinating domains with the typical properties of a β-sheet sandwich motif that is overall very similar to other known PA14/Flo5-like fungal lectins such as Epa1A (4ASL), Epa6A (4COU), Cea1A (5A3M), Lg-Flo1A (4GQ7), *Sc*Flo1A (4LHN) or *Sc*Flo5A (2XJP).

Fig. 57 Superimposition of PwpA homology models onto Pwp1/5A crystal structures. Structural models are based on Pwp5A and were generated with *SWISS-MODEL* [179]. Structural differences are most likely minor, apart from the expanded loop L1 of Pwp1A.

However, such structural similarities do not necessarily indicate similar ligand binding properties, since PA14/Flo5-like fungal lectins showed different modes of carbohydrate recognition and may confer either self-recognition (Flo) or host-recognition (Epa) [70,93,97,108,168,180,181].

A comparison of the electrostatic surface potentials of modeled PwpA domains reveals a high density of negative charges, largely surrounding the Ca^{2+}-binding site (Fig. 58), similar to Pwp1A and Pwp5A. The negative charges appear to dominate the overall surface potential, which correlates with the isoelectric points between 4.5 and 6.35. An accumulation of positively charged residues, which might be involved in the interaction with negatively charged sulfate groups in glycosaminoglycan chains, can be observed close to the *N*- and *C*-terminal regions.

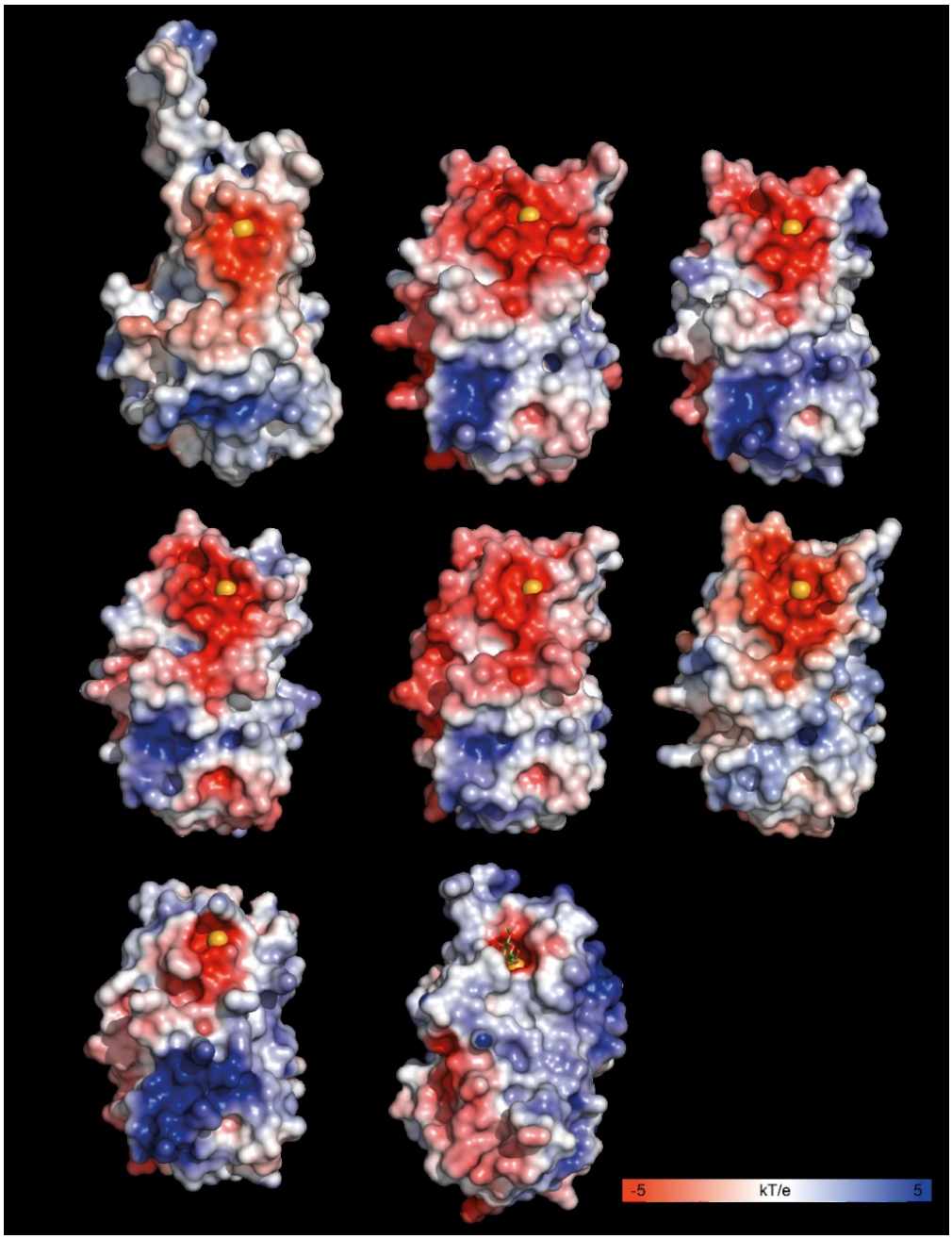

Fig. 58 Electrostatic surface potential of Pwp1A (A), Pwp5A (E), Epa1A (H) and homology models of Pwp2A (B), Pwp3A (C), Pwp4A (D), Pwp6A (F) and Pwp7A (G). The Pwp family is characterized by accumulations of negative charges around the Ca^{2+}-binding-site and positive charges in the *N*- and *C*-terminal region. Surface potentials were determined with *APBS* at pH 6.0 and 0.15 M NaCl.

Comparing the surface hydrophobicity of PwpA homology models to Pwp1A and Pwp5A reveals that the highest density of hydrophobic amino acid residues can be found on the front loops L1 and in particular L2 (Fig. 59). Such hydrophobic regions might indicate homotypic protein-protein interactions or heterotypic interactions, possibly with host proteins. Since glycosaminoglycans are usually linked to a proteoglycan core protein on the cell surface, hydrophobic interactions with the core protein might play a possible role in the effective binding of glycosaminoglycans by PwpA domains. In fact, heparan sulfate-binding adhesins do not show sequence

homology but usually contain domains rich in basic residues flanked by hydrophobic domains [136,182].

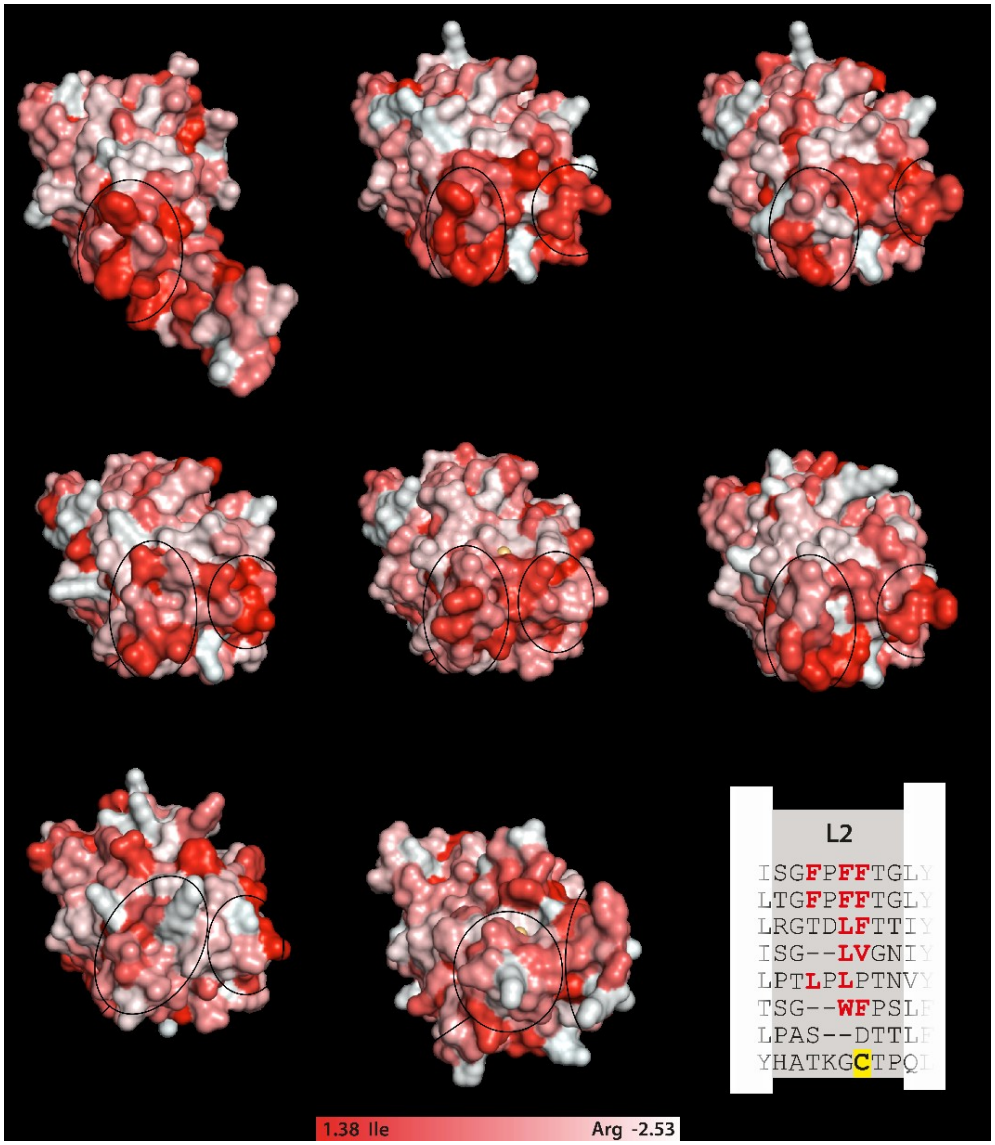

Fig. 59 Surface hydrophobicity of Pwp1A (A), Pwp5A (E), Epa1A (H) and homology models of Pwp2A (B), Pwp3A (C), Pwp4A (D), Pwp6A (F) and Pwp7A (G). A structural alignment of the loop L2 illustrates the most hydrophobic residues highlighted in red. Shown is the relative hydrophobicity, based on the EISENBERG scale [169] with hydrophobic areas in red and hydrophilic areas in white. The front loops L1 and in particular L2 of Pwp1-6A are dominated by hydrophobic residues. Pwp7A and Epa1A do not show such hydrophobic accumulations on the front loops.

Another possibility is dimerization or multimerization of PwpA domains, due to hydrophobic clustering. That might increase avidity by increasing the local adhesin concentration on the cell surface. Such adhesin clusters might increase the probability that a ligand, which dissociates from one adhesin due to low affinity, rapidly binds to another adhesin in the same cluster. As a result, the macroscopic dissociation constant would be expected to increase, which might enable effective cell adhesion. Such a clustering-effect might also enable binding to polysaccharide chains, such as glycosaminoglycans, and explain, why high-affinity binding to glycosaminoglycan-

monosaccharides was not observed in ITC measurements. A similar avidity-increasing effect of yeast adhesins has been postulated to be caused by β-aggregation of the repetitive B-regions [103].

Cell surface-hydrophobicity is another characteristic that is usually related with biofilm formation [183]. Such a correlation has been demonstrated for the related pathogen *Candida parapsilosis* [184,185]. Moreover, *C. glabrata* shows a four-fold greater cell surface hydrophobicity than for instance *C. albicans*. Cell surface hydrophobicity also positively correlates with a higher tendency to adhere to acrylic hydrophobic surfaces [186]. In this context, hydrophobic parts of Pwp proteins might be candidates that increase the global hydrophobicity of *C. glabrata* and thereby influence effective biofilm formation and adhesion to hydrophobic surfaces.

4.1.2. PwpA domains bind heparin-like compounds *in vitro*

In this study, different analytical approaches indicated an interaction of the PwpA domains with compounds of glycosaminoglycan, such as heparin and heparan sulfate or parts of it. An important question is if the observed interactions with glycosaminoglycan reflect a hitherto unknown subset of ligands in the context of a specific biological function for *C. glabrata*. Moreover, it needs to be clarified, if such interactions are of specific or of non-specific ionic nature, regarding the high negative charge density of glycosaminoglycans.

The glycosaminoglycans heparin and heparan sulfate are linear, anionic polysaccharides composed of disaccharide subunits, which consist of an amino sugar together with a uronic acid under varying degrees of sulfation [129] (1.4). Heparan sulfates are ubiquitous components of the cell surface and the extracellular matrix of multicellular animals, whereas heparin is a structurally similar but more sulfated compound produced by mast cells. In heparin, uronic acids are predominantly IdoA, whereas in heparan sulfate they are mainly GlcA, both subsequently *O*-sulfated. Moreover, in heparan sulfate, the glucosamine residues are predominantly *N*-acetylated, whereas in heparin, they are mostly *N*-sulfated [129].

In array-based glycan binding screens, Pwp3A and Pwp7A showed strong binding to oversulfated iduronic acid (IdoA-2,4-disulfate), whereas Pwp5A showed only weak interaction. IdoA-2,4-disulfate is an unnatural synthetic monosaccharide that contains two axial sulfate groups at the C2 and C4 positions and that has never been isolated from a natural source [170]. However, IdoA-2,4-disulfate has been shown to bind to the heparan sulfate-binding chemokine CCL20 with high affinity and in this way inhibits

the interaction of CCL20 to heparan sulfate and the endothelial cell surface [170]. CCL20 chemically attracts lymphocytes and is highly enriched on inflammatory epithelium [187]. Natural occurring iduronic acid appears to play a key role in binding site specificity for heparin/heparan sulfate binding proteins [132,188]. Therefore, it is striking to observe that several PwpA domains showed binding to natural heparin *in vitro*. Moreover, Pwp3A and Pwp7A showed binding to several heparin-like oligosaccharides, and in particular to those, which contained sulfated iduronic acid, whereas similar oligosaccharides without sulfation of the iduronic acids or with glucuronic acid substitutions did not bind to the screened PwpA domains. Moreover, Pwp7A showed binding to a heparin-like oligosaccharide, which contains terminal IdoA(2S). In contrast, Pwp3A did not bind to the heparin-like oligosaccharide compound with terminal IdoA(2S). However, Pwp3A showed binding to such heparin-like compounds, which contain a secondary, quaternary or sexternary IdoA(2S), indicating that the IdoA positioning has an impact of on the binding. Yet, the observed array-based heparin interactions remained partly inconsistent, since two heparin-like compounds, which contain sulfated iduronic acid, were not bound by PwpA domains. Moreover, the sulfation of the adjacent glucosamine appears to be necessary for binding, which might indicate for a high local degree of sulfation as possible cause for binding. In contrast, it was a remarkable observation that non-sulfated IdoA and GlcA induced several positive and negative significant thermal shifts in solution with PwpA domains, whereas Pwp3A and Pwp7A were not affected. This observation additionally indicates, that a sulfation of IdoA might be necessary for an effective recognition by Pwp3A and Pwp7A. Considering that neither the addition of GlcNAc, GlcN nor GlcNAc(6S) induced significant thermal shifts of Pwp3A and Pwp7A, all being part of heparin/heparan sulfate subunits, emphasizes this hypothesis.

Beside being part of heparin and heparan sulfate, IdoA and IdoA(2S) are also found as components of dermatan sulfate. Moreover, IdoA(2S) is a key component of the heparin-analog *Fondaparinux*, which binds with micromolar affinity to Pwp1A and Pwp5^{N131D}. In addition, GlcNAc(6S) and GlcA are components of heparin/heparin sulfate and induced significant thermal shifts for Pwp1A, whereas both glycans did not show binding to Pwp1A in ITC analysis, probably due to low affinity. However, affinity does not necessarily correlate with specificity, since the flocculin *Sc*Flo5A is highly specific for mannose but binds only with millimolar affinity [93]. Though, the glucosamine in the heparin analog *Fondaparinux* is either di- or tri-sulfated, which might be seen as a possible explanation for a low affinity of Pwp1A to mono-sulfated GlcNAc(6S), due to missing sulfate groups. Aside from heparin/heparan sulfate, GlcNAc(6S) is also part of the glycosaminoglycan keratan sulfate.

Surprisingly, heparin of low molecular weight (~5 kDa) led to a strong destabilization of all PwpA domains. This observation might be explained by an aggregation effect, caused by the high density of negative charge in the heparin polymer. Such an aggregation might also reflect a desired feature that would increase adhesion of *C. glabrata*. It has been suggested that the affinity of fungal adhesins might be amplified by clustering of hundreds or thousands of adhesins on the cell surface, possibly through amyloid multimerization of the repetitive B-regions [103]. A clustering effect might also explain the paradox that some putative fungal adhesins show weak ligand affinity *in vitro* but strong adhesion *in vivo.* However, *C. glabrata* might not encounter native heparin in high concentrations on host surfaces (with the exception of a direct granulocyte/mast cell encounter). Instead, heparan sulfate is highly distributed in connective tissues and has a lower charge density than heparin. In addition, most GAG-binding proteins interact with heparin and heparan sulfate [189]. Therefore, heparin/heparan sulfate interactions might also indicate the binding of other GAG structures.

4.1.3. Glycosaminoglycan - protein interactions and the role of heparin

GAG binding proteins are very diverse and do not share common folds in general [189]. The interaction of GAGs to endogenous proteins with different functional properties such as cytokines, growth factors, adhesion molecules and enzymes is usually enabled by their unique sulfation pattern and is mostly mediated through electrostatic interactions between the negatively charged sulfate groups, uronic acids and positively charged amino acid residues in the receptor protein [129]. In this context, GAGs may play roles in the guidance of protein ligands to ECM tissue locations or as cell surface receptors [129]. In some cases, the interaction has been shown to depend on rare but specific sequences of modified sugars in the GAG chain, such as binding of the blood coagulation inhibiting enzyme antithrombin III to heparin and heparan sulfate. Here, the interaction is mediated via a unique pentasaccharide sulfation sequence within the GAG polymer, which binds with nanomolar affinity to a specific binding site and causes a conformational change of antithrombin III that, in turn, results in its activation [190,191]. The complex then binds to coagulation factors such as thrombin and factor Xa, causing their inactivation [192].

Nowadays, heparin plays an important role in medicine since it is utilized as an anticoagulant drug for the inhibition of blood clotting. In this context, low molecular weight heparin from porcine origin is frequently applied. Furthermore, a variety of medical devices such as vascular catheters are coated with heparin in order to provide

improved blood compatibility [130], whereas the anticoagulant surface activity of vascular endothelium is naturally given by the presence of heparan sulfate [193]. While heparan sulfate can be synthesized by virtually all cells of the human body, the synthesis of heparin is restricted to mast cells found in connective tissue and mucosa [194,195] The synthetic heparin analog *Fondaparinux* (*Arixtra*®) has been developed to mimic the unique pentasaccharide sequence found in heparin and heparan sulfate. It binds to antithrombin III with high affinity, resulting in increased activity and high antithrombin III mediated selectivity to factor Xa [196-198].

Several known biological processes in humans are related to the interaction of GAGs with protein ligands. For example, GAGs are known to play significant roles in inflammation [199], cell proliferation and migration [200], cancer development [201] and diseases of the nervous system by the mediation of β-amyloid formation [202]. However, specific functions of GAGs and proteoglycans in the human body are exceptionally diverse and in many cases are not yet well understood. Their biological function mostly depends on the interaction of the glycosaminoglycan chains with different protein ligands, whereas the proteoglycan core determines the level of expression [129]. Beside their interaction with endogenous factors, GAGs have been demonstrated to function as receptors for a number of pathogenic bacteria, viruses, parasites and fungi [136-138,203-212].

4.1.4. Glycosaminoglycans in host-pathogen interactions

Beside their interaction with endogenous factors, GAGs have been demonstrated to function as receptors for a number of pathogens. Pathogenic bacteria, viruses and parasites employ host cell-surface GAGs as receptor molecules for the decisive cell attachment in the initial stage of infection, as well as eukaryotic cell invasion and intercellular migration processes [136-138]. A variety of viruses including herpes simplex and hepatitis B interact specifically with heparan sulfate and utilize it for the initial attachment to host cells [203-205]. Heparan sulfate is also important to several aspects of HIV infection [206]. Moreover, pathogenic bacteria such as *Helicobacter pylori*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* bind to heparan sulfate on cell surfaces for initial attachment [207-209]. Parasites such as members of the genus *Plasmodium* have been demonstrated to bind to heparan sulfate and chrondroitin sulfate in hepatocytes, erythrocytes and human epithelium during malaria infection [210,211]. The pathogenic fungus *Blastomyces dermatitis* harbors the essential virulence factor BAD-1, an adhesin, which has been shown to bind to heparin *in vivo* [212]. In addition, a sequence-based search identified several proteins with potential heparin binding motifs in *C. albicans*; one of them was the integrin-like protein Int1 [138]. The study demonstrated

that Int1 greatly contributes to biofilm formation in heparin-coated vascular catheters. The heparin coating of catheters is sometimes considered as a potential biofilm promoting factor for *Candida* colonization [213,214]. Int1 has also been shown to be a virulence factor that enables adhesion to human epithelial cells [215].

In contrast, glycosaminoglycans are in some cases believed to have a pathogen defense function. For example, the luminal surface of animal and human bladder epithelium is covered with a layer of sulfated glycosaminoglycans, which is thought to be a first line of defense against bacterial urinary tract infections by diminishing adherence [216,217].

A few pathogenic bacteria from the genera *Streptococcus*, *Pasteurella* and *Escherichia* also synthesize GAG polymers to build extracellular capsules, which have been shown to contribute to their virulence [218-220]. Potential functions for bacterial GAG capsules in this context include the enhancement of adhesion to host tissue and the modulation of host physiology [133]. These GAGs are chemically identical or very similar to animal GAGs, which is thought to enable a kind of molecular mimicry and thereby permit successive infections [133]. Moreover, bacterial GAG capsules contribute to avoid phagocytosis and to resist the host action of complement.

4.1.5. Functional models for Pwp-glycosaminoglycan interactions in *C. glabrata*

Although *Candida glabrata* is an important human pathogen, its pathogenicity mechanisms are largely unknown. Adhesion and immune evasion strategies seem to play key roles during infection [55,221]. Since glycosaminoglycans are ubiquitously distributed and make essential and part of the metazoan extracellular matrix, they are attractive targets for adherence and invasion by various microorganisms and this could also be the case for *C. glabrata* [222].

The binding of pathogens to host-GAGs might be promiscuous and mostly based on electrostatic and relatively non-specific interactions of host-GAGs to negatively charged parts of surface-exposed proteins. Nevertheless, *C. glabrata* might utilize such unspecific interactions for providing an initial adherence to the host cell surface, to which a stronger and more specific interaction follows, resulting in a cascade of events leading to infection [223]. Therefore, an interaction by surface-exposed Pwp proteins with host-GAGs, followed by a stronger adhesion mediated by for instance Epas might provide *C. glabrata* with a more flexible interface of adhesins, even if the Pwp-mediated interaction is of non-specific ionic nature and low affinity. Alternatively, GAG-pathogen interactions can be relatively strong with nanomolar affinity, such as the binding of *H. pylori* to heparan sulfate [224] or the specific interaction of *Plasmodium*

falciparum VAR2 to chrondroitin sulfate [209]. In this context, the saccharide sequence and positioning of the sulfate groups may be crucial for the recognition event. Anyway, GAG-binding by another subset of adhesins such as the Pwp family might make sense for *C. glabrata*, since the Epa family already covers a wide spectrum of epithelial glycans and GAGs might serve as alternative targets during adhesion and host invasion.

As a primary commensal, *C. glabrata* is in permanent contact with the innate immune system. Therefore, it is essential to be not recognized as a pathogen and to trigger an immune response. In humans, the key phagocytes that mediate the killing of fungal pathogens are macrophages and neutrophils. *C. glabrata* has been demonstrated to use the uptake of macrophages to evade other immune cells and to proliferate in macrophages [23,54,55]. In addition, studies have shown that *C. glabrata* induces only transient pro-inflammatory cytokine responses compared to *C. albicans* [225,226]. Multiple highly sulfated glycosaminoglycans are stored in large amounts in secretory granules of mast cells, which can be found within connective tissue and mucosa [227,228]. The GAG content is released by degranulation in a pro-inflammatory reaction to invading organisms [228]. Moreover, it has been reported by MARUYAMA *et al*., that the release of GAGs by mast cells has inhibitory effects on the invasion of parasitic worms [229]. It is also known that macrophages derived from mice synthesize and secrete glycosaminoglycans [230]. Such GAGs could be potential binding partners for Pwp family members in the context of the interaction with immune cells. Therefore the Pwp proteins might bind secreted glycosaminoglycan to mask the cell from being recognized by the immune system. However, this raises the question of how other fungal commensals remain unrecognized. In this context, *C. glabrata* might also benefit from binding to GAG-synthesizing bacteria by utilizing their extracellular GAGs for molecular mimicry or as a physical barrier.

4.1.6. Functional models for Pwp proteins that are not related to GAGs

Although several experiments in this study indicate for interaction of PwpA domains with glycosaminoglycan such as heparan sulfate, the actual mode of binding remains unclear and there is still demand for further studies to define the specificity and to proof a biological significance. This means that Pwp proteins are currently orphan lectins with an intact binding site.

Beside the proposed lectin function, Pwp proteins might act as a structural constituent of the cell wall, probably involved in cell wall integrity. However, the expression levels are rather low compared to other structural cell wall components of *C. glabrata* such as Pir1-4 [128]. Moreover, Pwp proteins might interact heterophilic with other components of

the *C. glabrata* cell wall. Although flocculation was not observed upon heterologous PwpA-expression in *S. cerevisiae*, Pwp proteins could bind to components in the cell wall that are absent in *S. cerevisiae*. In addition, environmental parameters such as pH might be highly specific.

The mode, in which *C. glabrata* interacts with cells of the primary immune system is still unclear. Therefore, Pwp proteins might have an invasin function e.g. for e.g. such as the adhesin Als3 of *C. albicans*, which also acts as an invasin that triggers induced endocytosis [92,231].

Pwp proteins might also bind to protein components of the extracellular matrix since Pwp-expressing *S. cerevisiae* cells showed moderate binding to collagen type IV. However, collagen type IV is most likely highly glycosylated, which in turn might also be responsible for the observed adhesion.

Furthermore, the Pwps could have a function that is not related to human associated commensalism or infection. Since a co-evolution of *C. glabrata* and humans seems unlikely, the Pwps might be resposible for interactions with other organisms found in a heterogenic biofilm, such as bacteria or other fungal pathogens.

4.1.7. Influence of the D*cis*D motif on glycosaminoglycan interactions

In this study, the Pwp5A domain showed reduced affinity to heparin-like compounds in comparison to the mutant Pwp5^{N131D}A, indicating the Ca^{2+}-ion is directly involved in the interaction. The asparagine substitution in the conserved D*cis*D motif on CBL1 might reduce the affinity to the Ca^{2+}-ion, which might result in reduced affinity to heparin/heparan sulfate. However, the removal of calcium with EDTA did not affect the ability of PwpA domains to bind heparin. Another possibility would be that the Ca^{2+}-ion has a structural purpose for the PwpA domains, such as an increase of stability.

However, the N*cis*D motif from Pwp5, obtained from the reference strain CBS138 seems to be conserved among several *C. glabrata* strains since homologous proteins with similar A-domain sequences from various clinical isolates show the same feature: (azole-resistant strains FFUL887 (ID: SLM17229) [232] and DSY565 (ID: OBX47934) [233]; azole susceptible strain DSY562 (ID: OBX42635) [233]; clinical isolate 3A from HÅVELSRUD *et al*. (ID: KTB16063) [234]). However, there is great variation in the length of the corresponding repetitive B-region between the different strains.

4.1.8. The Pwp family potentially binds glycosaminoglycan in a unique way

A well-formed binding pocket provides the advantage of a better shielding from the surrounding solvent atoms and a greater interface for the interaction with potential glycan ligands. Thus, a greater interaction interface implies a higher ligand affinity. The PwpA domains do not possess an apparent pocket, which would correspond to the glycan-binding site in EpaA domains and other fungal adhesins (Fig. 60).

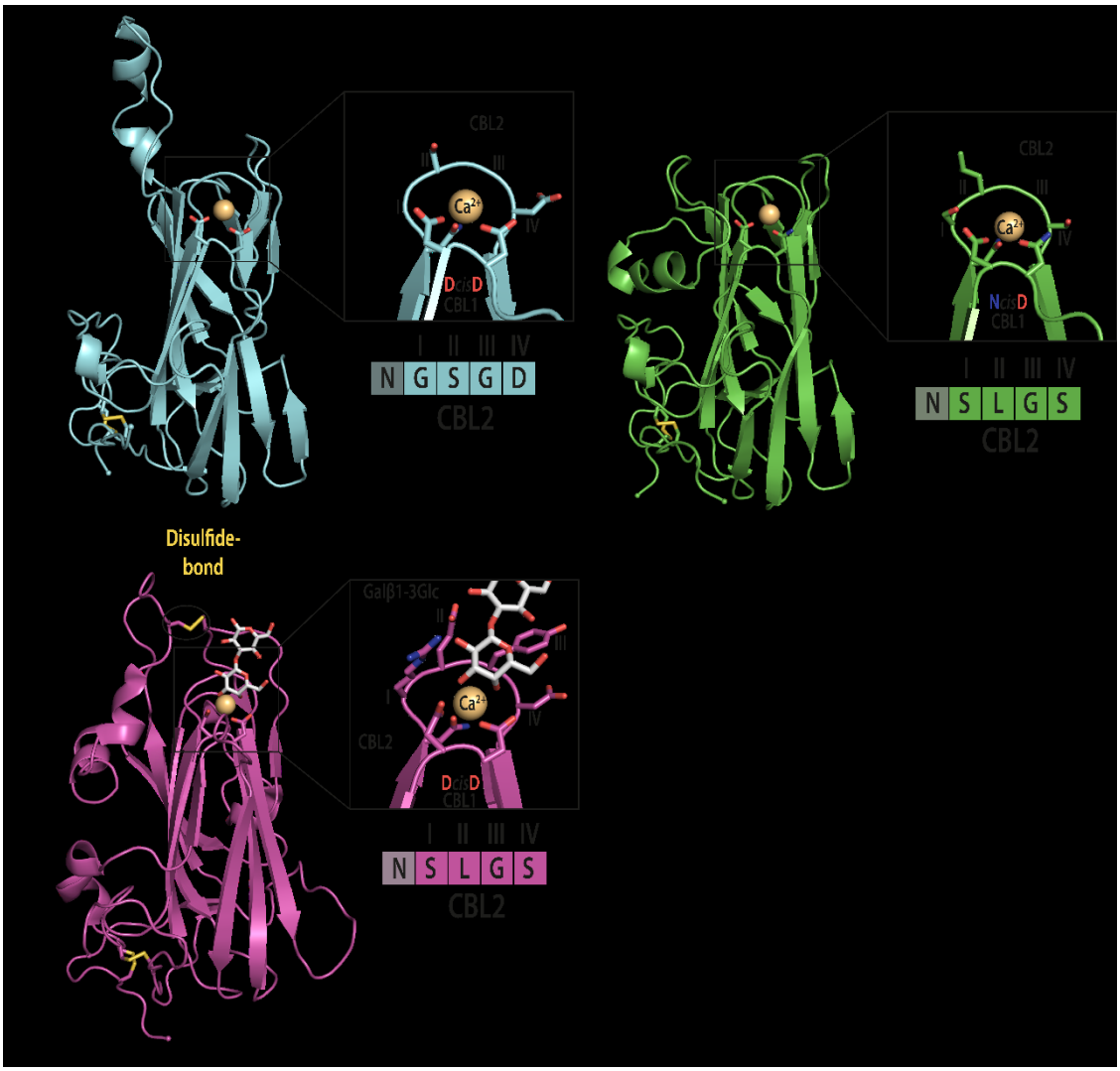

Fig. 60 Comparison of the calcium binding-sites in Pwp1A (A), Pwp5A (B) and Epa1A (C). CBL2 of Epa1A is densely covered with polar and aromatic side chains, whereas CBL2 of Pwp1A and Pwp5A are covered with serine residues and non-polar amino acid residues. CBL1 in Pwp5A shows an unusual N*cis*D peptide bond, which coordinates a Ca^{2+}-ion, as found in the D*cis*D motif of Pwp1 and the Epa family.

The Ca^{2+}-ion in PwpA domains is rather exposed to the surrounding solvent Moreover, aromatic residues are typical candidates to stabilize glycan ligands in the binding site, as it has been shown in Epa1A [70,108]. However, in PwpA domains, there are no aromatic residues that cover the Ca^{2+}-binding-site. These characteristics might explain why the PwpA domains did not bind to glycans on the mammalial glycan array and do not show a high affinity to monosaccharide components of heparin/ heparan sulfate, although

native heparin and *Fondaparinux* were bound by PwpA domains. In addition, the loop CBL2 is not densely covered with polar or aromatic amino acid residues, which are involved in glycan binding and discrimination in the Epa family [70]. In the Pwp family, the corresponding positions on CBL2 are primarily occupied by nonpolar residues or glycin. However, the mannose binding adhesin *Sc*Flo5 also has an open Ca^{2+}-binding-site and a CBL2 similar to that of the PwpAs but confers effective adhesion through relatively weak millimolar affinity [93]. A low affinity might be compensated *in vivo* by the avidity of many PwpA domains.

Since PwpA domains lack a potential glycan binding cavity around the Ca^{2+}-binding-site, the binding of heparin might also result from interactions with negatively charged sulfate groups elsewhere on the protein surface. In order to visualize and localize a potential interface for sulfate binding, the accumulation of sulfate ions was simulated on the surface of the Pwp1A structure (Fig. 61).

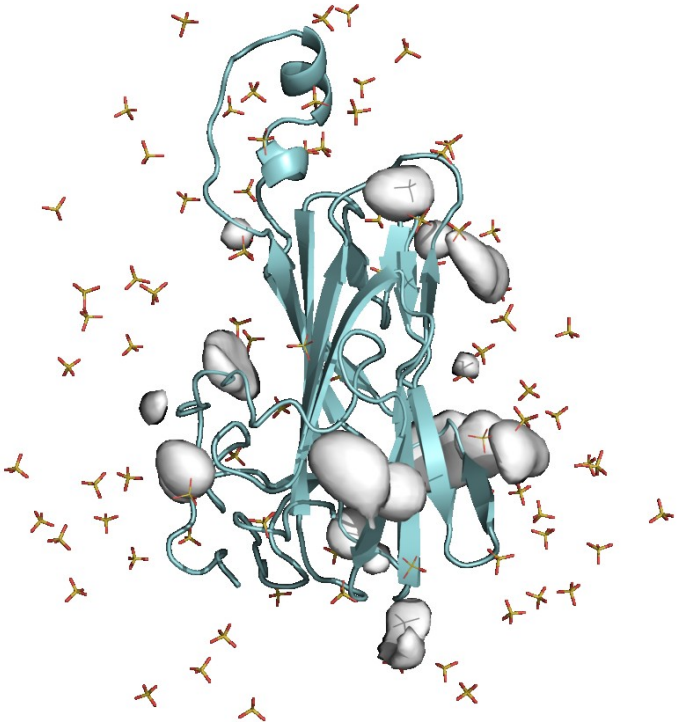

Fig. 61 Simulation of sulfate ions on the surface of the Pwp1A structure. The sulfates most likely accumulate in highest densities at the lower part of the domain (grey densities) closely to the positively charged *N*- and *C*-terminal region (Kindly provided by Prof. Dr. ESSEN).

The highest accumulation of sulfate groups was found in proximity to the lower end of Pwp1A, close to the *N*- and *C*-terminus. Consequently, Pwp1A has the most negatively charged surface area in this region. Here, according to an *AutoDock* simulation, sulfate groups of *Fondaparinux* most likely interact with Pwp1A via glutamine Q128, arginine R170, asparagine N196 and lysine K200 (Fig. 62). However, in its native constitution, this region of Pwp1 is probably sterically occupied by the linked B-domain. On the

other hand, the B-domain could also have a significant influence on the binding of sulfated polysaccharides, such as heparin, since it is supposed to bind huge amounts of calcium. The particularly long B-domain of Pwp1A, which is not specific for strain CBS138, is a noteworthy factor in this context. Another possibility is, that *Fondaparinux*/heparin oligosaccharides bind to Pwp1A via the left cavity, which is accessible through the turn of loop L1. Such a mode of binding might also include the interaction of the Ca^{2+}-ion with a terminal GlcNS6S. It is also possible, that such a mode of binding depends on the loose orientation of L1 or even a reverse closing of L1. The overlap of two Pwp1A domains might also be necessary for heparin binding [206].

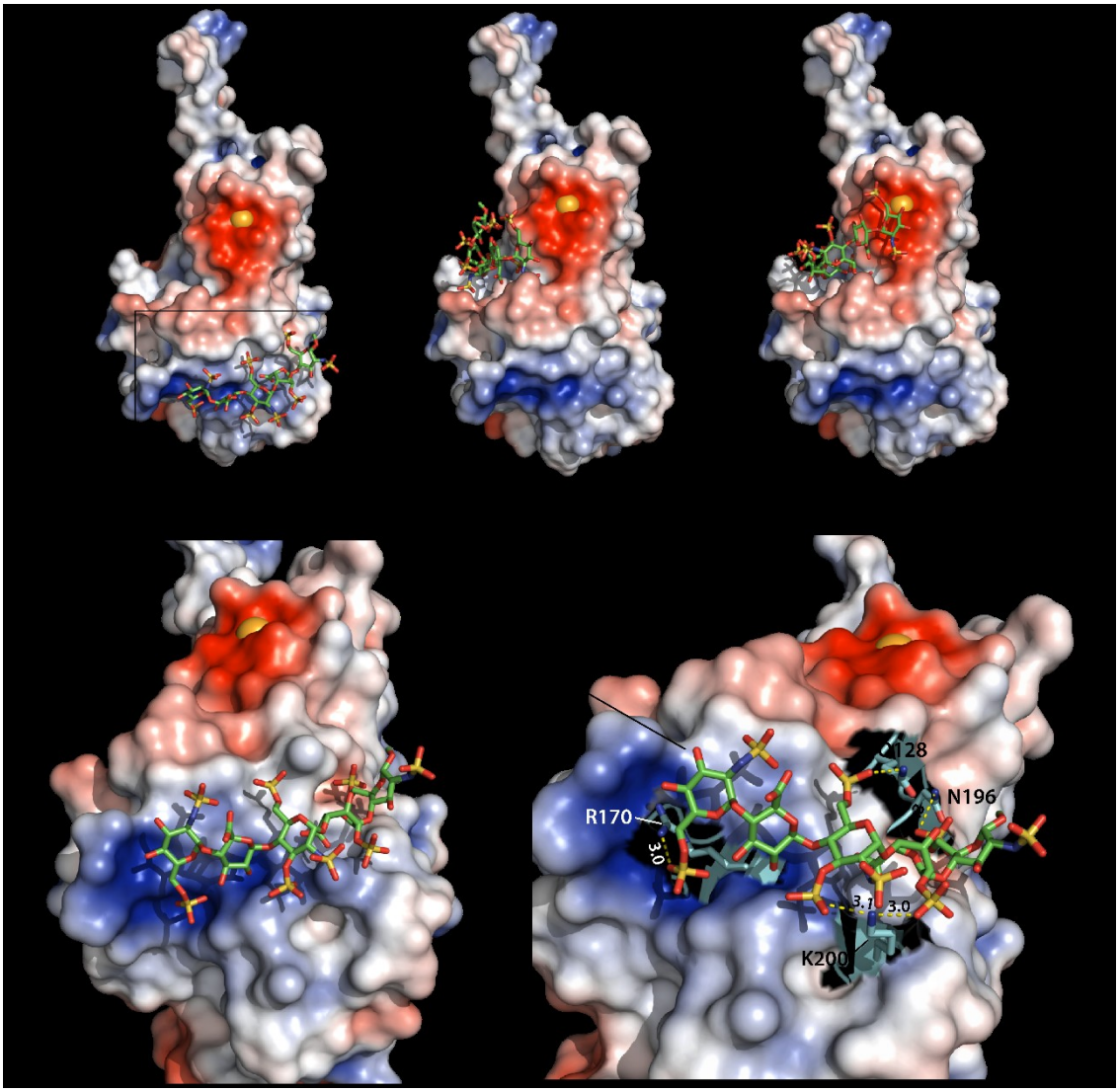

Fig. 62 Autodock-simulation of *Fondaparinux* on the surface of the crystal structure of Pwp1A. *Fondaparinux* docks with the highest affinity (docking energy: -6.5 kcal/mol) to the lower part of the Pwp1A domain (**A**, **D**). The 6-*O*-sulfate of the terminal GlcNS(6S) fits into a cavity with positive charge and might here interact with the guanidine group of arginine R170. Lysine K200 might interact with the 3-*O*-sulfate of the central GlcNS(3S,6S) and the 2-*O*-sulfate of IdoA(2S), whereas the 6-*O*-sulfates of the central a second terminal glucosamine might interact with Q128 and N196. Other putative binding sites with lower affinity (docking energy: -6.1/-6.0 kcal/mol) can be found in the cavity between the flexible loop L1 and the β-sheet core (**B**, **C**). Docking was simulated with *AutoDock Vina* [235].

The binding of GAGs by PwpA domains might via terminal monosaccharides but an unspecific type of binding to polysaccharide chains, which are variable in epimerization and sulfation. In this case, the Ca^{2+}-ion in the binding site might support the binding by covering the negative surface charge and reducing electrostatic repulsion.

4.1.9. Pwp-mediated adhesion *in vivo*

The expression of PwpA domains in an *S. cerevisiae* model system did not lead to adhesion in different adhesion assays, whereas moderate adhesion to type IV collagen has been observed. The level of glycosylation varies greatly between different types of collagen and between different tissues. Type IV collagen has been demonstrated to contain comparatively high levels of glycosylation [236-238]. Moreover, collagens are known to build a dense network with glycosaminoglycan and type IV collagen binds heparin with nanomolar affinity [239]. Therefore, collagen-bound glycosaminoglycan might be the cause of the observed adhesion. However, the addition of LMW heparin and *Fondaparinux* did not lead to apparent flocculation *in vivo*. The *in vivo* conditions for GAG binding are simply not known and the concentration of GAGs in the ECM and tissues varies greatly [189]. Moreover, specific parts and glycosylation in the B-regions of Pwp proteins might be essential for functionality *in vivo* and the Flo11B-domain, which was used for expression might have a negative effect on a proper function of the adhesion domain. Furthermore, the HA-tag might inhibit the function of PwpA-domains [109]. In addition, PwpA domains are sensitive to high salt concentrations. Interestingly, it has been shown, that high salt concentrations can inhibit floc formation in *S. cerevisiae* [240].

In contrast to the *EPA*s, the *PWP* genes are not localized in subtelomeric regions and therefore they are not subject to subtelomeric silencing, with exception of *PWP6*. However, until now a direct presence of Pwp proteins on the *C. glabrata* cell surface has not been observed and the expression is weak during laboratory cultivation [78,128]. This could be due to unknown conditions, which might be required for Pwp-expression. Such conditions might include a number of factors e.g. biofilm formation or direct contact with host cells. However, it would be challenging to simulate a possible host-related expression and binding to host-glycosaminoglycans under laboratory conditions, since the specific conditions that are present during the contact of the *C. glabrata* cell and host surfaces and adhesion remain unclear.

4.1.10. Putative medical applications of heparin coatings

Heparin coatings on medical devices, such as catheters, are used to improve blood compatibility [241]. Since there are indications for Pwp-heparin interactions, such coatings might support *Candida glabrata* biofilm formation, for example during hospital stays. This hypothesis requires further experimental research like *in vivo* adhesion assays on heparin coatings and under different conditions with subsequent microscopy.

It has been shown, that the malaria parasite *P. falciparum* uses heparan sulfate during adhesion to vascular endothelium and erythrocytes and that the adhesion can be disrupted *in vivo* and *in vitro* by the addition of GAGs without anticoagulant activity [242]. Therefore, GAGs have been suggested as candidates for adjunct therapy with an impact on malaria mortality. Furthermore, it has been shown, that the binding of herpes simplex virus to host cells via surface glycosaminoglycan can be inhibited by heparin fragments, composed of five disaccharide units, containing at least 1.5 sulfates per disaccharide [243]. Similar applications might demonstrate a promising example for a novel therapy with modified glycosaminoglycan upon an invasive infection with *Candida glabrata.*

4.2. The *S. cerevisiae* Wsc family

Membrane-spanning sensors, embedded in the yeast cell wall detect environmental stress and contribute to cell wall maintenance and integrity. In this context, the cysteine-rich domains in the CWI-stress-sensors Wsc1, Wsc2 and Wsc3 from *S. cerevisiae* have been suggested to play an essential role in stress sensing. The postulated sensor function derives from clustering and binding to glycans in the yeast cell wall in combination with a spring-like trigger that transmits cell wall stress to the cytoplasm. This work elucidates structural features of the cysteine-rich domains of the Wsc sensor family based in the crystal structure of the Wsc1-CRD. However, putative ligands and the mode of function remain subject of discussion.

4.2.1. Structural features of Wsc-CRDs indicate hydrophobic protein-protein interactions

Well diffracting crystals of the Wsc1-CRD allowed the successful elucidation of the structural basis, revealing a tightly packed and highly rigid globular domain, stabilized by four disulfide bonds. Wsc1-based homology modeling shows a highly similar fold in the *Sc*Wsc family with similar elements despite relatively low sequence similarity. The α-helix (α1) might be shortened in Wsc2 and Wsc3 (Fig. 63). However, putative glycan binding sites cannot be observed, similar to the homologous domain Kremen1. In line with this, the Wsc-CRDs did not bind to glycans in an array-based analysis and thermal shift analysis failed due to high structural rigidity. In contrast, the structure of the Wsc1-CRD and corresponding homology models revealed three clusters of mostly conserved aromatic residues, mainly from tyrosine.

Fig. 63 Homology models of Wsc domains from *Sc*Wsc1-3. A: Wsc1. **B**: Wsc2. **C**: Wsc3. Modelling was performed with *SWISS-MODEL* based in the crystal structure of the Wsc1 Wsc domain. Wsc2 and Wsc3 have a sequence identity to Wsc1 of ~ 35% each.

Clusters of solvent-exposed aromatic residues build hydrophobic regions on the protein surface, which might indicate a function that is based on hydrophobic protein-protein interactions (Fig. 64). However, dimerization was observed only in dependency with the 6×*His*-tag and the crystal packing of Wsc1 shows contacts via tags of two domains. The functional A-domain of the flocculin Flo11 from *S. cerevisiae* does not show any glycan binding capabilities but confers homotypic protein interactions via clusters of surface-exposed aromatic residues (tryptophan and tyrosine), similar to those of Wsc1 [94].

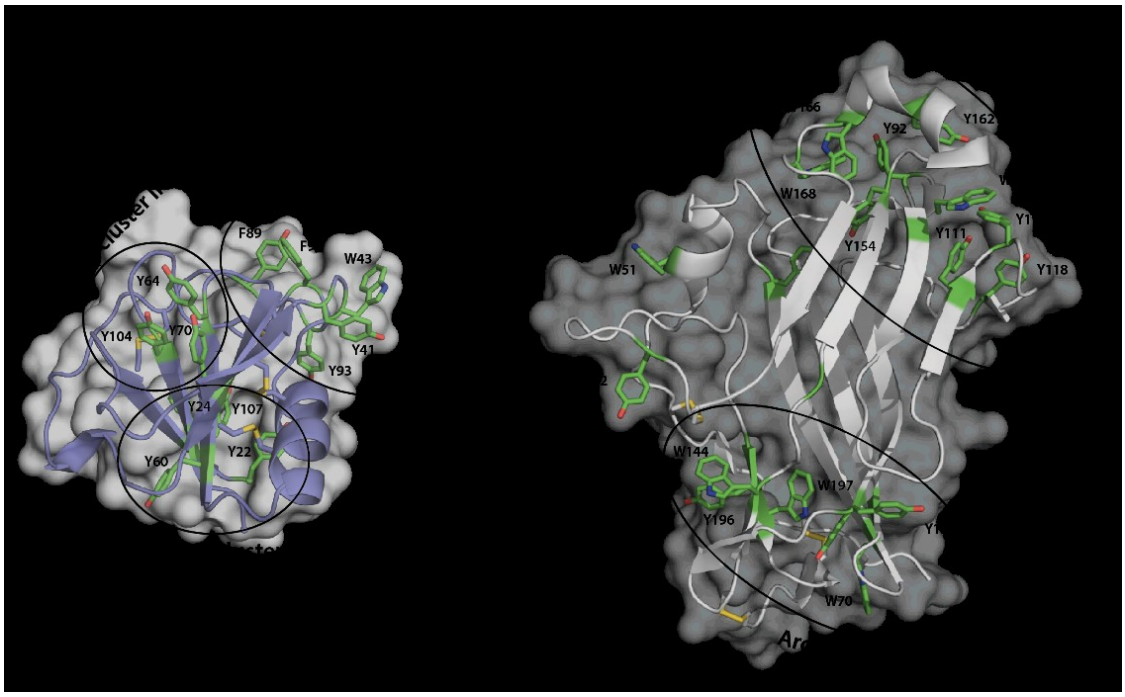

Fig. 64 Crystal structures of (A) Wsc1-CRD and (B) *Sc*Flo11A with solvent-exposed aromatic residues. In Wsc1, aromatic residues build three distinct clusters. In *Sc*Flo11, aromatic residues build two bands.

Homotypic Flo11 interactions have been shown to occur under strong pH-dependency [94]. A pH of 5.5, close to the isoelectric point of *Sc*Flo11A, results in Flo11-Flo11 interactions. However, these interactions disappear upon an increase to pH 6.8, most probably due to ionic repulsion of adjacent acidic amino acids. Such pH-dependent clustering effects might explain, that Wsc-CRD clustering has not been observed *in vivo*. Although the MST experiments were performed at pH 7.5 and pH 5.5, the more favorable pH for homotypic or heterotypic interactions might be more acidic since the isoelectric point of the Wsc1-CRD is pH 4.0 and the Wsc1-CRD domain harbors twelve surface-exposed acidic residues. In this context, conditions such as ionic strength might also have an effect on Wsc1-CRD clustering.

4.2.2. Putative functions of Wsc-CRD clustering for *S. cerevisiae*

The postulated functional model of Wsc sensors describes a spring-like nanosensor that includes the anchoring of the CRDs in the cell wall by binding of glycans. However, the structure of the Wsc1-CRD does not possess a potential glycan-binding site and glycan-binding was absent in an array based glycan screen. The array, however, does not contain β1-4 linked *N*-acetylglucosamine compounds, which are the subunits in chitin polymers, a key component of the yeast cell wall. An explanation might be, that Wsc-CRDs bind to and build clusters along with glycan polymers in the yeast cell wall, rather than binding to terminal glycan units. A specific binding affinity to terminal glycan ligands might also be too weak for detection by the used *in vitro* measurement methods. However, such a weak affinity might be sufficient for functional binding upon clustering *in vivo*, presuming appropriate solvent conditions. For instance, the PA14-like domain of the flocculin Flo5 from *S. cerevisiae* binds mannose with millimolar affinity, which cannot be detected on an array. However, Flo5 still induces strong cell-cell adhesion *in vivo* [93].

The CWI sensors Wsc1 and Mid2 have been suggested to be mainly responsible for CWI signaling since gene deletions resulted in growth deficiency [152]. Mid2 has no CRD and does not cluster upon cell wall stress *in vivo*. In contrast, Mid2 has an *N*-terminal, conserved and glycosylated asparagine residue, which has been shown to be essential for sensor function through putative interactions with glycan compounds in the yeast cell wall [87,150]. The CRDs of the Wsc family also contain such a conserved and surface-exposed asparagine residue (Fig. 65). This asparagine residue is potentially glycosylated in Wsc1 and Wsc3 and might therefore also be involved in a sensor function [164,244].

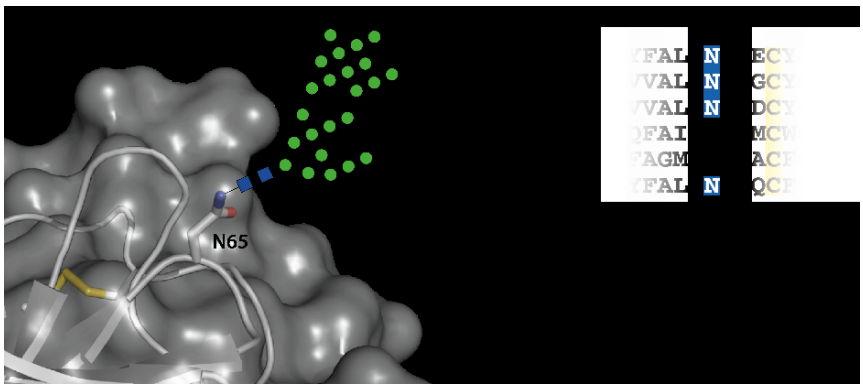

Fig. 65 Putative glycosylation of a surface-exposed asparagine N65 in the Wsc1-CRD.

Since glycosylation is missing during heterologous protein expression in bacteria, the lack of potential glycosylation of asparagine N65 might have a negative influence on the binding activity. Therefore an eukaryotic expression system might be necessary and more research needs to be done to determine the actual Wsc1 binding properties.

4.3. Outlook

4.3.1. Further studies of the *C. glabrata* Pwp family

The structural and functional characterization of the Pwp family revealed, that *Candida glabrata* might use an unexpected group of glycosaminoglycan as host substrates for adhesion. At least Pwp1A has a high affinity to the synthetic heparin pentasaccharide *Fondaparinux,* while Pwp5A does not bind in its natural form. However, Pwp5A binds *Fondaparinux* after a mutation on the calcium binding loop1 from N*cis*D to D*cis*D. Although this work sets the basis for the understanding of the Pwp family, a number of issues remain unclear. In order to fully characterize the Pwp family crystal structures of further PwpA domains remain to be solved. For a better understanding of the affinity of PwpA domains to glycosaminoglycan components, it would be interesting to further analyze the influence of the Ca^{2+}-ion in Pwp1A and other PwpA domains on the affinity during the binding to *Fondaparinux*. Especially, the affinity of Pwp2A to the heparan sulfate components IdoA and GalA should be further analyzed since these two uronic acids induced significant shifts in thermal shift assays. A possible inhibition of the Pwp1A-*Fondaparinux* interaction by the addition of IdoA2,4-disulfate could be tested in a titration measurement, as described by NONAKA *et al*. in 2014. In addition, the entire Pwp family should be screened for interaction with other glycosaminoglycan oligosaccharides, disaccharide- as well as monosaccharide components of glycosaminoglycan to clarify the target specificity. In the case of observed bindings, co-crystallization screens should be prepared to get valid information about an actual binding site. In order to further analyze the potential binding sites of *Fondaparinux* in Pwp1A, the potentially involved amino acids at the site could be mutated. A loss of affinity upon such mutations would verify that these amino acids contribute to ligand binding.

In order to get information about a possible influence of the Pwp-B domains, it would be interesting to repeat the *in vivo* adhesion experiments with the actual B-domains of the Pwp family instead of using the Flo11 system. The PwpB-domains might have a significant influence on the binding of long-chain polysaccharides with negative charges, such as heparan sulfate. That is particulary interesting since the B-domains are supposed to bind calcium.

In order to analyze the binding of Pwp expressing *S. cerevisiae* cells to heparin coated surfaces, a sensitive, microscopy-based method, described by JULIANNE GREEN *et al*. might be better than a standard adhesion test using wash assay.

4.3.2. Further studies of the *S. cerevisiae* Wsc family

Futher analysis are necessary to verify if the Wsc-CRDs are involved in potential pH-driven homo- or heterotypic interactions, similar to the function of *Sc*Flo11A (4.2.1). Therefore, thermophoresis- or SPR-measurements could be performed testing a wide range of different pH values.

Structure-based mutational analysis would be further helpful to understand the role of the aromatic clusters in Wsc-CRDs. The effect of these mutations on sensor clustering could be monitored *in vivo* under different stress conditions. In order to analyze the structural and functional significance of the aromatic clusters, single or multiple aromatic amino acids could be mutated to non-aromatic amino acids and the phenotypes could be compared *in vivo* under stress conditions. In addition, it might be interesting to further analyze the binding of Wsc-CRDs to chitin, either by an array or by ITC/MST with GlcNAcβ1-4GlcNAc or the monosaccharide GlcNAc under varying conditions such as pH, ionic strength or ligand concentration.

5. Methods

5.1. Bioinformatics

5.1.1. Sequence alignment, phylogenetic analysis and domain determination

For a comparative analysis of the A-domain regions of Pwp1-7 (*C. glabrata* CBS138), the sequences were aligned with known representatives of homologous fungal adhesins of the Epa-family (Epa1A - *C. glabrata* CBS138), the Flo-family (Flo5A - *S. cerevisiae* S288c) and Cea1A (*Komagataella pastoris*). Multiple sequences were aligned by the use of the *TCoffee Expresso* server, which combines primary sequence information with protein structural information [245]. The sequences used for the analysis were received from the *Candida Genome Database* (CGD) [123], the *Saccharomyces Genome Database* (SGD) [246] and the *UniProt* database [164].

The borders of the A-domain regions had to be determined precisely. *N*-terminal domain borders to the secretion signal sequences were determined by the use of the program *SignalP4.1*, which predicts cleavage sites of secretion signals by a combination of different artificial networks [247]. The *C*-terminal domain border to the B-region was determined by the occurrence of characteristic repetitive sequence [78]. In addition, the use of the protein domain prediction tool *GlobPlot 2* [248] contributed to theoretic discrimination of disordered inter domain regions and ordered domains.

In order to describe the degree of the phylogenetic relationship between the Pwps and other fungal adhesins of the PA14-type, the *maximum likelihood estimation* (ML) method was employed to generate a phylogenetic analysis based on the *N*-terminal A-domains. The ML is a computer-based comparison of gene- or protein sequences to derive evolutionary relationships between different biological entities like proteins. It uses a substitution model, which includes the probability of an exchange of two specific nucleotides or amino acids [249]. A structural phylogenetic tree was constructed by the use of the software *MEGA7* [250] The identity between the primary sequences of the PwpA domains was determined with the *BLAST* tool [251] The described methods were equally performed for Wsc-domains.

5.1.2. Protein parameters and glycosylation

Theoretical protein parameters were calculated by the use of the *ProtParam* [166] server. Putative eukaryotic glycosylation sites were predicted by the use of the

NetNGlyc 1.0 [244]and *NetOGlyc 4.0* [252] servers. Secondary structure content and positioning was predicted by use of the *JPred 4* server [253]

5.1.3. Primer design

Oligonucleotides for the amplification of specific DNA sequences were calculated with the program *Primer3* [254]. The primers were optimized for melting temperatures between 50-65°C, assuming a salt concentration of 50 mM and 300 nM for the oligonucleotide concentration. The primer- and amplicon sequences were checked for the formation of secondary structures by use of the *mfold* web server [255]. The directed cleavage of DNA by restriction enzymes and ligation of DNA fragments required a prior insertion of specific restriction sites next to the favored sequence. Such restriction sites were inserted by particular primers, which were complementary to the ends of the amplified DNA fragment and additionally had a sequence of the desired restriction site. The use of two different restriction sites at the ends of the DNA fragment ensured the correct orientation at the following ligation with the plasmids. Oligonucleotides in this study were synthesized by the *Microsynth AG*.

5.2. Molecular biology

5.2.1. Polymerase chain reaction: PCR

The polymerase chain reaction is a technique used for the amplification of specific DNA fragments *in vitro* [256]. Analytical colony-PCRs were carried out in 20 µl reactions by the use of a Taq DNA polymerase without proofreading function. DNA amplification for further use by cloning and mutagenesis PCRs was carried out in 50 µl batches by use of a Phusion DNA polymerase with proofreading function according to the provider's specifications. The number of PCR cycles, as well as the melting temperatures, were adjusted according to the specific primers and amplicons. Unspecific PCR products were avoided by a *touchdown* PCR program which includes a stepwise decrease of the annealing temperature with each cycle, starting from higher temperature and thereby minimizing unspecific pairing of primer and template [257]. The DNA polymerases elongate about 1000 base pairs in 15-30 s from 5' to 3' direction, while the annealing temperature should be about 3 °C above the melting temperature (T_m) of the primer with the lowest T_m.

In order to insert directed point mutations, deletions or insertions within a plasmid, mutagenesis PCRs were performed [258]. Therefore, primers with the desired mutation

were used to perform a whole-vector-PCR of the pJET1.2 vector with the according *PWP* inserts.

Colony-PCRs served as an analytical tool to verify the successful ligation of a PCR product into a vector. Therefore, an *E. coli* colony, carrying the template vector, was added to the PCR reaction. The denaturation temperature led to the release of the vector and hence the ligated insert could be amplified by the use of specific primers.

Tab. 18 Composition of PCR reactions

Component	PCR	Colony PCR	Mutagenesis PCR
SM buffer (10×)	5 µl	-	-
Phusion buffer (5×)	-	-	10 µl
Taq buffer (10×)	-	2 µl	-
dNTPs (10 mM)	1 µl	0.5 µl	1 µl
DMSO	1.5 µl	-	-
Forward primer (20 µM)	1 µl	0.25 µl	1 µl
Reverse primer (20 µM)	1 µl	0.25 µl	1 µl
Taq DNA Polymerase	-	0.5 µl	-
Phusion DNA Polymerase	0.5 µl	-	0.5 µl
Template DNA (~10 µg ml $^{-1}$)	1 µl	1 colony	1 µl
ddH_2O	39 µl	16.5 µl	35.5 µl

Tab. 19 Programs for normal PCR and touchdown PCR

Program step	Temp.	Cycles	PCR	Touchdown PCR
Initial denaturation	98 °C	1 ×	30 s	30 s
Denaturation	98 °C	30 ×	15 s	10 s
Hybridization	× °C		30 s	20 s (- 0.2 °C)
Elongation	72 °C		30 s	20 s
Final elongation	72 °C	1 ×	5 min	5 min
Store at	8 °C	∞	∞	∞

Tab. 20 Programs for colony PCR and mutagenesis PCR

Program step	Temp.	Cycles	Colony PCR	Temp.	Cycles	Mutagenesis PCR
Initial denaturation	94 °C	1 ×	5 min	98 °C	1 ×	10 s
Denaturation	94 °C	20 ×	30 s	98 °C	15 ×	30 s
Hybridization	× °C		40 s	× °C		30 s
Elongation	68 °C		50 s	72 °C		3 min
Final elongation	68 °C	1 ×	5 min	72 °C	-	-
Store at	8 °C	∞	∞	8 °C	∞	∞

5.2.2. Restriction of DNA

Restriction endonucleases catalyze the hydrolytic cleavage of DNA at specific short sequences. All used type II endonucleases in this study leave overhanging (*sticky*) ends, except *Eco*RV, which leaves blunt ends. Analytical restriction digestions of plasmids were prepared in 10 µl reaction volumes and incubated at 37 °C for 1 h. The appearing band pattern in the following agarose gel electrophoresis allowed an evaluation of the composition of plasmid and insert. Restriction digestions in preparative scale were prepared in 50 µl reaction volumes and incubated at 37 °C for 4 h. An agarose gel electrophoresis allowed the separation of inserts from the plasmid backbone and a following extraction from the gel.

Tab. 21 Composition of restriction digestion reactions.

Analytical digest	Component	Preparative digest
1 µl	10 × buffer	5 µl
0.1 µl	Enzyme A	1 µl
0.1 µl	Enzyme B	1 µl
1 µl	Plasmid DNA	40 µl
ad 10 µl	ddH_2O	ad 50 µl

5.2.3. Ligation of DNA fragments

In order to ligate plasmids with DNA inserts, the T4-DNA ligase from the bacteriophage T4 was used. The enzyme catalyzes the formation of phosphodiester bonds between 5' phosphate and 3' hydroxyl termini of blunt- and sticky ends in dsDNA and requires ATP and Mg^{2+} as a cofactor. The reaction was incubated for 1 h at room temperature at an insert to backbone ratio of 1:5. The ligated plasmids were subsequently stored at -20 °C and transformed with *E. coli* TOP10.

Tab. 22 Composition of ligation reactions.

Component	Amount
T4-DNA ligase	0.5 µl (400 U μl^{-1})
10 × buffer	2 µl
Plasmid backbone	25 fmol
Insert	125 fmol
ddH_2O	ad 20 µl

5.2.4. Preparation of DNA after PCR

For the preparation of PCR products, the *QIAquick PCR Purification Kit* was used. Its function is based on the binding of dsDNA to a silica matrix at high salt a concentration and a pH ≤ 7.5, followed by elution at a low salt concentration and alkaline pH at 8.5 [259], while nucleotides and primers are too small to bind to the matrix and can be

washed away. A size above 100 bp ensures the efficient adsorption to the matrix. The purification protocol was carried out following the supplier's instructions.

5.2.5. Agarose gel electrophoresis

In order to analyze or prepare restriction digest products, the fragments were separated by their size through an agarose gel. The structure of the gel equals a fine-pored matrix. The gel is located in a chamber filled with TAE buffer (Tab. 40) and connected to a power supply. Due to the negative charge of nucleic acids by their phosphate groups, the fragments move to the anode through the gel matrix, while their size determines the speed of movement and enables the separation. The samples were mixed with 6× loading dye and loaded into a 1 % - 1.5 % agarose gel. The use of a *GeneRuler*® DNA ladder ensured the correlation of fragment sizes. The separation was carried out for 45 min at 120 V and 300 mA, followed by a 20 min staining with *GelRed*®. The dye intercalates with nucleic acids and fluoresces under the exposition to ultraviolet light, allowing the optical detection in a *GelDoc*™ *XR* system at 254 nm.

5.2.6. Preparation of DNA from agarose gels

For the extraction of DNA fragments from agarose gels, the *E.Z.N.A.*® *Gel Extraction Kit* was used, following the suppliers' instructions. The basic function of this kit is described in 5.2.4. Corresponding gel bands were cut out under a transillumination at 365 nm wavelength for visualization.

5.2.7. Determination of DNA concentration

The concentration of DNA was measured by the use of a spectrophotometer *NanoDrop ND-1000* at 260 nm, as nucleic acids absorb light at this particular wavelength. By the multiplication with the factor 50, the concentration of dsDNA (ssDNA: 33; RNA: 40) in the solution can be determined [260].

$$\text{cDNA}\,[\text{ng}\,\mu\text{l}^{-1}] = A_{260} \cdot 50$$

The ratio of A_{260}:A_{280} indicates possible contaminations with proteins. A ratio of 1.8 – 2.0 indicates a pure DNA solution [260].

5.2.8. Production of chemically competent *E. coli* cells

Starting from a single *E. coli* colony, a 50 ml culture in LB-medium was inoculated and incubated at 37 °C; 225 rpm until an OD_{600} of 0.6. The cells were harvested at 15000 g and 4 °C for 20 min and the supernatant was discarded. Afterwards, the cells were resuspended in 15 ml TFBI buffer and incubated on ice for 2 h. After another centrifugation at 15000 g and 4 °C for 20 min, the supernatant was discarded and the cells were resuspended in 2 ml TFBII buffer. The cell suspension was divided into aliquots of 50 µl which were snap-frozen in liquid nitrogen and stored ad -80 °C [261].

5.2.9. Transformation of chemically competent *E. coli* cells

The transformation of chemically competent *E. coli* cells with plasmid DNA followed a modified protocol after MANDEL *et al* [262]. A frozen aliquot of competent cells was thawed on ice for 20 min with 1-3 ml plasmid DNA (50-250 ng). The cells were subjected to a heat shock at 42 °C for 60 s and subsequently diluted in 1 ml of LB-medium, followed by incubation at 37 °C and 125 rpm. Afterwards, the cells were centrifuged at 5000 g for 2 min, the supernatant was discarded while leaving 100 µl in which the cells were resuspended. The suspension was plated out on an LB agar plate containing the corresponding selection marker and incubated at 37 °C overnight. Cells carrying a plasmid formed colonies.

5.2.10. Preparation of plasmid DNA

Plasmids from E. coli were extracted with the *QIAprep Spin Miniprep Kit*. The method is based on the alkaline lysis by BIRNBOIM & DOLY combined with the adsorption of plasmid DNA to a silica matrix by VOGELSTEIN & GILLESPIE [259,263]. Cells of a 5 ml overnight culture were centrifuged at 15000 g for 2 min and the pellet was treated after the manufacturers' instructions. Finally, the plasmid DNA was eluted in 30-50 ml elution buffer and stored at -20 °C.

5.2.11. Long-term storage of DNA in *E. coli*

E. coli TOP10 transformed with the different plasmid constructs were stored for long-term use. Therefore, 500 µl of an overnight culture was mixed with 1 ml glycerol (87 %) and snap-frozen in liquid nitrogen. The cells were stored at -80 °C.

5.2.12. *E. coli* cultivation

E. coli cultures were produced in LB-medium [264] or TB-medium [265]. Solid media were produced by the addition of 2 % (w/v) agar. The media were sterilized by autoclaving at 121 °C and 2 bar for 20 min. After cooling down of the media to ca. 60 °C, the required antibiotics for selection were added. The growth of the culture was monitored by measuring the optical density at 600 nm with a photometer. An OD_{600} of 1 equates about $2 \cdot 10^8$ cells ml^{-1} [266].

5.2.13. Transformation of *S. cerevisiae*

For the transformation of *S. cerevisiae* with yeast-expression plasmids, 6 µl of purified plasmid solution was mixed with an aliquot of competent yeast cells, 300 µl PEG buffer and 5 µl freshly denatured salmon sperm DNA (10 mg ml^{-1}). The mixture was incubated for 30 min at 30 °C. Afterwards, the cells were subjected to a heat shock at 42 °C for 15 min, followed by centrifugation at 3000 g for 1 min and addition of 950 µl YPD-medium. The cells were centrifuged again at 3000 g for 1 min and 900 µl of the supernatant was discarded. The cells were resuspended in the remaining medium and plated out on a plate with solid selection medium (SC-Ura) [267].

5.2.14. *S. cerevisiae* cultivation

S. cerevisiae was cultivated in YPD-medium or SC-medium for selection by auxotrophic complementation. The SC-Ura medium contained all canonical amino acids but was lacking in uracil. Solid media were also produced by the addition 2 % (w/v) agar. The media were sterilized by autoclaving at 121 °C and 2 bar for 20 min, while glucose and agar were autoclaved separately for the SC-medium. Cultures were incubated at 30 °C, shaking. The growth of the culture was measured by the optical density at 600 nm with a photometer. OD_{600} of 1 equates about $1.5 \cdot 10^7$ cells ml^{-1} [268].

5.2.15. Construction of *E. coli* expression plasmids

Genomic sequences were codon-optimized for heterologous expression in *E. coli* and synthesized by *GeneArt® Life Technologies*. The obtained plasmids were amplified in *E.* coli TOP10, and the recombinant genes were subcloned into the pET-28(a)$^+$ expression vector by use of the restriction sites *XhoI* and *NdeI* (Tab. 46). An artificial Pwp5A(N131D) domain was constructed in addition to the natural Pwp5A domain. The sequences for the Wsc-domains originate from the *S. cerevisiae* S288c laboratory strain and were processed and cloned similarly.

5.3. Protein production

5.3.1. Heterologous test expression

A range of conditions and *E. coli* strains were tested in order to achieve optimal solubility for the expression of each recombinant PwpA- or Wsc domain (Fig. 66).

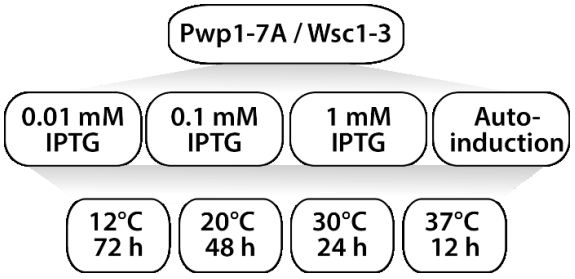

Fig. 66 Schematic overview of expression conditions that were tested for best solubility.

The expression was performed in culture-volumes of 50 ml. The strains *SHuffle®-T7 Express*, *BL21(DE3)Gold*, *Origami* or *Rosetta™* were tested for each construct. The expression was induced by IPTG (0.01 mM, 0.1 mM, 1 mM) in LB-Medium or by auto-induction in TB-Medium respectively. Each condition was tested for different temperatures and incubation times (12 °C for 72 h, 20 °C for 48 h, 30 °C for 24 h and 37 °C for 12 h). The cells were harvested at 15000 g and the pellets were resuspended in lysis-buffer (Tab. 40) with 5 mM EDTA, 5 mM PMSF, 1 µM Lysozyme). Glass beads (Ø 0.5 mm) were added to lyse the cells with a *FastPrep24®* homogenizer. The lysate was centrifuged 15000 g for 10 min and the soluble fraction was separated. The pellet was resuspended in lysis-buffer and both fractions were analyzed by SDS-PAGE. The protein solubility was estimated by the ratio of protein in the soluble and insoluble fraction.

5.3.2. Heterologous production of recombinant proteins

Recombinant proteins were overproduced at 12 °C for 72 h in 12 l LB-Medium (50 µg l^{-1} Kanamycin) by use of the *E. coli* strain *SHuffle®* T7 express. The pre-warmed medium was inoculated in a 1:100 ratio from a 100 ml overnight culture (LB-medium, 50 µg l^{-1} Kanamycin), starting from a single colony. The culture was incubated at 37 °C and 120 rpm until an OD_{600} of 0.6 was reached. Subsequently, the culture was cooled to 12 °C and the expression of the proteins was induced with 0.1 mM IPTG. After incubation, the cells were harvested by centrifugation (7000 rpm, 15 min, 4 °C) in an ultracentrifuge. The cell-pellets were resuspended in 20 mL lysis buffer (Tab. 40), shock-frozen in liquid nitrogen and stored at -80°C until preparation.

5.3.3. Cell lysis

Frozen cell pellets were thawed in a water bath at room temperature. 100 µg ml^{-1} lysozyme, 0.2 mM PMSF and a spatula tip of DNAse I were added, and the suspension was passed through a french pressure cell press 3 × at 4 °C and 1000 psi. By the arising shear forces at the exit of the pressure cell, the cells get ruptured. The flow-through was clarified by ultracentrifugation (35000 rpm, 45 min, 4 °C) and the supernatant was collected and sterile-filtered (0.22 µm).

5.3.4. Immobilized metal affinity chromatography

Recombinant proteins were first purified by Ni-NTA affinity. As the proteins were fused to an *N*-terminal hexa-histidine affinity tag, they could be immobilized on a Ni-NTA column. Within the column, divalent nickel ions are chelated by nitrilotriacetic acid (NTA) agarose. The hexa-histidine tag on the protein forms a complex with the nickel ions and thereby the protein remains in the column, while other proteins pass through [269]. By the addition of the histidine analog imidazole, the proteins can be competitively eluted from the column.

The Ni-NTA column was equilibrated with five column volumes of ddH_2O and the same volume of AM lysis buffer (Tab. 40) by a peristaltic pump. Then, the filtered lysate was loaded onto the column. The column was washed with ten column volumes AMI wash buffer and the protein was eluted with five column volumes of AMI elution buffer while collecting the flow through. The entire purification happened at 4 °C and a flow rate of 1 ml min^{-1}. After elution, the column was washed with each five column volumes of ddH_2O and 20 % (v/v) ethanol and stored at 4 °C.

5.3.5. Size exclusion chromatography

As the purification by affinity chromatography leaves some contaminations, a subsequent size exclusion chromatography (SEC) was performed in order to separate the recombinant protein. The SEC is a method in which proteins are separated by their apparent size [270]. In theory, proteins with a smaller hydrodynamic radius have a larger diffusion volume available while passing through porous column material than larger proteins. By their larger diffusion volume, smaller proteins have a higher retention time in comparison to larger proteins. Thereby, the globularity of the proteins also has an effect.

A 120 ml column with *HiLoad® Superdex® 75 pg* material was used as a stationary phase. The column was equilibrated with one column volume ddH_2O and one column

volume of either AM, AML or SEC buffer, depending on the type of protein. The choice of the buffer system was based on results of the thermal shift assays. The protein sample was concentrated to a volume of 1 ml and applied to the column. The sample was separated in the column with a flow rate of 1 ml min^{-1} and monitored using an Äkta® Purifier system. The elution was measured at 280 nm and collected in 2 ml samples and analyzed by SDS-PAGE. Afterwards, the column was washed with each one column volume ddH_2O and 20 % (v/v). The purification was performed at 4 °C and all buffers were degassed and filtered (0.22 µm) prior use.

5.3.6. Concentration of proteins

Protein solutions were concentrated in 10 kDa or 3 kDa *Amicon®-Ultra* concentrators depending on the molecular weight of the particular protein. A membrane within the concentrator holds back proteins with a molecular weight above the specific cut-off value, while the solution passes through during ultrafiltration. Therefore, the protein solution was filled into the concentrator of choice and centrifuged at 4000 rcf and 4°C in steps of 5-10 min until the desired volume was achieved.

5.3.7. Buffer exchange

In order to exchange the buffer system of a protein for different analytical methods, disposable *PD-10* desalting columns with a bed volume of 8.3 ml were used under gravity flow. The column was equilibrated with 5 CVs of the desired buffer and the 500 µl sample was filled in. The protein eluate in the new buffer whereas former buffer components remain in the column due to their small size.

5.4. Protein analytics *in vitro*

5.4.1. SDS-PAGE

The SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) is a method for the separation of proteins by their apparent mass in an electric field as described by LAEMMLI [271]. A polyacrylamide gel serves as a separation matrix. The protein samples are denaturated with SDS, high heat and a reducing agent like *β*-mercaptoethanol or dithiothreitol, which reduces disulfide bonds. The SDS also acts as surfactant, covering the proteins' intrinsic charge and conferring them a similar charge-to-mass ratio. The application of an electric field forces the proteins to migrate through the polyacrylamide matrix into the direction of the anode by their linear size. By staining with coomassie blue, the proteins are visualized and can be compared with a

molecular weight standard. The acrylamide concentration in the gel determines its separation performance and was adjusted to the size of the analyte. For the discontinuous gel electrophoresis, a two-phase gel with an acrylamide concentration of 4 % (stacking gel) and 12 % (dissolving gel) was poured (Tab. 37, Tab. 40). Samples were mixed with 2 × SDS loading dye and denatured at 95 °C for 5 min. The samples were transferred into the stacking gel in parallel with a *PageRuler*™ protein ladder and separated at 30 mA until the front reached the end of the dissolving gel. After separation, the gel was incubated in a bath of heated coomassie-staining buffer and destained in coomassie destaining buffer [272] (Tab. 40).

5.4.2. Determination of protein concentration

The concentration of protein solutions was determined with a photospectrometer at a wavelength of 280 nm. The absorption of proteins at this wavelength refers to the aromatic residues of tryptophan and phenylalanine, as well as disulfide bonds. That allows the determination of a protein solution by the LAMBERT-BEER law [273]. The extinction coefficient was calculated on the basis of the primary sequence by the *ProtParam* server [166].

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

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

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

E: extinction; $\boldsymbol{\varepsilon}$: molar extinction coefficient; m: mass; c: molar concentration; d: path length; V: volume; M: molecular weight; C_M: mass concentration

5.4.3. CD-spectroscopy

In order to get information about the secondary structure and thermal stability of the produced recombinant protein domains, melting curves and CD-spectra (circular dichroism) were measured. CD-spectroscopy is a method based on the optical isomers' property of interacting differentially with light. When an enantiomeric chromophore interacts with linearly polarized light, the two circularly polarized components of the wavelength will be differently absorbed, thus giving rise to circular dichroism. Although optical activity is given by the configuration of asymmetrical carbon centers, polymers composed of optically active monomers, such as proteins, also show distinct polarized-light-interaction behaviors that can be correlated with structural and conformational features. The chromophores responsible for protein circular dichroism are peptide bonds, which interact with far UV light (180-260 nm), and aromatic side

chains, which are responsible for absorption in the near UV range (260-320 nm). The asymmetric environment of peptide bonds is provided by secondary structures, while the tertiary structure is responsible for the differential absorption of aromatic amino acids. Structural information derived from far UV spectra are used in determining the secondary structure of a protein, while near UV spectra offer information on the general folding of the protein.

Melting curves provide information about the stability of a protein fold. Melting curves are executed by measuring ellipticity, the traditional signal of circular dichroism, at a single wavelength versus temperature. As protein folding is temperature-dependent, and absorption behavior is fold-dependent, it is possible to track protein denaturation by this method [124].

The measurement was performed with a J-810 spectropolarimeter (*Jasco*) in a quartz cuvette with a 1 mm gap (Tab. 23). The protein concentration was 200 µg ml^{-1} in 10 mM $NaHCO_3$ buffer (pH 8.3). The spectra were measured three times against the buffer spectrum. Collected data were converted to molar ellipticity per amino acid. The CD-Spectra were measured in the far UV range (190-260 nm). The *Jasco Secondary Structure Estimation* software was used for comparison with a reference spectrum to estimate secondary structure compositions. Subsequently, the melting curve was measured in a temperature range of 20 to 90°C, by increasing 1°C steps. Afterwards, the temperature was decreased in the same way. The wavelength for the measurement of the melting curve was selected by the minimum of the CD-spectrum. The data were normalized as follows:

$$[\Theta] = \frac{m \cdot \Theta}{(n-1) \cdot c \cdot d \cdot 10}$$

[Θ]: Ellipticity; Θ: Measured Ellipticity [mdeg]; m: Protein mass [Da]; n: Number of amino acids; d: cuvette gap [cm]; c: Protein concentration [g/mL]

Tab. 23 Settings for measurements on the CD-spectrometer.

Parameter	Settings CD spectrum	Settings melting curve
Sensitivity	100 mdeg	100 mdeg
Wavelength	/	depending
Start	260 nm (20 °C)	20 °C
Stop	195nm (20 °C)	90 °C
Data points	1/0.2 nm	1/1 °C
Mode	continuous	/
Speed	50 nm/min	50 nm/min
Integration time	2 s	2 s
Gap	1 mm	1 mm
Repeats	3	1

5.4.4. Thermal shift assay

The thermal shift assay (or Thermofluor) is a method that measures a shift in the thermal denaturation temperature of a protein under varying conditions. During the measurement, the melting point (T_m) of the protein is determined during a slowly increasing temperature by use of the fluorescent dye *SYPRO Orange*™ in a qPCR cycler. The dye interacts with hydrophobic regions of the protein that become exposed upon denaturation. The binding increases the fluorescence emission of the dye and serves as a readout of the denaturation of the protein. Changing buffer parameters like pH, salt or additives can have a stabilizing or destabilizing effect on the protein stability which is then represented by an increase or decrease of the melting point. In this way, the buffer system for a specific protein can be optimized by choosing the conditions under which the protein is most stable. In this way, the purification of recombinant proteins can be significantly improved. Moreover, increased thermal stability of a protein correlates with its ability for crystallization. Also, the binding of ligands like glycans, proteins or DNA usually increases the melting temperature. A thermal shift of ≥ 1.5 °C is considered to be significant.

The thermal shift assays were performed in a *Rotor-Gene Q* real-time PCR cycler (*Qiagen*) in volumes off 40 µl triplicates, using UV-permeable PCR cups (*Qiagen*) (Tab. 24). The temperature was increased linearly by 2 °C min^{-1} (Tab. 25). Additives that were screened for their effect on protein T_m: are summarized in Tab. 26 and Tab. 27.

Tab. 24 General composition of a thermal shift assay reaction.

Component	Concentration
Protein	10 µM
Sypro™ Orange in DMSO	7.8 ×
Additive	10-500 mM
ddH_2O or buffer	ad 40 µl

Tab. 25 Settings on the real-time PCR cycler for thermal shift assays.

Parameter	Setting
HRM (gain/source/detector)	(2.33/460 nm/510 nm)
Rotor	36-Well
Cycle	Melt
Ramp from [°C]	25 - 90
Rising by [°C / step]	0.5

Tab. 26 Buffers & pH screened for thermal shifts.

Buffer (20 mM)	pH range
Sodium acetate	4.0 / 5.0
MES	6.0
HEPES	7.0
Tris	8.0
CHES	9.0 / 10.0

Tab. 27 Additives that were screened in thermal shift assays.

Type	Additive	Concentration
Divalent cation	$CaCl_2$	10 mM
	$MgCl_2$	10 mM
	$NiCl_2$	10 mM
	$ZnCl_2$	10 mM
Salt	NaCl	0 / 50 /100 / 200 / 300 / 400 / 500 mM
Polyol	Glycerol	2.5 / 5 / 10 % (v/v)
Glycan	Glc	10 mM
	Man	10 mM
	Gal	10 mM
	IdoA	10 mM
	GlcA	10 mM
	Lac	10 mM
	GlcN	10 mM
	GlcNAc	10 mM
	GlcNAc(6S)	10 mM
	GalNAc	10 mM
	GalNAc(4S)	10 mM
	GalNAc(6S)	10 mM
	ΔIdoA(b1-3)GalNAc	10 mM
	ΔIdoA(b1-3)GalNAc(4S)	10 mM
	ΔIdoA(2S)(b1-3)GalNAc(4S)	10 mM
Polyanion	*Fondaparinux* pentasaccharide	10 mM
	Dalteparin LMW heparin	5 mM
	DNA (salmon sperm)	10 mM

5.4.5. Fluorescent labeling of proteins

For the glycan microarray analysis, purified PwpA-domains were fluorescently labeled with the tetrafluorophenyl (TFP) ester Alexa Fluor™ 488 (*Invitrogen*). Dyes of the Alexa™ Fluor family are frequently used in science and produced by sulfonation of coumarin, rhodamine, fluorescein or other dyes. Alexa Fluor™ 488 has an absorption maximum at 496 nm, while it's emission maximum is at 519 nm [274]. The TFP conjugate is a fluorescin derivate, which builds an amide linkage with primary amines in proteins, like the *N*-terminus or amino acid side chains, but also to buffer compounds like Tris.

The labeling buffer needs to have a slightly basic pH to ensure the deprotonation of the terminal amine of lysine.

For the labeling reaction, the proteins were transferred into a 0.1 M $NaHCO^3$ labeling buffer at pH 8.3 by the use of a PD10 column and concentrated to 2 mg ml^{-1}. 0.5 ml of the protein solution was incubated with 12.5 µl of Alexa Fluor™ 488-TFP (10 µg ml^{-1} in DMSO) for 1 h at room temperature on a roller. Afterwards, the labeled protein was separated from unbound dye by another exchange into a specific protein buffer (SEC) through a PD10 column. The sample was stored at 4°C in the dark until the measurement.

5.4.6. Glycan microarray analysis

In order to identify glycan ligands, microarray analysis were performed with PwpA- and Wsc domains. Therefore, the fluorescently labeled proteins were incubated on glass slides, on which a library of natural and synthetic glycans were printed. These glycans vary in length, linkage or carbohydrate composition. Three different types of glycan array were tested:

1. A mammalian printed glycan microarray (V 5.2) in cooperation with the *Consortium for Functional Glycomics* (CFG) [275] (Request #3059; #3171). The array has a library 609 natural and synthetic glycan targets, which are presented by mammalian cells and are covalently linked to the NHS chip surface by amino linkers, each in six-fold copies. The fluorescently labeled protein was diluted to concentrations of 2 µg ml^{-1}, 20 µg ml^{-1} and 200 µg ml^{-1} in CGF-buffer (0). The chip surface was incubated for 2 h. After incubation, the chips were repeatedly washed with the same buffer and the remaining fluorescence was detected and quantified as described by HEIMBURG-MOLINARO *et al.* [276]. The Epa1A domain served as a positive control, as it binds to terminal galactosides [70,108].

2. A pathogen printed glycan microarray in cooperation with Prof. Dr. SEEBERGER from the *MPI für Kolloid- und Grenzflächenforschung* (MPIKG) [171,172,277]. The array has a library of 140 synthetic glycans, which represent potential targets for glycan-binding proteins from microbial pathogens.

3. A glycosaminoglycan array in cooperation with Prof. Dr. SEEBERGER from the *MPI für Kolloid- und Grenzflächenforschung* (MPIKG). Beside natural heparin, the array has a library of defined oligosaccharide sequences from different types of glycosaminoglycans and synthetic oversulfated heparin glycans.

The array slides were incubated with fluorescently labeled protein in SEC-buffer for 1 h and repeatedly washed with the same buffer. The remaining fluorescence was detected, while Epa1A served as a positive control.

5.4.7. Isothermal titration calorimetry

The isothermal titration calorimetry (ITC) is a quantitative method for the analysis of thermodynamic reactions, based on the work of WISEMAN *et al.* [278]. Here, an ITC experiment is the titration of a biomacromolecule (protein) in solution by a reactant (ligand) solution at a constant temperature to obtain the exchanged heat of the protein-ligand reaction. The experiment is typically performed in a titration calorimeter by the injection of the ligand solution in 1-10 µl volumes to 200-1000 µl of the binding protein solution. The binding reaction consumes or generates heat, proportional to the amount of injected ligand and binding enthalpy (ΔH). Thus, measuring of the heat during binding enables the determination of the binding constant (K_D), reaction stoichiometry (n), enthalpy (ΔH) and entropy (ΔS) and provides a thermodynamic profile of the molecular interaction by:

$$\Delta G = -RT \cdot \ln K_a = \Delta H - T\Delta S$$

The operation principle of an ITC instrument is based on a feedback mechanism according to the dynamic power compensation principle. It measures the amount of power (µcal/sec) required to maintain a constant temperature difference between a sample and a reference cell. Initially, the feedback system continuously applies a small power to the sample cell, which determines the baseline level. Each injection of the ligand triggers the binding reaction and, depending on the binding affinity and the concentration of reactants in the cell, a certain amount of protein/ligand complex is formed. The formation of the complex is accompanied by the release (exothermic reaction) or the absorption (endothermic reaction) of heat that causes a difference in temperature between the two cells. Then, the feedback system either lowers or raises the thermal power applied to compensate the temperature unbalance. After each injection, the system reaches equilibrium and the temperature balance is restored. Therefore, the recorded signal shows a typical deflection pattern in the form of a peak. Integrating the area under the peak, assuming the baseline as a reference, provides the amount of heat associated with the injection. As the reactant in the cell becomes saturated, the heat signal diminishes until only the background heat is observed.

For this work, an ITC_{200} (*MicroCal*) with a 200 µl reaction- and reference cell and a 40 µl titration syringe was used. The reaction was measured at a protein concentration

of 200 µM in the reaction cell and a ligand concentration in tenfold excess of 2 mM in the syringe. The following settings were used:

Tab. 28 Settings for ITC experiments.

Parameter	Value
Injections	20
Temperature	20° C
Reference power	10 µcal/s
Initial delay	60 s
Volume first injection	0.4 µl
Duration first injection	0.8 s
Duration between injections	180 s
Filter	5 s
Volume other injections	2 µl
Duration other injections	4 s

5.4.8. Heparin-column binding assay

In order to analyze the binding of Pwp adhesion domains to heparin as a member of glycosaminoglycans, a heparin-agarose column binding assay was performed. Therefore, purified Pwp adhesion domains (0.25-1 mg) dissolved in 1 ml corresponding SEC buffer, were loaded on a 1 ml HiTrap® heparin-agarose column. The column was then flushed with 7.5 ml (1 ml min^{-1}) of 25 mM EDTA solution (in SEC buffer) in order to remove calcium ions and to check for calcium-dependent heparin binding in this way. Afterwards, the column was flushed with 1 M NaCl in order to elute the remaining protein. The elution of protein was monitored via an Äkta® purifier system at 260 nm wavelength. In order to estimate and compare binding affinity to the heparin-column a protein retention assay during an increasing NaCl gradient (50 mM to 1 M NaCl in SEC by 50 ml at 1 ml min^{-1}) was performed. In order to check for putative interactions of bound protein to the agarose-matrix in the heparin-column, purified Pwp adhesion domains were also applied to a GST-agarose column and equally submitted to a NaCl gradient. BSA and thrombin served as controls.

5.5. Protein structure determination

The three-dimensional structure determination of proteins with atomic resolution is a powerful tool to gain insight into biochemical and biomolecular processes. It contributed to a detailed understanding of many relevant questions in life science. Nowadays, two primary methods for protein structure determination at atomic resolution are used: X-ray crystallography [279] and nuclear magnetic resonance spectroscopy (NMR) [280]. Other methods to gain structural information with lower resolution are small-angle X-ray scattering (SAXS) [281] and cryo-electron microscopy (cryo-EM) [282]. These techniques are mainly used to study quaternary structures of protein complexes or protein-protein interactions. While X-ray crystallography utilizes the diffraction of hard X-rays in crystals to gain structural information with high resolution, in NMR ^{13}C- or ^{15}N- spin-labeled proteins are used to record multi-dimensional spectra. The coupling of nuclei together with restraints provide an ensemble of structures that represent flexible areas of the protein. Compared to X-ray crystallography, NMR has the advantage of being able to work with proteins in solution and is more suitable for high throughput usage [283]. Though, the method of NMR is limited to a maximum protein size of 30 kDa [284]. The advantage of X-ray crystallography, on the other hand, is the omitted labeling of the protein and its unlimited size. X-ray crystallography is still the most frequently used method for protein structure determinations. A look into the Protein Data Bank (PDB) [285] reveals 109653 solved protein structures by X-ray crystallography in contrast to 10398 protein structures solved by NMR.

5.5.1. Protein crystallization

For the structural determination of a protein by X-ray diffraction, the protein needs to be arranged in a mono-crystal. As it is impossible to derive the optimal crystallization conditions for a specific protein by its primary sequence, determining these conditions is the critical step of X-ray crystallography and basically depends on trial and error. Therefore, large scale sparse-matrix-screens were performed in order to cover a wide variety of different crystallization conditions and to identify the most appropriate. In these conditions, different chemical properties are variable and combined in different ways, like for example the pH, ionic strength, precipitant concentration, protein concentration or metal ions.

For the initial crystallization experiments, several commercially available sparse-matrix-screens were used. These screens comprise optimized conditions which often lead to crystal growth, based on the literature [286,287].

Tab. 29 Crystallization screens that were used in this study.

Screen	Specifications
JCSG Core I-IV (*Qiagen*)	Optimized conditions by the *Joint Center for Structural Genomics* (JCSG)
JCSG Core+ (*Qiagen*)	Precursor of the optimized Core screens
AmSO4 Suite (*Qiagen*)	Systematic screen with ammonium sulfate precipitant
PACT (*Qiagen*)	Different combinations of polyethylene glycol and ions
MbClass (*Qiagen*)	Optimized conditions for membrane proteins
Morpheus (*Mol. Dimensions*)	Polyethylene glycol based screen
Morpheus II (*Mol. Dimensions*)	Optimized follow-up of the Morpheus screen
Penta (*Jena Bioscience*)	Pentaerythritol polymers as precipitant

The aim of these screenings was to find conditions under which the protein undergoes nucleation and continues to grow to macroscopic crystals. The screens were prepared as a sitting drop vapor diffusion experiment. In vapor diffusion, a drop containing a mixture of precipitant and protein solution is sealed in a chamber with a pure precipitant solution in a reservoir. Due to the concentration gradient between both solutions, vapor diffuses out of the drop until the osmolarity of the drop and the precipitant are equal. The dehydration of the drop causes a slow concentration of both protein and precipitant, ideally into an unstable supersaturated zone where nuclei are formed. Then, the transition into the metastable zone with lower protein concentration forces the crystals to grow [288].

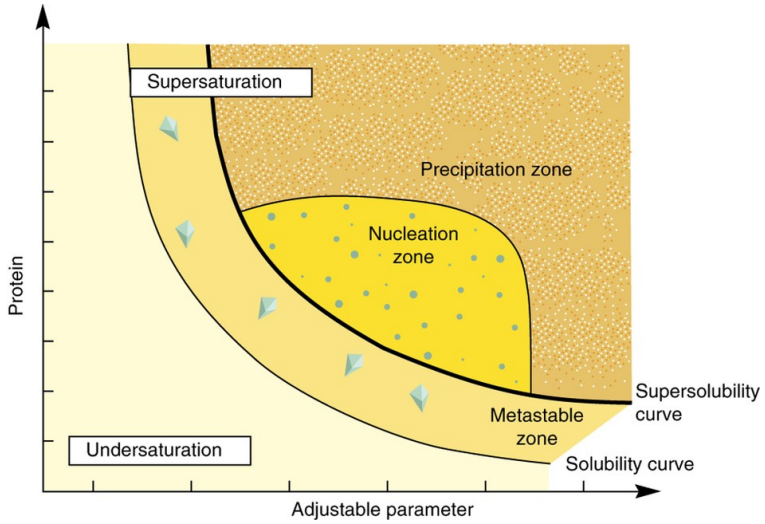

Fig. 67 Schematic illustration of a protein crystallization phase diagram. The adjustable parameter can be precipitant e.g. Solubility is defined as the concentration of protein in the solute that is in equilibrium with crystals. The supersolubility curve is defined as the line separating conditions where spontaneous nucleation (or phase separation, precipitation) occurs from conditions where the crystallization solution remains clear [289].

Therefore, the soluble proteins were concentrated to the maximum with centrifugal filters (Amicon Ultra, *Merck Millipore*). 300 nl of the concentrated protein solution with a concentration between 15 mg ml^{-1} and 50 mg ml^{-1}, depending on the specific protein, were mixed with the same volume of precipitant solution in 96 well plates (*Innovadyne*). This work was performed by the use of an automatic pipetting robot (Cartesian Microsys SQ 4004, *Genomic Solutions*). Beside the first drop, a second drop with half of the protein concentration was pipetted. The reservoir was filled with 80 µl precipitant solution. The plate was finally sealed with transparent plastic foil (VIEWseal, *Greiner BIOone*) and stored at either 4 °C or 18 °C or both. The development of the drops was documented in intervals by an automated light microscope (*Rockimager*, *Formulatrix*).

5.5.2. Optimization of protein crystallization

Crystalline material obtained from initial screenings is often not useful for protein structure determination due to poor crystal quality, improper morphology, small size or bad packing. Therefore, narrower optimization screenings were performed based on the conditions in which crystallization events were first observed. However sometimes crystals of sufficient quality form already under the initial screening conditions.

The optimization screens were performed by stepwise variations of specific chemical determinants from the initial condition like the pH, the concentration of salt and precipitant or additives. Additionally, different ratios (1:1, 1:2, 2:1) between protein and precipitant solution were tested in sitting-drop experiments in larger drop sizes of 2-3 µl in 24-well plates (*Qiagen*). Here, high concentrations (100 mM) of glucose, mannose and galactose were added in parallel drops. Furthermore, an additive screen (*Hampton Research*) containing different reagents of multivalent ions, salts, polymers, carbohydrates and organics was added to a duplicate of the prepared optimization screen, both prepared in 96-well plates (*Innovadyne*). If improper crystals occurred under the initial conditions, it was also used for micro-seeding. In that case, micro-crystals were produced from a larger crystal fragment by vortexing it together with a steel ball in the mother liquor. This liquid was then added to the optimization screens in order to induce crystals of better quality. The plates were stored under the same temperature were the first crystal growth was observed. The documentation of the crystallization was performed by either the *Rockimager* system (*Formulatrix*) for 96-well plates or manually under a light microscope for 24-well plates.

5.5.3. Co-crystallization

The determined crystallization conditions of Pwp1A and Pwp5A were basis for the co-crystallization with glycans and Fondaparinux, which resulted in positive thermal shifts (3.1.5). The crystallization conditions were reproduced in 24-well plates with 1 ml reservoir volumes. The concentrated protein solution was mixed with 5 mM glycan (in reservoir solution) in ratios of 1:1 µl, 1:2 µl and 2:1 µl and incubated according to the corresponding conditions that resulted in crystals.

5.5.4. X-ray diffraction experiments

Crystals for the diffraction measurements were picked with cryoloops (*Hampton Research*) or micromounts (*MiTeGen*) from the crystallization drops and immediately transferred into liquid nitrogen in order to avoid the destruction of the crystals. In order to avoid crystalline freezing of the mother liquor, the crystal is usually soaked in cryoprotectant solution before freezing. As all picked crystals for this study grew under conditions with either 20 % (v/v) glycerol or 20 % (w/v) PEG, sufficient cryoprotection was given by the mother liquor.

X-ray crystallography is based on the diffraction of X-rays by the electron clouds of atoms within the lattice of a crystal, which is defined by BRAGG's law [290]. Intensity and position of the emerging diffraction maxima are generated by constructive interference according to the content of the unit cell within the crystal. The diffraction pattern builds the basis for the generation of a structure model, while the amount of collected data strongly depends on the parameters and symmetry of the crystal. For this reason, not every measured crystal resulted in a sufficient dataset. The measurement was performed at the *European Synchrotron Radiation Facility* (ESRF) [291] in Grenoble (France) by members of the workgroups of Prof. Dr. ESSEN and Dr. BANGE.

5.5.5. Processing and data reduction & molecular replacement

The intensities of the detected reflections were integrated with the program XDS [292,293]. Thereby, the intensity of every reflection is measured and corrections for beam-position, detector-distance and mosaicity are included as well as the cell parameters. The completeness, which represents the number of measured reflections from theoretically possible reflections, differences of intensity between symmetry equivalent reflections (R_{merge}) and the signal-to-noise ratio ($I/\sigma(I)$) indicate the quality at this point. Scaling was performed with the program XSCALE [294] or SCALA [295] from the CCP4-package [296,297], where symmetry equivalent reflections are merged and scaled. The phase problem was solved by the method of molecular replacement with the structures

of either a Flo5A-based model of Pwp5A or Kremen1 as models for Pwp5A or Wsc1 respectively. The structure of Pwp1A was solved on the basis of the Pwp5A structure. Molecular replacement was performed by the use of the program Phaser [298] in the CCP4-package [296] and the Phenix suite [299]. Refinement was done automatically with the program Refmac5 [300] and manually with the program Coot [301]. The quality of the models was quantified with the R-factor (reliability-factor), which is discriminated between R_{work} and R_{free}. R_{work} indicates the value for the reflections used for refinement, while R_{free} indicates the value for test-reflections, selected during data reduction. The processing and refinement were done with the kind help of Prof. Dr. ESSEN.

5.5.6. Visualization and structure modeling

Protein structures and models were illustrated by the use of the visualization tool PyMOL[302]. The surface hydrophobicity was displayed by the definition of surface colors according to the normalized hydrophobicity consensus scale by EISENBERG [169] in PyMOL. The calculation of electrostatic surface potentials was performed with the Adaptive POISSON-BOLTZMANN Solver (APBS) [303] in PyMOL. Using the crystal structures of Pwp5A and Pwp1A as a template, structural models of the remaining Pwp adhesion domains were generated by the use of the homology-modeling server SWISS-MODEL[179,304]. Further figure labeling and editing were done with Adobe® Illustator.

5.6. Functional analysis *in vivo*

5.6.1. Construction of expression plasmids for *S. cerevisiae*

The yeast strain BY4741 that was used for *in vivo* experiments has an S288c background. It has a *flo8* mutation and is therefore not able to express the flocculation genes *FLO11* and *FLO1* [305]. Four additional *FLO* genes (*FLO1*, *FLO5*, *FLO9* and *FLO10*) are epigenetically silenced [126]. Thus, this strain is not able to flocculate and shows a weak ability of cell-cell sticking [306].

These constructed expression plasmids have a backbone of the *CEN*-plasmid YCplac33 and the expression level of the insert is under control of a *PGK1*-promotor. The plasmid constructs share a secretion signal (SS) of the flocculin $FLO11^{AA1\text{-}30}$ from *S. cerevisiae* Σ1278b, followed by a short spacer sequence (SP) (AGC GGT GGC GGC CGC ATC TTT) and a 3× hemagglutinin tag (3HA). The spacer sequence ensures a correct function of the 3HA tag [109] and the 3HA-tag enables immuno-fluorescence detection. The different *PWPA*-domain sequences from *C. glabrata* CBS138 were inserted into the

expression plasmid by use of *Sac*I and *Sac*II and fused to a *FLO11*$^{(AA214-1360)}$ BC domain sequence from *S. cerevisiae* Σ1278b, which ensures the presentation of the adhesion domains on the surface of the cell wall and therefore enables adhesive interactions.

5.6.2. Adhesion to mammalian cells

The adhesion domain Epa1A is known to mediate adhesion human epithelium [70]. In order to compare the relative adherence of Pwp adhesion domains to different human epithelial cell lines with Epa1A *in vivo*, wash tests were performed. Therefore, confluent monolayers of epithelial- and endothelial cell lines were cultivated on the flat bottom of 96-well plates (5.6.3). Different cancer cell lines served as a model system for human colon (Caco-2), oral (TR146) and vaginal epithelium (A431) as described in 6.16. Porcine aortic endothelial cells (PAOEC) served as a model system for mammalian endothelium. Yeast cells were incubated overnight in selective SC-Ura medium at 30 °C, washed twice and resuspended in 1× PBS. Mammalian cells were also washed twice with 1× PBS and covered with fresh DMEM. Either 10^4, 10^5 or 10^6 yeast cells were transferred into wells with confluent monolayers of epithelial- or endothelial cells and incubated for 2 h at 37 °C and 5 % CO_2. Afterwards, the DMEM together with non-adherend yeast cells was removed and the wells were washed twice with 1× PBS. Remaining cells were fixed with PLP buffer as described by McLean *et al*. [307], washed twice with 1× PBS blocked with Odyssey® Blocking Buffer (*Li-cor Biosciences*) for 1 h and washed twice with 1× PBS again. Fixed yeast cells were labeled with the primary antibody (Rabbit *α*-candida, 1:2000, *Acris* BP1006, 1 h, RT) in PBS-T. After another washing step, the secondary antibody was added to the cells (*Licor* goat *α*-rabbit, 1:1000, CW800, 1 h, RT) in PBS-T. The cells were washed again and stored in 1× PBS at 4 °C till fluorescence measurement in a *Licor*-Reader. Absolute fluorescence was measured in quadruplicates (Excision 778 nm, Emission 795 nm) and the relative fluorescence was determined by subtraction of background fluorescence in wells to which no yeast cells were added.

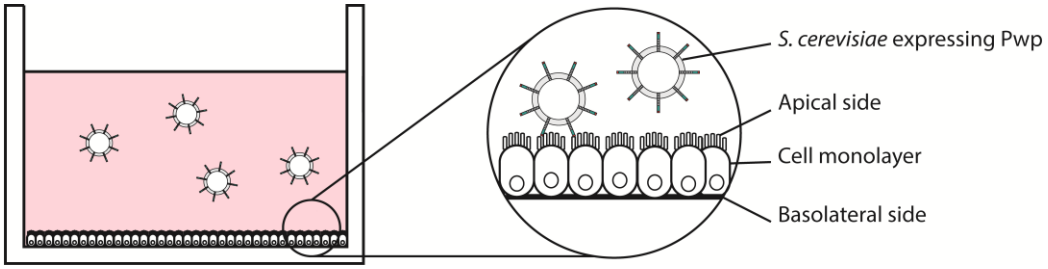

Fig. 68 Schematic illustration of the *in vivo* experimental set-up with mammalian cell monolayers in DMEM and added adhesin-presenting *S. cerevisiae* cells.

5.6.3. Cultivation of mammalian cells

Mammalian cell lines were cultivated in 75 cm² flasks for 72 h (37 °C, 5 % CO_2, DMEM + 10 % FCS, 4.5 g l^{-1} glucose, 1 mM sodium pyruvate, 100 µg ml^{-1} Penicillin, 100 µg ml^{-1} Streptomycin, 0.1 mM β-mercaptoethanol), starting at a concentration of 3×10^4 cells per well until passage. A passage was performed after 2-3 days of incubation upon the formation of approximately 80 % confluent monolayer or a color change of the medium from red to red-orange. All steps were done under a clean bench and with media and solutions that were pre-warmed to 37° C. The old medium was removed and the cell layer was washed with 10 ml PBS, which was in turn removed as well. Then 1 ml of a 0.05 % trypsin solution was disposed on the cells for 1-2 min in the CO_2-incubator in order to detach the cells from the flask. The addition of 9 ml DMEM stopped the trypsin reaction and led to a cell suspension that was then centrifuged at 500 g for 3 min. The supernatant was removed and the cells were resuspended in 1 ml DMEM and the cell density of viable cells was determined with a hemocytometer and trypan blue. The cells were diluted to 1×10^6 cells in 10 ml fresh DMEM and re-incubated.

5.6.4. Adhesion to agar, polystyrene and ECM coatings

In order to investigate the adhesion of Pwp1-7A domains *in vivo* different surface-bound substrates, wash test on agar, polystyrene and different ECM coatings were performed. Agar consists mainly of galactose polymers, agarose and agaropectin, synthesized by red algae. Knowing that Flo11 of *S. cerevisiae* mediates strong agar adhesion, it was used as positive control. *S. cerevisiae* cells, which were presenting Pwp adhesion domains on the cell surface were incubated on SC-Ura-Agar-plates for 5 days at 30°C and then washed off with a constant jet of water, referring to the assay of ROBERTS & FINK [308]. As negative controls for agar adhesion, the expression construct backbone without an adhesion domain (ΔA) and without adhesion and 3×HA-tag (ΔA ΔHA) was used.

Similar to the procedure described in (5.6.2), wash tests were performed to analyze adhesion to polystyrene. Here, specific amounts of Pwp-presenting yeast cells were incubated for 1 h in flat-bottom 48-well plates with 1× PBS. Afterwards, non-adherent cells were washed off with 1× PBS and remaining cells were dyed with 1 % crystal violet in 1×PBS for 15 min and the remaining dye was removed by repeated submerging in dH_2O. The wells were filled with 100 µl 1×PBS and the absorption at 600 nm was measured in a fluorimeter.

5.6.5. Flocculation assay

In *S. cerevisiae*, self-recognition and the formation of protective flocs is mediated by an adhesin family called flocculins (Flo). The flocculins of *S. cerevisiae* are closely related to the Pwp protein family from *C. glabrata* and have a PA14-like architecture of their *N*-terminal effector domain [93]. Flo5A binds calcium-dependent to mannose residues which are present in the cell wall of *S. cerevisiae*. In order to analyze a possible influence of the Pwp family on cell-cell dependent flocculation or binding to other yeast surface structures, a calcium-dependent flocculation test was performed as described previously [93]. The OD^{600} of Pwp-presenting *S. cerevisiae* cultures was measured before and after treatment with 50 mM EDTA.

5.6.6. Immunofluorescence microscopy

In order to ensure the presence of PwpA-domains on the cell surface of *S. cerevisiae* for the adhesion assays, the cells were analyzed by immunofluorescence microscopy. After the transformation of *S. cerevisiae* with the Pwp expression plasmids, they were inoculated in selection medium (SC-Uracil) and incubated overnight at 30 °C. Afterwards, the *PWP*-expressing *S. cerevisiae* cells were transferred into low-fluorescence medium to an OD_{600} of 0.1 and incubated for 4 h at 30 °C until an OD_{600} of 1. Then, 3 ml of the culture was centrifuged (4000 rpm, 3 min) and washed three times with 1×PBS (+1 % BSA). For subsequent labeling, the cells were resuspended in 1×PBS (+1 % BSA) with a primary antibody (mouse α-HA, monoclonal, 1:1000) and incubated for 30 min at RT. After three additional washing steps, the cells were resuspended in 1×PBS (+1 % BSA) with a secondary antibody (goat α-mouse Cy3-conjugated, 1:10000) and incubated for 20 min at RT in the dark. After three additional washing steps and resuspension in 0.5 ml BSA (+1 % BSA), 5 µl of the cell suspension was transferred on a glass slide with a layer of dried 1 % agarose. The presence of the A-domains on the cell surface was detected under a *Zeiss Axiovert 200 M* fluorescence microscope with a rhodamine filter (551 /573 nm).

6. Materials

6.1. Chemicals

Tab. 30 List of chemical compounds used in this study

Compound	Company
Acrylamide (30 %)	*Roth*
Agar agar	*Roth*
Agarose NEEO Ultra	*Roth*
Alexa-Fluor™ 488 5-SDP Ester	*Life Technologies*
Ammonium sulfate	*Roth*
Ampicillin	*AppliChem*
APS	*Roth*
Calcium chloride	*Roth*
CHES	*Roth*
Coomassie-Brilliant-Blue R250	*Roth*
Crystal violet dye	*Sigma-Aldrich*
p-Coumaric acid	*Roth*
DMSO	*Sigma-Aldrich*
DTT	*Roth*
EDTA	*Roth*
Ethidium bromide	*Roth*
Glycerol	*Roth*
Glycin	*Roth*
HEPES	*Roth*
Imidazole	*Roth*
IPTG	*Roth*
Kanamycin sulfate	*Roth*
2-Mercaptoethanol	*Roth*
Luminol	*Sigma*
Magnesium chloride	*Fluka*
MES	*Roth*
PageRuler™ Protein Ladder	*Thermo Fisher*
Paraformaldehyde	*Roth*
PMSF	*Sigma-Aldrich*
Penicillin	*Life Technologies*
SDS	*Roth*
Sodium chloride	*Roth*
Sodium hydrogen carbonate	*Roth*
Streptomycin	*AppliChem*
SYPRO™ Orange (5000×)	*Sigma*
TEMED	*Roth*
Tryptone	*Roth*
Tris	*Roth*
Tween®-20	*Roth*
Virkon® S	*Du Pont*
Yeast extract	*Roth*
Yeast nitrogen base	*Becton Dickson*

6.2. Carbohydrates

Tab. 31 Mono-, oligo- and polysaccharides used in this study

Carbohydrate	Abbreviation	Company
D-Glucose	Man	*Sigma-Aldrich*
D-Mannose	Man	*Sigma-Aldrich*
D-Galactose	Gal	*Sigma-Aldrich*
Glucosamine	GlcN	*Sigma-Aldrich*
N-Acetyl-D-glucosamine	GlcNAc	*Sigma-Aldrich*
N-Acetyl-D-galactosamine	GalNAc	*Sigma-Aldrich*
N-Acetyl-D-glucosamine 6-sulfate sodium salt	GlcNAc6S	*Sigma-Aldrich*
N-Acetyl-D-galactosamine 6-sulfate sodium salt	GalNAc6S	*Sigma-Aldrich*
N-Acetyl-D-galactosamine 4-sulfate sodium salt	GalNAc4S	*Sigma-Aldrich*
L-Iduronic acid sodium salt	IdoA	*Cayman Chemical*
D-Galacturonic acid sodium salt	GlcA	*Sigma*
Galactose-β-(1,3)-Glucose	Galβ1-3Glc	*Dextra*
Lactose	Galβ1-4Glc	*Roth*
Dermatan sulfate disaccharide	ΔIdoAβ1-3GalNAc	*Iduron*
Dermatan sulfate disaccharide	ΔIdoAβ1-3GalNAc4S	*Iduron*
Dermatan sulfate disaccharide	ΔIdoA2Sβ1-3GalNAc4S	*Iduron*
Fondaparinux sodium (Heparin-ATIII binding site)	GlcNS6Sα1-4GlcAα1-4 GlcNS3S6Sα1-4IdoA2S α1-4GlcNS6S	*Sigma-Aldrich*
Dalteparin sodium LMW Heparin	~5 kDa Heparin polymer	*Sigma-Aldrich*
HMW porcine heparin	~15 kDa Heparin polymer	*Sigma-Aldrich*

6.3. Technical equipment

Tab. 32 Technical laboratory equipment that was used in this study

Equipment	Company
ÄKTA™ Purifier System (UPC 10)	*GE Healthcare*
Autoclave (LTA 2 × 3 × 4)	*Zirbus*
Cell density meter (Ultrospec™ 10)	*Amersham Biosciences*
Centrifuge (Heraeus® Biofuge® Pico® / Fresco®) / (4K15)	*Thermo Fisher/Sigma*
Chromatography column (XK 16/60, Superdex® 75pg)	*GE Healthcare*
Crystallization robot (Cartesian MicroSys™ SQ 4000 / HoneyBee 963™)	*Genomic Solutions*
Digital camera (PowerShot A620)	*Canon*
Documentation of Crystallization (RockImager™ 1000)	*Formulatrix*
Documentation of DNA Gels (GelDoc® XR)	*Bio-Rad*
Fluorescence microscope (Axiovert 200M)	*Zeiss*
Fluorimeter (Synergy™ MX)	*Biotek*
French Pressure Cell Press	*American Instrument*
Homogenizer (FastPrep®-24)	*MP Biomedicals*
ITC-System (MicroCal™ iTC_{200})	*Malvern*
Microscope (Axioskop 40)	*Zeiss*
Real Time PCR Cycler (Rotor-Gene® Q)	*Qiagen*
PCR Thermocycler (Primus® 25 advanced)	*Peqlab*
Peristaltic pump (P1)	*Pharmacia Biotech*
pH meter (S20 SevenEasy™)	*Mettler Toledo*
Power supply (PowerPac™)	*Bio-Rad*
Scales (BL3100, CP124S)	*Sartorius*
Spectrophotometer (NanoDrop® ND-1000)	*Thermo Fisher*
Spectropolarimeter J-810	*Jasco*
Thermal Incubator (ThermoMixer® C)	*Eppendorf*
Transilluminator 312/365 nm (IL-20-M/L)	*H. Saur Laborbedarf*
Ultracentrifuge (L7-65)	*Beckmann Coulter*
Vacuum pump (MZ 2C)	*Vacuubrand*

6.4. Media

Tab. 33 Purchased media for the cultivation of mammalian cells.

Medium	Company
Gibco® DMEM/F12 (1×) [+] L-Glutamine	*Thermo Fisher*
Gibco® DMEM (1×) [+] L-Glutamine	*Thermo Fisher*
Gibco® FCS	*Thermo Fisher*
Gibco® PBS (1×) pH 7.4	*Thermo Fisher*

Tab. 34 Media and additives for the cultivation of *E. coli* & *S. cerevisiae.*

Medium	Composition
LB (lysogeny broth)	10 g l^{-1} Tryptone 10 g l^{-1} NaCl 5 g l^{-1} Yeast extract
TB (terrific broth)	24 g l^{-1} Yeast extract 12 g l^{-1} Tryptone 4 ml Glycerol 100 ml Phosphate buffer (10×)
YPD (yeast extract peptone dextrose)	20 g l^{-1} Tryptone 20 g l^{-1} Glucose 10 g l^{-1} Yeast extract
SC (synthetic complete)	20 g l^{-1} Glucose 5 g l^{-1} $(NH_4)_2SO_4$ 2 g l^{-1} Aminoacid mix 1.5 g l^{-1} Yeast Nitrogen Base w/o Amino Acids & $(NH_4)_2SO_4$
Phosphate buffer (10×)	0.17 M KH_2PO_4 0.72 M $K_2HPO_4 \cdot 3\ H_2O$
LFM (low fluorescence medium)	2 g l^{-1} Aminoacid mix 1/10 Salt stock solution 1/10000 Trace element solution 1/10000 Vitamin stock solution 2 % Glucose
Salt stock solution (10×)	50 g l^{-1} $(NH_4)_2SO_4$ 10 g l^{-1} KH_2PO_4 5 g l^{-1} MgSO4 1 g l^{-1} NaCl 1 g l^{-1} $CaCl_2$
Trace element solution	0.5 g l^{-1} H_3BO_4 0.4 g l^{-1} $MnSO_4$ 0.4 g l^{-1} $ZnSO_4$ 0.2 g l^{-1} $FeCl_3$ 0.2 g l^{-1} Na_2MoO_4 0.1 g l^{-1} KI 40 mg l^{-1} $CuSO_4$
Vitamin stock solution	2 g l^{-1} Inositol 0.4 g l^{-1} Calcium pantothenic acid 0.4 g l^{-1} Niacin 0.4 g l^{-1} Pyridoxine · HCl 0.4 g l^{-1} Thiamine · HCl 0.2 g l^{-1} *p*-Aminobenzoic acid 2 mg l^{-1} Biotin

6.5. Enzymes

Tab. 35 Enzymes used in this study

Enzyme	Company
*Bam*HI	*New England Biolabs*
BSA	*Roth*
DNAse I	*AppliChem*
Lysozyme	*Roth*
*Nde*I	*New England Biolabs*
*Pau*I	*New England Biolabs*
Phusion® DNA polymerase	*New England Biolabs*
T4-DNA-Ligase	*Roche*
Taq DNA polymerase	*New England Biolabs*
Thrombin from bovine plasma	*Sigma-Aldrich*
Trypsin-EDTA 0.05 % (1×)	*Life Technologies*
*Xho*I	*New England Biolabs*

6.6. Stock-Solutions

Tab. 36 Chemical stock solutions and appropriate concentrations in H_2O

Compound	Concentration
Ampicillin	100 mg ml^{-1}
Kanamycin	50 mg ml^{-1}
IPTG	1 M
PMSF	0.1 M in isopropanol
p-Coumaric acid	90 mM
Luminol	250 mM
Lysozyme	12 mg ml^{-1}
APS	10 %

6.7. Gels

Tab. 37 Composition of gels used for the separation of proteins or DNA

Gel	Composition
Agarose gel	1-2 % (w/v) Agarose in 1×TAE or 1×TBE
SDS-PAGE Stacking Gel (4 %)	1.57 ml H_2O, 2.08 ml 30 % Acrylamid, 1.25 ml Stacking Buffer, 50 µl 10 % (w/v) SDS, 50 µl 10 % (w/v) APS, 5 µl TEMED
SDS-PAGE Resolving Gel (12 %)	1.36 ml H_2O, 340 µl 30 % Acrylamid, 250 µl Stacking Buffer, 20 µl 10 % (w/v) SDS, 20 µl 10 % (w/v) APS, 2 µl TEMED

6.8. Consumables

Tab. 38 Purchased materials, kits and solutions

Material	Company
Amicon® Ultra-4, Amicon® Ultra-15 Centrifugal Filter Units	*Merck*
Cellstar® Cell Culture Plate / Tissue Coated 48 well	*Greiner BIOone*
Cellstar® Filter Top Cell Culture Flask 75 cm²	*Greiner BIOone*
Corning® 96 Well Assay Plate Black / Clear Bottom	*Corning*
E.Z.N.A.® Gel Extraction Kit	*Omega Bio-Tec*
E.Z.N.A.® Plasmid DNA Mini Kit I	*Omega Bio-Tec*
GelRed® Nucleic Acid Gel Stain (10.000×)	*Biotium*
GeneRuler™ DNA Ladder Mix	*Thermo Fisher*
HiTrap® Heparin HP Affinity Column (1 ml)	*GE Healthcare*
Millicoat™ Human CollagenI Cell Adhesion Strips 96 Well	*Merck*
Millicoat™ Human CollagenIV Cell Adhesion Strips 96 Well	*Merck*
Millicoat™ Human Laminin Cell Adhesion Strips 96 Well	*Merck*
Millicoat™ Human Fibronectin Cell Adhesion Strips 96 Well	*Merck*
NeXtal Crystallization Screens	*Qiagen*
Odyssey® Blocking Buffer (PBS)	*Li-Cor*
Omnifix® Luer Solo Syringe 1 ml, 5 ml, 50 ml	*B. Braun*
PageRuler™ Prestained Protein Ladder	*Thermo Fisher*
Protino® Ni-NTA Affinity Column 1 ml, 5 ml	*Macherey-Nagel*
RNeasy® Mini Kit	*Qiagen*
Rotilabo® Syringe Filter 0.22 µm	*Roth*
Sephadex™ PD-10 Desalting Columns	*GE Healthcare*
Sterican® Cannula Gr. 17	*B. Braun*
Steritop® Filter Unit 0.22 µm	*Merck*

6.9. Antibodies

Tab. 39 Primary- and secondary antibodies used for immunofluorescence microscopy

Antibody	Dilution	Company
Mouse *α*-HA Monoclonal	1:1000	*Sigma* (H3663)
Mouse *α*-StrepMAB Monoclonal	1:1000	*Iba* (2-1546-050)
Goat *α*-Mouse Cy3-label	1:10000	*Sigma* (C2181)
Goat *α*-Rabbit IgG Peroxidase label	1:10000	*Cayman* (10004301)
Rabbit *α*-Yeast Polyclonal	1:1000	*Acris* (BP1006)
Goat *α*-Rabbit IgG IRDye® 800CW	1:2000	*Licor* (32211)

6.10. Buffers

Tab. 40 Composition of buffers

Buffer	Composition
AM (Lysis)	20 mM Tris/HCl pH 8.0, 200 mM NaCl
AMI (Wash)	AM + 20 mM Imidazole
AMI (Elution)	AM + 250 mM Imidazole
AME	AM + 10 mM EDTA
AML	AM + 50 mM Lactose
Blotting	25 mM Tris/HCl, 192 mM Glycin, 20% (v/v) Methanol
Carbonate	100 mM $NaHCO_3$ pH 8.3
CFG buffer	20 mM Tris/HCl pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.05% (v/v) Tween 20, 1 %BSA (w/v)
Coomassie-Staining	0.025% (w/v) Coomassie-Brilliant-Blue R250, 10% (v/v) Acetic acid, 50% (v/v) Ethanol
Coomassie-Destaining	10% (v/v) Acetic acid, 50% (v/v) Ethanol
DNA-Loading Dye (6×)	10 mM Tris/HCl pH 7.6, 60% (v/v) Glycerol 60 mM EDTA, 0.03% (w/v) Bromphenol blue
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4 \cdot 2\ H_2O$, 1.8 mM KH_2PO_4, pH 7.4 (HCl), ddH_2O
PBST	PBS + 0.1% (v/v) Tween 20
PEG	10 mM Tris/HCl pH 8.0, 100 mM Lithium acetate, 1 mM EDTA, 40% (w/v) PEG 4000
PLP	PBS + 2% paraformaldehyde (w/v), 75 mM lysine, 10 mM $NaIO_4$
SDS Loading Dye (2×)	100 mM Tris/HCl pH 6.8, 200 mM DTT, 0.2% (w/v) Bromphenol blue, 20% (v/v) Glycerol, 4% (w/v) SDS
SDS-PAGE-Resolving	1.5 M Tris/HCl pH 8.8
SDS-PAGE-Running	25 mM Tris, 200 mM Glycin, 0.1% (w/v) SDS
SDS-PAGE-Stacking	0.5 M Tris/HCl pH 6.8
SORB	10 mM Tris/HCl pH 8.0, 100 mM Lithium acetate, 1 mM EDTA, 1 M Sorbitol
TAE	40 mM Tris, 20 mM Acetic acid, 1 mM EDTA
TBE	89 mM Tris, 89 mM Boric acid, 2 mM EDTA
TBS (10×)	100 mM Tris/HCl pH 7.5, 1.5 M NaCl
TBST	1× TBS, 0.1% (v/V) Tween 20
TE	10 mM Tris/HCl pH 8.0, 1 mM EDTA
TFBI	100 mM RbCl, 50 mM $MnCl_2$, 30 mM Potassium acetate, 10 mM $CaCl_2$, 15% (v/v) Glycerol, pH 5.8 (Acetate)
TFBII	10 mM MOPS, 10 mM RbCl, 75 mM CaCl2, 15% (v/v) Glycerol, pH 6.5 (KOH)
SEC	20 mM MES pH 6.0 (NaOH), 50 mM NaCl, 10 mM $CaCl_2$

6.11. Primers

Oligonucleotide synthesis and PCR-product sequencing were performed by the *Myrosynth Seqlab AG*.

Tab. 41 Primers that were used for the construction of yeast expression plasmids, including *Sac*II/*Sac*I restriction sites

Primer	Sequence 5'>3' (*Sac*II/*Sac*I)	Source
Pwp1A-fw	AAA**CCGCGG**GCCAACCCGATCAGTATAAC	This work
Pwp1A-rv	AAA**GAGCTC**GGTGTACGTACATGTTG	This work
Pwp2A-fw	AAA**CCGCGG**AAGAACCCAGTTGAATTAAAC	This work
Pwp2A-rv	AAA**GAGCTC**CGAATAAGTACAGGTCG	This work
Pwp3A-fw	AAA**CCGCGG**AAGAATCCAGTTGATTTTAC	This work
Pwp3A-rv	AAA**GAGCTC**CTTATAAGTGCAGGTAG	This work
Pwp4A-fw	AAA**CCGCGG**GTTGGCAATCCAATAACAG	This work
Pwp4A-rv	AAA**GAGCTC**TACATAAGAACATGTATC	This work
Pwp5A-fw	AAA**CCGCGG**GCAAATCCAGTACCTTTG	This work
Pwp5A-rv	AAA**GAGCTC**AGTGTATGTGCATGTGGAAC	This work
Pwp6A-fw	AAA**CCGCGG**GCCAATTTACCAAAGCTTAC	This work
Pwp6A-rv	AAA**GAGCTC**TTTATAGTCACAGTCTGTG	This work
Pwp7A-fw	AAA**CCGCGG**AAGAATCCGGTAGACTTTC	This work
Pwp7A-rv	AAA**GAGCTC**CGAATATTTACAGTTGTC	This work
Als3A-fw	**CCGCGG**AAGACAATCACTGGTGTTTTC	This work
Als3A-rv	**GAGCTC**AGTTTCCCCAATTGGTGGAG	This work

Tab. 42 Primers that were used for the construction of yeast expression plasmids with *Bam*HI/*Pau*I restriction sites.

Primer	Sequence 5'>3' (*Bam*HI/*Pau*I)	Source
Pwp1A-fw	**GGATCC**GCCAACCCGATCAGTATAAC	This work
Pwp1A-rv	**GCGCGC**GGTGTACGTACATGTTGAACAG	This work
Pwp2A-fw	**GGATCC**AAGAACCCAGTTGAATTAAACAATG	This work
Pwp2A-rv	**GCGCGC**CGAATAAGTACAGGTCGAACAG	This work
Pwp3A-fw	**GGATCC**AAGAATCCAGTTGATTTTACTACAG	This work
Pwp3A-rv	**GCGCGC**CTTATAAGTGCAGGTAGAACATTG	This work
Pwp4A-fw	**GGATCC**TACATAAGAACATGTATCACACG	This work
Pwp4A-rv	**GCGCGC**CTTATAAGTGCAGGTAGAACATTG	This work
Pwp5A-fw	**GGATCC**GCAAATCCAGTACCTTTGACC	This work
Pwp5A-rv	**GCGCGC**AGTGTATGTGCATGTGGAACATG	This work
Pwp6A-fw	**GGATCC**GCCAATTTACCAAAGCTTACAAAC	This work
Pwp6A-rv	**GCGCGC**TTTATAGTCACAGTCTGTGCATG	This work
Pwp7A-fw	**GGATCC**AAGAATCCGGTAGACTTTCCC	This work
Pwp7A-rv	**GCG CGC**CGAATATTTACAGTTGTCACAGC	This work

Tab. 43 Mutagenesis primers for the exchange of the D*cis*D region on CBL1 to Alanine.

Primer	Sequence 5'>3'	Source
Pwp1A-fw	CTGCAGGGCTCAGCTGCTGCCGCCATTCTG	This work
Pwp1A-rv	CAGAATGGCGGCAGCAGCTGAGCCCTGCAG	This work
Pwp2A-fw	CTCTTTCAGACTCAGCTGCTGCTGCAATCATCTTT CTTGG	This work
Pwp2A-rv	CCAAGAAAGATGATTGCAGCAGCAGCTGAGTCTG AAAGAG	This work
Pwp5A-fw	CATTAGAAAATACAGCTGCTGCCGCTATCATCTTC TTCGGC	This work
Pwp5A-rv	GCCGAAGAAGATGATAGCGGCAGCAGCTGTATTTTC TAATG	This work
Pwp7A-fw	CTTTGCTTTATATTTGACCGCTGCTGGTAGTGCCAT GTTTTTAGG	This work
Pwp7A-rv	CCTAAAAACATGGCACTACCAGCAGCGGTCAAA TATAAAGCAAAG	This work

Tab. 44 Constructed qPCR primers for the analysis of *PWP* expression levels in *C. glabrata* CBS138.

Primer	Sequence 5'>3'	Source
Pwp1-fw	GTACGGCTTAACAGTGTGGC	This work
Pwp1-rv	TGAGCCCTGCAGAGTGAAC	This work
Pwp2-fw	ATGGGATTGGAGCAGGTAGG	This work
Pwp2-rv	AGGGGTCGTACTGCCAAATG	This work
Pwp3-fw	ACAGGTTCCGATGATGCTGC	This work
Pwp3-rv	CCGTGATATCATTGGCAGTGG	This work
Pwp4-fw	AGGGTGTTGCTTATCCTATCCG	This work
Pwp4-rv	TCGACTCTCTTTACACCTTCGG	This work
Pwp5-fw	ACTGCCTCTACCAACGAACG	This work
Pwp5-rv	TATTGTCCGGAAACAGGCGG	This work
Pwp6-fw	TGTCCCAACTGAAACTGGTG	This work
Pwp6-rv	CACAAGGGAAAGCAGATGG	This work
Pwp7-fw	CGGTTTATATGGTCGCAGGTG	This work
Pwp7-rv	CTGGGTCAACAAAGGAAGCAG	This work
Epa1A-fw	CAAGCGCAGATCATTTCGG	This work
Epa1A-rv	GGTAATATACGCCCGCATC	This work
Epa6A-fw	CGTTTACAATTTCGGCCGAC	This work
Epa6A-rv	TACCAAAGTCATCAGCGCTG	This work
Act1-fw	GCGTTACCCAATCGAACACG	This work
Act1-rv	GTTCTTCTGGGGCGACTCTC	This work
Pgk1-fw	TTGTTCGCTGCAACTATCGC	This work
Pgk1-rv	ATAGCCTTGGTACCGGAAGC	This work

Tab. 45 Sequencing primers

Primer	Sequence 5'>3'	Source
PGK-Seq-Primer	GGGGGTGGTTTAGTTTAGTAGAACCTCG	Julia van der Felden
T7-fw	TAATACGACTCACTATAGGG	Rike Diderrich
M13-rv	CAGGAAACAGCTATGAC	Rike Diderrich

6.12. Plasmids

Tab. 46 *E. coli* expression plasmids.

Name	Construct	Source
pET-28(a)$^+$	*6×His lacI KanR*	*Merck*
BHUM1829	*EPA1*$^{(31\text{-}271)}$ in pET-28(a)$^+$	Maestre-R. 2012
BHUM2824	*PWP1*$^{(25\text{-}227)}$ in pET-28(a)$^+$	This study
BHUM2825	*PWP2*$^{(19\text{-}220)}$ in pET-28(a)$^+$	This study
BHUM2826	*PWP3*$^{(24\text{-}225)}$ in pET-28(a)$^+$	This study
BHUM2827	*PWP4*$^{(41\text{-}241)}$ in pET-28(a)$^+$	This study
BHUM2828	*PWP5*$^{(25\text{-}226)}$ in pET-28(a)$^+$	This study
BHUM2829	*PWP5*$^{(25\text{-}226;N131D)}$ in pET-28(a)$^+$	This study
BHUM2830	*PWP6*$^{(25\text{-}226)}$ in pET-28(a)$^+$	This study
BHUM2831	*PWP7*$^{(22\text{-}224)}$ in pET-28(a)$^+$	This study
BHUM3120	*WSC1*$^{(AA22\text{-}118)}$ in pET-28(a)$^+$	This study
BHUM3121	*WSC2*$^{(AA24\text{-}118)}$ in pET-28(a)$^+$	This study
BHUM3122	*WSC3*$^{(AA40\text{-}133)}$ in pET-28(a)$^+$	This study

Tab. 47 Temporary plasmids for the construction of yeast expression plasmids.

Name	Construct	Source
pJET1.2	*P_{T7} Eco47I AmpR*	*Thermo Fisher*
BHUM3058	*PWP2*$^{(19\text{-}220)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3059	*PWP3*$^{(24\text{-}225)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3060	*PWP4*$^{(41\text{-}241)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3061	*PWP6*$^{(25\text{-}226)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3066	*PWP1*$^{(25\text{-}227)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3067	*PWP5*$^{(25\text{-}226)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study
BHUM3068	*PWP7*$^{(22\text{-}224)}$ in pJET1.2 (*Bam*HI/*Pau*I)	This study

Tab. 48 Yeast expression plasmids.

Name	Construct	Source
B2445	YCplac33	Gietz1983
BHUM1964	P_{PGK1}-*FLO11*$^{(1-25)}$-*FLO11*$^{(214-1360)}$-T_{FLO11}	Diderrich
BHUM1983	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-EPA1*$^{(31–271)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	Diderrich
BHUM2157	P_{PGK1}-*FLO11*$^{(1-30)}$-*SP-3HA-FLO11*$^{(214-1360)}$-T_{FLO11}	Diderrich
BHUM2158	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-FLO11*$^{(31–257)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	Diderrich
BHUM3045	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP1*$^{(25-227)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3046	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP2*$^{(19-220)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3047	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP3*$^{(24-225)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3048	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP4*$^{(41-241)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3049	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP5*$^{(25-226)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3050	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP6*$^{(25-226)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM3051	P_{PGK1}-*FLO11*$^{(1–30)}$-*SP-3HA-PWP7*$^{(22-224)}$-*FLO11*$^{(214–1360)}$-T_{FLO11}	This study
BHUM2422	P_{PGK1}-*FLO11*$^{(1–30)}$-*FLO11*$^{(31–257)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	Kraushaar
BHUM3054	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-FLO11*$^{(214–1360)}T_{FLO11}$	Scheffer
BHUM3056	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-FLO11*$^{(31–257)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	Scheffer
BHUM3074	P_{PGK1}-*FLO11*$^{(1–30)}$-*PWP1*$^{(25-227)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3075	P_{PGK1}-*FLO11*$^{(1–30)}$-*PWP2*$^{(19-220)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3076	P_{PGK1}-*FLO11*$^{(1–30)}$-*PWP3*$^{(24-225)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3077	P_{PGK1}-*FLO11*$^{(1–30)}$-*PWP5*$^{(25-226)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3078	P_{PGK1}-*FLO11*$^{(1–30)}$- *PWP7*$^{(22-224)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3079	P_{PGK1}-*FLO11*$^{(1–30)}$- *EPA3*$^{(82-266)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3080	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP1*$^{(25-227)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3081	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP2*$^{(19-220)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3082	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP3*$^{(24-225)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3083	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP5*$^{(25-226)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3084	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP7*$^{(22-224)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3085	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-EPA3*$^{(82-266)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3086	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII-PWP1*$^{(25-227;D131A;D132A)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3087	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII- PWP2*$^{(19-220;D125A;D126A)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3088	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII- PWP5*$^{(25-226;N131A;D132A)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study
BHUM3089	P_{PGK1}-*FLO11*$^{(1–30)}$-*STREPII- PWP7*$^{(22-224;D127A;D128A)}$-*FLO11*$^{(214–1360)}T_{FLO11}$	This study

6.13. Vectors

6.13.1. pET-28(a)$^+$

The pET-28(a)$^+$ vector was used for the heterologous production of recombinant proteins. In combination with a compatible *Escherichia coli* strain that carries a genomic copy of the viral T7-RNA-Polymerase, the pET-28(a)$^+$ vector enables the transcription of a recombinant gene[309]. Therefore, the gene is cloned into a multiple cloning site (MSC) on the vector. The MSC is a segment, which enables the insertion of DNA by a variety of restriction sites. The pET-28(a)$^+$ additionally provides the fusion to an *N*- or *C*-terminal hexa histidine-tag (His6×) for affinity purification. Fusion proteins in this study were produced with an *N*-terminal His6× tag. The expression of the recombinant gene and the polymerase is under control of the LacI repressor which gene lies on the vector as well. Binding of the repressor to the *lac* operator sequence in the region of the *lac*- and T7-promoter represses the binding of the polymerase. The inducer isopropyl-*β*-D-thiogalactopyranoside (IPTG) binds to the LacI repressor and releases it from the operator in an allosteric manner[310], which in turn allows the induction of the expression of the recombinant gene. A kanamycin resistance on the vector (Kan^R) enables the positive selection for transformation.

6.13.2. YCplac33

The shuttle vector YCplac33 was used as a backbone for the construction of expression plasmids for *S. cerevisiae*. For cloning in *E. coli*, this vector carries an appropriate origin of replication (ORI) and an ampicillin resistance gene for positive selection. In order to ensure a proper function and replication in *S. cerevisiae*, the vector has an autonomously replicating sequence (ARS) and a yeast-centromer as well as the URA3 gene as a selection marker for auxotrophic complementation. Beside a high transformation efficiency, it shows mitotic stability and low copy numbers[311].

6.13.3. pJET1.2/blunt

The pJET1.2 vector was used for the cloning of blunt-end PCR products. The vector contains a lethal restriction enzyme gene (*eco47IR*) that is disrupted by ligation of a DNA insert into the cloning site. As a result, only bacterial cells with recombinant plasmids are able to form colonies. Recircularized pJET1.2/blunt vector molecules lacking an insert express a lethal restriction enzyme, which kills the host *E. coli* cell after transformation and therefore allows positive selection. A kanamycin resistance on the pJET1.2 ensures retaining of the vector.

6.14. Bacterial strains

6.14.1. *E. coli* TOP10

F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

The *E. coli* strain TOP10 originates from the company *Invitrogen*. It has a high rate of plasmid replication and was used for cloning, plasmid preparation and the transformation of ligations. Chemically competent cells have a high efficiency of transformation and a high rate of plasmid replication. They have a K12 strain background and are thus attenuated.

6.14.2. *E. coli* Shuffle® T7 Express

F′ *lac*, *pro*, *lacI*^{q} / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA) PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* (Spec^{R}, *lacI*^{q}) *ΔtrxB rpsL150*(Str^{R}) *Δgor Δ(malF)3*

The *E. coli* strain Shuffle® T7 Express was used for heterologous protein production. The strain was engineered by the company *NEB* and is optimized to promote disulfide bond formation in the cytoplasm. It constitutively expresses a chromosomal copy of the disufide bond isomerase DsbC which promotes the correction of mis-oxidized proteins into their correct form and is also a chaperone that assists in the folding of proteins[312,313]. The deletion of a glutathione reductase (*Δgor*) and thioreduktase (*ΔtrxB*) also promote the formation of stable disulfide bond through an oxidative environment in the cytoplasm[312]. The resulting lethality was solved by the insertion of a peroxidase (*ahpC*). A chromosomal copy of the T7 RNA polymerase allows the use as an expression system for pET-vectors.

6.14.3. *E. coli* BL21 Gold (DE3)

F– *ompT hsdS(rB– mB–) dcm+ TetR gal lambda (DE3) endA Hte*

This *E. coli* strain was used for test expressions. It is protease-deficient and has deletions of the proteases Lon und OmpT. The Hte phenotype increases the transformation efficiency. DE3 indicates that the host is a lysogen of λDE3, and therefore carries a chromosomal copy of the T7 RNA polymerase gene under control of the *lacUV5* promoter.

6.14.4. *E. coli* Origami 2 (DE3)

Δ(ara-leu)7697 ΔlacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsL F'[lac+ lacIq pro] (DE3) gor522::Tn10 trxB (StrR, TetR)

This *E. coli* strain is a derivate of the K-12 laboratory strain and has mutations in the reductases t*rxB* und *gor* which promotes the formation of stable disulfide bonds through an oxidative environment in the cytoplasm[312]. A chromosomal copy of the T7 RNA polymerase allows the use as an expression system for pET-vectors.

6.14.5. *E. coli* Rosetta™ (DE3)

F^- *ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) $[malB^+]_{K\text{-}12}(\lambda^S)$

The Rosetta strain is a BL21 derivative designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. Beside a chromosomal copy of the T7-RNA-Polymerase, this strain bears the plasmid pRARE, which harbors the genes of tRNAs for codons that are rare in *E. coli.*

6.15. Yeast strains

6.15.1. *S. cerevisiae* YHUM0719

S288C MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0

Former *EUROSCARF* strain BY4741[314], originally isolated from human feces. This haploid strain was transformed with yeast expression-plasmids and was used as a platform for *in vivo* adhesion studies. It is a non-adherent strain due to its S288C background. The cells have a mutation in the transcription factor *FLO8*, which leads to the absence of the flocculin Flo11 and thereby prevents cell-cell sticking. This allows the quantification of adherence by the expression of different PwpA domains or other lectins on the yeast cell surface.

6.15.2. *C. glabrata* CBS138

Wild-type

The genomic DNA of this haploid *Candida glabrata* strain served as a template for the amplification of the seven *PWPA* coding regions. The amplicons were used for the construction of *S. cerevisiae*-expression plasmids. It was kindly donated by Dr. MICHAEL WEIG (*Institut für Medizinische Mikrobiologie*, Göttingen).

6.15.3. *C. albicans* CAI4

ura3::imm434/*ura3*::imm434 *iro1/iro1*::imm434

Uridine auxotroph and haploid strain, constructed by the deletion of the second copy of *URA3*. This strain is avirulent in a mouse model of systemic infection unless complemented with *URA3*. This strain was kindly donated by Dr. MICHAEL WEIG (*Institut für Medizinische Mikrobiologie*, Göttingen).

6.16. Mammalian cells

6.16.1. Human oral epithelial cells TR146

The TR146 cell line was used as human model for oral epithelium in the *in vivo* adhesion experiments. It was derived from the neck node as a primary tumor in buccal mucosa of a 67 year-old female[315] in 1985. The cells show adherent growth and build a confluent monolayer while keeping their morphological and physiological properties. The capability to build a monolayer is a key factor for the *in vivo* adhesion experiments in this work. TR146 cells have established as a standard model for human oral mucosa in biomedical research[316,317] and in *Candida* infection studies[318].

6.16.2. Human vaginal epithelial cells A431

A431 is another human model cell line that was derived from the epidermoid carcinoma of an 85 year-old female[319] in 1973. The cells are hypertriploid. They show adherent growth and build a confluent monolayer while keeping their morphological and physiological properties. A431 cells have established as a standard model for human vulvovaginal mucosa especially in adhesion studies[320-322] and in *Candida* infection studies[318,320].

6.16.3. Human colorectal epithelial cells Caco-2

The Caco-2 cell line was isolated in 1977 from a colorectal adenocarcinoma of a 72 year old male caucasian[323]. The cells are also hypertriploid and show adherent growth while building a polarized confluent monolayer. Caco-2 cells have established as a standard model for human intestine mucosa in biomedical research and in *Candida* infection studies[318]. Thus, after extensive use over the years of use in different laboratories around the world, the characteristics of the cells have diverged significantly[324].

6.16.4. Porcine aortic endothelial cells PAOEC

Primary PAOEC cells are isolated from porcine aortas and provide a frequently used model for studies on cardiovascular function and disease or a model for endothelium in biomedical research. They also grow adherent as a confluent monolayer. This cell line was donated by Dr. SAUERHERING from the workgroup of Prof. Dr. MAISNER at the *Biomedical Research Center* (BMFZ) Marburg.

Epithelial cell lines were donated by Dr. BRAUNSDORF from the workgroup of Prof. Dr. SCHALLER at the *Universitäts-Hautklinik* Tübingen.

7. Supplements

7.1. Sequences of produced proteins

Pwp1A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MANPISITNV CQFPDNWART AGFQATTVRL GLTIASNTYK

        70         80         90        100        110        120
ELLNYPGLFN FATQKATVTQ PNFSISGFPF FTGLYGLTVW PINLLVELQG YFVPPESGNY

       130        140        150        160        170        180
TFTLQGSDDA AILFFANPST FACGSQNDWP QPREADITVT DTLVPSRTLY LVKGVAYPIR

       190        200        210        220
IAYYNGSGDG KLNVSFKDPS GTTHTDWKGY VWQYSNTCST CTYT
```

Pwp2A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MKNPVELNNV CTFPNDYGRI PGFTVLVGRL GLLISASDYK

        70         80         90        100        110        120
ELLRYPGAFL IATQDTMTTS PNFELTGFPF FTGLYGLTVF PLDLLVELRG YFVPPNSGQY

       130        140        150        160        170        180
TFTLSDSDDA AIIFLGNPSA FPCGDLQNWP KPSEEDITVT DHLFPSRTLY LIEGVAYPMR

       190        200        210        220
IAYYNGIGAG RLNAAFIDPS GTRHTNWEGY IWQYDPCSTC TYS
```

Pwp3A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MKNPVDFTTV CQFPDNFSKK PGFLATTARL GLWITGNSYK

        70         80         90        100        110        120
ELLNYPGTFT LTTQQASTTT PNIQLRGTDL FTTIYGLTVY PLYLLVELQG YFVPTQSGMY

       130        140        150        160        170        180
TFSLTGSDDA AIIFIGNPST FPCDAFKTWP QATANDITVT DSINPSRTLY MVKGVAYPLR

       190        200        210        220
IAYYNGLGNG VLNAAFTDPS GTKRTDWNGY IWQYDQCSTC TYK
```

Pwp4A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MVGNPITVSN VCSFPKNWHR TAGFRALTAR LGIGIQSDSY

        70         80         90        100        110        120
KELLNYPGTF QLTTQQSSTT KPNFQISGLV GNIYGLTIVP SYLLVELTGY FVPPVSGEYT

       130        140        150        160        170        180
FSLEGSDDAA IIFIANPSTF KCGHVEDWPT AYDNDVTVTE SIHATRTIYM IEGVAYPIRI

       190        200        210        220
AYYNGLSLGK LQASFTDPEG VKRVDWNGYI WRYDSCDTCS YV
```

Pwp5A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MANPVPLTNV CYFPEEYPRQ AGFRAITAGL ASTILQGSVD

        70         80         90        100        110        120
ELLNYSGLSS LTTQQTTTTT PNFQLPTLPL PTNVYGLTVV PLYLLVQLQG YFVPPVSGQY

       130        140        150        160        170        180
TFSLENTNDA AIIFFGNPSA FPCGDLSAWP QASADDITVT DGTFPSRTLF LKAGVAYPMR

       190        200        210        220
ISYYNSLGSG NLAVSFTDPG GTKRTNWNGY IWRYNSCSTC TYT
```

Pwp5A^{N131D}

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MANPVPLTNV CYFPEEYPRQ AGFRAITAGL ASTILQGSVD

        70         80         90        100        110        120
ELLNYSGLSS LTTQQTTTTT PNFQLPTLPL PTNVYGLTVV PLYLLVQLQG YFVPPVSGQY

       130        140        150        160        170        180
TFSLENTDDA AIIFFGNPSA FPCGDLSAWP QASADDITVT DGTFPSRTLF LKAGVAYPMR

       190        200        210        220
ISYYNSLGSG NLAVSFTDPG GTKRTNWNGY IWRYNSCSTC TYT
```

Pwp6A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MANLPKLTNI CQIPDNVKKT SGFKGEIRRI GLFPPGGDYT

        70         80         90        100        110        120
YLLKFPSTGS TASDTISLTT KPQLSTSGWF PSLFGHTYFI KADLLVYMVA YFVPTETGEY

       130        140        150        160        170        180
TFKLVNSDDA AVMYVGNPSA FPCGSIDTWP EPRKEDINVC DKISASTKLY MEKGLAYPVR

       190        200        210        220
LVWYNGGGDG QFNAQFIDPS GTTHTDWNGY VWQYDACTDC DYK
```

Pwp7A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MKNPVDFPNL CQYPNLDGAV AGFDTEIYNA DPSILINFSL

        70         80         90        100        110        120
WPILQFPKEW HKTIGRAITT QTNFKLPASD TTLFGVSIPT QNILIEFRGY FIPPKSGTYT

       130        140        150        160        170        180
FALYLTDDGS AMFLGNDAAF PCGNITSWSN PTNDQVSVSD YDPKKKTLTV YMVAGVAYPL

       190        200        210        220
RIAYYNREKG GSLTASFVDP DGLTHSDWNG YIWTLPSCDN CKYS
```

Epa1A

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MTSSNDISLA SKDPTTFPLG CSPDITTPKK GLSMELYSYD

        70         80         90        100        110        120
FRKKGSYPCW DAAYLDPNYP RTGYKSHRLL AKVDGVTGNI NFYYHATKGC TPQLGHLPAS

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

       190        200        210        220        230        240
GNYQAYAIWG SKTAKDELTV HLDAGVYYPI RLFYNNREYD GALSFTFKTE SNENTVSDFS

       250        260
EYFFSLDDTE EGCPGLISYD SS
```

Wsc1

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MYEYVNCFSS LPSDFSKADS YNWQSSSHCN SECSAKGASY

        70         80         90        100        110
FALYNHSECY CGDTNPSGSE STSSSCNTYC FGYSSEMCGG EDAYSVYQLD SDTNSNSI
```

Wsc2

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MDQFTYKACY SASDIRKLGL TYKGVYEYQS VSYCQNECPG

        70         80         90        100        110
QAVVALFNGT GCYCGGSVAQ LQSLTQVDSS KCDVSCAGWP YQNCGGSSAM NVYINN
```

Wsc3

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MFNYEGCYSA ADIQSAGLSL KNSYIYQSVS YCQNQCPESA

        70         80         90        100        110
VVALFNGSDC YCGNSVSFLT SLTKSTDSNC GTKCSGWPYQ MCGGSSYMNV YVNAE
```

7.2. CD-Spectroscopy

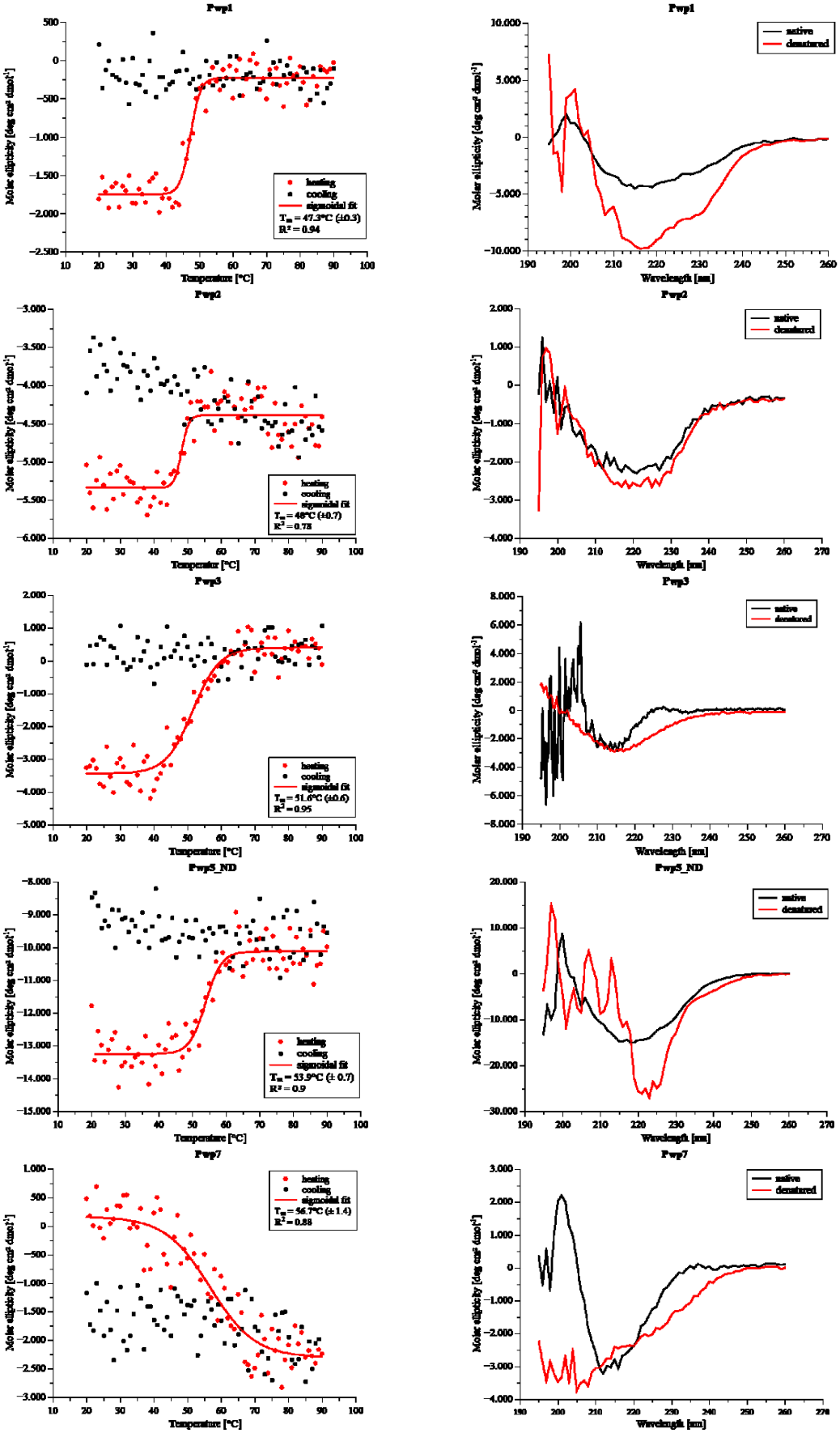

Fig. 69 CD-spectroscopy of PwpA-domains. Far-UV spectra and melting curves.

7.3. *In vitro* analytics

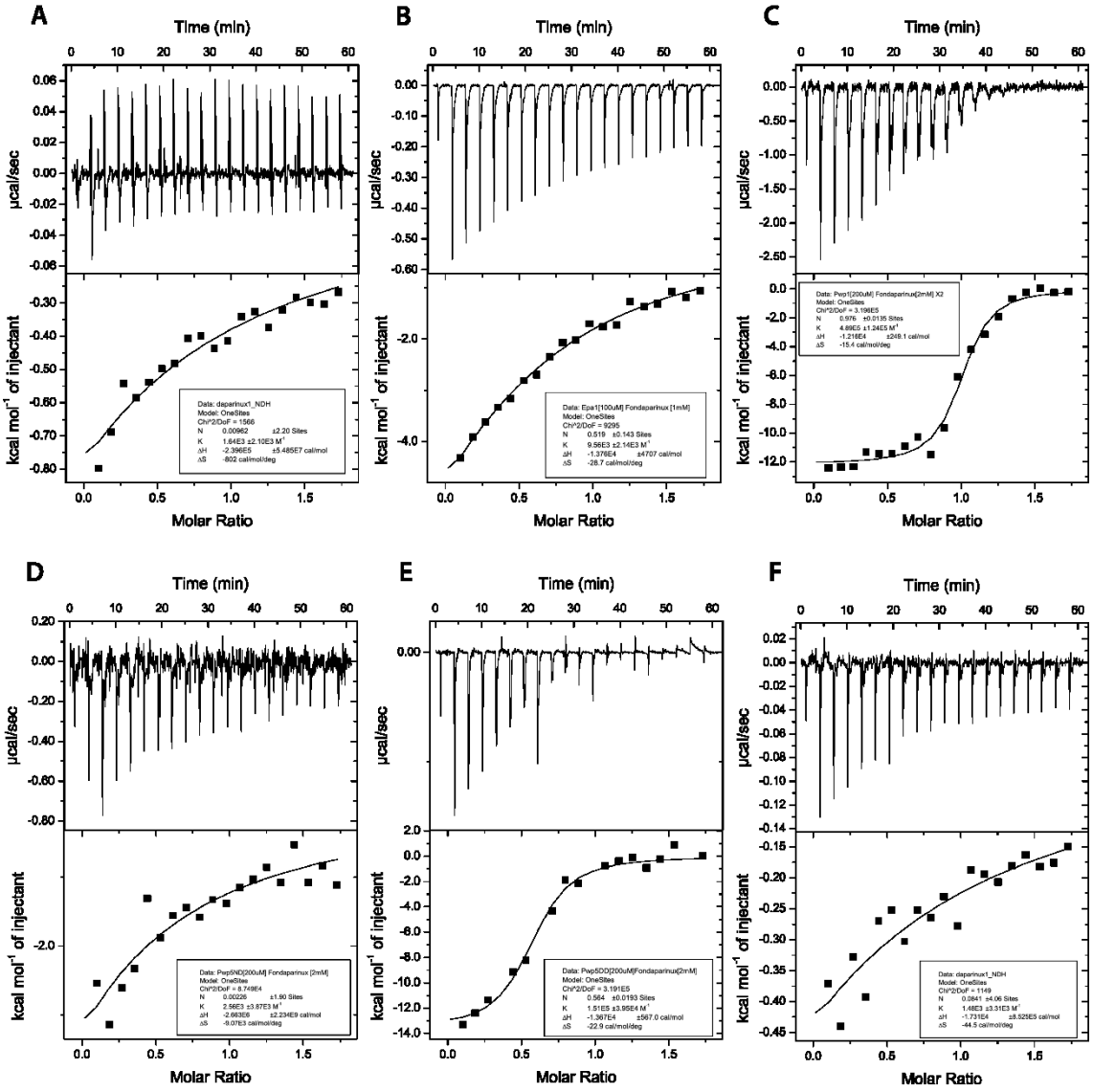

Fig. 70 ITC profiles for the binding of (A) BSA, (B) Epa1A, (C) Pwp1A, (D) Pwp5A, (E) Pwp5^{N131D} and (F) Thrombin with *Fondaparinux*. The upper panels represent the injections of 2 µl 2 mM *Fondaparinux* from the syringe into 200 µl of 200 µM protein sample in SEC buffer at 20°C. Heats of reactions were determined by integration of injection peaks with correction for heats of dilution. The lower panels represent the resulting titration curves of the fit to a single-class binding site model using *MicroCal Origin* software.

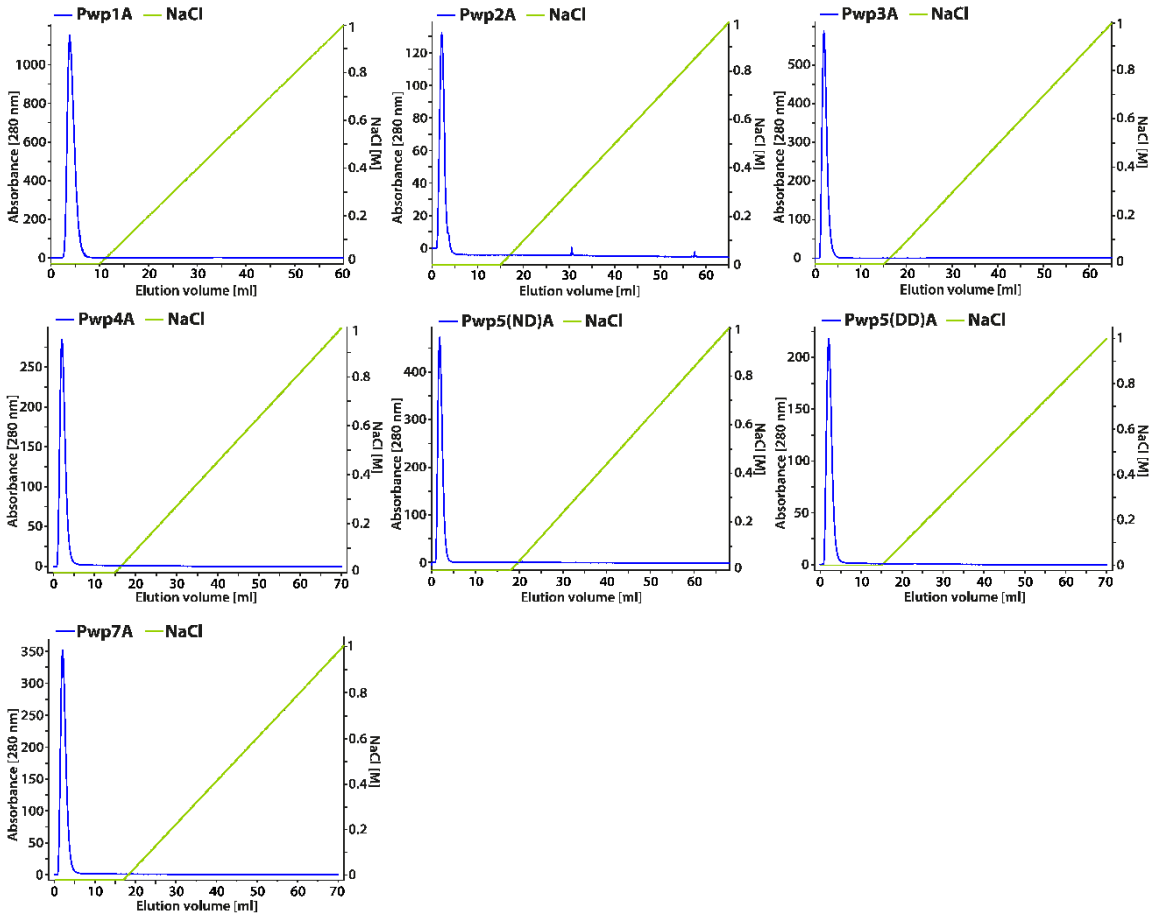

Fig. 71 Absent binding of PwpA domains to a GST column. Elution with a linearly NaCl gradient (0.05 – 1 M NaCl by 50 ml at 1 ml/min).

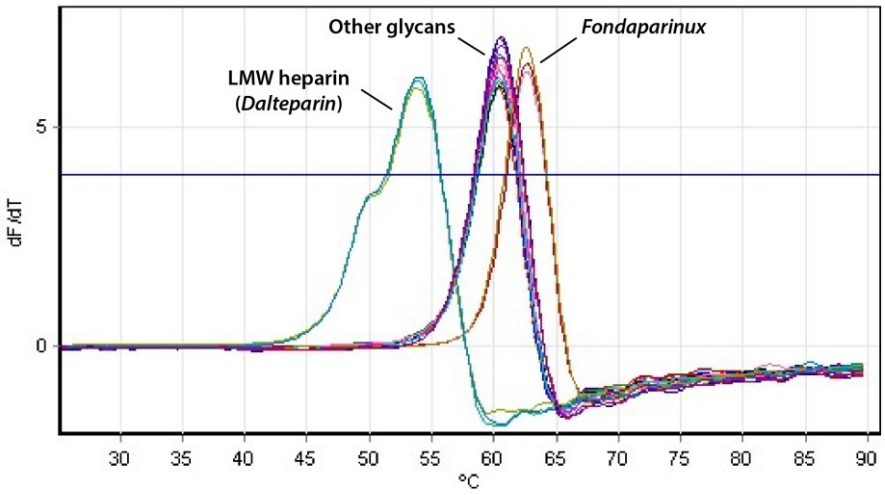

Fig. 72 Thermal shifts of Pwp1A in combination with different glycosaminoglycan compounds. Pwp1A shows a clear positive shift of 2.6 °C with the heparin pentasaccharide *Fondaparinux* and strong negative shift for LMW heparin (*Dalteparin*).

7.4. Crystal packing of Pwp1A, Pwp5A and Wsc1

The structure of Pwp1A was solved in space group $P4_12_12$ (1.85 Å) with two molecules per asymmetric unit (Fig. 73). Due to the truncated chain at the C-terminus of molecule B, molecule A (cyan) is used for further inspections.

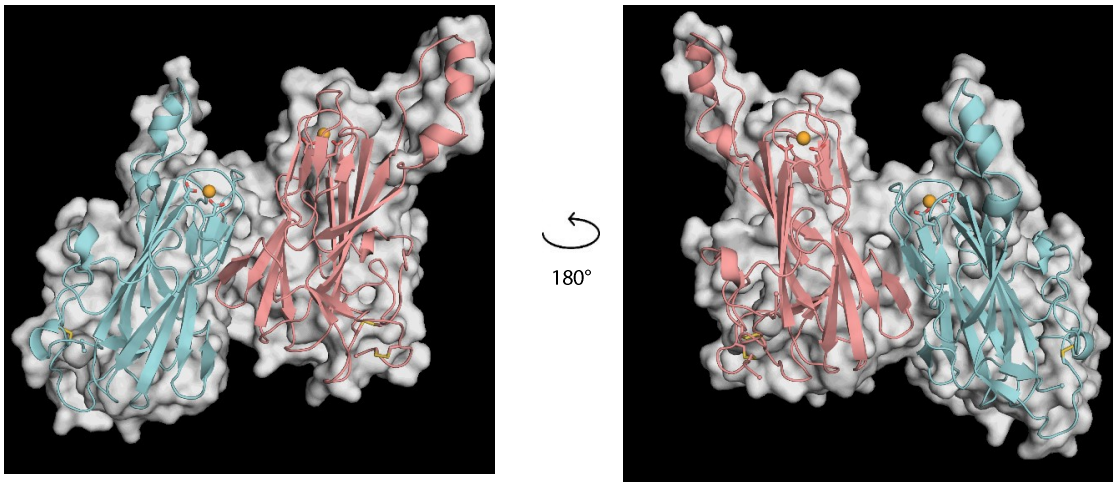

Fig. 73 Orientation of two Pwp1A molecules in the asymmetric unit of the protein crystal.

The structure of Pwp5A was solved in space group $P4_1$ (1.64 Å) also with two molecules per asymmetric unit (Fig. 74). Molecule A (green) is also used for further inspections, due to a truncated C-terminus of molecule B.

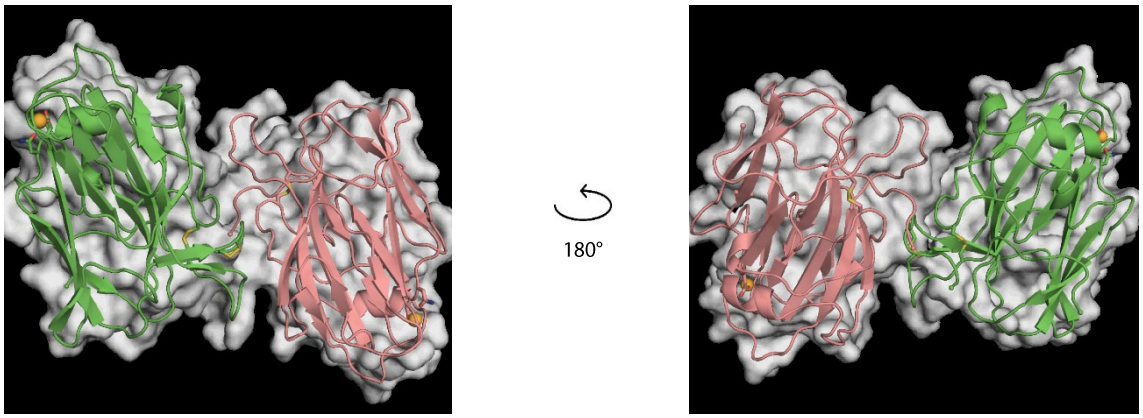

Fig. 74 Orientation of two Pwp5A molecules in the asymmetric unit of the protein crystal.

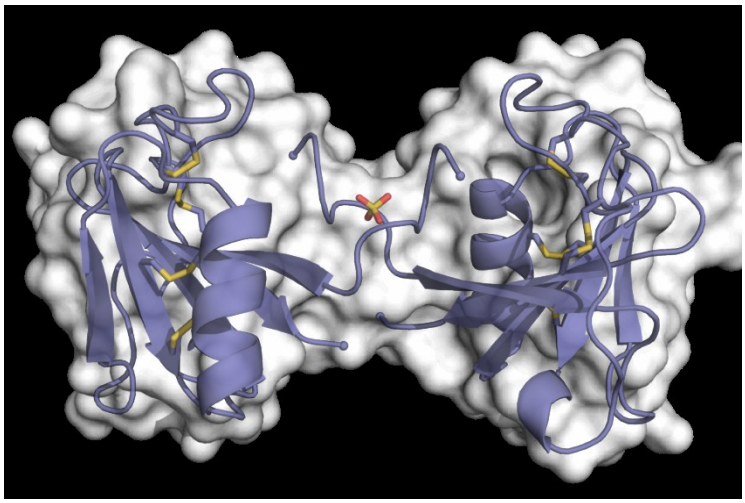

Fig. 75 Orientation of two Wsc1 molecules in the asymmetric unit. The Wsc1 molecules build a dimer by complexation of a sulfate ion via their hexa-histidine tags.

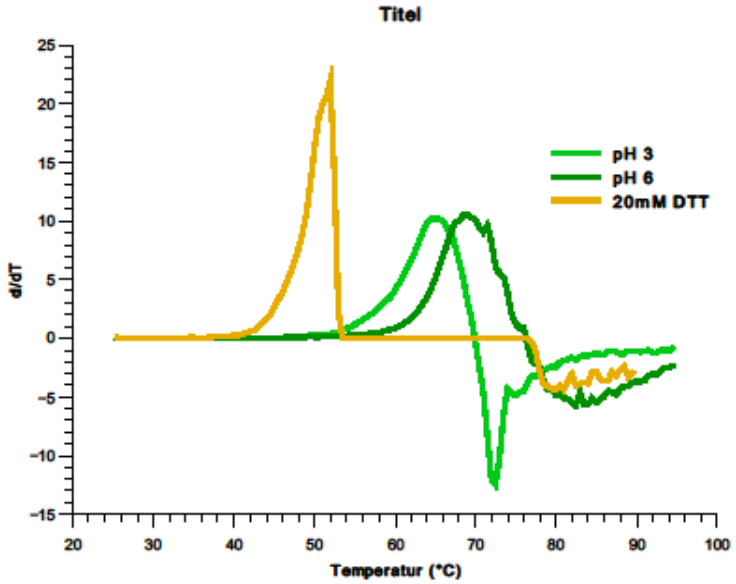

Fig. 76 Thermal shift analysis of the *Sc*Wsc3 CRD. The protein is strongly decreased in thermal stability by DTT whereas large pH shifts do not result in strong decrease of stability.

7.5. Homology modeling

Tab. 49 Parameter for protein homology modeling. Modeling was performed with *SWISS-MODEL* based on the crystal structure of Pwp5A and Wsc1.

Protein	Identity	GMQE	QMEAN
Pwp2A	59.9	0.8	-1.53
Pwp3A	62.69	0.81	-1.89
Pwp4A	59.3	0.8	-2.06
Pwp6A	38.31	0.75	-2.93
Pwp7A	39	0.74	-3.37
Wsc2	35	0.66	-1.51
Wsc3	36	0.7	-1.93

7.6. Glycan arrays

Tab. 50 Synthetic oligosaccharides immobilized on the pathogen glycan ligand microarray from the *MPI für Kolloid- und Grenzflächenforschung* (MPIKG). SPx:Typing serum for pneumococcal serotype x, NaP 8.5: (50 mM sodium phosphate buffer, pH 8.5), PBS: phosphate-buffered saline, Aminolinker2: 6-amino-N-(6-oxo-6-(pentylamino) hexyl) hexanamide.

Glycan ID	Name	Refer-ence	Detected by lectin	Detected by monoclonal antibody	Detected by defined antibody preparation	Signal from n human sera (out of 15)	Number saccharide units	Printed in
5	Neu5Ac(a2-6)Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol	1	SNL, RCA120, UEA, WGA, DBA			>3	5	NaP 8.5
6	Neu5Ac(a2-3)Gal(b1-3)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol	1	WGA, DBA			>3	5	NaP 8.5
7	Fuc(a1-3)[Neu5Ac(a2-3)Gal(b1-4)]GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol	2	WGA, DBA			>3	6	NaP 8.5
8	Neu5Ac(a2-6)Gal(b1-4)Glc(b1-1)aminohexanol	1	RCA120, SNL				3	NaP 8.5
9	Neu5Ac(a2-3)Gal(b1-4)Glc(b1-1)aminohexanol	1	WGA, MALI			1	3	NaP 8.5
10	Neu5Ac(a2-6)Gal(b1-4)GlcNAc-6-sulfate(b1-1)aminohexanol		SNL, ECA, WGA, RCA120, DBA,UEA			>3	3	NaP 8.5
11	Gal(b1-4)Glc(b1-1)aminohexanol		PNA, UEA, ECL, RCA120, MALI			>3	2	NaP 8.5
12	Gal(b1-4)GlcNAc-6-sulfate(b1-1)aminohexanol		ECL, RCA120, WGA, SBA			>3	2	NaP 8.5
69	Araf(a1-5)Araf(a1-1)aminopentanol	3				>3	2	NaP 8.5
70	Araf(a1-5)Araf(a1-3)[Araf(a1-5)Araf(a1-5)]Araf(a1-5)Araf(a1-1)aminopentanol	3	ConA			>3	6	NaP 8.5
71	Araf(a1-3)[Araf(a1-5)]Araf(a1-1)aminopentanol	3				>3	3	NaP 8.5
72	Araf(a1-5)Araf(a1-5)Araf(a1-5)Araf(a1-5)Araf(a1-5)aminopentanol	3				>3	6	NaP 8.5
73	Col(a1-3)[Col(a1-6)]Glc(a1-4)Gal(a1-3)GlcNAc(b1-1)aminopentanol	4				>3	5	NaP 8.5
74	ManNAc(b1-3)FucNAc(a1-3)GalNAc(a1-4)Gal(a1-1)aminopentanol	5		H16, B3		>3	4	NaP 8.5
75	GalNAc(a1-4)Gal(a1-1)aminopentanol	5	BSL, SBA, SNL, WGA, UEA, ECL	H16, B3	mouse anti-prevnar	>3	2	NaP 8.5
76	GalNAc(b1-4)Gal(a1-1)aminopentanol	5	DBA, SBA, WGA			>3	2	NaP 8.5
77	FucNAc(a1-3)GalNAc(a1-4)Gal(a1-1)aminopentanol	5	UEA, BSL	H16, B3		>3	3	NaP 8.5
78	FucNAc(b1-3)GalNAc(a1-4)Gal(a1-1)aminopentanol	5	UEA, BSL	H16, B3		>3	3	NaP 8.5
80	GalNAc(b1-1)aminoethanol	5	BSL, SBA, RCA120, WGA, ECL			>3	1	NaP 8.5
81	FucNAc(a1-		UEA, BSL	H16, B3		>3	1	NaP 8.5

	1)aminopentanol							
82	Man(a1-2)Man(a1-2)[Gal(b1-4)]Man(a1-1)aminopentanol	6	PNA, ECL, RCA120, ConA, UEA			>3	4	NaP 8.5
83	Man(a1-2)Man(a1-2)Man(a1-1)aminopentanol		ConA			>3	3	NaP 8.5
84	Gal(b1-4)Man(a1-1)aminopentanol	6	UEA, ECL, PNA, RCA120, SBA			>3	2	NaP 8.5
85	Man(a1-2)Man(a1-1)aminopentanol		ConA			>3	2	NaP 8.5
90	Glc(b1-1)aminoethanol						1	NaP 8.5
91	GlcNAc(a1-2)Hep(a1-3)Hep(a1-5)Kdo(a2-1)aminopentanol	7, 8	WGA	1A5		>3	4	NaP 8.5
92	Hep(a1-3)Hep(a1-5)Kdo(a2-1)aminopentanol	8		1A5		2	3	NaP 8.5
93	Hep(a1-3)Hep(a1-5)[L-Ara4N(b1-8)]Kdo(a2-1)aminopentanol	8		1A5		2	4	NaP 8.5
94	Hep(a1-7)Hep(a1-3)Hep(a1-5)Kdo(a2-1)aminopentanol	8		1A5		1	4	NaP 8.5
95	Hep(a1-2)Hep(a1-3)Hep(a1-5)Kdo(a2-1)aminopentanol	8		1A5			4	NaP 8.5
96	Hep(a1-5)Kdo(a2-1)aminopentanol	8	BSL				2	NaP 8.5
97	Hep(a1-7)Hep(a1-3)Hep(a1-1)aminopentanol	9					3	NaP 8.5
98	Kdo(a2-8)Kdo(a2-4)Kdo(a2-1)aminopentanol					>3	3	NaP 8.5
99	Kdo(a2-1)aminopentanol	8		1A5		>3	1	NaP 8.5
100	Hep(a1-1)aminopentanol	8				>3	1	NaP 8.5
101	Glc(b1-1)aminopentanol	10				>3	1	NaP 8.5
102	D-FucNAc(b1-1)aminopentanol	10	DBA, RCA120, UEA, WGA, BSL, SNL			>3	1	NaP 8.5
103	FucNAc(b1-1)aminopentanol	10	UEA	H16, B3		>3	1	NaP 8.5
104	Glc(b1-3)D-FucNAc(b1-1)aminopentanol	10				>3	2	NaP 8.5
105	Glc(b1-3)FucNAc(b1-1)aminopentanol	10				>3	2	NaP 8.5
153	Gal(b1-3)GalNAc(a1-1)aminopentanol		PNA, BSL, RCA120			1	2	NaP 8.5
154	Fuc(a1-3)[Gal(b1-4)]GlcNAc(b1-1)aminopentanol		BSL		mouse anti-prevnar		2	NaP 8.5
155	Neu5Ac(a2-6)GalNAc(a1-1)aminopentanol		BSL, RCA120, WGA, PNA, SBA			2	2	NaP 8.5
156	Gal(b1-4)[Gal(b1-4)Glc(b1-6)]GlcNAc(b1-1)aminopentanol		PNA, ECL, UEA, WGA, DBA, MAL I, RCA120		mouse anti-prevnar, SP14	>3	4	NaP 8.5
157	Fuc(a1-3)[Fuc(a1-2)Gal(b1-4)]GlcNAc(b1-1)aminopentanol		BSL, UEA, SNL				4	NaP 8.5
158	Gal(b1-3)[Fuc(a1-4)]GlcNAc(b1-1)aminopentanol		UEA			2	3	NaP 8.5
159	Fuc(a1-2)Gal(b1-3)[Fuc(a1-4)]GlcNAc(b1-1)aminopentanol		UEA, BSL			2	3	NaP 8.5
160	Gal-2,3-Pyruvate(a1-1)aminopentanol (mixture of R/S pyruvate)	5				>3	1	NaP 8.5

161	Gal(a1-3)Gal(b1-4)Glc(b1-1)aminopentanol	11	BSL, RCA120			>3	3	NaP 8.5
162	Gal(a1-3)Gal(b1-4)GlcNAc(b1-1)aminopentanol		BSL, WGA, RCA120, SNL, SBA, ECL			3	3	NaP 8.5
163	Gal(a1-3)Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminopentanol		BSL, WGA, RCA120, PNA, SBA			3	5	NaP 8.5
164	Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminopentanol		UEA			>3	4	NaP 8.5
165	Fuc(a1-2)Gal(b1-3)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminopentanol		PNA, RCA120, ECL, DBA, BSL			>3	5	NaP 8.5
166	Gal(b1-3)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminopentanol					>3	4	NaP 8.5
167	Rha(a1-1)aminopentanol	12		1A10		>3	1	NaP 8.5
168	Rha(a1-3)Glc(b1-1)aminopentanol	12		1A10		>3	2	NaP 8.5
169	Glc(a1-2)Glc(a1-1)aminopentanol	12	ConA			>3	2	NaP 8.5
170	Glc(b1-4)Glc(a1-2)Glc(a1-1)aminopentanol	12	ConA	1A10		>3	3	NaP 8.5
171	Rha(a1-3)Glc(b1-4)Glc(a1-1)aminopentanol	12				>3	3	NaP 8.5
172	Gal(b1-3)GalNAc(b1-3)Gal(a1-4)Gal(b1-4)Glc(b1-1)aminopentanol		PNA, BSL			>3	5	NaP 8.5
173	Neu5Ac(a2-8)Neu5Ac(a2-3)[GalNAc(b1-4)]Gal(b1-4)Glc(b1-1)aminopentanol		SBA, DBA				5	NaP 8.5
174	Gal(a1-4)Gal(b1-4)Glc(b1-1)aminopentanol	11	BSL, SBA			>3	3	NaP 8.5
175	GalNAc(a1-1)AminoLinker2	13	BSL, SBA, WGA, DBA, SNL			1	1	NaP 8.5
176	Fuc(a1-3)[Gal(b1-4)]GlcNAc(b1-1)AminoLinker2	13	BSL				3	NaP 8.5
177	GlcNAc(a1-2)Hep(a1-3)Hep(a1-1)aminopentanol		WGA			>3	3	NaP 8.5
178	Hep(a1-3)Hep(a1-1)aminopentanol	14				2	2	NaP 8.5
179	Gal(b1-4)Glc(b1-1)aminopentanol	14	PNA, RCA120, ECL, MALI			>3	2	NaP 8.5
180	GalNAc(b1-4)Gal(b1-4)Glc(b1-1)aminopentanol		SBA, DBA			>3	3	NaP 8.5
181	Neu5Ac(a2-3)Gal(b1-4)Glc(b1-1)aminopentanol		WGA			1	3	NaP 8.5
182	GalNAc-4-sulfate(b1-1)aminopentanol	15					1	NaP 8.5
183	IdoA-2,4-disulfate(a1-1)aminopentanol	15					1	NaP 8.5
184	IdoA(a1-3)GalNAc-4-sulfate(b1-1)aminopentanol	15					2	NaP 8.5

185	IdoA-2-sulfate(a1-3)GalNAc-4-sulfate(b1-1)aminopentanol	15					2	NaP 8.5
186	IdoA(a1-3)GalNAc(b1-1)aminopentanol	15					2	NaP 8.5
187	GlcA(b1-4)Glc(b1-3)GlcA(b1-4)Glc(b1-1)aminoethanol				mouse anti-prevnar, SP3	>3	4	NaP 8.5
188	Glc(b1-3)GlcA(b1-4)Glc(b1-1)aminoethanol	16			mouse anti-prevnar, SP3	>3	3	NaP 8.5
189	GalNAc(a1-1)Thr-Linker		BSL, SBA, WGA, DBA			3	1	NaP 8.5
190	Glc(b1-3)Glc(b1-3)[Glc(b1-6)]Glc(b1-3)Glc(b1-1)aminopentanol					>3	5	NaP 8.5
191	Glc(b1-3)Glc(b1-3)[Glc(b1-6)]Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-1)aminopentanol					>3	9	NaP 8.5
192	Glc(b1-3)Glc(b1-3)[Glc(b1-6)]Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-1)aminopentanol					>3	13	NaP 8.5
193	Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-3)Glc(b1-1)aminopentanol	17				>3	12	NaP 8.5
194	L-PneNAc(a1-2)GlcA(b1-3)FucNAc(a1-3)D-FucNAc(b1-1)aminopentanol			H16, B3	SP5	>3	4	NaP 8.5
195	Mixture of: D-6d-xylHexpNAc-4-ulo(b1-1)aminopentanol and D-FucNAc(b1-1)aminopentanol		WGA, DBA, SNL				1	NaP 8.5
196	FucNAc(a1-3)D-6d-xylHexpNAc-4-ulo(b1-1)aminopentanol and FucNAc(a1-3)D-FucNAc(b1-1)aminopentanol			H16, B3	mouse anti-prevnar	>3	2	NaP 8.5
197	FucNAc(a1-3)D-FucNAc(b1-1)aminopentanol			H16, B3		>3	2	NaP 8.5
198	GlcA(b1-4)FucNAc(a1-1)aminopentanol					>3	2	NaP 8.5
199	Glc(b1-3)FucNAc(a1-1)aminopentanol					>3	2	NaP 8.5
200	L-PneNAc(a1-2)GlcA(b1-1)aminopentanol				SP5	>3	2	NaP 8.5
201	L-PneNAc(a1-1)aminopentanol			H16, B3	SP5	>3	1	NaP 8.5
202	L-PneNAc(b1-1)aminopentanol					2	1	NaP 8.5
203	Gal(b1-4)[Glc(b1-6)]GlcNAc(b1-3)Gal(b1-1)aminopentanol		ECL, RCA120		SP14	>3	4	NaP 8.5
204	Glc(a1-4)Gal(a1-4)GlcA(b1-4)Glc(b1-1)aminoethanol				SP8	3	4	NaP 8.5
205	Glc(a1-4)Gal(a1-1)aminoethanol					>3	2	NaP 8.5
206	GlcA(b1-4)Glc(b1-				SP8	>3	4	NaP 8.5

	4)Glc(a1-4)Gal(a1-1)aminoethanol							
207	Glc(a1-4)Gal(a1-4)GlcA(b1-4)Glc(b1-1)aminopentanol				SP8	>3	4	NaP 8.5
208	Gal(a1-4)GlcA(b1-4)Glc(b1-4)Glc(a1-1)aminopentanol		BSL		SP8	>3	4	NaP 8.5
209	GlcA(b1-4)Glc(b1-4)Glc(a1-4)Gal(a1-1)aminopentanol				SP8	>3	4	NaP 8.5
210	Glc(b1-4)Glc(a1-4)Gal(a1-4)GlcA(b1-1)aminopentanol				SP8	>3	4	NaP 8.5
211	Xyl(b1-4)Xyl(b1-1)aminopentanol	18		LM10			2	NaP 8.5
212	Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-1)aminopentanol	18		LM10		1	4	NaP 8.5
213	Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-1)aminopentanol	18		LM10		>3	6	NaP 8.5
214	Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-4)Xyl(b1-1)aminopentanol	18		LM10			8	NaP 8.5
215	Glc(b1-4)Glc(b1-4)Glc(b1-4)Glc(b1-1)aminopentanol					>3	4	NaP 8.5
216	Glc(b1-3)GlcA(b1-4)Glc(b1-1)aminopentanol				mouse anti-prevnar, SP3	>3	3	NaP 8.5
217	GlcA(b1-4)Glc(b1-1)aminoethanol						2	NaP 8.5
218	Glc(b1-3)GlcA(b1-1)aminoethanol					>3	2	NaP 8.5
219	ManNAc(b1-3)FucNAc(a1-3)GalNAc(a1-4)Gal-2,3-pyruvate(a1-1)aminopentanol	5		H16, B3	mouse anti-prevnar, SP4	>3	4	NaP 8.5
220	GlcA(b1-1)aminoethanol						1	NaP 8.5
225	Glc(a1-4)GalNAc(b1-4)Man(a1-1)aminopentanol	19	ConA			>3	3	NaP 8.5
226	Glc(a1-4)GalNAc(b1-4)[Man(a1-2)Man(a1-6)]Man(a1-1)aminopentanol	19	ConA			>3	5	NaP 8.5
227	Glc(a1-4)GalNAc(b1-4)[Man-6-PEtN(a1-2)Man(a1-6)]Man(a1-1)aminopentanol	19	ConA			>3	5	NaP 8.5
229	GalNAc(b1-4)Man(a1-1)aminopentanol		SBA, WGA, DBA, ECL, SNL			>3	2	NaP 8.5
230	GalNAc(b1-4)[Man-6-PEtN(a1-2)Man(a1-6)]Man(a1-1)aminopentanol		SBA, ConA, WGA, DBA, ECL			3	4	NaP 8.5
231	GalNAc(b1-4)[Man(a1-2)Man(a1-6)]Man(a1-1)aminopentanol		SBA, ConA, WGA			2	4	NaP 8.5
232	GalNAc(b1-4)[Man-6-PEtN(a1-2)Man(a1-6)]Man-2-PEtN(a1-1)aminopentanol		SBA, ConA				4	NaP 8.5
233	GalNAc(b1-4)Man(a1-1)aminododecanol		SBA, WGA			3	2	NaP 8.5
234	GalNAc(b1-4)Man(a1-1)p-aminocyclohexanol		SBA, WGA			>3	2	NaP 8.5

235	GlcA(b1-4)Glc(b1-3)GlcA(b1-1)aminoethanol				SP3		3	NaP 8.5
236	GlcA(b1-4)Glc(b1-3)Glc(b1-4)Glc(b1-1)aminoethanol and Glc(b1-4)Glc(b1-3)GlcA(b1-4)Glc(b1-1)aminoethanol				SP3	>3	4	NaP 8.5
237	Man(a1-1)aminopentanol		ConA				1	NaP 8.5
238	GlcNAc-6-P-phosphoaminopentanol(a1-3)GlcNAc-6-P-phosphoaminopentanol(a1-2)glyceric acid	20				3	2	NaP 8.5
239	GlcA(a1-3)Gal(a1-3)ManNAc(b1-4)Glc(b1-4)Glc(a1-1)aminopentanol					1	5	NaP 8.5
240	GlcA(a1-3)Gal(a1-3)ManNAc-6-acetate(b1-4)Glc(b1-4)Glc(a1-1)aminopentanol		DBA		SP9	1	5	NaP 8.5, PBS
241	GlcA(a1-3)Gal(a1-1)aminopentanol					1	2	NaP 8.5
242	Glc(b1-4)Glc(a1-1)aminopentanol					2	2	NaP 8.5
243	ManNAc(b1-4)Glc(b1-4)Glc(a1-1)aminopentanol					>3	3	NaP 8.5
244	GalNAc(b1-3)GalNAc(b1-1)aminopentanol		SBA				2	NaP 8.5
245	Glc(b1-4)Gal(b1-4)Glc(b1-1)aminopentanol	11				2	3	NaP 8.5
247	Rha(a1-3)[Rha(a1-3)Glc(b1-4)]Glc(a1-2)Glc(a1-1)aminopentanol	12				>3	5	NaP 8.5
248	GlcNAc(a1-3)GlcNAc-6-P-phosphoaminopentanol(a1-2)glyceric acid	20				>3	2	NaP 8.5
249	GlcNAc(a1-3)GlcNAc[(a1-2)glyceric acid](6-P-6)GlcNAc(a1-3)GlcNAc-6-P-phosphoaminopentanol(a1-2)glyceric acid	20				>3	4	NaP 8.5
250	Man(a1-2)Man(a1-2)[Gal(b1-4)]Man(a1-1)aminoethanol	6	ConA			1	4	NaP 8.5
251	Glc(b1-3)Gal(b1-4)Man(a1-1)aminopentanol	6				1	3	NaP 8.5
252	Rha(a1-2)Rha(a1-2)Rha(a1-1)aminopentanol					>3	3	NaP 8.5
253	GalNAc-2,3-Oxazolidinone(a1-4)GalNAc-2,3-Oxazolidinone(a1-1)aminopentanol		BSL, SBA, WGA			1	2	NaP 8.5
254	Glc(a1-2)Glc(a1-3)[FucNAc(a1-3)GalNAc(b1-4)]ManNAcA(b1-1)aminopentanol					>3	5	NaP 8.5
255	Glc(a1-2)Glc(a1-3)[Gal(a1-3)FucNAc(a1-3)GalNAc(b1-4)]ManNAcA(b1-1)aminopentanol					3	6	NaP 8.5

256	Araf(a1-3)[Araf(a1-5)]Araf(a1-5)Araf(a1-1)aminopentanol					>3	4	NaP 8.5
257	Man(a1-5)Araf(a1-3)[Man(a1-5)Araf(a1-5)]Araf(a1-5)Araf(a1-1)aminopentanol		ConA			1	6	NaP 8.5
258	GalNAc-3,4-diacetate(a1-4)GalNAc-3-acetate(a1-4)GalNAc-3-acetate(a1-4)GalNAc-3-acetate(a1-1)aminopentanol						4	PBS

Tab. 51 Heparin-like compounds printed on the heparin/glycosaminnoglycan glycan microarray from the *MPI für Kolloid- und Grenzflächen-forschung* (MPIKG).

Compound	Trivial-Name	Name
1	Heparin I	GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-4)GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-4)GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-1)AminoLinker1
2	Heparin II	GlcNAc-6-sulfate(a1-4)IdoA-2-sulfate(a1-4)GlcNAc-6-sulfate(a1-4)IdoA-2-sulfate(a1-4)GlcNAc-6-sulfate(a1-4)IdoA-2-sulfate(a1-1)aminopentanol
3	Heparin III	GlcN-N-sulfate(a1-4)IdoA(a1-4)GlcN-N-sulfate(a1-4)IdoA(a1-4)GlcN-N-sulfate(a1-4)IdoA(a1-1)aminopentanol
4	Heparin IV	GlcNAc(a1-4)IdoA(a1-4)GlcNAc(a1-4)IdoA(a1-4)GlcNAc(a1-4)IdoA(a1-1)aminopentanol
5	Heparin V	IdoA-2-sulfate(a1-4)GlcN-N-sulfate(a1-4)IdoA-2-sulfate(a1-4)GlcN-N-sulfate(a1-4)IdoA-2-sulfate(a1-4)GlcN-N-sulfate(a1-1)
6	Heparin VI	IdoA-2-sulfate(a1-4)GlcNAc(a1-4)IdoA-2-sulfate(a1-4)GlcNAc(a1-4)IdoA-2-sulfate(a1-4)GlcNAc(a1-1)aminopentanol
7	Heparin VII	GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-4)GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-1)AminoLinker1
8	Heparin VIII	IdoA(a1-4)GlcN-N-sulfate(a1-4)IdoA(a1-4)GlcN-N-sulfate(a1-1)aminopentanol
9	Heparin IX	IdoA(a1-4)GlcN(a1-4)GlcA(b1-3)GlcNAc(a1-1)AminoLinker1
10	Heparin X	GlcN-6,N-disulfate(a1-4)IdoA-2-sulfate(a1-1)AminoLinker1
11	Heparin XI	Ido-2,4-disulfate(a1-1)AminoLinker1
12	Heparin XII	GlcN-6,N-disulfate(a1-1)AminoLinker1
13	Heparin XIII	IdoA(a1-1)AminoLinker1
14	5 kDa native Heparin	
15	Negative control	Printing buffer
16	Neu5Aca(2-6)Galb(1-4)penta	Neu5Ac(a2-6)Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol
17	Neu5Aca(2-3)Galb(1-3)penta	Neu5Ac(a2-3)Gal(b1-3)GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol
18	Sialyl LewisX	Fuc(a1-3)[Neu5Ac(a2-3)Gal(b1-4)]GlcNAc(b1-3)Gal(b1-4)Glc(b1-1)aminohexanol
19	Neu5Aca(2-6)Gal GM3	Neu5Ac(a2-6)Gal(b1-4)Glc(b1-1)aminohexanol
20	Neu5Aca(2-3)Gal GM3	Neu5Ac(a2-3)Gal(b1-4)Glc(b1-1)aminohexanol
21	Neu5Aca(2-6)GalSO3H	Neu5Ac(a2-6)Gal(b1-4)GlcNAc-6-sulfate(b1-1)aminohexanol
22	Sialyl-Tn (STn) Antigen	Neu5Ac(a2-6)GalNAc(a1-1)aminopentanol
23	GD2 (Aminolinker)	Neu5Ac(a2-8)Neu5Ac(a2-3)[GalNAc(b1-4)]Gal(b1-4)Glc(b1-1)aminopentanol
24	GM3 (Aminolinker)	Neu5Ac(a2-3)Gal(b1-4)Glc(b1-1)aminopentanol
25	Dermatan GalNAc monosaccharide, 6 sulfated	GalNAc-4-sulfate(b1-1)aminopentanol
26	Iduronic acid disulfate monosaccharide	IdoA-2,4-disulfate(a1-1)aminopentanol
27	Dermatan monosulfated disaccharide	IdoA(a1-3)GalNAc-4-sulfate(b1-1)aminopentanol
28	Dermatan disulfated disaccharide	IdoA-2-sulfate(a1-3)GalNAc-4-sulfate(b1-1)aminopentanol
29	Dermatan nonsulfated disaccharide	IdoA(a1-3)GalNAc(b1-1)aminopentanol

30	LacNAc repeat dimer	Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)GlcNAc(b1-1)aminopentanol
31	LacNAc repeat branched trimer	Gal(b1-4)GlcNAc(b1-3)[Gal(b1-4)GlcNAc(b1-6)]Gal(b1-4)GlcNAc(b1-1)aminopentanol
32	LacNAc repeat linear trimer	Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)GlcNAc(b1-1)aminopentanol
33	Keratan sulfate 2 repeating units	Gal-6-sulfate(b1-4)GlcNAc(b1-3)Gal-6-sulfate(b1-4)GlcNAc(b1-1)aminopentanol
34	Keratan sulfate 2 repeating units	Gal(b1-4)GlcNAc-6-sulfate(b1-3)Gal(b1-4)GlcNAc-6-sulfate(b1-1)aminopentanol
35	Keratan sulfate 2 repeating units	Gal-3,6-disulfate(b1-4)GlcNAc(b1-3)Gal-6-sulfate(b1-4)GlcNAc(b1-1)aminopentanol
36	Keratan sulfate 2 repeating units	Gal-6-sulfate(b1-4)GlcNAc-6-sulfate(b1-3)Gal-6-sulfate(b1-4)GlcNAc-6-sulfate(b1-1)aminopentano
37	negative control	Printing buffer

7.7. Abbrevations

Å	Ångstrom (1 Å = 100 pm = 10^{-10} m)
CD	Circular dichroism
C-terminal	Carboxyterminal
CV	Column Volume
Da / kDa	Dalton (1 Da = 1 u = 1,660539040(20)·10^{-27} kg)
DMSO	Dimethyl sulfooxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GPI	Glycosylphosphatidyl-inositol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
M	Molar (mol $^{-1}$)
Mb	Mega base pairs
N-terminal	Aminoterminal
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonylfluorid
r.m.s.d.	Root mean square deviation
rpm.	Revolutions per minute
RT	Room temperature

8. References

1 Kurtzman, C., Fell, J. W. & Boekhout, T. *The yeasts: a taxonomic study*. (Elsevier, 2011).

2 Kurtzman, C. & Piškur, J. Taxonomy and phylogenetic diversity among the yeasts. *Comparative Genomics*, 29-46 (2006).

3 Sláviková, E. & Vadkertiová, R. The diversity of yeasts in the agricultural soil. *Journal of basic microbiology* **43**, 430-436 (2003).

4 Bass, D. *et al.* Yeast forms dominate fungal diversity in the deep oceans. *Proceedings of the Royal Society of London B: Biological Sciences* **274**, 3069-3077 (2007).

5 Raspor, P. & Zupan, J. Yeasts in extreme environments. *Biodiversity and ecophysiology of yeasts*, 371-417 (2006).

6 Spribille, T. *et al.* Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* **353**, 488-492 (2016).

7 Martini, A. Biodiversity and conservation of yeasts. *Biodiversity and Conservation* **1**, 324-333 (1992).

8 Cauchie, M., Desmet, S. & Lagrou, K. Candida and its dual lifestyle as a commensal and a pathogen. *Research in Microbiology* (2017).

9 Barnett, J. A. A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850–1880. *Yeast* **16**, 755-771 (2000).

10 De Deken, R. The Crabtree effect: a regulatory system in yeast. *Microbiology* **44**, 149-156 (1966).

11 Piškur, J. & Compagno, C. *Molecular mechanisms in yeast carbon metabolism*. (Springer, 2014).

12 Piškur, J., Rozpędowska, E., Polakova, S., Merico, A. & Compagno, C. How did Saccharomyces evolve to become a good brewer? *TRENDS in Genetics* **22**, 183-186 (2006).

13 Balasubramanian, M. K., Bi, E. & Glotzer, M. Comparative analysis of cytokinesis in budding yeast, fission yeast and animal cells. *Current Biology* **14**, R806-R818 (2004).

14 Neiman, A. M. Ascospore formation in the yeast Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews* **69**, 565-584 (2005).

15 Hornsey, I. S. *A history of beer and brewing*. (Royal Society of Chemistry, 2007).

16 Mattanovich, D., Sauer, M. & Gasser, B. Yeast biotechnology: teaching the old dog new tricks. *Microbial cell factories* **13**, 34 (2014).

17 Kurtzman, C. P. & Robnett, C. J. Phylogenetic relationships among yeasts of the 'Saccharomyces complex'determined from multigene sequence analyses. *FEMS yeast research* **3**, 417-432 (2003).

18 Gabaldón, T. *et al.* Comparative genomics of emerging pathogens in the Candida glabrata clade. *BMC genomics* **14**, 623 (2013).

19 Dujon, B. *et al.* Genome evolution in yeasts. *Nature* **430**, 35-44 (2004).

20 Gabaldón, T., Naranjo-Ortíz, M. A. & Marcet-Houben, M. Evolutionary genomics of yeast pathogens in the Saccharomycotina. *FEMS yeast research* **16** (2016).

21 Gabaldón, T. & Fairhead, C. Genomes shed light on the secret life of Candida glabrata: not so asexual, not so commensal. *Current genetics*, 1-6 (2018).

22 Marcet-Houben, M. & Gabaldón, T. The tree versus the forest: the fungal tree of life and the topological diversity within the yeast phylome. *PloS one* **4**, e4357 (2009).
23 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).
24 Fidel, P. L., Vazquez, J. A. & Sobel, J. D. Candida glabrata: review of epidemiology, pathogenesis, and clinical disease with comparison toC. albicans. *Clinical microbiology reviews* **12**, 80-96 (1999).
25 Cole, G. T., Halawa, A. A. & Anaissie, E. J. The role of the gastrointestinal tract in hematogenous candidiasis: from the laboratory to the bedside. *Clinical Infectious Diseases* **22**, S73-S88 (1996).
26 Cafarchia, C., Romito, D., Coccioli, C., Camarda, A. & Otranto, D. Phospholipase activity of yeasts from wild birds and possible implications for human disease. *Sabouraudia* **46**, 429-434 (2008).
27 de Melo Pereira, G. V. *et al.* Isolation, selection and evaluation of yeasts for use in fermentation of coffee beans by the wet process. *International journal of food microbiology* **188**, 60-66 (2014).
28 Al-Yasiri, M. H. *et al.* Opportunistic fungal pathogen Candida glabrata circulates between humans and yellow-legged gulls. *Scientific reports* **6**, 36157 (2016).
29 Bialkova, A. & Šubík, J. Biology of the pathogenic yeastCandida glabrata. *Folia microbiologica* **51**, 3-20 (2006).
30 Rodrigues, C. F., Silva, S. & Henriques, M. Candida glabrata: a review of its features and resistance. *European journal of clinical microbiology & infectious diseases* **33**, 673-688 (2014).
31 Roetzer, A., Gabaldón, T. & Schüller, C. From Saccharomyces cerevisiae to Candida glabrata in a few easy steps: important adaptations for an opportunistic pathogen. *FEMS microbiology letters* **314**, 1-9 (2011).
32 Inglis, D. O. *et al.* The Candida genome database incorporates multiple Candida species: multispecies search and analysis tools with curated gene and protein information for Candida albicans and Candida glabrata. *Nucleic acids research* **40**, D667-D674 (2012).
33 Cherry, J. M. *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* **40**, D700-D705 (2012).
34 Perlroth, J., Choi, B. & Spellberg, B. Nosocomial fungal infections: epidemiology, diagnosis, and treatment. *Medical mycology* **45**, 321-346 (2007).
35 Pfaller, M. A. *et al.* Epidemiology and outcomes of invasive candidiasis due to non-albicans species of Candida in 2,496 patients: data from the Prospective Antifungal Therapy (PATH) registry 2004–2008. *PLoS One* **9**, e101510 (2014).
36 Pfaller, M. & Diekema, D. Epidemiology of invasive candidiasis: a persistent public health problem. *Clinical microbiology reviews* **20**, 133-163 (2007).
37 Pfaller, M. & Diekema, D. Twelve years of fluconazole in clinical practice: global trends in species distribution and fluconazole susceptibility of bloodstream isolates of Candida. *Clinical Microbiology and Infection* **10**, 11-23 (2004).
38 Calcagno, A. M. *et al.* Candida glabrata STE12 is required for wild-type levels of virulence and nitrogen starvation induced filamentation. *Molecular microbiology* **50**, 1309-1318 (2003).

39 Wargo, M. J. & Hogan, D. A. Fungal—bacterial interactions: a mixed bag of mingling microbes. *Current opinion in microbiology* **9**, 359-364 (2006).

40 Cohen, Y. *et al.* Early prediction of Candida glabrata fungemia in nonneutropenic critically ill patients*. *Critical care medicine* **38**, 826-830 (2010).

41 Caston-Osorio, J., Rivero, A. & Torre-Cisneros, J. Epidemiology of invasive fungal infection. *International journal of antimicrobial agents* **32**, S103-S109 (2008).

42 Hof, H. Mycoses in the elderly. *European journal of clinical microbiology & infectious diseases* **29**, 5-13 (2010).

43 Hitchcock, C., Pye, G., Troke, P., Johnson, E. & Warnock, D. Fluconazole resistance in Candida glabrata. *Antimicrobial agents and chemotherapy* **37**, 1962-1965 (1993).

44 Klepser, M. E. Candida resistance and its clinical relevance. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* **26**, 68S-75S (2006).

45 Moran, C., Grussemeyer, C. A., Spalding, J. R., Benjamin, D. K. & Reed, S. D. Comparison of costs, length of stay, and mortality associated with Candida glabrata and Candida albicans bloodstream infections. *American journal of infection control* **38**, 78-80 (2010).

46 Douglas, L. J. Candida biofilms and their role in infection. *Trends in microbiology* **11**, 30-36 (2003).

47 Pisa, D., Alonso, R., Rábano, A., Rodal, I. & Carrasco, L. Different Brain Regions are Infected with Fungi in Alzheimer's Disease. *Scientific reports* **5** (2015).

48 Fabre, E. *et al.* Comparative genomics in hemiascomycete yeasts: evolution of sex, silencing, and subtelomeres. *Molecular biology and evolution* **22**, 856-873 (2004).

49 Csank, C. & Haynes, K. Candida glabrata displays pseudohyphal growth. *FEMS microbiology letters* **189**, 115-120 (2000).

50 Vila, T. V. M. & Rozental, S. in *Fungal Pathogenicity* (InTech, 2016).

51 Dalle, F. *et al.* Cellular interactions of Candida albicans with human oral epithelial cells and enterocytes. *Cellular microbiology* **12**, 248-271 (2010).

52 Zakikhany, K. *et al.* In vivo transcript profiling of Candida albicans identifies a gene essential for interepithelial dissemination. *Cellular microbiology* **9**, 2938-2954 (2007).

53 Brunke, S. & Hube, B. Two unlike cousins: Candida albicans and C. glabrata infection strategies. *Cellular microbiology* **15**, 701-708 (2013).

54 Roetzer, A., Gratz, N., Kovarik, P. & Schüller, C. Autophagy supports Candida glabrata survival during phagocytosis. *Cellular microbiology* **12**, 199-216 (2010).

55 Seider, K. *et al.* The facultative intracellular pathogen Candida glabrata subverts macrophage cytokine production and phagolysosome maturation. *The Journal of Immunology* **187**, 3072-3086 (2011).

56 Charlier, C. *et al.* Evidence of a role for monocytes in dissemination and brain invasion by Cryptococcus neoformans. *Infection and immunity* **77**, 120-127 (2009).

57 Tati, S. *et al.* Candida glabrata binding to Candida albicans hyphae enables Its development in oropharyngeal candidiasis. *PLoS pathogens* **12**, e1005522 (2016).

58 Coco, B. *et al.* Mixed Candida albicans and Candida glabrata populations associated with the pathogenesis of denture stomatitis. *Oral microbiology and immunology* **23**, 377-383 (2008).

59 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**, S226-S232, doi:10.1007/s00239-003-0031-2 (2003).

60 Martini, A. Origin and domestication of the wine yeast Saccharomyces cerevisiae. *Journal of Wine research* **4**, 165-176 (1993).

61 Goffeau, A., Barrell, B. G., Bussey, H. & Davis, R. Life with 6000 genes. *Science* **274**, 546 (1996).

62 Liu, W. *et al.* From Saccharomyces cerevisiae to human: The important gene co-expression modules. *Biomedical reports* **7**, 153-158 (2017).

63 Mager, W. H. & Winderickx, J. Yeast as a model for medical and medicinal research. *Trends in pharmacological sciences* **26**, 265-273 (2005).

64 Buijs, N. A., Siewers, V. & Nielsen, J. Advanced biofuel production by the yeast Saccharomyces cerevisiae. *Current opinion in chemical biology* **17**, 480-488 (2013).

65 Sanchez-Garcia, L. *et al.* Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microbial cell factories* **15**, 33 (2016).

66 Herskowitz, I. Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiological reviews* **52**, 536 (1988).

67 Brückner, S. & Moesch, H.-U. Choosing the right lifestyle: adhesion and development in Saccharomyces cerevisiae. *FEMS microbiology reviews* **36**, 25-58 (2011).

68 Hoyer, L., Clevenger, J., Hecht, J., Ehrhart, E. & Poulet, F. Detection of Als Proteins on the Cell Wall ofCandida albicans in Murine Tissues. *Infection and immunity* **67**, 4251-4255 (1999).

69 Sundstrom, P. Adhesion in Candida spp. *Cellular microbiology* **4**, 461-469 (2002).

70 Diderrich, R. *et al.* Structural hot spots determine functional diversity of the Candida glabrata epithelial adhesin family. *Journal of Biological Chemistry* **290**, 19597-19613 (2015).

71 Tumbarello, M. *et al.* Biofilm production by Candida species and inadequate antifungal therapy as predictors of mortality for patients with candidemia. *Journal of clinical microbiology* **45**, 1843-1850 (2007).

72 Wheeler, R. T. & Fink, G. R. A drug-sensitive genetic network masks fungi from the immune system. *PLoS pathogens* **2**, e35 (2006).

73 Smits, G. J., Kapteyn, J. C., van den Ende, H. & Klis, F. M. Cell wall dynamics in yeast. *Current opinion in microbiology* **2**, 348-352 (1999).

74 Orlean, P. Biogenesis of yeast wall and surface components. *The Molecular and Cellular Biology of the Yeast Saccharomyces. Cell, Cycle and Cell Biology*, 229-362 (1997).

75 Lipke, P. N. & Ovalle, R. Cell wall architecture in yeast: new structure and new challenges. *Journal of bacteriology* **180**, 3735-3740 (1998).

76 Orlean, P. Architecture and biosynthesis of the Saccharomyces cerevisiae cell wall. *Genetics* **192**, 775-818 (2012).

77 Gabius, H.-J. *The sugar code: fundamentals of glycosciences*. (John Wiley & Sons, 2011).

78 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).
79 Kollár, R. *et al.* Architecture of the yeast cell wall β (1→ 6)-glucan interconnects mannoprotein, β (1→ 3)-glucan, and chitin. *Journal of Biological Chemistry* **272**, 17762-17775 (1997).
80 Osumi, M. The ultrastructure of yeast: cell wall structure and formation. *Micron* **29**, 207-233 (1998).
81 Marquardt, H. Der Feinbau von Hefezellen im Elektronenmikroskop. *Zeitschrift für Naturforschung B* **17**, 42-48 (1962).
82 Kapteyn, J. C., Van Den Ende, H. & Klis, F. M. The contribution of cell wall proteins to the organization of the yeast cell wall. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1426**, 373-383 (1999).
83 Lesage, G. & Bussey, H. Cell wall assembly in Saccharomyces cerevisiae. *Microbiology and molecular biology reviews* **70**, 317-343 (2006).
84 Cappellaro, C., Baldermann, C., Rachel, R. & Tanner, W. Mating type-specific cell-cell recognition of Saccharomyces cerevisiae: cell wall attachment and active sites of a-and alpha-agglutinin. *The EMBO journal* **13**, 4737 (1994).
85 de Groot, P. W., Bader, O., de Boer, A. D., Weig, M. & Chauhan, N. Adhesins in human fungal pathogens: glue with plenty of stick. *Eukaryotic cell* **12**, 470-481 (2013).
86 Verna, J., Lodder, A., Lee, K., Vagts, A. & Ballester, R. A family of genes required for maintenance of cell wall integrity and for the stress response in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences* **94**, 13804-13809 (1997).
87 Rodicio, R. & Heinisch, J. J. Together we are strong—cell wall integrity sensors in yeasts. *Yeast* **27**, 531-540 (2010).
88 Hoyer, L. L. The ALS gene family of Candida albicans. *Trends in microbiology* **9**, 176-180 (2001).
89 Alberti-Segui, C. *et al.* Identification of potential cell-surface proteins in Candida albicans and investigation of the role of a putative cell-surface glycosidase in adhesion and virulence. *Yeast* **21**, 285-302 (2004).
90 Gaur, N. K., Klotz, S. A. & Henderson, R. L. Overexpression of the Candida albicans ALA1 gene in Saccharomyces cerevisiae results in aggregation following attachment of yeast cells to extracellular matrix proteins, adherence properties similar to those of Candida albicans. *Infection and immunity* **67**, 6040-6047 (1999).
91 Sheppard, D. C. *et al.* Functional and structural diversity in the Als protein family of Candida albicans. *Journal of Biological Chemistry* (2004).
92 Liu, Y. & Filler, S. G. Candida albicans Als3, a multifunctional adhesin and invasin. *Eukaryotic cell* **10**, 168-173 (2011).
93 Veelders, M. *et al.* Structural basis of flocculin-mediated social behavior in yeast. *Proceedings of the National Academy of Sciences* **107**, 22511-22516 (2010).
94 Kraushaar, T. *et al.* Interactions by the fungal Flo11 adhesin depend on a fibronectin type III-like adhesin domain girdled by aromatic bands. *Structure* **23**, 1005-1017 (2015).
95 Verstrepen, K. J. & Klis, F. M. Flocculation, adhesion and biofilm formation in yeasts. *Molecular microbiology* **60**, 5-15 (2006).

96 Kock, M. A. Vergleichende Struktur-/Funktionsanalyse von Adhäsionsdomänen pathogener und nicht-pathogener Hefen. (2015).
97 Kock, M. *et al.* Structural and functional characterization of PA14/Flo5-like adhesins from Komagataella pastoris. *Frontiers in Microbiology* **9**, 2581 (2018).
98 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).
99 Frieman, M. B., McCaffery, J. M. & Cormack, B. P. Modular domain structure in the Candida glabrata adhesin Epa1p, a β1, 6 glucan-cross-linked cell wall protein. *Molecular microbiology* **46**, 479-492 (2002).
100 Verstrepen, K. J., Reynolds, T. B. & Fink, G. R. Origins of variation in the fungal cell surface. *Nature reviews. Microbiology* **2**, 533 (2004).
101 Jentoft, N. Why are proteins O-glycosylated? *Trends in biochemical sciences* **15**, 291-294 (1990).
102 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).
103 Ramsook, C. B. *et al.* Yeast cell adhesion molecules have functional amyloid-forming sequences. *Eukaryotic cell* **9**, 393-404 (2010).
104 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).
105 Cummings, R. D. & McEver, R. P. in *Essentials of Glycobiology. 2nd edition* (Cold Spring Harbor Laboratory Press, 2009).
106 Calderone, R. A. & Fonzi, W. A. Virulence factors of Candida albicans. *Trends in microbiology* **9**, 327-335 (2001).
107 Weig, M. *et al.* Systematic identification in silico of covalently bound cell wall proteins and analysis of protein–polysaccharide linkages of the human pathogen Candida glabrata. *Microbiology* **150**, 3129-3144 (2004).
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* **109**, 16864-16869 (2012).
109 Diderrich, R. Strukturelle und funktionelle Charakterisierung der epithelialen Adhäsine aus Candida glabrata. (2014).
110 Iraqui, I. *et al.* The Yak1p kinase controls expression of adhesins and biofilm formation in Candida glabrata in a Sir4p-dependent pathway. *Molecular microbiology* **55**, 1259-1271 (2005).
111 Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic acids research* **42**, D222-D230 (2013).
112 Kraneveld, E. A. *et al.* Identification and differential gene expression of adhesin-like wall proteins in Candida glabrata biofilms. *Mycopathologia* **172**, 415-427 (2011).
113 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).
114 Hoyer, L. & Hecht, J. The ALS5 gene of Candida albicans and analysis of the Als5p N-terminal domain. *Yeast* **18**, 49-60 (2001).
115 Hoyer, L., Payne, T. & Hecht, J. Identification of Candida albicans ALS2 andALS4 and Localization of Als Proteins to the Fungal Cell Surface. *Journal of bacteriology* **180**, 5334-5343 (1998).

116 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 (1997).

117 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).

118 Kintzer, A. F. *et al.* The protective antigen component of anthrax toxin forms functional octameric complexes. *Journal of molecular biology* **392**, 614-629 (2009).

119 De Groot, P. W. & Klis, F. M. The conserved PA14 domain of cell wall-associated fungal adhesins governs their glycan-binding specificity. *Molecular microbiology* **68**, 535-537 (2008).

120 Domergue, R. *et al.* Nicotinic acid limitation regulates silencing of Candida adhesins during UTI. *Science* **308**, 866-870 (2005).

121 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 & development* **17**, 2245-2258 (2003).

122 Castaño, I. *et al.* Telomere length control and transcriptional regulation of subtelomeric adhesins in Candida glabrata. *Molecular microbiology* **55**, 1246-1258 (2005).

123 Skrzypek, M. S. *et al.* The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. *Nucleic Acids Research* **45**, D592-D596 (2017).

124 Thierry, A., Dujon, B. & Richard, G.-F. Megasatellites: a new class of large tandem repeats discovered in the pathogenic yeast Candida glabrata. *Cellular and molecular life sciences* **67**, 671-676 (2010).

125 Christiaens, J. F. *et al.* Functional divergence of gene duplicates through ectopic recombination. *EMBO reports* **13**, 1145-1151 (2012).

126 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).

127 Hwang-Wong, E. *TWO EARLY PROCESSES DURING INFECTION BY THE FUNGAL PATHOGEN CANDIDA GLABRATA: ADHERENCE AND ALKALINIZATION*, Johns Hopkins University, (2016).

128 Linde, J. *et al.* Defining the transcriptomic landscape of Candida glabrata by RNA-Seq. *Nucleic acids research* **43**, 1392-1406 (2015).

129 Lindahl, U., Couchman, J., Kimata, K. & Esko, J. Proteoglycans and sulfated glycosaminoglycans. (2015).

130 Nelson, D. L., Lehninger, A. L. & Cox, M. M. *Lehninger principles of biochemistry*. (Macmillan, 2008).

131 Funderburgh, J., Caterson, B. & Conrad, G. Distribution of proteoglycans antigenically related to corneal keratan sulfate proteoglycan. *Journal of Biological Chemistry* **262**, 11634-11640 (1987).

132 Trowbridge, J. M. & Gallo, R. L. Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **12**, 117R-125R (2002).

133 DeAngelis, P. L. Microbial glycosaminoglycan glycosyltransferases. *Glycobiology* **12**, 9R-16R (2002).

134 Bi, Y., Patra, P. & Faezipour, M. in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE.* 3949-3952 (IEEE).

135 Mattson, J. M., Turcotte, R. & Zhang, Y. Glycosaminoglycans contribute to extracellular matrix fiber recruitment and arterial wall mechanics. *Biomechanics and modeling in mechanobiology* **16**, 213-225 (2017).

136 Rostand, K. S. & Esko, J. D. Microbial adherence to and invasion through proteoglycans. *Infection and immunity* **65**, 1 (1997).

137 Bartlett, A. H. & Park, P. W. Proteoglycans in host–pathogen interactions: molecular mechanisms and therapeutic implications. *Expert reviews in molecular medicine* **12** (2010).

138 Green, J. V. *et al.* Heparin-binding motifs and biofilm formation by Candida albicans. *The Journal of infectious diseases* **208**, 1695-1704 (2013).

139 Kock, C., Dufrêne, Y. F. & Heinisch, J. J. Up against the wall: is yeast cell wall integrity ensured by mechanosensing in plasma membrane microdomains? *Applied and environmental microbiology* **81**, 806-811 (2015).

140 Levin, D. E. Regulation of cell wall biogenesis in Saccharomyces cerevisiae: the cell wall integrity signaling pathway. *Genetics* **189**, 1145-1175 (2011).

141 Heinisch, J. J., Lorberg, A., Schmitz, H. P. & Jacoby, J. J. The protein kinase C-mediated MAP kinase pathway involved in the maintenance of cellular integrity in Saccharomyces cerevisiae. *Molecular microbiology* **32**, 671-680 (1999).

142 Kamada, Y., Jung, U. S., Piotrowski, J. & Levin, D. E. The protein kinase C-activated MAP kinase pathway of Saccharomyces cerevisiae mediates a novel aspect of the heat shock response. *Genes & development* **9**, 1559-1571 (1995).

143 Jacoby, J., Nilius, S. & Heinisch, J. A screen for upstream components of the yeast protein kinase C signal transduction pathway identifies the product of the SLG1 gene. *Molecular and General Genetics MGG* **258**, 148-155 (1998).

144 Ketela, T., Green, R. & Bussey, H. Saccharomyces cerevisiae Mid2p Is a Potential Cell Wall Stress Sensor and Upstream Activator of thePKC1-MPK1 Cell Integrity Pathway. *Journal of bacteriology* **181**, 3330-3340 (1999).

145 Rajavel, M., Philip, B., Buehrer, B. M., Errede, B. & Levin, D. E. Mid2 is a putative sensor for cell integrity signaling in Saccharomyces cerevisiae. *Molecular and cellular biology* **19**, 3969-3976 (1999).

146 Mamoun, C. B., Beckerich, J.-M., Gaillardin, C. & Kepes, F. Disruption of YHC8, a member of the TSR1 gene family, reveals its direct involvement in yeast protein translocation. *Journal of Biological Chemistry* **274**, 11296-11302 (1999).

147 Vay, H. A., Philip, B. & Levin, D. E. Mutational analysis of the cytoplasmic domain of the Wsc1 cell wall stress sensor. *Microbiology* **150**, 3281-3288 (2004).

148 Lommel, M., Bagnat, M. & Strahl, S. Aberrant processing of the WSC family and Mid2p cell surface sensors results in cell death of Saccharomyces cerevisiae O-mannosylation mutants. *Molecular and cellular biology* **24**, 46-57 (2004).

149 Dupres, V. *et al.* The yeast Wsc1 cell surface sensor behaves like a nanospring in vivo. *Nature chemical biology* **5**, 857-862 (2009).

150 Hutzler, F., Gerstl, R., Lommel, M. & Strahl, S. Protein N-glycosylation determines functionality of the Saccharomyces cerevisiae cell wall integrity sensor Mid2p. *Molecular microbiology* **68**, 1438-1449 (2008).

151 Heinisch, J. J., Dupres, V., Wilk, S., Jendretzki, A. & Dufrêne, Y. F. Single-molecule atomic force microscopy reveals clustering of the yeast plasma-membrane sensor Wsc1. *PLoS One* **5**, e11104 (2010).

152 Levin, D. E. Cell wall integrity signaling in Saccharomyces cerevisiae. *Microbiology and molecular biology reviews* **69**, 262-291 (2005).

153 Straede, A. & Heinisch, J. J. Functional analyses of the extra-and intracellular domains of the yeast cell wall integrity sensors Mid2 and Wsc1. *FEBS letters* **581**, 4495-4500 (2007).

154 Petkova, M. I., Pujol-Carrion, N., Arroyo, J., García-Cantalejo, J. & de la Torre-Ruiz, M. A. Mtl1 is required to activate general stress response through Tor1 and Ras2 inhibition under conditions of glucose starvation and oxidative stress. *Journal of Biological Chemistry* **285**, 19521-19531 (2010).

155 Petkova, M. I., Pujol-Carrion, N. & de la Torre-Ruiz, M. A. Mtl1 O-mannosylation mediated by both Pmt1 and Pmt2 is important for cell survival under oxidative conditions and TOR blockade. *Fungal genetics and biology* **49**, 903-914 (2012).

156 Kock, C., Arlt, H., Ungermann, C. & Heinisch, J. J. Yeast cell wall integrity sensors form specific plasma membrane microdomains important for signalling. *Cellular microbiology* **18**, 1251-1267 (2016).

157 Dupres, V., Heinisch, J. r. J. & Dufrêne, Y. F. Atomic force microscopy demonstrates that disulfide bridges are required for clustering of the yeast cell wall integrity sensor Wsc1. *Langmuir* **27**, 15129-15134 (2011).

158 Letunic, I., Doerks, T. & Bork, P. SMART 7: recent updates to the protein domain annotation resource. *Nucleic acids research* **40**, D302-D305 (2011).

159 Rodicio, R., Buchwald, U., Schmitz, H.-P. & Heinisch, J. J. Dissecting sensor functions in cell wall integrity signaling in Kluyveromyces lactis. *Fungal Genetics and Biology* **45**, 422-435 (2008).

160 Cruz, S., Muñoz, S., Manjón, E., García, P. & Sanchez, Y. The fission yeast cell wall stress sensor-like proteins Mtl2 and Wsc1 act by turning on the GTP ase Rho1p but act independently of the cell wall integrity pathway. *MicrobiologyOpen* **2**, 778-794 (2013).

161 Zebisch, M., Jackson, V. A., Zhao, Y. & Jones, E. Y. Structure of the Dual-Mode Wnt Regulator Kremen1 and Insight into Ternary Complex Formation with LRP6 and Dickkopf. *Structure* **24**, 1599-1605 (2016).

162 Mao, B. *et al.* Kremen proteins are Dickkopf receptors that regulate Wnt/β-catenin signalling. *Nature* **417**, 664 (2002).

163 Stayner, C. & Zhou, J. Polycystin channels and kidney disease. *Trends in pharmacological sciences* **22**, 543-546 (2001).

164 Apweiler, R. *et al.* UniProt: the universal protein knowledgebase. *Nucleic acids research* **32**, D115-D119 (2004).

165 Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology* **157**, 105-132 (1982).

166 Gasteiger, E. *et al.* in *The proteomics protocols handbook* 571-607 (Springer, 2005).

167 Rehm, H. & Letzel, T. *Der Experimentator: Proteinbiochemie/Proteomics*. (Springer, 2010).

168 Sim, L., Groes, M., Olesen, K. & Henriksen, A. Structural and biochemical characterization of the N-terminal domain of flocculin L g-F lo1p from S accharomyces pastorianus reveals a unique specificity for phosphorylated mannose. *The FEBS journal* **280**, 1073-1083 (2013).

169 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).

170 Nonaka, M. *et al.* Synthetic di-sulfated iduronic acid attenuates asthmatic response by blocking T-cell recruitment to inflammatory sites. *Proceedings of the National Academy of Sciences* **111**, 8173-8178 (2014).

171 de Paz, J. L. *et al.* Profiling heparin–chemokine interactions using synthetic tools. *ACS chemical biology* **2**, 735-744 (2007).

172 de Paz, J. L., Noti, C., Böhm, F., Werner, S. & Seeberger, P. H. Potentiation of fibroblast growth factor activity by synthetic heparin oligosaccharide glycodendrimers. *Chemistry & biology* **14**, 879-887 (2007).

173 Olson, S. T., Halvorson, H. & Björk, I. Quantitative characterization of the thrombin-heparin interaction. Discrimination between specific and nonspecific binding models. *Journal of Biological Chemistry* **266**, 6342-6352 (1991).

174 Rajalingam, D., Kathir, K. M., Ananthamurthy, K., Adams, P. D. & Kumar, T. K. S. A method for the prevention of thrombin-induced degradation of recombinant proteins. *Analytical biochemistry* **375**, 361-363 (2008).

175 Maestre-Reyna, M. Structural and functional studies of mucin-interacting adhesion domains from Candida glabrata and Helicobacter pylori. (2011).

176 Bamford, C. V., Nobbs, A. H., Barbour, M. E., Lamont, R. J. & Jenkinson, H. F. Functional regions of Candida albicans hyphal cell wall protein Als3 that determine interaction with the oral bacterium Streptococcus gordonii. *Microbiology* **161**, 18-29 (2015).

177 Nobbs, A. H., Vickerman, M. M. & Jenkinson, H. F. Heterologous expression of Candida albicans cell wall-associated adhesins in Saccharomyces cerevisiae reveals differential specificities in adherence and biofilm formation and in binding oral Streptococcus gordonii. *Eukaryotic cell* **9**, 1622-1634 (2010).

178 Basak, T. *et al.* Comprehensive characterization of glycosylation and hydroxylation of basement membrane collagen IV by high-resolution mass spectrometry. *Journal of proteome research* **15**, 245-258 (2015).

179 Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research*, gku340 (2014).

180 Goossens, K. V. *et al.* Molecular mechanism of flocculation self-recognition in yeast and its role in mating and survival. *MBio* **6**, e00427-00415 (2015).

181 Ielasi, F. S., Decanniere, K. & Willaert, R. G. The epithelial adhesin 1 (Epa1p) from the human-pathogenic yeast Candida glabrata: structural and functional study of the carbohydrate-binding domain. *Acta Crystallographica Section D: Biological Crystallography* **68**, 210-217 (2012).

182 Bernfield, M. *et al.* Functions of cell surface heparan sulfate proteoglycans. *Annual review of biochemistry* **68**, 729-777 (1999).

183 Hazen, K. C. Participation of yeast cell surface hydrophobicity in adherence of Candida albicans to human epithelial cells. *Infection and immunity* **57**, 1894-1900 (1989).

184 Panagoda, G., Ellepola, A. & Samaranayake, L. Adhesion of Candida parapsilosis to epithelial and acrylic surfaces correlates with cell surface hydrophobicity. *Mycoses* **44**, 29-35 (2001).

185 Silva-Dias, A. *et al.* Adhesion, biofilm formation, cell surface hydrophobicity, and antifungal planktonic susceptibility: relationship among Candida spp. *Frontiers in microbiology* **6**, 205 (2015).

186 Luo, G. & Samaranayake, L. Candida glabrata, an emerging fungal pathogen, exhibits superior relative cell surface hydrophobicity and adhesion to denture acrylic surfaces compared with Candida albicans. *Apmis* **110**, 601-610 (2002).

187 Dieu-Nosjean, M.-C. *et al.* Macrophage inflammatory protein 3α is expressed at inflamed epithelial surfaces and is the most potent chemokine known in attracting Langerhans cell precursors. *Journal of Experimental Medicine* **192**, 705-718 (2000).
188 Rabenstein, D. L. Heparin and heparan sulfate: structure and function. *Natural product reports* **19**, 312-331 (2002).
189 Varki, A., Cummings, R. & Esko, J. Proteins That Bind Sulfated Glycosaminoglycans.
190 Lin, P.-H., Sinha, U. & Betz, A. Antithrombin binding of low molecular weight heparins and inhibition of factor Xa. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1526**, 105-113 (2001).
191 Olson, S. *et al.* Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement. *Journal of Biological Chemistry* **267**, 12528-12538 (1992).
192 Lindahl, U. *et al.* Structure of the antithrombin-binding site in heparin. *Proceedings of the National Academy of Sciences* **76**, 3198-3202 (1979).
193 De Agostini, A. I., Watkins, S. C., Slayter, H. S., Youssoufian, H. & Rosenberg, R. D. Localization of anticoagulantly active heparan sulfate proteoglycans in vascular endothelium: antithrombin binding on cultured endothelial cells and perfused rat aorta. *The Journal of cell biology* **111**, 1293-1304 (1990).
194 Kjellén, L. & Lindahl, U. Proteoglycans: structures and interactions. *Annual review of biochemistry* **60**, 443-475 (1991).
195 Wernersson, S. & Pejler, G. Mast cell secretory granules: armed for battle. *Nature Reviews Immunology* **14**, 478 (2014).
196 Walenga, J. M. *et al.* Fondaparinux: a synthetic heparin pentasaccharide as a new antithrombotic agent. *Expert opinion on investigational drugs* **11**, 397-407 (2002).
197 Büller, H. R. *et al.* Fondaparinux or enoxaparin for the initial treatment of symptomatic deep venous thrombosis: a randomized trial. *Annals of Internal Medicine* **140**, 867-873 (2004).
198 Choay, J. *et al.* Structure-activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor Xa activity. *Biochemical and biophysical research communications* **116**, 492-499 (1983).
199 Parish, C. R. The role of heparan sulphate in inflammation. *Nature Reviews Immunology* **6**, 633 (2006).
200 Wight, T. N., Kinsella, M. G. & Qwarnström, E. E. The role of proteoglycans in cell adhesion, migration and proliferation. *Current opinion in cell biology* **4**, 793-801 (1992).
201 Afratis, N. *et al.* Glycosaminoglycans: key players in cancer cell biology and treatment. *The FEBS journal* **279**, 1177-1197 (2012).
202 Díaz-Nido, J., Wandosell, F. & Avila, J. Glycosaminoglycans and β-amyloid, prion and tau peptides in neurodegenerative diseases. *Peptides* **23**, 1323-1332 (2002).
203 Shieh, M.-T., WuDunn, D., Montgomery, R. I., Esko, J. D. & Spear, P. G. Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *The Journal of cell biology* **116**, 1273-1281 (1992).
204 WuDUNN, D. & Spear, P. G. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *Journal of virology* **63**, 52-58 (1989).

205 Leistner, C. M., Gruen-Bernhard, S. & Glebe, D. Role of glycosaminoglycans for binding and infection of hepatitis B virus. *Cellular microbiology* **10**, 122-133 (2008).
206 Capila, I. & Linhardt, R. J. Heparin–protein interactions. *Angewandte Chemie International Edition* **41**, 390-412 (2002).
207 Bucior, I., Mostov, K. & Engel, J. N. Pseudomonas aeruginosa-mediated damage requires distinct receptors at the apical and basolateral surfaces of the polarized epithelium. *Infection and immunity* **78**, 939-953 (2010).
208 Guzman-Murillo, M. A., Ruiz-Bustos, E., Ho, B. & Ascencio, F. Involvement of the heparan sulphate-binding proteins of Helicobacter pylori in its adherence to HeLa S3 and Kato III cell lines. *Journal of medical microbiology* **50**, 320-329 (2001).
209 Srivastava, A. *et al.* Full-length extracellular region of the var2CSA variant of PfEMP1 is required for specific, high-affinity binding to CSA. *Proceedings of the National Academy of Sciences* **107**, 4884-4889 (2010).
210 Rathore, D. *et al.* Molecular mechanism of host specificity in Plasmodium falciparum infection role of circumsporozoite protein. *Journal of Biological Chemistry* **278**, 40905-40910 (2003).
211 Clausen, T. M. *et al.* Structural and functional insight into how the Plasmodium falciparum VAR2CSA protein mediates binding to chondroitin sulfate A in placental malaria. *Journal of Biological Chemistry* **287**, 23332-23345 (2012).
212 Brandhorst, T. T. *et al.* Structure and function of a fungal adhesin that binds heparin and mimics thrombospondin-1 by blocking T cell activation and effector function. *PLoS pathogens* **9**, e1003464 (2013).
213 Carrasco, M. *et al.* Evaluation of a triple-lumen central venous heparin-coated catheter versus a catheter coated with chlorhexidine and silver sulfadiazine in critically ill patients. *Intensive care medicine* **30**, 633-638 (2004).
214 Mermel, L. A. *et al.* Clinical practice guidelines for the diagnosis and management of intravascular catheter-related infection: 2009 Update by the Infectious Diseases Society of America. *Clinical infectious diseases* **49**, 1-45 (2009).
215 Gale, C. A. *et al.* Linkage of adhesion, filamentous growth, and virulence in Candida albicans to a single gene, INT1. *Science* **279**, 1355-1358 (1998).
216 Parsons, C. L., Mulholland, S. & Anwar, H. Antibacterial activity of bladder surface mucin duplicated by exogenous glycosaminoglycan (heparin). *Infection and immunity* **24**, 552-557 (1979).
217 Parsons, C. L., Stauffer, C. & Schmidt, J. D. Bladder-surface glycosaminoglycans: an efficient mechanism of environmental adaptation. *Science* **208**, 605-607 (1980).
218 Wessels, M. R., Goldberg, J. B., Moses, A. E. & DiCesare, T. J. Effects on virulence of mutations in a locus essential for hyaluronic acid capsule expression in group A streptococci. *Infection and immunity* **62**, 433-441 (1994).
219 Chung, J. Y. *et al.* Role of capsule in the pathogenesis of fowl cholera caused by Pasteurella multocida serogroup A. *Infection and immunity* **69**, 2487-2492 (2001).
220 Svanborg-Edén, C. *et al.* Bacterial virulence versus host resistance in the urinary tracts of mice. *Infection and immunity* **55**, 1224-1232 (1987).
221 Kuhn, D. M. & Vyas, V. K. The Candida glabrata adhesin Epa1p causes adhesion, phagocytosis, and cytokine secretion by innate immune cells. *FEMS yeast research* **12**, 398-414 (2012).

222 Singh, B., Fleury, C., Jalalvand, F. & Riesbeck, K. Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS microbiology reviews* **36**, 1122-1180 (2012).
223 Kamhi, E., Joo, E. J., Dordick, J. S. & Linhardt, R. J. Glycosaminoglycans in infectious disease. *Biological Reviews* **88**, 928-943 (2013).
224 Ascencio, F., Fransson, L. & WadstrÖum, T. Affinity of the gastric pathogen Helicobacter pylori for the N-sulphated glycosaminoglycan heparan sulphate. *Journal of medical microbiology* **38**, 240-244 (1993).
225 Jacobsen, I. *et al.* Candida glabrata persistence in mice does not depend on host immunosuppression and is unaffected by fungal amino acid auxotrophy. *Infection and immunity* **78**, 1066-1077 (2010).
226 Jacobsen, I. D., Große, K., Berndt, A. & Hube, B. Pathogenesis of Candida albicans infections in the alternative chorio-allantoic membrane chicken embryo model resembles systemic murine infections. *PLoS One* **6**, e19741 (2011).
227 Rönnberg, E., Melo, F. R. & Pejler, G. Mast cell proteoglycans. *Journal of Histochemistry & Cytochemistry* **60**, 950-962 (2012).
228 Mulloy, B., Lever, R. & Page, C. Mast cell glycosaminoglycans. *Glycoconjugate journal* **34**, 351-361 (2017).
229 Maruyama, H., Yabu, Y., Yoshida, A., Nawa, Y. & Ohta, N. A role of mast cell glycosaminoglycans for the immunological expulsion of intestinal nematode, Strongyloides venezuelensis. *The Journal of Immunology* **164**, 3749-3754 (2000).
230 Michelacci, Y. M. & Petricevich, V. L. Glycosaminoglycan profile of peritoneal and bone marrow-derived macrophages. Changes associated with macrophage activation. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **100**, 617-625 (1991).
231 Phan, Q. T. *et al.* Als3 is a Candida albicans invasin that binds to cadherins and induces endocytosis by host cells. *PLoS biology* **5**, e64 (2007).
232 Salazar, S. B. *et al.* Comparative genomic and transcriptomic analyses unveil novel features of azole resistance and adaptation to the human host in Candida glabrata. *FEMS yeast research* **18**, fox079 (2017).
233 Vale-Silva, L., Beaudoing, E. & Sanglard, D. Comparative Genomics of Two Sequential Candida glabrate Clinical Isolates. *G3: Genes, Genomes, Genetics*, g3. 117.042887 (2017).
234 Håvelsrud, O. E. & Gaustad, P. Draft genome sequences of Candida glabrata isolates 1A, 1B, 2A, 2B, 3A, and 3B. *Genome announcements* **5**, e00328-00316 (2017).
235 Vina, A. Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading Trott, Oleg; Olson, Arthur J. *Journal of Computational Chemistry* **31**, 455-461 (2010).
236 Kresina, T. F. & Miller, E. J. Isolation and characterization of basement membrane collagen from human placental tissue. Evidence for the presence of two genetically distinct collagen chains. *Biochemistry* **18**, 3089-3097 (1979).
237 Sage, H., Woodbury, R. G. & Bornstein, P. Structural studies on human type IV collagen. *Journal of Biological Chemistry* **254**, 9893-9900 (1979).
238 Sage, H. & Bornstein, P. Characterization of a novel collagen chain in human placenta and its relation to AB collagen. *Biochemistry* **18**, 3815-3822 (1979).
239 Tsilibary, E. *et al.* Heparin type IV collagen interactions: equilibrium binding and inhibition of type IV collagen self-assembly. *Journal of Biological Chemistry* **263**, 19112-19118 (1988).

240 Stratford, M. & Assinder, S. Yeast flocculation: Flo1 and NewFlo phenotypes and receptor structure. *Yeast* **7**, 559-574 (1991).
241 Biran, R. & Pond, D. Heparin coatings for improving blood compatibility of medical devices. *Advanced drug delivery reviews* **112**, 12-23 (2017).
242 Vogt, A. M. *et al.* Release of sequestered malaria parasites upon injection of a glycosaminoglycan. *PLoS pathogens* **2**, e100 (2006).
243 Lycke, E., Johansson, M., Svennerholm, B. & Lindahl, U. Binding of herpes simplex virus to cellular heparan sulphate, an initial step in the adsorption process. *Journal of General Virology* **72**, 1131-1137 (1991).
244 Gupta, R., Jung, E. & Brunak, S. Prediction of N-glycosylation sites in human proteins. 2004. *Ref Type: Unpublished Work* (2016).
245 Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment1. *Journal of molecular biology* **302**, 205-217 (2000).
246 Cherry, J. M. *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research*, gkr1029 (2011).
247 Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785-786 (2011).
248 Linding, R., Russell, R. B., Neduva, V. & Gibson, T. J. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic acids research* **31**, 3701-3708 (2003).
249 Scholz, F. Maximum likelihood estimation. *Encyclopedia of statistical sciences* (1985).
250 Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution* **33**, 1870-1874 (2016).
251 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).
252 Steentoft, C. *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO journal* **32**, 1478-1488 (2013).
253 Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic acids research* **43**, W389-W394 (2015).
254 Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic acids research* **40**, e115-e115 (2012).
255 Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31**, 3406-3415 (2003).
256 Mullis, K. B. The unusual origin of the polymerase chain reaction. *Scientific American* **262**, 56-65 (1990).
257 Don, R., Cox, P., Wainwright, B., Baker, K. & Mattick, J. 'Touchdown'PCR to circumvent spurious priming during gene amplification. *Nucleic acids research* **19**, 4008 (1991).
258 Braman, J., Papworth, C. & Greener, A. in *In Vitro Mutagenesis Protocols* 31-44 (Springer, 1996).
259 Vogelstein, B. & Gillespie, D. Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences* **76**, 615-619 (1979).
260 Desjardins, P. & Conklin, D. NanoDrop microvolume quantitation of nucleic acids. *Journal of visualized experiments: JoVE* (2010).

261 Morrison, D. Transformation in Escherichia coli: cryogenic preservation of competent cells. *Journal of bacteriology* **132**, 349-351 (1977).

262 Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *Journal of molecular biology* **53**, 159-162 (1970).

263 Bimboim, H. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* **7**, 1513-1523 (1979).

264 Bertani, G. STUDIES ON LYSOGENESIS I.: The Mode of Phage Liberation by Lysogenic Escherichia coli1. *Journal of bacteriology* **62**, 293 (1951).

265 Tartof, K. Improved media for growing plasmid and cosmid clones. *Bethesda Res Lab Focus* **9**, 12 (1987).

266 Sherman, F., Fink, G. R. & Hicks, J. B. Laboratory course manual for methods in yeast genetics. (1986).

267 Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Current genetics* **16**, 339-346 (1989).

268 Amberg, D. C., Burke, D. J. & Strathern, J. N. Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, 2005 Edition (Cold Spring). (2005).

269 Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598 (1975).

270 Porath, J. & Flodin, P. Gel filtration: a method for desalting and group separation. *Nature* **183**, 1657-1659 (1959).

271 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature* **227**, 680 (1970).

272 Bennett, J. & Scott, K. Quantitative staining of fraction I protein in polyacrylamide gels using Coomassie brilliant blue. *Analytical biochemistry* **43**, 173-182 (1971).

273 Gill, S. C. & Von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry* **182**, 319-326 (1989).

274 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 & Cytochemistry* **51**, 1699-1712 (2003).

275 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).

276 Heimburg-Molinaro, J., Song, X., Smith, D. F. & Cummings, R. D. Preparation and analysis of glycan microarrays. *Current protocols in protein science*, 12.10. 11-12.10. 29 (2011).

277 Geissner, A., Anish, C. & Seeberger, P. H. Glycan arrays as tools for infectious disease research. *Current opinion in chemical biology* **18**, 38-45 (2014).

278 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).

279 Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-666 (1958).

280 Wüthrich, K. The way to NMR structures of proteins. *Nature Structural & Molecular Biology* **8**, 923 (2001).

281 Lipfert, J. & Doniach, S. Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307-327 (2007).

282 Kühlbrandt, W. Cryo-EM enters a new era. *Elife* **3**, e03678 (2014).

283 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).

284 Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *Journal of Biological Chemistry* **265**, 22059-22062 (1990).

285 Bernstein, F. C. *et al.* The protein data bank. *European Journal of Biochemistry* **80**, 319-324 (1977).

286 Jancarik, J. & Kim, S. H. Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of applied crystallography* **24**, 409-411 %@ 0021-8898 (1991).

287 Stevens, R. C., Yokoyama, S. & Wilson, I. A. Global efforts in structural genomics. *Science* **294**, 89-92 (2001).

288 Saridakis, E. & Chayen, N. E. Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion. *Protein Science* **9**, 755-757 (2000).

289 Khurshid, S., Saridakis, E., Govada, L. & Chayen, N. E. Porous nucleating agents for protein crystallization. *Nature protocols* **9**, 1621-1633 (2014).

290 Drenth, J. *Principles of protein X-ray crystallography*. (Springer Science & Business Media, 2007).

291 Pernot, P. *et al.* Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of synchrotron radiation* **20**, 660-664 (2013).

292 Kabsch, W. Xds. *Acta Crystallographica Section D: Biological Crystallography* **66**, 125-132 (2010).

293 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).

294 Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D: Biological Crystallography* **66**, 133-144 (2010).

295 Evans, P. Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography* **62**, 72-82 (2006).

296 Collaborative, C. P. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760 (1994).

297 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**, 235-242 (2011).

298 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674 (2007).

299 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221 (2010).

300 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography* **53**, 240-255 (1997).

301 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126-2132 (2004).

302 DeLano, W. L. The PyMOL molecular graphics system. (2002).

303 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* **98**, 10037-10041 (2001).

304 Bordoli, L. *et al.* Protein structure homology modeling using SWISS-MODEL workspace. *Nature protocols* **4**, 1-13 (2009).
305 Fichtner, L., Schulze, F. & Braus, G. H. Differential Flo8p-dependent regulation of FLO1 and FLO11 for cell–cell and cell–substrate adherence of S. cerevisiae S288c. *Molecular microbiology* **66**, 1276-1289 (2007).
306 Bester, M. C., Pretorius, I. S. & Bauer, F. F. The regulation of Saccharomyces cerevisiae FLO gene expression and Ca2+-dependent flocculation by Flo8p and Mss11p. *Current genetics* **49**, 375-383 (2006).
307 McLean, I. W. & NAKANE, P. K. Periodate-lysine-paraformaldehyde fixative a new fixative for immunoelectron microscopy. *Journal of Histochemistry & Cytochemistry* **22**, 1077-1083 (1974).
308 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 & development* **8**, 2974-2985 (1994).
309 Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. [6] Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in enzymology* **185**, 60-89 (1990).
310 Dubendorf, J. W. & Studier, F. W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of molecular biology* **219**, 45-59 (1991).
311 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).
312 Bessette, P. H., Åslund, F., Beckwith, J. & Georgiou, G. Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. *Proceedings of the National Academy of Sciences* **96**, 13703-13708 (1999).
313 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**, 753 (2012).
314 Brachmann, C. B. *et al.* Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *YEAST-CHICHESTER-* **14**, 115-132 %@ 0749-0503X (1998).
315 Rupniak, H. T. *et al.* Characteristics of four new human cell lines derived from squamous cell carcinomas of the head and neck. *Journal of the National Cancer Institute* **75**, 621-635 (1985).
316 Jacobsen, J., van Deurs, B., Pedersen, M. & Rassing, M. R. TR146 cells grown on filters as a model for human buccal epithelium: I. Morphology, growth, barrier properties, and permeability. *International journal of pharmaceutics* **125**, 165-184 (1995).
317 Nielsen, H. M. & Rassing, M. R. TR146 cells grown on filters as a model of human buccal epithelium: IV. Permeability of water, mannitol, testosterone and β-adrenoceptor antagonists. Comparison to human, monkey and porcine buccal mucosa. *International journal of pharmaceutics* **194**, 155-167 (2000).
318 Hernandez, R. & Rupp, S. Human epithelial model systems for the study of Candida infections in vitro: part II. Histologic methods for studying fungal invasion. *Host-Pathogen Interactions: Methods and Protocols*, 105-123 (2009).
319 Giard, D. J. *et al.* In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *Journal of the National Cancer Institute* **51**, 1417-1423 (1973).

320 Schaller, M., Zakikhany, K., Naglik, J. R., Weindl, G. & Hube, B. Models of oral and vaginal candidiasis based on in vitro reconstituted human epithelia. *Nature protocols* **1**, 2767-2773 (2006).

321 Gaffney, R. A. *et al.* Effect of vaginal fluid on adherence of type 1 piliated Escherichia coli to epithelial cells. *Journal of Infectious Diseases* **172**, 1528-1535 (1995).

322 Strus, M., Kukla, G., Rurańska-Smutnicka, D., Przondo-Mordarska, A. & Heczko, P. Surface properties of Lactobacillus strains. II. Adherence to tissue culture surfaces. *Medycyna doswiadczalna i mikrobiologia* **53**, 253-258 (2000).

323 Fogh, J., Fogh, J. M. & Orfeo, T. One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *Journal of the National Cancer Institute* **59**, 221-226 (1977).

324 Sambuy, Y. *et al.* The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell biology and toxicology* **21**, 1-26 (2005).