

Summary

Candida albicans is a common fungal pathogen of humans that resides in the gastrointestinal tracts of most healthy individuals. This commensal relationship is a complex interplay of candidial and human factors. However, in certain conditions, when the balance between the fungus and host is disturbed, infection called candidiasis may develop. One of the first steps during infection is attachment to the surface of host tissues, where microbial cells organize into communities known as biofilms. The adhesion is coordinated through specialized proteins, called adhesins. The agglutinin-like sequence (Als) proteins from *C. albicans* are members of a family involved in binding to human cells. These proteins possess a modular configuration consisting of three basic domains: the N-terminal part, the tandem repeat domain and the C-terminal region. Between the N-terminus and the tandem repeats, the threonine-rich region (T) is found and it is often associated with the N-terminal domain. The N-terminal region is responsible for binding to host cells. The epithelial and endothelial cells produce extracellular matrix, which plays a structural role by supporting the cells, but also performs other important metabolic and biochemical functions. To understand the molecular mechanisms of adhesion and protein interaction, the relationship between ligand and receptor has to be examined. This can be achieved by performing affinity studies.

In this work, the N-terminal, binding parts of Als1 and Als3 proteins were used to study the molecular mechanisms of protein interaction. For this purpose, the affinity of Als1p-N to the extracellular matrix (ECM) proteins such as fibronectin and laminin was measured. These ECM proteins were selected, because they are the main components of the extracellular matrix. Since the N-terminal domain of Als proteins contains two immunoglobulin-like domains and a sequence (IVIVATT) with high propensity to aggregate, we have also examined the self-association of Als1p-N through surface plasmon resonance and calculated the affinity for this interaction. The obtained affinity constants for all tested interactions were in the micromolar range, which is typical for most protein-protein interactions.

Samenvatting

Candida albicans is een veel voorkomende pathogene schimmel die zich in het maagdarmkanaal van de meeste gezonde mensen bevindt. In deze symbiose bestaat een complex samenspel van *Candida* en menselijke factoren. In bepaalde omstandigheden, wanneer de balans tussen de schimmel en de gastheer verstoord is, kan een infectie - candidiasis - ontstaan. Een van de eerste stappen tijdens de infectie is de hechting aan het oppervlak van de gastheerweefsels, waarbij microbiële cellen gegroepeerd worden tot biofilms. De hechting wordt mogelijk gemaakt door gespecialiseerde eiwitten, de zogenaamde adhesinen. De “agglutinin-like sequence” (Als) eiwitten van *C. albicans* zijn leden van een eiwitfamilie betrokken bij de binding aan menselijke cellen. Deze eiwitten hebben een modulaire configuratie die bestaat uit drie domeinen: het N-terminale deel, het tandem repeat domein en het C-terminale domein. Tussen het N-terminale domein en de tandem repeats, bevindt zich het threonine-rijke gebied (T) en dit wordt vaak geassocieerd met het N-terminale domein. Het N-terminale domein is verantwoordelijk voor de binding aan gastheercellen. De epitheel- en endotheelcellen van de gastheer produceren een extracellulaire matrix (ECM), die een structurele rol speelt in de stevigheid van de cellen, maar ook andere belangrijke metabolische en biochemische functies vervult. Om de moleculaire mechanismen van adhesie en eiwit interactie te begrijpen, moet de binding tussen de ligand en de receptor worden onderzocht. Hiervoor werden affiniteit studies uitgevoerd.

In dit onderzoek werden de N-terminale bindingsdomeinen van de eiwitten Als1 en Als3 gebruikt om de moleculaire mechanismen van eiwit-eiwit interacties te bestuderen. Hiertoe werd de affiniteit van Als1p-N voor eiwitten van de ECM zoals fibronectine en laminine gemeten. Deze ECM eiwitten werden geselecteerd omdat ze de belangrijkste componenten zijn van de extracellulaire matrix. Het N-terminale domein van de Als-eiwitten bevat twee immunoglobuline-achtige domeinen en een sequentie (IVIVATT) met sterke neiging tot aggregatie. Daarom hebben we ook de zelfassociatie van Als1p-N onderzocht door “Surface Plasmon Resonance” (SPR) en hebben we de affiniteit voor deze interactie berekend. De verkregen affiniteitsconstanten voor alle geteste interacties

bevonden zich in het micromolaire gebied, wat typisch is voor de meeste eiwit-eiwit interacties.

Menselijke cellen zijn bedekt door een glycoconjugaat laag, de glycocalyx. De meeste ziekteverwekkers zijn in staat complexe oligosacchariden aanwezig in het gastheerweefsel te herkennen. Er is echter weinig gekend over de interactie van *C. albicans* met koolhydraten van de gastheer. Om de carbohydraat specificiteit van Als1p-N te onderzoeken, werd de binding van het eiwit aan honderden glycanen aanwezig op een glycan array getest. De resultaten toonden voor het eerst aan dat Als1p-N specifiek kan binden aan fucose bevattende glycanen die gewoonlijk worden aangetroffen in bloedgroepantigenen. Vervolgens werd de affiniteit van Als1p-N voor fucose bepaald door SPR en de bindingsconstante bevond zich in het millimolaire gebied, typisch voor de meeste eiwit-koolhydraat interacties.

Evenzo werd de affiniteit tussen Als3p-N en de eerder genoemde ECM eiwitten gekwantificeerd. Bovendien werden de krachten nodig voor het verbreken van de binding tussen Als3p-N en fibronectine, bepaald met AFM (Atomic Force Microscope) krachtspectroscopie. Als3p-N werd ook gescreend op binding met koolhydraat-bevattende moleculen met behulp van de glycan-array-methode. De data van de glycan array onthulde een groot aantal koolhydraat bevattende liganden en de meest voorkomende structuur waarmee Als3p-N kon binden, was het N-acetyllactosamine (LacNAc) motief. Met biofysische technieken, werd de affiniteit tussen Als3p-N en Als1p-N en N-acetylglucosamine (GlcNAc), een bestanddeel van een LacNAc, bepaald. De affiniteit voor GlcNAc, geraamd door fluorescentie spectroscopie was in het millimolaire gebied. Bovendien werd de affiniteit van Als3p-N aan fucose gekwantificeerd door SPR en de schijnbare K_D was in het millimolaire gebied. Net als bij Als1p-N, werd de homotypische interactie voor Als3p-N geanalyseerd door SPR, en gevonden dat de affiniteit zich in het micromolaire gebied bevond. Deze resultaten leveren voor het eerst kwantitatieve informatie over Als-gemedieerde interacties met verschillende substraten.

Om de basis van de eiwit-eiwit interacties te begrijpen, zouden structurele studies moeten worden uitgevoerd. Aangezien de pogingen om zowel Als1p-N als N-Als3p te kristalliseren mislukten, werden verschillende lengtes van Als3p-N geproduceerd om te

zoeken naar de koolhydraat-bindingsplaats. Ondersteund door homologie modellen en “docking” van koolhydraten, werd een bindingsplaats voorspeld dat gelegen is in de fibrinogeen peptide-bindende holte, eerder beschreven voor Als9-2p-N.

Om de resultaten op moleculair niveau te bevestigen op cellulair niveau, werd de binding van *C. albicans* met Als1p en Als3p tot overexpressie, op menselijke epitheelcellen getest. In het bijzonder werd de invloed van fucose op de interactie van wild-type stammen en Als1p overexpressiestammen met menselijke cellen onderzocht. Slechts een klein inhiberend effect werd waargenomen bij de wild-type stam en geen effect bij de Als overexpressiestam. Deze resultaten bevestigen dat de Als-gemedieerde interacties mogelijk een combinatie van eiwit-koolhydraat en eiwit-eiwit interacties omvatten.

Tenslotte, gebaseerd op de resultaten van de glycan array screening, waarbij enkel de terminale koolhydraten werden gebruikt, werd een glycan interactie netwerk voor Als1p-N en N-Als3p, gegenereerd. Koolhydraten die specifiek kunnen binden aan beide Als eiwitten, werden gekoppeld aan eiwitten van het lichaam, dan aan organen en uiteindelijk aan menselijke lichaamssystemen. Als gevolg hiervan werd een groot aantal nieuwe doelen en nieuwe functies voor beiden adhesinen geïdentificeerd. De meeste potentiële liganden werden gevonden in het urogenitale en in het bloedsysteem. Deze resultaten breiden de kennis over de Als-gemedieerde processen uit, die een centrale rol in de pathogenese van *C. albicans* spelen.

Dagmara Szesztaowska Donohue

Human cells are covered by glycoconjugate coat called glycocalyx. Most pathogens have the capacity to specifically recognize complex oligosaccharides present on host tissue. However, little detail is available on *C. albicans* interaction with host carbohydrates. In search for the carbohydrate specificity, Als1p-N was screened for binding to hundreds of glycans presented in a glycan array format. The results revealed, for the first time, the specificity of Als1p-N towards fucose-containing glycans that are commonly found in blood group antigens. Subsequently, the affinity of Als1p-N for fucose was tested by surface plasmon resonance and revealed the apparent binding constant to be in the millimolar range, which is typical for most protein-carbohydrate interactions.

Likewise, the affinity between Als3p-N and previously mentioned ECM proteins was quantified. Additionally, the unbinding forces required for the rupture of the link between single Als3p-N and fibronectin molecules were obtained using AFM (Atomic Force Microscope) force spectroscopy. The Als3p-N was also screened for carbohydrate-binding partners using the glycan array method. The glycan array data revealed a large set of carbohydrate ligands and the most common carbohydrate structure that was bound by Als3p-N was N-acetyllactosamine (LacNAc) motif. Using biophysical techniques, the affinity of Als3p-N and Als1p-N to N-acetylglucosamine (GlcNAc), which is a constituent of a LacNAc, was estimated. The affinity towards GlcNAc, as estimated by fluorescence spectroscopy was in the millimolar range. Additionally, the affinity of Als3p-N to fucose was quantified by surface plasmon resonance and the apparent K_D was in the millimolar range. Similarly to Als1p-N, the homotypic interaction for Als3p-N was analyzed, by surface plasmon resonance, and showed the affinity in micromolar range. These results provide, for the first time, the quantitative information regarding the Als-mediated interactions with different substrates.

In order to understand the basis of protein-protein interaction, structural studies should be performed. Since the attempts to crystallize both Als1p-N and Als3p-N were unsuccessful, different lengths of Als3p-N were produced to search for the carbohydrate-binding site. Supported by homology modeling that involved carbohydrate docking, one binding site was predicted that was situated in the fibrinogen peptide-binding cavity previously described for the Als9-2p-N.

To confirm the results obtained at the molecular level, the binding of *C. albicans* cells overexpressing whole-length Als1p and Als3p to human epithelial cells was tested. In particular, the influence of fucose on the interaction of wild-type strain and Als1p strains with human cells was investigated. Only a slight inhibitory effect for the wild-type and no effect on Als-mediated interaction were observed. These results confirm that the Als-mediated interactions possibly involve a combination of protein-carbohydrate and protein-protein interactions.

Finally, based on the results obtained from glycan screenings, using only terminal carbohydrates, a glycan interaction network for Als1p-N and Als3p-N was generated. Carbohydrates that were bound specifically by both Als proteins were linked to proteins, then organs and finally human body systems. As a result, a large set of new targets and novel functions were identified for these adhesins. Most potential ligands were found in the urogenital system and in the hemic system. These results extend the knowledge over the Als-mediated processes, which play a central role in *C. albicans* pathogenesis.