The Perfect Adhesive

Stephen A. Klotz and Peter N. Lipke

Department of Medicine, University of Arizona, Tucson, Arizona, and Department of Biology, Brooklyn College of the City University of New York, Brooklyn, New York

Corresponding Author: Stephen A. Klotz; sklotz@u.arizona.edu, 1501 N. Campbell Ave., Tucson, AZ 85724

Candida albicans is a eukaryotic commensal of man that survives on mucous membranes. The fungus attaches to the host surface and other microorganisms through Als and other proteins on the cell surface. Following adhesion the Als proteins initiate a process leading to cellular aggregation, thus forming small microcolonies of fungi. Mature Als proteins possess four regions, an Immunoglobulin-like region, a Threonine-rich region that forms amyloids, a Tandem Repeat region with hydrophobic interacting surfaces and a Stalk region. Als adhesion begins as a receptor-ligand interaction with a target peptide. Cellular aggregation occurs later, the modular hydrophobic Tandem Repeat domains associating with one another while amyloid is forming in the Threonine-rich region. Formation of amyloid leads to high-strength, stable associations between cells in the fungal aggregate. Als proteins contribute to make C. albicans a highly successful commensal microorganism and a worrisome opportunistic pathogen.

Keywords: Candida albicans, Als proteins, adhesion, adherence

1. Introduction

Candida albicans is a commensal microorganism of humans and other vertebrates that infrequently causes opportunistic infections in its hosts. Human commensal microorganisms are preponderantly prokaryotes, Candida albicans being an exception. It is one of a few eukaryotes commonly isolated in low numbers from mucous membranes of healthy people. Opportunistic infections in humans often follow immune suppression or neutropenia [1]. Like all infectious diseases adhesion of a microbe to host tissue must occur before infection can become established. Infection with Candida species is invariably opportunistic, where a breach in host defenses occurs such as the loss of mucous membrane epithelial cells or a reduction in the number of circulating polymorphonuclear leukocytes. Candida albicans is the most commonly encountered fungus in hospital microbiology laboratories and is the fourth most common cause of positive blood cultures in hospitalized patients in North America [2]. Furthermore, 20% or more of cases of candidemia are polymicrobial involving one or more bacteria or several species of Candida [3, 4]. Lastly, candidiasis is the most common opportunistic infection in HIV-infected individuals [5-7].

Candida albicans is well known in the clinical realm for its ability to “stick” to damaged human tissue whether that be mucous membrane, bladder epithelium, heart valve or retinal vessel. It possesses cell surface proteins known as Agglutinin-like sequence (Als) proteins that are necessary for continuous fungal residence in the dense microbial milieu on mucous membranes. Evolutionary pressures leading to the present day Als proteins were undoubtedly exerted on C. albicans in its commensal state. The functional role of the Als proteins in the attachment of C. albicans to host tissue is the subject of this review.

2. Candida albicans Als proteins

2.1 General observations on adhesion

As with many microorganisms, the ability of C. albicans to adhere to surfaces is presumed to be an important survival trait [8, 9]. The fungus elaborates biofilm after adhesion and aggregation occur. Biofilm is the preferred mode of social behavior and protects its members against host defense mechanisms such as neutrophils with their contents of calprotectin and defensins. King et al. demonstrated the prototypical C. albicans adhesion phenotype [10] of yeast cells covering buccal epithelial cells and adhering to one another forming clumps of cells (aggregation). Adhesion followed by aggregation has since been documented in numerous studies involving a myriad of biological targets. When C. albicans encounters an anchored protein, cell, or tissue, the fungus adheres and then aggregates.

The molecular explanation of this phenotype, i.e., adhesion and aggregation, awaited the discovery of Als protein function. Within a year, two functionally different Als proteins, Als5p [11] and Als1p [12] were described. Genome sequencing has shown that most strains have eight ALS alleles, and each locus shows high heterozygosity [13]. Thus, there are potentially 16 different Als proteins that can be expressed on any cell and the different loci are expressed under different conditions. Adhesion is an intrinsic property of all Als proteins. However, the various members of the Als protein family mediate several modes of adhesion, for example, adhesion alone, both adhesion and aggregation, or aggregation alone. Which specific “sticky” behaviors are exhibited in any individual circumstance depends on the Als proteins expressed and the nature of target proteins.
2.2 The Als tool kit: the regions of Als proteins
Like many adhesion proteins, Als proteins are mosaics of several types of domains, and each type of domain may be present in multiple copies [13]. All Als proteins possess a signal piece at the N-terminus allowing for passage of the protein to the cell surface.

2.2.1 Immunoglobulin (Ig)-like region
The N-terminal region is the Ig-like region, a globular region containing ligand binding sites. This region initiates adhesion of yeast cells to biological targets [14-16]. The Ig-like domains have structures similar to bacterial adhesins and invasins of the immunoglobulin superfamily, and the region consists of three tandem β-sheet domains homologous to the Ig-like region of α-agglutinin in S. cerevisiae. The region possesses cysteine and tryptophan residues in the canonical positions of Ig-like domains. Each domain has two apposed β-sheets of three to five β-strands, with each pair of sheets bound together by the hydrophobic interaction of the amino acid side chains and disulfide bonds. Amino acid residues in the loops of Als1p and Als5p between β-strands bind ligands and determine the specificity of these Als proteins [16-18].

The prevalence of β-sheet domain structures appears to result from their intrinsic acid stability, which is required in environments typical for growth of Candida albicans [19]. For example, the fungus grown in an industrial fermenter results in extreme acidification of the medium, from pH 1 to 2. Although CD studies show unfolding of Als5p at these low pH values, the unfolding is reversible as is acid inhibition of Als-mediated adhesion [20].

2.2.2 Threonine-rich region
This region, situated between the Ig-like and Tandem Repeat regions is predicted to fold into a single domain [16]. Like the Ig-like domains, it is an all β-sheet structure. Modeling algorithms predict that the Threonine-rich region folds as a single Ig-like domain, although it lacks the characteristic cysteine and tryptophan residues [16]. The region may be necessary for determining the configuration of the Ig-like region and support its binding specificity. Similar regions are not found in the homologous adhesin, α-agglutinin of S. cerevisiae or the Epa adhesins in C. glabrata [13].

The Threonine-rich region is the most strongly conserved part of Als proteins. The synonymous mutation rate for this region matches that of the rest of the protein, implying that it is of similar age as the rest of the protein, but the rate of non-synonymous substitutions is extremely low, and most of the allowed substitutions are among highly similar amino acids [21]. This implies strong evolutionary selection for retention of structure and function of this region. The Threonine-rich region of Als5p is involved in aggregation [14]. In most Als proteins this region consists of 127 amino acids, whereas, in Als7p and Als9p there are 128 amino acids. Interestingly, Als7p and Als9p do not aggregate after adhesion [22, 23].

About 28% of the region’s residues are threonine, denoting potential for extensive H-bonding and amyloid formation. A highly conserved heptapeptide found in the Threonine-rich region forms characteristic amyloid fibers as shown by transmission electron microscopy [21, 24]. The fibers bind amyloidophilic dyes which intercalate between the β-sheets of amyloid. Furthermore, large fragments from Als5p (residues 20-431, 20-664, or 20-1351) formed characteristic amyloid fibers and bound amyloid dyes [21]. Similar sequences exist in many other yeast adhesins [25].
2.2.3 Tandem Repeat region

This region is composed of 36-amino acid repeats in tandem, with as few as two repeats in some Als5p molecules to as many as thirty-seven repeats in Als2p. The aggregation property of various *C. albicans* Als proteins is roughly proportional to the number of Tandem Repeats in the molecule [14] similar to findings for *S. cerevisiae* FLO1 [26]. For example, Als1p with 20 Tandem Repeats forms aggregates larger than those found with Als5p with only 5 Tandem Repeats [23]. Repeats in the Als molecules are similar. In most Als proteins 14 of the 36 residues are threonine and other β-branched amino acids interspersed with other residues common to protein loops and turns. Such amino acids have a strong preference for extended β-strand conformation [14, 27]. Modeling studies predict that each 36 amino acid residue repeat forms a compact three strand antiparallel β-sheet domain [28, 29]. In addition, many O-mannosylations in the repeats form a ring or corona of sugar residues on the surface of each domain [14, 28, 29]. The corona surrounds the exposed hydrophobic surface on each domain. Such structures are “classic” protein-protein interaction surfaces allowing for a broad range of interactions with other exposed proteins within the same molecule or adjacent molecules [27]. Other fungal adhesins such as the Flo lectins in *S. cerevisiae* and Eap1p in *C. glabrata* have Tandem Repeats so this motif is not unique to *C. albicans* Als proteins.

2.2.4 Stalk region

The carboxy-terminal region of Als proteins is the least conserved element [22] and yet the region constitutes over half the residue length with 35 to 55% of the amino acids being serine and threonine. In the few cases where yeast adhesins have been analyzed for carbohydrate content, all Stalk serine and threonine residues have been found to be O-glycosylated [13]. Thus, they are predicted to form an extended conformation [30]. Such an extended structure is consistent with the length of the Stalk visible in an electron photomicrograph of wall-bound α-agglutinin [31, 32]. The length allows the ligand-binding regions of Als proteins to be displayed far enough from the wall surface to be free to engage a ligand [32]. Deletion of some amino acids from this region results in inactive adhesins [15, 33, 34]. The Stalk region has substantial clusters of hydrophobic residues which may limit wall permeability [35]. The C-termini of Als proteins possess a sequence that specifies for the addition of a GPI anchor. The anchors initially integrate into membranes surrounding the lumen of vesicles. On the cell surface the anchors are cleaved, then transglycosylated to create covalent linkage to wall polysaccharides. Deletions of the GPI anchoring peptide piece cause unanchored Als proteins to be secreted [36].

<table>
<thead>
<tr>
<th>Function(s):</th>
<th>Deletion of region leads to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-like</td>
<td>Loss of adhesion</td>
</tr>
<tr>
<td>Threonine-rich</td>
<td>Loss of adhesion</td>
</tr>
<tr>
<td>Tandem Repeat</td>
<td>Loss of aggregation</td>
</tr>
<tr>
<td>Stalk</td>
<td>Loss of adhesion</td>
</tr>
</tbody>
</table>

Table 1 Properties of the Als regions.

2.3 The ALS gene family

The similarity of *C. albicans* ALS genes to the *Saccharomyces cerevisiae* α-agglutinin gene, SAG1, led to their denomination as Agglutinin-Like Sequences (ALS). The ALS family has eight members: ALS1, 2, 3, 4, 5, 6, 7, and 9 and each locus is often heterozygous [37]. ALS genes are found in other *Candida* species including *C. tropicalis* and *C. dublinsiensis* [38]. Both these species show adhesion and aggregation like *C. albicans* [10, 39]. ALS genes are not found in *Candida glabrata*, which adheres primarily through adhesins of the EPA family, but does not aggregate [23, 38, 40].

3. How do the proteins work?

3.1 Adhesion is intrinsic to Als proteins

The major regions and domains of Als proteins, with the exception of the Stalk, have adhesion activity resulting in an extremely broad range of protein-protein interactions mediated by these proteins. While specific ligand substructures bind to an individual region or binding site, the overall binding activity displayed by Als molecules is the sum of interactions of all binding sites. Thus, binding complexity increases as ligands increase in complexity from peptides to proteins to cells.
Adhesion in *Candida* has been measured by deletion of molecular components of Als proteins and the results of such assays have been recently reviewed [41]. Other assays measure activities of specific adhesins in isolation, either as surface-displayed cell-associated molecules or as soluble fragments with their properties assayed in biochemical binding assays. Both types of assays have been extensively reported for *C. albicans* Als adhesins, with complementary results.

There are three basic interactions. First is the specific binding to peptide ligands by the Ig-like region. This region mediates much of the adhesion to biological substrates [15, 16]. Second, non-specific hydrophobic interactions within the Tandem Repeats are important in aggregation and reinforcing the interaction of the Ig-like region and likely enhance binding to plastic surfaces. Third, amyloid interactions in the Thr-ric region increase Als avidity for adhesion through formation of Als multimers and likely enhance aggregation. The sum result of these activities leads to adhesion and aggregation to target peptides and proteins. For reasons that are not yet clear, it appears that adhesion often precedes aggregation [11].

Adhesion, like aggregation to be discussed later, is an intrinsic property of Als protein structure. This is proved by the observation that adhesion does not require cellular energy since heat-killed cells readily undergo adhesion and aggregation [20] as do cells treated with metabolic and signal transduction inhibitors [42]. Adhesion is a protein-protein interaction and is not inhibited by non-ionic detergents, EDTA, salt, carbohydrates or pH values between pH 4 to 9 [20]. Shear force actually increases aggregation. Hydrogen bond disruptants such as formamide and urea inhibit binding to various proteins and peptides and disrupt aggregates. Upon washing out the chemicals, cells spontaneously re-associate, adhere and aggregate.

Binding of Als proteins to peptides, proteins and cells is not competed by the addition of carbohydrates. This observation means that neither the ligand (target of Als protein) nor Als protein carbohydrate moieties are directly involved in the process of adhesion. However, carbohydrates in the fungal cell wall, particularly in the Tandem Repeats, reinforce the hydrophobic interactions through weak, non-specific complementary interactions [43].

A recent finding using single molecule force spectroscopy provided insight into the mechanics of Als adhesion [44]. Stretching a single Als5p molecule led to the unfolding of the individual Tandem Repeat domains. The modular protein structure was able to resist high mechanical force demonstrating the strength and toughness of the protein. The unfolding probability increased with the number of Tandem Repeat domains and correlated with yeast adhesion suggesting that the process is intrinsic in the mechanical properties of Als proteins. This finding may have broad implications for many adhesins as modular repeat structures are found in cell adhesion molecules of other pathogens [44].

### 3.1.1 Adhesion ligands: peptides and proteins

Three classes of peptide ligands are engaged by the Ig-like region: homopolymers of serine, threonine, or alanine or peptides with three successive amino acids that conform to a broad sequence motif of tφ+ and some natural sequences that do not fall into either class. Target peptides must include a sufficient number of exposed peptide bonds that are sterically available for adhesion [45]. Small peptides do not possess stable folded structures and therefore, the peptide backbone is readily accessible in these peptides. Immobilized tetramers of threonine residues serve as ligands, but smaller peptides do not. Denatured proteins are ideal ligands for adhesion [18]. A preference for unstructured peptides is consistent with studies demonstrating that *C. albicans* adheres to damaged proteins and cells over that of native proteins [46, 47]. This concept is consistent with clinical observations of candidal lesions occurring at damaged tissue where cleaved and/or denatured peptides occur [9].

*C. albicans* or *S. cerevisiae* expressing Als1p and Als5p bind to peptides containing the three amino acid structural motif of “tφ+”: a turn-like residue, a bulky hydrophobic residue, followed by a lysine or arginine [18]. This motif is present in multiple copies in most proteins, hence the ability to adhere to all proteins tested thus far. For example, bovine serum albumin has 16 such structures, only three of which are available on the surface of the native protein.
whereas, 13 more sites become available for adhesion when the protein is denatured. In all proteins investigated, the binding of *C. albicans* is greatest to denatured proteins, because denaturation exposes otherwise cryptic targets of adhesion. Peptides with sequences other than the three amino acid motif are also bound by Als proteins [17, 45].

Initial adhesion to target peptides is localized to the Ig-like domains, but the Threonine-rich and Tandem Repeat regions facilitate adhesion. In solid phase binding assays the presence of the Tandem Repeat region increased the maximum binding of Als5p for fibronectin about 1000-fold in solid phase binding assays over that of the Ig-like and Threonine-rich regions alone [14, 29].

3.2 Cellular aggregation: structure dictates function

Aggregation begins after adhesion of a fungal cell to target. Aggregates can be quite large and are only broken apart with H-bond perturbers such as formamide and 6M urea—treatments that disaggregate amyloids. After cells are disaggregated by a perturber and the chemical washed away, cells quickly reform into aggregates. Similar to adhesion, aggregation is inherent in protein structure and occurs spontaneously. It is not an energy requiring process. Aggregation does not require Ca$^{2+}$ and therefore, is different from the process of flocculation of *Saccharomyces* which requires Ca$^{2+}$. Tandem Repeats play an important role in aggregation [14]. For example, Als1p with 20 Tandem Repeats forms the largest microcolonies *in vitro*, whereas, Als5p with 5 Tandem Repeats forms much smaller microcolonies. Furthermore, aggregation increases in proportion to the number of repeats in Als1p [15] and Als5p [14]. *Saccharomyces cerevisiae* expressing Als5p loses the property of aggregation when Tandem Repeats are deleted. Interestingly, aggregation occurs when the Tandem Repeat region is expressed alone at the distal end of the stalk [14].

![Fig. 3](image_url)

**Fig. 3** Structure explains Als function [29]. A shows 3 Tandem Repeats extended with their polysaccharide cuffs (pink) and central hydrophobic plates. B shows the repeats in relaxed position. C demonstrates a *trans* interaction of 2 molecules from separate cells with repeats lining up to interact leading to cell to cell aggregation. Note the molecules may be either parallel or anti-parallel. D shows the two molecules now bound to one another through hydrophobic binding of the repeat cores with complementary weak binding being contributed by the hydrophilic cuffs of polysaccharide. Each repeat has exposed hydrophobic surfaces which interact between domains, as do the fringing polar mannosyl groups. Thus, there are interactions within individual Als proteins or between Als molecules on apposed cells.

Promote inter-cellular aggregation, whereas the repeats in other Als proteins tend to form more intra-molecular associations.

Aggregation, but not adhesion, is inhibited in the presence of amyloidophilic dyes [25, 42]. Extensive H-bonding occurs in Als proteins in the process of aggregation, as it does in amyloid formation. Rauceo et al. documented amyloid-like structures forming within cell aggregates [42] where a phase transition occurred over the yeast surface once they were incorporated into aggregates. Congo red dye and ANS bound to the cell surface of the yeasts that had undergone phase transition (detected by polarized light microscopy). Phase transition was detectable in real time and cell surfaces
became birefringent and hydrophobic denoting that surface proteins adopted an ordered array. This process takes approximately 30 minutes to complete.

4. Concluding remarks

Als proteins have great versatility and affect aspects of Candida biology in ways we are just beginning to discover. Als proteins are products of a large conserved gene family and play overlapping roles. Their expression is regulated by complex environmental cues. Deletion of some ALS genes results in compensatory up-regulation of others. Furthermore, the adhesins not only mediate adhesion, but cause changes in cell wall architecture [41]. Als proteins have a large repertoire of ligand binding specificities including peptides with a broad three amino acid motif [18, 45]. Binding is of moderate affinity to sterically accessible peptide bonds. In addition, Als adhesins interact with each other and other substrates through non-specific hydrophobic and hydrophilic interactions within the Tandem Repeats [14, 29]. The binding is secured by the formation of amyloid within the Threonine-rich regions. Some of these modes of adhesion illustrate new principles in cell adhesion, for example, amyloid formation and the non-specific, hydrophobic interactions within the Tandem Repeat regions leading to aggregation. The avidity for macromolecular ligands increases greatly due to the binding at multiple sites on individual Als proteins. Ligands can bind to sites on different Als molecules clustered on the cell surface through amyloid-like and Tandem-repeat interactions [21, 25, 29]. These properties are important for C. albicans commensal lifestyle, including formation of biofilm, resisting shear, and also in the pathogenesis of disease with this fungus. These multiple, novel activities make these proteins truly the perfect biological adhesive.

Acknowledgements. Work at Hunter College and Brooklyn College was supported by NIGMS SCORE grants S06 GM60654 and S06 GM76168, and SCI GM83765.

References


