Structure and conservation of amyloid spines from the *Candida albicans* Als5 adhesin including similarity to human LARKS

Nimrod Golan¹, Sergei Schwartz Perov¹, Meytal Landau^{1,2*}, and Peter N. Lipke^{3*}

¹ Department of Biology, Technion-Israel Institute of Technology, Haifa 3200003, Israel

² European Molecular Biology Laboratory (EMBL), Hamburg 22607, Germany

³ Biology Department, Brooklyn College of the City University of New York, New York, USA

*Correspondence to: PLipke@brooklyn.cuny.edu; mlandau@technion.ac.il

Running Title: Amyloid spines from a fungal adhesin

Fig. 1. Als5 Sequence analysis and 3D model.

Fig.2 Spine str.

Fig. 3. ThT and amyloid fibers formation.

Fig. 4 Seq & Cross- β conservation of the segments:

Fig. 5 Seq & Cross- β conservation.in the Ig/invasin and T regions of six species.

Supp Figure 1 Crystal structures of ³⁶⁹TSYVGV³⁷⁴

Supp Figure 2 evolutionary conservation map.

ABSTRACT

Candida Als family adhesins mediate adhesion to biological and abiotic substrates, as well as fungal cell aggregation and fungal-bacterial co-aggregation. The activity of at least two family members, Als5 and Als1, is dependent on amyloid-like protein aggregation that is initiated by shear force. Each Als adhesin has a ~300-residue N-terminal Ig-like/invasin region. The following 108-residue, low complexity, threonine-rich (T) domain unfolds under shear to expose a critical amyloid-forming segment ³²²SNGIVIVATTRTV³³⁴ at the interface between the Ig-like/invasin domain 2 and the T domain of *Candida albicans* Als5. Amyloid prediction programs identified six potential amyloidogenic sequences in the Ig/invasin region and three others in the T domain of *C. albicans* Als5. Peptides derived from four of these sequences formed fibrils that bound thioflavin T, the amyloid indicator dye, and three of these revealed atomic-resolution structures of cross- β spines. These are the first atomic-level structures for fungal adhesins. One of these segments, from the T domain, revealed kinked β -sheets, similarly to LARKS (Low-complexity, Amyloid-like, Reversible, Kinked segments) found in human functional amyloids. Based on the cross- β structures in Als proteins, we use evolutionary arguments to identify functional amyloidogenic sequences in other fungal adhesins. Thus, cross- β structures are often involved in fungal pathogenesis and potentially in antifungal therapy.

Importance. Fungal adhesins form cell-to-cell bonds in biofilms. Many of the cellular interactions are dependent on formation of amyloid-like cross- β protein aggregates. Such structures are called ‘functional amyloids’ because they perform physiological activities and they are dependent on the same types of protein interactions that form the more familiar amyloid

deposits in neurodegenerative diseases. We have identified sequence segments that form cross- β structures in the Als5 adhesin from the human pathogen *Candida albicans*. Such sequences are widespread among ALS family adhesins, including those from other human pathogens including *Candida auris*. Moreover, we revealed a structural similarity in a segment originating from Als5 threonine-rich low complexity region to human LARKS, pointing on a common structural motif coding for functional amyloids in different kingdoms of life.

INTRODUCTION

Potential amyloid-like segments in fungal adhesins are functional in the sense that they mediate microbial adhesion, aggregation, and biofilm formation (1-4). They are amyloid in the sense that they mediate protein assembly into structured β -rich aggregates and fibrils, the same molecular interactions as those in amyloids in neurodegenerative diseases. Amyloid definition is based on the formation of cross- β fibrils composed of tightly mated β -sheets and interdigitating side chains (5-7). This structure leads to formation of birefringent structures that bind amyloidophilic dyes like Congo red and Thioflavins. Many microbial aggregates and biofilms display these dye-binding properties, and several adhesins have been demonstrated to form amyloid fibers *in vitro* (4, 8, 9). Among the best characterized, gram negative curli are pili composed of proteins assembled through cross- β amyloid-like interactions, and they mediate bacterial adhesion and biofilm formation. Like classical amyloids, biofilm formation can be inhibited by anti-amyloid compounds (9, 10). Thus, the characteristics of microbial functional amyloid mirror those of the better-known pathological amyloid in neurodegenerative disease and serum amyloidoses.

Classical amyloids form fibers that have a characteristic cross- β pattern under X-ray diffraction (7). This pattern shows strong orthogonal reflections of β -strands and β -sheets, with the β -strands orthogonal to the fiber axis and the sheets parallel to the axis. The β -sheets are stacked, and at least one interface between β -sheets is composed of tightly packed interdigitated sidechains, interacting through van der Waals forces and anhydrous H-bonds. This close interdigitation and dry interface give rise to the name “steric zipper,” and the crystalline cross- β structures of short segments are called “amyloid spines”(7).

A second type of homologous crystalline assembly has been recently described: Low-complexity, Amyloid-like, Reversible, Kinked segments (LARKS)(11). These structures also have a cross- β like arrangement, but the β -strands are kinked, the sheet interactions are dependent on smaller interfaces, and the structures are less stable and more evanescent. These LARKs spines are common in low-complexity unstructured regions of proteins (11). Low complexity domains are regions in protein sequences that differ from the composition and complexity of most globular proteins. Such LARKS sequences are involved in RNA-dependent phase separations that influence cellular stress response ((11, 12). LARKS have not been previously reported in adhesins.

Among the best studied fungal adhesins are the paralogs Als1, Als3, and Als5 from *Candida albicans* (4, 12). Als1 and Als3 are important for fungal virulence in mice, and Als5 can attenuate immune response on the host (13, 14). Als adhesins are covalently attached to the cell wall glycans and are encoded at 8 loci in the genome (15, 16). They are typically 1200-2200 residues long (15). The N-terminal region is distal to the wall and consists of a 300-residue β -sheet-rich Ig-like/invasin region with two domains with Ig-like Greek key folds stabilized by disulfide bonds (17, 18). In each Als protein, the Ig-like/invasin region is followed by a 108-residue Threonine-

rich low-complexity domain (T domain) that is highly conserved in all paralogs and includes a strong amyloidogenic sequence GIVIVA (19). C-terminal to T domains are 5-30 copies of a 36-residue tandem repeat. This region mediates hydrophobic effect interactions with abiotic substrates and with homologous structures. The C-terminal 500-1000 residues of Als proteins are also low complexity and are rich in O-glycosylated Ser and Thr residues, as well as in N-glycosylated Asn-Xaa-Ser/Thr motifs. This region forms a long flexible stalk that extends from the surface of the wall. The proteins are synthesized with C-terminal glycosyl phosphatidyl inositol (GPI) anchors, which are subsequently modified to form covalent crosslinks to cell wall glucans (20-22).

Als5 and other fungal adhesins have characteristics of functional amyloids. Under extension force applied by liquid flow or atomic force microscopy (AFM) extension, Als adhesin T domains unfold and expose the amyloidogenic sequence GIVIVA from the interface between the Ig/invasin region and the T domain (4). Peptides with this sequence form amyloid fibers, as do soluble forms of Als5, but peptides of fragments with a non-amyloid substitution GINIVA do not form amyloids (8, 19). When anchored to the cell surface, exposure of the GIVIVA sequence leads to formation of high avidity patches of adhesins on the cell surface. These patches consist of adhesins aggregated through amyloid-like bonds that are birefringent and bind amyloidophilic dyes. Recent evidence also shows that cell-to-cell binding is due to amyloid-like interactions between cells for Als5 and Als1 (23-25). Thus, formation of amyloid-like interactions mediates strong bonding between fungal cells, especially in biofilms formed under flow. Furthermore, amyloid-like structures coat the surface of fungi in fatal invasive fungal infections caused by many species. In superficial and invasive candidiasis, Als proteins are major components of these surface

amyloid-like structures (14, 26). The fungal surface amyloid attenuates macrophage response to the invading fungi (13). Thus, extensive amyloid-like interactions are common on adhering fungi and in fungal biofilms and can influence host response.

Amyloidogenic sequences in fungal adhesins have been widely predicted, and force-dependent formation of amyloid-like surface nanodomains has been demonstrated in several cases (4, 27, 28). However, no atomic-level structures of amyloid spines or steric zippers from this type of protein have been reported. Therefore, we screened the Als5 Ig-like/invasin and T domains for predicted amyloidogenic sequences and determined structural characteristics of several of these. We have also analyzed the conservation and potential *in vivo* roles of these sequence segments in Als5 and homologous proteins.

RESULTS

Potential amyloid core sequences in *C. albicans* Als5.

Sequence- and structure-based amyloid predictors use a variety of criteria including solubility and physicochemical properties, geometry, homology, and secondary structure propensity to predict occurrence of amyloidogenic core regions that can form cross- β aggregates and steric zippers. Accordingly, we scanned the primary structure of Als5 Ig/invasin and T domains (residues 20-433) with three amyloid prediction programs: TANGO, which has been useful for predictions in Als adhesins, is a thermodynamics state-based predictor (4, 29). AmylPred 2 is a consensus method based on eleven predictors (30). Lastly, FiSH Amyloid detects co-occurrence patterns in sequence data based on machine-learning techniques (31). Nine segments scored above threshold in at least two of the three predictors, marked on the Als5 sequence in Figure 1A. Six segments are in the Ig-like/invasin region: ⁶⁵DTFILN⁷⁰, ¹³²IAFNVG¹³⁷, ¹⁵⁶NTVTFN¹⁶¹, ¹⁶⁸SIAVNF¹⁷³, ¹⁹⁶IATLYV²⁰¹, and ²⁵⁹FGISIT²⁶⁴, and three sequences are in the T domain: ³²⁴GIVIVA³²⁹, a section from the longer segment ³²²SNGIVIVATTRTV³³⁴ at the Ig/T domain interface, ³⁶⁹TSYVGV³⁷⁴, and ³⁸⁸TATVIV³⁹³.

Figure 1. Predicted amyloidogenic segments within Als5²⁰⁻⁴³³ sequence and modeled structure

(A) Annotated N-terminal residues 20-433 of the *Candida albicans* Als5 protein containing the ig-like/invasin and T domains, indicated by blue and yellow backgrounds, respectively. Segments

with strong amyloidogenic predictions are denoted with colored bars under the sequence: TANGO (29) in red, AmylPred 2 (30) in green, and FiSH Amyloid in blue (31). Segments for which the crystal structure was determined are marked magenta, and the known amyloid segment ³⁰³SNGIVIVATTRTV³¹⁵ is marked green. Disulfide bonds are illustrated by yellow dashed arrow lines. (B) A three-dimensional (3D) model of the Als5 (residues 20-433) predicted by AlphaFold2 (32). The Ig/invasin region is based on homologous structures for Als3 with the two Ig subdomains indicated, the T domain is an *ab initio* model. Coloring schemes are as in panel A.

Structures of amyloidogenic spines from Als5

We determined the atomic structures of ¹⁵⁶NTVTFN¹⁶¹, ¹⁹⁶IATLYV²⁰¹ and ³⁶⁹TSYVG³⁷⁴ segments (PDB IDs 6RHA, 6RHB, and 6RHD, respectively; Figure 2, Table 1 and Supplemental Table S1). The locations of the amyloid segments on the AlphaFold2 (32) predicted structure of Als5 are shown in Figure 1B. The crystal structure of the ¹⁵⁶NTVTFN¹⁶¹ spine segment from the Ig-like domain formed a canonical steric-zipper structure of tightly mated parallel β -sheets with a dry interface. The ¹⁹⁶IATLYV²⁰¹ spine segment, also from the Ig-like domain 2, revealed a partial mating between antiparallel β -sheets. The ³⁶⁹TSYVG³⁷⁴ spine segment from the T domain of Als5 exhibited an atypical amyloid structure, with antiparallel β -sheets and kinked β -strands (Fig. 2 and Supplemental Fig S1).

These three peptides induced thioflavin-T (ThT) fluorescence with a typical amyloid curve of lag time followed by rapid aggregation (Fig. 3A) and formed amyloid-like fibrils visualized by transmission electron microscopy (TEM) (Fig. 3B). The peptides ¹⁵⁶NTVTFN¹⁶¹ and ¹⁹⁶IATLYV²⁰¹ each formed abundant and long fibrils, while ³⁶⁹TSYVG³⁷⁴ formed abundant short plump

159 bundled fibrils. The longer peptide ³²²SNGIVIVATTRTV³³⁴ segment, known to be critical for
160 biological activity of the adhesin and crucial for fibrillation of soluble forms of Als5 protein,
161 formed long and relatively straight fibrils that also induced ThT fluorescence (4). This sequence
162 contains a strongly predicted amyloid core sequence ³²⁴GIVIVA³²⁹. Although we did not
163 successfully crystallize either the 6-residue or the 13-residue segments, the X-ray fiber diffraction
164 of ³²²SNGIVIVATTRTV³³⁴ showed a canonical amyloid cross- β pattern with orthogonal reflection
165 arches at 4.6 Å and ~9 Å (Fig. 2).

166

167 **Figure 2. Crystal structures and fiber diffraction of Als5 amyloidogenic segments**
168 High-resolution crystal structure of three predicted amyloidogenic segments of Als5 and the

169 known amyloid segment ³⁰³SNGIVIVATTRTV³¹⁵. The left panel of each structure is viewed down
 170 the fibril axis, with β -strands shown as ribbons and residues as sticks, while the right panel is
 171 viewed perpendicular to the fibril axis displaying β -strands running horizontally. Fibrils contain
 172 thousands of layers of β -strands, but only seven layers are shown here. Segment backbones are
 173 colored either gray or purple; heteroatoms are colored according to their atom type (nitrogen in
 174 blue, oxygen in red) and water molecules in cyan. The X-ray fiber diffraction pattern with major
 175 reflections labeled showing a cross- β pattern with orthogonal reflections at 4.6 Å and ~9 Å
 176 spacings. Some reflections originating from ice also appear.

177 **Sup Figure 1. Crystal structures of the T-domain ³⁶⁹TSYVGV³⁷⁴ segments**

178 The high-resolution crystal structure of ³⁶⁹TSYVGV³⁷⁴ is shown at different angles to better
 179 demonstrate its unique LARKS-like structure. The left panel is viewed down the fibril axis, with β -
 180 strands shown as ribbons and residues as sticks, while the middle and right panels respectively
 181 display a tilted view (middle) and a view perpendicular to the fibril axis. The different view
 182 demonstrate the unique packed anti-parallel β -sheets composed of twisted strands, reminiscing
 183 the LARKS structures of kinked β -strands. β -sheet carbons are colored either gray or purple, and

184 heteroatoms are colored according to their atom type (nitrogen in blue, oxygen in red) and water
185 molecules in cyan.
186

187

188 **Figure 3. Predicted segments exhibiting amyloidogenic properties**

189 (A) Segments for which the crystal structure was determined and the amyloid spines
190 ³⁰³SNGIVIVATTRTV³¹⁵ show fibrillation kinetics when incubated with ThT fluorescence dye.
191 Measurements were conducted in 50 mM Tris-HCl buffer, 20 mM ThT, and 300 μM of peptide.
192 The graph shows the average of three technical repeats, with error bars representing standard
193 deviations. (B) Electron micrographs of the four segments show fibril formation after 48 hours of
194 incubation. Scale bar represents 300 nm.

195

Because steric-zipper fibrils are unusual in that pairs of β -sheets mate more closely than the adjoining surfaces in other protein complexes, quantitative measures of amyloid stability are based on solvent-accessible surface area buried at the interface between the mating sheets, and shape complementarity indicating on the closeness of fit of two protein surfaces(33). The shape complementarity, inter-strand distance, and solvent-exposed surface area buried were calculated for Als5 spine structures and compared to the NNQQNY segment from yeast prion Sup35 steric zipper (PDB ID 1YJO) (34). The NNQQNY was chosen for this comparison as it shows one of the highest values of shape complementarity and surface area buried among steric zipper structures ((7)). The shape complementarity was similar for all Als5 spines segment, all lower than prion NNQQYY (Table S1). However, the T domain ³⁶⁹TSYVGV³⁷⁴ spine segment, displaying kinked β -strands, showed a significantly lower solvent-exposed surface area buried compared to the other structures (Table S1), indicating on a less stable structure that might also suggest reversible fibril formation.

The amyloid sequences are mapped onto an AlphaFold2 (32) based model of Als5 (Fig. 1B). The ¹⁵⁶NTVTFN¹⁶¹ and ¹⁹⁶IATLYV²⁰¹ segments in the Ig-like/invasin region are each in close proximity to a disulfide bonded Cys residue (Figs. 1). Therefore, it is likely that these sequences would not have the conformational freedom to lead to solvent exposure or interaction necessary for amyloid formation. In contrast, the amyloid-forming segments in the T region, which is conformationally variable *in vivo*, are not constrained by disulfides or other structured regions. Thus, the amyloid propensity in this region is likely to be expressed *in vivo*, as demonstrated before for ³²²SNGIVIVATRTV³³⁴ ((4).

Patterns of evolutionary conservation and amyloidogenic propensities of amyloid segments in fungal adhesins

Evolution can shift and alter the functions of traits in all biological systems. Selection pressure will work to conserve traits that possess an important beneficial function. It is also true for functional amyloids, where purifying selection acts to conserve the amyloidogenic potential, amino acid sequence, and accessible conformation; however, the amyloidogenic potential and amino acid sequence are not always conserved equally. Occasionally, amino acid sequences can change considerably without greatly affecting their amyloidogenic potential (35). In light of this difference, we have examined the evolutionary conservation patterns of the amyloidogenic segments correlated with the trajectory of the change in amyloidogenic potential from close to more distant homologs of *C. albicans* Als5.

To evaluate the conservation of the amyloidogenic segments in protein homologous to Als5 we used the ConSurf webserver (36). First, we searched for homologs for Als5 residues 20-433 which include the Ig-like/invasin and T domains. Seventy-one unique sequences of homologs with 35–95% sequence identity were retrieved from the UniProt (UniRef90) database. We excluded homologs with large gaps in the multiple sequence alignment. The segments ¹⁵⁶NTVTFN¹⁶¹, ¹⁹⁶IATLYV²⁰¹, ³²²SNGIVIVATTRTV³³⁴, ³²⁴GIVIVA³²⁹ and ³⁶⁹TSYVGV³⁷⁴ showed diverse pattern of sequence conservation across the Als5 homologs. In the Ig-like/invasin region, the sequence ¹⁵⁶NTVTFN¹⁶¹ was remarkably conserved across the homologs (Fig. 4A), with five strongly conserved positions (Supplemental Fig S2), implying a conserved functional or structural role. In contrast, ¹⁹⁶IATLYV²⁰¹ was not well conserved and showed low average conservation value with four highly variable positions in its sequence. In the T domain, the ³²²SNGIVIVATTRTV³³⁴ region

showed intermediate to above-average conservation. This observation is in agreement with previous reports showing that the functional amyloid-forming sequence ³²²SNGIVIVA³²⁹ is strongly conserved within the *C. albicans* ALS gene family (19, 25). The LARKS-like sequence ³⁶⁹TSYVGV³⁷⁴ also showed intermediate average conservation value with three highly conserved positions, with highest value at Gly³⁷³.

We also evaluated the conservation of amyloidogenic potential within the Als5p sequence homologs (Fig. 4B). The spine-forming segments ¹⁵⁶NTVTFN¹⁶¹, ¹⁹⁶IATLYV²⁰¹, ³²²SNGIVIVATTRTV³³⁴, and ³⁶⁹TSYVGV³⁷⁴ presented a wide range of differences in amyloidogenic potential calculated by three predictors: TANGO, AmylPred 2, and FiSH Amyloid. For further analyses, we selected as controls two segments with a predicted low amyloidogenic propensity, ¹⁷⁷TVDQSG¹⁸² and ²³⁸VNDWNH²⁴³, showing low and intermediate evolutionary conservation, respectively. The ³²²SNGIVIVATTRTV³³⁴ segment already implicated in amyloidogenic role of Als5 (4) showed a high amyloidogenic propensity in close homologs, which declined in more distant homologs. The LARKS-like sequence ³⁶⁹TSYVGV³⁷⁴, which is relatively conserved (Fig. 4A), shows the highest stability in maintaining amyloidogenic potential across homologs according to Amylpred2 and FiSH amyloid, with more fluctuations in propensity towards more distant homologs (Fig. 4B). In contrast, the Ig-like/invasin region segments ¹⁵⁶NTVTFN¹⁶¹ and ¹⁹⁶IATLYV²⁰¹ show a less clear pattern of amyloidogenic propensity across homologs, similar to the control sequences ¹⁷⁷TVDQSG¹⁸² and ²³⁸VNDWNH²⁴³. Control sequence ¹⁷⁷TVDQSG¹⁸² showed high fluctuations in amyloidogenic propensity, despite its low value in the *C. albicans* proteins. This finding corresponds to its low sequence conservation (Fig. 4). Overall, the ⁶⁹TSYVGV³⁷⁴ and ³²²SNGIVIVATTRTV³³⁴ segments in the Als5 T domain appear to have the most conserved

261 amyloidogenic propensity among examined segments, at least across close homologs of *C.*
262 *albicans*.

Figure 4. Amyloidogenic propensity and the evolutionary conservation of *C. albicans* Als5 segments.

(A) Evolutionary conservation of Als5 segments. The graph displays evolutionary conservation values for seven segments (three structured segments, the amyloid spine ³⁰³SNGIVIVATTRTV³¹⁵, and two control segments) from *C. albicans* Als5. These values were calculated based on a multiple sequence alignment of 71 homologs of Als5 generated using the ConSurf webserver ((36). The columns' colors correspond to the average evolutionary conservation score (1 being the least conserved and 9 being the most conserved). Error bar represents the standard deviation of the average value of the residues in the sequence. (B) The amyloidogenic propensities of the above seven segments in Als5 orthologs (homologs from different species). Fourteen representative orthologs of Als5p were selected to represent a diverse group (as described in the Method section). The orthologs are plotted on the X-axis by their sequence similarity to *C. albicans* (evaluated based on ConSurf *e*-value parameter) from the most to the least similar homolog. The Y-axis shows the average amyloidogenic propensity of each Als5 segment and their equivalent regions in the fourteen orthologs. The amyloidogenic values were calculated using three servers and displayed using different colors: TANGO (orange) (29, 37) AmylPred2 (green) ((30), FiSH Amyloid (blue) (31). The different methods have a different score range, thus in each graph, a horizontal dashed line indicates the amyloidogenic value of the relevant segment from *C. albicans* Als5.

283

284 **Supplemental Figure S2. Sequence conservation scores mapped on the Als5(20-433) sequence.**

285 Evolutionary conservation scores were calculated using the ConSurf webserver (36, 38) based on

286 a multiple sequence alignment and phylogenetic tree of 71 homologous proteins of the of N-

287 terminal region of *Candida albicans* Als5 (UniProt ID Q5A8T7), residues 20-433. The conservation

288 score of each residue is represented by a color scheme ranging from dark blue to dark purple

289 (indicating variable to conserved). Segments for which the crystal structure was determined are

290 marked with magenta boxes, the known amyloid segment ³⁰³SNGIVIVAT³¹TRTV³¹ is marked with a

291 bright green box, and predicted amyloidogenic segments which didn't crystalize are marked with

292 black boxes. Segments that were not predicted to have amyloidogenic potential and used as

293 control are marked with red boxes.

Amyloidogenic propensity within the Als5 Ig-like/invasin and T domains in six species

We compared the amyloidogenic propensity of the entire Als5 Ig-like/invasin and T domains in Als5 homologs from seven species: *Candida dubliniensis*, *Candida tropicalis*, *Candida maltosa*, *Candida viswanathii*, *Spathaspora passalidarum*, *Spathaspora sp. JA1*, and *Candida auris* (39). The amyloidogenic propensities predicted by AmylPred2 are summarized in Figure 5. Each of the adhesins was predicted to have amyloid-forming sequence at the equivalent region to the ³²²SNGIVIVATTRTV³³⁴ segment of *C. albicans* Als5 (marked green in Fig. 5), except for one homolog from *Spathaspora sp. JA1*. The equivalent region to the LARKS-like segment ³⁶⁹TSYVGV³⁷⁴ also showed a strong amyloidogenic propensity (marked dark yellow in Fig. 5). In the *C. auris* homolog, there was no equivalent sequence to TSYVGV, but other segments in this region (residues 343-370) were predicted to be amyloidogenic (marked purple). In the Ig-like/invasin region, the equivalent sequence to ¹⁹⁶IATLYV²⁰¹ (marked blue in Fig. 5) showed no amyloidogenic propensity in *C. maltosa* and *S. passalidarum* or *Spathaspora sp. JA1*, but this segment did show some amyloidogenic propensity in *C. auris*. The predicted amyloidogenic propensity of this region was the least conserved among the distant homologs examined in Figure 4B. The second segment in the Ig-like/invasin region, equivalent to ¹⁵⁶NTVTFN¹⁶¹, showed amyloidogenic propensity predictions (marked red in Fig. 5) across homologs except for *S. passalidarum* and *C. auris*. Overall, each ortholog shown in Figure 5 contains various regions with strong prediction for amyloidogenic propensity. Among the identified segments of Als5, we observed that the T-domain segments display more conserved amyloidogenic traits compared to segment in the Ig-like/invasin region.

Figure 5. Predicted segments with amyloidogenic propensities in the Als5 Ig-like/invasin and T domains of *C. albicans* and seven orthologs

The amyloidogenic propensity, calculated using the AmylPred2 server (30), of Als5 Ig-like/invasin and T domains of *C. albicans* and seven orthologs. These specific orthologs were chosen as a diverse group among the 71 homologs collected by the ConSurf webserver along with *C. auris*, as described in the Method section. Each Ortholog is identified by the name of the species, their Swissprot ID, residue range of the Ig-like/invasin and T domains and their similarity to *C. albicans* ranked among 71 homologs (determined using the ConSurf-generated multiple sequence

alignment). The regions with predicted amyloidogenic propensity are marked purple on the sequences, represented as a linear stretch with residue position marks. Regions equivalent to the *C. albicans* sequences for which the crystal structure was determined and the known amyloid segment ³⁰³SNGIVIVATTRTV³¹ are marked with different colors as indicated:¹⁵⁶NTVTFN¹⁶¹ (red), ¹⁹⁶IATLYV²⁰¹ (blue), and ³⁶⁹TSYVGV³⁷⁴ (yellow) and ³⁰³SNGIVIVATTRTV³¹⁵ (green).

DISCUSSION

The structures and analyses presented here support the predictions of functional amyloid formation for fungal adhesins. Of the segments identified as potentially forming amyloids, we argue on the bases of structural and evolutionary constraints that two of them, in the T domain, are likely to function *in vivo*. We also showed atomic-level cross- β spine structures with known and novel geometries, including one LARKS-like structure reminiscent of functional and reversible human amyloids (11, 40).

Amyloid predictions. Sequence- and structure- based amyloid predictors generally identify amyloidogenic or β -aggregation potential in many protein sequences. Indeed, systematic studies show that the majority of small peptides with high potential amyloidogenic sequences do form amyloids *in vitro* (7, 11, 41, 42). However, only few of the proteins containing these sequences form amyloids *in vivo*, probably because peptides are constrained by being part of a stable protein fold *in situ*. Formation of cross- β structures depends on conformation, solvent accessibility, and compatible geometry in flanking regions (5, 7, 42). Therefore, amyloids can often form only after proteins denature to expose the amyloidogenic sequences and allow flexibility in the peptide chains. In Als5, we identified 9 segments with high potential to form

amyloid-like cross- β structures. We argue here two of these sequences are known to or may be involved in functional amyloid formation, while 6 others probably do not form cross- β structures *in situ* in the protein. The arguments are based on structural and evolutionary arguments, and on positive evidence for functionality of one segment.

Potential amyloidogenic sequences in the Ig-like/invasin domain

Of the six potential amyloidogenic β -aggregating sequences in the Ig-like/invasin domain of Als5, three are within disulfide-stapled regions of the protein and so these segments are unlikely to be solvent-exposed and flexible enough to form cross- β structures (Figure 1). Of the three potential amyloidogenic sequences between that are not disulfide-constrained, ¹⁵⁶NTVFN¹⁶⁰ and ¹⁹⁶IATLYV²⁰¹ formed cross- β spines *in vitro*. However, these segments are unlikely to do so *in vivo*. Although the sequence ¹⁵⁶NTVFN¹⁶¹ is relatively highly evolutionary conserved (Fig. 4A and Supplemental Fig. S2), its TANGO amyloidogenic propensity is low and not conserved across homologs. It has greater potential in the other two predictors. (Figs. 4B and 5). In Ig/invasin domain II ¹⁵⁶NTVFN¹⁶¹ seals the edge of a β -sheet (Fig. 1), and so it contributes to the hydrophobic core of the fold. Therefore, this sequence is expected to be key for the high stability of this region, and this region of the protein unfolds only at high extension forces > 350 pN ((43, 44). In addition, ¹⁵⁶NTVFN¹⁶¹ is near disulfide-bonded Cys¹⁵⁰ as well as the domain I/II interface, which both constrain its flexibility and geometry in the native protein. Therefore, this sequence is unlikely to form a cross- β structure *in vivo* in the intact protein.

The sequence ¹⁹⁶IATLYV²⁰¹ is in a β -strand (Fig. 1B) and is poorly conserved in sequence (Fig. 4A). Its equivalent region is predicted to have amyloidogenic propensity in some of the homologs, but without the specific trend of amyloidogenic propensity conservation along relatively close

and distant homologs (Figs. 4B and 5). Therefore, it seems unlikely that this segment participates in the formation of functional β -aggregates in Als adhesins. The segment is extremely close to disulfide bonded Cys204, and so it is likely to be too geometrically constrained *in vivo* to form cross- β structures.

The four other potential amyloid sequences in the Ig-like/invasin region failed to form cross- β spines. Therefore, there is little support for formation of cross- β structures *in situ* within the Ig/invasin region of the protein.

Potential amyloidogenic sequences in the T domain

Two of the potential amyloid sequences in the T domain show conservation of their sequence and amyloidogenic propensity. Figures 4B and 5 demonstrate the strong cross- β potential of ³²²SNGIVIVATTRTV³³⁴ and the conservation of this propensity in this region across homologs. Similarly, the LARKS-like segment ³⁶⁹TSYVGV³⁷⁴ also shows high conservation of amyloidogenic propensity (Figs. 4B and 5) and conservation in sequence as well (Fig. 4A and Supplemental Fig S2). Both of these segments showed strong cross- β properties (Fig. 3). Also contrasting with the Ig/Invasin domain, this region is unconstrained by disulfide bonds and becomes unstructured under low shear stress. Thus, these sequences are likely to be exposed and relatively unconstrained *in vivo* (4, 43, 45). These features are in accord with the known T domain functional amyloid formation.

Functions of T domain amyloid sequences

The T domain and the sequence ³²²SNGIVIVATTRTV³³⁴ is essential for many activities of Als5 and its close paralog Als1 (4, 25). This sequence is necessary for strong cell-to-cell binding, as well as for clustering the adhesins on the cell surface. The cross- β nature of the interactions is supported

by amyloid-like properties including birefringence, Congo red and thioflavin binding. A single site mutation V326N reduces the TANGO β -aggregation potential by about 20-fold, and reduces most aggregation activity in Als5 without affecting ligand binding to the Ig-like/invasin domain (9). Furthermore, adhesion activity is enhanced in the presence of a homologous peptide, and activity is reduced in the presence of a homologous non-amyloid peptide with the Val->Asn mutation. These are properties expected from a sequence interacting through β -aggregation. The T region unfolds easily under shear, including under physiological shear stresses (4, 28). Thus, extension forces unfold the T domain, rendering it unstructured giving the conformational flexibility needed for formation of cross- β structures.

The ³⁶⁹TSYVGV³⁷⁴ potential amyloid sequence is located ~50 residues downstream in the T domain. This sequence should also be exposed when the domain unfolds under extension force (4, 43, 45, 46). It is conserved in sequence, and all homologs show similar amyloid-prone sequences at nearby positions (Figure 4A, Figure 5 yellow bars and Supplemental Figure S2). Under crystallization conditions, this peptide formed a structure like evanescent LARKS structures shown within low complexity regions (11). The conservation of amyloid potential implies that it may play a role in Als protein function, possibly to promote alignment of ³²⁴GIVIVA³²⁹ sequences in different proteins to facilitate formation of strong amyloid-like β -aggregates both *in cis* on a cell surface and *in trans* to form intercellular bonds. The third potential cross- β forming sequence in the T domain is ³⁸⁸TATVIV³⁹³, which is highly conserved in sequence and amyloid potential (Fig. 1 and Supplemental Fig. S2). However, we have no information on the relevance of this segment to Als5 function or to cross- β structures.

Conclusions

Atomic resolution structures show that sequences in Als5 can form amyloid-like cross- β aggregates. Of the 4 sequences for which we have structural information, two in the Ig-like/invasin region have characteristics of fortuitous amyloid-formers, *i.e.* these segments have key roles in a hydrophobic domain core, coupled with propinquity to disulfide bonds and poor conservation of amyloid potential. In contrast, two sequences in the T domain show characteristics of force-dependent functional amyloids: conservation of sequence and amyloid potential, as well as demonstrated roles in activity of the protein for the ³²⁴GIVIVA³²⁹ sequence of Als5 and Als1 (4, 25). These results thus demonstrate a role for conservation of amyloid potential as well as sequence in determining functionality of naturally occurring sequences that form β -aggregates. Thus, evolutionary considerations for both sequence and amyloid potential are valuable tools to identify functional and pathological amyloid-forming sequences. Conversely, the same arguments can be used to identify many of the sequences that have high amyloid potential but nonetheless do not form amyloid structures *in vivo*.

Table 1. Data collection and refinement statistics

	Als5 ₁₉₆ IATLYV ₂₀₁	Als5 ₁₅₆ NTVTFN ₁₆₁	Als5 ₃₆₉ TSYVGV ₃₇₃
	6RHB	6RHA	6RHD
Beamline	ESRF ID23 2	ESRF ID23 2	ESRF ID23 2
Date	September 7 th 2015	May 10 th 2015	December 11 th 2015
PDB ID			
Data collection			
Space group	P 2 ₁ 2 ₁ 2 ₁	P 1 2 ₁ 1	P 1 2 ₁ 1
Cell dimensions			
a, b, c (Å)	9.49 17.87 22.42	4.87 39.81 20.26	11.77 9.36 16.79
α , β , γ (°)	90.0 90.0 90.0	90.0 90.0 90.0	90.0 95.1 90.0
Wavelength (Å)	0.8729	0.8729	0.8729

Resolution (Å)	22.4- 1.26 (1.32-1.26)	39.8-1.60 (1.69-1.60)	16.73-1.20 (1.25-1.20)
R-factor observed (%)	24.6 (75.7)	27.0 (100.5)	20.0 (85.5)
^a R _{meas} (%)	25.8 (79.5)	28.0 (108.4)	20.4 (90.0)
<i>I</i> / σI	6.8 (2.9)	7.3 (2.2)	13.7 (2.5)
Test set size [%], selection	10, random	10, random	10, random
Total reflections	12416 (984)	14151 (893)	31912 (1290)
Unique reflections	1076 (90)	991 (126)	1178 (139)
Completeness (%)	91.1 (58.8)	98.2 (90.0)	96.2 (97.9)
Redundancy	11.5 (10.9)	14.3 (7.1)	27.1 (9.3)
^b CC1/2 (%)	99.5 (97.1)	97.5 (72.7)	99.9 (84.2)
Refinement			
Resolution (Å)	13.97-1.26 (1.41-1.26)	19.91-1.60 (1.64-1.60)	16.73-1.20 (1.23-1.20)
Completeness (%)	91.11 (76.6)	97.9 (77.9)	96.2 (98.8)
^c No. reflections	968 (218)	891 (54)	1059 (75)
^d R _{work} (%)	9.97 (17.5)	19.8 (29.0)	10.6 (21.7)
R _{free} (%)	10.39 (24.1)	21.3 (22.1)	12.6 (22.3)
No. atoms	49	103	51
Protein	48	Chain A: 49 Chain B: 49	44
Water	1	5	7
B-factors			
Protein	4.8	Chain A: 11.1 Chain B: 12.4	7.25
Water	7.2	19.4	24.1
R.m.s. deviations			
Bond lengths (Å)	0.007	0.014	0.011
Bond angles (°)	1.560	1.634	1.545
Clash score (67)	0.00	0.00	0.00
Molprobrity score (67)	0.5	0.5	0.5
Molprobrity percentile (67)	100th percentile	100th percentile	100th percentile
Number of xtals used for scaling	One crystal, one data set	Two crystals, one data set from each	One crystal, six data sets

Values in parentheses are for highest-resolution shell.

^(a) R-meas is a redundancy-independent R-factor defined in Reference (47).

^(b) CC_{1/2} is percentage of correlation between intensities from random half-datasets (48) .

^(c) Number of reflections corresponds to the working set.

^(d) Rwork corresponds to working set

Table S1. Features of the Als5 spine structures compared to the NNQQNY steric zipper structure

	Als5 196 ^{IATLYV} 201	Als5 156 ^{NTVTFN} 161	Als5 ^{369TSYVG} 374	Sup35 NNQQNY

Shape complementarity ^a	0.74	0.78	0.79	0.86
Inter-strand distance along the sheet	4.745 Å (9.49 Å between antiparallel strands)	4.87 Å	4.68 Å (9.36 Å between antiparallel strands)	4.87 Å
Area buried of one strand within pairs of sheets ^b	Strand 1: 563 Å ² Strand 2: 606 Å ² Average: 585 Å ²	589 Å ²	Strand 1: 581 Å ² Strand 2: 401 Å ² Average: 491 Å ²	624 Å ²

The values of shape complementarity, inter-strand distance and solvent exposed surface area buried calculated for the Als5 spine structures are compared with those of the NNQQNY segment from yeast prion Sup35 (PDB code 1YJO) (49) steric zipper structure. The NNQQNY was chosen for this comparison as it shows one of the highest values of shape complementarity and surface area buried among steric zipper structures (50).

^a Shape complementarity of 0 indicates no complementarity of the two surfaces and approaches 1 for atomic surfaces that fit perfectly together (34).

^b The solvent accessible surface area buried is the average area buried of one strand within two β -sheets (total area buried from both side is double the reported number). The surface area buried was calculated with Chimera (UCSF) (51) with default probe radius and vertex density are

1.4 Å and 2.0/Å², respectively. IATLYV and TSYVGV assemble into antiparallel β-sheets, hence each of the antiparallel β-strands forms a difference interface that was calculated separately.

Materials and Methods

Peptides and reagents

Peptide segments ¹⁵⁶NTVTFN¹⁶¹, ¹⁶⁸SIAVNF¹⁷³, ¹⁹⁶IATLYV²⁰¹, ³²⁴GIVIVA³²⁹ and ³⁶⁹TSYVGV³⁷⁴ from Als5 (UniProt accession number Q5A8T7) were synthesized at >98% purity and purchased from GL Biochem (Shanghai) Ltd. The peptides were synthesized with unmodified termini for crystallography or with fully capped termini (acetylated in the N-terminus and amidated in the C structural terminus) for fibrillation assays. Thioflavin T (ThT) was purchased from Sigma-Aldrich. Dimethyl-sulfoxide (DMSO) was purchased from Merck. Ultra-pure water was purchased from Biological Industries.

Computational prediction of amyloid spine segments

Amyloidogenic propensities of Als5 segments were predicted using combined information from several computational methods, including TANGO (29, 37), AmylPred 2 (30) and FiSH Amyloid (31). TANGO and FiSH Amyloid were used in their default settings, and AmylPred 2 was used with the default consensus option of only 4 of 9 predictors (TANGO and Amyloid Mutants were omitted from its prediction). We used TANGO, AmylPred 2, and FiSH Amyloid because they were the most sensitive and consistent predictors with quantifiable scores. The spine segments were defined as short segments of 6 residues, detected by at least two methods out of three at the epicenter of the predicted amyloidogenic area in the sequence (like in the case of ³⁶⁹TSYVGV³⁷⁴).

Only in the case of the segment ³²²SNGIVIVATTRTV³³⁴, which is known to be critical to fibrillation of the Als5 protein (9), we worked with both its full length and a shorter version of 6 residues.

Modelling of the Ig-like domain of Als5

The three-dimensional (3D) model was generated by AlphaFold2 (32) for the entire Als5 sequence from *Candida albicans*, taken directly from the UniProt accession number Q5A8T7. Using Chimera (51) we focused only on residues 20-433 which include the Ig-like/invasin and T domains. The per-residue confidence score (pLDDT) of the Ig-like domain ranged above 90 indicating a reliable prediction. For the T domain, the per-residue confidence score ranged between 70-90 indicating lower confidence in the prediction.

Calculation of evolutionary conservation using the ConSurf/CenSeq webserver

We used ConSurf (36, 38) with default settings to examine the evolutionary conservation patterns of *Candida albicans* Als5 residues 20-433 which include the Ig-like and T domains. The calculation was run using ConSeq (52) (no structure was provided). One hundred and twenty-one unique sequences of homologs with 35-95% sequence identity were retrieved from the UniRef90 database. Homologs with large gaps (more than 5 residues) in their alignment were filtered using Jalview2 multiple sequence alignment editing program ((53). The alignment of the remaining 71 homologs was used for the calculations presented here. The alignment is accessible at [Link1](#). The alignment of the seven orthologs used to generate the information presented in Fig. 5 is accessible at [Link2](#).

Amyloidogenic propensity correlated to the evolutionary conservation of the segments

Comparison between the segments ¹⁵⁶NTVTFN¹⁶¹, ¹⁹⁶IATLYV²⁰¹, ³⁶⁹TSYVGV³⁷⁴, and ³²²SNGIVIVATTRTV³³⁴ evolutionary conservation was done by averaging the amino acid

conservation scores attained from the ConSurf results for each segment. The segments ¹⁷⁷TVDQSG¹⁸² and ²³⁸VNDWNH²⁴³ were taken as controls due to their lack of amyloidogenic propensity and their different conservation values, namely low and high respectively. The 71 retrieved homologs were ranked by their sequence similarity to *C. albicans* Als5, with the lowest *e*-value corresponding to the closer homolog to *C. albicans* Als5. Out of the 71 homologs we selected one along every five ranked homologs (in descending order of similarity to *C. albicans* Als5) for a total of fourteen sequences representing diverse sequences in the multiple sequence alignment. We then calculated the amyloidogenic propensity for each of the fourteen homologs using TANGO, AmylPred 2, and FiSH Amyloid (Figs 4B and 5). The multiple sequence alignment was used, via Jalview2 (53), to identify equivalent regions the segments of *C. albicans* discussed here. We then calculated the average predicted amyloidogenic propensity values in these corresponding sequences in each of the 14 homologs. TANGO averages were calculated by dividing the sum of each residue's amyloidogenic propensity by its length. AmylPred 2 averages were calculated by assigning to each residue of the segments either a Boolean value of zero or one (based on consensus recognition of the method) and calculating the average similarly to TANGO averages. Using the default sliding window settings of the method, FiSH Amyloid averages were calculated by dividing the sum of the values assigned to all residues covering the segment by its length.

For the calculations of the amyloidogenic propensity of the entire Ig-like/invasin and T domains, we selected seven orthologs (homologs with a common ancestral parent in different species). Six of those were chosen along the ranked 71 homologs in gaps of 10 (starting from *C. dubliniensis* orthologs located five places away from the original *C. albicans* Als5). To these six, we added the

emerging pathogen *Candida auris*, which was not a part of the homologs collected by the ConSurf webserver due to low sequence identity to the *C. albicans* Als5p (54, 55). The multiple sequence alignment of the 71 homologs along with the *C. auris* sequence was obtained by re-producing the ConSurf search with similar parameters but lower sequence identity of 25-95%, and then filtering out the irrelevant homologs using Jalview2.

Thioflavin T kinetic assays. Thioflavin T (ThT) is a commonly used for identifying and investigating the formation of amyloid fibrils *in vivo* and *in vitro*. In the presence of ThT, fibrillation curves commonly show delayed nucleation followed by rapid aggregation. Fibrillation kinetics of the peptides ¹⁵⁶NTVTFN¹⁶¹, ¹⁹⁶IATLYV²⁰¹, ³⁶⁹TSYGVG³⁷⁴, and ³²²SNGIVIVATTRTV³³⁴ were monitored using ThT. All of the above peptides were synthesized with fully capped termini (acetylated in the N-terminus and amidated in the C structural terminus) to mimic its chemical nature in the full-length protein, and were dissolved to 10 mM in dimethyl sulfoxide (DMSO). Each of the freshly dissolved peptides was mixed with 50 mM Tris-HCl buffer with pH 7.3 and with 2 mM filtered ThT stock (made in ultra-pure water) to reach final concentrations of 300μM of peptide and 20 μM ThT in the final volume of 100 μl reaction. The reaction mixture was carried out in a black 96-well flat-bottom plate (Greiner bio-one) covered with a thermal seal film (EXCEL scientific) and incubated in a plate reader (CLARIOstar, BMG Labtech), at 37°C, with orbital shaking at 300 rpm, for 30 sec before each measurement. ThT fluorescence was recorded every two minutes using an excitation of 438±20 nm and an emission of 490±20 nm. The measurements were conducted in triplicates, and the entire experiment was repeated at least three times.

Transmission electron microscopy (TEM). TEM was used to visualize fibrils. Samples for TEM were taken directly from the ThT kinetic assay plate, which was left to incubate at 37 C for 3 days

with 300 rpm shaking in the plate reader. The TEM grids were prepared by applying 5 µl samples of each 300µM sample of peptide on 400 mesh copper grids with support films of Formvar/Carbon (Ted Pella), that were charged by high-voltage, alternating current glow-discharge (PELCO easiGlow, Ted Pella) immediately before use. Samples were allowed to adhere for 1 min followed by negative staining with 1% uranyl acetate for 1 min. Micrographs were recorded using a FEI Tecnai G2 T20 S-Twin transmission electron microscope at an accelerating voltage of 200 KeV at the MIKA electron microscopy center of the Department of Material Science & Engineering at the Technion.

Fiber X-ray diffraction of the SNGIVIVATTRTV segment from Als5

The peptide SNGIVIVATTRTV was dissolved to 10 mg/ml in ultra-pure water. A few microliters were applied between two sealed glass capillaries until completely dried. X-ray diffraction of the samples was collected at the micro-focused beam P14 at the high brilliance 3rd Generation Synchrotron Radiation Source at DESY: PETRA III, Hamburg, Germany.

Crystallization of Als5 segments

Peptides synthesized with free (unmodified) termini were used for crystallization experiments to facilitate crystal contacts. All peptides were dissolved to 10 mM in ultrapure water if possible, and in dimethyl sulfoxide (DMSO) in case they were water-insoluble. ¹⁹⁶IATLYV²⁰¹ was dissolved in 95% DMSO, ¹⁵⁶NTVTFN¹⁶¹ ¹⁶⁸SIAVNF¹⁷³ and ³²⁴GIVIVA³²⁹ were dissolved in 100% DMSO. ³⁶⁹TSYVGV³⁷⁴ was dissolved in 100% ultrapure water. Peptide solution drops (100 nL) were dispensed onto crystallization screening plates, using the Mosquito automated liquid dispensing robot (TTP Labtech, UK) located at the Technion Center for Structural Biology (TCSB). Crystallization using the hanging drop method, was performed in 96-well plates, with 100 µL

solution in each well. The drop volumes were 150-300 nL. All plates were incubated in a Rock imager 1000 robot (Formulatrix), at 293 °K. Micro-crystals grew after few days and were mounted on glass needles glued to brass pins. No cryogenic protection was used. Crystals were kept at room temperature prior to data collection. Structures were obtained from drops that were a mixture of the following peptide and reservoir solutions: ¹⁹⁶IATLYV²⁰¹: 10 mM IATLYV, 0.1 M tri-Sodium citrate pH 5.6 and 1.0 M Ammonium phosphate. ¹⁵⁶NTVTFN¹⁶¹: 10 mM NTVTFN, 0.1M HEPES 7.5pH, 0.8M NaH₂PO₄, 0.8M KH₂PO₄. ³⁶⁹TSYVGV³⁷⁴: 10 mM TSYVGV, 0.2M Sodium chloride, 0.1M Sodium acetate pH 4.6, 30% (v/v) 2-Methyl-2,4-pentanediol (MPD).

Structure determination and refinement

X-ray diffraction data were collected at 100°K, using 5° oscillation. The X-ray diffraction data were collected at the micro-focus beamline ID23-EH2 of the European Synchrotron Radiation Facility (ESRF) in Grenoble, France; wavelength of data collection was 0.8729 Å. Data indexation, integration and scaling were performed using XDS/XSCALE (56). Molecular replacement solutions for all segments were obtained using the program Phaser within the CCP4 suite (57-59). The search models consisted of geometrically idealized β-strands. Crystallographic refinements were performed with the program Refmac5 (59). Model building was performed with Coot (60) and illustrated with Chimera (51). There were no residues that fell in the disallowed region of the Ramachandran plot. Crystallographic statistics are listed in Table 1.

Calculations of structural properties

The Lawrence and Colman's shape complementarity index (34) was used to calculate the shape complementarity between pairs of sheets forming the dry interface (Supplementary Table S1). The buried surface area was calculated with Chimera (UCSF), with a default probe radius and

vertex density of 1.4 Å and 2.0/Å², respectively. The number of solvent accessible buried surface areas was calculated as the average area buried of one strand within two β-sheets (total area buried from both sides is there double the reported number in (Supplementary Table S1).

Acknowledgments

We acknowledge technical support provided by Yael Pazy-Benhar and Dikla Hiya at the Technion Center for Structural Biology (TCSB). We acknowledge support from Yaron Kauffmann from the MIKA electron microscopy center of the Department of Material Science & Engineering at the Technion, and Na'ama Koifman from the Russell Berrie Electron Microscopy Center of Soft Matter at the Technion, Israel. This research was supported by Israel Science Foundation (grant no. 2111/20), Israel Ministry of Science, Technology & Space (grant no. 78567), U.S.-Israel Binational Science Foundation (BSF) (grant no. 2017280), BioStruct-X, funded by FP7, and the iNEXT consortium of Instruct-ERIC. The synchrotron MX data collection experiments were performed at beamlines ID23-EH2 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, and at beamline P14, operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We are grateful to the teams at ESRF and EMBL Hamburg for their assistance.

Author contributions

ML and PNL conceived the study. ML provided all support. All authors generated data, analyzed results, and contributed to writing and editing.

Literature Cited

1. Tayeb-Fligelman E, Tabachnikov O, Moshe A, Goldshmidt-Tran O, Sawaya MR, Coquelle N, Colletier JP, Landau M. 2017. The cytotoxic *Staphylococcus aureus* PSMalpha3 reveals a cross-alpha amyloid-like fibril. *Science* 355:831-833. 10.1126/science.aaf4901.
2. Evans ML, Gichana E, Zhou Y, Chapman MR. 2018. Bacterial Amyloids. *Methods Mol Biol* 1779:267-288. 10.1007/978-1-4939-7816-8_17.
3. Jain N, Chapman MR. 2019. Bacterial functional amyloids: Order from disorder. *Biochim Biophys Acta Proteins Proteom* . S1570-9639(19)30104-9.
4. Lipke PN, Klotz SA, Dufrene YF, Jackson DN, Garcia-Sherman MC. 2017. Amyloid-Like beta-Aggregates as Force-Sensitive Switches in Fungal Biofilms and Infections. *Microbiol Mol Biol Rev* 82:10.1128/MMBR.00035-17. Print 2018 Mar. e00035-17.
5. Shewmaker F, McGlinchey RP, Wickner RB. 2011. Structural insights into functional and pathological amyloid. *J Biol Chem* 286:16533-16540. 10.1074/jbc.R111.227108.
6. Eisenberg D, Jucker M. 2012. The amyloid state of proteins in human diseases. *Cell* 148:1188-1203. 10.1016/j.cell.2012.02.022.
7. Eisenberg DS, Sawaya MR. 2017. Structural Studies of Amyloid Proteins at the Molecular Level. *Annu Rev Biochem* 86:69-95. 10.1146/annurev-biochem-061516-045104.
8. Ramsook CB, Tan C, Garcia MC, Fung R, Soybelman G, Henry R, Litewka A, O'Meally S, Otoo HN, Khalaf RA, Dranginis AM, Gaur NK, Klotz SA, Rauceo JM, Jue CK, Lipke PN. 2010. Yeast cell adhesion molecules have functional amyloid-forming sequences. *Eukaryot Cell* 9:393-404. 10.1128/EC.00068-09.
9. Garcia MC, Lee JT, Ramsook CB, Alsteens D, Dufrene YF, Lipke PN. 2011. A role for amyloid in cell aggregation and biofilm formation. *PLoS One* 6:e17632. 10.1371/journal.pone.0017632.
10. Cegelski L, Pinkner JS, Hammer ND, Cusumano CK, Hung CS, Chorell E, Aberg V, Walker JN, Seed PC, Almquist F, Chapman MR, Hultgren SJ. 2009. Small-molecule inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nat Chem Biol* 5:913-919. 10.1038/nchembio.242.
11. Hughes MP, Sawaya MR, Boyer DR, Goldschmidt L, Rodriguez JA, Cascio D, Chong L, Gonen T, Eisenberg DS. 2018. Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. *Science* 359:698-701. 10.1126/science.aan6398.
12. Hughes MP, Goldschmidt L, Eisenberg DS. 2021. Prevalence and species distribution of the low-complexity, amyloid-like, reversible, kinked segment structural motif in amyloid-like fibrils. *J Biol Chem* 297:101194. S0021-9258(21)00996-0.
13. Behrens NE, Lipke PN, Pilling D, Gomer RH, Klotz SA. 2019. Serum Amyloid P Component Binds Fungal Surface Amyloid and Decreases Human Macrophage Phagocytosis and Secretion of Inflammatory Cytokines. *MBio* 10:10.1128/mBio.00218-19. e00218-19.

- 637 14. Garcia-Sherman MC, Lundgren T, Sobonya R, Lipke PN, Klotz SA. 2015. A Unique Biofilm in
638 Human Deep Mycoses: Fungal Amyloid is bound by Host Serum Amyloid P Component. NPJ
639 Biofilms and Microbiomes Article number 15009. 10.1038/npjbiofilms.2015.9.
- 640 15. Hoyer LL. 2001. The ALS gene family of *Candida albicans*. Trends Microbiol 9:176-80.
- 641 16. Hoyer LL, Payne TL, Bell M, Myers AM, Scherer S. 1998. *Candida albicans* ALS3 and insights
642 into the nature of the ALS gene family. Curr Genet 33:451-459.
- 643 17. Lin J, Oh SH, Jones R, Garnett JA, Salgado PS, Rusnakova S, Matthews SJ, Hoyer LL, Cota E.
644 2014. The peptide-binding cavity is essential for Als3-mediated adhesion of *Candida albicans* to
645 human cells. J Biol Chem 289:18401-18412. 10.1074/jbc.M114.547877.
- 646 18. Salgado PS, Yan R, Taylor JD, Burchell L, Jones R, Hoyer LL, Matthews SJ, Simpson PJ, Cota E.
647 2011. Structural basis for the broad specificity to host-cell ligands by the pathogenic fungus
648 *Candida albicans*. Proc Natl Acad Sci U S A 108:15775-15779. 10.1073/pnas.1103496108.
- 649 19. Otoo HN, Lee KG, Qiu W, Lipke PN. 2008. *Candida albicans* Als adhesins have conserved
650 amyloid-forming sequences. Eukaryot Cell 7:776-782. EC.00309-07.
- 651 20. Lipke PN, Kurjan J. 1992. Sexual agglutination in budding yeasts: structure, function, and
652 regulation of adhesion glycoproteins. Microbiol Rev 56:180-94.
- 653 21. Kapteyn JC, Hoyer LL, Hecht JE, Muller WH, Andel A, Verkleij AJ, Makarow M, Van Den Ende
654 H, Klis FM. 2000. The cell wall architecture of *Candida albicans* wild-type cells and cell wall-
655 defective mutants. Mol Microbiol 35:601-611.
- 656 22. Roh DH, Bowers B, Riezman H, Cabib E. 2002. Rho1p mutations specific for regulation of
657 beta(1-->3)glucan synthesis and the order of assembly of the yeast cell wall. Mol Microbiol
658 44:1167-1183. 10.1046/j.1365-2958.2002.02955.x.
- 659 23. Dehullu J, Valotteau C, Herman-Bausier P, Garcia-Sherman M, Mittelviefhaus M, Vorholt JA,
660 Lipke PN, Dufrene YF. 2019. Fluidic Force Microscopy Demonstrates That Homophilic Adhesion
661 by *Candida albicans* Als Proteins Is Mediated by Amyloid Bonds between Cells. Nano Lett
662 19:3846-3853. 10.1021/acs.nanolett.9b01010.
- 663 24. Dehullu J, Vorholt JA, Lipke PN, Dufrene YF. 2019. Fluidic Force Microscopy Captures
664 Amyloid Bonds between Microbial Cells. Trends Microbiol . S0966-842X(19)30156-8.
- 665 25. Ho V, Herman-Bausier P, Shaw C, Conrad KA, Garcia-Sherman MC, Draghi J, Dufrene YF,
666 Lipke PN, Rauceo JM. 2019. An Amyloid Core Sequence in the Major *Candida albicans* Adhesin
667 Als1p Mediates Cell-Cell Adhesion. mBio 10:10.1128/mBio.01766-19. e01766-19.
- 668 26. Klotz SA, Sobonya RE, Lipke PN, Garcia-Sherman MC. 2016. Serum Amyloid P Component
669 and Systemic Fungal Infection: Does It Protect the Host or Is It a Trojan Horse? Open Forum
670 Infect Dis 3:ofw166. 10.1093/ofid/ofw166.
- 671 27. Lipke PN. 2018. What We Do Not Know about Fungal Cell Adhesion Molecules. J Fungi
672 (Basel) 4:10.3390/jof4020059. E59.
- 673 28. Chan CX, Lipke PN. 2014. Role of force-sensitive amyloid-like interactions in fungal catch
674 bonding and biofilms. Eukaryot Cell 13:1136-1142. 10.1128/EC.00068-14.

29. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L. 2004. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302-1306. 10.1038/nbt1012; nbt1012.
30. Tsois AC, Papandreou NC, Iconomidou VA, Hamodrakas SJ. 2013. A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *PLoS One* 8:e54175. 10.1371/journal.pone.0054175.
31. Gasior P, Kotulska M. 2014. FISH Amyloid - a new method for finding amyloidogenic segments in proteins based on site specific co-occurrence of aminoacids. *BMC Bioinformatics* 15:54-54. 10.1186/1471-2105-15-54.
32. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583-589. 10.1038/s41586-021-03819-2.
33. Tayeb-Fligelman E, Landau M. 2017. X-Ray Structural Study of Amyloid-Like Fibrils of Tau Peptides Bound to Small-Molecule Ligands. *Methods Mol Biol* 1523:89-100. 10.1007/978-1-4939-6598-4_5.
34. Lawrence MC, Colman PM. 1993. Shape complementarity at protein/protein interfaces. *J Mol Biol* 234:946-950. S0022-2836(83)71648-7.
35. Maji SK, Wang L, Greenwald J, Riek R. 2009. Structure-activity relationship of amyloid fibrils. *FEBS Lett* 583:2610-2617. 10.1016/j.febslet.2009.07.003.
36. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44:344. 10.1093/nar/gkw408.
37. Rousseau F, Serrano L, Schymkowitz JW. 2006. How evolutionary pressure against protein aggregation shaped chaperone specificity. *J Mol Biol* 355:1037-1047. 10.1016/j.jmb.2005.11.035.
38. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 33:299. 33/suppl_2/W299.
39. Munoz JF, Welsh RM, Shea T, Batra D, Gade L, Howard D, Rowe LA, Meis JF, Litvintseva AP, Cuomo CA. 2021. Clade-specific chromosomal rearrangements and loss of subtelomeric adhesins in *Candida auris*. *Genetics* 218:10.1093/genetics/iyab029. iyab029.
40. Guenther EL, Cao Q, Trinh H, Lu J, Sawaya MR, Cascio D, Boyer DR, Rodriguez JA, Hughes MP, Eisenberg DS. 2018. Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nat Struct Mol Biol* 25:463-471. 10.1038/s41594-018-0064-2.
41. Porat Y, Stepensky A, Ding FX, Naider F, Gazit E. 2003. Completely different amyloidogenic potential of nearly identical peptide fragments. *Biopolymers* 69:161-164. 10.1002/bip.10386.

714 42. Bera S, Arad E, Schnaider L, Shaham-Niv S, Castelletto V, Peretz Y, Zaguri D, Jelinek R, Gazit
715 E, Hamley IW. 2019. Unravelling the role of amino acid sequence order in the assembly and
716 function of the amyloid-beta core. *Chem Commun (Camb)* . 10.1039/c9cc03654g.

717 43. Alsteens D, Dupres V, Klotz SA, Gaur NK, Lipke PN, Dufrene YF. 2009. Unfolding Individual
718 Als5p Adhesion Proteins on Live Cells. *ACS Nano* 3:1677-1682. 10.1021/nn900078p.

719 44. Alsteens D, Van Dijck P, Lipke P, Dufrene YF. 2013. Quantifying the forces driving cell-cell
720 adhesion in a fungal pathogen. *Langmuir : the ACS journal of surfaces and colloids* 29:13473-80.
721 10.1021/la403237f.

722 45. Alsteens D, Garcia MC, Lipke PN, Dufrene YF. 2010. Force-induced formation and
723 propagation of adhesion nanodomains in living fungal cells. *Proc Natl Acad Sci U S A* 107:20744-
724 20749. 10.1073/pnas.1013893107.

725 46. Lipke PN, Garcia MC, Alsteens D, Ramsook CB, Klotz SA, Dufrene YF. 2012. Strengthening
726 relationships: amyloids create adhesion nanodomains in yeasts. *Trends Microbiol* 20:59-65.
727 10.1016/j.tim.2011.10.002.

728 47. Diederichs K, Karplus PA. 1997. Improved R-factors for diffraction data analysis in
729 macromolecular crystallography. *Nat Struct Biol* 4:269-275. 10.1038/nsb0497-269.

730 48. Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science*
731 336:1030-1033. 10.1126/science.1218231.

732 49. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ,
733 Balbirnie M, Wiltzius JJ, McFarlane HT, Madsen AO, Riekel C, Eisenberg D. 2007. Atomic
734 structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 447:453-457.
735 10.1038/nature05695.

736 50. Sammond DW, Eletr ZM, Purbeck C, Kimple RJ, Siderovski DP, Kuhlman B. 2007. Structure-
737 based protocol for identifying mutations that enhance protein-protein binding affinities. *J Mol*
738 *Biol* 371:1392-1404. 10.1016/j.jmb.2007.05.096.

739 51. Goddard TD, Huang CC, Ferrin TE. 2007. Visualizing density maps with UCSF Chimera. *J*
740 *Struct Biol* 157:281-287. 10.1016/j.jsb.2006.06.010.

741 52. Berezin C, Glaser F, Rosenberg J, Paz I, Pupko T, Fariselli P, Casadio R, Ben-Tal N. 2004.
742 ConSeq: the identification of functionally and structurally important residues in protein
743 sequences. *Bioinformatics* 20:1322-1324. 10.1093/bioinformatics/bth070.

744 53. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a
745 multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189-1191.
746 10.1093/bioinformatics/btp033.

747 54. Willaert RG, Kayacan Y, Devreese B. 2021. The Flo Adhesin Family. *Pathogens (Basel)*
748 10:1397. 10.3390/pathogens10111397.

749 55. Willaert RG. 2018. Adhesins of Yeasts: Protein Structure and Interactions. *J Fungi (Basel)*
750 4:10.3390/jof4040119. E119.

751 56. Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr* 66:125-132.
752 10.1107/S0907444909047337.

753 57. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser
754 crystallographic software. *J Appl Crystallogr* 40:658-674. 10.1107/S0021889807021206.

755 58. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB,
756 Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read
757 RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta*
758 *Crystallogr D Biol Crystallogr* 67:235-242. 10.1107/S0907444910045749.

759 59. Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by
760 the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240-255.
761 10.1107/S0907444996012255.

762 60. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta*
763 *Crystallogr D Biol Crystallogr* 66:486-501. 10.1107/S0907444910007493.

764
765