SapL1: A New Target Lectin for the Development of Antiadhesive Therapy Against *Scedosporium apiospermum*

Dania Martínez-Alarcón[1, 2], Jean-Philippe Bouchara[3], Roland J. Pieters[2], Annabelle Varrot[1*].

1 Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France. dania.martinez.alarcon@gmail.com, annabelle.varrot@cermav.cnrs.fr

2 Utrecht University, 3584 CG Utrecht, The Netherlands. R.J.Pieters@uu.nl

3 Host-Pathogen Interaction Study Group, GEIHP, EA 3142, SFR ICAT 4208, UNIV Angers, UNIV Brest, Institut de Biologie en Santé, IRIS, CHU d'Angers, Angers, France

* Corresponding author

Email: annabelle.varrot@cermav.cnrs.fr

Abstract

Scedosporium apiospermum is an emerging opportunistic fungal pathogen responsible for life-threatening infections in immunocompromised patients. This fungus exhibits limited susceptibility to all current antifungals and, due its emerging character, its pathogeny and virulence factors remain largely unknown. Carbohydrate binding proteins such as lectins are involved in host-pathogen interactions and may constitute valuable therapeutic targets to inhibit microbial adhesion to the host cells by using carbohydrate mimics. However, such lectins are still unidentified in *S. apiospermum*. Here, we present the first report of the identification and characterization of a lectin from *S. apiospermum* named SapL1. SapL1 is homologous to the conidial surface lectin FleA from *Aspergillus fumigatus* known to be involved in the adhesion to host glycoconjugates present in human lung epithelium. The present report includes a detailed strategy to achieve SapL1 soluble expression in bacteria, its biochemical characterization, an analysis of its specificity and affinity by glycan arrays and isothermal titration calorimetry (ITC), as well as the structural characterization of its binding mode by X–ray crystallography. The information gathered here contribute to the understanding of glycosylated surface recognition by *Scedosporium* species and is essential for the design and development of antiadhesive glycodrugs targeting SapL1.

Keywords: fucolectin, glycodrug, host recognition, glycans, opportunistic pathogen.

Introduction

During the last decades, an increased incidence of invasive infections, especially in immunosuppressed patients, has been caused by previously rare fungal pathogens.[1] Among those pathogenic fungi, species from the *Scedosporium* genus are particularly dangerous since their mortality rate may be over 80%.[2] *Scedosporium* infections mainly affect immunocompromised patients such as solid organ transplant (SOT) or hematopoietic stem cell transplant (HSCT) recipients, or patients with chronic granulomatous disease. Together with *Fusarium* species and some dematiaceous fungi, *Scedosporium* species are responsible for approximately 10% of the mycosis caused by filamentous fungi in hematopoietic cell transplant recipients and up to 19% in solid organ transplant patients.[3] In addition, *Scedosporium* species rank the second among the filamentous fungi colonizing the respiratory tract of people with cystic fibrosis (CF), after *Aspergillus fumigatus*, and because of the propensity of these fungi to disseminate in case of immune deficiency, this fungal colonization of the airways is considered in some centers as a contraindication to lung transplantation which remains the ultimate treatment in CF.[4] Apart from lung transplantation, this fungal colonization of the airways which usually starts in the adolescence and becomes chronic in up to 19% of the patients, may also lead to a chronic inflammation and sometimes to an allergic bronchopulmonary mycosis.[2] Furthermore, it has been reported that some species of this genus can also affect immunocompetent patients.[5]

The genus *Scedosporium* comprises more than ten species *(Scedosporium angustum, S. apiospermum, S. aurantiacum, S. boydii, S. cereisporum, S. dehoogii, S. desertorum, S. ellipsoideum, S. fusoideum, S. minutisporum, S. rarisporum and S. sanyaense)* whose taxonomy has been previously described based on molecular phylogenetic by using the sequences of four genetic loci.[2,6] However only the genome of *S. boydii, S. aurantiacum, S. dehoogii* and *S. apiospermum* have been completely sequenced to date.[7-9] Among *Scedosporium* species, *S. apiospermum* (formerly considered as the asexual form of *Pseudallescheria boydii*) is the most outstanding. This cosmopolitan microfungus is responsible for localized to severe or fatal disseminated infections in humans and is one of the most common *Scedosporium* species (with *S. boydii* and *S. aurantiacum*) capable to colonize chronically the lungs of CF patients.[3,10]

The treatment of *Scedosporium* infections is challenging because its efficacy depends on timely diagnosis, which is based on morphological detection by microscopy, histological analysis, and culture on selective media. This process is time consuming and may lead to false negative results, especially from respiratory secretions of patients with CF because of the common co-colonization by other fungi or bacteria.[2] Furthermore, *Scedosporium* species display a primary resistance to classical antifungals such as 5-flucytosine, amphotericin B, and the first-generation triazole drugs, fluconazole or itraconazole and exhibit a limited susceptibility to the newest generations of antifungal drugs, *i.e.* echinocandins, voriconazole and isavuconazole.[2,11] Nowadays, the first-line treatment for *Scedosporium* infections involves a combination therapy that includes the use of voriconazole in conjunction with other antifungals.[11,12] However, due to common recurrences, even without interruption of treatment, the recovery rates are poor and mortality remains over 65% while it is almost 100% when dissemination occurs, reason why it has aroused special interest despite its emerging character.[12,13]

Scedosporium infections begin with conidial adherence to tissues, followed by germination and hyphal elongation.[14] This adherence process allows it to avoid cleansing mechanisms aimed to eradicate the invading pathogens and represents the initial step towards infection.[14-17] Conidial adhesion is mediated by cell surface molecules (CSM), including different types of carbohydrates such as polysaccharides and glycoconjugates. Some of the most important carbohydrates CSM described to date in *Scedosporium* species include peptidorhamnomannans (PRMs),[18-20] α-glucan,[21] melanin,[22] ceramide monohexosides,[20,23] N-acetyl-D-glucosamine-containing molecules[24] and mannose/glucose-rich glycoconjugates.[20] The presence and/or abundance of these molecules on the cell surface varies according to the stage of development and is of great relevance to understand fungal pathobiology.[20] The carbohydrates binding proteins known as lectins also act as CSMs and were shown to have an essential role during pathogenesis. However, the knowledge about lectins in emerging microfungi is very limited to date.

Since some lectins from microorganisms have been demonstrated to be involved in the host recognition and adhesion process, they have been used to develop anti-adhesive therapy through the use of carbohydrates and their mimics.[15-17,25-31] This approach is particularly promising since it does not kill the pathogen nor arrest its cell cycle. Consequently, resistance frequencies are rare [16,26] and expected side effects

99 in non-target tissues are lower than those caused by conventional antifungal
100 compounds. This may allow its implementation as a prophylactic therapy for
101 immunocompromised patients.[17,27]

102 Due to the emerging character of *Scedosporium apiospermum*, there is very limited
103 information on its mechanisms of recognition and anchoring to the host.
104 Furthermore, lectins from this microorganism have not been characterized to date,
105 what hinders the development of an anti-adhesive therapy. Conversely, closely
106 related filamentous fungi were investigated leading to the identification and
107 characterization of their host binding modes. For example, in *A. fumigatus,* which is
108 a saprophytic mold also responsible for bronchopulmonary infections in receptive
109 hosts, the lectin FleA (or AFL) was identified and revealed a role in host-pathogen
110 interactions.[32] FleA is a six-bladed β-propeller homodimer located on the conidial
111 surface that recognize human blood group antigens and mediates *A. fumigatus*
112 binding to airway mucins and macrophages glycoproteins in a fucose-dependent
113 manner.[33] In healthy individuals, this anchorage is critical for the mucociliary
114 clearance process and the macrophagy, in fact it has been described that fleA-
115 deficient (ΔfleA) conidia are even more pathogenic than wild type (WT) conidia
116 both, in healthy and chemically immunocompromised mice.[33,34] However, CF
117 patients represent a very particular scenario because the mucus in their lungs is
118 thicker, in relation with mucin overproduction and the high content of calcium ions,
119 which modulates the supramolecular organization of MUC5B by protein cross-
120 linkages.[34,35] This contributes to the suboptimal transport properties of mucus and
121 compromises the pathogen clearance mechanisms [34]. Furthermore, the aberrant
122 glycosylation of the mucins MUC5B and MUC5B in CF lungs causes, among other
123 things, an increase in the abundance of sialyl-Lewis X and Lewis X determinants in
124 the non-reducing end of mucins carbohydrate chains, [36,37]which can also be
125 translated as an increase in the fucose content. Therefore, in this context, the FleA
126 (and homologous proteins) anchoring to the mucus layer plays an essential role in
127 the colonization of the CF lungs by *A. fumigatus*.

128 Here, we have used the recently sequenced genome of *S. apiospermum*,[9] to identify a
129 putative homologue of FleA that we have called SapL1 for *Scedosporium apiospermum*
130 Lectin 1. The present report comprises SapL1 identification, its production in
131 bacteria, an analysis of the fine specificity and affinity of the recombinant protein, as
132 well as its structural characterization by X–ray crystallography.

133 # Results

134 **Production and purification of SapL1**

135 The hypothetical protein XP_016640003.1 (EMBL accession number), encoded by
136 SAPIO_CDS9261 and from now on referred as SapL1, was identified through data
137 mining using the FleA (pdb entry 4D4U[38]) sequence as bait into the genome of the
138 reference strain *S. apiospermum* IHEM 14462. [9] During the sequence analysis, we
139 identified that the first 74 amino acids of the putative protein (Uniprot A0A084FYP2)
140 are a very disordered region that is not present in any other protein of the same
141 family. We suspected that this peptide could derived from a misplacement of the
142 open reading frame (ORF) of SapL1 during the genome assembly and that its
143 sequence actually starts at methionine-75 (Figure S1). The SapL1 coding sequence
144 (from Met-75) was fused to an N-terminal 6xHis tag cleavable by the Tobacco Etch
145 virus protease (TEV) under regulation of *trc* and *T7* promoters into pProNde and
146 pET-TEV vectors, respectively (Figure 1A). Expression was performed in *Escherichia*
147 *coli* and purification was carried out using immobilized metal affinity
148 chromatography (IMAC). Unfortunately, the original expression yield with both
149 vectors (~0.35 $mg \cdot L^{-1}$ of culture) was too low to proceed with characterization
150 studies. Therefore, we explored new alternatives to enhance the expression. First,
151 the thioredoxin protein (Trx) as well as 6-His tag and TEV cleavage site were fused
152 at the N-terminus of SapL1 by subcloning into the pET32-TEV vector (Figure 1A).
153 This strategy substantially increased the SapL1 production yield but most of the
154 protein remained insoluble as part of inclusion bodies (Figure 1B). A wide range of
155 expression conditions were subsequently assayed to achieve sufficient soluble
156 expression by modification of various parameters such as growth temperature, host
157 strain, inducer concentration, optical density of the culture at induction, culture
158 duration, etc. Sixty-nine different sets of parameters were assayed (S1 Table) and it
159 was possible to improve the yield up to 4 $mg \cdot L^{-1}$ (Figure 1C). The best set of
160 conditions for SapL1 expression was using *E. coli* strain TRX, pProNde vector, LB
161 media, growth at 37°C and 160 rpm until OD_{600} = 0.4, before a switch of the
162 temperature to 16°C and overnight induction at OD 0.8 with 0.05 mM IPTG and 1%
163 L-rhamnose (Rh).

164 **Rhamnose influence on SapL1 solubility**

165 Interestingly, during the optimization of the expression conditions, we found that
166 rhamnose (Rh) plays an essential role in the solubility and stability of SapL1.
167 Therefore, we performed a new set of experiments to demonstrate the correlation
168 between the amount of protein recovered after purification and the Rh concentration
169 in the media. For this, we investigated SapL1 expression under the previously
170 described conditions modified by supplementation of the culture medium with
171 different concentrations of Rh (0.1%, 0.15%, 0.2%, 0.5% and 1%). Purification
172 parameters were set to obtain high purity product and kept identical for all
173 experiments. Figure 2 shows the chromatograms obtained for representative
174 concentrations accompanied by their respective SDS-PAGE profile. To quantify
175 SapL1 expression in these experiments, we integrated the area under peaks
176 corresponding to the protein (Figure 2, black bars). Thus we confirmed that SapL1
177 recovery is directly proportional to the Rh concentration in the culture medium.
178 Then, the experiment was repeated at three different concentrations (0.1%, 0.15%,
179 0.2%) using 20°C as the induction temperature, and production behavior was similar
180 to that found at 16°C, but with lower performance.

181 Additional experiments also showed that SapL1 expression can be induced at high
182 concentration of rhamnose in a dose dependent manner, even without addition of
183 IPTG when the pProNde vector is used (data not shown). For this reason, we initially
184 consider that the higher production rates of SapL1 could be attributed to the
185 dissociation of the *LacI*-DNA complex caused by an increase in cytosolic potassium
186 concentration since the accumulation of this ion is one of the main adaptive
187 responses of *E. coli* to hypertonic shocks, and its concentration has been correlated
188 to the loss of interactions between DNA and proteins in *in vitro* assays.[39-41] However,
189 previous reports have demonstrated that, *in vivo*, the effects of macromolecular
190 crowding caused by plasmolysis (the reduction of intracellular water content, as an
191 adaptive response to osmotic shock) on the activity of cytoplasmic proteins, is
192 sufficient to buffer the kinetics of association of DNA-protein against changes in
193 $[K^+]$.[41,42] Thus, it seems that the higher recovery rates of SapL1 is actually caused by
194 a combination of two main parameters: 1) the leak of genetic repression by
195 dissociation of the *LacI-DNA* complex (characteristic of *lac* promoters) and 2) a
196 positive effect on the solubility and folding of the protein due to the increased
197 concentration of organic osmoregulators, whose synthesis is induced in *E. coli* as an
198 adaptive response to the external hypertonicity caused, in this case, by the addition

199 of rhamnose to the culture medium[39-47] This hypothesis was later supported by the
200 finding that the addition of glycerol, instead of rhamnose, displays the same effect
201 on production and solubility of SapL1 (data not shown). Plasmolysis may also play
202 an essential role in SapL1 stabilization, since it has been shown that macromolecular
203 crowding has positive effects on protein folding and in some cases it can drive self-
204 association of improper folded proteins into functional oligomers.[41]

205 **Biochemical characterization**

206 We assayed the thermal stability of SapL1 in 26 different buffers in a pH range of 5
207 to 10 through Thermal Shift Assay (TSA). The most suitable condition for this
208 protein was MES buffer 100 mM pH 6.5, where a single denaturing event at T_m of
209 55°C was observed (Figure S2). Then, to estimate the molecular size of the native
210 protein, we performed size exclusion chromatography using an ENrich™ SEC 70
211 column (Bio-Rad) and 20 mM MES, 100 mM NaCl, pH 6.5 as mobile phase.
212 However, the protein displayed strong non-specific interactions with the matrix of
213 the column, and SapL1 could not be eluted even using 5 M of NaCl. Interestingly, it
214 could be recovered when the buffer was supplemented with 20 mM α-methyl
215 fucoside or L-rhamnose, evidencing similar effects on SapL1 elution for both sugars.
216 Due to the impossibility to estimate the molecular weight of SapL1 by size exclusion
217 chromatography on this resin, we performed measurements in solution using
218 Dynamic Light Scattering (DLS). We obtained a monodisperse peak corresponding
219 to a protein of 72 ± 29.4 kDa corroborating that SapL1 forms dimers as FleA and
220 other proteins of this family (monomer MW: 40 kDa, data not shown).[32] The range
221 of the standard deviation also suggested an ellipsoidal shape, which is characteristic
222 for the dimers in this lectin family[48].

223 **Carbohydrate binding properties**

224 A hemagglutination assay showed that recombinant SapL1 agglutinated rabbit red
225 blood cells at 0.97 $\mu g \cdot ml^{-1}$ (Figure 3A). It confirms that the recombinant lectin is
226 active and that its heterologous production in *E. coli* did not alter its
227 hemagglutinating properties.

228 To identify the potential ligands of SapL1 on epithelial cell surfaces, we submitted
229 SapL1 to the glycan array version 5.4 of the Consortium for Functional Glycomics
230 (USA) consisting of 585 mammalian glycans. It was labelled with Fluorescein
231 Isothiocyanate (FITC) in a molar ratio of 0.426 and its binding properties were

232 analysed at two different concentrations (5 and 50 μg·mL^{-1}). As expected from its
233 homology with FleA, SapL1 recognizes fucosylated oligosaccharides independently
234 of the fucose linkage. The α1-2 and α1-3/4 linked fucosides displayed the highest
235 affinity whilst α1-6 linked ones, the lowest. The weakest interactions with
236 fucosylated compounds were reported for branched oligosaccharides (Figure S3).
237 The affinity of SapL1 for L-fucose and α-methyl-fucoside was determined by
238 Isothermal Titration Calorimetry (ITC) and the K_d was found to be 225 μM and 188
239 μM, respectively with stoichiometry fixed to 1 since the measurements were done
240 in the presence of an excess of ligand (Figure 3B). These values are in agreement
241 with the affinity constant (around 110 μM) reported for FleA for α-methyl-
242 fucoside.[49] No binding interaction was observed for SapL1 with rhamnose by ITC
243 (data not shown).

244 Analysis of the glycans constituting blood group determinants revealed that SapL1
245 binds to all epitopes with a preference for H type 2 (Fucα1-2Galβ1-4GlcNAcβ) then
246 Lewis^a (Galβ1-3(Fucα1-4)GlcNAcβ) and Lewis^x (Galβ1-4(Fucα1-3)GlcNAcβ).
247 However, most of the recognized branched oligosaccharides contained the core
248 fucose Fucα1-6. Epitopes with two fucose units, such as Lewis^b and fucosylated
249 polylactosamine, were also well recognized. Addition of a galactose or a GalNAc as
250 in blood group B or A antigens did not impair Fucα1-2 recognition (Figure 3C).
251 Spacers used to join the carbohydrates to the chip also display a strong influence on
252 binding. It is remarkable that 70% of the 90 positive binders contained either the
253 spacer Sp0 ($CH_2CH_2NH_2$) or Sp8 ($CH_2CH_2CH_2NH_2$). Those spacers also display a
254 strong influence on binding, especially for small glycans such as Galα1-3(Fucα1-
255 2)Galβ1-4(Fucα1-3)GlcNAcβ which was recognized when attached to Sp0 but not to
256 Sp8. This may be due to a steric hindrance caused by the modification of
257 carbohydrate presentation on the surface of the chip.

258 **Overall structure of SapL1**

259 SapL1 was co-crystallized with α-methyl-fucoside and the structure of the complex
260 was solved by molecular replacement at 2.3 Å resolution in the $P2_1$ space group using
261 the coordinates of FleA (PDB code 4D4U[38]) as the search model. The asymmetric unit
262 contained two monomers, assembled as a dimer with all 295 amino acids visible
263 apart of the N-terminal methionine. SapL1 folds into the canonical six-bladed β-
264 propeller with six-binding sites at the interface between blades typical for this family
265 of lectin (Figure 4). A fucose moiety was found in 5 and 4 of the six binding pockets

266 of chains A and B, respectively while glycerol originating from the crystallizing
267 solution was found in the other binding sites.

268 The overall fold of SapL1 and FleA as well as the overall dimer are very similar with
269 a rsmd of 1.2 and 1.26 Å, respectively. They share 43% of sequence identity and both
270 proteins present the same distribution of β-strands except for the lack of the last
271 blade of strand 4 in SapL1, and the external face of blade 5, in which FleA displays
272 an elongated β-strand (β22) that is split in two in SapL1 (β22/23). It is remarkable
273 that the structure of the first 3 blades is highly conserved in both proteins, while the
274 second half (blades 3-6) displays the largest discrepancies. Furthermore, FleA also
275 presents two additional small α-helices (ή3 and ή4), located in the loops between
276 sheets that serve as connectors for blades (2/3 and 5/6, respectively, Figure 5). Similar
277 conclusions can be drawn when comparing SapL1 to its homologue AOL in
278 *Aspergillus oryzae*.[50] Those proteins have similar dimerization interface and we
279 noticed differences on the other side of the propeller where sugar binding occurs.
280 The lectins present surface loops of different size and sequence leading to change in
281 the architecture of the related binding sites that impacts their affinity as discussed
282 in more details below. The first member structurally characterized in this lectin
283 family was AAL from the mushroom *Aleuria aurantia*.[51] It presents differences on
284 both sides of the β-propeller with the microfungal members of this family mainly at
285 the level of the surface loops. This leads to a different dimerization interface and the
286 dimer cannot overlay. This also affects fucose binding since AAL only has five
287 functional binding sites. The role of these lectins is still unknown, in particular for
288 the mushroom members.
289

290 **Protein-ligand interactions**

291 Due to divergence in the tandem repeat sequence forming each blade of the
292 propeller, the six binding sites of SapL1 monomer are not equivalent, but they share
293 important conserved features. Hydrophobic interactions are observed between the
294 C6 of the fucose and at least three residues of the protein (mainly isoleucine,
295 tryptophan/tyrosine and leucine). The O2 and O3 hydroxyls make strong hydrogen
296 bonds with a conserved triad of amino acid consisting of an arginine, a glutamic acid
297 and a tryptophan. In the cases where glycerol was found in the binding pocket, its
298 own hydroxyls mimic the interactions of fucose with these same residues (Figure 6).
299 It is to be noted that SapL1 binding sites are more conserved than FleA binding site

300 where a glutamine can replace the glutamic acid and a tyrosine the tryptophan in
301 the triad (Figure S4).

302 The O2 hydroxyl seemed to be the most versatile position, since it established
303 interactions with a loop adjacent in four of the six binding sites (BPs 2, 4, 5 and 6).
304 This is particularly interesting since the binding pockets that do not contain this loop
305 (BPs 1 and 3) were mostly occupied by glycerol instead of fucose, indicating that
306 those interactions could be responsible for enhancing the affinity and might be
307 explored for development of inhibitors.

308 Discussion

309 The development novel techniques for diagnosis has exposed that, until now, the
310 incidence of fungal infections has been underestimated.[1,52] In comparison with
311 pathogenic bacteria, which have been the focus of attention for several years, there
312 is a lack of information about the virulence factors and host recognition systems of
313 pathogenic fungi.[53] This hinders the development of new drugs against this
314 microorganisms, while their resistance to the current ones is progressing
315 quickly.[13,24,52] Therefore, there is a rush in obtaining information that may contribute
316 to the development of new antifungal agents.

317 In this study, we present the identification, production and characterization of a new
318 carbohydrate binding protein from the emerging microfungus *S. apiospermum*,
319 whose host-anchoring mechanism is completely unknown. To the best of our
320 knowledge, this is the first studied lectin for this opportunistic pathogen. SapL1 is
321 homologous to the conidial surface lectin FleA from *A. fumigatus* known to be
322 involved in adhesion to the host glycoconjugates present in mucins and human lung
323 epithelium.[33] During our analysis, we have identified a possible shift in the ORF of
324 SapL1 and we have relayed this information to corroborate this finding and to
325 update the entry if necessary.

326 In our initial attempts to produce recombinant SapL1, we observe that the
327 conventional parameters of expression in *E. coli*, result in the production of insoluble
328 protein that was mainly found in inclusion bodies. Later, experimental data
329 indicated that the presence of rhamnose or glycerol was required for SapL1
330 solubilization during bacterial expression and that the production yield was directly

331 proportional to their concentration in the culture medium. This suggests that the
332 stabilization of SapL1 derives from an osmomodulator effect that, either directly or
333 indirectly, enhances its proper folding. This osmomodulator effect may be due to
334 the leak of genetic repression by dissociation of the repressor (*LacI-DNA)* complex[39-
335 41] in the first case, or in the second case to the increased concentration of organic
336 osmoregulators, which are known to enhance protein folding and solubility,
337 presumably also reinforced by the macromolecular crowding produced by
338 plasmolysis.[39-47] These findings, and the extensive research carried out to produce
339 SapL1 in its soluble form, provide important insights for the heterologous
340 expression of eukaryotic lectins tending to be produced as insoluble proteins in *E.*
341 *coli.* Besides, our findings highlight the useful role of organic osmoregulators during
342 heterologous expression of proteins to favour their proper folding and to avoid the
343 formation of inclusion bodies.

344 We demonstrated that SapL1 is strictly specific for fucosylated carbohydrates and
345 recognized all blood group types present in the glycan array screening. These results
346 are particularly interesting since it has been shown that there is six times more
347 content of fucose α1,3/4 linked to glycoproteins in the CF airways.[36] This
348 phenomenon is mainly due to an increased expression of the α1,3-
349 fucosyltransferase, which is involved in the synthesis of sialyl-Lewis X and Lewis X
350 determinants attached to mucus mucins.[37] Therefore, the fact that SapL1 recognizes
351 equally the α1-2 and α1-/3/4-linked fucosides may explain the high incidence of
352 scedosporiosis in CF patients. Besides, it correlates with the presence of *SapL1* gene
353 in all the pathogenic strains of *Scedosporium* whose genome has been sequenced to
354 date. However, there is no evidence suggesting that the phenotype of blood group
355 could have an influence in this recognition conversely to previous reports on other
356 pathogens like *Pseudomonas aeruginosa* and *Haemophilus influenza*.[54-57]

357 Finally, the specificity and affinity of each one of the six binding pockets of SapL1
358 has been deeply analyzed to highlight the features that must be explored for the
359 design of efficient inhibitors. Within our analysis, we have found that the binding
360 pockets are non-equivalent but they all share the features necessary for fucose
361 recognition. This example of divergence can also be found in the other lectins of the
362 same family with FleA being the member with the greatest differences between
363 pockets known to date.[38] It is remarkable that although SapL1 has a relatively low
364 sequence identity with other members of this family, its specificity and affinity are

365 very similar to those reported for FleA from *A. fumigatus*,[49] AAL from the orange
366 peel mushroom *Aleuria aurantia*[51] and the bacterial lectins BambL and RSL, from
367 *Burkholderia ambifaria* and *Ralstonia solanacearum*, respectively.[58,59] Together with the
368 crystallographic data, this confirms that the structure and function of this lectin
369 family are highly conserved, settling the possibility for development of a broad-
370 spectrum therapy.

371 Overall, our research has revealed the first insights about the recognition of human
372 glycoconjugates by *S. apiospermum* lectins and contributes to the general
373 understanding of the host-binding process during the early stages of infection. The
374 detailed information exposed here, places SapL1 as a promising target for the
375 diagnosis and/or treatment of *Scedosporium* infections and it will be of great value to
376 guide the development of antiadhesive glycodrugs against this pathogen.

377 Material and methods

378 Production

379 The coding sequence for SapL1 was optimized for expression in *Escherichia coli* and
380 ordered at Eurofins Genomics (Ebersberg, Germany) before cloning into the
381 expression vectors pET-TEV,[60] pET32-TEV[49] and pProNde. pProNde is a homemade
382 vector where the NcoI restriction site of pProEx HTb (EMBL, Heidelberg) has been
383 replaced by a NdeI restriction site by PCR. Then, plasmids were introduced into *E.*
384 *coli* strains by thermal shock at 42°C and different expression parameters were
385 assayed (S1 Table) until soluble expression was achieved. The best set of conditions
386 for SapL1 expression were: *E. coli* TRX strain, pProNde vector, LB medium, cells
387 grown at 37°C and 160 rpm until 0.4 DO_{600}, then cultures transferred to 16°C until
388 0.8 OD_{600nm} and protein induction carried out overnight by the addition of isopropyl-
389 β-D-1-thiogalactoside (IPTG) to a final concentration of 0.05 mM and 1% L-
390 rhamnose (Rh).

391 Purification

392 Cells were subsequently harvested by centrifugation at 5000 g for 10 min and
393 resuspended in buffer A (50 mM Tris-HCl, 500 mM NaCl, pH 8.5). After addition of
394 1 µl of Denarase (C-LEcta GmbH, Leipzig, Germany) and moderate agitation during
395 30 minutes at room temperature, cells were disrupted at 1.9 kbar using a cell

396 disruptor (Constant Ltd Systems). Cell debris were removed by centrifugation at
397 22000 g for 30 minutes. The supernatant was filtered using 0.45 micrometer
398 membranes (PES, ClearLine) and the protein purification was carried out by IMAC
399 using 1 mL His-Trap FF columns (GE Healthcare Life Sciences) and a NGC™
400 chromatography system (Bio-Rad). After loading the supernatant, the column was
401 rinsed thoroughly with buffer A until stabilization of the baseline. Bound proteins
402 were eluted through the addition of buffer B (50 mM Tris-HCl, 500 mM NaCl, 500
403 mM Imidazole, pH 8.5) in an 0-500 mM imidazole gradient over 20 mL. Fractions
404 containing SapL1 were pooled and concentrated by ultrafiltration (Pall, 10kDa cut-
405 off) prior buffer exchange to 50 mM Tris-HCl, 100 mM NaCl, pH 8.5 using PD10
406 desalting columns (GE Healthcare Life Sciences). Then, the fusion was cleaved off
407 overnight at 19°C using the TEV protease produced in the lab in 1:50 ratio and
408 addition of 0.5 mM EDTA and 0.25 mM TCEP. The sample was loaded on the His-
409 trap column to separate the cleaved protein collected in the flow-through from the
410 TEV protease and potential uncleaved sample retained and eluted with imidazole.
411 Sapl1 containing fractions were concentrated by centrifugation to the desired
412 concentration.

413 **Size exclusion chromatography**

414 Size exclusion chromatography (SEC) was performed using a High-Resolution
415 ENrich™ SEC 70 column (Bio-Rad) on NGC™ chromatography system (Bio-Rad).
416 Column was equilibrated with 50 ml of buffer D (20 mM MES, 100 mM NaCl, pH
417 6.5) and 200 µl of sample at 10 mg mL^{-1} were injected into the system followed by 40
418 ml isocratic elution on buffer D supplemented with 20 mM α-methyl fucoside, 0.1%
419 L-rhamnose, 5 M NaCl, 2 M NaCl or 500 mM NaCl, according to the experiment.
420 Fractions were monitored by the absorbance at 280 nm and 0.5 mL fractions were
421 collected in the resolving region of the column.

422 **Dynamic light scattering (DLS)**

423 DLS analyses were performed using a Zetasizer™ NanoS (Malvern Panalytical) with
424 a 40 µl quartz cuvette. Measurements were performed in triplicate on protein sample
425 at 1 mg·ml^{-1} in buffer C (50 mM Tris-HCl, 100 mM NaCl, pH 8.5) after centrifugation.

426 **Thermal Shift Assay (TSA)**

427 The thermal stability of SapL1 was analyzed by TSA with the MiniOpticon real-time
428 PCR system (Bio-Rad). Prior assay, buffer stocks at 100 mM and a mixture

429 containing 70 µl of SapL1 at 1 mg ml^{-1}, 7 µl of 500x Sypro Orange (Merk Sigma-
430 Aldrich,) and 63 µl of ultrapure H_2O were prepared. Then, 7.5 µl of H_2O, 12.5 µl of
431 the corresponding buffer and 5 µl of the protein/Sypro mixture were mixed in 96-
432 well PCR microplates. The heat exchange test was then carried out from 20°C to
433 100°C with a heating rate of 1°C·min^{-1}. Fluorescence intensity was measured with
434 Ex/Em: 490/530 nm and the data processing was performed with the CFX Manager
435 software.

436 **Isothermal Titration Calorimetry (ITC)**

437 Experiments were performed using a Microcal ITC200 calorimeter (Malvern
438 Panalytical) with 40 µl of L-fucose 2.5 mM in the syringe and 200 µl of protein 0.04
439 mM in the sample cell. 2 µl injections in a range of 120 seconds while stirring at 1000
440 rpm were performed. Experimental data were adjusted to a theoretical titration
441 curve by the Origin ITC Analysis software.

442 **Hemagglutination assay**

443 Agglutination test was performed with fresh rabbit erythrocytes (bioMérieux, Lyon)
444 in U 96-well plates (Nalgene). For the test, 150 mM NaCl was used as a negative
445 control and 1 mg ml^{-1} LecB of *P. aeruginosa* as a positive control. 50 µl of sample was
446 prepared at 0.1 mg·ml^{-1} and submitted to serial double dilutions. 50 µl of rabbit
447 erythrocytes 3% were added to each well prior incubation of the plate at room
448 temperature. After 2 hours, the result of the experiment was evaluated and
449 agglutination activity was calculated according to the dilution of the protein.

450 **Glycan arrays**

451 Protein was labeled with Fluorescein Isothiocyanate (FITC) (Merk Sigma-Aldrich)
452 according to the supplier's instructions with slight modifications. Briefly two
453 milligrams of protein were dissolved in 1 ml of buffer E (100 mM Na_2CO_3, 100 mM
454 NaCl, pH 9); then, 40 µl of FITC at 1 mg·ml^{-1}, previously dissolved in DMSO, were
455 gradually added to the protein solution and the mixture was gently stirred at room
456 temperature overnight. Next day, the solution was supplemented with NH_4Cl to a
457 final concentration of 50 mM and free FITC was removed using PD10 column with
458 PBS as mobile phase. Protein concentration was determined at A_{280} and FITC at A_{490}
459 using a NanoDrop™ 200 (Thermo Scientific) and Fluorescein/Protein molar ratio
460 (F/P) was estimated by the following formula:

461 $$\text{Molar}\frac{\text{F}}{\text{P}} = \frac{\text{MW}}{389} \times \frac{\frac{A_{495}}{195}}{[A_{280} - (0.35 \times A_{495})]E^{0.1\%}}$$

462 Where MW is the molecular weight of the protein, 389 is the molecular weight of FITC, 195 is the
463 absorption E 0.1% of bound FITC at 490 nm at pH 13.0, (0.35 X *A*495) is the correction factor due to
464 the absorbance of FITC at 280 nm, and E 0.1% is the absorption at 280 nm of a protein at 1.0 mg·ml^{-1},
465 Being an ideal F/P should be 0.3 > 1.

466 Labeled lectin was sent to the Consortium for Functional Glycomics (CFG; Boston,
467 MA, USA) and its binding properties were assayed at 5 and 50 µg·ml^{-1} on a
468 "Mammalian Glycan Array version 5.4" which contain 585 glycans in replicates of
469 6. The highest and lowest signal of each set of replicates were eliminated and the
470 average of the remaining data was normalized to the percentages of the highest RFU
471 value for each analysis. Finally, the percentages for each glycan were averaged at
472 different lectin concentrations.

473 **Crystallization and data collection**

474 Crystal screening was performed using the hanging-drop vapour diffusion
475 technique by mixing equal volumes of pure protein at 5 mg·ml^{-1} and precipitant
476 solutions from commercial screenings of Molecular Dimensions (Newmarket, UK).
477 2 µl drops were incubated at 19°C until crystals appeared. A subsequent
478 optimization of positive conditions for SapL1 crystallization was carried out and
479 crystals suitable for X-ray diffraction analysis were obtained under solution
480 containing 100 mM Bicine pH 8.5, 1.5 M ammonium sulfate $(NH_4)_2SO_4$ and 12% v/v
481 glycerol. Crystals were soaked in mother liquor supplemented with 10% (v/v)
482 glycerol, prior to flash cooling in liquid nitrogen. Data collection was performed on
483 PX1 beamline at SOLEIL Synchrotron (Saint Aubin, FR) using a Pixel detector.

484 **Structure determination**

485 Data were processed using XDS [61]software and were converted to structure factors
486 using the CCP4 program package v.6.1,[62] with 5% of the data reserved for Rfree
487 calculation. The structure was determined using the molecular-replacement method
488 with Phaser v.2.5,[63] using the structure of FleA dimer (PDB entry 4D4U[38]) as starting
489 model. Model refinement was performed using REFMAC 5.8 [64] alternated with
490 manual model building in Coot v.0.7.[65] Sugar residues and other compounds that
491 were present were placed manually using Coot and validated using Privateer.[66] The

492 final model has been validated and deposited in the PDB Database with accession
493 number 6TRV.

494 **Figures**

495 Figures were created using PyMOL Version 1.8.4 (Schrödinger), ChemDraw Version
496 15, ESPript Version 3.0[67] and PowerPoint 16.

497 **Accession numbers**

498 The structure of SapL1 was deposited as PDB ID: 6TRV.

499 **Author Contributions:** Conceptualization, A.V.; formal analysis, A.V. and D.M.A;
500 investigation, D.M.A.; methodology A.V. and D.M.A; visualization A.V. and D.M.A;
501 writing—original draft preparation, D.M.A.; writing—review and editing, A.V, B.J-
502 P, R.P and D.M.A; supervision, A.V. and R.P.; project administration, A.V.; funding
503 acquisition, A.V. All authors have read and agreed to the published version of the
504 manuscript.

505 **Funding:** This project has received funding from the European Union's Horizon
506 2020 research and innovation program under the Marie Skłodowska-Curie grant
507 (H2020-MSCA-ITN-2017-EJD-765581), from the French cystic fibrosis association
508 Vaincre la Mucoviscidose. The project was also supported by Glyco@Alps (ANR-15-
509 IDEX02).

510 **Acknowledgments:** We would like to thank to Valérie Chazalet and Emilie Gillon
511 for their technical assistance. We also would like to thank for access to the beamlines
512 BM30A-FIP at the European Synchrotron Radiation Facility (ESRF), Grenoble,
513 France where initial tests were performed and Proxima 1 at SOLEIL Synchrotron,
514 Saint Aubin, France (proposal number 20170827), where final data were collected.
515 Thanks to our local contacts Jean-Luc Ferrer, Serena Sirigu, and Pierre Legrand for
516 their assistance and technical support.

517 **Conflicts of Interest:** The authors declare no conflict of interest.

518

519 References

520 1. Tuite NL, Lacey K. Overview of invasive fungal infections. In: O'Connor L, Glynn
521 B, editors. NY: Humana Press, Totowa, NJ; 2013. p. 1-23.

522 2. Ramirez-Garcia A, Pellon A, Rementeria A, Buldain I, Barreto-Bergter E, Rollin-
523 Pinheiro R, et al. (2018). *Scedosporium* and *Lomentospora*: An updated overview of
524 underrated opportunists. *Med Mycol.* 56:S102-S25.
525 3. Goldman C, Akiyama MJ, Torres J, Louie E, Meehan SA (2016). *Scedosporium*
526 *apiospermum* infections and the role of combination antifungal therapy and GM-
527 CSF: a case report and review of the literature. *Medical Mycol Case Rep.* 11:40-3.
528 4. Rammaert B, Puyade M, Cornely OA, Seidel D, Grossi P, Husain S, et al. (2019).
529 Perspectives on Scedosporium species and Lomentospora prolificans in lung
530 transplantation: Results of an international practice survey from ESCMID fungal
531 infection study group and study group for infections in compromised hosts, and
532 European Confederation. *Transpl Infect Dis.* 21:1-8.
533 5. Luplertlop N (2018). *Pseudallescheria/Scedosporium* complex species: from saprobic
534 to pathogenic fungus. *J Mycol Med.* 28:249-56.
535 6. Gilgado F, Cano J, Gené J, Guarro J (2005). Molecular phylogeny of the
536 *Pseudallescheria boydii* species complex: proposal of two new species. *J Clin*
537 *Microbiol.* 43:4930-42.
538 7. Duvaux L, Shiller J, Vandeputte P, Bernonville TDD, Thornton C, Papon N, et al.
539 (2017). Draft genome sequence of the human-pathogenic fungus *Scedosporium*
540 *boydii*. *Genome Announc.* 5:e00871-17.
541 8. Pérez-Bercoff Å, Papanicolaou A, Ramsperger M, Kaur J, Patel HR, Harun A, et
542 al. (2015). Draft genome of Australian environmental strain WM 09.24 of the
543 opportunistic human pathogen *Scedosporium aurantiacum*. *Genome Announc.*
544 3:e01526-14.
545 9. Vandeputte P, Ghamrawi S, Rechenmann M, Iltis A, Giraud S, Fleury M, et al.
546 (2014). Draft genome sequence of the pathogenic fungus *Scedosporium*
547 *apiospermum*. *Genome Announc.* 2:e00988-14.
548 10. Pereira De Mello T, Aor AC, Santiago S, De Oliveira C, Branquinha MH, Luis A
549 (2016). Conidial germination in *Scedosporium apiospermum, S . aurantiacum, S .*
550 *minutisporum* and *Lomentospora prolificans*; influence of growth conditions and
551 antifungal susceptibility profiles. *Mem Inst Oswaldo Cruz.* 111:484-94.
552 11. Tortorano AM, Richardson M, Roilides E, van Diepeningen A, Caira M, Munoz
553 P, et al. (2014). ESCMID and ECMM joint guidelines on diagnosis and
554 management of hyalohyphomycosis: *Fusarium spp., Scedosporium spp.* and others.
555 *Clin Microbiol Infect.* 20:27-46.
556 12. Troke P, Aguirrebengoa K, Arteaga C, Ellis D, Heath CH, Lutsar I, et al. (2008).
557 Treatment of scedosporiosis with voriconazole: clinical experience with 107
558 patients. *Antimicrob Agents Chemother.* 52:1743-50.
559 13. Rodriguez-tudela JL, Guarro J, Kantarcioglu AS, Horre R, Estrella MC, Berenguer
560 J, et al. (2006). *Scedosporium apiospermum:* changing clinical spectrum of a therapy-
561 refractory opportunist. *Med Mycol.* 44:295-327.

562 14. Cortez KJ, Roilides E, Quiroz-telles F, Meletiadis J, Antachopoulos C, Knudsen
563 T, et al. (2008). Infections Caused by Scedosporium spp. *Clin Microbiol Rev.* 21:157-
564 97.
565 15. Theuretzbacher U, Piddock LJV (2019). Non-traditional antibacterial therapeutic
566 options and challenges. *Cell Host Microbe.* 26:61-72.
567 16. Krachler AM, Orth K (2013). Targeting the bacteria-host interface strategies in
568 anti-adhesion therapy. *Virulence.* 4:284-94.
569 17. Ofek I, Hasty DL, Sharon N (2003). Anti-adhesion therapy of bacterial diseases:
570 prospects and problems. *FEMS Immunol Med Microbiol.* 38:181-91.
571 18. Figueiredo RT, Fernandez PL, Dutra FF, Gonza Y, Cristina L, Bittencourt VCB, et
572 al. (2010). TLR4 recognizes *Pseudallescheria boydii* conidia and purified
573 rhamnomannans. *J Biol Chem.* 285:40714-23.
574 19. Pinto MR, Gorin PAJ, Wait R, Mulloy B, Barreto-Bergter E (2005). Structures of
575 the O-linked oligosaccharides of a complex glycoconjugate from *Pseudallescheria*
576 *boydii*. *Glycobiology.* 15:895-904.
577 20. Pinto MR, De Sá ACM, Limongi CL, Rozental S, Santos ALS, Barreto-Bergter E
578 (2004). Involvement of peptidorhamnomannan in the interaction of
579 *Pseudallescheria boydii* and HEp2 cells. *Microbes Infect.* 6:1259-67.
580 21. Bittencourt VCB, Figueiredo RT, Silva RB, Moura DS, Fernandez PL, Sassaki GL,
581 et al. (2006). An a-glucan of *Pseudallescheria boydii* is involved in fungal
582 phagocytosis and Toll-like receptor activation. *J Biol Chem.* 281:22614-23.
583 22. Ghamrawi S, Renier G, Saulnier P, Cuenot S, Zykwinska A, Dutilh BE, et al.
584 (2014). Cell wall modifications during conidial maturation of the human
585 pathogenic fungus *Pseudallescheria boydii*. *PLoS ONE.* 9:e100290-e.
586 23. Rollin-Pinheiro R, Liporagi-lopes LC, Meirelles JVD, Souza LMD (2014).
587 Characterization of *Scedosporium apiospermum* glucosylceramides and their
588 involvement in fungal development and macrophage functions. *PLoS ONE.*
589 9:e98149-e.
590 24. De Mello TP, Aor AC, Gonçalves DDS, Seabra SH, Branquinha MH, Luis A
591 (2018). *Scedosporium apiospermum, Scedosporium aurantiacum, Scedosporium*
592 *minutisporum* and *Lomentospora prolificans*: a comparative study of surface
593 molecules produced by conidial and germinated conidial cells. *Mem Inst Oswaldo*
594 *Cruz.* 113:1-8.
595 25. Shoaf-sweeney KD, Hutkins RW. Adherence, anti-adherence, and
596 oligosaccharides: preventing pathogens from sticking to the host. 2009. p. 101-61.
597 26. Sharon N (2006). Carbohydrates as future anti-adhesion drugs for infectious
598 diseases. *Biochim Biophys Acta.* 1760:527-37.
599 27. Sharon N, Ofek I (2000). Safe as mother's milk: carbohydrates as future anti-
600 adhesion drugs for bacterial diseases. *Glycoconj J.* 17:659-64.

601 28. Tamburrini A, Colombo C, Bernardi A (2020). Design and synthesis of
602 glycomimetics: Recent advances. *Med Res Rev.* 40:495-531.
603 29. Sattin S, Bernardi A (2016). Glycoconjugates and glycomimetics as microbial anti-
604 adhesives. *Trends Biotechnol.* 34:483-95.
605 30. Pieters RJ (2011). Carbohydrate mediated bacterial adhesion. *Adv Exp Med Biol.*
606 715:227-40.
607 31. Visini R, Jin X, Bergmann M, Michaud G, Pertici F, Fu O, et al. (2015). Structural
608 Insight into Multivalent Galactoside Binding to *Pseudomonas aeruginosa* Lectin
609 LecA. *ACS Chem Biol.* 10:2455-62.
610 32. Houser J, Komarek J, Kostlanova N, Cioci G, Varrot A, Kerr SC, et al. (2013). A
611 soluble fucose-specific lectin from *Aspergillus fumigatus* conidia - structure,
612 specificity and possible role in fungal pathogenicity. *PLoS ONE.* 8:e83077-e.
613 33. Kerr SC, Fischer GJ, Sinha M, McCabe O, Palmer M, Choera T, et al. (2016). FleA
614 expression in *Aspergillus fumigatus* is recognized by fucosylated structures on
615 mucins and macrophages to prevent lung infection. *PLOS Pathog.*e100555-e.
616 34. Sakai K, Hiemori K, Tateno H, Hirabayashi J, Gonoi T (2019). Fucose-specific
617 lectin of *Aspergillus fumigatus:* binding properties and effects on immune response
618 stimulation. *Med Mycol.* 57:71-83.
619 35. Thornton DJ, Rousseau K, McGuckin MA (2008). Structure and function of the
620 polymeric mucins in airways mucus. *Annu Rev Physiol.* 70:459-86.
621 36. Glick MC, Kothari VA, Liu A, Stoykova LI, Scanlin TF (2001). Activity of
622 fucosyltransferases and altered glycosylation in cystic fibrosis airway epithelial
623 cells. *Biochimie.* 83:743-7.
624 37. Lamblin G, Degroote S, Perini J-mM, Delmotte P, Scharfman A, Davril M, et al.
625 (2001). Human airway mucin glycosylation: a combinatory of carbohydrate
626 determinants which vary in cystic fibrosis. *Glycoconj J.* 18:661-84.
627 38. Houser J, Komarek J, Cioci G, Varrot A, Imberty A, Wimmerova M (2015).
628 Structural insights into *Aspergillus fumigatus* lectin specificity : AFL binding sites
629 are functionally non-equivalent. *Acta Crystallogr D Biol Crystallogr.*442-53.
630 39. Frank DE, Saecker RM, Bond JP, Capp MW, Tsodikov OV, Melcher SE, et al.
631 (1997). Thermodynamics of the interactions of Lac repressor with variants of the
632 symmetric Lac operator: effects of converting a consensus site to a non-specific
633 site. *J Mol Biol.* 267:1186-206.
634 40. Capp MW, Cayley DS, Zhang W, Guttman HJ, Melcher SE, Saecker RM, et al.
635 (1996). Compensating effects of opposing changes in putrescine (2+) and K+
636 concentrations on lac repressor-lac operator binding: in vitro thermodynamic
637 analysis and in vivo relevance. *J Mol Biol.* 258:25-36.
638 41. Cayley S, Lewis BA, Guttman HJ, Record MT (1991). Characterization of the
639 cytoplasm of *Escherichia coli K-12* as a function of external osmolarity implications
640 for protein-DNA interactions in vivo. *J Mol Biol.* 222:281-300.

641 42. Richeys B, Cayley DS, Mossing MC, Kolka C, Anderson CF, Farrar TC, et al.
642 (1987). Variability of the intracellular ionic environment of *Escherichia coli.*
643 Differences between *in vitro* and *in vivo* effects of ion concentrations on protein-
644 DNA interactions and gene expression. *J Biol Chem.* 262:7157-64.
645 43. Harries D, Rosgen J. A practical guide on how osmolytes modulate
646 macromolecular properties. Elsevier; 2008. p. 679-735.
647 44. Massiah MA, Wright KM, Du H (2016). Obtaining soluble folded proteins from
648 inclusion bodies using sarkosyl, Triton X-100, and CHAPS: application to LB and
649 M9 minimal media. *Curr Protoc Protein Sci.* 84:1-24.
650 45. Pettitt BM, Bolen DW (2005). Protein folding, stability, and solvation structure in
651 osmolyte solutions. *Biophys J.* 89:2988-97.
652 46. Singh LR, Poddar NK, Dar TA, Rahman S, Kumar R, Ahmad F (2011). Forty years
653 of research on osmolyte-induced protein folding and stability. *J Iran Chem Soc.* 8:1-
654 23.
655 47. Ajito S, Iwase H, Takata S-i, Hirai M (2018). Sugar-mediated stabilization of
656 protein against chemical or thermal denaturation. *The Journal of Physical Chemistry*
657 *B.* 122:8685-97.
658 48. Bonnardel F, Kumar A, Wimmerova M, Lahmann M, Perez S, Varrot A, et al.
659 (2019). Architecture and evolution of blade assembly in β-propeller lectins.
660 *Structure.* 27:764-75.e3.
661 49. Lehot V, Brissonnet Y, Dussouy C, Brument S, Cabanettes A, Gillon E, et al.
662 (2018). Multivalent fucosides with nanomolar affinity for the *Aspergillus fumigatus*
663 lectin FleA prevent spore adhesion to pneumocytes. *Chem Eur J.* 24:19243-9.
664 50. Makyio H, Shimabukuro J, Suzuki T, Imamura A, Ishida H, Kiso M, et al. (2016).
665 Six independent fucose-binding sites in the crystal structure of *Aspergillus oryzae*
666 lectin. *Biochem Biophys Res Commun.* 477:477-82.
667 51. Wimmerova M, Mitchell E, Sanchez JF, Gautier C, Imberty A (2003). Crystal
668 structure of fungal lectin. Six-bladed β-propeller fold and novel fucose
669 recognition mode for *Aleuria aurantia* lectin. *J Biol Chem.* 278:27059-67.
670 52. Debourgogne A, Dorin J, Machouart M (2016). Emerging infections due to
671 filamentous fungi in humans and animals: only the tip of the iceberg? *Env*
672 *Microbiol Rep.* 8:332-42.
673 53. Hogan LH, Klein BS, Levitz SM (1996). Virulence factors of medically important
674 fungi. *Clin Microbiol Rev.* 9:469-88.
675 54. Kuo KC, Kuo HC, Huang LT, Lin CS, Yang SN (2013). The clinical implications
676 of ABO blood groups in *Pseudomonas aeruginosas* sepsis in children. *J Microbiol*
677 *Immunol Infect.* 46:109-14.
678 55. Gilboa-Garber N, Sudakevitz D, Sheffi M, Sela R, Levene C (1994). PA-I and PA-
679 II lectin interactions with the ABO(H) and P blood group glycosphingolipid

680 antigens may contribute to the broad spectrum adherence of *Pseudomonas*
681 *aeruginosa* to human tissues in secondary infections. *Glycoconj J.* 11:414-7.
682 56. Scanlin TF, Glick MC (1999). Terminal glycosylation in cystic fibrosis. *Biochim*
683 *Biophys Acta.* 1455:241-53.
684 57. Taylor-Cousar JL, Zariwala MA, Burch LH, Pace RG, Drumm ML, Calloway H,
685 et al. (2009). Histo-blood group gene polymorphisms as potential genetic
686 modifiers of infection and cystic fibrosis lung disease severity. *PLoS ONE.*
687 4:e4270-e.
688 58. Audfray A, Claudinon J, Abounit S, Ruvoe N, Smith DF, Pendu JL, et al. (2012).
689 Fucose-binding lectin from opportunistic pathogen *Burkholderia ambifaria* binds to
690 both plant and human oligosaccharidic epitopes. *J Biol Chem.* 287:4335-47.
691 59. Kostlanova N, Mitchell EP, Lortat-Jacob H, Oscarson S, Lahmann M, Gilboa-
692 Garber N, et al. (2005). The fucose-binding lectin from *Ralstonia solanacearum*. A
693 new type of β-propeller architecture formed by oligomerization and interacting
694 with fucoside, fucosyllactose, and plant xyloglucan. *J Biol Chem.* 280:27839-49.
695 60. Houben K, Marion D, Tarbouriech N, Ruigrok RWH, Blanchard L (2007).
696 Interaction of the C-terminal domains of sendai virus N and P proteins:
697 comparison of polymerase-nucleocapsid interactions within the paramyxovirus
698 family. *J Virol.* 81:6807-16.
699 61. Kabsch W (2010). XDS. *Acta Crystallographica.* 66:125-32.
700 62. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. (2011).
701 Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol*
702 *Crystallogr.* 67:235-42.
703 63. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ
704 (2007). Phaser crystallographic software. *J Appl Cryst.* 40:658-74.
705 64. Murshudov GN, Nicholls RA (2011). REFMAC 5 for the refinement of
706 macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 67:355-67.
707 65. Emsley P, Lohkamp B (2010). Features and development of Coot. *Acta Crystallogr*
708 *D Biol Crystallogr.* 66:486-501.
709 66. Aguirre J, Iglesias-Fernandez J, Rovira C, Davies GJ, Wilson KS, Cowtan KD
710 (2015). Privateer: software for the conformational validation of carbohydrate
711 structures. *Nat Struct Mol Biol.* 22:833-4.
712 67. Robert X, Gouet P (2014). Deciphering key features in protein structures with the
713 new ENDscript server. *Nucleic Acids Res.* 42:320-4.

714

715

716 **Figure captions**

717 **Figure 1. SapL1 production.** A) Schematic representation of the genetic constructs
718 used for the expression of SapL1. B) SDS-PAGE of fractions collected from SapL1
719 purification using pET32-TEV vector. IF: Insoluble fraction, SF: Soluble fraction, FT:
720 Flow through, W: wash and elution, from right to left respectively. Red arrow
721 indicates inclusion bodies from insoluble fractions. C) Representative
722 chromatograms of SapL1 purification before and after the process with their
723 respective profile on 15% SDS-polyacrylamide gels (insoluble fraction, soluble
724 fraction, flow through, washing and elution, from left to right respectively). The
725 fractions containing SapL1 are delimited.

726 **Figure 2. Rhamnose influence on SapL1 solubility.** A) SapL1 purification
727 chromatograms at 0.1%, 0.2%, and 1% rhamnose with their corresponding SDS-
728 PAGE profile. IF: Insoluble fraction, SF: Soluble fraction, FT: Flow through, W: wash
729 and elution, from left to right respectively. Black rectangles indicate the elution
730 peaks and elution fractions containing SapL1 on SDS-polyacrylamide gels. B)
731 Numerical values derived from the integration of areas under the peaks
732 corresponding to SapL1 for each experiment. Black and striped bars represent the
733 results for experiments performed at 16°C and 20°C, respectively.

734 **Figure 3. SapL1 carbohydrate binding properties.** A) Hemagglutination assay of
735 SapL1 on fresh rabbit erythrocytes. Negative and positive controls consist of 150 mM
736 NaCl and the lectin LecB from *Pseudomonas aeruginosa*, respectively. B) Titration of
737 SapL1 with L-fucose and α-methyl-fucoside with the thermogram and the
738 integration displayed at the top and bottom, respectively. C) Analysis of the
739 interactions of SapL1 with blood group epitopes. The graph shows the relative
740 binding of SapL1 to glycans containing Lewis and ABH blood group antigens from
741 the 90 hits identified as binders.

742 **Figure 4: SapL1 overall structure.** A) Surface and cartoon representation of SapL1
743 monomer colored from blue (*N*-terminal end) to red (*C*-terminal end). BP: Binding
744 pockets. B) Representation of SapL1 dimer colored by monomer with ligand
745 depicted in sticks.

746 **Figure 5: SapL1/FleA comparison.** A) Overlay of SapL1 (blue, PDB: 6TRV)) and
747 FleA (green, PDB: 4D4U) structures. Red arrows indicate the main differences
748 between structures. B) Sequence alignment of SapL1 and FleA with display of their

749 secondary structure elements. The residues involved in ligand interactions of each
750 pocket are indicated by circles for hydrophobic and stacking interactions and by
751 triangles for hydrogen bonds colored according to the binding pocket: BP-1, blue;
752 BP-2 cyan; BP-3 green; BP-4, orange, BP-5, brown; BP-6, red.

753 **Figure 6: SapL1-ligand interactions.** A) The sequence alignment of the 6 blades of
754 SapL1. The residues involved in ligand binding interactions are indicated in solid
755 boxes: BP-1, blue; BP-2 cyan; BP-3 green; BP-4, orange, BP-5, brown; BP-6, red.
756 Striped green boxes indicate the four additional residues expected to be involved in
757 fucose binding within BP-3 that were not visible in the structure since only glycerol
758 was attached to this pocket. B) Zoom on the interactions of ligand with each binding
759 site of SapL1 and their schematic representation.

760

761

762

763 **Table 1.** Data-collection and refinement statistics.

Data collection		
Beamline	SOLEIL Proxima 1	
Wavelength (A°)	0.97857	
Space group	$P2_1$	
Unit cell dimensions a, b, c (Å), α, β, γ (°)	76.06, 45.66, 83.48, 90, 105.05, 90	
No. of monomers in ASU	2	
Resolution (Å)	40.0 – 2.4 (2.46-2.4)	
R*merge*	0.085 (0.428)	
R*pim*	0.07 (0.356)	
Mean *I/σ (I)*	5.3 (1.6)	
Completeness (%)	98.2 (98.4)	
Multiplicity	2.7 (2.7)	
CC1/2	0.991(0.7)	
No. reflections /No. Unique reflections	58305/ 21570	
Refinement		
Resolution (Å)	20.0-2.3	
No. of reflections in working set / Free set	21561 / 1089	
R work/ R free	17.5 / 24.0	
R.m.s Bond lengths (Å)	0.017	
Rmsd Bond angles (°)	1.94	
Rmsd Chiral ($Å^3$)	0.087	
No. atoms / Bfac ($Å^2$)	Chain A	Chain B
Protein	2222/32.64	2220/40.83
Ligand and heterogen	89/31.56	66/36.914
Waters	111/31.018	69/33.206
Ramachandran Allowed / Favored/ Outliers (%)	96 / 4 / 0	95 / 4 / 1
PDB Code	6TRV	

764 *Values in parentheses are for the outer shell

765

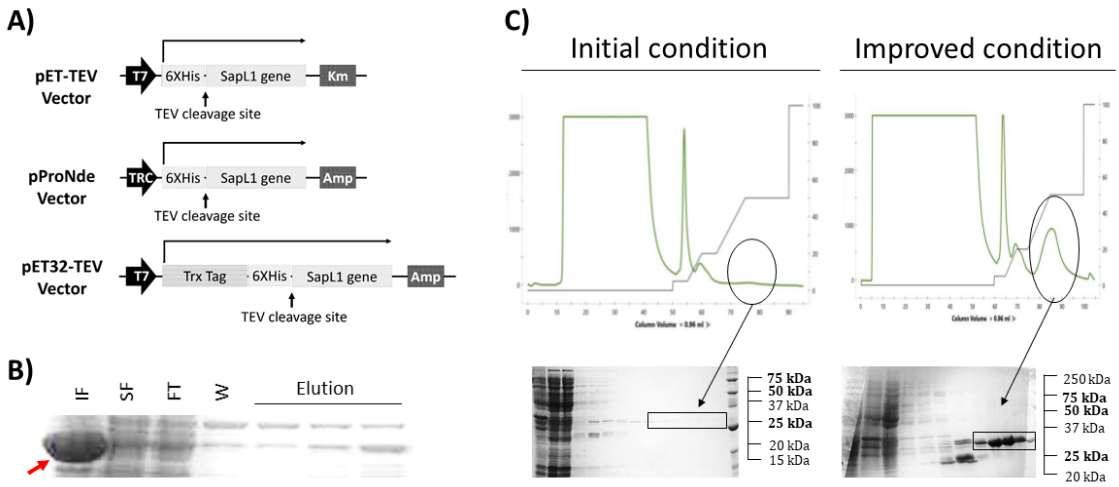


Figure 1

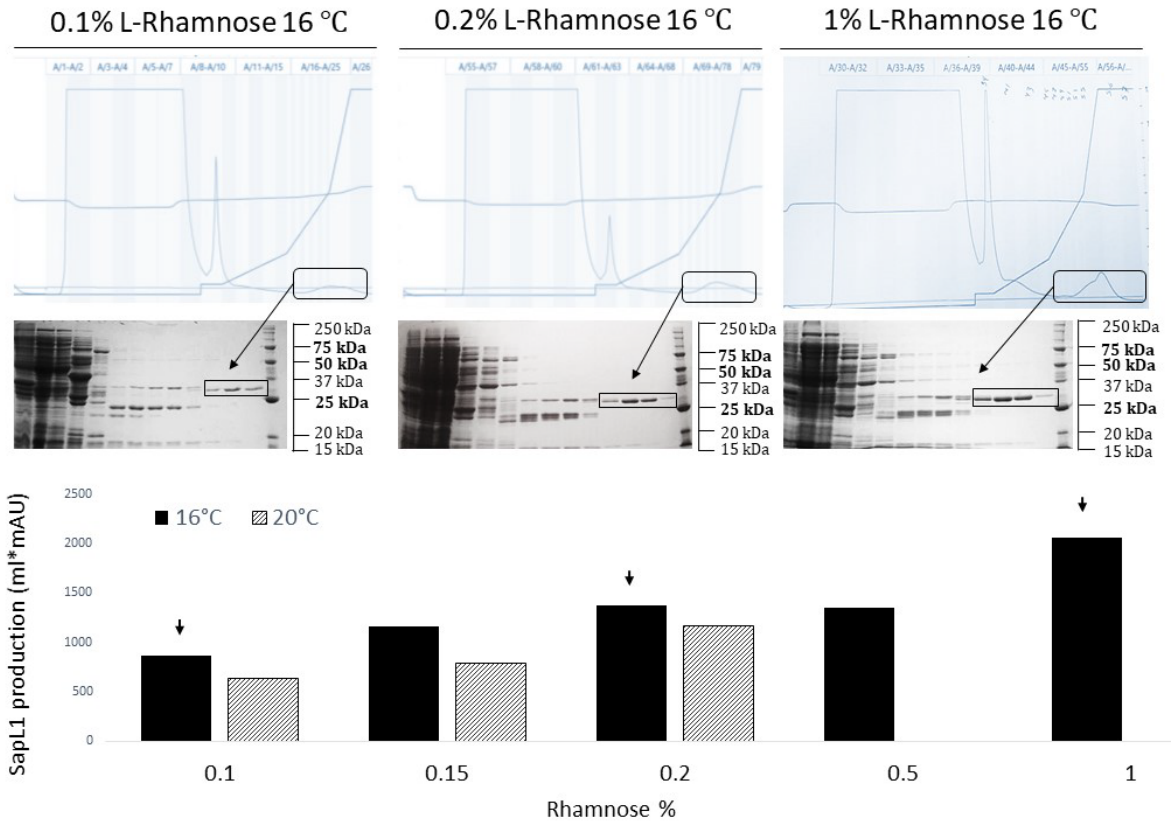
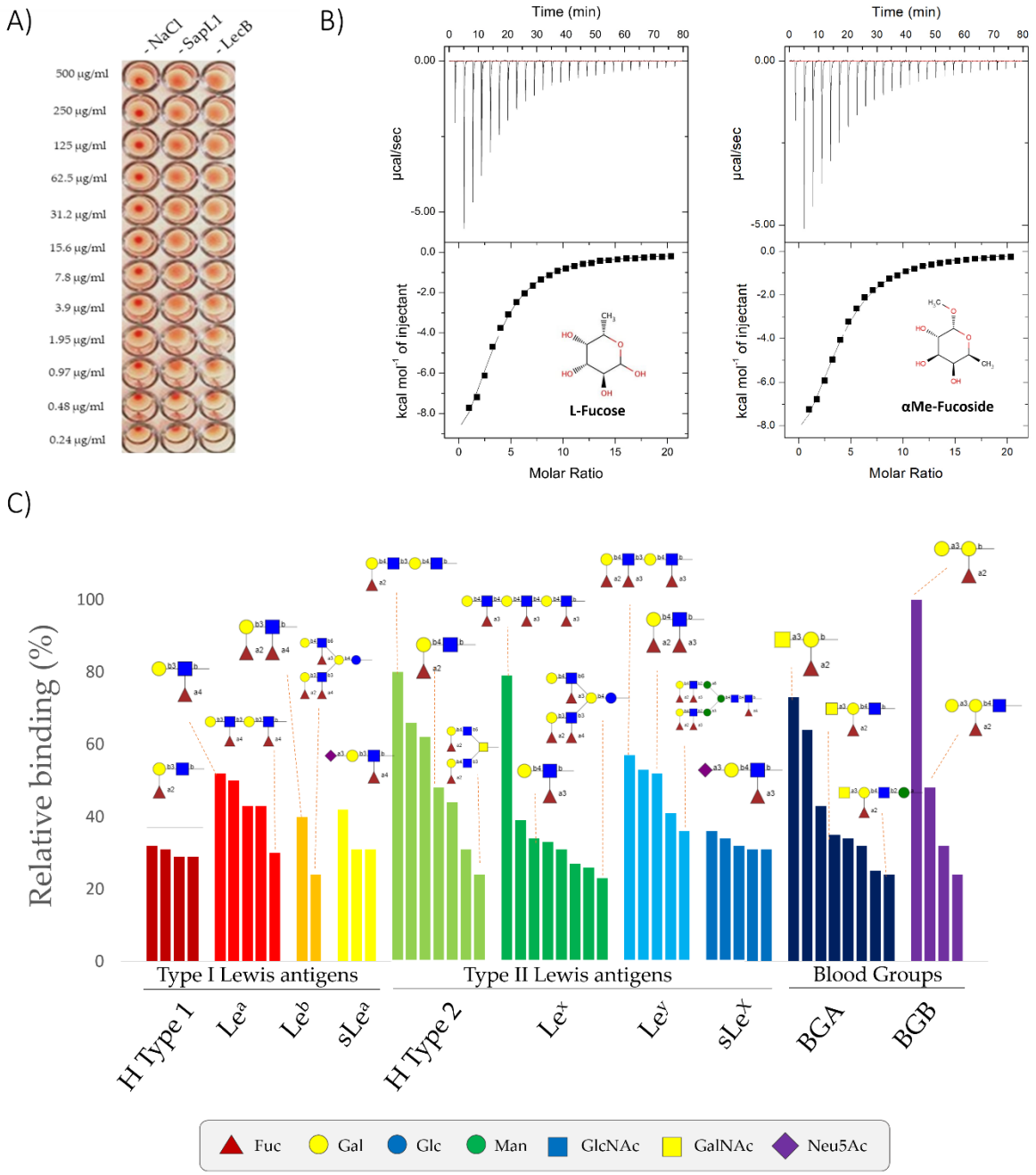


Figure 2


A)
- NaCl
- SapL1
- LecB
500 µg/ml
250 µg/ml
125 µg/ml
62.5 µg/ml
31.2 µg/ml
15.6 µg/ml
7.8 µg/ml
3.9 µg/ml
1.95 µg/ml
0.97 µg/ml
0.48 µg/ml
0.24 µg/ml
B)
Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Molar Ratio
L-Fucose
αMe-Fucoside
C)
Relative binding (%)
Type I Lewis antigens
Type II Lewis antigens
Blood Groups
H Type 1
Le^a
Le^b
sLe^a
H Type 2
Le^x
Le^y
sLe^x
BGA
BGB
Fuc
Gal
Glc
Man
GlcNAc
GalNAc
Neu5Ac

772

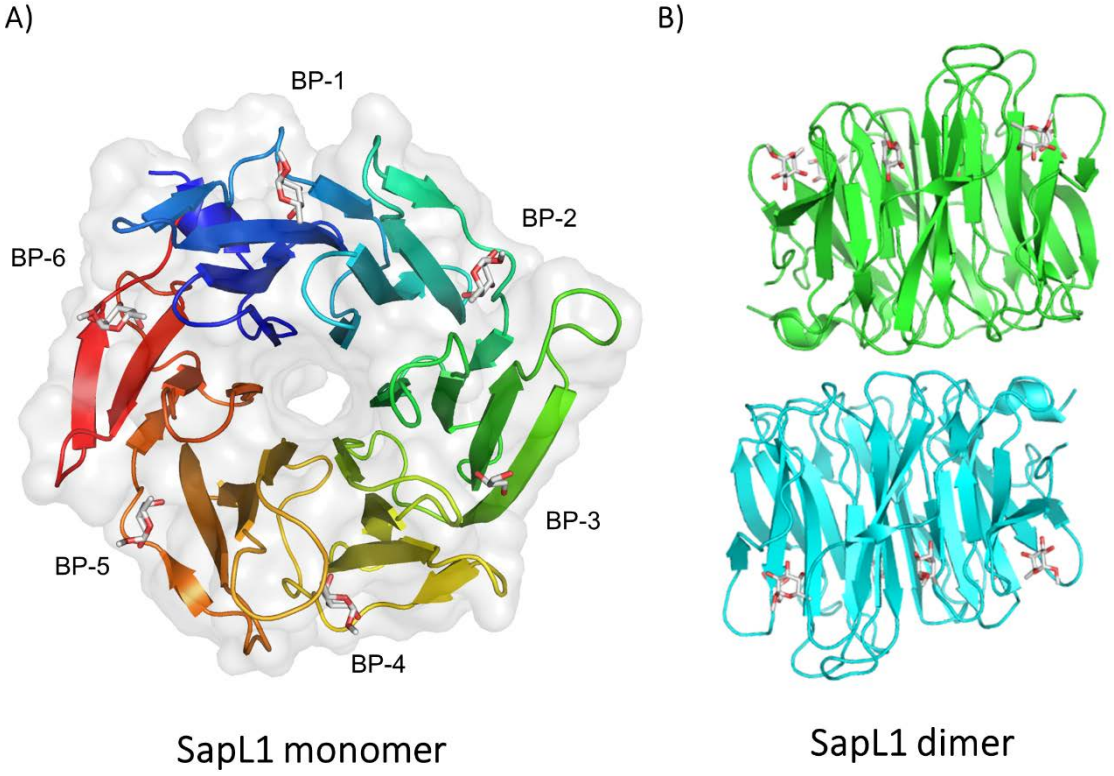


Figure 4

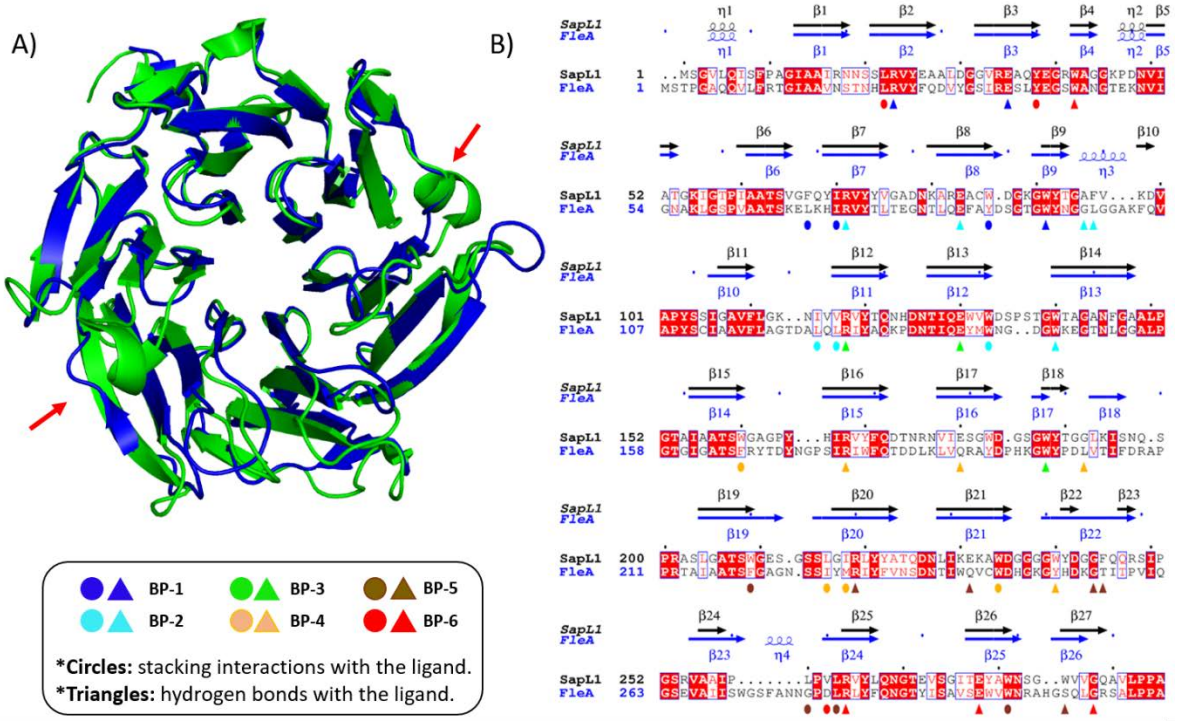


Figure 5

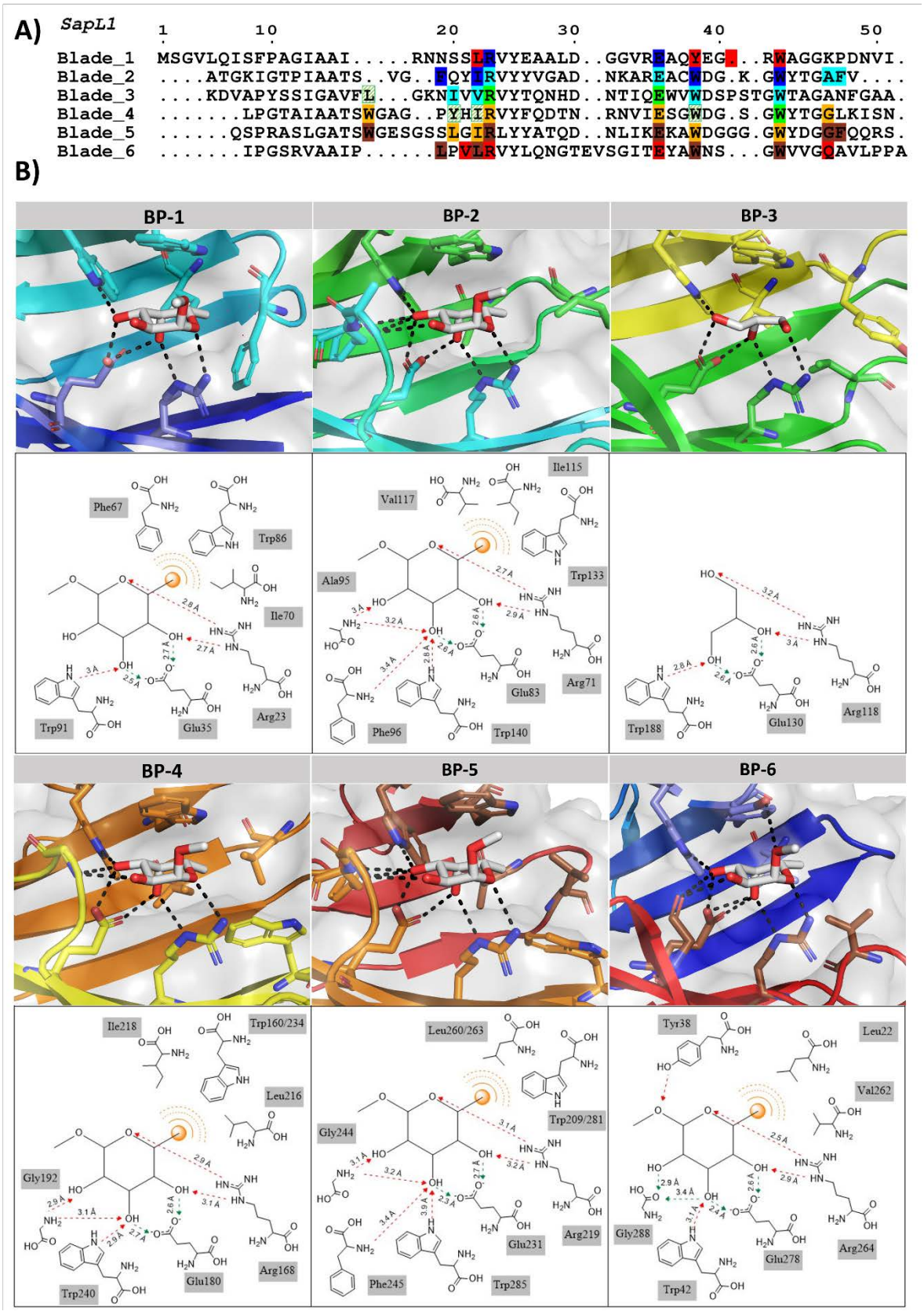

A)
SapL1
1 10 20 30 40 50
Blade_1 MSGVLQISFPAGIAAI.....RNNSSLRVYEAALD..GGVREAQYEG...RWAGGKPDNVI.
Blade_2 ....ATGKIGTPIAATS..VG..HQYIRVYYVGAD..NKAREACWDG.K.GWYTGAFV....
Blade_3 ....KDVAPYSSIGAVFL...GKNIVVRVYTQNHD..NTIQEWVWDSPSTGWTAGANFGAA.
Blade_4 .....LPGTAIAATSWGAG..PYHIRVYFQDTN..RNVIESGWDG.S.GWYTGGLKISN.
Blade_5 .....QSPRASLGATSWGESGSSLGIRLYYATQD..NLIKEKAWDGGG.GWYDGGFQQRS.
Blade_6 .....IPGSRVAAIP.....LPVLRVYLQNGTEVSGITEYAWNS...GWVVGQAVLPPA
B)
BP-1
BP-2
BP-3
Phe67
Trp86
Ile70
Trp91
Glu35
Arg23
Ile115
Val117
Trp133
Ala95
Arg71
Glu83
Phe96
Trp140
Trp188
Glu130
Arg118
BP-4
BP-5
BP-6
Ile218
Trp160/234
Leu216
Gly192
Trp240
Glu180
Arg168
Leu260/263
Trp209/281
Gly244
Arg219
Glu231
Phe245
Trp285
Tyr38
Leu22
Val262
Gly288
Glu278
Arg264
Trp42

Figure 6