

# Structure–function relationship of a novel fucoside-binding fruiting body lectin from *Coprinopsis cinerea* exhibiting nematotoxic activity

Silvia Bleuler-Martinez<sup>1</sup>, Annabelle Varrot<sup>2</sup>, Vincent Olieric<sup>3</sup>, Mario Schubert<sup>4,5</sup>, Eva Vogt<sup>1</sup>, Céline Fetz<sup>1</sup>, Therese Wohlschlager<sup>1</sup>, David Fernando Plaza<sup>1,6</sup>, Martin Wälti<sup>1</sup>, Yannick Duport<sup>1</sup>, Guido Capitani<sup>3</sup>, Markus Aebi<sup>1</sup>, Markus Künzler<sup>1,\*</sup> 

<sup>1</sup>Institute of Microbiology, Department of Biology, Eidgenössische Technische Hochschule (ETH) Zürich, 8093, Zürich, Switzerland, <sup>2</sup>Univ. Grenoble Alpes, CNRS, CERMAV, 38000, Grenoble, France, <sup>3</sup>Swiss Light Source (SLS), Paul Scherrer Institute (PSI), 5232, Villigen, Switzerland,

<sup>4</sup>Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zürich, 8093, Zürich, Switzerland, <sup>5</sup>Department of Biosciences, University of Salzburg, 5020, Salzburg, Austria, <sup>6</sup>Division of Infectious Diseases, Karolinska University Hospital, 171 64, Solna, Sweden

\*Corresponding author: Email: mkuenzle@ethz.ch

Lectins are non-immunoglobulin-type proteins that bind to specific carbohydrate epitopes and play important roles in intra- and inter-organismic interactions. Here, we describe a novel fucose-specific lectin, termed CML1, which we identified from fruiting body extracts of *Coprinopsis cinerea*. For further characterization, the coding sequence for CML1 was cloned and heterologously expressed in *Escherichia coli*. Feeding of CML1-producing bacteria inhibited larval development of the bacterivorous nematode *Caenorhabditis tropicalis*, but not of *C. elegans*. The crystal structure of the recombinant protein in its apo-form and in complex with H type I or Lewis A blood group antigens was determined by X-ray crystallography. The protein folds as a sandwich of 2 antiparallel  $\beta$ -sheets and forms hexamers resulting from a trimer of dimers. The hexameric arrangement was confirmed by small-angle X-ray scattering (SAXS). One carbohydrate-binding site per protomer was found at the dimer interface with both protomers contributing to ligand binding, resulting in a hexavalent lectin. In terms of lectin activity of recombinant CML1, substitution of the carbohydrate-interacting residues His54, Asn55, Trp94, and Arg114 by Ala abolished carbohydrate-binding and nematotoxicity. Although no similarities to any characterized lectin were found, sequence alignments identified many non-characterized agaricomycete proteins. These results suggest that CML1 is the founding member of a novel family of fucoside-binding lectins involved in the defense of agaricomycete fruiting bodies against predation by fungivorous nematodes.

Key words: defense; fucose; mushroom; nematode; toxin.

## Introduction

Lectins are defined as non-immunoglobulin proteins that bind to carbohydrate epitopes on polysaccharides, glycoproteins, and glycolipids without modifying them (Sharon and Lis 2004). These proteins occur in all kingdoms of life where they play important roles in intra- and inter-organismal interactions (Taylor and Drickamer 2019). The increasing availability of genome sequences from fungi (<https://mycocosm.jgi.doe.gov/mycocosm/home>) and the general interest in bioactive proteins from mushrooms (Erjavec et al. 2012) have led to the identification and characterization of a growing number of fungal lectins from the phyla Basidiomycota and Ascomycota (Varrot et al. 2013; Kobayashi and Kawagishi 2014; Sabotic et al. 2016).

Lectins from macrofungi are often referred to as fruiting body lectins, because they are usually highly expressed and/or have been isolated from reproductive or resting structures such as fruiting bodies and sclerotia, respectively (Goldstein and Winter 2007). Besides their expression pattern, fruiting body lectins share several other features, such as a small molecular weight, high water solubility, predicted cytoplasmic localization, sustained stability and activity under a wide range of temperature and pH conditions, and considerable

resistance towards proteolytic degradation (Khan and Khan 2011). Since many of these lectins recognize particular carbohydrate epitopes that have not been described in fungi, it has been suggested that they function in inter-organismal interactions such as symbiosis and defense rather than intraorganismal interactions (Kunzler 2015). In accordance with latter function, several fruiting body lectins were demonstrated to be toxic towards insects and nematodes (Bleuler-Martinez et al. 2011; Sabotic et al. 2016).

In addition to the diversity of their carbohydrate specificities, fungal lectins adopt a variety of 3D structures. Indeed, several lectin folds have so far only been identified in the fungal kingdom (Varrot et al. 2013; Hassan et al. 2015), e.g. the novel  $\beta$ -prism III fold (Cabanettes et al. 2018). To date, 12 structural lectin families have been identified in higher fungi based on protein folds, such as  $\beta$ -propeller,  $\beta$ -trefoil,  $\beta$ -sandwich,  $\alpha/\beta$  mixed fold, or  $\beta$ -prism among others. Classification of lectins based on 3D structures has recently been updated in the new database *Unilectin3D*. In the MycoLec subsection of this database, 68 lectin classes were predicted in fungal genomes (Bonnardel et al. 2019; Lebreton et al. 2021).

Structural and biochemical characterization of fungal lectins has offered valuable clues to their biological function and laid the basis for their use as tools in glycobiological

research. For example, lectins may be applied for the physical separation and structural and functional characterization of glycans, glycoconjugates and cells (Wu et al. 2009; Singh et al. 2020; Tsaneva and Van Damme 2020). Fucose-binding lectins are of special interest as fucose residues commonly occur as terminal or core modifications of cell surface glycans including both N- and O-linked glycans in glycoproteins and glycolipids (Becker and Lowe 2003; Aeschbacher et al. 2017; Paschinger and Wilson 2019; Thomes and Bojar 2021). In mammals, fucosylated glycans, as ABO and Lewis blood group antigens or core fucose, play a vital role in the modulation of immunity, in cell adhesion, in host-pathogen or host-microbiome interactions. Aberrant fucosylation is associated with diseases such as cancer (Li et al. 2018; Holdener and Haltiwanger 2019; Galeev et al. 2021; Kononova et al. 2021).

X-ray crystallography is so far the most commonly used method to determine the 3D structure of a protein and its interactions with specific ligands at the atomic level. The affinity and thermodynamics of protein–ligand interactions is usually obtained by isothermal titration calorimetry (ITC; Dam and Brewer 2007). With the exception of some bacterial examples, lectins generally bind to monosaccharides with low affinity (dissociation constant  $K_D$  in millimolar range). This is often compensated by multivalency, that is the occurrence of multiple binding sites with identical or different glycan specificities, resulting in increased affinity or avidity of the lectin to glycoconjugates displaying multiple glycoepitopes (Cecioni et al. 2015). In addition, affinity of an individual binding site tends to be higher towards oligosaccharides than monosaccharides ( $K_D$  in micromolar range) as a result of additional interactions (Turnbull et al. 2004; Schubert 2017; Bermeo et al. 2020). This is especially true for oligosaccharides with restricted flexibility due to stabilization of a certain conformer as observed in many fucosylated oligosaccharides (Aeschbacher et al. 2017; Lepsik et al. 2019).

*Coprinopsis cinerea* is the mushroom with the largest number of structurally characterized fruiting body lectins up to date: two members of the galectin family (CGL1 and CGL2) specific for  $\beta$ -galactosides (Walser et al. 2004), 1 galectin-related lectin (CGL3) specific for chitooligosaccharides and LacdiNAc (Walti et al. 2008), and the two Ricin-B lectins CCL1 and CCL2 specific for  $\alpha$ 1,3-fucosylated chitobiose (Schubert et al. 2012; Bleuler-Martinez et al. 2017). In this article, a high-resolution crystal structure of a novel fucoside-binding lectin, CML1, from *C. cinerea* fruiting bodies is described in its free form and in complex with blood group H type I ( $Fuc\alpha 1-2Gal\beta 1-3GlcNAc$ ) trisaccharide or Lewis A ( $Gal\beta 1-3[Fuc\alpha 1-4]GlcNAc\beta 1-3Gal$ ) tetrasaccharide. CML1 adopts a novel fold that has so far not been described for any other lectin so far: the protein forms trimers of homodimers with the six carbohydrate-binding sites found on the peripheral side of the dimer interfaces, creating a novel kind of trefoil arrangement. This hexameric arrangement of CML1 has been confirmed in solution by different biophysical methods. The lectin is produced exclusively in the fruiting bodies and homologues are encoded in the genomes of other Agaricomycetes. The toxicity of CML1 towards the bacterivorous nematode *Caenorhabditis tropicalis* suggests a biological function in the defense of Agaricomycetes against predation by fungivorous nematodes.

## Results

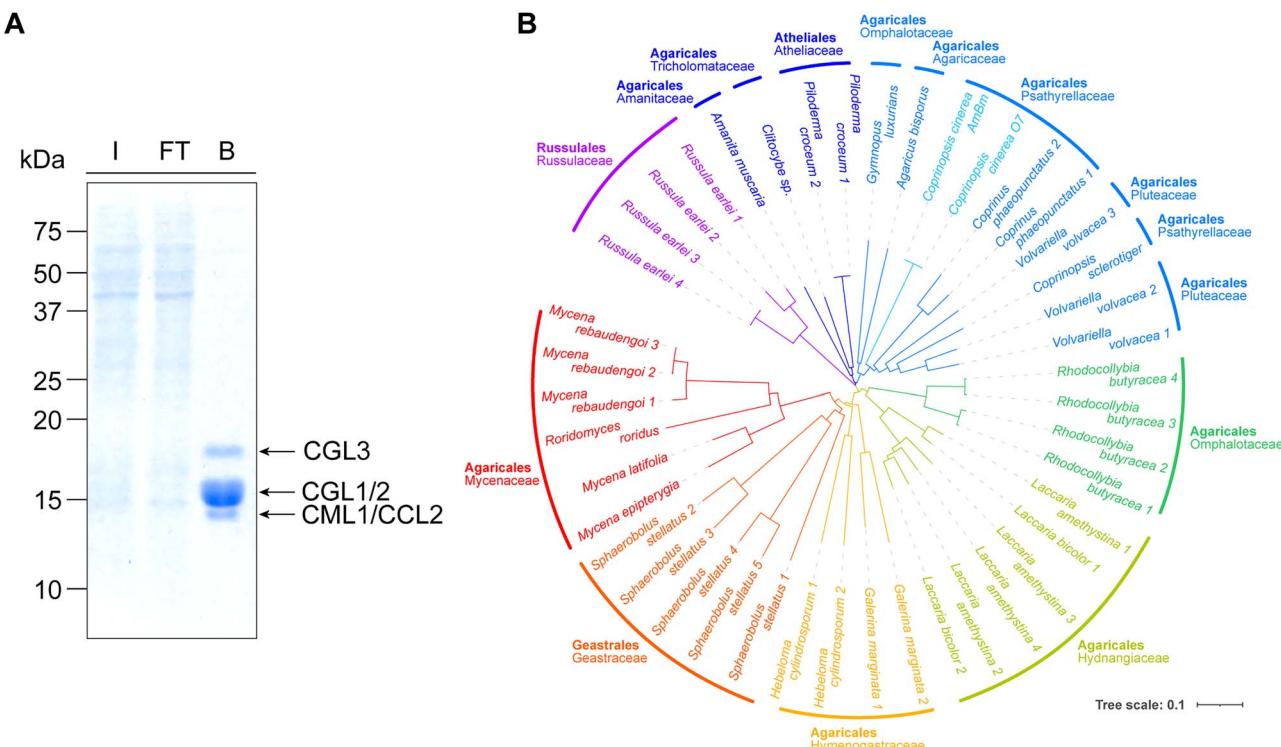
### Identification, cloning, and heterologous expression of CML1

In a search for novel fungal lectins, we employed affinity chromatography of soluble protein extracts from *C. cinerea* AmutBmut fruiting bodies over porcine stomach mucin immobilized on sepharose. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis yielded four distinct Coomassie-stainable bands in the beads fraction containing proteins that specifically bound to the affinity matrix (B: beads in Fig. 1A). Subsequent analysis by mass spectrometry (MS) revealed that the upper three bands corresponded to the previously described lectins CGL1, 2, and 3 (Walser et al. 2004; Walti et al. 2008). Within the lowest band, corresponding to a size of  $\sim$ 14 kDa, peptides from 2 different proteins were identified: the formerly characterized lectin CCL2 (Schubert et al. 2012; Bleuler-Martinez et al. 2017) and a hypothetical protein (JGI ID 439037) encoded in the *C. cinerea* AmutBmut genome. Latter protein, which we termed CML1 for *C. cinerea* Mucin-binding Lectin 1 (GenBank ACF17561), is composed of 127 amino acids with a predicted molecular weight of 13.7 kDa. The CML1 coding sequence from *C. cinerea* strain AmutBmut (Genbank EU796933) is identical to the one from the monokaryotic strain Okayama 7 (CC1G\_10558) and does not contain any intron.

Basic Local Alignment Search Tool (BLAST) searches with the amino acid sequence of CML1 did not reveal similarity to any characterized protein. However, homologues in several other basidiomycetes were identified with the largest number of hits obtained when searching the Mycosom database (Grigoriev et al. 2014). According to this analysis, CML1 homologues are restricted to the class of Agaricomycetes with most examples encoded by members of the order Agaricales and some homologues identified within Gasterales, Atheliiales, and Russulales (Fig. 1B). Although most examples are single-domain proteins, two homologues comprise additional domains: a tandem repeat protein from *Hebeloma cylindrosporum* contains two CML1 domains and an N-terminal domain predicted to adopt a  $\beta$ -trefoil fold according to Phyre 2 (Kelley et al. 2015). Furthermore, one of the two homologues encoded by the *Laccaria bicolor* genome comprises a long N-terminus (1–116) that does not show sequence or structural similarity to any known protein.

### Expression of CML1 in *C. cinerea*

We next assessed the relative expression level of the *cml1* gene in *C. cinerea* during fruiting body formation in immunoblots. For this purpose, the coding region of the gene was cloned and N-terminally His8-tagged CML1 was expressed in *Escherichia coli* BL21(DE3). The recombinant protein was purified by metal affinity chromatography and a polyclonal antiserum against this protein was raised in rabbits. Expression of CML1 in *C. cinerea* was examined by probing immunoblots of soluble protein extracts from vegetative mycelium and fruiting bodies of strain AmutBmut with the antiserum. While the lectin was highly expressed in fruiting bodies, no expression was observed in vegetative mycelium grown on rich medium (Fig. 2A). These results correlate very well with RNA sequencing data previously obtained when investigating fruiting body formation in *C. cinerea* (Plaza et al. 2014; Muraguchi et al. 2015). Interestingly, analysis of *cml1*



**Fig. 1. A)** Identification of CML1 by affinity chromatography. Coomassie-stained SDS-PAGE of fractions obtained by affinity chromatography of *Coprinopsis cinerea* fruiting body extracts using immobilized porcine stomach mucin. I: input, FT: flow through, and B: beads. The mobility and molecular mass of the marker proteins is indicated. **B)** Phylogenetic tree of CML1 homologues based on the Mycocosm database (Grigoriev et al. 2014), using the Simple Phylogeny tool from EMBL-EBI ([https://www.ebi.ac.uk/Tools/phylogeny/simple\\_phylogeny/](https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/); Madeira et al. 2019) and iTOL (Letunic and Bork 2019) for analysis and visualization.

gene expression across all stages of fruiting body development revealed a peak of expression in the stipes of intermediate stage fruiting bodies (Supplementary Fig. S1; Muraguchi et al. 2015).

## In vitro carbohydrate-binding specificity of CML1

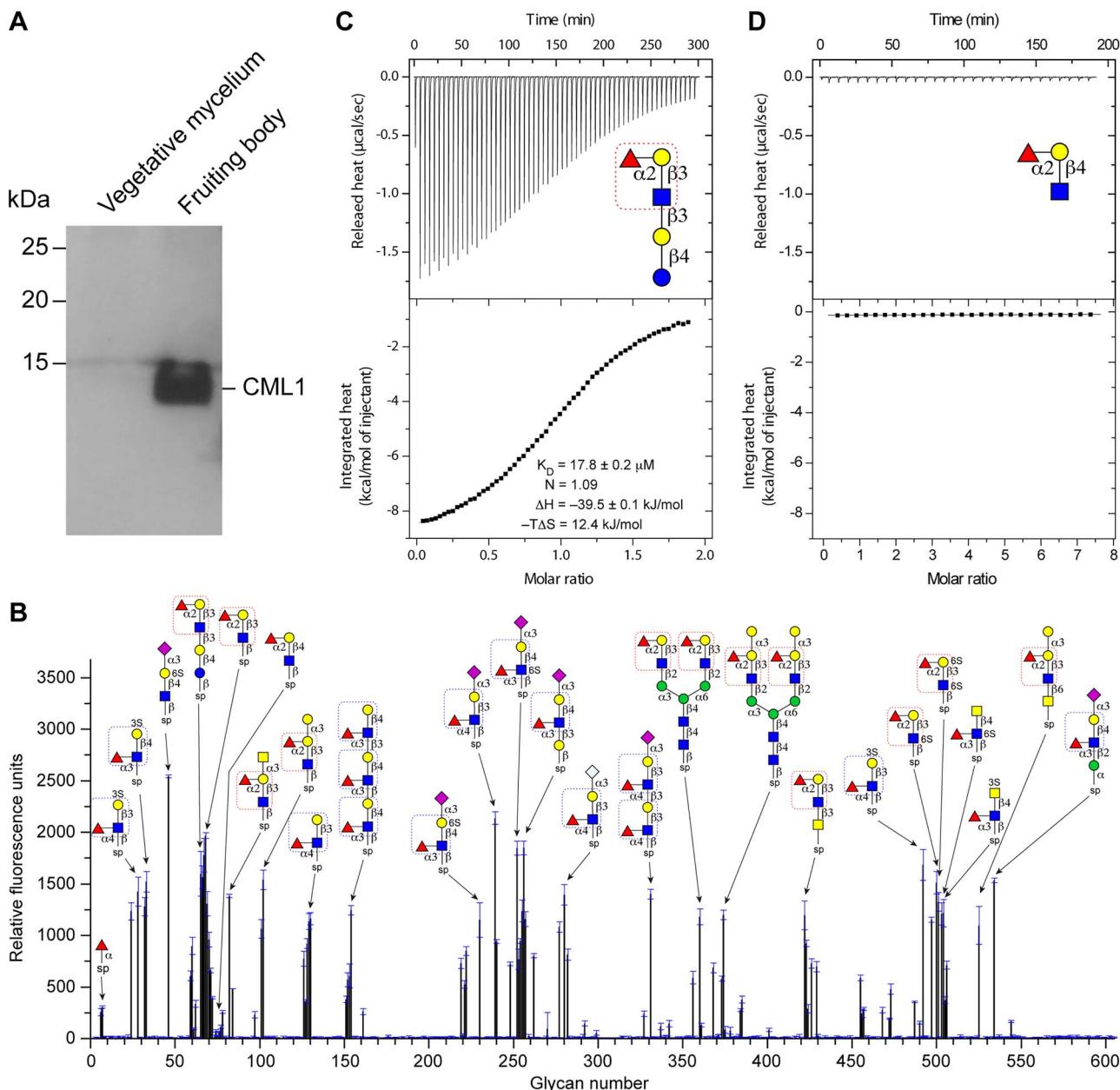
In order to determine the carbohydrate-binding specificity of CML1, the mammalian glycan array by the Consortium of Functional Glycomics (CFG; V5.1, comprising 610 different glycans) was probed with recombinant, fluorescently labeled untagged protein. In this assay, CML1 showed a strict binding specificity for fucose-containing oligosaccharides such as human ABO blood group determinants. The only exception is glycan #46 (Neu5Ac $\alpha$ 2-3(6S)Gal $\beta$ 1-4GlcNAc $\beta$ -Sp8), which does not contain any fucose and yet yielded the highest fluorescence signal (Fig. 2B, Supplementary data). The observed carbohydrate specificity is independent of the CML1 concentrations used, as different dilutions of the labeled lectin resulted in comparable results in glycan array analysis (Supplementary data).

High fluorescence signals were observed for complex oligo/polysaccharides containing the blood group H-type I trisaccharide, Le<sup>x</sup>, Le<sup>a</sup>, as well as their sialylated and/or sulphated forms. Except for glycan #46, all bound oligosaccharides contained fucose in  $\alpha$ 1-2,  $\alpha$ 1-3, and  $\alpha$ 1-4, but not  $\alpha$ 1-6 linkage. Binding was also observed to L-fucose monosaccharide, albeit with low fluorescence (Fig. 2B, Supplementary data). This observation is well in accordance with the binding of the recombinant protein to fucosylsepharose beads.

ITC was used to determine the association/dissociation constants and the thermodynamic parameters of the interactions of untagged CML1 with lacto-N-fucopentaose I ( $\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}3\text{GlcNAc}\beta 1\text{-}3\text{Gal}\beta 1\text{-}4\text{Glc}$ , represented e.g. by glycan #65 in the CFG array), a pentasaccharide that contains the H-type I blood group motif, and the H-type II blood group trisaccharide ( $\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}4\text{GlcNAc}$ , represented e.g. by glycan #76 in the CFG array) (Fig. 2C and D). In accordance with the glycan array data, high affinity of CML1 was observed to lacto-N-fucopentaose, whereas no interaction was detected with H-type II trisaccharide. The interaction of CML1 with lacto-N-fucopentaose is enthalpy-driven, with a binding affinity in the micromolar range ( $K_d = 17.8 \mu\text{M}$ ) and a stoichiometry of 1:1, indicating that there is one binding site per protomer and that this site is completely occupied by the ligand. The observed affinity is in the expected range for the recognition of oligosaccharides by lectins. For example, the N-terminal fucose binding domain of Bc2L-C from *Burkholderia cenocepacia* displays a  $K_d$  of 26  $\mu\text{M}$  for the globo-H hexasaccharide (Bermeo et al. 2020).

## Crystal structure and oligomerization state of CML1

We solved the crystal structure of the untagged CML1 protein in its apo form using a platinum derivative at 1.35 Å resolution in *P*<sub>2</sub>1 space group ([Supplementary Table SIV](#)). Six CML1 protomers were found in the asymmetric unit arranged as a trimer of dimers ([Fig. 3A and B](#)). Each protomer adopts a  $\beta$ -sandwich fold consisting of one 5-stranded and one 6-stranded antiparallel  $\beta$ -sheet with one helix perpendicular to the  $\beta$ -sheets ([Fig. 3A](#)). The 5-stranded  $\beta$ -sheet is mainly involved in the formation of the hexamer and is buried,

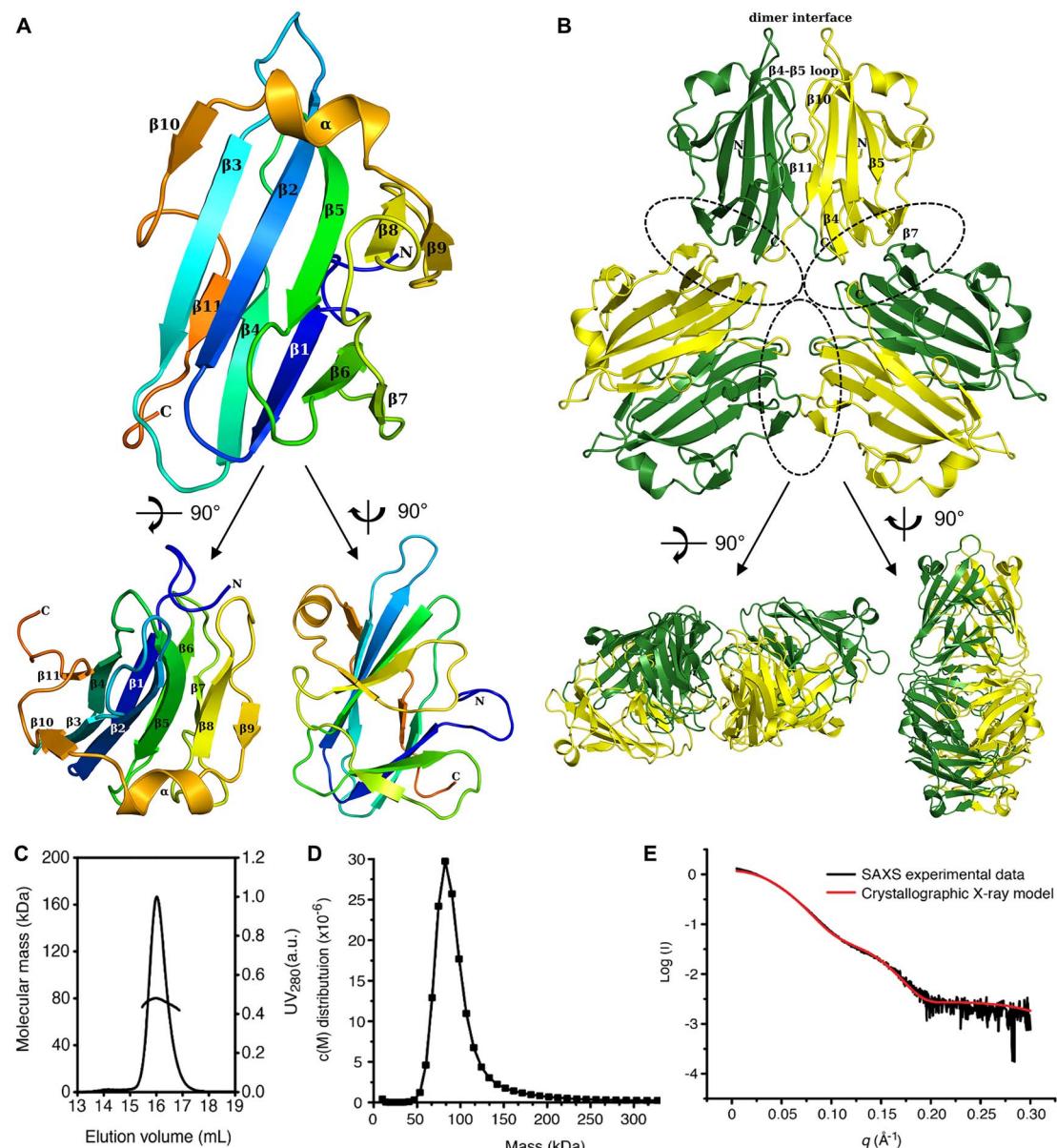


**Fig. 2.** **A**) Immunoblot showing the differential expression of *cml1* in vegetative mycelium and fruiting bodies of *Coprinopsis cinerea* (strain AmutBmut). Specific anti-CML1 polyclonal antiserum was used for detection and equal amounts of total soluble proteins were loaded per sample. **B**) Glycan array data of CML1. Fluorescently labeled CML1 was analyzed for binding to the mammalian glycan array (V5.1) of the Consortium for Functional Glycomics (CFG). Results shown are averages of triplicate measurements of fluorescence intensity at a lectin concentration of 50  $\mu\text{g}/\text{mL}$ . Error bars indicate the standard deviations of the mean. Glycan structures are depicted for representative epitopes with highest relative fluorescence. The raw data and the entire list of glycans with the respective spacers can be found on the CFG homepage [<http://functionalglycomics.org>] or in the Supplementary data. **C** and **D**) Isothermal titration calorimetry (ITC) measurements of CML1 with lacto-N-fucopentaose containing H-type I blood group antigen and H-type II blood group trisaccharide ( $\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}4\text{GlcNAc}$ ), respectively. Titration curves at 25°C and schematic ligand structures are shown in the upper panel; the lower panel displays the total heat released as a function of total ligand concentration. The solid line represents the best least squares fit.

whereas the 6-stranded  $\beta$ -sheet and the helix are exposed and facing outward. The dimers result from a 2-fold axis along an interface involving the N-terminal loop, the surface loop between the  $\beta 2$  and  $\beta 3$  strands, the  $\beta 4$  strand and its connecting loop with the  $\beta 5$  strand and residues from the  $\beta 10$  strand to the C-terminus leading to buried surface area of  $\sim 1,565 \text{ \AA}^2$  (Fig. 3B). The trimer results from three 2-fold axes on the dimers involving  $\beta 7$  strands and the C-terminal loops leading to a buried surface area of  $\sim 545 \text{ \AA}^2$  (Fig. 3B). The interior of the hexamer is formed by all strands

from the 5-stranded  $\beta$ -sheet and the C-terminus of each protomer.

To confirm the hexameric state in solution, we determined the mass of native CML1 by multi-angle light scattering (MALS), analytical ultracentrifugation (AUC), sedimentation velocity, and small-angle X-ray scattering (SAXS) methods. Based on the theoretical molecular mass of one protomer (13.5 kDa, UniProt B3VS76), which was confirmed by denaturing SDS-PAGE (data not shown), we expected a mass of  $6 \times 13.5 \text{ kDa} = 81 \text{ kDa}$  for the hexamer. MALS, coupled with



**Fig. 3.** Structure and oligomerization state of CML1 in crystal and in solution. **A)** Views of CML1 protomer rainbow colored from N to C-terminus with labelling of the secondary structure. **B)** views of CML1 hexamer from 3 different angles. Protomers are colored in green and yellow for each dimer; the interfaces creating the trimer are circled. **C)** MALS analysis. The UV absorbance profile of size exclusion chromatography is overlaid with the molecular weight determination by MALS. **D)** Sedimentation velocity analysis. The continuous molar mass distribution of purified CML1 is shown. **E)** Small-angle X-ray scattering analysis. The theoretical scattering curve calculated from the crystallographic structure of CML1 (red) is superimposed with the experimental X-ray scattering curve (black).

size exclusion chromatography, yielded a molecular mass of 80 kDa (Fig. 3C). The AUC experiment provided a value of 84.5 kDa (Fig. 3D). Finally, the theoretical SAXS curve calculated from the hexameric crystal structure of CML1 fitted very well the X-ray experimental scattering curve (Fig. 3E) confirming that the hexameric state of CML1 found in 2 different crystal forms corresponds to the one in solution.

#### Structural homology of CML1 to other proteins

A classical search for structural homologs with a lower acceptable match of secondary structure elements of 70% was performed in the Protein Data Bank with PDBeFold (Krissinel and Henrick 2005). This search yielded only one match, namely the intracellular growth loci protein E (IGLE) from the

Gram-negative coccobacillus *Francisella tularensis* (Robb et al. 2010). IGLE is an outer membrane-associated lipoprotein, essential for intracellular survival and murine virulence in type A tularemia and associated with Francisella pathogenicity island (FPI) encoding a type VI secretion system (T6SS). This protein is monomeric and shows a root mean square deviation (RMSD) of about 4 Å from CML1. Its fold essentially matches for the 4 central β-strands of the 6-stranded β-sheet of CML1. However, the surface loops, helices and terminal region are very different (Supplementary Fig. S2A). For a PDBeFold search applying a lower acceptable match of 50%, only structures of the subunit of the *Bacillus stearothermophilus* TRP RNA-binding attenuation protein (TRAP) were found, where 3 strands in each β-sheet aligned (Hopcroft et al. 2002;

Chen et al. 2011). TRAP is only composed of a  $\beta$ -sandwich with 4 and 3 antiparallel stranded  $\beta$ -sheets connected by short loops (Supplementary Fig. S2B). A search against the PDB with the DaliLite server (Holm 2020) revealed matches with proteins containing a  $\beta$ -sandwich domain or antiparallel  $\beta$ -sheets such the carbohydrate-binding domain of carbohydrate-active enzymes and several uncharacterized or hypothetical proteins. The top hits (Z-score around 5) were structures of bacterial proteins belonging to the telluride resistance family (TerD, PDB IDs: 2QNG, 2KXT, and 2QZ7), which also adopt a  $\beta$ -sandwich fold (Supplementary Fig. S2C; Pan et al. 2011). Only 54% of the amino acids can be aligned with a RMSD of 2.8 or 3.2 Å. The arrangement and the length of the  $\beta$ -strands, however, differ with strong discrepancies in the surface loops. Similar conclusions can be drawn for the hypothetical protein TM1070 from *Thermotoga maritima* exhibiting 13% sequence identity and aligning with an RMSD of 3 Å (Supplementary Fig. S2D). Taken together, these results suggest that CML1 is the first member of a novel family of  $\beta$ -sandwich lectins.

### Structural basis of CML1 ligand specificity

CML1 structures in complex with H-type I trisaccharide or Lewis A tetrasaccharide were solved by molecular replacement at 1.55 and 1.95 Å resolution, respectively in *P3*<sub>1</sub> space group. Statistics on data collection and refinement are provided in Supplementary Table SIV. Again, one hexamer is found in the asymmetric unit; no significant differences are observed in the tertiary structure compared with the apo monoclinic structure, even in loop regions (RMSD values < 0.4 Å). Analysis of the electron density unambiguously revealed one sugar binding site per protomer lying at the dimer interface (Fig. 4A–C). All 3 sugar moieties of the H-type I blood group trisaccharide could be identified and modelled in each binding site. Apart from the fucose residue located within a pocket created by the  $\beta$ 10 strand and the  $\beta$ 2– $\beta$ 3,  $\beta$ 4– $\beta$ 5, and  $\beta$ 10– $\beta$ 11 surface loops from one protomer and the  $\beta$ 2– $\beta$ 3,  $\beta$ 4– $\beta$ 5, and  $\beta$ 8– $\beta$ 9 loops from the neighboring protomer, the other sugar moieties are quite exposed to the solvent. The binding surface is positively charged, as shown by the electrostatic surface potential, with no significant differences being observed between the 6 binding sites (Fig. 4B). This surface property might explain the preference for negatively charged ligands such as sialylated or sulfated oligosaccharides in the glycan array. Nevertheless, there is not clear conservation of these residues, suggesting that the carboxyl- or sulfo-groups are not specifically recognized, but rather electrostatic attraction will lead to an increase in the affinity due to faster  $k_{ON}$  kinetics.

Direct contacts involved residues His54, Asn55, and Trp94 from one protomer and Arg114\* (\*from the adjacent protomer of each dimer, Fig. 4D). Arg114\* is the most important residue as its side chain forms simultaneously 4 hydrogen bonds with the fucose. Arg114\* N $\gamma$ 2 forms one hydrogen bond with Fuc O2 and one with the oxygen from the glycosidic linkage between the fucose and the galactose. The side chain nitrogen N $\epsilon$  of Arg114\* forms one hydrogen bond with Fuc O2 and one with Fuc O3. His54 N $\delta$ 1 forms one hydrogen bond with Fuc O3 and the Asn55 main chain nitrogen atom forms one with the axial hydroxyl Fuc O4. Trp94 N $\epsilon$ 1 forms a hydrogen bond with GlcNAc O6. In addition, the Tyr56 and Val117\* main chain oxygens and Asn55 O $\delta$ 1 interact with Fuc O4, Fuc O3, and Fuc O2, respectively, through a water molecule. The galactose does not form any direct contact with

the protein but there is a conserved water mediated interaction between Gal O3 and Asn55 O $\delta$ 1 (Fig. 4D). Hydrophobic contacts are observed with the side chains of Leu28, Val34\*, Thr57, and Trp94. The  $\Phi$  and  $\psi$  torsion angles observed for each glycosidic bond are very close to those calculated for the main energy minimum described for each disaccharide in Glyco3D (Perez et al. 2015).

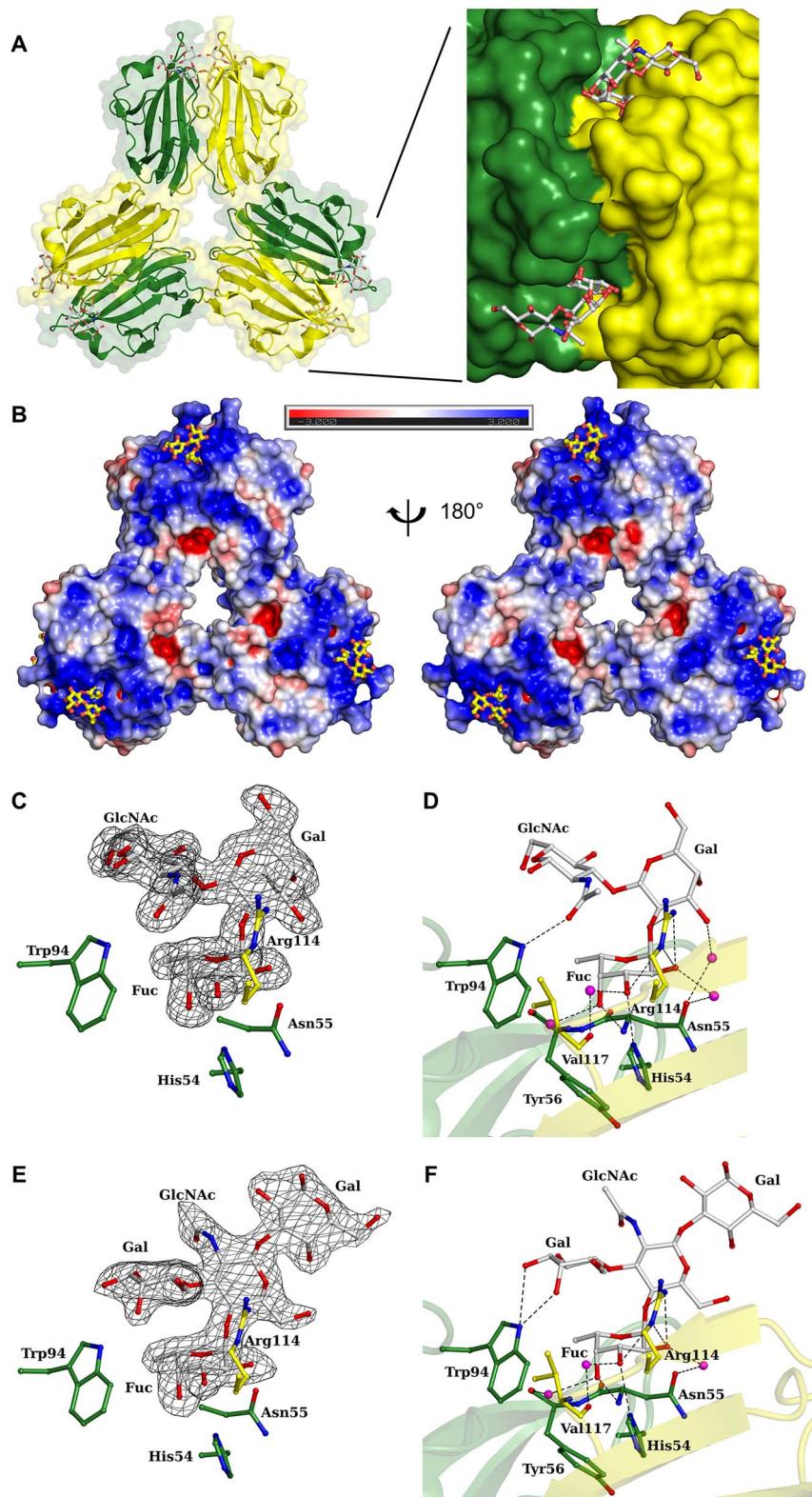
In the complex structure with Lewis A tetrasaccharide, electron density for the reducing galactose was only visible in the binding site of protomers C and D, whereas the trisaccharide was modelled in all binding sites (Fig. 4E). The interactions with the fucose moiety are the same as described above. Direct hydrogen bonds are observed for Trp94 N $\delta$ 1 atom and the O3 and O4 hydroxyls of the nonreducing galactose (Fig. 4F). The  $\Phi$  and  $\psi$  torsion angles observed for each glycosidic bond are in the periphery of the main energy minimum. When the 2 complex structure are overlaid, the sugar ring of galactose in H-type I or GlcNAc in Lewis A are 180° apart. This allows the reducing sugar to be correctly positioned for interaction with Trp94. The position of GlcNAc O1 and GlcNAc O7 is very similar to the one observed for Gal O3 and Gal O4, respectively.

### Functional confirmation of ligand-binding residues by mutagenesis

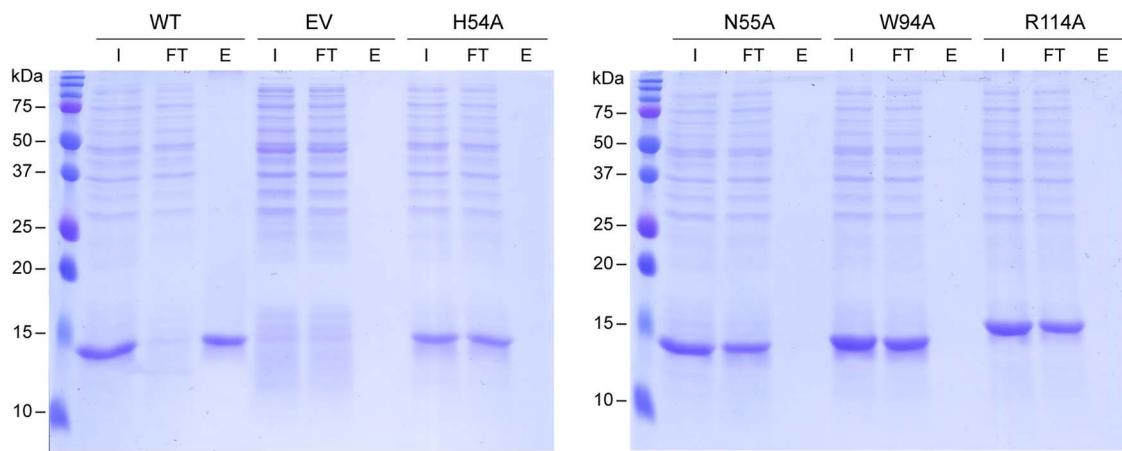
Based on the observed interactions in the structure of the complexes between the CML1 lectin and fucosylated oligosaccharides, 4 mutants involving the residues mentioned above were constructed by site-directed mutagenesis: His54Ala (H54A), Asn55Ala (N55A), Trp94Ala (W94A), and Arg114Ala (R114A). The expression levels of the various CML1 variants in the respective BL21 transformants were comparable with the wildtype protein except for H54A, for which it was slightly lower (Supplementary Fig. S3). The carbohydrate-binding ability of these mutant CML1 proteins compared with the wild type protein was assessed by binding tests to fucosyl-sepharose beads using soluble protein extracts of bacteria expressing the respective CML1 variants. In accordance with the structural predictions, all 4 mutations abolished binding of the protein to the matrix, whereas the wildtype protein was quantitatively bound and eluted using L-fucose (Fig. 5). The W94A and R114A variants of CML1 showed a slightly reduced mobility in SDS-PAGE compared with the wildtype protein (Fig. 5 and Supplementary Fig. S3). The results confirm the involvement of these 4 residues in the binding of CML1 to carbohydrates.

### Nematotoxicity of CML1

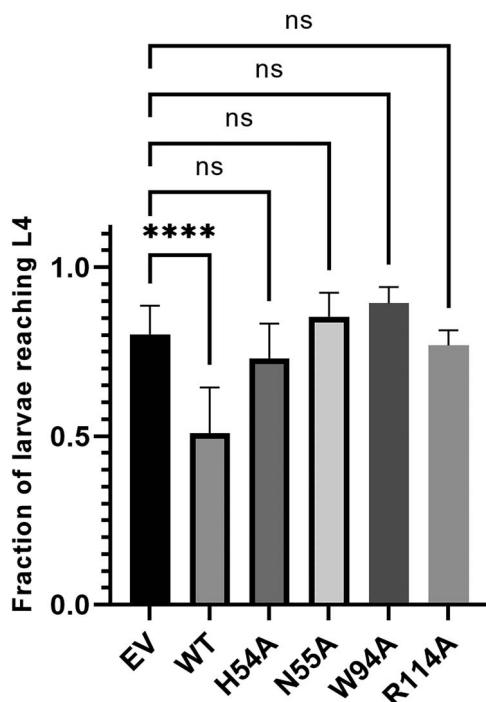
Based on the nematotoxicity of previously characterized fruiting body lectins (Kunzler 2015; Sabotic et al. 2016), including mucin-binding CGL2 and CCL2 from *C. cinerea* (Butschi et al. 2010; Schubert et al. 2012; Bleuler-Martinez et al. 2017), the toxicity of CML1 towards nematodes was assessed. For this purpose, the development of L1 larvae of bacterivorous nematodes feeding on *E. coli* transformants containing empty vector or plasmids directing the expression of CML1 in the bacterial cytoplasm was examined as previously described (Kunzler et al. 2010). Due to the recently reported differences in the susceptibility of different bacterivorous nematode species towards different fungal defense effector proteins (Tayyrov et al. 2019; Tayyrov et al. 2021), 2 nematode species, *C. elegans* and *C. tropicalis*, were included in the



**Fig. 4.** Binding interactions of fucosylated oligosaccharide with CML1. **A)** Surface representation of CML1 complexed with H-type 1 blood group trisaccharide ( $\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}3\text{GlcNAc}$ ) with a zoom on the dimer interface. Protomers are colored in green and yellow for each dimer. **B)**  $\pm 3 \text{ kT/e}$  electrostatic surface potential of hexameric CML1 calculated with the AMBER force field using PBD2PQR (Dolinsky et al. 2004) and APBS (Baker et al. 2001) in PyMOL for the 2 views of the CML1 hexamer (top and bottom). Blue and red denote positively and negatively charged surface; respectively. The carbon atoms of the ligand are colored in yellow. **C)** 2mFo-DFc calculated electron density map contoured at 1 sigma level ( $0.39 \text{ e}\AA^3$ ) around H-type I blood group antigen with close up view of the interactions **D).** **E)** 2mFo-DFc calculated electron density map contoured at 1 sigma level around Lewis A tetrasaccharide ( $\text{Gal}\beta 1\text{-}3[\text{Fuc}\alpha 1\text{-}4]\text{GlcNAc}\beta 1\text{-}3\text{Gal}$ ,  $0.34 \text{ e}\AA^3$ ) with close up view of the interactions **F).** Amino acids are displayed in balls and sticks and color in green and yellow to differentiate protomers. H-bonds are represented as dashed lines and water as pink spheres.



**Fig. 5.** Binding of CML1 wildtype (WT) and variants thereof (H54A, N55A, W94A, and R114A) to fucosyl-sepharose. Affinity chromatography was performed using soluble protein extracts of respective CML1-producing and empty vector control (EV)-containing *Escherichia coli* BL21 transformants. Samples of input (I), flow through (FT) and eluate (E) were analyzed by SDS-PAGE followed by Coomassie-blue staining. Eluate fractions were 3 times as concentrated as input and flow-through fractions.



**Fig. 6.** Toxicity of CML1 and variants thereof towards bacterivorous nematode *Caenorhabditis tropicalis*. The development of L1 larvae feeding on *Escherichia coli* expressing CML1 wildtype protein (WT) or variants thereof (H54A, N55A, W94A, and R114A), or containing an empty vector (EV) to the L4 stage was assessed. The error bars indicate the standard deviation of 5 replicates. Comparisons between conditions were performed using Dunnett's multiple comparison test (ns, not significant; \*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001; \*\*\*\*P < 0.0001).

assay. CML1 was toxic to *C. tropicalis* but not to *C. elegans* (Supplementary Fig. S4). Only 50% of *C. tropicalis* larvae fed on bacteria expressing CML1 reached adulthood, compared to 80% in the negative control (Fig. 6).

In order to find out whether the toxicity of CML1 towards *C. tropicalis* was dependent on carbohydrate-binding, the nematotoxicity assay was performed with *E. coli* transformants producing CML1 variants with above mentioned mutations in the carbohydrate-binding residues.

Indeed, all 4 CML1 variants exhibited reduced nematotoxicity compared with wildtype CML1 (Fig. 6). Considering the inability of the investigated CML1 mutant proteins to bind to fucosyl-sepharose (Fig. 5), these results suggest that nematotoxicity of CML1 is dependent on its ability to bind carbohydrates.

## Discussion

In this study, we report the identification and the structural, biochemical and functional characterization of a novel fucoside-binding fruiting body lectin, termed CML1, from the mushroom *C. cinerea*. Homologues of CML1 are encoded in the genomes of other fungi, specifically in the class of Agaricomycetes within the phylum Basidiomycota, suggesting that the protein is the founding member of a novel family of fungal lectins. Not only the amino acid sequence but also the regulation appears to be conserved, as, similar to the *cml1* gene, the gene coding for the *L. bicolor* CML1 homologue is expressed specifically at the fruiting body stage (Martin et al. 2008).

The tertiary structure of CML1 shows a novel  $\beta$ -sandwich architecture not previously described for lectins. Low structural homology was observed with other proteins adopting a  $\beta$ -sandwich fold. This fold represents, similar to the  $\beta$ -propeller fold, a very versatile fold that is able to accommodate different types of ligands and functions (Notova et al. 2020). The disposition and number of the strands and surface loops of CML1, however, substantially differ from other  $\beta$ -sandwich fold proteins, justifying the claim of a novel family within this superfamily.

Crystal structures of the complex between CML1 and H-type I or Lewis A blood group antigens revealed that the carbohydrate-binding sites of the hexameric lectin are localized at the dimer interface with amino acid residues from both protomers contributing to ligand-binding. Ligand-binding at the oligomerization interface has also been observed in other, structurally unrelated lectins, in particular in 2 bacterial fucoside-binding lectin families represented by the TNF-like lectin Bc2L-C from *B. cenocepacia* (Sulak et al. 2010) and the  $\beta$ -propeller lectin RSL from *Ralstonia solanacearum* (Kostlanova et al. 2005), suggesting convergent

evolution of this binding mode. Similar to these bacterial fucoside-binding lectins, an arginine residue is one of the critical amino acid residues involved in fucoside-binding of CML1. This is also true for the fungal  $\beta$ -propeller fucose-binding lectin AAL (Wimmerova et al. 2003), but not for the  $\alpha$ 1–3 fucose binding lectin CCL2 from *C. cinerea* where the binding site is mainly formed by the protein backbone and other residues such as valine and glycine (Schubert et al. 2012). The overall binding site topology and correct orientation of the oligosaccharide are necessary to provide additional interactions to increase affinity and define fine specificity. For example, CML1 is able to discriminate between H-type I and H-type II blood group antigens, which only differ in the linkage between the galactose and GlcNAc moieties ( $\beta$ 1–3 in H-type I compare with  $\beta$ 1–4 in H-type II). This enables the N-acetyl group of GlcNAc to interact with Trp94 in H-type I blood group, whereas in H-type II, the GlcNAc will be more exposed to the solvent and unable to interact with the protein. The preference for negatively charged ligands such as sialylated or sulfated oligosaccharides (in one case even lacking a fucose residue) may be due to the positively charged binding surface.

An amino acid sequence alignment of the CML1 homologues shows, that the 4 residues that were identified to be important for the binding of fucose and the nematotoxicity, are not strictly conserved (Fig. 7). Instead of His54, a tyrosine and sometimes a phenylalanine are found resulting in a loss of hydrogen bonding but perhaps in a gain in hydrophobic interaction. Asn55 is often replaced by a threonine, which could maintain all interactions of this residue. Trp94 and Arg112 are essentially conserved, apart from the 5 CML1-like proteins from *Sphaerobolus stellatus*, the 3 proteins from *Mycena rebaudengois* and one protein from *Roridomyces roridus* (Fig. 7). Thus, these proteins are expected not to recognize fucosylated oligosaccharides. In addition, differences are observed in residues on the oligomerization interface. The function of these CML1 homologues remains to be determined.

Many fucoside-binding lectins from all kingdoms have been characterized. In terms of protein structure, several characterized fucoside-binding lectins from the fungal kingdom adopt a 6-bladed  $\beta$ -propeller fold: *Aleuria aurantia* agglutinin (AAL) and *Aspergillus oryzae* lectin (AOL; Wimmerova et al. 2003; Matsumura et al. 2004) as well as 2 AOL-homologues that have recently been characterized in the pathogenic molds *Aspergillus fumigatus* (FleA) and *Scedosporium apiospermum* (SapL1; Houser et al. 2015; Martinez-Alarcon et al. 2021). These lectins are produced mainly by the asexual spores (conidia) of their ascomycete hosts. In contrast, the Ricin-B-type ( $\beta$ -trefoil-fold) isolectins CCL1 and CCL2 were isolated from fruiting bodies of the basidiomycete *C. cinerea* (Schubert et al. 2012; Bleuler-Martinez et al. 2017). In terms of protein function, fucoside-binding lectins often recognize exogenous glycans and play roles in inter-kingdom interactions. As examples, the previously mentioned AOL and homologues play a role in plant and human pathogenesis. Similarly, RSL, a lectin from the plant pathogenic bacterium *Ralstonia solanacearum* is suggested to bind to fucosylated xyloglucans of plant cell walls during the infection process (Kostlanova et al. 2005), and LecB, a lectin of *Pseudomonas aeruginosa*, binds to mucins in the lung and seems to play a key role in host recognition, adhesion, and formation of biofilms by this opportunistic bacterial pathogen (Tielker et al. 2005). In contrast, AAA, a fucoside-binding lectin in the serum of the European eel *Anguilla anguilla*, recognizes bacterial polysaccharides as an

effector of the innate immune system of this animal against bacterial pathogens (Bianchet et al. 2002). Similarly, AAL and CCL1/2 displayed toxicity towards amoebae, nematodes, and insects (Bleuler-Martinez et al. 2011; Schubert et al. 2012; Bleuler-Martinez et al. 2017; Tayyrov et al. 2018) and were, thus, proposed to function as defense effectors of fungal conidia and fruiting bodies against predators and parasites (Kunzler 2015). The toxicity of these lectins was demonstrated to depend on the binding to fucosylated glycoconjugates in insects and nematodes and, more specifically, in the case of CCL1/2, on binding to  $\alpha$ 1–3 fucose modifications in the core of N-glycans in the nematode (Bleuler-Martinez et al. 2011; Schubert et al. 2012).

The differential toxicity of CML1 towards *C. tropicalis* and *C. elegans* is reminiscent of CGL2; however, CGL2 was described to be less toxic to *C. tropicalis* than to *C. elegans* (Tayyrov et al. 2019). Although these bacterivorous nematodes will never encounter fungal lectins under natural conditions, these results suggest nevertheless that there is a constant arms race between the glycoepitopes displayed by the nematode digestive tract and the carbohydrate-specificity of lectins taken up from their prey (bacteria in case of *C. tropicalis* and *C. elegans*). In this regard, the difference in toxicity towards *C. elegans* between AAL (toxic), CCL1/2 (toxic) and CML1 (nontoxic) may be explained by the difference in the fine specificities between these fucose-binding lectins. Although AAL from *A. aurantia* binds fucose and oligosaccharides displaying terminal fucose in  $\alpha$ 1–2,  $\alpha$ 1–3,  $\alpha$ 1–4, or  $\alpha$ 1–6 linkage (Fukumori et al. 1990; Matsumura et al. 2007), CCL1/2 from *C. cinerea* only recognize  $\alpha$ 1–3 linked fucosides (Schubert et al. 2012; Bleuler-Martinez et al. 2017). The carbohydrate specificity of CML1 lies somewhere in between as the protein binds to glycans of the mammalian glycan array containing terminal fucose in  $\alpha$ 1–2,  $\alpha$ 1–3, and  $\alpha$ 1–4 but not  $\alpha$ 1–6 linkage. Binding of CML1 to the monosaccharide L-fucose was observed both in the glycan array and by binding to fucosyl-sepharose. However, when co-crystallization was tried with fucose, no fucose could be detected in the crystals, suggesting very weak or no binding.

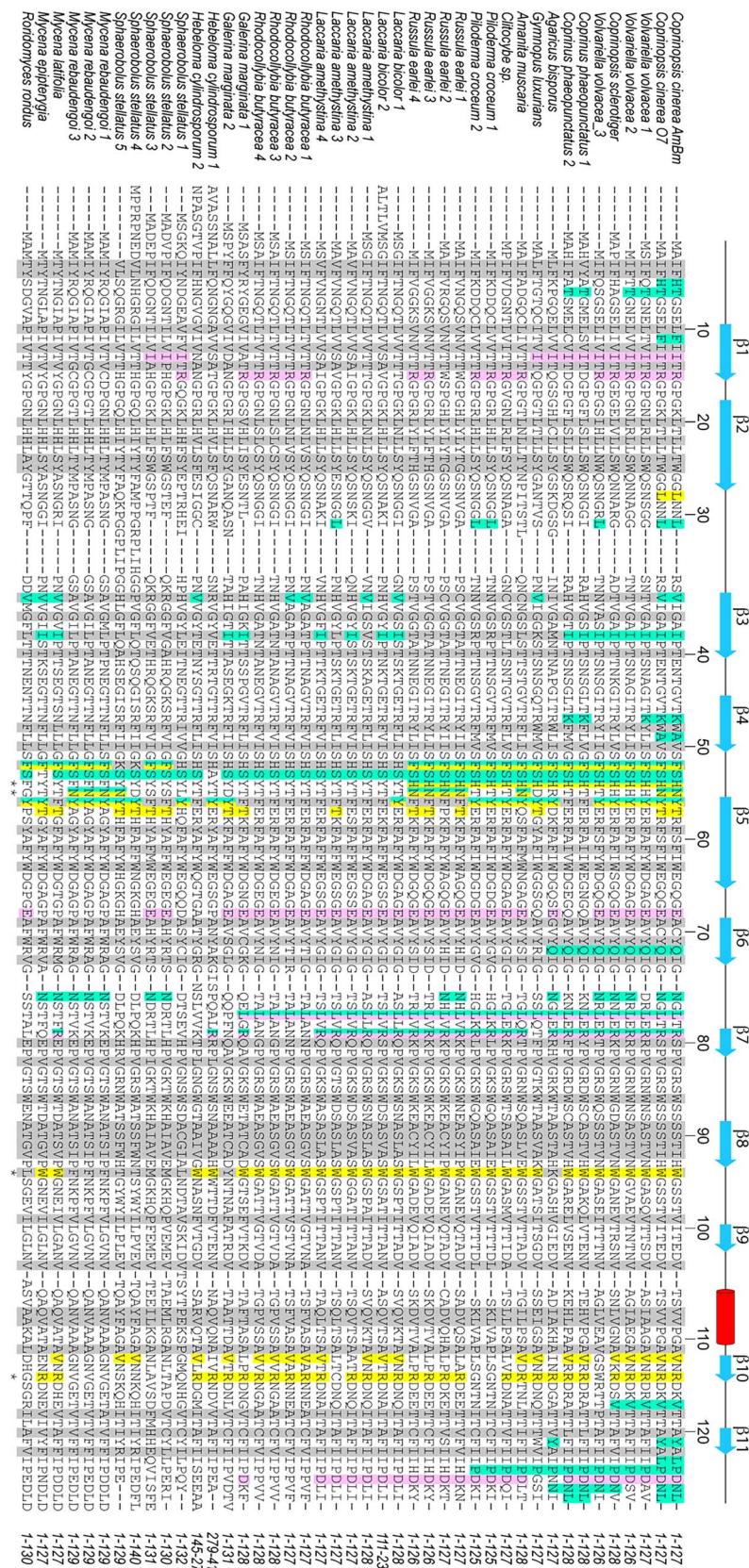
In conclusion, we present a novel lectin family extending the hitherto described portfolio of structural families of fungal lectins. The protein forms trimers of dimers in an unusual trefoil arrangement demonstrating the versatility of lectins to gain multivalency and, concomitantly, avidity. The protein hexamer contains 6 fucoside-binding sites, which are located at the dimer interfaces of each protomer. Interestingly, the lectin strongly prefers H-type I blood group glycoepitopes over H-type II revealing the importance of the interactions with the third sugar unit and explaining the recognition of other fucosylated oligosaccharides like Lewis A. Site-directed mutagenesis of the carbohydrate-binding sites along with toxicity assays demonstrated the importance of carbohydrate-binding for the bioactivity of the protein. Thus, the novel lectin family adds up to the list of lectins used by organisms to recognize and attack antagonists via their glycoepitopes. Based on the specificity of many of these lectins, fucosides appear to play a prominent role in this defense mechanism.

## Materials and methods

### Strains and cultivation conditions

*Escherichia coli* strain DH5 $\alpha$  (Merck Millipore) was used for cloning and amplification of plasmids, strain BL21 (DE3) (Merck Millipore) was used for bacterial expression of fungal

## Novel agaricomycete fruiting body lectin



**Fig. 7.** Amino acid sequence alignment of CML1 homologues. The alignment was generated by Clustal Omega and manually adjusted using the secondary structure elements ( $\beta$ -sheets and  $\alpha$ -helices) from the 3D structure (top). Residues important for the fold, which are located in the hydrophobic core, are highlighted in grey. Conserved residues involved in carbohydrate binding or present at the dimer interface or the contact surface of the trimer of dimers are highlighted in color as indicated. Residue numbers of the aligned sequences are indicated on the right. Asterisks at the bottom indicate residue positions of point mutations. Detailed information for each sequence is given in Supplementary Table SIII; a phylogenetic tree of the aligned sequences is provided in Fig. 1B.

proteins and nematotoxicity assays and strain OP50 (Brenner 1974) was used to feed bacterivorous nematodes for propagation. All *E. coli* strains were cultivated on LB at 37°C for propagation. Special cultivation media and temperatures were used for selection of plasmid-containing strains, protein expression, and nematotoxicity assays (see below).

*Coprinopsis cinerea* strain AmutBmut (Swamy et al. 1984) was maintained on YMG solid medium (0.4% (w/v) yeast extract, 1% (w/v) malt extract, 0.4% glucose (w/v), and 1.5% (w/v) agar) at 37°C. Cultivation conditions and techniques for harvesting vegetative mycelium and fruiting bodies have been described previously (Walti et al. 2006).

*Caenorhabditis elegans* N2 and *C. tropicalis* JU1373 were purchased from the Caenorhabditis Genetics Center (CGC) at the University of Minnesota, United States. Worms were grown on NGM plates (51 mM NaCl, 2.5 g/L bacteriological peptone, 13 mM cholesterol, 1 mM CaCl<sub>2</sub>, 25 mM KPO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, and 1.7% (w/v) agar) pre-seeded with *E. coli* OP50 at 20°C.

### Identification, cloning, and heterologous expression of CML1

Soluble proteins of *C. cinerea* fruiting bodies (strain AmutBmut) were extracted as previously described (Schubert et al. 2012). For affinity chromatography, mucin from porcine stomach type II (Sigma) was coupled to CNBr-activated Sepharose 4B (Cytiva) according to the manufacturer's protocol. 150 μL of the mucin-sepharose beads were equilibrated with at least 10 column volumes of PBS and incubated with 500 μL of the fruiting body soluble extract (input) at 4°C. The flow through was collected and beads were washed with at least 10 column volumes of PBS prior re-suspension and boiling in Lämmli buffer to release proteins bound to the matrix. Proteins from input, flow through and beads were separated by SDS-PAGE. Peptide mass fingerprinting of isolated protein bands was done as previously described (Schubert et al. 2012). Results were analyzed with MASCOT ([www.matrixscience.com](http://www.matrixscience.com)) using the *C. cinerea* AmutBmut genome database ([https://mycocosm.jgi.doe.gov/Copci\\_Amu tBmut1/Copci\\_AmutBmut1.home.html](https://mycocosm.jgi.doe.gov/Copci_Amu tBmut1/Copci_AmutBmut1.home.html); Muraguchi et al. 2015).

As the identified coding sequence of *cml1* (JGI protein ID 439037) contained no predicted introns, the sequence was amplified by PCR from genomic DNA of *C. cinerea* strain AmutBmut using oligonucleotides (locus-fwd and locus-rev, Supplementary Table SI) hybridizing to the untranslated region up and downstream of the open reading frame. The obtained PCR product was cloned into vector pGEM-T Easy (Promega), according to the manufacturer's protocol, resulting in plasmid pGEM-CML1. The sequence of the insert was verified by DNA sequencing. Oligonucleotides (NdeI-CML1-N and BamHI-CML1-C, Supplementary Table SI) containing NdeI and BamHI restriction sites were designed based on the obtained sequence and used to amplify the open reading frame from pGEM-CML1. The PCR product was cloned into vector the pGEM-T Easy and expression plasmids were produced by recloning the insert as NdeI-BamHI fragment into pET22 (Novagen-Merck Millipore) resulting in the plasmid pET22-CML1 (Supplementary Table SII). For protein purification, the plasmid pET22-NHisCML1 encoding an N-terminal His-tagged version of CML1 was constructed accordingly but using the NdeI-CML1-NHis oligonucleotide as forward primer (Supplementary Tables SI and SII).

For heterologous expression of the lectin, the expression plasmids (pET22-CML1 and pET22-NHisCML1) were transformed and expressed in *E. coli* BL21(DE3) cells (Merck Millipore). Cells were grown in LB or TB liquid medium supplemented with 100-μg/mL ampicillin and induced with IPTG at OD<sub>600</sub> = 0.5–1 at 23°C for 16 h. Solubility of the recombinant protein was verified as previously described (Kunzler et al. 2010).

### Phylogeny

The Mycocosm database (Grigoriev et al. 2014) was used for a BLAST search within Agaricomycotina resulting in CML1 and 42 homologous proteins that spanned the entire primary sequence of CML1. For proteins with variable N- or C-terminal additions, the protein-domain borders were manually adjusted to the CML1 sequence visible in the crystal structure. Based the amino acid alignment (Clustal Omega) an unrooted phylogenetic tree was constructed by the Simple Phylogeny tool from EMBL-EBI ([https://www.ebi.ac.uk/Tools/phylogeny/simple\\_phylogeny/](https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/); Madeira et al. 2019), which was visualized using iTOL (Letunic and Bork 2019).

### Protein purification

For the purification of untagged and His-tagged CML1, plasmids pET22-CML1 and pET22-NHisCML1, respectively, were expressed in BL21(DE3) as described above. Cells were harvested by centrifugation, resuspended in cold PBS supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by French Press (SLM Aminco; SLM instruments, Inc. United Kingdom) or broken using a cell disruptor at 1.9 kbar (Constant Systems Ltd) after incubation with 0.5 μL Denarase (c-Lecta) 15 min at room temperature on a rotating wheel. Untagged CML1 was purified on fucosyl-sepharose resin (L-fucose was bound to CNBr-activated sepharose (GE healthcare) according the manufacturer's protocol), eluted with 200 mM fucose in PBS and concentrated on Microsep 3 kDa centrifugal device (PALL). Metal-affinity chromatography of His-tagged CML1 was performed with TALON resin (Clontech) following the manufacturer's instructions, but using PBS for column equilibration, PBS + 5 mM imidazole for washing and PBS + 200 mM imidazole for elution. Purity of the eluted protein was verified by SDS-PAGE and Coomassie-blue staining. The protein was further purified by size exclusion chromatography using a Superdex 75 HR 10/30 column (GE Healthcare) or Enrich650 column (Bio-Rad) during which the buffer was exchanged to TBS (10 mM Tris pH 7.5, 100 mM NaCl). In the absence of gel filtration, the buffer was exchanged to 10 mM Tris pH 7.5, 100 mM NaCl and the protein was concentrated by ultracentrifugation using a Vivaspin 4 concentrator with a 3 kDa cut-off (Sartorius).

### Production of anti-CML1 antiserum and probing of CML1 expression in *C. cinerea*

Purified recombinant His-tagged CML1 was used to immunize 2 rabbits, yielding 2 equally specific polyclonal antisera (Pineda Antikörper Service, Berlin, Germany).

CML1 production in *C. cinerea* was examined by immunoblotting. Cultivation of vegetative mycelium and fruiting bodies of *C. cinerea* and preparation of soluble protein extracts were performed as described (Bleuler-Martinez et al. 2011). Protein extracts were separated on a 12% SDS-PAGE and probed with the specific antisera. Bound

antibodies were detected using horseradish peroxidase (HRP)-conjugated secondary anti-rabbit antibodies.

### Glycan array analysis

Purified recombinant untagged CML1 was fluorescently labeled with Alexa Fluor 488 (Invitrogen) according to the manufacturer's instructions and used to probe the mammalian plate glycan array (V5.1) offered by the Core H of the Consortium for Functional Glycomics (CFG, Emory University, Atlanta). The screening of the array was performed at lectin concentrations of 200, 50, and 5 µg/mL.

### Isothermal titration microcalorimetry (ITC)

ITC experiments were performed on a VP-ITC instrument (Malvern Panalytical, Malvern, UK) at 25°C. Samples of protein and carbohydrate were dialyzed against the same buffer (1× PBS). For a typical titration the sample cell (1.4 mL) was loaded with 140 µM protein and carbohydrate concentration in the syringe was 1.5–4 mM. A titration experiment typically consisted of 60 injections, each of 4 µL volume and 8 s duration, with a 300–400 s interval between additions. Stirring rate was 307 rpm. Raw data were integrated, corrected for non-specific heats, normalized for the molar concentration, and analyzed by nonlinear regression using a single-site binding model with Microcal Origin software. Lacto-N-fucopentaose I and H-type II blood group trisaccharide ( $\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}4\text{GlcNAc}$  (Carbosynth Limited, Compton, UK) were used as ligands for ITC. More details are found in the figure caption.

### Crystallization and structure determination of CML1 apo-protein

For crystallization, the purified recombinant untagged protein was concentrated to 10 mg/mL. Crystallization conditions were screened with the Crystal Screen I (Hampton Research, Aliso Viejo, USA) using the hanging-drop vapor-diffusion technique and mixing 2 µL of the protein solution with 2 µL of the mother liquor. Crystals were grown at 18°C. Crystals of ligand-free CML1 were obtained with 0.2 M ammonium acetate, 0.1 M sodium citrate tribasic dehydrate pH 5.6, 30% polyethylene glycol 4,000. They were directly flash-frozen in liquid nitrogen.

Data were collected at X06DA beamline at the Swiss Light Source (Paul Scherrer Institut, Villigen PSI) and processed with XDS (Kabsch 2010). Phase information was obtained by SAD using Pt derivatized crystals with SHELX package (Sheldrick 2010). The structure was refined with Phenix (Adams et al. 2010) and iterative model rebuilding with Coot (Emsley et al. 2010) to final  $R_{\text{work}}/R_{\text{free}}$  values of 14.41%/17.55% (Supplementary Table SIV). Model statistics were obtained with MolProbity (Chen et al. 2010). Molecular visualizations and structures illustrations were performed using Pymol (Version 2.4.1; Schrödinger Inc., Portland, USA). Coordinates and structure factors have been deposited with the PDB code 6ZRW.

### Crystallization and structure determination of CML1-carbohydrate complexes

Recombinant untagged CML1 in 20 mM Tris-HCl pH 7.5, 100 mM NaCl at 9.8 mg/mL was incubated with 2.5 mM of blood group H-type I trisaccharide or 5 mM of Lewis A tetrasaccharide (Elicity, Crolles, France) at room temperature for at least 1 h prior crystallization. The vapor diffusion

method was used with 2 µL hanging drops of a 1:1 mix of the protein complex and reservoir solutions. Parallelepiped crystals were obtained in a few days for the blood group H type I complex using 1.5 M ammonium sulfate and 100 mM trisodium citrate pH 5.6 as buffer whilst bipyramidal crystals were obtained for Le<sup>a</sup> complex using 1.6 M ammonium sulfate, 100 mM trisodium citrate pH 5.6 and 1% glycerol or 25% polyethylene glycol 4K, 200 mM ammonium acetate, and 100 mM trisodium citrate pH 5.6. Crystals were mounted in a cryoloop (Molecular Dimensions Ltd, Sheffield, UK) after transfer in a cryoprotecting solution containing either 2.5 M Li<sub>2</sub>SO<sub>4</sub> or in a solution where the PEG concentration was increased to 30% prior to flash-freezing in liquid nitrogen. Diffraction data were collected at 100 K at the SOLEIL synchrotron in Saint Aubin, France on the Proxima 1 beamline using an EIGER-X 16M detector (Dectris Ltd.).

The data were processed using XDS and XDSME (Kabsch 2010; Legrand 2019). All further computing was performed using the CCP4 suite (Winn et al. 2011). The structure was solved by molecular replacement using as search model the hexamer coordinates of the apo structure for the H-type I complex or the hexamer of the H-type I complex for Lewis A complex in PHASER 2.8.3 (McCoy et al. 2007). Restrained maximum likelihood refinement using REFMAC 5.8.0257 and local NCS restraints (Murshudov et al. 2011) was iterated with manual rebuilding in Coot (Emsley et al. 2010) to refine the structure. Twin refinement in REFMAC was applied for the Lewis A complex at the end of the refinement. Cross-validation analyses were performed using 5% of the observations set aside. Hydrogen atoms were added in their riding positions and used for geometry and structure factor calculations. Carbohydrate moieties were introduced after inspection of the electron density maps and checked using Privateer (Agirre et al. 2015). Prior deposition in the Protein Data Bank, the models were validated using both the wwPDB Validation server (<http://wwpdb-validation.wwpdb.org>) and Molprobity (Williams et al. 2018). Coordinates and structure factors have been deposited under PDB accession codes 6ZU2 and 6ZU5 for CML1 in complex with H-type I and Lewis A, respectively. Data quality and refinement statistics are described in Supplementary Table SIV.

### Biophysical characterization

#### Multi-angle light scattering (MALS)

The experiment was performed on a miniDawn TriStar system equipped with an Optilab rEX refractometer (Wyatt Technology Corp.) coupled to a Superdex 200 10/30 (Cytiva) size-exclusion chromatography column run on an Agilent 1100 HPLC. 65 µL of a 1 mg/mL protein solution was injected at a constant flow rate of 0.5 mL/min onto the column equilibrated in 20 mM TrisHCl pH 7.5, and 50 mM TrisHCl. The molecular mass of the eluted protein was determined with the Wyatt Astra Version 5.3.4.14 software package (Wyatt Technology Corp.).

### Analytical ultracentrifugation

Sedimentation velocity was measured in a Beckman-Coulter XL-A analytical ultracentrifuge. Sample (370 µL) and buffer (400 µL) solutions were loaded into the double sector centerpiece separately and built up in a Beckman An-50Ti rotor. Experiments were performed at 20°C and rotor speed of 40,000 rpm. The protein sample was monitored by UV absorbance at 280 nm in a continuous mode with a time

interval of 8 min and a step size of 0.003 cm. Multiple scans at different time points were fitted to a continuous size distribution model by the program SEDFIT (Schuck et al. 2002).

#### Small-angle X-ray scattering (SAXS)

SAXS experiments were collected at the cSAXS (X12SA) beamline at the Swiss Light Source (Paul Scherrer Institut, Villigen) at a wavelength of  $\lambda = 1 \text{ \AA}$  on a Pilatus 2 M detector. The scattering vector range was  $0.008\text{--}0.3 \text{ \AA}^{-1}$ , where the length of the scattering vector is defined as  $q = 4\pi \sin \theta/\lambda$ ,  $2\theta$  being the scattering angle. The  $q$  range was calibrated using a standard silver behenate sample (Huang et al. 1993). Protein scattering at 4 concentrations (0.5, 2.5, 5, and 10 mg/mL) was measured in quartz capillaries of 1-mm diameter (Hilgenberg, Malsfeld, Germany), and a buffer control was measured before each protein sample in the same capillary. The measurements were made at several positions along the capillary using 0.5 s exposures. The total exposure time was 50 s. Individual frames were checked for radiation damage (i.e. for changes in intensity at low- $q$  region) and averaged. After initial data integration and background subtraction, data evaluation and processing was performed using the ATSAS program package (Petoukhov et al. 2007). CRYSTAL (Svergun et al. 1995) was used to determine the theoretical scattering curves based on the CML1 crystal structure.

#### Functional analysis of carbohydrate-binding residues by mutagenesis

Plasmids coding for a total of 4 mutant versions of CML1, with mutations in residues involved in carbohydrate binding (H54, N55, W94, and R114), were constructed by site-directed mutagenesis of pET22-CML1 and pET22-NHisCML1 (Supplementary Table SIII). The codons of these residues were mutated to alanine-encoding codons using the QuickChange Site Directed Mutagenesis Kit (Stratagene) and oligonucleotides listed in Supplementary Table SI. Carbohydrate-binding of the various CML1 variants was tested by performing the following binding assay using fucosyl-sepharose beads that were prepared as described above: *E. coli* BL21(DE3) transformed with the various plasmids encoding CML1 wildtype and above mentioned variants thereof were cultivated in 20 mL LB medium containing 100  $\mu\text{g}/\text{mL}$  ampicillin, induced at  $\text{OD}_{600} = 1$  with 1 mM IPTG and grown for 16 h at 23°C. The induced bacterial cells were resuspended in 1 mL of cold PBS containing 1 mM PMSF and lysed using glass beads (diameter 0.1 mm) and a Fastprep cell disruptor. Cell lysates were cleared by centrifugation at  $1,000 \times g$  for 5 min (4°C), and  $1,600 \times g$  for 30 min at 4°C, and 300  $\mu\text{L}$  of the cleared lysates were applied to 150  $\mu\text{L}$  fucosyl-sepharose beads equilibrated with PBS. Binding was performed for 1 h at 4°C on a turning wheel. Fucosyl-sepharose beads were washed with 8 column volumes of PBS and eluted with 100  $\mu\text{L}$  of 200 mM L-fucose in PBS. Samples of input, flow through and eluate were separated by SDS-PAGE and stained with Coomassie Blue.

#### Nematotoxicity assays

Nematotoxicity assays with the 3 different nematode species were performed as described previously with small modifications (Kunzler et al. 2010). For the preparation of nematode

L1 larvae, NGM plates pre-seeded with *E. coli* OP50 were used to grow worms at 20–24°C until most of the bacteria were depleted and a major fraction of the worms represented gravid hermaphrodites. To purify eggs, gravid hermaphrodites were treated with a solution containing 0.5 M NaOH and 1% NaClO for  $\approx 10$  min in 15 mL conical bottom tubes. Eggs were washed twice with 5 mL distilled deionized water (dH<sub>2</sub>O) and resuspended in 100–200  $\mu\text{L}$  dH<sub>2</sub>O before being transferred to a 1.5% agar plate for overnight hatching. Subsequently, L1 larvae were collected in sterile dH<sub>2</sub>O, counted and adjusted to 1,500 worms/mL.

Toxicity assays of CML1 towards *C. tropicalis* and *C. elegans* were performed on NGM plates (Kunzler et al. 2010). For this purpose, *E. coli* BL21 harboring the empty vector pET22b and the derived expression plasmids for wildtype CML1 and the four CML1 mutants H54A, N55A, W94A, and R114A were cultivated in LB medium containing 100  $\mu\text{g}/\text{mL}$  ampicillin at 37°C until an OD of 0.5–1 was reached. The cells were then put on ice for 20 min, followed by induction of protein expression by addition of 0.5 mM IPTG and incubation of the cells at 16°C for 2 h. The OD of all cultures was adjusted to 0.7 and 300  $\mu\text{L}$  of cells were transferred to 10 cm NGM plates with 100  $\mu\text{g}/\text{mL}$  ampicillin and 1 mM IPTG. Five replicates of each condition were prepared and incubated overnight at 23°C. On the next day,  $\sim 30$  L1 larvae of *C. tropicalis* or *C. elegans* were added to each plate and the exact number of nematodes per plate was counted under a stereo microscope. After 72 h at 23°C, all stage L4 nematodes were counted and the percentage of L1 nematodes that reached the L4 stage was determined for each plate individually. The error bars indicate the standard deviation of 5 biological replicates. The statistical analysis was done using Dunnett's multiple-comparison test (Dunnett 1964). Bacterial protein expression was verified by preparing whole protein extracts of bacterial cells scraped from the NGM plates, and supernatants after high speed centrifugation (20 min at 14,000g) thereof, and analyzing these samples using SDS-PAGE with Coomassie Blue staining.

#### Authors' contributions

M.K. conceptualized and designed the study. A.V., V.O., M.S., and G.C. were involved in study design methodological development. S.B., A.V., V.O., M.S. E.V., C.F., T.W., D.P., M.W., and Y.D. were involved in investigation. A.V., V.O., M.S., and M.K. took the responsibility of data curation. S.B., A.V., V.O., and M.S. were involved in writing—original draft preparation. A.V., M.S., and M.K. did writing—review and editing. M.K. and M.A. did supervision. M.K. took the responsibility of funding acquisition. All authors have read and agreed to the published version of the manuscript.

#### Acknowledgements

The authors thank Simon Flückiger and the staff of the PROXIMA-1 beamline (in particular Serena Sirigu) for excellent technical assistance as well as PSI (Villigen, Switzerland) and SOLEIL (St Aubin, France) for the provision of synchrotron radiation facilities.

#### Supplementary material

Supplementary material is available at Glycobiology Journal online.

## Funding

This work was financially supported by ETH Zürich, the Swiss National Science Foundation (grants nos. 31003A\_149512 to MK and CRSII3\_127333 to MS and MA), the NIGMS (grant no. GM62116 to the Consortium for Functional Glycomics) and Glyco@Alps (grant no. ANR-15-IDEX02 to AV).

*Conflict of interest statement:* None declared.

## Data availability statement

The data underlying this article are available in the article, its online supplementary material, public databases (Genbank at <https://www.ncbi.nlm.nih.gov/genbank/> and DOE Joint Genome Institute at <https://mycocosm.jgi.doe.gov> for nucleotide and protein sequences, and Protein Data Bank at <https://www.rcsb.org> for protein structures, using the identifiers indicated in the article) and upon reasonable request to the corresponding author.

## References

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010;66: 213–221.
- Aeschbacher T, Zierke M, Smiesko M, Collot M, Mallet JM, Ernst B, Allain FH, Schubert M. A secondary structural element in a wide range of fucosylated glycoepitopes. *Chem-Eur J.* 2017;23: 11598–11610.
- Agirre J, Iglesias-Fernandez J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. Privateer: software for the conformational validation of carbohydrate structures. *Nat Struct Mol Biol.* 2015;22:833–834.
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A.* 2001;98:10037–10041.
- Becker DJ, Lowe JB. Fucose: biosynthesis and biological function in mammals. *Glycobiology.* 2003;13:41R–53R.
- Bermeo R, Bernardi A, Varrot A. BC2L-C N-terminal lectin domain complexed with Histo blood group oligosaccharides provides new structural information. *Molecules.* 2020;25:248.
- Bianchet MA, Odom EW, Vasta GR, Amzel LM. A novel fucose recognition fold involved in innate immunity. *Nat Struct Biol.* 2002;9: 628–634.
- Bleuler-Martinez S, Butschi A, Garbani M, Walti MA, Wohlschlager T, Potthoff E, Sabotic J, Pohleven J, Luthy P, Hengartner MO, et al. A lectin-mediated resistance of higher fungi against predators and parasites. *Mol Ecol.* 2011;20:3056–3070.
- Bleuler-Martinez S, Stutz K, Sieber R, Collot M, Mallet JM, Hengartner M, Schubert M, Varrot A, Kunzler M. Dimerization of the fungal defense lectin CCL2 is essential for its toxicity against nematodes. *Glycobiology.* 2017;27:486–500.
- Bonnardel F, Mariethoz J, Salentin S, Robin X, Schroeder M, Perez S, Lisacek F, Imbert A. UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. *Nucleic Acids Res.* 2019;47:D1236–D1244.
- Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics.* 1974;77: 71–94.
- Butschi A, Titz A, Walti MA, Olieric V, Paschinger K, Nobauer K, Guo X, Seeberger PH, Wilson IB, Aebi M, et al. *Caenorhabditis elegans* N-glycan core beta-galactoside confers sensitivity towards nematotoxic fungal galectin CGL2. *PLoS Pathog.* 2010;6:e1000717.
- Cabanettes A, Perkams L, Spies C, Unverzagt C, Varrot A. Recognition of complex Core-Fucosylated N-Glycans by a mini lectin. *Angew Chem Int Ed.* 2018;57:10178–10181.
- Cecioni S, Imbert A, Vidal S. Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. *Chem Rev.* 2015;115:525–561.
- Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProB: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010;66:12–21.
- Chen CS, Smits C, Dodson GG, Shevtsov MB, Merlini N, Gollnick P, Antson AA. How to change the oligomeric state of a circular protein assembly: switch from 11-subunit to 12-subunit TRAP suggests a general mechanism. *PLoS One.* 2011;6:e25296.
- Dam TK, Brewer CF. Applications of isothermal titration calorimetry to lectin-carbohydrate interactions. In: Nilsson CL, editors. *Lectins: analytical technologies.* Amsterdam: Elsevier; 2007. pp. 75–101.
- Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 2004;32:W665–W667.
- Dunnett CW. New tables for multiple comparisons with control. *Biometrics.* 1964;20:482–491.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of coot. *Acta Crystallogr Sect D.* 2010;66: 486–501.
- Erjavec J, Kos J, Ravnikar M, Drešo T, Sabotić J. Proteins of higher fungi—from forest to application. *Trends Biotechnol.* 2012;30: 259–273.
- Fukumori F, Takeuchi N, Hagiwara T, Ohbayashi H, Endo T, Kochibe N, Nagata Y, Kobata A. Primary structure of a fucose-specific lectin obtained from a mushroom, Aleuria aurantia. *J Biochem.* 1990;107: 190–196.
- Galeev A, Suwandi A, Cepic A, Basu M, Baines JF, Grassl GA. The role of the blood group-related glycosyltransferases FUT2 and B4GALNT2 in susceptibility to infectious disease. *Int J Med Microbiol.* 2021;311:151487.
- Goldstein IJ, Winter HC. Mushroom lectins. In: Kamerling JP, editors. *Comprehensive Glycoscience: from chemistry to systems biology.* Amsterdam: Elsevier Ltd.; 2007.
- Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 2014;42: D699–D704.
- Hassan MA, Rouf R, Tiralongo E, May TW, Tiralongo J. Mushroom lectins: specificity, structure and bioactivity relevant to human disease. *Int J Mol Sci.* 2015;16:7802–7838.
- Holdener BC, Haltiwanger RS. Protein O-fucosylation: structure and function. *Curr Opin Struct Biol.* 2019;56:78–86.
- Holm L. DALI and the persistence of protein shape. *Protein Sci.* 2020;29:128–140.
- Hopcroft NH, Wendt AL, Gollnick P, Antson AA. Specificity of TRAP-RNA interactions: crystal structures of two complexes with different RNA sequences. *Acta Crystallogr D Biol Crystallogr.* 2002;58: 615–621.
- Houser J, Komarek J, Cioci G, Varrot A, Imbert A, Wimmerova M. Structural insights into Aspergillus fumigatus lectin specificity: AFL binding sites are functionally non-equivalent. *Acta Crystallogr D Biol Crystallogr.* 2015;71:442–453.
- Huang TC, Toraya H, Blanton TN, Wu Y. X-ray-powder diffraction analysis of silver behenate, a possible low-angle diffraction standard. *J Appl Crystallogr.* 1993;26:180–184.
- Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr.* 2010;66:125–132.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10:845–858.
- Khan F, Khan MI. Fungal lectins: current molecular and biochemical perspectives. *Int J Biol Chem.* 2011;5:1–20.
- Kobayashi Y, Kawagishi H. Fungal lectins: a growing family. *Methods Mol Biol.* 2014;1200:15–38.
- Kononova S, Litvinova E, Vakhitov T, Skalinskaya M, Sitkin S. Acceptive immunity: the role of fucosylated Glycans in human host-microbiome interactions. *Int J Mol Sci.* 2021;22:3854.
- Kostlanova N, Mitchell EP, Lortat-Jacob H, Oscarson S, Lahmann M, Gilboa-Garber N, Chambat G, Wimmerova M, Imbert A. The fucose-binding lectin from Ralstonia solanacearum. A new

- type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. *J Biol Chem.* 2005;280:27839–27849.
- Krissinel E, Henrick K. Multiple alignment of protein structures in three dimensions. In: Berthold MR, Glen RC, Diederichs K, Kohlbacher O, Fischer I, editors. *Computational life sciences. CompLife 2005. Lecture notes in computer science.* Berlin, Heidelberg: Springer; 2005
- Kunzler M. Hitting the sweet spot: Glycans as targets of fungal defense effector proteins. *Molecules.* 2015;20:8144–8167.
- Kunzler M, Bleuler-Martinez S, Butschi A, Garbani M, Luthy P, Hengartner MO, Aeby M. Biotoxicity assays for fruiting body lectins and other cytoplasmic proteins. *Methods Enzymol.* 2010;480:141–150.
- Lebreton A, Bonnardel F, Dai YC, Imbert A, Martin FM, Lisacek F. A comprehensive phylogenetic and bioinformatics survey of lectins in the fungal kingdom. *J Fungi (Basel).* 2021;7:453.
- Legrand P. XDSME: XDS made easier. GitHub; 2019. <https://doi.org/10.5281/zenodo.837885>.
- Lepsik M, Sommer R, Kuhaudomlarp S, Lelimousin M, Paci E, Varrot A, Titz A, Imbert A. Induction of rare conformation of oligosaccharide by binding to calcium-dependent bacterial lectin: X-ray crystallography and modelling study. *Eur J Med Chem.* 2019;177: 212–220.
- Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47:W256–W259.
- Li J, Hsu HC, Mountz JD, Allen JG. Unmasking fucosylation: from cell adhesion to immune system regulation and diseases. *Cell Chem Biol.* 2018;25:499–512.
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019;47:W636–W641.
- Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussay F, Gibon J, Kohler A, Lindquist E, Pereda V, et al. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature.* 2008;452:88–92.
- Martinez-Alarcon D, Balloy V, Bouchara JP, Pieters RJ, Varrot A. Biochemical and structural studies of target lectin SapL1 from the emerging opportunistic microfungus *Scedosporium apiospermum*. *Sci Rep.* 2021;11:16109.
- Matsumura K, Higashida K, Ishida H, Hata Y, Kawato A, Abe Y, Kato M, Ueda M, Yamamoto K. Carbohydrate-binding specificity of a fucose-specific lectin from *Aspergillus oryzae*. *Glycobiology.* 2004;14:1092–1092.
- Matsumura K, Higashida K, Ishida H, Hata Y, Yamamoto K, Shigeta M, Mizuno-Horikawa Y, Wang X, Miyoshi E, Gu J, et al. Carbohydrate binding specificity of a fucose-specific lectin from *Aspergillus oryzae*: a novel probe for core fucose. *J Biol Chem.* 2007;282: 15700–15708.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007;40:658–674.
- Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori K, Sakamoto Y, et al. Strand-specific RNA-Seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One.* 2015;10:e0141586.
- Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 2011;67:355–367.
- Notova S, Bonnardel F, Lisacek F, Varrot A, Imbert A. Structure and engineering of tandem repeat lectins. *Curr Opin Struct Biol.* 2020;62:39–47.
- Pan YR, Lou YC, Seven AB, Rizo J, Chen C. NMR structure and calcium-binding properties of the tellurite resistance protein TerD from *Klebsiella pneumoniae*. *J Mol Biol.* 2011;405:1188–1201.
- Paschinger K, Wilson IBH. Comparisons of N-glycans across invertebrate phyla. *Parasitology.* 2019;146:1733–1742.
- Perez S, Sarkar A, Rivet A, Breton C, Imbert A. Glyco3D: a portal for structural glycosciences. *Methods Mol Biol.* 2015;1273:241–258.
- Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI. ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *J Appl Crystallogr.* 2007;40:S223–S228.
- Plaza DF, Lin CW, van der Velden NS, Aebi M, Kunzler M. Comparative transcriptomics of the model mushroom *Coprinopsis cinerea* reveals tissue-specific armories and a conserved circuitry for sexual development. *BMC Genomics.* 2014;15:492.
- Robb CS, Nano FE, Boraston AB. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of intracellular growth locus E (IglE) protein from *Francisella tularensis* subsp. novicida. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2010;66:1596–1598.
- Sabotic J, Ohm RA, Kunzler M. Entomotoxic and nematotoxic lectins and protease inhibitors from fungal fruiting bodies. *Appl Microbiol Biotechnol.* 2016;100:91–111.
- Schubert M. Insights into carbohydrate recognition by 3D structure determination of protein-carbohydrate complexes using NMR. *NMR Glycosci Glycotechnol.* 2017;10:101–122.
- Schubert M, Bleuler-Martinez S, Butschi A, Walti MA, Egloff P, Stutz K, Yan S, Wilson IB, Hengartner MO, Aeby M, et al. Plasticity of the beta-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. *PLoS Pathog.* 2012;8:e1002706.
- Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J.* 2002;82: 1096–1111.
- Sharon N, Lis H. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology.* 2004;14:53R–62R.
- Sheldrick GM. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr.* 2010;66:479–485.
- Singh RS, Walia AK, Kennedy JF. Mushroom lectins in biomedical research and development. *Int J Biol Macromol.* 2020;151: 1340–1350.
- Sulak O, Cioci G, Delia M, Lahmann M, Varrot A, Imbert A, Wimmerova M. A TNF-like trimeric lectin domain from *Burkholderia cenocepacia* with specificity for fucosylated human histo-blood group antigens. *Structure.* 2010;18:59–72.
- Svergun D, Barberato C, Koch MHJ. CRYSTOL-A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr.* 1995;28:768–773.
- Swamy S, Uno I, Ishikawa T. Morphogenetic effects of mutations at the a and B incompatibility factors in *Coprinus-Cinereus*. *J Gen Microbiol.* 1984;130:3219–3224.
- Taylor ME, Drickamer K. Mammalian sugar-binding receptors: known functions and unexplored roles. *FEBS J.* 2019;286:1800–1814.
- Tayyrov A, Schmieder SS, Bleuler-Martinez S, Plaza DF, Kunzler M. Toxicity of potential fungal defense proteins towards the fungivorous nematodes *Aphelenchus avenae* and *Bursaphelenchus okinawaensis*. *Appl Environ Microbiol.* 2018;84:e02051–e02018.
- Tayyrov A, Stanley CE, Azevedo S, Kunzler M. Combining microfluidics and RNA-sequencing to assess the inducible defensome of a mushroom against nematodes. *BMC Genomics.* 2019;20: 243.
- Tayyrov A, Wei C, Fetz C, Goryachkin A, Schächle P, Nyström L, Kunzler M. Cytoplasmic lipases - a novel class of fungal defense proteins against nematodes. *Front Fungal Biol.* 2021;2:696972.
- Thomes L, Bojar D. The role of fucose-containing glycan motifs across taxonomic kingdoms. *Front Mol Biosci.* 2021;8:755577.
- Tielker D, Hacker S, Loris R, Strathmann M, Wingender J, Wilhelm S, Rosenau F, Jaeger KE. *Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiol-Sgm.* 2005;151:1313–1323.
- Tsaneva M, Van Damme EJM. 130 years of plant lectin research. *Glycoconj J.* 2020;37:533–551.

- Turnbull WB, Precious BL, Homans SW. Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry. *J Am Chem Soc.* 2004;126:1047–1054.
- Varrot A, Basheer SM, Imbert A. Fungal lectins: structure, function and potential applications. *Curr Opin Struct Biol.* 2013;23:678–685.
- Walser PJ, Haebel PW, Kunzler M, Sargent D, Kues U, Aebi M, Ban N. Structure and functional analysis of the fungal galectin CGL2. *Structure.* 2004;12:689–702.
- Walti MA, Villalba C, Buser RM, Grunler A, Aebi M, Kunzler M. Targeted gene silencing in the model mushroom *Coprinopsis cinerea* (*Coprinus cinereus*) by expression of homologous hairpin RNAs. *Eukaryot Cell.* 2006;5:732–744.
- Walti MA, Walser PJ, Thore S, Grunler A, Bednar M, Kunzler M, Aebi M. Structural basis for chitotetraose coordination by CGL3, a novel galectin-related protein from *Coprinopsis cinerea*. *J Mol Biol.* 2008;379:146–159.
- Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* 2018;27:293–315.
- Wimmerova M, Mitchell E, Sanchez JF, Gautier C, Imbert A. Crystal structure of fungal lectin: six-bladed beta-propeller fold and novel fucose recognition mode for *Aleuria aurantia* lectin. *J Biol Chem.* 2003;278:27059–27067.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011;67:235–242.
- Wu AM, Lisowska E, Duk M, Yang ZG. Lectins as tools in glycoconjugate research. *Glycoconj J.* 2009;26:899–913.