



**HAL**  
open science

## Structure and engineering of tandem repeat lectins

Simona Notova, François Bonnardel, Frédérique Lisacek, Annabelle Varrot,  
Anne Imberty

► **To cite this version:**

Simona Notova, François Bonnardel, Frédérique Lisacek, Annabelle Varrot, Anne Imberty. Structure and engineering of tandem repeat lectins. *Current Opinion in Structural Biology*, 2020, 62, pp.39-47. 10.1016/j.sbi.2019.11.006 . hal-02463790

**HAL Id: hal-02463790**

**<https://hal.science/hal-02463790v1>**

Submitted on 1 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Structure and engineering of tandem repeat lectins

Simona Notova <sup>1,§</sup>, François Bonnardel <sup>1,2,3,§</sup>, Frédérique Lisacek <sup>2,3,4</sup>,  
Annabelle Varrot <sup>1</sup>, and Anne Imberty <sup>1\*</sup>

1. Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France.
2. Swiss Institute of Bioinformatics, CH-1227 Geneva, Switzerland.
3. Computer Science Department, UniGe, CH-1227 Geneva, Switzerland.
4. Section of Biology, UniGe, CH-1205 Geneva, Switzerland.

§ Both first authors contributed equally to this work

\* To whom correspondence should be addressed. Anne Imberty (anne.imberty@cermav.cnrs.fr, Tel: +33 476 03 76 40, Twitter: @AnneImberty) Frédérique Lisacek (frederique.lisacek@sib.swiss, Tel: +4122 379 01 95)

## Abstract (100 – 120 words)

Through their ability to bind complex glycoconjugates, lectins have unique specificity and potential for biomedical and biotechnological applications. In particular, lectins with short repeated peptides forming carbohydrate-binding domains are not only of high interest for understanding protein evolution but can also be used as scaffold for engineering novel receptors. Synthetic glycobiology now provides the tools for engineering the specificity of lectins as well as their structure, multivalency and topologies. This review focuses on the structure and diversity of two families of tandem-repeat lectins, i.e.  $\beta$ -trefoils and  $\beta$ -propellers, demonstrated as the most promising scaffold for engineering novel lectins.

## Keywords

lectins, carbohydrate-binding protein, database,  $\beta$ -trefoil,  $\beta$ -propeller

## 1 Introduction

Lectins are ubiquitous protein receptors able to bind complex oligosaccharides (Lis & Sharon, 2002). They display a large panel of biological functions and many lectins are useful in biotechnology and biomedicine applications. Lectins are generally multivalent, that is, they contain several binding sites for the simultaneous attachment to several target carbohydrates. While multivalency is usually associated with strong avidity for glycoconjugates, the topology of the binding sites in space is also of importance for generating high specificity, for example, towards the branches of N-glycans and for biological function. Recently, it was demonstrated that specific arrangements of lectin binding sites favour glycolipid clusters and therefore affect the structure and dynamics of the cell membranes [1].

The multivalent properties of proteins relate to their highly symmetrical shapes. Most of the time, the symmetry of a protein structure stems from the oligomerisation of identical protomers. In

oligomeric lectins, the carbohydrate-binding sites are generally spread in different directions resulting in structures with dihedral (D2 to D8) or tetrahedral symmetry (Fig. 1). In contrast, the cyclic symmetry  $C_n$  generates structures where all binding sites are on the same side, therefore perfectly fitted to bind glycosylated surfaces, such as cell membranes. Few oligomeric cyclic lectins are observed and they adopt  $C_2$  to  $C_7$  symmetry (Fig. 1). Other lectins generate multivalency yet through a different strategy, using tandem repeats of a single motif to create cyclic and dihedral symmetries. In such case, the evolutionary origin of structure symmetry has been demonstrated to occur through gene duplication and fusion events [2].

The Lectin3D database [3] that gathers all known 3D-structures of lectins has been recently updated and remodeled as Unilectin3D [4]\*\*. It contains more than 2000 structures in over 100 lectin families (i.e. unrelated sequences). While most are oligomeric, several different lectin families contain tandem repeats. A two-fold repeat with  $C_2$  symmetry is observed in some members of the  $\beta$ -prism I (e.g. banana lectin, griffithsin), cyanovirin and few other families. Tandem repeats with  $C_3$  cyclic symmetry are observed in the  $\beta$ -prism II (e.g., piocyn, monocot lectins) and  $\beta$ -trefoil families and is expected to occur in many lectins with the novel  $\beta$ -prism III fold [5]. Higher cyclic symmetry ( $C_n$ ) is the hallmark of  $\beta$ -propellers lectins. Combination of tandem repeat and oligomerization result in even higher valency (Fig. 1).

Tandem repeat lectins and their symmetry are of high interest for understanding protein evolution but also for the design of “neo-lectins” with engineered structures and binding sites. In the emerging domain of synthetic glycobiology [6], engineering lectins provide powerful tools for deciphering the glycode, but can also be applied in biotechnology for quality control of recombinant therapeutic glycoproteins, or in clinics as biomarkers for histopathology. Engineering strategies have been developed with some success [7] [8]\*\*. In this review, we focus on the two families of tandem repeat lectins that have been demonstrated to be the most promising scaffold for engineering novel lectins, i.e.  $\beta$ -propellers and  $\beta$ -trefoils.

## 2 $\beta$ -Propeller lectins

### 2.1 $\beta$ -propellers lectins: large panel of structure

The  $\beta$ -propeller fold is widely spread and results in proteins with four to ten repeats of four-stranded antiparallel  $\beta$ -sheets (i.e., blades of the propeller) assembled in a donut shape [9-11]. Members of the  $\beta$ -propeller family have diverse functions, ranging from signalling to enzymatic activity, and they are characterized by their structural rigidity. The known  $\beta$ -propeller lectins have been recently collected, compared and grouped into five classes based on the HMM signature of their blades [12]\*\*. They are characterized by the positioning of five to seven carbohydrate binding sites on the same side of the donut, resulting in very efficient binding to glycoconjugates (Fig. 2).  $\beta$ -propeller lectins are present in some bacterial and fungal pathogens and involved in infection by targeting glycans on host tissue, such as AFL/FleA in *Aspergillus fumigatus* that binds fucose and triggers a response from the host [13,14]. They are also involved in the innate immunity of several mushrooms and invertebrates and act by recognizing PAMPs (*pathogen-associated molecular patterns*) such as N-acetylated or methylated carbohydrates [15]\*\*. Variations also occur at the supramolecular level: the donut-shaped  $\beta$ -propeller frequently associates back-to-back in solution, therefore presenting the carbohydrate-binding sites on opposite faces. Tetramer assemblies of such toruses can result in box-like shapes as in the lectin of *Photobacterium luminescens* (PLL) [16] or virus-like shapes as in Tectonin 2 from *Laccaria bicolor* [15]\*\*.

The PropLec6A class of  $\beta$ -propellers contains fungal and bacterial proteins and is rather unique with respect to the blade assembly that forms the propeller. While all fungal structures display the classical 6-repeat architecture, the bacterial ones are much shorter and their sequences contain only two or three blades. The  $\beta$ -propeller of lectins of *Ralstonia solanacearum* (RSL) and *Burkholderia ambifaria* (BambL) result from the trimerization of short 90 amino acid-long peptides comprising two repeats [17,18] whereas the one found in *Kordia zhangzhouensis* (KozL) assembles as a dimer of longer peptides comprising three repeats [12]\*\*. Most recent studies have concluded that all  $\beta$ -propellers have a common monophyletic origin, in spite of the different families stemming from

independent amplification and diversification pathways from an ancestral peptide [19]. The  $\beta$ -prism-II proteins, which include several lectin classes, have been proposed to also share this ancestral peptide [20]. Interestingly, the only oligomeric  $\beta$ -propellers, i.e. possible fossil forms, identified so far are the microbial PropLec6A lectins (RSL, BambL and KozL). It has been suggested that they represent less stable evolutionary intermediates that have been replaced by fully amplified peptides during evolution [19].

$\beta$ -Propeller lectins are robust proteins, easy to produce in bacterial fermentation and to handle in the lab, which explains their wide use in biotechnology. Fucose-specific lectins from *Aleuria aurantia* (AAL) and *Aspergillus oryzae* (AAO) are available from companies and included in lectin arrays [21]. The specificity for terminal GlcNAc of the *Psathyrella velutina* lectin (PVL) can be used for labeling some cancer cells and tissues [22] and both PVL and its homolog in *Agrocybe aegerita* (AANL) are very interesting tools for the glycoproteomics of O-GlcNAcylation [23] [24]\*.

## 2.2 $\beta$ -Propellers: scaffolds for engineering symmetry and multivalency

The  $\beta$ -propeller fold has been an excellent scaffold for understanding and reproducing Evolution and several groups demonstrated the possibility to reconstruct symmetrical proteins. The Pizza protein was computationally designed and produced as a perfectly symmetrical 6-fold propeller [25], followed by the Tako and Ika peptides that self-assemble as symmetrical 8-fold propellers [26]. Tachylectin-2, the only propeller lectin with five blades, has been also used due to the remarkable conservation of amino acids among the blades. Genetic library screening by phage display resulted in the identification of 2-blade fragments that could assemble as new lectins with sugar-binding activity [27]. A novel approach was proposed through the alignment of Tachylectin-2 and *Nematostella vectensis* lectin sequences and the reconstruction of an ancestor blade, a 43 amino acid sequence that self-assembles in a 5-blade lectin able to efficiently bind glycoconjugates (Fig. 3) [28]\*\*.

Directed mutagenesis has been used for lowering the number of binding sites in  $\beta$ -propeller lectins, particularly in the LecB6 class. An AAL variant produced by site-directed mutagenesis to present

only one functional fucose-binding site displayed a decrease in agglutination activity [29]. Furthermore,  $\beta$ -propeller lectins are powerful tools for studying membrane dynamics as they can cluster glycan heads and invaginate the membrane. Decreasing the valency of the RSL lectin from six to three binding sites resulted in the loss of capacity to invaginate the cell membrane of giant vesicles decorated with Lewis<sup>a</sup>-pseudolipids [30]. The construction of a library of neolectins with systematic variations of the number and localization of binding sites demonstrated the crucial importance of distances between the binding sites, i.e. the effect of the lectin topology on the dynamics of targeted glycolipids [1].

The LecB6 family displays strong affinity for fucose, partly due to the interaction between aromatic amino acids (Trp or Tyr) stacking with the hydrophobic face of the monosaccharides. In the AAL fungal lectin, only five of the six binding sites bind fucose, and different occupancy of oligosaccharides indicate differences among the five sites. Alanine-mutation mapping on Trp and Tyr residues established that sites 1, 2 and 4 are more efficient for binding [31]. Synthetic biology was used for introducing fluorinated indol analogs of the tryptophan residues at different positions in RSL. Introduction of non-canonical amino acids resulted in lectin variants with altered specificity towards fucosylated blood group oligosaccharides [32]\*\*.

Mutations of other amino acids in the binding sites were also explored and AAL mutants displaying altered binding affinities to different fucosylated oligosaccharides were obtained [33]. The Q224N mutant has higher affinity for core fucosylated N-glycan and is of special interest as biomarker for hepatocellular carcinoma [34]\*. Chemical functionalization of RSL generated conjugates with 3, 6, or 12 polyethyleneglycol molecules (PEG) that displayed altered carbohydrate-binding properties [35]. The RSL peptide has also been used for the development of novel thermo-responsive affinity sugar binders by fusion with elastin-like polypeptides [36]. The fusion of the RSL peptide with a sialic acid binding domain of another bacterium resulted in the first Janus-lectin, able to bind six fucoses on one face and three sialic acids on the other face. Such supramolecular assembly was demonstrated to form an organized layer with bivalent glyco-compounds and to associate glyco-vesicles in proto-tissues [37]\*\*.

### 3 $\beta$ -Trefoil lectins

$\beta$ -Trefoil proteins are composed of a three-lobed architecture consisting of peptide repeats forming a barrel structure at one extremity and a triangular arrangement of hairpins at the other [38].  $\beta$ -Trefoil lectins are largely spread in bacteria, fungi, plants and animals and often attached to other proteins with toxic or enzymatic activity. They are classified as family 13 of the carbohydrate-binding modules when attached to carbohydrate active enzymes [39]. Whilst very conserved in shape, the sequence conservation is very weak and only a few hydrophobic residues are conserved in the core. This involves the QxW motif that is repeated in each foil in most classes of  $\beta$ -trefoil lectins [40], albeit with some exceptions. The 3-fold symmetry can be degenerated and many  $\beta$ -trefoil lectins present only one or two carbohydrate active binding sites (Fig. 4). While originally described as specific for galactose,  $\beta$ -trefoil lectins have been evidenced to bind a variety of sugars, including rhamnose in the structure of a recently solved protozoan lectin [41].

#### 3.1 $\beta$ -Trefoil lectin: structural plasticity for multiple function

The function of  $\beta$ -trefoil lectins can vary depending on the organism and ligands but many of them serve as a recognition module for an attached toxin domain, as for ricin-B, the first structure to be determined [42]. The  $\beta$ -trefoil domains often play a role in innate immunity as demonstrated by CCL2 from the *Coprinopsis cinerea* fungus which is toxic towards nematodes. The dimeric lectin acts by binding a fucosylated epitope present in the worm digestive system [43]\*\*.  $\beta$ -Trefoil lectins are also useful in biotechnology. Some have fine specificity that can be used for glycan sorting and labeling, such as the *Marasmius oreades* mushroom MOA lectin recognizing specifically  $\alpha$ -galactose in blood group B epitope [44]. Actinohivin in Actinomycetes is a broadly-neutralizing lectin against HIV-1 infection that targets the gp120 envelope glycoprotein [45]. Indeed, the crystal structure of the lectin confirms the presence of oligomannose in the three bindings sites of the peptide [46].

Lectins from marine invertebrates are of special interest due to their role in the protection against pathogens. Mytilec from the mussel *Mytilus galloprovincialis* has antibacterial and antifungal

functions. Both Mytilec and the related *Crenomytilus grayanus* CGL lectin bind with high avidity the  $\alpha$ Gal1-4Gal epitope present on globotriaosyl ceramide in cancer cells. Both lectins have anti-cancer properties since this binding downregulates cell growth and leads to apoptosis [47]\* [48]\*.

#### 3.2 Engineering $\beta$ -trefoil structures and specificity

The plasticity of  $\beta$ -trefoil lectins makes them a convenient tool to study evolution. The construction of artificial lectins, mimicking their evolutionary pathways, was attempted for both plant and invertebrate proteins. A consensus sequence was designed based on the lactose-binding ricin B. The sequence was repeated to obtain the three-fold globular protein named ThreeFoil [49]. The resulting structure is almost perfectly symmetrical and exhibits high thermal stability. Interestingly, the carbohydrate ligand, i.e. lactose, acts as a chaperone and helps in the folding of the artificial lectin [50]\*.

MytiLec is also an excellent candidate for engineering artificial lectins thanks to its high internal similarity and the presence of one active sugar-binding site in each lobe of the trefoil structure. In its native form, Mytilec consists of a tight homodimer and its specificity towards cancer antigen is of interest in biotechnology. The substitution of apolar residues by polar ones at the dimerization interface, resulted in a monomeric mutant, albeit with low stability [48]\*. The loss of cytotoxic activity demonstrated that dimerization is crucial for cellular impact, as confirmed for the toxicity of CGL2 [43]\*\*. The ancestral reconstruction method was used to create a three-fold perfectly symmetrical protein derived from Mytilec and called Mitsuba (three-leaf in Japanese) [51]\*\*. Mitsuba interacts with cells which express globotriose on their surfaces and shows good stability but only weak anti-cancer activity.

Yabe *et al.* demonstrated the process of natural evolution mimicry by changing the specificity of earthworm EW29Ch lectin from galactose to sialic acid. The lectin was engineered by error-prone PCR with a reinforced ribosome display system and the original lactose-binding pocket was extended to adapt sialylation on the galactose ligand. With this approach, the authors created the first sialic acid specific R-type lectin [52]. In 2012, this method was further developed by Hu *et al.*

[53], who improved the process by introducing a high-throughput screening system with glycoconjugate microarrays. The engineered mutant is the first receptor specific for 6-sulfo-galactose terminated glycans, and is a potential marker for various disease. Further analysis showed that the change of specificity was reached just by a single E20K point mutation.

## 4 . Conclusion and discussion

Tandem-repeat lectins originate from an evolution strategy based on gene duplication resulting in the multiplication of binding sites with ideal topology for binding to substrates. The identification of repeated short peptides with carbohydrate-binding activities is difficult and it is likely that novel lectin families are yet to be identified in Nature. For the known folds, such as  $\beta$ -propellers, appropriate bioinformatics tools turned out essential for efficient search in genomic databases [12]\*\*.

Tandem-repeat lectins such as  $\beta$ -propellers and  $\beta$ -trefoils are excellent tools as bio-markers due to their strong avidity for glycoconjugates presented on cell membranes. Engineering strategies, from single point mutation to incorporation of non-canonical amino acids, or directed evolution, are now available for designing and controlling the specificity and the topology of such lectins. Creating such multivalent receptors through artificially constructed tandem repeats is a route of interest. It generally strongly enhances the receptor affinity for its ligand [54,55]. It has also been demonstrated that expression as tandem-repeat sequences improves the production of recombinant algal lectins in bacterial systems [56]. This is a definite option for the future development of artificial lectins, molecules that will have many applications in health science and biotechnology.

## Acknowledgments

The authors acknowledge support by the ANR PIA Glyco@Alps (ANR-15-IDEX-02), the Alliance Campus Rhodanien Co-funds (<http://campusrhodanien.unige-cofunds.ch>) and the H2020 Marie Skłodowska-Curie Action n° 814029 (synBIOcarb).

## Legends for figures

**Figure 1.** Schematic representation of different types of symmetry observed in A/ oligomeric lectins, B/ tandem-repeat lectins, and C/ hybrid types. Lectin domain with identical peptide sequences are represented by identical color.

**Figure 2.** Representation of selected  $\beta$ -propeller lectins. **A)** Schematic representation of the different folds and assemblies observed in lectins. **B)** Structure of representative lectins in each family with carbohydrate ligands. **C)** Different types of functional assemblies of the  $\beta$ -propeller lectins

**Figure 3.** Engineering of propellers. Schematic representation of deconstruction/construction of neo-lectin (left) and janus lectin (right) from the propeller architecture

**Figure 4.** Representation of selected  $\beta$ -trefoil lectins. **A)** Variations of structure and specificity observed in  $\beta$ -trefoil lectins for bacteria (botulin toxin and actinohivin), fungi (CCL2 and MOA), plants (amaranthus agglutinin and ricin) and animals (mussel CGL and sea cucumber CEL-III). Ligands are represented using the Symbol Nomenclature for Glycans [57]. **B)** Sugar binding sites of the same lectins as in section A with hydrogen bonds as blue lines, water molecule and calcium ions as red and green balls, respectively. **C)** Schematic representation of different architecture associated with  $\beta$ -trefoil lectins, including monomer, covalent or non-covalent dimers, and multimer (hemolytic lectin CEL-III).

## References

1. Arnaud J, Tröndle K, Claudinon J, Audfray A, Varrot A, Römer W, Imberty A: **Membrane deformation by neolectins with engineered glycolipid binding sites.** *Angew. Chem., Int. Ed. Engl.* 2014, **53**:9267-9270.
2. Blaber M, Lee J, Longo L: **Emergence of symmetric protein architecture from a simple peptide motif: evolutionary models.** *Cell. Mol. Life Sci.* 2012, **69**:3999-4006.
3. Pérez S, Rivet A, Imberty A: **3D-Lectin Database.** In *Glycoscience: Biology and Medicine.* Edited by Taniguchi N, Hart GW, Seeberger P, Wong CH: Springer; 2015:283-289.

4. \*\* Bonnardel F, Mariethoz J, Salentin S, Robin X, Schroeder M, Pérez S, Lisacek F, Imberty A: **UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands.** *Nucleic Acids Res.* 2019, **47**:D1236–D1244.
- Description of the database of 2000 structures of lectins classified on origin and fold, with cross-links to literature, other databases in glycosciences and functional data such as known specificity.*
5. 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. Engl.* 2018, **57**:10178-10181.
6. Turnbull WB, Imberty A, Blixt O: **Synthetic glycobiology.** *Interface Focus* 2019, **9**:20190004.
7. Hu D, Tateno H, Hirabayashi J: **Lectin engineering, a molecular evolutionary approach to expanding the lectin utilities.** *Molecules* 2015, **20**:7637-7656.
8. \*\* Hirabayashi J, Arai R: **Lectin engineering: The possible and the actual.** *Interface Focus* 2019, **9**:20180068.
- Up-to-date review on strategies for engineering lectin specificities. It also discusses possible approaches to confer sugar-binding properties on synthetic proteins and peptides.*
9. Chen CK, Chan NL, Wang AH: **The many blades of the beta-propeller proteins: conserved but versatile.** *Trends Biochem. Sci.* 2011, **36**:553-561.
10. Fülöp V, Jones DT:  **$\beta$  Propellers: Structural rigidity and functional diversity.** *Curr. Opin. Struct. Biol.* 1999, **9**:715-721.
11. Jawad Z, Paoli M: **Novel sequences propel familiar folds.** *Structure* 2002, **10**:447-454.
12. \*\* Bonnardel F, Kumar A, Wimmerova M, Lahmann M, Perez S, Varrot A, Lisacek F, Imberty A: **Architecture and evolution of blade assembly in  $\beta$ -propeller lectins.** *Structure* 2019, **27**:764–775.
- Bioinformatics approach for identification of  $\beta$ -propeller lectins in genomes, that reveals their widespread distribution of across species.*
13. Houser J, Komarek J, Kostlanova N, Cioci G, Varrot A, Kerr SC, Lahmann M, Balloy V, Fahy JV, Chignard M, et al.: **A soluble fucose-specific lectin from *Aspergillus fumigatus* conidia - Structure, specificity and possible role in fungal pathogenicity.** *PLoS ONE* 2013, **8**:e83077.
14. Richard N, Marti L, Varrot A, Guillot L, Guitard J, Hennequin C, Imberty A, Corvol H, Chignard M, Balloy V: **Human bronchial epithelial cells inhibit *Aspergillus fumigatus* germination of extracellular conidia via FleA recognition.** *Sci. Rep.* 2018, **8**:15699
15. \*\* Sommer R, Makshakova ON, Wohlschlagler T, Hutin S, Marsh M, Titz A, Kunzler M, Varrot A: **Crystal structures of fungal tectonin in complex with O-methylated glycans suggest key role in innate immune defense.** *Structure* 2018, **26**:391-402.
- An unusual capsid-shape assembly of lectins with direct link to the function as defense against pathogens and predator*
16. Kumar A, Sykorova P, Demo G, Dobes P, Hyrs I P, Wimmerova M: **A novel fucose-binding lectin from *Photorhabdus luminescens* (PLL) with an unusual heptabladed beta-propeller tetrameric structure.** *J. Biol. Chem.* 2016, **291**:25032-25049.
17. Audfray A, Claudinon J, Abounit S, Ruvoën-Clouet N, Larson G, Smith DF, Wimmerová M, Le Pendu J, Römer W, Varrot A, et al.: **Fucose-binding lectin from opportunistic pathogen *Burkholderia ambifaria* binds to both plant and human oligosaccharidic epitopes.** *J. Biol. Chem.* 2012, **287**:4335-4347.
18. Kostlanová N, Mitchell EP, Lortat-Jacob H, Oscarson S, Lahmann M, Gilboa-Garber N, Chambat G, Wimmerová M, Imberty A: **The fucose-binding lectin from *Ralstonia solanacearum*: a new type of -propeller architecture formed by oligomerisation and interacting with fucoside, fucosyllactose and plant xyloglucan.** *J. Biol. Chem.* 2005, **280**:27839-27849.
19. Chaudhuri I, Soding J, Lupas AN: **Evolution of the beta-propeller fold.** *Proteins* 2008, **71**:795-803.
20. Kopec KO, Lupas AN: **beta-Propeller blades as ancestral peptides in protein evolution.** *PLoS One* 2013, **8**:e77074.
21. Hirabayashi J, Yamada M, Kuno A, Tateno H:

- Lectin microarrays: concept, principle and applications.** *Chem. Soc. Rev.* 2013, **42**:4443-4458.
22. Audfray A, Beldjoudi M, Breiman A, Hurbin A, Boos I, Unverzagt C, Bouras M, Lantuejoul S, Coll JL, Varrot A, et al.: **A recombinant fungal lectin for labeling truncated glycans on human cancer cells.** *PLoS ONE* 2015, **10**.
  23. Machon O, Baldini SF, Ribeiro JP, Steenackers A, Varrot A, Lefebvre T, Imberty A: **Recombinant fungal lectin as a new tool to investigate O-GlcNAcylation processes.** *Glycobiology* 2017.
  24. \* Liu W, Han G, Yin Y, Jiang S, Yu G, Yang Q, Yu W, Ye X, Su Y, Yang Y, et al.: **AANL (*Agrocybe aegerita* lectin 2) is a new facile tool to probe for O-GlcNAcylation.** *Glycobiology* 2018.
- The AANL lectin is proven as a useful tool for enrichment and identification of O-GlcNAcylated proteins and peptides; a fine example of the use of lectin in glycoproteomics*
25. Voet AR, Noguchi H, Addy C, Simoncini D, Terada D, Unzai S, Park SY, Zhang KY, Tame JR: **Computational design of a self-assembling symmetrical beta-propeller protein.** *Proc. Natl. Acad. Sci. U. S. A.* 2014, **111**:15102-15107.
  26. Noguchi H, Addy C, Simoncini D, Wouters S, Mylemans B, Van Meervelt L, Schiex T, Zhang KYJ, Tame JRH, Voet ARD: **Computational design of symmetrical eight-bladed  $\beta$ -propeller proteins.** *IUCrJ* 2019, **6**:46-55.
  27. Yadid I, Tawfik DS: **Functional beta-propeller lectins by tandem duplications of repetitive units.** *Protein Eng. Des. Sel.* 2011, **24**:185-195.
  28. \*\* Smock RG, Yadid I, Dym O, Clarke J, Tawfik DS: **De Novo evolutionary emergence of a symmetrical protein is shaped by folding constraints.** *Cell* 2016, **164**:476-486.
- First example of reconstruction of artificial functional lectin from 47-residue ancestral motifs that form five-bladed propellers via oligomeric assembly*
29. Olausson J, Astrom E, Jonsson BH, Tibell LA, Pahlsson P: **Production and characterization of a monomeric form and a single-site form of *Aleuria aurantia* lectin.** *Glycobiology* 2011, **21**:34-44.
  30. Arnaud J, Claudinon J, Tröndle K, Trovaslet M, Larson G, Thomas A, Varrot A, Römer W, Imberty A, Audfray A: **Reduction of lectin valency drastically changes glycolipid dynamics in membranes but not surface avidity.** *ACS Chemical Biology* 2013, **8**:1918-1924.
  31. Amano K, Fujihashi M, Ando A, Miki K, Nagata Y: **Involvement of tyrosines at fucose-binding sites of *Aleuria aurantia* lectin: non-equal response to site-directed mutagenesis among five sites.** *Biosci., Biotechnol., Biochem.* 2004, **68**:841-847.
  32. \*\* Tobola F, Lelimosin M, Varrot A, Gillon E, Darnhofer B, Blixt O, Birner-Gruenberger R, Imberty A, Wiltschi B: **Effect of non-canonical amino acids on protein-carbohydrate interactions: Structure, dynamics and carbohydrate affinity of a lectin engineered with fluorinated tryptophan analogs.** *ACS Chemical Biology* 2018, **13**:2211-2219.
- First inclusion of fluoro-tryptophan analogs, i.e. non-canonical amino acids, in a lectin binding site, resulting in change of specificity towards histo-blood group oligosaccharides*
33. Romano PR, Mackay A, Vong M, DeSa J, Lamontagne A, Comunale MA, Hafner J, Block T, Lec R, Mehta A: **Development of recombinant *Aleuria aurantia* lectins with altered binding specificities to fucosylated glycans.** *Biochem. Biophys. Res. Commun.* 2011, **414**:84-89.
  34. \* Norton P, Comunale MA, Herrera H, Wang M, Houser J, Wimmerova M, Romano PR, Mehta A: **Development and application of a novel recombinant *Aleuria aurantia* lectin with enhanced core fucose binding for identification of glycoprotein biomarkers of hepatocellular carcinoma.** *Proteomics* 2016, **16**:3126-3136.
- Development of a candidate oncofetal marker for liver cancer by engineering a fungal lectin and modifying specifically its binding affinity to fucose linkages*
35. Ramberg KO, Antonik PM, Cheung DL, Crowley PB: **Measuring the impact of PEGylation on a protein-polysaccharide interaction.** *Bioconjug. Chem.* 2019, **30**:1162-1168.



36. Arnold L, Chen R: **Novel thermo-responsive fucose binding ligands for glycoprotein purification by affinity precipitation.** *Biotechnol. Bioeng.* 2014, **111**:413-417.
37. \*\* Ribeiro JP, Villringer S, Goyard D, Coche-Guerente L, Höferlin M, Renaudet O, Römer W, Imberty A: **Tailor-made Janus lectin with dual avidity assembles glycoconjugate multilayers and crosslinks protocells.** *Chem. Sci.* 2018, **9**:7634 - 7641
- Design of the first chimeric bispecific lectin, with two rationally oriented and distinct recognition surfaces that bind independently to both fucosylated and sialylated glycoconjugates.*
38. Murzin AG, Lesk AM, Chothia C: **beta-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors.** *J. Mol. Biol.* 1992, **223**:531-543.
39. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ: **Carbohydrate-binding modules: fine-tuning polysaccharide recognition.** *Biochem. J* 2004, **382**:769-781.
40. Hazes B: **The (QxW)<sub>3</sub> domain: a flexible lectin scaffold.** *Protein Sci.* 1996, **5**:1490-1501.
41. Khan F, Suguna K: **Crystal structures of a  $\beta$ -trefoil lectin from *Entamoeba histolytica* in monomeric and a novel disulphide bond-mediated dimeric forms.** *Glycobiology* in press.
42. Rutenber E, Robertus JD: **Structure of ricin B-chain at 2.5 Å resolution.** *Proteins* 1991, **10**:260-269.
43. \*\* 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.
- Illustration of the protective function of lectin in fungi and of the requirement of multivalency for maintaining its toxicity towards invertebrates.*
44. Grahn E, Askarieh G, Holmner A, Tateno H, Winter HC, Goldstein IJ, Krenzel U: **Crystal structure of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope.** *J. Mol. Biol.* 2007, **369**:710-721.
45. Hoorelbeke B, Huskens D, Ferir G, Francois KO, Takahashi A, Van Laethem K, Schols D, Tanaka H, Balzarini J: **Actinohivin, a broadly neutralizing prokaryotic lectin, inhibits HIV-1 infection by specifically targeting high-mannose-type glycans on the gp120 envelope.** *Antimicrob. Agents Chemother.* 2010, **54**:3287-3301.
46. Zhang F, Hoque MM, Jiang J, Suzuki K, Tsunoda M, Takeda Y, Ito Y, Kawai G, Tanaka H, Takenaka A: **The characteristic structure of anti-HIV actinohivin in complex with three HMTG D1 chains of HIV-gp120.** *Chembiochem* 2014, **15**:2766-2773.
47. \* Liao JH, Chien CT, Wu HY, Huang KF, Wang I, Ho MR, Tu IF, Lee IM, Li W, Shih YL, et al.: **A multivalent marine lectin from *Crenomytilus grayanus* possesses anti-cancer activity through recognizing globotriose Gb3.** *J. Am. Chem. Soc.* 2016, **138**:4787-4795.
- A promising example for development of lectin-based biosensor based on the specificity to globotriose (Gb3). Gb3 represents marker for various types of cancer and in this work, CGL reduced viability of breast cancer cell line*
48. \* Terada D, Kawai F, Noguchi H, Unzai S, Hasan I, Fujii Y, Park SY, Ozeki Y, Tame JRH: **Crystal structure of MytiLec, a galactose-binding lectin from the mussel *Mytilus galloprovincialis* with cytotoxicity against certain cancer cell types.** *Sci. Rep.* 2016, **6**:28344.
- The prototype of mussel lectin, showing dose dependent cytotoxic activity to certain cancer cell lines, whereas the engineered monomer loses haemagglutination and cytotoxic activity*
49. Broom A, Doxey AC, Lobsanov YD, Berthin LG, Rose DR, Howell PL, McConkey BJ, Meiering EM: **Modular evolution and the origins of symmetry: reconstruction of a three-fold symmetric globular protein.** *Structure* 2012, **20**:161-171.
50. \* Broom A, Ma SM, Xia K, Rafalia H, Trainor K, Colon W, Gosavi S, Meiering EM: **Designed protein reveals structural determinants of extreme kinetic stability.** *Proc. Natl. Acad. Sci. U. S. A.* 2015, **112**:14605-14610.
- ThreeFoil, as an example of how rational design can improve solubility and folding kinetics. Moreover, lactose, as a ligand, served as*

*chaperone in the process of ThreeFoil folding which supported the advantage of designing the fold and function together*

51. \*\* Terada D, Voet ARD, Noguchi H, Kamata K, Ohki M, Addy C, Fujii Y, Yamamoto D, Ozeki Y, Tame JRH, et al.: **Computational design of a symmetrical  $\beta$ -trefoil lectin with cancer cell binding activity.** *Sci. Rep.* 2017, 7:5943.

*Design of Mitsuba (three leaf in Japanese), an artificial protein engineered from MytiLec. This  $\beta$ -trefoil lectin is composed of 3 identical tandem repeats and retains sugar binding activity*

52. Yabe R, Suzuki R, Kuno A, Fujimoto Z, Jigami Y, Hirabayashi J: **Tailoring a novel sialic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry.** *J. Biochem.* 2007.
53. Hu D, Tatenno H, Kuno A, Yabe R, Hirabayashi J: **Directed evolution of lectins with sugar-binding specificity for 6-sulfo-galactose.** *J. Biol. Chem.* 2012.

54. Connaris H, Crocker PR, Taylor GL: **Enhancing the receptor affinity of the sialic acid-binding domain of *Vibrio cholerae* sialidase through multivalency.** *J. Biol. Chem.* 2009, 284:7339-7351.
55. Ribeiro J, Pau W, Pifferi C, Renaudet O, Varrot A, Mahal LK, Imberty A: **Characterization of a high-affinity sialic acid specific CBM40 from *Clostridium perfringens* and engineering of divalent form.** *Biochem. J* 2016, 473:2109-2118.
56. Hwang HJ, Han JW, Jeon H, Han JW: **Induction of recombinant lectin expression by an artificially constructed tandem repeat structure: A case study using *Bryopsis plumosa* mannose-binding lectin.** *Biomolecules* 2018, 8.
57. Neelamegham S, Aoki-Kinoshita K, Bolton E, Frank M, Lisacek F, Lutteke T, O'Boyle N, Packer NH, Stanley P, Toukach P, et al.: **Updates to the Symbol Nomenclature for Glycans guidelines.** *Glycobiology* 2019, 29:620-624.

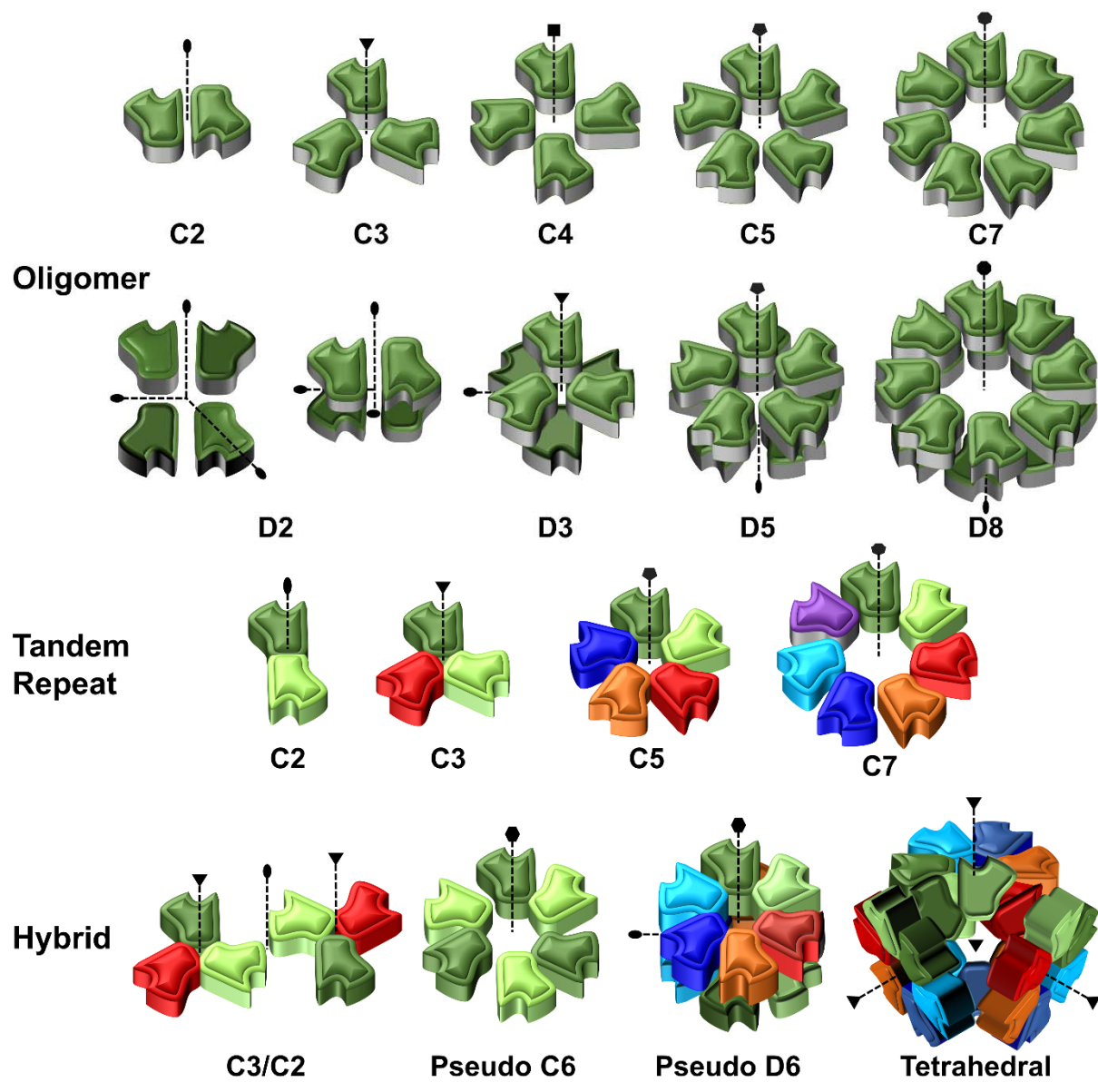


Figure 1

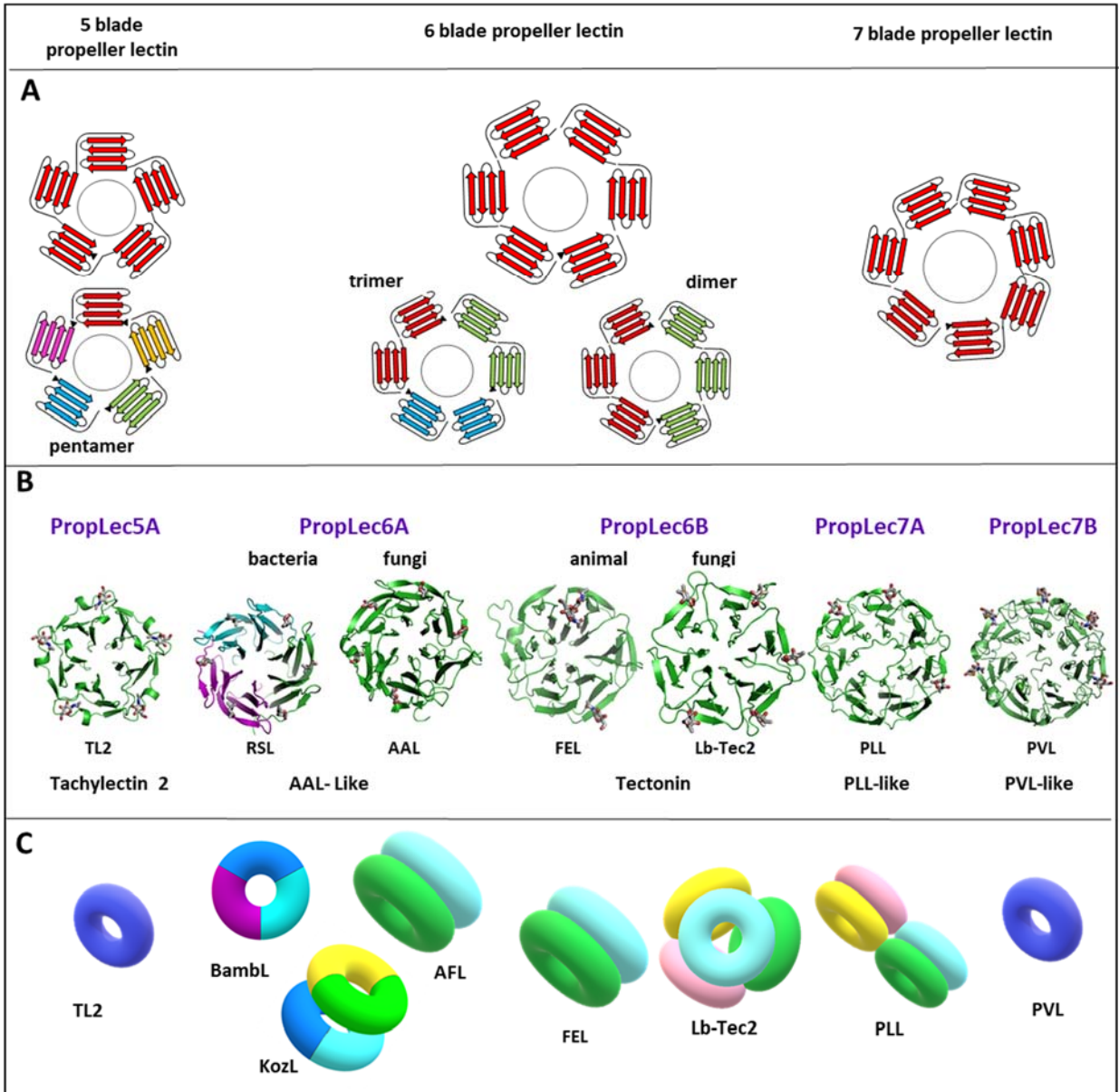


Figure 2

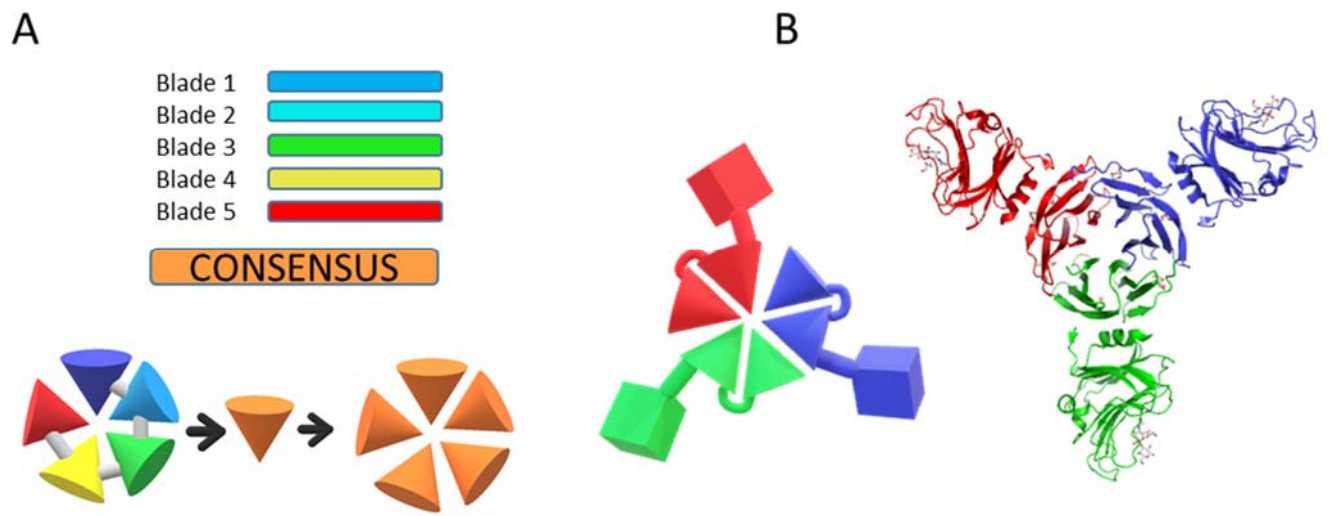


Figure 3

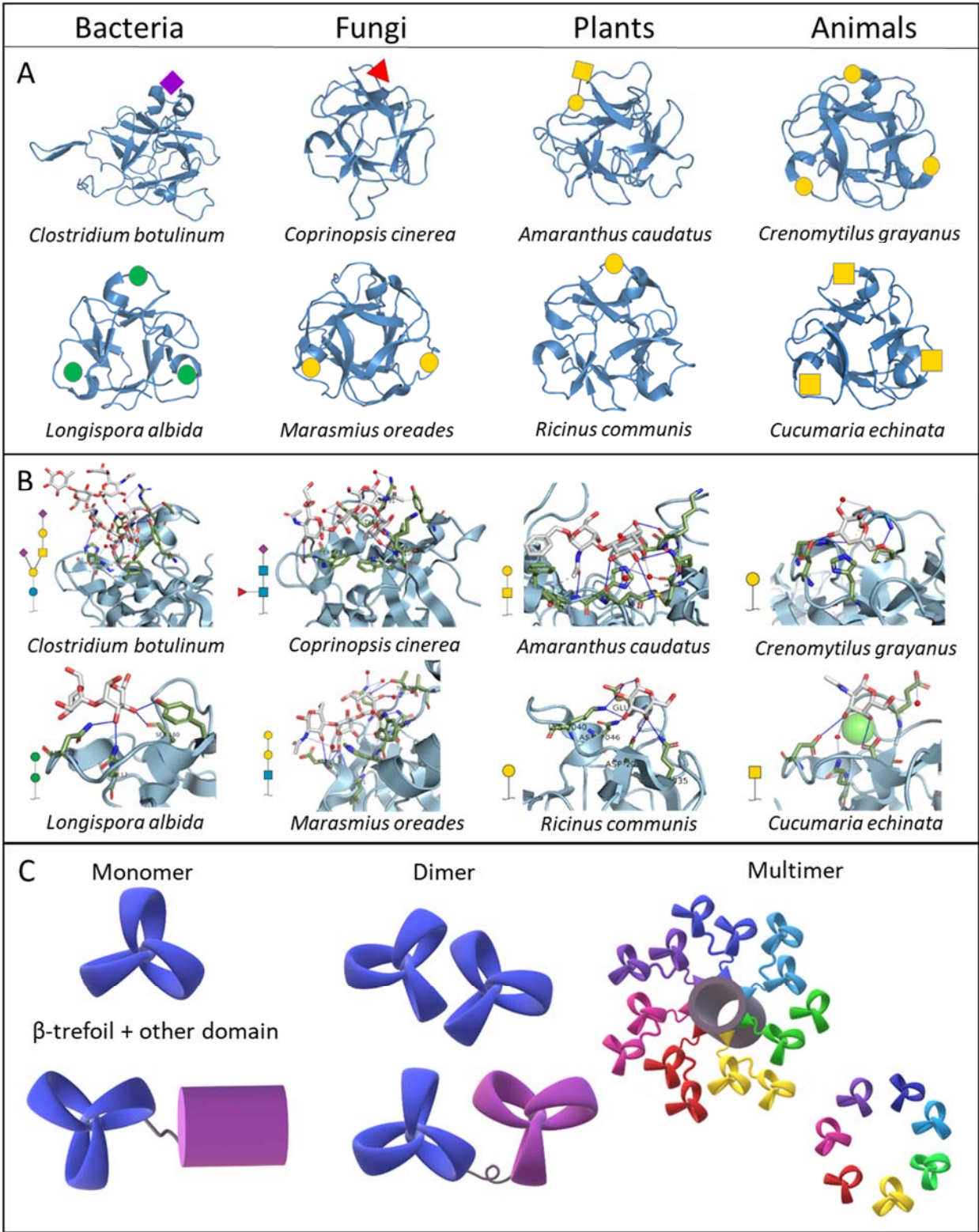


Figure 4