

| Ligand                        | Terminal Epitope                                      | Affinity ( $\mu\text{M}$ ) | Reference                   |
|-------------------------------|-------------------------------------------------------|----------------------------|-----------------------------|
| <b>BC2L-C-C<sub>ter</sub></b> |                                                       |                            |                             |
| D-Mannose                     | Man                                                   | 37.4                       |                             |
| $\alpha\text{MeMan}$          | Man                                                   | 27.6                       | Sulák and co-workers (2011) |
| Trimannose                    | Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man                 | 28.8                       |                             |
| $\alpha\text{MeHept}$         | L,D-manHep                                            | 236                        |                             |
| Diheptose                     | L,D-manHep $\alpha$ 1-3L,D-manHep                     | 88.1                       |                             |
| <b>BC2L-C-N<sub>ter</sub></b> |                                                       |                            |                             |
| $\alpha\text{Me-L-Fuc}$       | Fuc                                                   | 2700                       |                             |
| H-type 2                      | Fuca1-2Gal $\beta$ 1-4GlcNAc                          | 1236                       |                             |
| Lewis b                       | Fuca1-2Gal $\beta$ 1-3[ $\alpha\text{Fuc1-4}$ ]GlcNAc | 213                        | Sulák and co-workers (2010) |
| Lewis x                       | Gal $\beta$ 1-4[ $\alpha\text{Fuc1-3}$ ]GlcNAc        | 196                        |                             |
| Lewis a                       | Gal $\beta$ 1-3[ $\alpha\text{Fuc1-4}$ ]GlcNAc        | 132.1                      |                             |
| H-type 1                      | Fuca1-2Gal $\beta$ 1-3GlcNAc                          | 77.2                       |                             |
| Lewis y                       | Fuca1-2Gal $\beta$ 1-4[ $\alpha\text{Fuc1-3}$ ]GlcNAc | 53.9                       |                             |
| <b>BC2L-C</b>                 |                                                       |                            |                             |
| D-Mannose                     | Man                                                   | 21.8                       | Sulák and co-workers (2011) |
| $\alpha\text{MeMan}$          | Man                                                   | 18.3                       |                             |
| Lewis y                       | Fuca1-2Gal $\beta$ 1-4[ $\alpha\text{Fuc1-3}$ ]GlcNAc | 47.5                       |                             |

## The superlectin BC2L-C: state of the art

### Description

Like BC2L-A, the protein BC2L-C is a “Lec-B” like lectin: its C-terminal domain is 116 residues long and shares 43% identity with LecB. (Sulak et al., 2011) Much like BC2L-A, this domain assembles itself as a homodimeric lectin, featuring two calcium-dependent binding sites. Continuing the similarities, the specificity loop in the binding site bears alanine residues, ensuring specific affinity for mannosides and mannosylated structures in the low micromolar range. Nevertheless, beyond the 116 residues of its C-terminal domain, BC2L-C departs from BC2L-A and becomes a unique lectin.

BC2L-C was initially identified from *B. cenocepacia strain J2315*: its gene *bclC* (NCBI-GI 206562055) codes for 272 amino acids and has been consistently found in other *B. cenocepacia* strains. (Lameignere et al., 2008; Sulak et al., 2010) The C-terminal Lec-B like lectin that led to its discovery accounts for 116 residues. The following 26 amino acids form a serine- and glycine-rich flexible region considered a linker; the remaining amino acids of BC2L-C form its 130 residues-long N-terminal domain. As previously mentioned, lectins presenting many domains are not uncommon: multivalency occurs by repeating the same lectin unit. Nevertheless, the N-terminal domain of BC2L-C (BC2L-C-Nter) was found to be a lectin domain structurally different to BC2L-C-Cter, with well-defined carbohydrate specificity for fucosides. (Sulak et al., 2010) This dual carbohydrate specificity is exceptional and defines the chimeric BC2L-C as a superlectin. (Sulak et al., 2011)

Two seminal studies characterized BC2L-C. In 2010, Šulák and co-workers expressed the 28 kDa native protein, then designed a gene coding for the 156 N-terminal residues of the protein and cloned it in *E. coli* for recombinant production. (Sulak et al., 2010) Indeed, this initial design included the linker region, which was only characterized as such in retrospective. The protein construct thus obtained was labelled BC2L-C-nt and was 187 residues-long due to the un-cleavable 31 residues-long C-terminal histidine tag (HisTag) that was engineered for purification. Size exclusion chromatography revealed the first structural feature of this domain: its elution size corresponded to a 58 kDa protein rather than the expected 19 kDa, signifying a homotrimeric assembly. Assuming it to be a lectin, Šulák and co-workers characterized the new construct by probing it against different monosaccharides. BC2L-C-nt showed specific millimolar affinity towards L-fucose by surface plasmon resonance (SPR). The construct was probed against a glycan array to define specificity further, resulting in a marked preference for

fucosylated histo-blood group epitopes. Indeed, the N-terminal of the BC2L-C superlectin is specific for a well-known lectin target for adhesion, present in the glycocalyx of human epithelial cells. Isothermal titration calorimetry (ITC) allowed further characterization of affinity against human oligosaccharides, returning micromolar values detailed in Table 1.2. The best affinity for BC2L-C-nt occurred for Lewis y (Ley:  $\alpha$ Fuc1-2 $\beta$ Gal1-4[ $\alpha$ Fuc1-3] $\beta$ GlcNAc), with a  $K_D$  of 54  $\mu$ M. The data collected confirmed one binding site per monomer, meaning three per trimer.

| Ligand                        | Terminal Epitope                                        | Affinity ( $\mu$ M) | Reference                   |
|-------------------------------|---------------------------------------------------------|---------------------|-----------------------------|
| <b>BC2L-C-C<sub>ter</sub></b> |                                                         |                     |                             |
| D-Mannose                     | Man                                                     | 37.4                | Šulák and co-workers (2011) |
| $\alpha$ MeMan                | Man                                                     | 27.6                |                             |
| Trimannose                    | Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man                   | 28.8                |                             |
| $\alpha$ MeHept               | L,D-manHep                                              | 236                 |                             |
| Diheptose                     | L,D-manHep $\alpha$ 1-3L,D-manHep                       | 88.1                |                             |
| <b>BC2L-C-N<sub>ter</sub></b> |                                                         |                     |                             |
| $\alpha$ Me-L-Fuc             | Fuc                                                     | 2700                | Šulák and co-workers (2010) |
| H-type 2                      | Fuc $\alpha$ 1-2Gal $\beta$ 1-4GlcNAc                   | 1236                |                             |
| Lewis b                       | Fuc $\alpha$ 1-2Gal $\beta$ 1-3[ $\alpha$ Fuc1-4]GlcNAc | 213                 |                             |
| Lewis x                       | Gal $\beta$ 1-4[ $\alpha$ Fuc1-3]GlcNAc                 | 196                 |                             |
| Lewis a                       | Gal $\beta$ 1-3[ $\alpha$ Fuc1-4]GlcNAc                 | 132.1               |                             |
| H-type 1                      | Fuc $\alpha$ 1-2Gal $\beta$ 1-3GlcNAc                   | 77.2                |                             |
| Lewis y                       | Fuc $\alpha$ 1-2Gal $\beta$ 1-4[ $\alpha$ Fuc1-3]GlcNAc | 53.9                |                             |
| <b>BC2L-C</b>                 |                                                         |                     |                             |
| D-Mannose                     | Man                                                     | 21.8                | Šulák and co-workers (2011) |
| $\alpha$ MeMan                | Man                                                     | 18.3                |                             |
| Lewis y                       | Fuc $\alpha$ 1-2Gal $\beta$ 1-4[ $\alpha$ Fuc1-3]GlcNAc | 47.5                |                             |

Table 2. Affinities measured by ITC for different ligands of the two domains of BC2L-C. Standard deviations are below 5%. Adapted from Šulák and co-workers (2010 and 2011). (Sulak et al., 2010; Sulak et al., 2011)

In the same study, the first crystal structure of this domain (PDB-ID: 2WQ4) was solved at 1.42 Å, as seen in Figure 11. It revealed a trimeric structure presenting  $\beta$  sheets in a jellyroll – Greek-key architecture, which was unprecedented for lectins. It closely resembled the structure of the human tumor necrosis factor (TNF), heavily studied for its role in signalling and immunity, despite no sequence identity. It is worth mentioning at this point that the sequence coding for BC2L-C-Nter did not match any other known sequence, except for a putative protein from the unrelated *Photobacterium luminescens*, an insect pathogen. Three fucoside-populated binding sites occurred at the protomeric interfaces, presented on the same structure face, which facilitates interactions with surface-bound epitopes, as observed in many lectins.

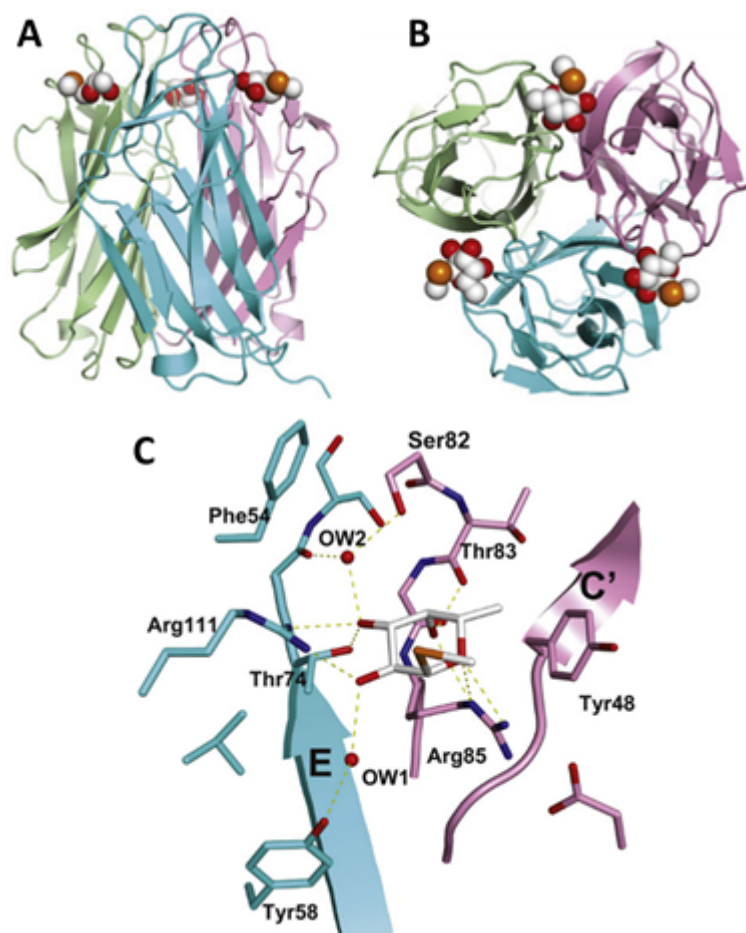


Figure 11. Crystal structure of the N-terminal domain of BC2L-C. A and B: Side view of the homotrimer, ligand is  $\alpha$ -Me-Seleno-L-Fucoside. C: Details of the binding site of BC2L-C-Nter and interactions with the ligand. Water molecules are depicted as red spheres, ligands as spheres or sticks, H-bonds as yellow dashed lines. Adapted from Šulák and co-workers (2010). (Sulak et al., 2010)

The study of the binding interaction provided a rationalization of the observed L-fucose selectivity. Arginine residues Arg111 and Arg85 belong to the two adjacent protomers, encase the monosaccharide from the side and below. They establish hydrogen bonds (H-bonds) with oxygen atoms O2, O3, and O4, O6, respectively. Other noteworthy interactions involve a water molecule buried between ligand and protein, establishing H-bonds with the ligand's O3 atom and residues Tyr75 (carbonyl) and Ser82 (side chain). The remaining interactions appear in Figure 11. They define a novel fucose binding mode previously unseen in other lectins. The selectivity for L-fucosides and related L-galacto-configured structures can be condensed to the substituent at the C4 position: residue Arg85 allows a downward axial substituent, whereas an equatorial substituent would generate steric conflict with the side chain of Ser82.

Although the crystal structure was solved at high resolution, the residues corresponding to the linker region were not visible due to disorder and with high mobility. This study highlights the main limitations of this study of BC2L-C-Nter: the unaccounted flexible tail of the construct was detrimental for stability, with precipitation being a common problem. (Houser et al., 2021) Similarly, attempts to co-crystallize

the protein with larger ligands were unsuccessful since accessibility to the binding side. Thus, structural information of the carbohydrate/ligand interaction with human oligosaccharides could not be obtained, although their affinity for the lectin is several orders of magnitude stronger than the monosaccharide. Molecular modelling led to predicted binding modes of oligosaccharides such as Ley and H-type 1, to be verified by future studies.

In 2011, a subsequent study from Šulák and co-workers uncovered the superlectin as a whole: recombinant versions of the C-terminal domain and the entire protein were produced and probed. The similarities mentioned earlier with BC2L-A were defined, and the affinity for different mannoside and manno-configured heptose ligands were quantified (see Table 2). (Sulak et al., 2011) A crystal structure of the recombinant BC2L-C-Cter dimeric lectin domain was obtained (PDB: 2XR4, Figure 10), and was used for successful computational docking of mannosides as ligands. (2011) Apart from the characterization of the LecB-like domain, the entire protein was characterized in terms of affinity and structure. Glycan array, SPR and ITC technologies were reprinted to confirm the dual-specificity, revealing no overlaps between the binding abilities of both domains. Structural analysis by small-angle X-ray scattering (SAXS) and electron microscopy (EM) revealed a flexible hexameric structure, which accounted for the (three) dimeric and (two) trimeric C- and N-terminal domains (see Figure 12).

This arrangement supported the working theory that the superlectin acts as a 'cellular bridge' or cross-linker of human and bacterial cells by simultaneously engaging with the surface-bound epitopes recognized specifically by each terminal. Indeed, the hypothesis is further supported by three additional findings. Firstly, BC2L-C and -A and -B are secreted by the bacterial cell into the extracellular medium by a yet unknown mechanism. Secondly, these lectins can be found at the bacterial surface, later confirmed for BC2L-A and -B in separate studies. (Inhulsen et al., 2012; Marchetti et al., 2012) Lastly, BC2L-C was released into the extracellular matrix only upon incubation of the cells with mannose, hinting heavily at its regulation by quorum sensing and involvement in virulence. (Sulak et al., 2011)

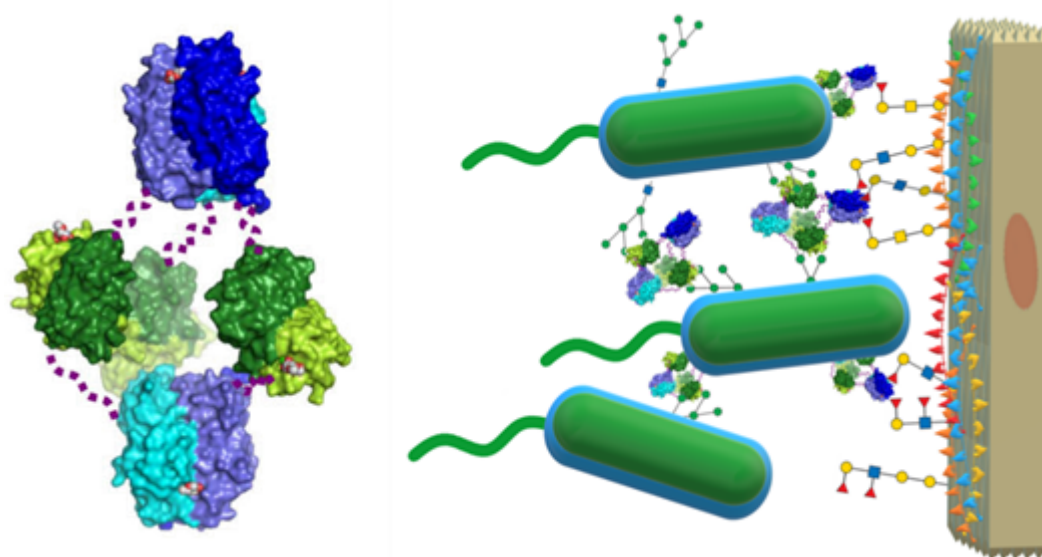


Figure 12. Left: Likely hexameric arrangement of BC2L-C from the SAXS and EM reconstructions. Right: Schematic depiction of the 'cellular bridge' hypothesis: BC2L-C cross-links *B. cenocepacia* and human epithelial cells by binding their LPS mannoside and histo-blood fucoside epitopes, respectively. Adapted from Šulák and co-workers

(2011).(Sulak et al., 2011)

This study evaluated the recently discovered structural relation of BC2L-C-Nter to inflammatory elicitor TNF on a different note. In particular, the study assessed whether exposing epithelial cells to the superlectin would trigger an immune response. A marked increase in interleukin 8 (IL-8) secretion upon treatment with the whole protein or its N-terminus proved the hypothesis. Nevertheless, the inflammatory pathway remained obscured because the obvious candidate, TNF receptor 1 (TNFR1), was not engaged by BC2L-C-Nter. Alternatively, it was demonstrated that the inflammatory response was not linked to carbohydrate binding. The capacity of a virulence factor from *B. cenocepacia* to elicit inflammation through a cytokine-like structure can be tied to the heavy inflammation seen in patients with cepacia syndrome. Demanding proof of the relationship between BC2L-C, inflammation and cepacia syndrome remains to be obtained by further study.

In the decade since its initial characterization as superlectin, BC2L-C has been studied and implemented by many. Across many works, Tateno, Ito and co-workers have established the lectin affinity for fucosylated oligosaccharide epitopes on human pluripotent stem cells. Their new construct of BC2L-C-Nter, called rBC2LC-N, was produced and implemented for fluorescence-based techniques to detect induced pluripotent stem cells and embryonic stem cells against differentiated stem cells (iPSCs, ESCs, etc. SCs, respectively). (Onuma et al., 2013; Tateno et al., 2020; Tateno et al., 2011) The glycoprotein Podocalyxin was identified as a cell-surface ligand of rBC2LC-N, through its H-type three epitopes. (Tateno et al., 2013) This discovery allowed the development of a method for detecting and eliminating tumorigenic pluripotent cells, with a direct use for safety in stem cell therapy. (Tateno et al., 2014; Tateno et al., 2015) In recent years, their efforts have developed chimeric proteins featuring rBC2LC-N and various toxins: lectin-drug conjugates (LDC), aimed at varied cell targets recognized explicitly by the lectin domain. (Shimomura et al., 2018; Tateno et al., 2015) Finally, they have highlighted the usefulness of BC2L-C-Nter to detect specific populations of cancer cells. (Mawaribuchi et al., 2020; Mawaribuchi et al., 2019) Among the many discoveries from this line of research, more support for the 'cellular bridge' hypothesis can be found: BC2L-C-Nter probes bind to human cells via the histo-blood groups in their glycocalyx with antibody-level sensitivity, and they seem to bind to cell lines with epithelial characteristics specifically. (Breiman et al., 2016; Mawaribuchi et al., 2019; Onuma et al., 2013; Shimomura et al., 2018)

Apart from the work of this group on stem and cancer cells, others have benefitted from BC2L-C as a tool for varied endeavours: detection of histo-blood epitopes for cell characterization, development of protein stability screening kits, validation of microbe-oriented glycan arrays, and validation of Fragment Molecular Orbital (FMO) tools for the analysis of protein/ligand interactions. (Geissner et al., 2019; Houser et al., 2021; Sugahara et al., 2017; Tokiwa et al., 2019; Ziganshina et al., 2020) Finally, some groups have challenged antagonizing the superlectin with an early array of fucoside ligands, reaching some degree of success thanks to multivalency. (Kasakova et al., 2018; Thai Le et al., 2019) From these campaigns, the best synthetic ligand for BC2L-C-Nter was a calix[4]arene-based tetravalent fucoside, which showed a 256-fold increase of potency compared to L-fucose for inhibition of hemagglutination of red blood cells. A cross-linking test confirmed the capacity of this inhibitor to aggregate *B. cenocepacia* cells by engaging the surface-bound lectin. Bearing mostly unmodified C-fucosides, this compound proves a successful multivalent effect: the potency per sugar corresponds to a 64-fold increase. With multivalency validated as a viable strategy to inhibit this virulence factor, S-fucoside glycomimetics were put forward in an attempt to develop glycomimetic monovalent inhibitors. Nevertheless, this optimization was hindered by a lack of biophysical techniques to assess affinity

constants and structural data to rationalize the relative potency observed in hemagglutination assays. (Thai Le et al., 2019)

### **Category**

1. News