Heliyon 8 (2022) e10396

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

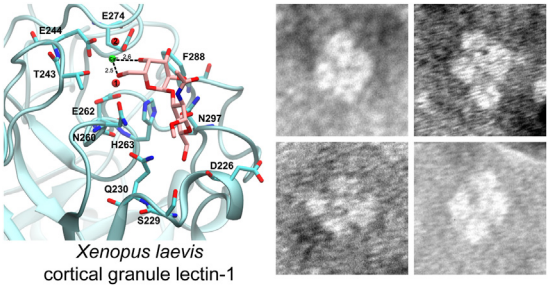
Research article

Biochemical and ligand binding properties of recombinant *Xenopus laevis* cortical granule lectin-1

Peerapon Deetanya [a,b], Thassanai Sitthiyotha [c], Nusara Chomanee [d], Surasak Chunsrivirot [c], Kittikhun Wangkanont [a,b,*]

[a] *Center of Excellence for Molecular Biology and Genomics of Shrimp, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand*
[b] *Center of Excellence for Molecular Crop, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand*
[c] *Structural and Computational Biology Research Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, 10330, Thailand*
[d] *Electron Microscopy Unit, Department of Pathology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand*

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Xenopus laevis cortical granule lectin-1
XCGL-1
XL35
Intelectin
Omentin

ABSTRACT

Intelectins are putative innate immune lectins that are found throughout chordates. The first intelectin reported was *Xenopus laevis* cortical granule lectin-1 (XCGL-1 or XL-35). XCGL-1 is critical in fertilization membrane development in *Xenopus*. Here, we explored the biochemical properties of XCGL-1. The cysteines responsible for forming intermolecular disulfide bonds were identified. XCGL-1 adopted a four-lobed structure as observed by electron microscopy. The full-length XCGL-1 and the carbohydrate recognition domain (CRD) bind galactose-containing carbohydrates at nanomolar to micromolar affinities. Molecular modeling suggested that galactoside ligands coordinated the binding site calcium ion and interacted with residues around the groove made available by the non-conserved substitution compared to human intelectin-1. Folding conditions for production of recombinant XCGL-1 CRD were also investigated. Our results not only provide new biochemical insights into the function of XCGL-1, but may also provide foundation for further applications of XCGL-1 as glycobiology tools.

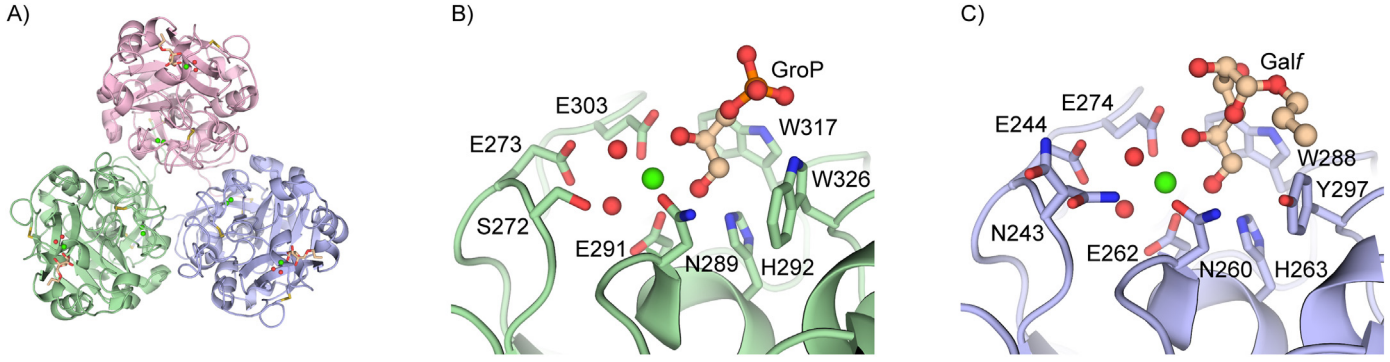
* Corresponding author.
E-mail address: kittikhun.w@chula.ac.th (K. Wangkanont).

https://doi.org/10.1016/j.heliyon.2022.e10396
Received 13 October 2021; Received in revised form 5 March 2022; Accepted 15 August 2022

1. Introduction

Intelectins are proposed to be innate immune lectins that are conserved in chordate species [1, 2]. Intelectins are upregulated upon microbial or parasitic infection [2, 3, 4, 5, 6, 7]. Opsonization activity has been demonstrated in ascidian and blunt snout bream intelectins [8, 9]. Human intelectin-1 (hIntL-1) is a disulfide-linked trimeric lectin that recognizes microbe-specific carbohydrate residues (Figure 1A) [10, 11]. *Xenopus* embryonic epidermal lectin (XEEL) is expressed in *Xenopus* embryo and is proposed to bind microbes to prevent infection [12, 13]. XEEL is likely a dimer of trimers that has similar ligand specificity to hIntL-1 [14]. In addition to XEEL, *Xenopus laevis* has at least 6 other intelectin paralogs. *X. laevis* cortical granule lectin-1 (XCGL-1), or XL35, is the founding member of the intelectin family [15]. Cortical granule is also known to contain a similar protein, XCGL-2, as well [16]. *Xenopus* serum contains two intelectins: XCL-1 and XCL-2 [17, 18]. These intelectins are upregulated upon injection of lipopolysaccharide (LPS). In addition, the gut of *X. laevis* contains two intelectins that could bind bacteria and are upregulated upon LPS stimulation [19]. Despite the growing interest in the intelectin family, the biochemistry and the structure of the first intelectin discovered, XCGL-1, still has details to be explored.

XCGL-1, a galactose-binding intelectin, is found in *Xenopus* oocyte and is responsible for the formation of the fertilization membrane to prevent polyspermy [15, 16, 20, 21]. Glycan array results also suggested that XCGL-1 bound α-galactosides, especially α-D-Gal-(1→3)-α/β-D-GalNAc [22]. It is unclear from earlier studies whether the protein preparation also contains XCGL-2 or other lectins that may have different



D)

```
                                                  18
XCGL-1      ---MLV---HILLLLVTGGLSQSCEPVVIVASKNMVK-QL--------------------    33
XEEL        MGKMLSYSLLLLALAFPAGHAGSCEQASISEKKEKILNLLACWTEGNADNSLSRSGGSPT    60
hIntL-1     ---MNQLSFLLFLIATTRGWSTDE-----------ANTYFKEWTCS--------------    32
               *      :: :    * : .                 :

              35
XCGL-1      --DCDKFRSCKEIKDSNEEAQDGIYTLTSSDGISYQTFCDMTTNGGGWTLVASVHENNMA    91
XEEL        GDMNYGYRSCNEIKSSDSRAPDGIYTLATEDGESYQTFCDMTTNGGGWTLVASVHENNMF    120
hIntL-1     -SSPSLPRSCKEIKDECPSAFDGLYFLRTENGVIYQTFCDMTSGGGGWTLVASVHENDMR    91
                   ***:***..   * **:* * :.:*  ********:.*************:*

XCGL-1      GKCTIGDRWSSQQGNRADYPEGDGNWANYNTFGSAGGATSDDYKNPGYYDIEAYNLGVWH    151
XEEL        GKCTVGDRWSTQQGNMLQNPEGDGNWANYATFGLPEGATSDDYKNPGYYDIEAKNLALWH    180
hIntL-1     GKCTVGDRWSSQQGSKAVYPEGDGNWANYNTFGSAEAATSDDYKNPGYYDIQAKDLGIWH    151
            ****:*****:***.    ********** ***   .**************:* :*.:**

XCGL-1      VPNKTPLSVWRNSSLQRYRTTDGILFKHGGNLFSLYRIYPVKYGIGSCSKDSGPTVPVVY    211
XEEL        VPNKTPMVMWRNSSILRYRTQNGFLTEEGGNLFELYKKYPVKYDIGKCLADNGPAVPVVY    240
hIntL-1     VPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYGEGKCWTDNGPVIPVVY    211
            ****:*:  *****: **** *:*   * *** :*: *****. *.*  *.**.:****
                                                243 260
                                                244 262
                                                    263
XCGL-1      DLGSAKLTASFYSPDFRSQFTPGYIQFRPINTEKAALALCPGMKMESCNVEHVCIGGGGY    271
XEEL        DLGSAEKTASLYSPNGRSEFTPGFVQFRAVNSERATLALCAGVKVKGCNVEHHCIGGGGY    300
hIntL-1     DFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCNTEHHCIGGGGY    271
            *:*.*: *** ***  : :** *::*** .*.*:*: *** *::: .**.** *******

             274          288      297
XCGL-1      FPEADPRQCGDFAAYDFNGYGTKKFNSAGIEITEAAVLLFYL     313
XEEL        IPEGSPRQCGDFAALDWDGYGTNLGWSASKQIIEAAVMLFYR     342
hIntL-1     FPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR     313
            :**..*:*****:. *:.****:   *:. :* ****:***
```

Figure 1. A) Previously reported crystal structure of trimeric hIntL-1 (PDB ID 4WMY) [10]. B) Ligand binding site in a previously reported crystal structure of XEEL (PDB ID 4WN0) [14]. C) Ligand binding site in a previously reported crystal structure of hIntL-1 (PDB ID 4WMY) [10]. D) Sequence alignment between XCGL-1, XEEL, and hIntL-1. Cysteine residues are labeled in yellow. Residues in the ligand binding sites are labeled in cyan. *,:, and. denote conserved, highly similar, and similar residues, respectively.

ligand specificity and function. Therefore, recombinant expression and characterization of XCGL-1 may yield more conclusive information. Recombinant full-length XCGL-1 was expressed and refolded from *Escherichia coli*-produced inclusion bodies, but the disulfide-link oligomeric state of the refolded XCGL-1 was not demonstrated [20, 23]. As mentioned earlier, oocyte-derived XCGL-1 could bind galactose. In contrast, hIntL-1 and XEEL are reported to bind the exocyclic 1,2-diol moiety on the saccharide ligand, such as in galactofuranose (Gal*f*) or glycerol phosphate (GroP) [10, 14]. In the previously published crystal structures of hIntL-1 and XEEL, the exocyclic 1,2-diol moiety is surrounded by aromatic residues (Figures 1B and 1C). These aromatic residues were proposed to provide steric constraints and also stereoelectronic interactions for ligand binding [24]. However, sequence alignment indicates that XCGL-1 has a phenylalanine at position 288 instead of a tryptophan in the equivalent position in XEEL and hIntL-1 (Figure 1D). The non-conservative asparagine substitution at position 297, instead of a tryptophan in XEEL or a tyrosine in hIntL-1, was also observed. Thus, these amino acid differences in the aromatic box may alter the ligand specificity. Therefore, we aim to further explore the mechanism of ligand binding.

Like other intelectins, XCGL-1 is a disulfide-linked oligomers of various sizes [15, 21, 25, 26]. The oligomeric state of intelectins could be crucial for cellular recognition, especially by multivalent interactions [27, 28, 29]. Therefore, we aim to determine the oligomeric architecture of XCGL-1. Non-conserved cysteines in intelectins are known to form intermolecular disulfide bonds in intelectins [10, 11, 14]. By aligning the sequence of XCGL-1 with hIntL-1 and XEEL, of which crystal structures are available, it is possible to predict which cysteines are likely involved in the formation of intermolecular disulfide linkages. From this type of analysis, C18 and C35 are the candidates (Figure 1D). Therefore, we aim to investigate the function of these cysteine residues by site-directed mutagenesis.

In addition to the carbohydrate specificity, structure determination of XCGL-1 is of particular interest to elucidate its ligand binding mechanism. However, significant amount of correctly-folded protein is required. Because intelectins are oligomeric, XCGL-1 likely need to be truncated to obtain the conserved carbohydrate recognition domain (XCGL-1 CRD) without the N terminal fragment that is responsible for the inhomogeneous disulfide-linked oligomerization. The conserved intelectin CRD fold is a highly decorated fibrinogen-like domain that is

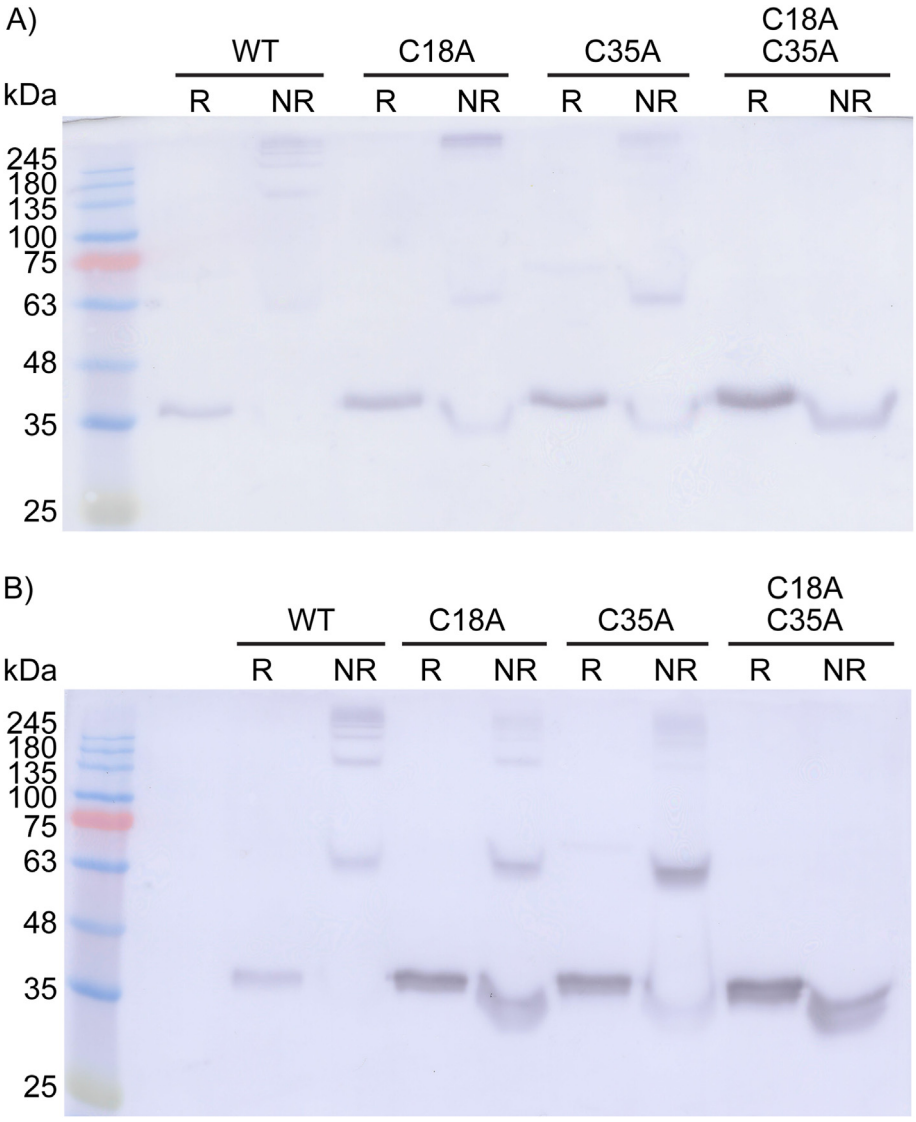


Figure 2. Western blot analysis of wild-type (WT) XCGL-1 and XCGL-1 with mutations at non-conserved cysteine residues using a pan-intelectin antibody. R = reducing conditions and NR = non-reducing conditions. A) XCGL-1 expressed in HEK293T cells. The uncropped image is shown in Figure S8. B) XCGL-1 expressed in *T. ni* cells. The uncropped image is shown in Figure S9.

distinct from its clostest fibrinogen-like relative, the ficolin family, and other calcium-dependent lectins, such as the C-type lectins [14]. From sequence alignment, it is likely that the conserved domain starts from D36, right after the non-conserved C35 (Figure 1D). Therefore, expression and purification of XCGL-1 CRD were explored. We also compared ligand specificity and affinity, as well as ligand recognition in the cellular context, between full-length XCGL-1 and XCGL-1 CRD. These experiments will not only yield basic biochemical knowledge of XCGL-1 and the intelectin family, but will also provide methods to produce active XCGL-1 for further applications.

2. Results and discussion

2.1. Disulfide-linked oligomeric state of XCGL-1

Full length XCGL-1 expressed in the HEK293T mammalian cells (Figure 2A) and the *Trichoplusia ni* insect cells (Figure 2B) revealed multiple disulfide-linked oligomeric states, evident by ladders in non-reducing conditions, as previously described for oocyte-derived XCGL-1 [15, 21, 26]. To identify the cysteine residues that form the intermolecular disulfide bridges, full length XCGL-1 harboring C18A, C35A, and C18A-C35A mutations, expressed from both HEK293T and *T. ni* cells were examined by western blotting. The C18A or C35A mutation alone still retained intermolecular disulfide bonds under non-reducing conditions. However, the C18A-C35A double mutation resulted in the complete loss of disulfide-linked oligomers. Therefore, both C18 and C35 are responsible for intermolecular disulfide bonds in XCGL-1, in both expression systems. These results also suggest that the signal peptide cleavage site is upstream of C18. Indeed, Edman degradation of the XCGL-1 expressed from *T. ni* cells and purified with lactose resin affinity chromatography revealed a blank first cycle and EPVV in the subsequent four cycles (Table S1). Because the phenylthiohydantoin derivative of cysteine cannot be detected [30], it is likely that the first amino acid of XCGL-1 is a cysteine. In conjunction with the observation that the second to fifth cleavage cycle aligns with E19 to V22, we conclude that the first amino acid of XCGL-1 is C18.

2.2. Expression, refolding, and purification of XCGL-1 and XCGL-1 CRD

Due to the ease of scaling up, the insect cell expression system was chosen for subsequent experiments. After the establishment of the signal peptide cleavage site, His_6 XCGL-1 and His_6 XCGL-1 CRD constructs were made for insect cell expression. XEEL signal peptide was used as it is known to be an efficient secretion signal [14, 31, 32]. Ni-NTA affinity chromatography is a superior approach due to higher binding capacity compared to in-house-generated lactose resin. In addition, baculovirus-derived chitinase is a major contaminant when carbohydrate resin is used for affinity chromatography [33].

To produce large quantity of XCGL-1 CRD for biochemical characterization and further utilization, expression of XCGL-1 CRD in *Escherichia coli* was explored. An XCGL-1 CRD expression construct comprising of residue 36 to 313 was cloned into pET28a. The protein was produced as inclusion bodies in Tuner (DE3) (Figure S1). The molecular weight of the expressed XCGL-1 CRD was around 30 kDa as expected. The molecular weight calculated from the amino acid sequence is 30.7 kDa.

XCGL-1 CRD inclusion bodies were then dissolved in urea, and screened for optimal refolding buffers (Table S2). The two conditions that resulted in minimal protein precipitation (Figure S2) were identified: Arginine buffer (50 mM Tris pH 10.0, 10 mM $CaCl_2$, 2 mM GSSG, 0.2 mM GSH, and 0.3 M L-arginine) and NDSB-201 buffer (50 mM Tris pH 10.0, 10 mM $CaCl_2$, 2 mM GSSG, 0.2 mM GSH, and 0.75 M NDSB-201). Due to the higher yield, as judged by band intensity, and the lower cost of L-arginine compared to NDSB-201, the arginine buffer was chosen for a larger scale production of XCGL-1 CRD. The refolded protein could be purified by lactose-agarose affinity chromatography (Figure S3).

The disulfide-linked oligomeric states of the purified recombinant proteins were then investigated under reducing and non-reducing conditions. Purified His_6 XCGL-1 exhibited the same oligomeric state pattern under non-reducing conditions (Figure 3A) as the untagged protein (Figure 2B). Therefore, it is likely that the tagging strategy did not affect protein folding. Larger disulfide-linked oligomeric states was abolished when His_6 XCGL-1 was truncated to His_6 XCGL-1 CRD (Figure 3A). However, a less intense band around 70 kDa was observed under both reducing and non-reducing conditions. This band was observed to some extent in non-tagged XCGL-1 expressed in HEK293T and *T. ni* cells in Figure 2, but was not observed in the refolded *E. coli*-derived XCGL-1 CRD. This observation suggests that there are potential differences in the protein structure despite the same ability to bind lactose. Previously, the equivalent band was observed in hIntL-1 and also XEEL, even though the crystal structure of the trimeric XEEL CRD showed no intermolecular disulfide bonds [14]. These observations suggested that the observed 70-kDa band was not likely disulfide-linked and that the CRD of intelectins could form oligomeric states in the presence of SDS. This behavior under SDS-PAGE was also observed with the soluble snake venom Phospholipase A_2 that could oligomerize under SDS-PAGE conditions even though no disulfide linkage was present in the crystal structure [34].

In our hands, recombinant XCGL-1 variants were prone to aggregation and they precipitated out of solution within a few days. The proteins solution could be frozen at -80 °C, but they precipitated within a few days after thawing as well. The proteins also precipitated if concentrated beyond 0.3–0.4 mg/mL. These behavior hindered characterization of the native complexes with techniques that required high protein concentration and extensive amount of time, such as analytical ultracentrifugation or gel filtration chromatography. Because the natural role of this protein is to form the fertilization membrane to prevent polyspermy [15, 16, 20], it may not be surprising that these proteins are prone to aggregation and they may not form homogeneous oligomeric states, especially at high concentrations.

The molecular weight of His_6 XCGL-1 bands were further analyzed (Figures 3B and 3C). The heaviest band had the estimated molecular weight of 390 kDa, which was comparable to 12 times the expected molecular weight of XCGL-1 from the amino acid sequence (12×33.6 kDa = 403.2 kDa). The 137, 252, and 333 kDa bands were roughly 4-, 8-, and 10-meric XCGL-1, respectively. The significance of these disulfide bonding patterns remains elusive.

In addition to the disulfide-linked oligomeric state, it is known that oocyte-derived XCGLs contain N-linked glycans [16]. Therefore, we examined the N-linked glycosylation status of XCGL-1 expressed in *T. ni* insect cells by PNGase F and Endo H digestion under denaturing conditions. A lower band was observed for both His_6 XCGL-1 and His_6 XCGL-1 CRD (Figure S4A), suggesting that the *T. ni* cells did add N-linked glycans to XCGL-1. However, deglycosylation was incomplete, even though RNase B was completely deglycosylated under the same experimental conditions. Although it was possible that more optimal deglycosylation conditions were yet to be found, the results also fit with our earlier observation that XCGL-1 might be able to oligomerize in the presence of SDS and this could impede enzymatic deglycosylation.

2.3. Electron microscopy and dynamic light scattering of XCGL-1

Although analytical ultracentrifugation experiments suggested that the oocyte-derived full-length XCGLs, which are likely a mixture of XCGL-1 and XCGL-2, are 12-mer [15, 21, 26], the exact molecular arrangement of the monomers are unclear. Therefore, we performed negative staining transmission electron microscopy (TEM) to investigate the oligomeric arrangement of His_6 XCGL-1 (Figure 4A). The results revealed that His_6 XCGL-1 had a 4-lobed structure of 100–120 Å in diameter (Figure 4B). If the CRD adopts a trimeric structure as in XEEL and hIntL-1, and XCGL-1 is a 12-mer as previously suggested, we propose that XCGL-1 is a tetramer of trimeric CRDs.

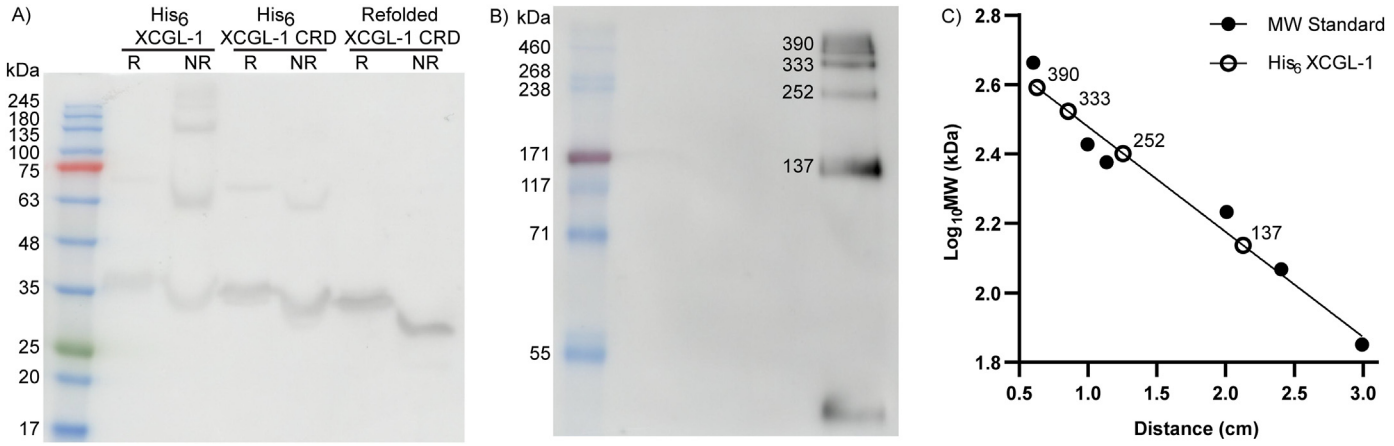


Figure 3. Molecular weight of XCGL-1 species. A) Western blot of XCGL-1 species under reducing (R) and non-reducing (NR) conditions probed with a pan-intelectin antibody. The uncropped image is shown in Figure S10. B) Western blot of His_6 XCGL-1 under non-reducing conditioned probed with an anti-His_6 antibody. The uncropped image is shown in Figure S11. C) Molecular weight standard plot with the molecular weight of the marker (solid circles) and molecular weight of His_6 XCGL-1 bands (white circles).

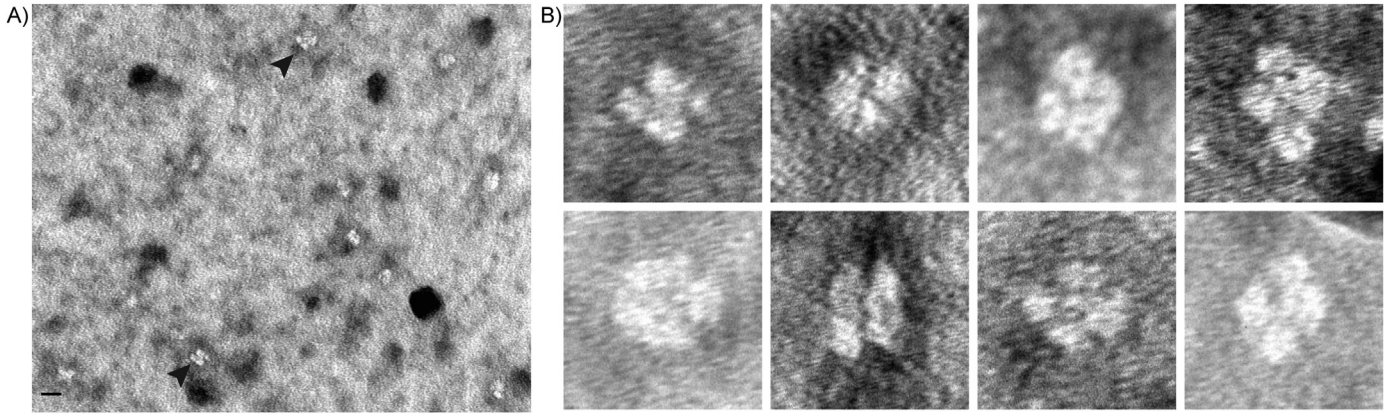


Figure 4. Transmission electron micrograph of His_6 XCGL-1. A) A representative field. The black bar at the bottom left corner is 200 Å long. The 4-lobed particles are indicated by arrows. B) Representation particles, not including those in A). Each particle box is 200 × 200 Å.

Dynamic light scattering data were also obtained to determine the hydrodynamic size of the protein. Size distribution by intensity from all measurements showed a >100 nm peak that we attributed as dust particles because the peak was absence when the data was transformed into the size distribution by number. The hydrodynamic size of His_6 XCGL-1 was 15.9 ± 0.7 nm (159 ± 7 Å), which was consistent with the size measured from TEM (Figure 5A). The hydrodynamic size, which includes a surrounding solvent layer, is expected to be bigger than the size determined from dried negative stain samples by TEM. The size of each lobe of the 4-lobed structure in the TEM images was around 35–45 Å in diameter, which was similar to the hydrodynamic size of 5.9 ± 0.2 nm (59 ± 2 Å) of His_6 XCGL-1 CRD (Figure 5B). The size of His_6 XCGL-1 CRD was also comparable to that of the trimeric XEEL CRD determined by small-angle X-ray scattering (radius of gyration of 27.9 Å or diameter of 55.8 Å) [14]. Therefore, truncation of His_6 XCGL-1 to His_6 XCGL-1 CRD did reduce the size of the protein. Note that, discounting the dust peak in

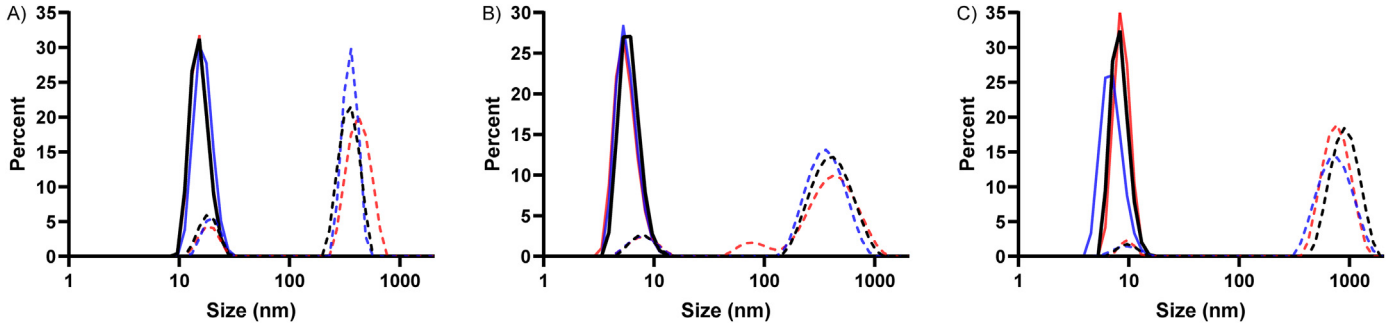


Figure 5. Size measurement of XCGL-1 variants by dynamic light scattering. A) His_6-XCGL-1. B) His_6-XCGL-1 CRD. C) Refolded XCGL-1 CRD. Three independent data sets were measured for each protein and shown as three different colors. The dash lines showed size distribution by intensity and the solid lines of the same color showed size distribution by number from the same experiment.

the DLS data, there was only one peak for His_6 XCGL-1 CRD. Thus, the 70-kDa band observed in Figure 3 was either an insignificant population in solution or merely an SDS-induced artifact as described earlier. In contrast to the His_6 XCGL-1 CRD, the refolded XCGL-1 CRD had the hydrodynamic size of 8.1 ± 0.8 nm (81 ± 8 Å) (Figure 5C). The results were consistent with our observation that the refolded XCGL-1 CRD was different from the *T. ni* cell-derived XCGL-1 CRD despite the ability to bind lactose.

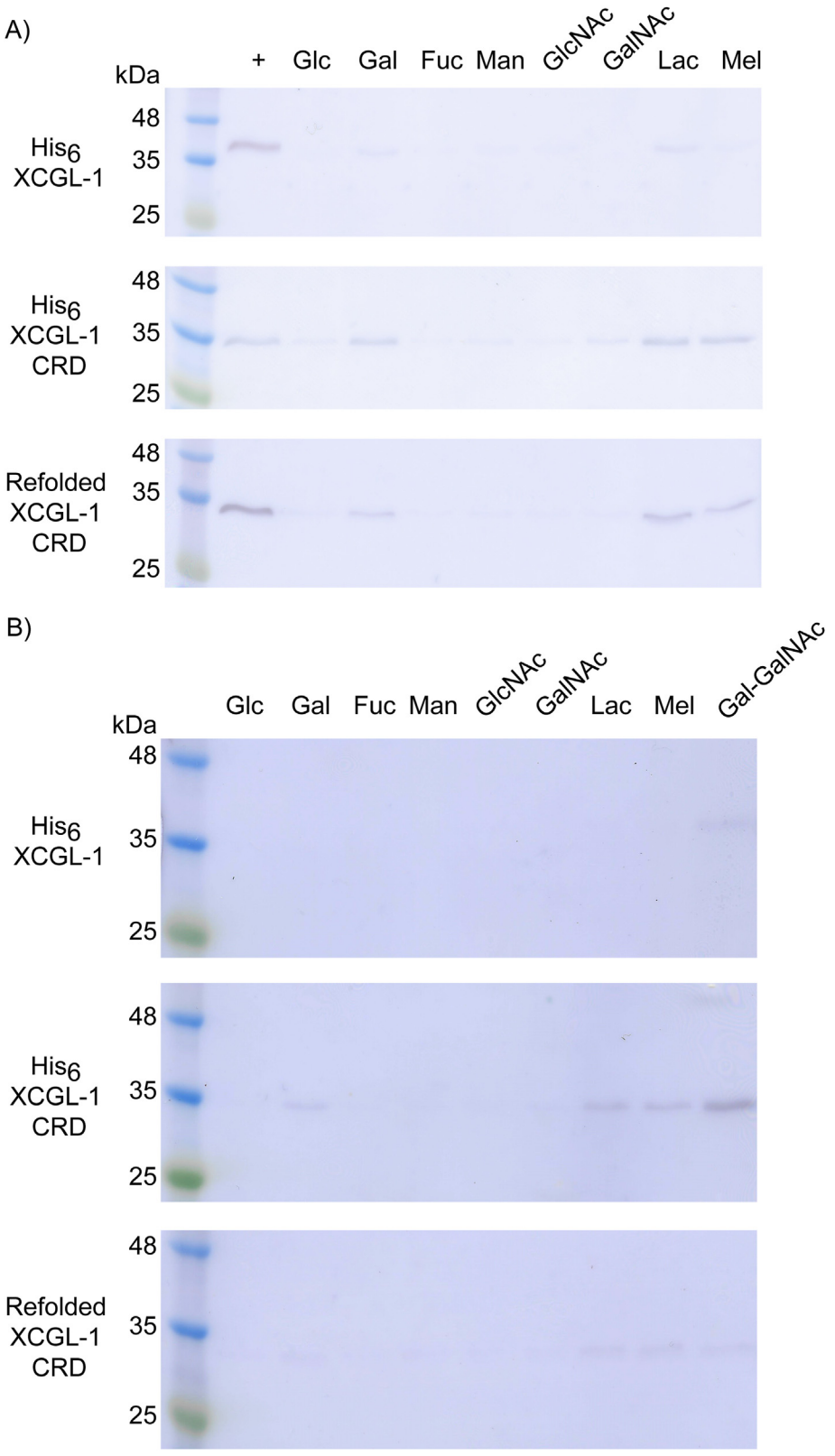


Figure 6. A) Elution of XCGL-1 bound to different carbohydrate resins by EDTA. The uncropped images are shown in Figure S12, Figure S13, and Figure S14. B) Competitive elution of XCGL-1 bound to lactose resin by various carbohydrates. The uncropped images are shown in Figure S15, Figure S16, and Figure S17. XCGL-1 CRD was detected with a pan-intelectin antibody. His_6-tagged proteins were detected with an anti-His tag antibody (pan-intelectin antibody yielded similar results). Glc = D-glucose, Gal = D-galactose, Fuc = L-fucose, Man = D-mannose, GlcNAc = N-acetyl-D-glucosamine, GalNAc = N-acetyl-D-galactosamine, Lac = lactose (β-D-Gal-(1→4)-D-Glc), Mel = melibiose (α-D-Gal-(1→6)-D-Glc), Gal-GalNAc = α-D-Gal-(1→3)-D-GalNAc.

2.4. Carbohydrate specificity of XCGL-1

His_6 XCGL-1 and His_6 XCGL-1 CRD, both expressed in *T. ni* cells, and the XCGL-1 CRD refolded from inclusion bodies, were tested for binding to different carbohydrate-functionalized resin. The bound protein was eluted by EDTA and detected by western blotting (Figure 6A). The results showed that all forms of XCGL-1 were specific to galactose or galactose containing disaccharides, such as lactose (β-galactoside) or melibiose (α-galactoside). This result is consistent with the property of the cortical granule lectins purified from *Xenopus* oocytes [15, 21] and the full length

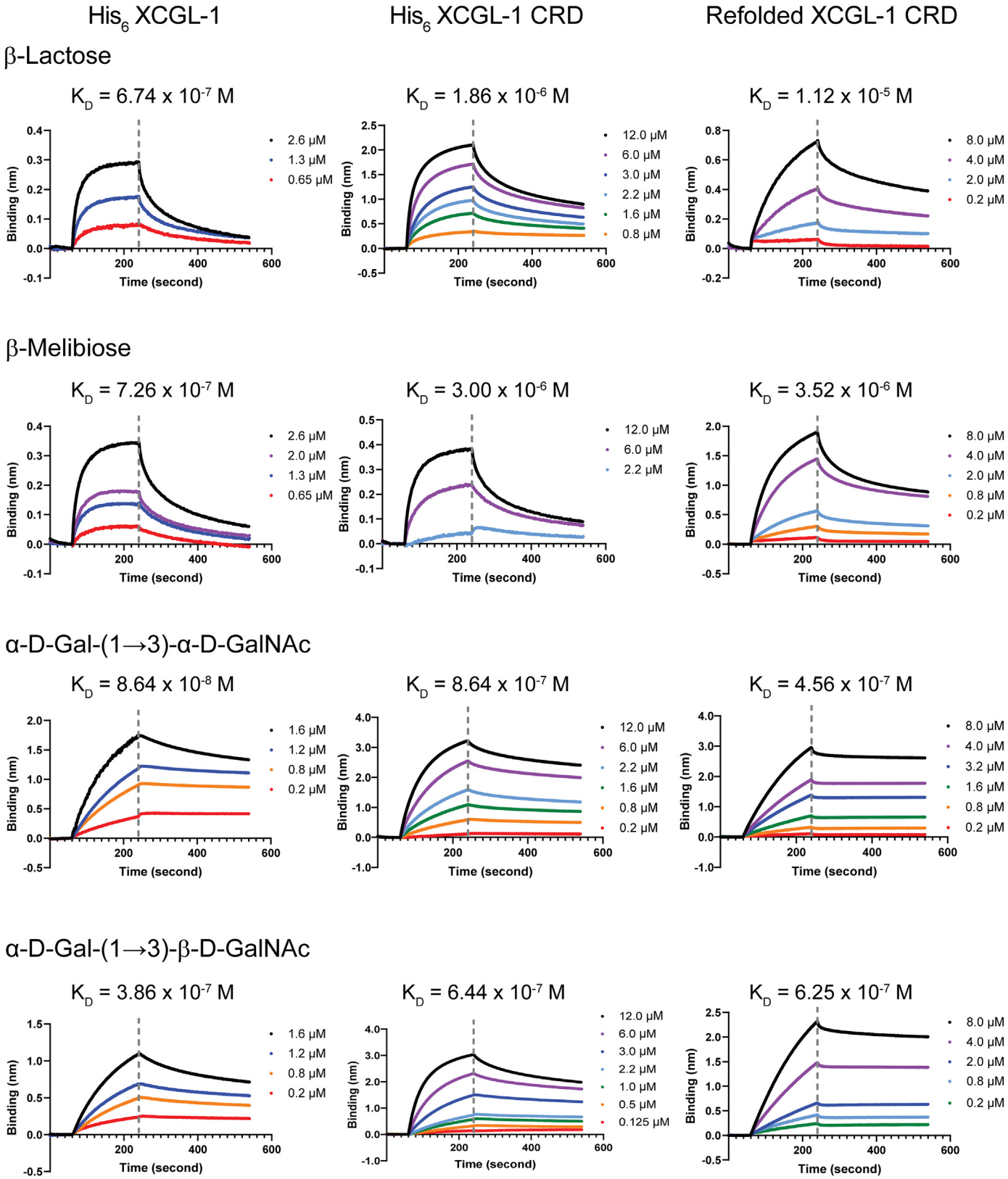


Figure 7. Biolayer interferometry sensorgram XCGL-1 variants toward immobilized carbohydrates.

XCGL-1 that was refolded from inclusion bodies produced in *E. coli.* [20, 23] We further verified our results by competitively eluting XCGL-1 from the lactose resin with various carbohydarte solutions, including α-D-Gal-(1→3)-D-GalNAc that was previously identified to be a potent XCGL-1 ligand (Consortium for Functional Glycomics, primscreen_758) (Figure 6B) [22]. As expected, galactose, lactose, melibiose, and α-D-Gal-(1→3)-D-GalNAc were able to elute XCGL-1 competitively. Interestingly, compared to other carbohydrates, α-D-Gal-(1→3)-D-GalNAc appear to elute *T. ni* cell-derived XCGL-1 more than the refolded *E. coli* XCGL-1 CRD.

The α-D-Gal-(1→3)-D-GalNAc is similar to the T antigen (β-D-Gal-(1→3)-D-GalNAc) found in mucin [35], which is also upregulated in breast cancer [36]. The T antigen was present in the glycan array, but was not recognized by XCGL-1. Thus, for the D-Gal-(1→3)-D-GalNAc disaccharide, the stereochemistry of the glycosidic bond between Gal and GalNAc was important for XCGL-1 binding. It is unclear if the α-D-Gal-(1→3)-D-GalNAc motif is present in *Xenopus* oocytes, although both α and β terminal galactosides are known to be present in the egg jelly [37, 38]. The α-D-Gal-(1→3)-D-GalNAc motif is not unknown to nature. It is the Core 8 in O-linked glycans [35, 39, 40]. It is also found in the conotoxin tx5a from *Conus textile* with an α linkage (α-D-Gal-(1→3)-α-D-GalNAc) to a threonine residue [41]. The terminal α-D-Gal-(1→3)-α-D-GalNAc motif is present in the egg jelly coat mucins of the amphibian *X. tropicalis* [42] and *Bufo bufo* [43], thus this motif is likely present in *X. laevis* as well. Therefore, the interaction between XCGL-1 and α-D-Gal-(1→3)-α-D-GalNAc is likely biologically relevant in *X. laevis* egg. Moreover, the α-D-Gal-(1→3)-β-D-GalNAc disaccharide is found in the cell wall of bacteria *Streptomyces antibioticus* 39 [44]. Although in a non-terminal position, the α-D-Gal-(1→3)-β-D-GalNAc motif exists in the lipopolysaccharide of *Escherichia coli* O 55 [45] and 128 [46]. While it is not known whether the non-terminal α-D-Gal-(1→3)-β-D-GalNAc could bind XCGL-1, its presence in bacteria mentioned suggests that other bacteria might be able to biosynthesize such a motif as well. Therefore, XCGL-1 has the potential to be utilized as a bacteria identification or typing tool, when more bacteria containing this motif are discovered in the future.

2.5. Binding affinity of XCGL-1 towards carbohydrate ligands

Because the binding affinities of XCGL-1 towards galactoside ligands have not been reported, the binding kinetic parameters between XCGL-1 variants against galactosides ligands were acquired using biolayer interferometry (BLI) for dissociation constant (K_D) determination (Figure 7). The biotinylated ligands were immobilized onto the biosensor tip pre-coated with streptavidin. The tip was then dipped into a protein solution to measure the association rate (k_a). The tip was subsequently immersed in a buffer solution to measure the dissociation rate (k_d). The binding kinetic parameters and the K_D values were provided in Table S3. Overall, the K_D values of XCGL-1 variants were in a similar range compared to hIntL-1 (85 ± 14 nM) [10] and XEEL CRD (4.1 ± 0.5 μM) [14]. His_6 XCGL-1 bound β-lactose, β-melibiose, and α-D-Gal-(1→3)-β-D-GalNAc with roughly the same affinity (0.386–0.726 μM). However, the K_D value towards α-D-Gal-(1→3)-α-D-GalNAc, the same linakge found in the conotoxin tx5a and egg jelly mucins mentioned earlier, was about an order of magnitude lower (0.0864 μM). These results agreed with the previous glycan array results that higher fluorescence levels were observed with α-D-Gal-(1→3)-α-D-GalNAc compared to α-D-Gal-(1→3)-β-D-GalNAc [22]. In addition, the anomericity of the disaccharide is important and both carbohydrate rings likely interact with XCGL-1. The glycan array also contained β-lactose and β-melibiose, but no binding was observed even though the K_D values were similar to α-D-Gal-(1→3)-β-D-GalNAc. We attributed this to the faster dissociation rates of both β-lactose and β-melibiose that were about one and two order of magnitude faster compared to α-D-Gal-(1→3)-β-D-GalNAc and α-D-Gal-(1→3)-α-D-GalNAc, respectively. Thus, XCGL-1 bound to β-lactose and β-melibiose may not survive glycan array washing.

The binding affinity of His_6 XCGL-1 CRD was about an order of magnitude weaker compared to His_6 XCGL-1 for β-lactose and β-melibiose. The affinity towards α-D-Gal-(1→3)-α-D-GalNAc and α-D-Gal-(1→3)-β-D-GalNAc were similar. Compared with His_6 XCGL-1, His_6 XCGL-1 CRD bound an order of magnitude and two fold weaker towards, α-D-Gal-(1→3)-α-D-GalNAc and α-D-Gal-(1→3)-β-D-GalNAc, respectively. From our TEM and DLS experiments, truncation of the His_6 XCGL-1 to His_6 XCGL-1 CRD likely reduced the tetrameric association of the CRD to merely the trimeric CRD. The mere order of magitude difference in K_D between His_6 XCGL-1 and His_6 XCGL-1 CRD suggested that the tetravalent display of the CRD did somewhat enhance the binding affinity, but not to the >100 fold extent observed with other lectins [29, 47].

The refolded XCGL-1 CRD bound to β-lactose poorly, compared to other ligands, with the K_D of 11.2 μM. Still, this was strong enough to observe binding and competitive elution from the lactose resin. The binding affinities of the refolded XCGL-1 CRD towards β-melibiose and α-D-Gal-(1→3)-β-D-GalNAc were similar to His_6 XCGL-1 CRD. The refolded XCGL-1 CRD bound α-D-Gal-(1→3)-α-D-GalNAc two times weaker compared to His_6 XCGL-1 CRD. The discrepancy between ligand binding afiinities between the *T. ni* cell-derived His_6 XCGL-1 CRD and the *E. coli*-derived XCGL-1 CRD, with the latter binding much poorer to lactose, suggested that the refolded and affinity purified XCGL-1 CRD did not fold as correctly as the *T. ni* cell-derived XCGL-1.

2.6. Hemagglutination activity of XCGL-1

Because XCGL-1 is specific towards galactosides, agglutination of pig red blood cells that contain galactose residues [48] was performed to determine whether XCGL-1 could recognize glycans in a cellular context. Indeed, all forms of XCGL-1 could agglutinate pig red blood cells at nM range (Figure 8A). This is consistent with the oligomeric nature of XCGL-1 and XCGL-1 CRD that could engage in multivalent binding. Thus, the protein could bind multiple red blood cells simultaneously, resulting in hemagglutination. Agglutination could be inhibited by EDTA and galactose, suggesting that the interaction was mediated by calcium ions and that binding was specific to galactose as expected. The larger oligomeric His_6 XCGL-1 could agglutinate red blood cells at only slighlty lower concentrations than the CRDs, suggesting that assembly of the CRDs into the 4-lobed structure did not significantly enhance the binding avidity towards the red blood cells. Overall, the minimum red blood cell agglutination concentration of XCGL-1 (1.56–6.25 nM) is lower than the K_D values obtained from the BLI experiments (86.4 nM–11.2 μM). This could be a result of the difference in surface ligand density. However, another possibility is that XCGL-1 could recognize galactoside ligands in glycans with higher affinity due to secondary interactions with neighboring carbohydrate residues.

Inhibition of hemagglutination by various carbohydrates were also examined (Figure 8B). As expected, glucose, fucose, mannose, N-acetylglucosamine, and N-acetylgalactosamine could not inhibit hemagglutination as they were not XCGL-1 ligands. The minimum inhibitory concentrations (MIC) of galactose towards His_6 XCGL-1 and His_6 XCGL-1 CRD were 15 and 10 mM, respectively. Lactose and melibiose were both able to inhibit His_6 XCGL-1 and His_6 XCGL-1 CRD at 15 and 10 mM, respectively. The α-D-Gal-(1→3)-D-GalNAc was a more potent inhibitor towards His_6 XCGL-1 and His_6 XCGL-1 CRD than other galactosides, thus a lower concentration range was explored (Figure 8C). The disaccharide was able to inhibit hemagglutination by His_6 XCGL-1 and His_6 XCGL-1 CRD at 5 and 2.5 mM, respectively. Overall, XCGL-1 ligands were able to inhibit His_6 XCGL-1 CRD at slightly low concentrations compared to His_6 XCGL-1, suggesting that His_6 XCGL-1 was a slightly better red blood cell binder. The results were consistent with earlier observations that XCGL-1 oligomerization beyond the putative trimeric CRD did not drastically improve multivalent binding toward red blood cells.

Galactose is not efficient at inhibiting hemagglutination of the refolded XCGL-1 CRD at 25 mM (Figure 8B). Inhibition by lactose

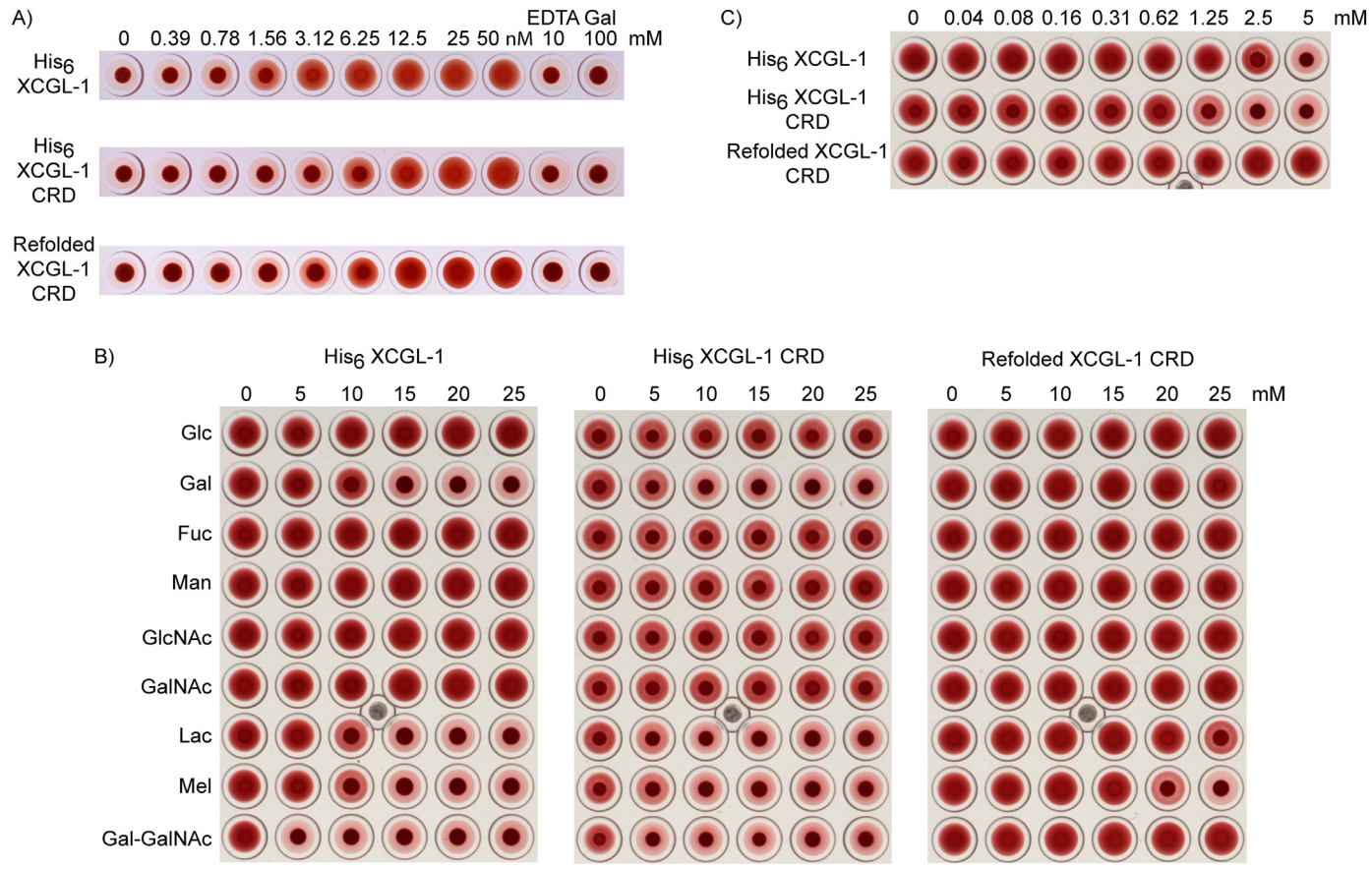


Figure 8. A) Hemagglutination activity of XCGL-1 variants. The number indicates XCGL-1 concentrations in nanomolar. EDTA and galactose was added as indicated at millimolar concentrations in the presence of 50 nM XCGL-1. B) Inhibition of hemagglutination by various carbohydrates in the presence of 25 nM XCGL-1. C) Inhibition of hemagglutination by α-D-Gal-(1→3)-D-GalNAc at lower concentrations than in B). Glc = D-glucose, Gal = D-galactose, Fuc = L-fucose, Man = D-mannose, GlcNAc = N-acetyl-D-glucosamine, GalNAc = N-acetyl-D-galactosamine, Lac = lactose (β-D-Gal-(1→4)-D-Glc), Mel = melibiose (α-D-Gal-(1→6)-D-Glc), Gal-GalNAc = α-D-Gal-(1→3)-D-GalNAc.

and melibiose was observed at >20 mM. In addition, α-D-Gal-(1→3)-D-GalNAc showed no inhibition at 25 mM. These results suggested that despite the ability to purify the refolded XCGL-1 CRD by lactose affinity chromatography, the protein did not have the same biochemical activity as the XCGL-1 produced from the *T. ni* insect cells. The results were also inconsistent with the previous glycan array results that XCGL-1 binds α-D-Gal-(1→3)-D-GalNAc well [22]. The poor inhibition results hinted that the refolded XCGL-1 CRD may interact with red blood cells through other mechanisms not involving the ligand binding site. For example, the protein may not fold entirely correctly and the exposed hydrophobic regions may bind non-specifically to red blood cells. The lack of glycosylation may play a role as well. Thus, the refolded XCGL-1 CRD may not be suitable for further investigations and applications.

2.7. Structural model of galactoside-XCGL-1 interactions

Because XCGL-1 binds galactosides with no exocyclic 1,2-diol, the ligand binding mechanism is expected to be different than that of hIntL-1 and XEEL. Given that the ligand binding was still calcium ion dependent, the ring hydroxyl groups of galactose likely interact with the protein-bound calcium ion in the ligand binding site. To investigate the potential ligand binding mechanism, we built a homology model of XCGL-1 CRD using a crystal structure of hIntL-1 (PDB ID 4WMY) [10] as the template. Methyl α-D-galactopyranoside, methyl β-D-galactopyranoside, α-D-Gal-(1→3)-α-methyl-D-GalNAc, and α-D-Gal-(1→3)-β-methyl-D-GalNAc were then docked into the ligand binding site. The docked structures were clustered according to the distances of the galactose ring to the calcium ion. A short molecular dynamics (MD) was performed on the structure of each complex to allow more structural movements and relaxation. The structures of the complexes are shown in Figure 9 and Table S4. The coordinate files are provided as supplementary data. For methyl α-D-galactopyranoside (Supplementary file methyl_a-galactoside.pdb) and methyl β-D-galactopyranoside (Supplementary file methyl_b-galactoside.pdb) (Table S4), few interactions were observed which explained why galactose was not as good of a ligand compared to galactose-containing disaccharides. There are two stable binding orientations of α-D-Gal-(1→3)-α-methyl-D-GalNAc, named orientation A (Supplementary file a-D-Gal-1-3-a-methyl-D-GalNAc_A.pdb) and B (Supplementary file a-D-Gal-1-3-a-methyl-D-GalNAc_B.pdb), respectively (Table S4). Orientation B (Figure 9A) had more extensive interactions, thus taken as a better solution. The 4-OH group and the 6-OH group of the galactose residue coordinated the calcium ion. The disaccharide bound in the groove formed between N260 and N297. The smaller size of F288 and N297, compared to the equivalent residues in XEEL and hIntL-1, likely allowed accommodation of the pyranose ring. N297 formed a hydrogen bond with the 2-OH group of galactose, as well as the 4-OH group of GalNAc. The 6-OH group of GalNAc could potentially form hydrogen bonds with S229, Q230, and the backbone carbonyl of D226. These additional interactions explained why α-D-Gal-(1→3)-α-methyl-D-GalNAc was a high affinity ligand. For α-D-Gal-(1→3)-β-methyl-D-GalNAc (Supplementary file a-D-Gal-1-3-b-methyl-D-GalNAc.pdb) (Figure 9B), the 2-OH group of the galactose residue does not form a hydrogen bond with N297 and GlcNAc cannot form a hydrogen bond with S229, while the 6-OH group of GalNAc could still interacts with Q230 and the backbone carbonyl of D226. These fewer interactions explained why α-D-Gal-(1→3)-β-D-GalNAc did not bind XCGL-1 as tightly as α-D-Gal-(1→3)-α-D-GalNAc.

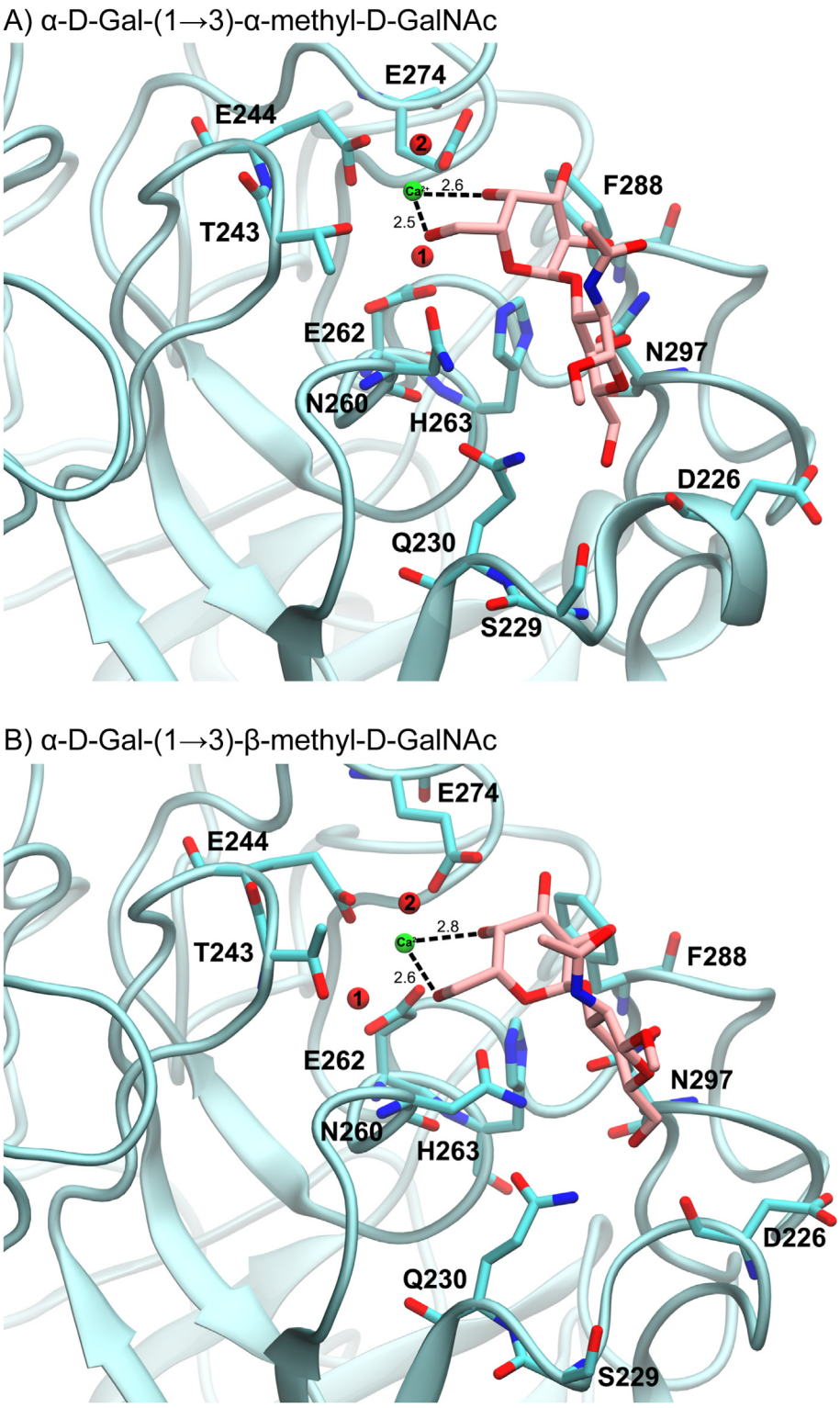


Figure 9. Model of galactosides bound in the ligand binding site of XCGL-1. The distances are in Å. A) α-D-Gal-(1→3)-α-methyl-D-GalNAc in orientation B. B) α-D-Gal-(1→3)-β-methyl-D-GalNAc.

3. Conclusion

Full length XCGL-1 was successfully expressed in HEK293T and *T. ni* cells. The disulfide-linked oligomers were observed as previously reported with XCGLs isolated from *Xenopus* eggs. We determined that both C18 and C35 were essential in the formation of disulfide-linked oligomers in XCGL-1. XCGL-1 CRD was successfully expressed as insoluble inclusion bodies in bacteria. Two refolding buffer that resulted in active XCGL-1 CRD are arginine buffer (50 mM Tris pH 10.0, 10 mM $CaCl_2$, 2 mM GSSG, 0.2 mM GSH, and 0.3 M L-arginine) and NDSB-201 buffer (50 mM Tris pH 10.0, 10 mM $CaCl_2$, 2 mM GSSG, 0.2 mM GSH, and 0.75 M NDSB-201). All forms of XCGL-1 recognized galactose-containing ligands and could agglutinate pig red blood cells at nM range in a calcium ion-dependent manner. However, the refolded XCGL-1 CRD did not bind

ligands as well as the *T. ni* cell-derived XCGL-1. Molecular modeling suggested that galactosides bind to XCGL-1 through calcium ion coordination with 4-OH and 6-OH, and additional interactions with the GalNAc ring makes α-D-Gal-(1→3)-D-GalNAc a potent ligand of XCGL-1. These results further advance the basic biochemical knowledge of XCGL-1 that will be essential for further structural and functional studies of XCGL-1, as well as applications of XCGL-1 as a carbohydrate recognition tool.

4. Materials and methods

4.1. Construction of XCGL-1 expression plasmids

The cDNA coding for full length XCGL-1 was purchased from Source BioScience (GenBank BC170087, MGC:196814, IMAGE:9041722). The open reading frame was then amplified using the forward primer 5′-GCTTGGTACCATGCTGGTGCACATTCTTCTCCTGC-3′ and the reverse primer 5′-GCTCGGATCCTCATTATAGATAGAAAAGTAATACAGCGGCCTCAGTTATCTC-3'. The gel-purified PCR product was subsequently digested and cloned into the KpnI and BamHI sites of the plasmid pcDNA4/myc-His A. The sequence of the gene was verified by DNA sequencing. This plasmid is designated as pcDNA4 XCGL-1. The plasmid was subsequently used for expression in HEK293T cells.

For XCGL-1 expression in insect cells, the cDNA template was amplified with the primer 5′-GCGCGGATCCATGCTGGTGCACATTCTTCTCCTGCTG-3′ and primer R (5′-CTACGTCGACTCATTATAGATAGAAAAGTAATACAGCGGCCTCAG-3′). The PCR product was then subsequently digested and cloned into the BamHI and SalI restriction sites of the plasmid pFastBac1. DNA sequencing was performed to verify the gene sequence. The resulting plasmid was designated as pFastBac1 XCGL-1. For the construction of the plasmid for His_6 XCGL-1 expression, the cDNA template was amplified with the primer CCAGCAGGGCACGCTGGTTCA-CATCACCACCATCACCACGGTAGCTGTGAACCTGTTGTAA-TAGTAGCCTCAAAAAAC and the primer R. The PCR product was re-amplified with the primer XSPF (GCGCGGATCCATGTTGTCATATAGCCTGTTGCTTTTTGCACTTGCATTTCCAGCAGGGCACGCTGGTTCA) and the primer R, and then digested and cloned into the BamHI and SalI restriction sites of the plasmid pFastBac1. The resulting plasmid was designated pFastBac His_6 XCGL-1. For the construction of His_6 XCGL-1 CRD expression plasmid, the cDNA template was amplified with the primer CCAGCAGGGCACGCTGGTTCACATCACCACCATCACCACGATAAATTCAGAAGCTGCAAGGAGATC and the primer R. The PCR product was re-amplified with the primer XSPF and the primer R, and then digested and cloned into the BamHI and SalI restriction sites of the plasmid pFastBac1. The resulting plasmid was designated pFastBac His_6 XCGL-1 CRD.

To construct a bacterial expression plasmid for XCGL-1 CRD, the XCGL-1 CRD cDNA corresponding to amino acid residue 36 to 313 was amplified using the forward primer 5′-TATACCATGGATAAATTCAGAAGCTGCAAGGAGATCAAAG-3′ and the primer R. The DNA fragment were then cloned into the NcoI and SalI sites of pET28a. The plasmid is designated pET28a XCGL-1 CRD. The amino acid sequences of the expression construct in pcDNA XCGL-1, pFastBac1 XCGL-1, pFastBac1 His_6 XCGL-1, pFastBac1 His_6 XCGL-1 CRD, and pET28a XCGL-1 CRD are summarized Table S5.

4.2. Protein expression in HEK293T cells

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM). Cells were seeded into a 6-well plate at the density of 1×10 [6] cells per well and cultured overnight before transfection. For each transfection reaction, 2 μg of plasmid was combined with 4 μg of the linear polyethylenimine (PEI, molecular weight 25,000) (Polysciences) in 200 μL of Opti-MEM I. After the incubation period of 15 min, Opti-MEM I (800 μL) was then added. The transfection mixture was subsequently applied to aspirated cells. The cells were exposed to the transfection mixture for 4 h at 37 °C. SFM4HEK293 (GE Healthcare) (1.5 mL) was added. Three days were allowed for protein expression as secreted protein. The conditioned media was then analyzed by western blotting using the pan-intelectin antibody (Proteintech 11770-1-AP) as the primary antibody [14, 32].

4.3. Protein expression in insect cells and purification

Recombinant baculovirus construction from pFastBac1 vectors in *Escherichia coli* DH10Bac and Sf21 insect cells, and protein expression from *Trichoplusia ni* insect cells, were performed exactly as previously described [14, 32]. For His_6-tagged protein purification, the conditioned culture media were dialyzed against the 20 mM HEPES pH 7.5, 150 mM NaCl, and 25 mM imidazole buffer. The solution was then applied onto a Ni-NTA column and eluted with an imidazole gradient up to 250 mM. For protein purification with a lactose resin, the spent culture media was dialyzed into the binding buffer (20 mM HEPES pH 7.5, 150 mM NaCl, and 10 mM $CaCl_2$), and then applied to a lactose-agarose column [49]. The protein was eluted from the column by the binding buffer that contained 100 mM lactose.

For Edman degradation to determine the signal peptide cleavage site, XCGL-1 expressed from *T. ni* cells was purified using a lactose resin as described above. The protein was then separated by SDS-PAGE and blotting onto a polyvinylidene fluoride (PVDF) membrane. The Coomassie Blue-stained band was then sent for 5 cycles of Edman degradation at the Tufts University Core Facility (Table S1).

For the enzymatic deglycosylation experiments, 0.1 μM of His_6 XCGL-1 and His_6 XCGL-1 CRD were treated with Endo H and PNGase F (New England Biolabs) under denaturing conditions following the manufacturer directions. The deglycosylation reactions were analyzed by western blotting with an anti-His_6 tag antibody as the primary antibody. RNase B was used as a control to verify the PNGase F and Endo H activities.

4.4. Site-directed mutagenesis of XCGL-1

Site-directed mutagenesis of the pFastBac1 XCGL-1 and pcDNA4 XCGL-1 plasmids were performed following the protocol of the QuikChange II Site-Directed Mutagenesis kit (Agilent Technologies). However, the Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific) was used for mutagenesis. The C18A mutation was made using the primers 5′-TGACTGGTGGGCTCTCTCAGTCTGCTGAACCTGTTGT-3′ and 5′-ACAA-CAGGTTCAGCAGACTGAGAGAGCCCACCAGTCA-3'. The C35A mutation was made using the primers 5′-CAAAAAA-CATGGTGAAGCAGCTGGATGCTGATAAATTCAGAAGCTG-3′ and 5′-CAGCTTCTGAATTTATCAGCATCCAGCTGCTTCACCATGTTTTTTG-3'. The C18A and C35A double mutation was made by performing a C35A mutagenesis on the previously mutated C18A plasmid. The mutations were verified by DNA sequencing. Western blot was performed using a pan-intelectin antibody as described above.

4.5. Expression and purification of XCGL-1 CRD in Escherichia coli

The pET28a XCGL-1 CRD plasmid was transformed into *E. coli* Tuner (DE3). The bacteria were grown at 37 °C in LB to OD600 = 0.6, then induced with 0.1 mM IPTG. Cells were then pelleted after 6 h and stored at -80 °C until use. To isolate the inclusion bodies, the frozen pellet was thawed and resuspended in 100 mM Tris pH 8.0, 10 mM EDTA. Lysozyme was then added, and the mixture sonicated. The inclusion bodies were collected by centrifugation and then washed with the resuspension buffer containing 1% Triton X-100. The washed inclusion bodies were washed the second time with the resuspension buffer containing 1 M NaCl. The washed inclusion bodies were then stored at -20 °C until use.

4.6. XCGL-1 CRD refolding assay

The washed inclusion bodies were dissolved in 20 mM Tris pH 8.0, 6 M urea, 10 mM DTT. The concentration was measured by UV absorbance at 280 nm. The protein solution was aliquoted and stored at -20 °C until use. For the refolding assay, the protein sample was diluted to 0.1 mg/mL in various buffer (1 mL) according to Table S2. The refolding buffer consisted of Tris buffer at either pH 8.0 or 10.0, $CaCl_2$, one type of additive, or one ratio of reduced and oxidized glutathione. Conditions that did not result in protein precipitation were selected for further examination. For XCGL-1 CRD purification, the refolded protein solution was dialyzed against 20 mM HEPES pH 7.5, 150 mM NaCl, and 10 mM $CaCl_2$, and purified with a lactose resin as described above.

4.7. Negative staining transmission electron microscopy and dynamic light scattering

A His_6 XCGL-1 sample (30 μg/mL) was applied onto a carbon-coated grids for 10 min. The grids were washed with water and stained for 2 min with 1% uranyl acetate (EMS). After air-drying overnight, the grids were imaged with a transmission electron microscope (JEM1230, JEOL, JAPAN) at 100 kV. Dynamic light scattering data were collected on Malvern Zetasizer Nano ZS at the protein concentration of 0.3 mg/mL.

4.8. XCGL-1 carbohydrate binding assay

Carbohydrate resin was generated by activation of agarose resin by divinyl sulfone and subsequent reaction with a carbohydrate solution [49]. Protein solution was dialyzed against 20 mM HEPES pH 7.5, 150 mM NaCl, and 10 mM $CaCl_2$ (binding buffer) then applied to the resin that had been equilibrated with the same buffer. The resin was then washed with the binding buffer then the bound protein was eluted with 20 mM HEPES pH 7.5, 150 mM NaCl, and 10 mM EDTA. The eluted protein was then analyzed by western blotting using a pan-intelectin antibody (refolded XCGL-1 CRD) or an anti-His tag antibody (Sigma-Aldrich SAB4200620) (His_6 XCGL-1 and His_6 XCGL-1 CRD).

For the competitive elution assay, 250 μL of lactose-agarose resin was settled in a spin column by centrifugation at 100 x g for 10 s. The resin was washed three times with the binding buffer before applying 250 μL (0.1 μM) of protein. The protein was left to bind the resin for 5 min before washing for 3 times with the binding buffer. The protein was then eluted by various carbohydrates at 6.5 mM in the binding buffer and analyzed by western blotting.

4.9. Biolayer interferometry

Bio-layer interferometry (BLI, BLItz, ForteBio) was used to determine the binding dissociation constant (K_D) values of XCGL-1 variants to carbohydrate ligands. The ligands used were biotinylated glycosides (GlycoNZ, New Zealand): β-lactose (0046a-BM), β-melibiose (0069-BM), α-D-Gal-(1→3)-α-D-GalNAc (0053-BM), and α-D-Gal-(1→3)-β-D-GalNAc (0038-BM). For ligand immobilization, the ligands were dissolved in the BLI buffer (20 mM HEPES, 150 mM NaCl, 10 mM $CaCl_2$, 0.005% Tween-20) at 300 nM. For each ligand, the streptavidin probe (18–5019, ForteBio) was dipped into the ligand solution for 5 min. After washing, a baseline measurement was obtained by dipping the probe in the BLI buffer for 1 min. Association phase measurement was performed by dipping the probe into 4 μL of the protein solution for 3 min. Dissociation phase measurement was recorded by dipping the probe into 250 μL of the BLI buffer for 5 min. The entire process was performed at 25 °C and 2,200 RPM shaking speed. Between each measurement, the probe was regenerated three times by dipping the probe into 250 μL of the regeneration buffer (20 mM HEPES, 150 mM NaCl, 10 mM EDTA, 0.005% Tween-20) for 5 s and then 250 μL of the BLI buffer for 5 s. The probe was reused for another concentration of the same protein. Data were analyzed using the BLItz Pro software. The data was fitted globally by generating the same kinetic constants for the same protein at various concentrations.

4.10. Hemagglutination assay

Fresh pig blood was purchased from a local slaughterhouse and immediately added to medical-grade blood collection tubes containing sodium citrate. The blood was then centrifuged and the red blood cells were washed in 20 mM HEPES pH 7.5, 150 mM NaCl, and 10 mM $CaCl_2$. The washed red blood cells (2.5% suspension) were then mixed with XCGL-1 variants at different concentrations in a 72-well agglutination plate (Terasaki plate). The plate was photographed after 1 h of incubation. For inhibition of hemagglutination by various carbohydrates, red blood cells were incubated with 25 nM of XCGL-1 in the presence of carbohydrate at various concentrations.

4.11. Model construction of the methyl α/β-galactoside-XCGL-1 complex and α-D-Gal-(1→3)-α/β-methyl-D-GalNAc-XCGL-1 complex

The structures of α- and β-galactose were obtained from the crystal structure of *Pseudomonas aeruginosa* lectin I (PA-IL) (PDB ID: 1OKO), and they were modified to build the structures of methyl α/β-galactoside using LEaP module in AMBER18 [50] with the force field parameters of GLYCAM06j-1 [51]. The structures of α-D-Gal-(1→3)-α/β-methyl-D-GalNAc were also constructed using LEaP module. Each structure was solvated in a truncated octahedral isomeric box of TIP3P water molecules. Each structure was minimized by 2,500 steps of steepest descent and 2,500 steps of conjugate gradient. SWISS-MODEL server [52, 53, 54, 55] was employed to construct the homology model of XCGL-1 from *Xenopus* egg using the human intelectin-1 structure that contains β-galactofuranose (PDB ID: 4WMY [10]) as the template. The quality of the constructed homology model was evaluated by the Ramachandran plot (Figure S5), the majority of amino acid residues (94.58%) are in the favored region and allowed region (0.36%), suggesting adequate qualities of this homology model. H^{++} server [56] was used to protonate all ionized amino acids at the experimental pH of 7.5. Vina-Carb [57] was employed to predict the binding orientations of methyl α/β-galactoside/XCGL-1 complex and α-D-Gal-(1→3)-α/β-methyl-D-GalNAc/XCGL-1 complex, and 20 independent docking runs were performed to dock the structure of each substrate into the binding site of XCGL-1. For each ligand, the total of 180 binding orientations were obtained. The reasonable binding orientations with the distances from the closest 4-OH and 6-OH groups of the galactose ring to the calcium atom less than or equal to 5.0 Å were selected, and they were later clustered into groups based on their structural similarities as measured by their RMSD values of heavy atoms. The binding orientation that is most similar to the average structure of all members of each group was selected to be a centroid/representative binding orientation, and short MD was performed to identify the best binding orientation for each complex that has the distances from the closest 4-OH and 6-OH groups of the galactose ring to the calcium atom less than or equal to 5.0 Å during and after short MD.

LEaP module in AMBER18 with GLYCAM06j-1 force field and ff14SB parameters [58] was used to immerse the representative binding orientations of all systems in isomeric truncated octahedral boxes of TIP3P water molecules with the 13 Å buffer distance. The five-step minimization procedure [58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73] was employed to reduce the unfavorable interaction of the system. In this study, all steps used 2,500 steps of steepest descent and 2,500 steps of conjugate gradient. Initially, a force constant of 10 kcal/(mol.$Å^2$) was applied to restrain the heavy atoms of protein, while minimizing water molecules and hydrogen atoms. Subsequently, the backbone of the protein was restrained with the force constants of 10, 5 and 1 kcal/(mol $Å^2$), respectively. Finally, the entire system was minimized without any restraining force. Short MD was performed on each system to refine and allow structural relaxation to a more favorable orientation as well as to optimize and correct the first binding mode and minimize local steric clashes between protein and ligand. The GPU (CUDA) version of PMEMD

module [74, 75, 76] was used to simulate this system with the periodic boundary condition. All bonds with hydrogen atoms were constrained using the SHAKE algorithm [77], allowing the time step of 0.002 ps. The Langevin dynamics method [78] was employed to control the temperature with a 1.0 per ps collision frequency. The system was heated from 0 K to the experimental temperature of 298 K for 200 ns in the NVT ensemble, as the protein backbones restrained at a 10 kcal/(mol Å^2) force constant. The system in the NVT ensemble was then equilibrated for 300 ps without any restraint. Lastly, the system was simulated for 10 ns in the NPT ensemble at 298 K and 1 atm. For analyses, the root mean square deviation (RMSD) values were calculated to assess the stability of the system. As shown in Figure S6, all systems were found to be quite stable during the simulation period. The interatomic distances required for calcium ion coordination were measured and shown in Figure S7 and Table S6. Furthermore, the centroid, which is the structure the most similar to the last 5 ns average structure of the MDs trajectory was minimized and used to be a representative structure (Table S4). The best binding orientation of each complex was selected from the binding orientation that has the distances from the closest 4-OH and 6-OH groups of the galactose ring to the calcium atom less than or equal to 5.0 Å during and after short MD. The best binding orientations are group I for methyl α-galactoside-XCGL-1 complex, group I for methyl β-galactoside-XCGL-1 complex, group II and III for α-D-Gal-(1→3)-α-methyl-D-GalNAc-XCGL-1 complex (named orientation A and B, respectively) and group I for α-D-Gal-(1→3)-α/β-methyl-D-GalNAc-XCGL-1 complex.

Declarations

Author contribution statement

Peerapon Deetanya: conceived and designed the experiments; performed the experiments; analyzed and interpreted the data.

Thassanai Sitthiyotha and Surasak Chunsriviro: performed the experiments; analyzed and interpreted the data.

Nusara Chomanee: performed the experiments.

Kittikhun Wangkanont: conceived and designed the experiments; performed the experiments; analyzed and interpreted the data; wrote the paper.

Funding statement

This work was supported by Institute for the Promotion of Teaching Science and Technology (IPST) under the Research Fund for DPST Graduate with First Placement (08/2559), Grants for Development of New Faculty Staff, Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University (GDNS 59-059-23-020 and DNS 61-011-23-003-2), Chulalongkorn University grant to the Center of Excellence for Molecular Biology and Genomics of Shrimp, Chulalongkorn University grant to the Center of Excellence for Molecular Crop, 90th Anniversary of Chulalongkorn University Scholarship, Chulalongkorn University grant to the Structural and Computational Biology Research Unit and Science Achievement Scholarship of Thailand (SAST).

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of interest's statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e10396.

Acknowledgements

The authors would like to thank DKSH (Thailand) Limited for access to the dynamic light scattering (DLS) equipment. We thank Khomkrit Sappakhaw and Pansilp Aungkasukho for assistance in DLS data collection. The authors would like to thank Rushmore Precision Co., Ltd. for access to the BLItz equipment and Supattara Suwanpairoj for assistance in data collection.

References

[1] J. Yan, L. Xu, Y. Zhang, C. Zhang, C. Zhang, F. Zhao, L. Feng, Comparative genomic and phylogenetic analyses of the intelectin gene family: implications for their origin and evolution, Dev. Comp. Immunol. 41 (2013) 189–199.

[2] B. Lin, Z. Cao, P. Su, H. Zhang, M. Li, Y. Lin, D. Zhao, Y. Shen, C. Jing, S. Chen, A. Xu, Characterization and comparative analyses of zebrafish intelectins: highly conserved sequences, diversified structures and functions, Fish Shellfish Immunol. 26 (2009) 396–405.

[3] A.D. Pemberton, P.A. Knight, J. Gamble, W.H. Colledge, J.K. Lee, M. Pierce, H.R. Miller, Innate BALB/c enteric epithelial responses to Trichinella spiralis: inducible expression of a novel goblet cell lectin, intelectin-2, and its natural deletion in C57BL/10 mice, J. Immunol. 173 (2004) 1894–1901.

[4] S. Russell, M.A. Hayes, J.S. Lumsden, Immunohistochemical localization of rainbow trout ladderlectin and intelectin in healthy and infected rainbow trout (Oncorhynchus mykiss), Fish Shellfish Immunol. 26 (2009) 154–163.

[5] S. Russell, K.M. Young, M. Smith, M.A. Hayes, J.S. Lumsden, Identification, cloning and tissue localization of a rainbow trout (Oncorhynchus mykiss) intelectin-like protein that binds bacteria and chitin, Fish Shellfish Immunol. 25 (2008) 91–105.

[6] R. Datta, M.L. deSchoolmeester, C. Hedeler, N.W. Paton, A.M. Brass, K.J. Else, Identification of novel genes in intestinal tissue that are regulated after infection with an intestinal nematode parasite, Infect. Immun. 73 (2005) 4025–4033.

[7] A.T. French, P.A. Knight, W.D. Smith, J.K. Brown, N.M. Craig, J.A. Pate, H.R. Miller, A.D. Pemberton, Up-regulation of intelectin in sheep after infection with Teladorsagia circumcincta, Int. J. Parasitol. 38 (2008) 467–475.

[8] Y. Abe, M. Tokuda, R. Ishimoto, K. Azumi, H. Yokosawa, A unique primary structure, cDNA cloning and function of a galactose-specific lectin from ascidian plasma,, Eur. J. Biochem. 261 (1999) 33–39.

[9] Z. Ding, X. Zhao, J. Wang, F. Zhang, W. Wang, H. Liu, Intelectin mediated phagocytosis and killing activity of macrophages in blunt snout bream (Megalobrama amblycephala), Fish Shellfish Immunol. 87 (2019) 129–135.

[10] D.A. Wesener, K. Wangkanont, R. McBride, X. Song, M.B. Kraft, H.L. Hodges, L.C. Zarling, R.A. Splain, D.F. Smith, R.D. Cummings, J.C. Paulson, K.T. Forest, L.L. Kiessling, Recognition of microbial glycans by human intelectin-1, Nat. Struct. Mol. Biol. 22 (2015) 603–610.

[11] S. Tsuji, J. Uehori, M. Matsumoto, Y. Suzuki, A. Matsuhisa, K. Toyoshima, T. Seya, Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall, J. Biol. Chem. 276 (2001) 23456–23463.

[12] S. Nagata, Isolation, characterization, and extra-embryonic secretion of the Xenopus laevis embryonic epidermal lectin, XEEL, Glycobiology 15 (2005) 281–290.

[13] S. Nagata, M. Nakanishi, R. Nanba, N. Fujita, Developmental expression of XEEL, a novel molecule of the Xenopus oocyte cortical granule lectin family, Dev. Gene. Evol. 213 (2003) 368–370.

[14] K. Wangkanont, D.A. Wesener, J.A. Vidani, L.L. Kiessling, K.T. Forest, Structures of Xenopus embryonic epidermal lectin reveal a conserved mechanism of microbial glycan recognition, J. Biol. Chem. 291 (2016) 5596–5610.

[15] M.M. Roberson, S.H. Barondes, Lectin from embryos and oocytes of Xenopus laevis. Purification and properties, J. Biol. Chem. 257 (1982) 7520–7524.

[16] H. Shoji, K. Ikenaka, S. Nakakita, K. Hayama, J. Hirabayashi, Y. Arata, K. Kasai, N. Nishi, T. Nakamura, Xenopus galectin-VIIa binds N-glycans of members of the cortical granule lectin family (xCGL and xCGL2), Glycobiology 15 (2005) 709–720.

[17] T. Ishino, T. Kunieda, S. Natori, K. Sekimizu, T. Kubo, Identification of novel members of the Xenopus Ca2+ -dependent lectin family and analysis of their gene expression during tail regeneration and development, J. Biochem. 141 (2007) 479–488.

[18] S. Nagata, S. Nishiyama, Y. Ikazaki, Bacterial lipopolysaccharides stimulate production of XCL1, a calcium-dependent lipopolysaccharide-binding serum lectin, in Xenopus laevis, Dev. Comp. Immunol. 40 (2013) 94–102.

[19] S. Nagata, Identification and characterization of a novel intelectin in the digestive tract of Xenopus laevis, Dev. Comp. Immunol. 59 (2016) 229–239.

[20] J.K. Lee, P. Buckhaults, C. Wilkes, M. Teilhet, M.L. King, K.W. Moremen, M. Pierce, Cloning and expression of a Xenopus laevis oocyte lectin and characterization of its mRNA levels during early development, Glycobiology 7 (1997) 367–372.

[21] T. Nishihara, R.E. Wyrick, P.K. Working, Y.H. Chen, J.L. Hedrick, Isolation and characterization of a lectin from the cortical granules of Xenopus laevis eggs, Biochemistry 25 (1986) 6013–6020.

[22] Y.C. Lee, M. Pierce, X-lectins: a new family with homology to the Xenopus laevis oocyte lectin XL35, in: G.R. Vasta, H. Ahmed (Eds.), Animal Lectins: A Functional View, first ed., CRC Press, 2008 ed.

[23] S. Sharma, T.N.C. Ramya, Saccharide binding by intelectins, Int. J. Biol. Macromol. 108 (2018) 1010–1016.

[24] C.M. McMahon, C.R. Isabella, I.W. Windsor, P. Kosma, R.T. Raines, L.L. Kiessling, Stereoelectronic effects impact glycan recognition, J. Am. Chem. Soc. 142 (2020) 2386–2395.

[25] R.L. Outenreath, M.M. Roberson, S.H. Barondes, Endogenous lectin secretion into the extracellular matrix of early embryos of Xenopus laevis, Dev. Biol. 125 (1988) 187–194.
[26] S.M. Chamow, J.L. Hedrick, Subunit structure of a cortical granule lectin involved in the block to polyspermy in Xenopus laevis eggs, FEBS Lett. 206 (1986) 353–357.
[27] T.K. Dam, C.F. Brewer, Multivalent lectin-carbohydrate interactions energetics and mechanisms of binding, Adv. Carbohydr. Chem. Biochem. 63 (2010) 139–164.
[28] L.L. Kiessling, Chemistry-driven glycoscience, Bioorg. Med. Chem. 26 (2018) 5229–5238.
[29] R.T. Lee, Y.C. Lee, Affinity enhancement by multivalent lectin-carbohydrate interaction, Glycoconj. J. 17 (2000) 543–551.
[30] R.A. Jue, J.E. Hale, Identification of cysteine residues alkylated with 3-bromopropylamine by protein sequence analysis, Anal. Biochem. 210 (1993) 39–44.
[31] N. Singrang, S. Laophetsakunchai, B.N. Tran, P.T. Matsudaira, A. Tassanakajon, K. Wangkanont, Biochemical and structural characterization of a recombinant fibrinogen-related lectin from Penaeus monodon, Sci. Rep. 11 (2021) 2934.
[32] N. Singrang, T. Sitthiyotha, N. Chomanee, C. Watthanasak, S. Chunsrivirot, K. Wangkanont, Molecular properties and ligand specificity of zebrafish intelectin-2, Fish Shellfish Immunol. 123 (2022) 528–536.
[33] T. Fukamizo, H. Sato, M. Mizuhara, T. Ohnuma, T. Gotoh, K. Hiwatashi, S. Takahashi, Chitinase from Autographa californica multiple nucleopolyhedrovirus: rapid purification from Sf-9 medium and mode of action, Biosci. Biotechnol. Biochem. 75 (2011) 1763–1769.
[34] T. Matsui, S. Kamata, K. Ishii, T. Maruno, N. Ghanem, S. Uchiyama, K. Kato, A. Suzuki, N. Oda-Ueda, T. Ogawa, Y. Tanaka, SDS-induced oligomerization of Lys49-phospholipase A(2) from snake venom, Sci. Rep. 9 (2019) 2330.
[35] H. van Halbeek, A.M. Strang, M. Lhermitte, H. Rahmoune, G. Lamblin, P. Roussel, Structures of monosialyl oligosaccharides isolated from the respiratory mucins of a non-secretor (O, Lea+b-) patient suffering from chronic bronchitis. Characterization of a novel type of mucin carbohydrate core structure, Glycobiology 4 (1994) 203–219.
[36] J. Taylor-Papadimitriou, J. Burchell, D.W. Miles, M. Dalziel, MUC1 and cancer, Biochim. Biophys. Acta 1455 (1999) 301–313.
[37] Y. Guerardel, O. Kol, E. Maes, T. Lefebvre, B. Boilly, M. Davril, G. Strecker, O-glycan variability of egg-jelly mucins from Xenopus laevis: characterization of four phenotypes that differ by the terminal glycosylation of their mucins, Biochem J 352 Pt 2 (2000) 449–463.
[38] G. Strecker, J.M. Wieruszeski, Y. Plancke, B. Boilly, Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran Xenopus laevis by reductive beta-elimination, Glycobiology 5 (1995) 137–146.
[39] D.J. Gill, H. Clausen, F. Bard, Location, location, location: new insights into O-GalNAc protein glycosylation, Trends Cell Biol. 21 (2011) 149–158.
[40] W.R. Alley Jr., B.F. Mann, M.V. Novotny, High-sensitivity analytical approaches for the structural characterization of glycoproteins, Chem. Rev. 113 (2013) 2668–2732.
[41] J. Kang, W. Low, T. Norberg, J. Meisenhelder, K. Hansson, J. Stenflo, G.P. Zhou, J. Imperial, B.M. Olivera, A.C. Rigby, A.G. Craig, Total chemical synthesis and NMR characterization of the glycopeptide tx5a, a heavily post-translationally modified conotoxin, reveals that the glycan structure is alpha-D-Gal-(1–>3)-alpha-D-GalNAc, Eur. J. Biochem. 271 (2004) 4939–4949.
[42] Y. Guérardel, D. Petit, T. Madigou, B. Guillet, E. Maes, A. Maftah, D. Boujard, G. Strecker, O. Kol, Identification of the blood group Lewis(a) determinant in the oviducal mucins of Xenopus tropicalis, FEBS Lett. 554 (2003) 330–336.
[43] W. Morelle, G. Strecker, Structural analysis of oligosaccharide-alditols released by reductive beta-elimination from oviducal mucins of Bufo bufo: characterization of the carbohydrate sequence Gal(alpha1-3)GalNAc(alpha1-3)[Fuc(alpha1-2)]Gal, Glycobiology 7 (1997) 777–790.
[44] A.S. Shashkov, M.S. Zaretskaya, S.V. Yarotsky, I.B. Naumova, O.S. Chizhov, Z.A. Shabarova, On the structure of the teichoic acid from the cell wall of Streptomyces antibioticus 39. Localization of the phosphodiester linkages and elucidation of the monomeric units structure by means of 13C-nuclear-magnetic-resonance spectroscopy, Eur. J. Biochem. 102 (1979) 477–481.
[45] B. Lindberg, F. Lindh, J. Lönngren, Structural studies of the O-specific side-chain of the lipopolysaccharide from Escherichia coli O 55, Carbohydr. Res. 97 (1981) 105–112.
[46] P. Sengupta, T. Bhattacharyya, A.S. Shashkov, H. Kochanowski, S. Basu, Structure of the O-specific side chain of the Escherichia coli O128 lipopolysaccharide, Carbohydr. Res. 277 (1995) 283–290.
[47] G.A. Snyder, J. Ford, P. Torabi-Parizi, J.A. Arthos, P. Schuck, M. Colonna, P.D. Sun, Characterization of DC-SIGN/R interaction with human immunodeficiency virus type 1 gp120 and ICAM molecules favors the receptor's role as an antigen-capturing rather than an adhesion receptor, J. Virol. 79 (2005) 4589–4598.
[48] Y. Tan, F. Gong, S. Li, S. Ji, Y. Lu, H. Gao, H. Xu, Y. Zhang, Brief report: a new profile of terminal N-acetyllactosamines glycans on pig red blood cells and different expression of alpha-galactose on Sika deer red blood cells and nucleated cells, Glycoconj. J. 27 (2010) 427–433.
[49] N. Fornstedt, J. Porath, Characterization studies on a new lectin found in seeds of Vicia ervilia, FEBS Lett. 57 (1975) 187–191.
[50] D. Case, I. Ben-Shalom, S. Brozell, D. Cerutti, T. Cheatham III, V. Cruzeiro, T. Darden, R. Duke, D. Ghoreishi, M. Gilson, University of California, san Francisco, AMBER 2018 (2018) 2018.
[51] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. González-Outeiriño, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (2008) 622–655.
[52] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling,, Bioinformatics 22 (2006) 195–201.
[53] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni, L. Bordoli, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res. 42 (2014) W252–W258.
[54] N. Guex, M.C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective,, Electrophoresis 30 (2009) S162–S173.
[55] F. Kiefer, K. Arnold, M. Künzli, L. Bordoli, T. Schwede, The SWISS-MODEL Repository and associated resources,, Nucleic Acids Res. 37 (2009) D387–D392.
[56] J.C. Gordon, J.B. Myers, T. Folta, V. Shoja, L.S. Heath, A. Onufriev, H++: a server for estimating p K as and adding missing hydrogens to macromolecules, Nucleic Acids Res. 33 (2005) W368–W371.
[57] A.K. Nivedha, D.F. Thieker, S. Makeneni, H. Hu, R.J. Woods, Vina-Carb: improving glycosidic angles during carbohydrate docking, J. Chem. Theor. Comput. 12 (2016) 892–901.
[58] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theor. Comput. 11 (2015) 3696–3713.
[59] T. Sitthiyotha, R. Pichyangkura, S. Chunsrivirot, Molecular dynamics provides insight into how N251A and N251Y mutations in the active site of Bacillus licheniformis RN-01 levansucrase disrupt production of long-chain levan, PLoS One 13 (2018), e0204915.
[60] P. Punnatin, C. Chanchao, S. Chunsrivirot, Molecular dynamics reveals insight into how N226P and H227Y mutations affect maltose binding in the active site of α-glucosidase II from European honeybee, Apis mellifera, PLoS One 15 (2020), e0229734.
[61] M. Klaewkla, R. Pichyangkura, T. Charoenwongpaiboon, K. Wangpaiboon, S. Chunsrivirot, Computational design of oligosaccharide producing levansucrase from Bacillus licheniformis RN-01 to improve its thermostability for production of levan-type fructooligosaccharides from sucrose, Int. J. Biol. Macromol. 160 (2020) 252–263.
[62] W. Mokmak, S. Chunsrivirot, A. Assawamakin, K. Choowongkomon, S. Tongsima, Molecular dynamics simulations reveal structural instability of human trypsin inhibitor upon D50E and Y54H mutations, J. Mol. Model. 19 (2013) 521–528.
[63] W. Mokmak, S. Chunsrivirot, S. Hannongbua, Y. Yuthavong, S. Tongsima, S. Kamchonwongpaisan, Molecular dynamics of interactions between rigid and Flexible antifolates and dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine-resistant plasmodium falciparum, Chem. Biol. Drug Des. 84 (2014) 450–461.
[64] T. Sitthiyotha, S. Chunsrivirot, Computational design of 25-mer peptide binders of SARS-CoV-2, J. Phys. Chem. B 124 (2020) 10930–10942.
[65] J. Manissorn, T. Sitthiyotha, J.R.E. Montalban, S. Chunsrivirot, P. Thongnuek, K. Wangkanont, Biochemical and structural investigation of GnnA in the lipopolysaccharide biosynthesis pathway of acidithiobacillus ferrooxidans, ACS Chem. Biol. 15 (2020) 3235–3243.
[66] T. Sitthiyotha, S. Chunsrivirot, Computational design of SARS-CoV-2 peptide binders with better predicted binding affinities than human ACE2 receptor, Sci. Rep. 11 (2021) 1–14.
[67] K. Wangpaiboon, T. Sitthiyotha, S. Chunsrivirot, T. Charoenwongpaiboon, R. Pichyangkura, Unravelling regioselectivity of leuconostoc citreum ABK-1 alternansucrase by acceptor site engineering, Int. J. Mol. Sci. 22 (2021) 3229.
[68] Na Ayutthaya, P. P, C. Chanchao, S. Chunsrivirot, Insight into the substrate specificity change caused by the Y227H mutation of α-glucosidase III from the European honeybee (Apis mellifera) through molecular dynamics simulations, PLoS One 13 (2018), e0198484.
[69] P. Kanjanatanin, R. Pichyangkura, T. Sitthiyotha, T. Charoenwongpaiboon, K. Wangpaiboon, S. Chunsrivirot, Computational design of Bacillus licheniformis RN-01 levansucrase for control of the chain length of levan-type fructooligosaccharides, Int. J. Biol. Macromol. 140 (2019) 1239–1248.
[70] T. Charoenwongpaiboon, M. Klaewkla, S. Chunsrivirot, K. Wangpaiboon, R. Pichyangkura, R.A. Field, M.H. Prousoontorn, Rational re-design of Lactobacillus reuteri 121 inulosucrase for product chain length control, RSC Adv. 9 (2019) 14957–14965.
[71] T. Charoenwongpaiboon, T. Sitthiyotha, P.P.N. Ayutthaya, K. Wangpaiboon, S. Chunsrivirot, M.H. Prousoontorn, R. Pichyangkura, Modulation of fructooligosaccharide chain length and insight into the product binding motif of Lactobacillus reuteri 121 inulosucrase, Carbohydr. Polym. 209 (2019) 111–121.
[72] T. Charoenwongpaiboon, P. Punnatin, M. Klaewkla, P. Pramoj Na Ayutthaya, K. Wangpaiboon, S. Chunsrivirot, R.A. Field, R. Pichyangkura, Conserved calcium-binding residues at the Ca-I site involved in fructooligosaccharide synthesis by lactobacillus reuteri 121 inulosucrase, ACS Omega 5 (2020) 28001–28011.
[73] M. Klaewkla, R. Pichyangkura, S. Chunsrivirot, Computational design of oligosaccharide-producing levansucrase from Bacillus licheniformis RN-01 to increase its stability at high temperature, J. Phys. Chem. (2021).
[74] A.W. Gotz, M.J. Williamson, D. Xu, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born, J. Chem. Theor. Comput. 8 (2012) 1542–1555.
[75] S. Le Grand, A.W. Götz, R.C. Walker, SPFP: speed without compromise—a mixed precision model for GPU accelerated molecular dynamics simulations, Comput. Phys. Commun. 184 (2013) 374–380.
[76] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald, J. Chem. Theor. Comput. 9 (2013) 3878–3888.
[77] D.M. York, T.A. Darden, L.G. Pedersen, The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods,, J. Chem. Phys. 99 (1993) 8345–8348.
[78] X. Wu, B.R. Brooks, Self-guided Langevin dynamics simulation method, Chem. Phys. Lett. 381 (2003) 512–518.