BBA Advances 7 (2025) 100145

Contents lists available at ScienceDirect

BBA Advances

journal homepage: www.sciencedirect.com/journal/bba-advances

The future of plant lectinology: Advanced technologies and computational tools

Vinicius J.S. Osterne [a,b], Kyria S. Nascimento [b], Benildo S. Cavada [b], Els J.M. Van Damme [a,*]

[a] Laboratory of Biochemistry and Glycobiology, Department of Biotechnology, Faculty of Bioscience Engineering, Ghent University, Proeftuinstraat 86, 9000 Ghent, Belgium
[b] BioMol-Lab, Campus do Pici, Universidade Federal do Ceará, Fortaleza, Ceará 60.440-970, Brazil

ARTICLE INFO

Keywords:
Lectins
Glycobiology
Omics
Computational techniques

ABSTRACT

Lectins play crucial roles in many biological processes and serve as tools in fields ranging from agriculture to biomedicine. While classical methods for lectin discovery and characterization were foundational for the field, they often lack sensitivity and throughput, limiting the detection of less abundant or weakly binding lectins, such as the stress-inducible or monovalent lectins. This review focuses on recent advancements in plant lectin research, particularly novel technologies that complement traditional approaches. Techniques such as glycan microarrays allow rapid assessment of lectin specificity across a diverse range of glycans by evaluating interactions with immobilized glycans on solid surfaces. Phage display libraries enable the identification of carbohydrate-mimetic peptides and the development of ligands for lectins by presenting diverse peptide libraries on bacteriophages. Genomic and transcriptomic analyses facilitate the exploration of the lectome in various plant species by scanning entire datasets to identify genes that contain lectin motifs—specific conserved amino acid sequences involved in carbohydrate recognition—and lectin domains, the larger structural regions that facilitate and stabilize these interactions. Additionally, computational methods—including molecular docking, molecular dynamics simulations, and machine learning pipelines—support predictions of lectin structures and binding properties, underpinning experimental efforts. These advanced techniques bring increased efficiency, accuracy, and a broader scope to lectin studies, with potential impacts across multiple fields. However, challenges such as data complexity and the need for experimental validation for computational methods remain. The future of lectin research will depend on the integration of these methods and the strengthening of interdisciplinarity to unlock the full potential of lectins.

1. Introduction

Lectins are defined as proteins capable of specific and reversible binding to carbohydrates and glycoconjugates, and thus represent an important group of decoders of the glycocode [1,2]. Considering the importance of protein-carbohydrate interactions for numerous processes related to structural and molecular recognition it is not surprising that the binding between lectins and glycoconjugates is directly involved in various functions and biological activities displayed by these carbohydrate-binding proteins [3]. While traditional methods like hemagglutination assays have laid the foundation for lectin research, they often suffer from limitations in sensitivity and specificity, hindering the discovery of less abundant or weakly binding lectins. To overcome these challenges, a new generation of technologies has emerged, revolutionizing the way we identify, characterize, and understand lectins.

The growing interest in lectins gave rise to the lectin field, referred to as lectinology, and over the years the advances within the field resulted in large amounts of biochemical and structural data of a wide array of lectins [4], particularly those from plant origin, by far the most well-researched group. This interest in plant lectins was sparked by the many functions of lectins, ranging from roles in plant defense against pathogens to plant growth and development, in addition to the diversity of biological activities and biotechnological applications of these molecules including their antiviral, anticancer, antibacterial, antifungal properties, their use as tools in glycobiology [3,5–7].

Driven by this interest, the pipeline for lectin discovery and characterization has evolved, albeit slowly, as the lectinology field has

* Corresponding author.
E-mail address: ElsJM.VanDamme@UGent.be (E.J.M. Van Damme).

https://doi.org/10.1016/j.bbadva.2025.100145
Received 21 November 2024; Received in revised form 26 January 2025; Accepted 27 January 2025
Available online 28 January 2025

benefitted from a shift in focus to better align with modern advancements. Early lectin research primarily focused on purifying lectins from sources like plant tissues with high lectin concentrations, such as seeds. These efforts revealed a broad spectrum of lectins with varying specificities and biochemical properties, enabling the first attempts to classify plant lectins [4,8]. Lectins were initially categorized based on their reactivity with specific monosaccharides, determined by the hydroxyl group configuration at specific carbons of the pyranose ring. While later refined using oligosaccharide interactions, these early categorizations of lectins based on simple sugar interactions were fundamental to advancing the field [9]. Later, other classifications also emerged, such as grouping based on the quaternary structure (mero-, holo-, chimero- and superlectins) or based on their constitutive or stress-inducible expression patterns and subcellular location [10].

Literature analysis shows that research projects aimed at detecting, purifying, and biochemically characterizing plant lectins have generally followed similar steps, utilizing standard techniques such as hemagglutination assays, sugar-inhibition assays, and affinity chromatography [11]. Over time, lectin research has evolved from relying on qualitative and semi-quantitative affinity assays using mono- and disaccharides to obtain precise quantitative interaction data for a broad spectrum of carbohydrate ligands and biologically relevant glycans. This change was driven by advancements in techniques like glycan arrays, surface plasmon resonance, and microscale thermophoresis. Furthermore, progress in structural determination and molecular biology has driven the field towards a more structurally focused approach, which has been instrumental in revealing the diversity of structural folds and binding modes for plant lectins [7,12]. This structural focus, combined with sequence data, later allowed the classification of plant lectins into 12 families based on their lectin domains. These families include: the *Agaricus bisporus* lectin family, Amaranthin family, Class V chitinase-related agglutinin (CRA) family, Cyanovirin domain-containing family, *Euonymus europaeus* lectin (EUL) domain-containing family, *Galanthus nivalis* agglutinin (GNA) domain-containing family, Hevein domain-containing family, Jacalin-related lectin (JRL) family, Legume lectin family, LysM domain-containing family, *Nicotiana tabacum* agglutinin (Nictaba) family, and Ricin-B domain-containing family. Collectively, these 12 families showcase the remarkable diversity of plant lectins. While each

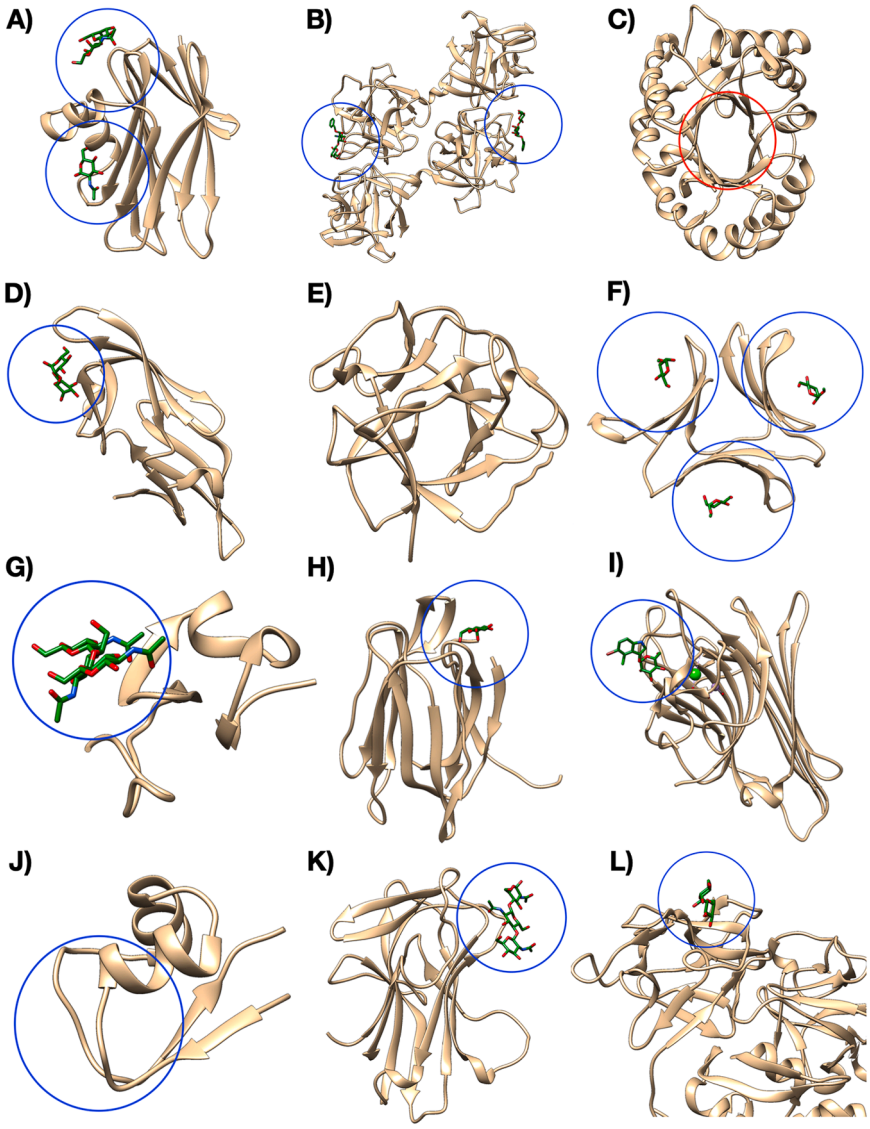

Fig. 1. 3D structure of representative structures of the different families of plant lectins with the carbohydrate-recognition domain highlighted in blue (experimentally confirmed location) or red (predicted location). A) *Agaricus bisporus* agglutinin (PDB id: 1y2 w); B) Amaranthin dimer (PDB id: 1jlx); C) *Parkia platycephalla* lectin II (PDB id: 2gsj); D) Cyanovirin (PDB id: 2pys); E) Modeled *Euonymus europaeus* lectin; F) *Galanthus nivalis* agglutinin (PDB id: 1msa); G) Hevein (PDB id: 1hev); H) Jacalin (PDB id: 1ugw); I) *Dioclea megacarpa* lectin (PDB id: 8ux7); J) LysM domain (PDB id: 4uz3); K) Nictaba (PDB id: 8ad2) and L) Ricin-B domain of Ebulin (PDB id: 1hwo).

family shares a characteristic structural fold, lectins within one family can exhibit a wide range of carbohydrate-binding specificities, recognizing structures from simple monosaccharides to complex oligosaccharides and glycans. This diversity in binding specificity is mirrored by the variety of structural folds adopted by their lectin domains, including the prevalent β-sandwich, the β-trefoil, the β-prism, as well as some unique folds in some lectin families. The quaternary structures also vary, encompassing monomers, dimers, tetramers, and even higher-order oligomers. Functionally, these families are implicated in a wide range of plant processes. Some lectins, such as the legume lectins and GNA-related lectins, are abundant in seeds and are thought to play roles in storage and defense. Other lectins, including the LysM domain-containing lectins and chitinase-related agglutinins, are involved in recognizing pathogen-associated molecular patterns and triggering immune responses. The stress-inducible Nictaba-related lectin family and the *Euonymus europaeus* lectin family highlight the roles of lectins in responding to environmental challenges [8]. Fig. 1 displays the overall fold of representative proteins for each of these 12 plant lectin families. Most images represent the monomeric form of the lectins with the position of the carbohydrate-binding site highlighted, except for Amaranthin where the binding site is located at the interface of two subunits. Please refer to [8] for detailed information on each lectin family.

Classical lectin detection techniques, such as hemagglutination and sugar-inhibition assays, remain easy tools for the initial detection and characterization of plant lectins [13]. While these techniques are widely used due to their simplicity and effectiveness, they have several inherent drawbacks. A key limitation is their low sensitivity, requiring high lectin concentrations for detectable agglutination. Additionally, these techniques primarily provide qualitative data without precise quantification, and the results can be influenced by factors like the specific glycans on the surface of the red blood cells, temperature and pH, which affect lectin activity. Non-specific agglutination and cross-reactivity further complicate the interpretation of the results, since contaminants and other proteins may bind to red blood cells, potentially leading to inaccuracies in lectin specificity assessment. Considering these limitations, advanced approaches can be applied standalone or in combination with classical methods leading to a significant increase in impact on the research. An overview of the techniques used for lectin discovery and characterization is presented in Fig. 2.

This review focuses on the latest advances in plant lectin research, summarizing novel approaches such as glycan microarrays, phage display, and computational methods, while contextualizing them within existing literature to highlight their potential to advance the field of lectinology.

2. Advanced methods for lectin discovery and carbohydrate-specificity analysis

Novel screening methods for lectin discovery, such as a search for lectin genes in genomic and transcriptomics databases, phage display libraries, and glycan microarrays represent interesting approaches that can be complementary to the traditional techniques used for lectin discovery.

2.1. Genome-Wide studies

As shown in Fig. 2, genomic and transcriptomic database searches provide an opportunity to explore and identify the lectome for each plant species. This approach involves making use of publicly available datasets to provide a comprehensive overview of lectin sequences encoding representatives from each lectin family within a plant.

For instance, a genome-wide study for *Arabidopsis thaliana* by Eggermont and colleagues [14] resulted in the identification of 217 genes, encoding putative lectins from nine out of twelve plant lectin families identified. These nine families include CRA (9 genes), EUL (1 gene), GNA (49 genes), hevein (10 genes), JRL (50 genes), legume lectin (54 genes), LysM (12 genes), Nictaba (30 genes), and ricin-B (2 genes). Interestingly, most lectin genes identified in the genome of *A. thaliana* encode putative chimeric lectins in which the lectin domain is linked to one or more protein domains, many of which are associated with roles in stress signaling and plant defense. The study also mapped the lectin genes across the plant genome and investigated their evolutionary relationships. Phylogenetic analyses indicated that lectin sequences sharing similar domain architectures evolved together, with conserved amino acids in the carbohydrate-binding site for different lectin families.

Lectomes for *Citrus sinensis* (sweet orange) and *Cucumis sativus* (Cucumber) were investigated using genome-wide identification of all

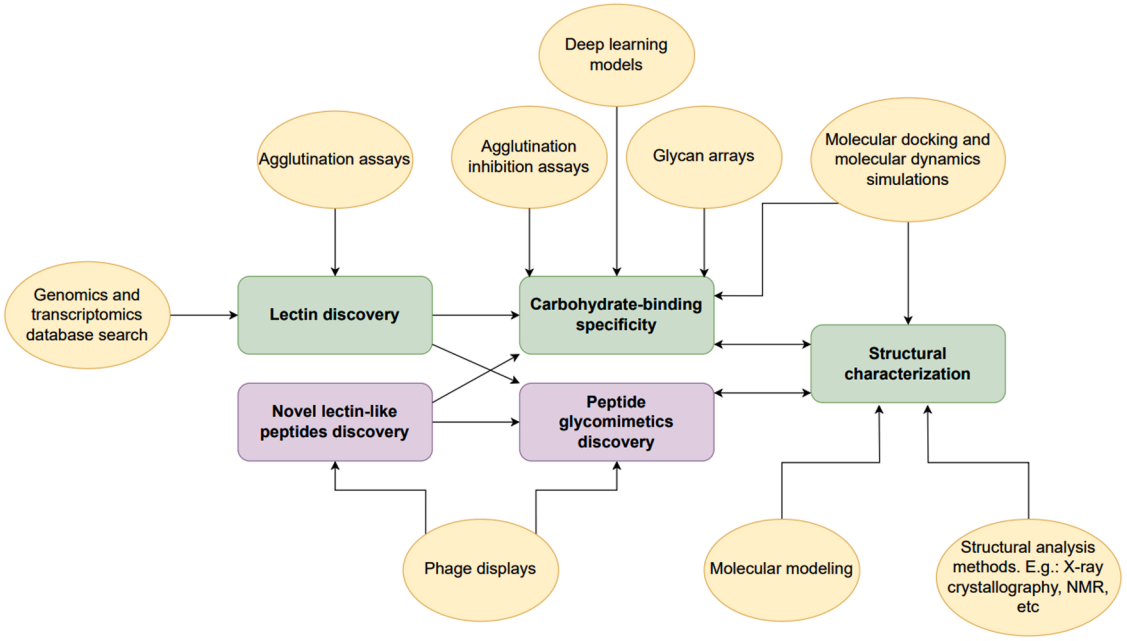

Fig. 2. Flowchart summarizing different experimental techniques used for lectin discovery, determination of carbohydrate-binding specificity and structural characterization.

putative lectin sequences [15,16]. The study in *Citrus sinensis* identified 141 genes encoding putative lectins from 10 distinct lectin families. The researchers used comparative genomic approaches, applying known lectin gene sequences from *Arabidopsis thaliana* to identify and classify the lectin genes in sweet orange. Phylogenetic analyses revealed that these genes share a significant level of sequence homology with other plant species, revealing their conserved nature across the plant kingdom. In addition to gene identification, this study explored the regulatory landscape of lectin genes, identifying key transcription factors such as ERF, MYB, NAC, WRKY, bHLH, and TCP that regulate these genes. Cis-acting regulatory elements were found to be involved in stress-responsive, light-responsive, and hormone-responsive pathways, suggesting that lectins play a role in the adaptation of sweet orange to environmental stressors [32]. Similarly, the genome-wide analysis of *Cucumis sativus* identified 146 putative lectin genes distributed across the cucumber genome, encoding lectins from 10 families. The study placed significant emphasis on the domain architecture of these genes, revealing that many lectin sequences encode chimeric proteins, where lectin domains are fused to other protein domains, such as kinase or glycosyl hydrolase domains. This structural complexity suggests that cucumber lectins have dual functions, particularly in plant defense mechanisms and signaling pathways. Moreover, the study highlighted that the expansion of lectin genes in cucumber was primarily driven by tandem and segmental duplication events, suggesting that these processes may have enabled the species to adapt to a variety of environmental stresses. The analysis of carbohydrate-binding sites within these lectin domains indicated that many retained their functional roles, further supporting the idea that these proteins are integral to cucumber's biological responses to external stimuli [16].

Different from the previously cited studies, the work of Quan et al. [17] focused on a single lectin family. The authors investigated the Jacalin-related lectins in barley (*Hordeum vulgare*) and lectin involvement in low-nitrogen stress responses. The authors identified 32 genes encoding putative Jacalin-related lectins. Transcriptomics studies employing two barley genotypes with contrasting tolerances to low-nitrogen stress identified 9 differently expressed lectins. Of particular interest was the gene HvHorcH, which was strongly upregulated in the tolerant genotype. Functional characterization of HvHorcH in transgenic *Arabidopsis* plants confirmed that overexpression of this gene enhanced low nitrogen tolerance, highlighting its role as a key regulator in the stress signaling.

2.2. Glycomimetics and phage display

Phage display has had limited use in plant lectin research up to this point, which is unexpected given that it is one of the few techniques available for finding glycomimetic and lectin-like peptides (Fig. 2). In this technique, bacteriophages are used to display a multitude of peptides on their surfaces, enabling researchers to identify peptides that act as glycan mimics or carbohydrate-binding peptides. These peptides can then be developed into ligands that target lectins or lectin-like molecules [18].

Glycomimetics, molecules that emulate the structure and function of natural glycans, represent powerful tools for their ability to mimic natural carbohydrates, making them exceptionally effective ligands for lectins [19]. This is relevant in the study of plant lectins, which play crucial roles in cell signaling, pathogen recognition, and plant defense [20]. Glycomimetics provide a means to investigate and manipulate lectin-carbohydrate interactions with precision. By developing structures that resemble the natural ligands of plant lectins, researchers can achieve enhanced binding affinity, stability, and selectivity. This allows for detailed investigation of lectin activity and facilitates the exploration of their potential applications in agriculture, biotechnology, and medicine. Furthermore, glycomimetics can function as either inhibitors or agonists of plant lectin-mediated processes, offering the potential to control pathogen-plant interactions or enhance crop stress responses.

For instance, Scott and colleagues [21] applied a large hexapeptide library displayed on filamentous phages, containing approximately 200 million unique peptide sequences covering a significant portion of all possible combinations. The authors aimed to find sequences able to mimic the ligands of ConA and went about it by putting the library in contact with ConA and washing out the non-binders. By subjecting the peptides to 4 more rounds of selection, the researchers were able to pinpoint those sequences containing the YPY sequences presented strong interaction with the lectin and proved the potential of the technique to find glycomimetic sequences.

In a follow-up study, Yu and colleagues [22] built upon this approach by attempting to discover glycomimetic peptides that could bind to multiple lectins with shared sugar selectivity. They used two libraries, one displaying 12 random residues in a linear format ('X12') and the other displaying seven random amino acids flanked by two cysteine residues that cyclize the peptide ('C–X7–C'). Using these libraries, phage display screening was performed against *Canavalia ensiformis* lectin (concanavalin A, ConA), *Lens culinaris* agglutinin (LCA), and *Pisum sativum* lectin (PSA). The screening strategy was designed to select phages that bind to multiple targets while reducing the number of hits that bind non-specific binding sites. The authors reported that, although the resulting sequences from their screens did not converge on a YPY consensus, they are rich in threonine (T) and serine (S) residues providing hydrogen bonding interactions. Proline (P) residues, which appear in many of the convergent sequences from the screens, may serve an important structural role by spatially orienting key side chains. Nine of the eleven convergent phage clones were selected for further analysis. Binding analysis using a phage based ELISA experiment revealed that, except for STLHALDSHLAL, all the selected clones bound strongly to ConA, LCA, and PSA.

Another innovative application of phage display is the selection of glycopeptide ligands, which are composed of peptides chemically modified with sugars. In this context, Chang et al. [23] employed a α-helical peptide phage library to find strong ligands for ConA by modifying the peptides with mannose. The authors observed that the addition of the peptide region increased the binding affinity reaching a minimum of 1.2 μM dissociation constant for the p3-Man peptide and the high affinity was a result of hydrophobic and polar residues.

2.3. Glycan arrays

As depicted in Fig. 2, glycan arrays are a central method for determining carbohydrate-binding specificity. This technique is crucial technology for the lectin field, enabling the rapid assessment of lectin specificity towards diverse glycan structures. Constructed by immobilizing oligosaccharides onto slides in a spatially defined manner, these arrays allow for the profiling of fluorescently labeled lectins against a wide range of carbohydrate motifs [24]. Advancements in the technology, particularly the incorporation of structurally diverse glycans, including branched and multi-antennary structures, have expanded the scope of research into complex glycan-binding events [25,26]. For over a decade, glycan arrays have been widely used for determining the fine carbohydrate-binding specificity of plant lectins. While traditional methods like sugar-inhibition assays offer insights into lectin specificity for mono- and disaccharides, microarrays refine this analysis, providing detailed binding motifs for a given lectin.

An example of this application was reported by Benevides and colleagues [27], who investigated the fine specificity of the legume lectin from *Platypodium elegans* (PELa) using a combination of glycan array analysis and isothermal titration calorimetry. These methods confirmed the previously reported mannose-binding specificity of the lectin, but also revealed a unique specificity towards asymmetric complex *N*-glycans. Specifically, PELa demonstrated a particularly strong affinity for N-glycans where one arm contains only one mannose residue, while the other arm carries extended structures such as GlcNAc, Gal, and sialic acid. This property distinguishes PELa from other mannose-specific

legume lectins, which often prefer oligomannose-type glycans. In a separate study, Oliveira et al. [28] combined agglutination inhibition assays and glycan array analysis to define the specificity of *Collaea speciosa* lectin. The authors employed a glycan array displaying mammalian, plant, and insect glycans, and reported that the lectin binds to GlcNAc, GlcNAc oligomers, and specific *N*-glycan structures containing the trisaccharide GlcNAcβ1-2(GlcNAcβ1-4)Man. Complex *N*-glycans featuring these motifs displayed significant binding, whereas those with galactose substitutions or lacking terminal GlcNAc exhibited reduced or no binding. This result contrasts with other lectins from the same subtribe, which are typically classical ConA-like lectins.

Lannoo et al. [29] employed this approach to investigate the *in situ* interactions of the tobacco lectin Nictaba, a protein known to reside in the cytoplasm and the nucleus of tobacco leaf cells. They used glycan array screening, using arrays containing hundreds of structurally diverse glycans, to determine Nictaba's binding preferences. This analysis revealed that Nictaba exhibits a strong affinity for both high-mannose and complex *N*-glycans. Further analysis of the array data suggested that the binding motif preferred by Nictaba is the core structure Manβ1-4GlcNAcβ1-4GlcNAcβ-N-Asn. The preference for this common structure provides an explanation for Nictaba's interaction with numerous glycans in the array. The array data also confirmed Nictaba's previously reported interaction with chito-oligosaccharides, but highlighted a significantly stronger binding to the *N*-glycan structures.

Glycan arrays can elucidate the carbohydrate-binding specificity of novel lectins without requiring prior knowledge. For example, Lundstrøm et al. [30] determined the specificity and structure of an R-type lectin from *Cucumis melo* using a combination of techniques, including multiple glycan arrays comprising a total of 1046 unique glycan sequences. Microarray data indicated a strong preference of the lectin for Fucα1-2 Gal and GalNAc epitopes, present in various common structures such as blood group H, LacdiNAc, and other important antigen motifs. Further analysis revealed the lectin's high specificity for C2-substituted galactose, where the galactose is modified at position 2 with either a fucosyl or an N-acetyl group. Notably, β-linked GalNAc exhibited stronger binding than its α-counterpart. Binding to glycosaminoglycans was also observed, particularly to chondroitin sulfate types A and C, with affinity increasing with chain length.

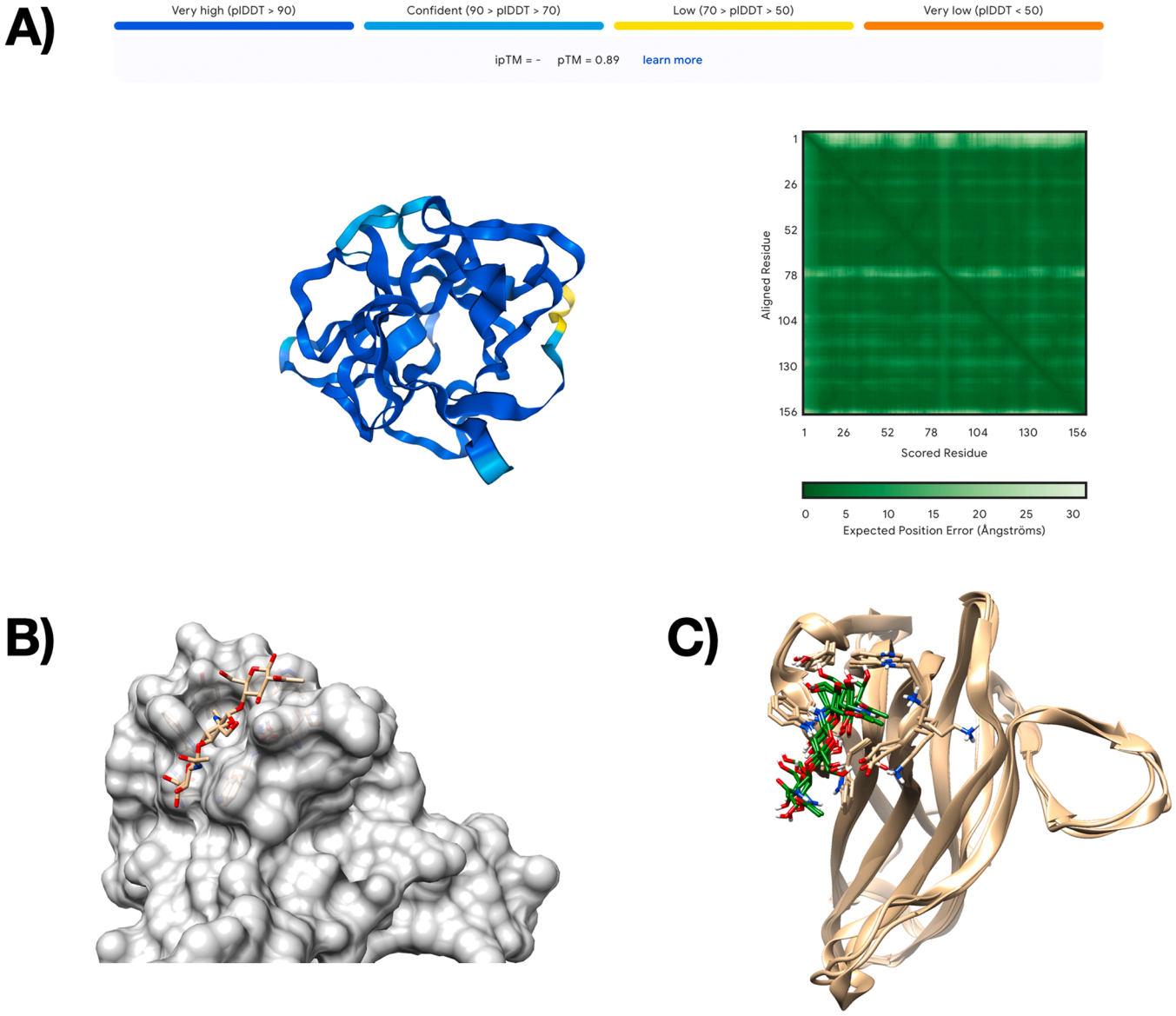

Fig. 3. Examples of outputs from computational tools. A) Molecular modeling output of the *Euonymus europaeus* lectin using AlphaFold3; B) Molecular docking of Nictaba in complex N-acetyl chitotriose and C) Snapshots of molecular dynamics simulations of Nictaba in complex with N-acetyl chitotriose.

3. Predictive computational approaches for lectin discovery and characterization

Computational and bioinformatics approaches have been applied for a long time in lectin research and can be applied in multiple steps of the lectin characterization process, including prediction of structural conformation, lectin interactions and binding specificity. Predictive approaches are made more accurate by increasing the amount of available structural and binding data and have a significant advantage of being cost-effective for the data they generate. Structural prediction tools like AlphaFold, along with functional annotation via computational methods, help researchers to explore the binding mechanisms and link expression patterns to potential functions [31]. These approaches streamline discovery, reduce time and costs, and enhance accuracy through virtual screening and automation. Examples of outputs of molecular modeling, molecular docking and molecular dynamics for lectin studies are presented in Fig. 3.

3.1. Structural modeling

The computational methods that are most applied in lectin research are tools that allow the prediction of the three-dimensional (3D) structure of lectins using their amino acid sequences as inputs, in combination with molecular docking pipelines to predict the binding with carbohydrates [32]. Given the abundance of experimental data for plant lectins, homology modeling has been used most for structural prediction of lectins although machine learning-based methods are taking over. Understanding the 3D structure of proteins is crucial for understanding their carbohydrate-binding properties. In the absence of experimental structure data, structural prediction of plant lectins has been used to gain this understanding. A representation of the output from a structural prediction simulation can be seen in Fig. 3A.

For instance, the work of Imberty and colleagues [33] employed knowledge-based homology modeling to predict the structure of the GalNAc-specific lectin from *Dolichos biflorus* using the conserved regions of five template structures of similar proteins, sharing 37 % - 50 % identity. The authors successfully predicted the β-sandwich structure of the legume lectin and its expected interactions with GlcNAc through molecular docking. 85 % of the lectin structure was covered by similar lectins, with loop structures being manually edited based on the template's averages to match the predicted structure. Computational prediction with the goal to understand the 3D structure and binding of a lectin was also used in the work of Agostino and colleagues [34] who employed hidden Markov models to identify templates for homology modeling of the *Euonymus europaeus* lectin (EUL). The authors managed to build the structure and concluded that EUL adopts a ricin-B-like fold (R-type) with a β-trefoil architecture stabilized by hydrophobic interactions and a tightly packed core. Molecular docking revealed that EUL binds fucosylated structures related to blood groups B and H antigens with conserved residues such as Asp130 and Gln149 forming H-bonds with the ligands and CH-π stacking interactions being especially important for the binding.

Recently, a new type of molecular modeling has emerged and designated as machine learning-based methods joining the traditional template-based modeling and template-free modeling methods [35]. A game-changing tool in structural biology is AlphaFold, a deep-learning model developed by DeepMind to predict protein structures. It applies multiple sequence alignments and co-evolutionary signals that can predict amino acid positions in space using neural networks [31]. Although the adoption of AlphaFold in publications involving plant lectins is still low, it is likely to grow fast. Bobili et al. [36] used AlphaFold2 to predict the structure of several phloem lectins and compared their fold to the structure from *Cucumis sativus* (Cus17), a lectin structure that was solved experimentally using X-ray diffraction analyses. The prediction data revealed that all phloem lectins tested displayed a conserved β-sandwich fold and a carbohydrate recognition domain containing residues similar to those identified in Cus17. This investigation expanded the experimental insights towards similar lectins and provided a structural basis for the study of phloem lectins in plant defense mechanisms. In another work, AlphaFold3 has been used to predict the 3D structure of a lectin from *Bauhinia holophylla* revealing a jelly roll motif for the lectin monomer, typical for legume lectins [37].

3.2. Molecular docking and molecular dynamics

Molecular docking and molecular dynamics are also important computational techniques that are widely used in the study of plant lectins and enable us to better understand the interactions between lectins and carbohydrates, glycoconjugates and other molecules [38]. Molecular docking is normally used to predict the preferred orientation of the carbohydrates within the binding site (Fig. 3B) whereas molecular dynamics simulations complement docking studies or experimental data by providing a dynamic view into the protein and complexes over time (Fig. 3C) allowing the evaluation of properties including structural stability, binding mechanisms while also providing a layer of validation to the docking results [38,39].

Publications reporting the use of molecular docking and molecular dynamics simulations to study lectins are plentiful in literature. In the paper of Pol-Fachin [40], the author explores molecular dynamics simulations to study the dynamics of the lectin from *Crataeva tapia* (CrataBL), specifically focusing on how the glycosylation of the lectin affects the conformation, dynamics and carbohydrate-binding activity of the lectin. The results indicated that the two identified N-glycans contribute to local and distant stabilization of the protein structure. Additionally, the study proposed the location of two potential carbohydrate recognition domains and the structural basis for the lectin specificity for glucose. Molecular docking and molecular dynamics have also been used to study the effect of posttranslational processing of lectin polypeptides. The work of Osterne et al. [41] investigated the effect of circular permutation on ConA. Circular permutation is a rare post-translational modification that significantly impacts protein stability without altering carbohydrate-binding properties. The study applied molecular dynamics simulations and molecular docking to compare ProConA (unprocessed ConA) and mature ConA in their binding interactions with d-mannose and a high-mannose *N*-glycan. The findings reveal that while the carbohydrate-binding properties remain consistent between the two forms, the stability of ProConA is lower than that of mature ConA, especially in maintaining its tetrameric oligomerization.

A publication focused on applications of lectins, investigated the potential of the banana lectin (BanLec) as a therapeutic agent against SARS-CoV-2. BanLec is a mannose/glucose-binding lectin, and the study aimed to understand its interactions with the receptor-binding domain of SARS-CoV-2 through computational approaches, including molecular docking and molecular dynamics simulations. The results demonstrated that both wild-type and modified BanLec proteins have strong affinities for the binding domain of various SARS-CoV-2 variants, including the Omicron variant. The binding was stable throughout the simulations, and the study suggested that BanLec could inhibit viral entry into host cells by interacting with the key residues of the receptor binding domain that are critical for viral attachment to the human ACE2 receptor [42].

3.3. Bioinformatics

Lectin databases are invaluable for lectin researchers and provide a publicly available catalog for lectins. The primary purpose of lectin databases is to provide researchers with structured, searchable information on structural properties, functions, sources and binding properties of lectins with each database focusing on one or more of these aspects (Table 1).

Table 1
Non-exhaustive list of online databases specialized in various aspects of lectin research.

Database	Description	Website
Unilectin [43]	Comprehensive platform that integrates information about lectin sequences, 3D structures, and carbohydrate-binding specificity, offering tools to explore lectins across various species.	https://unilectin.eu
GlyTouCan [44]	While not exclusive to lectins, this database focuses on glycans and includes lectin-related data, especially in terms of glycan recognition.	https://glytoucan.org
CAZy Database [45]	Database containing extensive information on carbohydrate-active enzymes, including lectins and other glycan-binding proteins, organized according to their functional domains.	http://www.cazy.org
LFDB Database [46]	Database offers quantitative interaction data for lectins and glycans, focusing on affinity constants obtained through frontal affinity chromatography.	http://riodb.ibase.aist.go.jp/rcmg/lfdb
CFG Database	Database focusing on experimental data for glycan-lectin interactions, helping to identify glycan binding preferences for various lectins.	https://research.bidmc.org/ncfg
GlyCosmos Lectins [47]	Part of the GlyCosmos portal integrates lectin data from various sources, including UniProt and PDB, providing a comprehensive view of glycan-binding proteins.	https://glycosmos.org/lectins/
CarboGrove [48]	Database designed to analyze glycan-array data to determine the glycan-binding specificities of lectins.	https://carbogroove.org

3.4. Novel methods

A recent revolution involves the introduction of machine learning methods, this approach is becoming more widespread in the study of lectins, particularly in the context of investigations focusing on lectin-glycan binding interactions. Carbohydrates, or glycans, participate in critical biological processes such as immune responses, cell signaling, and pathogen recognition [49]. However, due to the structural diversity of lectins and the complexity of their interactions with carbohydrate structures, traditional experimental methods such as nuclear magnetic resonance and X-ray crystallography have struggled to accurately map these interactions. Recent innovations in machine learning, specifically deep learning models, have provided robust solutions to these challenges, offering a more scalable, precise, and efficient way to predict glycan-protein interactions.

A significant development in this field has been achieved in the work of Lundstrøm et al. [50] which revealed LectinOracle, a generalizable deep learning model that predicts lectin-glycan interactions. The LectinOracle algorithm combines transformer-based representations for protein sequences with a graph convolutional neural network that models glycan structures. Unlike previous models, LectinOracle is generalizable to novel lectins, making it a tool for predicting the specificity of lectins that have not yet been characterized experimentally. The model uses protein sequence data as input rather than protein structures, which significantly broadens its applicability, as sequence data are far more abundant and accessible. Additionally, the program ability to predict glycan-binding specificity with high accuracy enables its use in diverse applications, such as lectin classification, understanding host-pathogen interactions, and therapeutic development. By learning from a vast dataset of protein-glycan interactions, the model can accurately predict both well-documented and novel interactions, expanding the resources available for glycan research.

A relevant work has been published by Bojar et al. [51], where the authors reported on the application of machine learning to annotate the glycan-binding specificity of 57 lectins, using a combination of machine learning and manual annotation to analyze glycan microarray data and evaluate the binding specificity of commercially available lectins. By employing machine learning algorithms to probe the microarray data, they were able to define predominant binding motifs and reveal subtle binding characteristics that had previously been overlooked. The integration of machine learning into this process allowed for the processing and interpretation of complex glycan-binding behaviors on a large scale, which would have been impossible with manual methods alone. Their work demonstrated how computational tools could expand the applications of lectins in glycan research by enhancing the specificity and the precision of identification of the binding motif(s).

Another milestone in lectin research was the introduction of DeepGlycanSite by He et al. [52], a deep learning model designed to predict carbohydrate-binding sites on protein structures. DeepGlycanSite integrates both geometric and evolutionary features of proteins using a deep equivariant graph neural network combined with transformer architecture. This architecture allows the model to outperform existing methods significantly. DeepGlycanSite can predict binding sites for a wide variety of glycans, including monosaccharides and oligosaccharides. Its use of geometric features, such as residue orientations and distances within protein structures, combined with evolutionary information from protein sequences, enables the model to generalize across different carbohydrate-binding proteins.

4. Advantages and limitations of advanced lectin discovery and characterization methods

As highlighted throughout the review, the discovery and characterization of novel plant lectins is accelerating due to novel approaches. Collectively these methodologies have enhanced the efficiency, accuracy and scope of lectin research while also enabling a deeper view into the carbohydrate-binding specificities, an essential step in the understanding of the biological roles and applications of lectins. However, these techniques are not without limitations; challenges related to protocol and data complexity, interpretation, and the need for experimental validation for computational approaches persist, indicating that the combination of novel and traditional techniques is likely to remain important for several years. A summary of advantages and limitations of the techniques reviewed in the current work can be accessed in Table 2.

4.1. Advantages of current techniques

Genomic and Transcriptomic Databases: The main advantage of genomic and transcriptomic approaches is their efficiency. Bioinformatics tools significantly accelerate lectin discovery by enabling rapid, genome-wide identification of potential lectin sequences. By screening vast datasets across plant species, researchers can easily identify putative lectin genes based on conserved domains and sequence homology [53]. This high-throughput approach facilitates comprehensive cataloging of lectin sequences and allows for initial predictions about their structures and functions. Importantly, this method is relatively fast and inexpensive compared to traditional experimental methods.

Phage Display: Phage display technology offers the key advantage of rapidly screening a vast library of peptides to identify those with high affinity for a target lectin. This is particularly valuable for the discovery of glycomimetic peptides that can mimic natural carbohydrate ligands. By presenting a diverse array of peptide motifs and iteratively selecting for high-affinity binders, researchers can accelerate the identification of compounds for further study. This technique has been successfully employed to identify peptides that interact with specific lectins [19].

Glycan Arrays: A major advantage of glycan arrays is their ability to provide a high-throughput platform for detailed profiling of lectin-glycan interactions. By immobilizing hundreds of diverse glycan structures on a solid surface, these arrays allow for the simultaneous analysis of lectin binding to a wide range of potential ligands [54]. This not only saves time but also generates comprehensive data on the fine specificity

Table 2
Summary of advantages and limitations of techniques used in lectin research.

Technique	Advantages	Limitations
Genomic/ Transcriptomic Database Search	High-throughput screening of vast databases. Identification of putative lectins based on conserved domains. Evolutionary insights. Cost-effective.	Limited to sequence-based predictions. Overlooks post-translational modifications and conformational changes. Requires experimental validation.
Phage Display	Rapid screening of millions of peptide variants. Identification of glycomimetic peptides and lectin-like peptides.	In vitro conditions may not reflect *in vivo* interactions. Subsequent assays needed to verify functional relevance.
Glycan Arrays	High-throughput analysis of lectin-glycan interactions. Detailed specificity profiling. Semi-quantitative binding data.	Challenges related to glycan immobilization, orientation, and density. Variability across platforms complicates comparisons. May not reflect the native context of glycans on cell surfaces.
Computational Modeling	Cost-effective prediction of lectin structure and interactions. Exploration of structural and energetic aspects of binding. Good for hypothesis generation.	Dependence on high-quality protein models. Challenges in accurately predicting interactions with complex glycans. Computationally demanding. Requires experimental validation.
Machine Learning	Prediction of glycan-binding specificities from large datasets. Pattern recognition in complex data. Streamlines discovery process.	"Black box" nature limits interpretability. Heavily reliant on large, high-quality datasets. Predictions may be biased or incomplete if training data is not representative. Requires experimental validation.

of lectins, revealing subtle differences in binding preferences that would be difficult to discern using traditional methods.

Computational Modeling: Computational tools, including molecular modeling, docking, and molecular dynamics simulations, offer the significant advantage of being able to predict lectin structure and lectin-glycan interactions in a cost-effective manner. These methods allow researchers to explore the structural and energetic aspects of binding interactions, providing insights into binding affinities and specificities that can guide experimental design [55–57]. For instance, computational techniques can identify potential binding sites and predict how structural modifications to lectins or glycans might influence their interaction. This is particularly useful for refining hypotheses and narrowing down potential candidates for experimental follow-up.

Machine Learning: The application of machine learning, particularly deep learning, to lectin research has the advantage of enabling the analysis of vast and complex datasets. Models trained on large volumes of lectin and glycan interaction data can predict glycan-binding specificities, even for lectins that have not been previously characterized [31, 50]. Tools like LectinOracle and AlphaFold have demonstrated the ability to accurately predict lectin-glycan interactions, which can greatly enhance the efficiency of directed evolution experiments and high-throughput annotation of lectin-glycan binding profiles. This predictive power helps streamline experimental efforts by focusing research on the most promising lectins and glycans.

4.2. Limitations of current techniques

Genomic and Transcriptomic Approaches: While efficient, genomic and transcriptomic approaches are limited by their reliance on sequence-based predictions. These methods primarily identify putative lectin sequences based on homology and conserved domain structures, often overlooking post-translational modifications, conformational changes, and specific glycan-binding properties that are crucial for function. Moreover, the presence of a lectin domain does not guarantee that the protein is a functionally active lectin. Thus, experimental validation of bioinformatically-predicted lectins remains essential.

Phage Display and Glycan Arrays: Both phage display and glycan arrays, despite their high-throughput capabilities, are limited by their *in vitro* nature. The conditions in these assays may not fully respresent the complex, multivalent interactions and structural diversity of glycans in their native biological context. Consequently, peptides or glycans that show strong binding *in vitro* may not exhibit the same specificity or affinity *in vivo* [58]. Glycan arrays, in particular, face challenges related to glycan immobilization, orientation, and density, which can affect binding outcomes and may not accurately reflect the natural presentation of glycans on cell surfaces. Variability across different glycan array platforms also complicates the cross-comparison of results and the generalization of binding specificities [59,60].

Computational Modeling: Computational modeling, while powerful, is limited by its dependence on high-quality protein structural models and the inherent complexity of glycan structures. Large, branched, and flexible glycans pose significant challenges for accurate prediction of their interactions with lectins [61,62]. Furthermore, these methods are computationally intensive and often rely on algorithmic simplifications that may not fully capture the intricate conformational dynamics and binding interactions of glycans under physiological conditions [63,64].

Machine Learning: Machine learning models, including deep learning approaches, are limited by their “black box” nature, making it difficult to understand the underlying mechanisms driving their predictions. This can hinder the development of testable hypotheses and mechanistic insights [65,66]. Moreover, the accuracy of these models is highly dependent on the quality, diversity, and representativeness of the training data. Biased or incomplete datasets can lead to inaccurate or incomplete predictions [67]. Similar to the broader challenges facing artificial intelligence (AI) in glycobiology, the complexity of glycan biosynthesis and the limited availability of comprehensive datasets linking glycosylation to biological processes pose significant hurdles for developing robust and reliable AI models in lectin research. These limitations indicate the need for careful curation of training data, transparent model development, and experimental validation of AI-generated predictions [68].

In summary, while every approach provides valuable insights, each method is individually constrained by limitations in accuracy, data representation, and the necessity of thorough validation.

5. Future directions and potential impacts

The integration of novel experimental approaches with advanced computational methods is transforming the field of lectinology. While traditional methods have limitations, the synergy of newer technologies offers unprecedented potential. This combination overcomes constraints in sensitivity and throughput, enabling the discovery and characterization of less abundant or weakly binding plant lectins, such as stress-inducible or monovalent lectins. Techniques like glycan microarrays and phage display, combined with computational tools for large-scale data analysis, are accelerating discoveries and enhancing understanding of lectin-carbohydrate interactions. For instance, integrating glycan array data with machine learning improves lectin specificity predictions, allowing rapid screening of lectin libraries against diverse glycan structures [69]. Computational tools also enable the rational design of lectins with desired specificities, leading to engineered lectins with enhanced affinity or altered specificity [70]. AI-driven databases and predictive tools promise to standardize and globally share lectin data by integrating sequence information, structural data, and binding specificities, fostering innovation and collaboration in lectinology research

and biotechnology.

The implications of these advancements extend beyond basic research into diverse fields. In agriculture, understanding and manipulating plant lectins can enhance pest and disease resistance in crops. Lectins, integral to plant defense mechanisms, can be engineered to improve resistance, reducing dependence on chemical pesticides. This includes exploiting the insecticidal properties of lectins like GNA and Nictaba to create transgenic plants resistant to various insect orders. Similarly, antifungal lectins, particularly those with chitin-binding specificity, could combat fungal pathogens that threaten crop yields [10,71].

In medicine, lectins hold promise for diagnostics and therapeutics due to their ability to specifically bind glycans. This makes them valuable for detecting diseases characterized by abnormal glycosylation, such as cancer and viral infections. Altered glycosylation patterns create unique glycocodes for each cancer type, and lectins can be tailored to recognize these patterns [72]. Lectins like ConA and Viscumin exhibit antiproliferative activity against cancer cells [73]. Moreover, lectins show antiviral potential against glycosylated viruses such as HIV and coronaviruses by targeting viral glycoproteins, like gp120 of HIV or the spike protein of SARS-CoV-2, thereby blocking viral entry [74]. Lectins are also used in diagnostic assays, targeted therapies, drug delivery systems, and biosensors for detecting disease-associated carbohydrates or glycoproteins. Additionally, plant lectins serve as essential research tools in glycobiology, such as in lectin microarrays for glycan profiling, in histochemistry and cytochemistry for visualizing glycosylated structures, and in chromatographic systems for glycoconjugates purification [75,76].

The future of lectin research depends on integrating advanced methodologies and fostering interdisciplinary collaboration. AI and machine learning will play critical roles in data analysis, structural prediction, and the development of comprehensive lectin databases. Collaboration among biologists, computational scientists, and engineers will drive innovation and propel the field forward. Addressing existing challenges and adhering to ethical research practices will be essential to fully harness these advancements and ensure their responsible application across sectors.

6. Conclusion

The integration of novel technologies represents a transformative approach in lectin research. These innovations have the potential to overcome current limitations, enabling rapid discovery and detailed characterization of lectins with unprecedented efficiency. By harnessing these technologies, researchers can accelerate the development of lectin-based applications that address critical challenges in agriculture, medicine, environmental science, and beyond.

Interdisciplinary collaboration will be key to advancing the field. Bringing together expertise in biology, computational science, and engineering will foster innovative solutions to drive the next generation of discoveries in lectinology. As we move forward, careful consideration of the challenges and a commitment to responsible research practices will ensure that the benefits of these advancements can be fully realized.

CRediT authorship contribution statement

Vinicius J.S. Osterne: Writing – review & editing, Writing – original draft, Conceptualization. **Kyria S. Nascimento:** Writing – review & editing. **Benildo S. Cavada:** Writing – review & editing. **Els J.M. Van Damme:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) for their support. B.S.C. and K.S.N. acknowledge their positions as senior investigators of CNPq. V.J.S.O. expresses gratitude to FWO-Vlaanderen for the post-doctoral fellowship (12T4622 N). The authors also extend their appreciation to all collaborators and institutions that contributed to this work.

Data availability

All data supporting this study are included within this manuscript and the references cited herein.

References

[1] W.J. Peumans, E.J.M. Van Damme, Lectins as plant defense proteins, Plant Physiol. 109 (1995) 347–352.
[2] M. Ambrosi, N.R. Cameron, B.G. Davis, Lectins: tools for the molecular understanding of the glycocode, Org. Biomol. Chem. 3 (2005) 1593–1608.
[3] D. Chettri, M. Boro, L. Sarkar, A.K. Verma, Lectins: biological significance to biotechnological application, Carbohydr. Res. 506 (2021) 108367.
[4] E.J.M. Van Damme, 35 years in plant lectin research: a journey from basic science to applications in agriculture and medicine, Glycoconj. J. 39 (2022) 83–97.
[5] T. De Coninck, E.J.M. Van Damme, Review: the multiple roles of plant lectins, Plant Sci. 313 (2021) 111096.
[6] N. Sharon, H. Lis, A century of lectin research (1888–1988), Trends Biochem. Sci 12 (1987) 488–491.
[7] N. Sharon, H. Lis, History of lectins: from hemagglutinins to biological recognition molecules, Glycobiology 14 (2004) 53R–62R.
[8] V.J.S. Osterne, E.J.M. Van Damme, Structural and Functional Diversity of Plant lectins, in: Defense-Related Proteins in Plants, Elsevier, 2024, pp. 167–203.
[9] A. Pusztai, Lectins and Their specificity, in: Plant Lectins, Cambridge University Press, Cambridge, 1992, pp. 3–31.
[10] M. Tsaneva, E.J.M. Van Damme, 130 years of Plant Lectin Research, Glycoconj. J. 37 (2020) 533–551.
[11] K.S. Nascimento, A.I. Cunha, K.S. Nascimento, B.S. Cavada, A.M. Azevedo, M. R. Aires-Barros, An overview of lectins purification strategies, J. Mol. Recognit. 25 (2012) 527–541.
[12] E.J.M. Van Damme, History of plant lectin research, Methods Mol. Biol. 1200 (2014) 3–13.
[13] R. Katoch, A. Tripathi, Research advances and prospects of legume lectins, J. Biosci. 46 (2021) 104.
[14] L. Eggermont, B. Verstraeten, E.J.M. Van Damme, Genome-wide screening for Lectin Motifs in *Arabidopsis thaliana*, Plant Genome 10 (2017) plantgenome2017.02.0010.
[15] F.F. Ahmed, F.S. Dola, F.T. Zohra, S.M. Rahman, J.N. Konak, M.A.R. Sarkar, Genome-wide identification, classification, and characterization of lectin gene superfamily in sweet orange (*Citrus sinensis* L.), PLoS One 18 (2023) e0294233.
[16] L. Dang, E.J.M. Van Damme, Genome-wide identification and domain organization of lectin domains in cucumber, Plant Physiol. Biochem. 108 (2016) 165–176.
[17] X. Quan, C. Meng, C. Xie, H. Sun, B. Xu, R.Santos Bermudez, W. He, Genome-wide and transcriptome analysis of jacalin-related lectin genes in Barley and the functional characterization of HvJRL1 in low-nitrogen tolerance in *Arabidopsis*, Int. J. Mol. Sci. 24 (2023) 16641.
[18] W. Jaroszewicz, J. Morcinek-Orłowska, K. Pierzynowska, L. Gaffke, G. Węgrzyn, Phage display and other peptide display technologies, FEMS Microbiol. Rev. 46 (2022) fuab052.
[19] S. Leusmann, P. Ménová, E. Shanina, A. Titz, C. Rademacher, Glycomimetics for the inhibition and modulation of lectins, Chem. Soc. Rev. 52 (2023) 3663–3740.
[20] S. Naithani, S.S. Komath, A. Nonomura, G. Govindjee, Plant lectins and their many roles: carbohydrate-binding and beyond, J. Plant Physiol. 266 (2021) 153531.
[21] J.K. Scott, D. Loganathan, R.B. Easley, X. Gong, I.J. Goldstein, A family of concanavalin A-binding peptides from a hexapeptide epitope library, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 5398–5402.
[22] L. Yu, P.S. Yu, E.Y.M. Mui, J.C. McKelvie, T.P.T. Pham, Y.W. Yap, W.Q. Wong, J. Wu, W. Deng, B.P. Orner, Phage display screening against a set of targets to establish peptide-based sugar mimetics and molecular docking to predict binding site, Bioorg. Med. Chem. 17 (2009) 4825–4832.
[23] I.V. Chang, H. Tsutsumi, H. Mihara, Screening for concanavalin A binders from a mannose-modified α-helix peptide phage library, Mol. Biosyst. 13 (2017) 2222–2225.
[24] S. Serna, J. Etxebarria, N. Ruiz, M. Martin-Lomas, N.-C. Reichardt, Construction of N-glycan microarrays by using modular synthesis and on-chip nanoscale enzymatic glycosylation, Chemistry (Easton) 16 (2010) 13163–13175.

[25] K. Brzezicka, B. Echeverria, S. Serna, A. van Diepen, C.H. Hokke, N.C. Reichardt, Synthesis and microarray-assisted binding studies of core xylose and fucose containing N-glycans, ACS Chem. Biol. 10 (2015) 1290–1302.

[26] E.J.M. Van Damme, D.F. Smith, R.D. Cummings, W.J. Peumans, Glycan arrays to decipher the specificity of plant lectins, Adv. Exp. Med. Biol. 705 (2011) 757–767.

[27] R.G. Benevides, G. Ganne, R. da C. Simões, V. Schubert, M. Niemietz, C. Unverzagt, V. Chazalet, C. Breton, A. Varrot, B.S. Cavada, A. Imberty, A lectin from *Platypodium elegans* with unusual specificity and affinity for asymmetric complex N-glycans, J. Biol. Chem. 287 (2012) 26352–26364.

[28] M.V. Oliveira, V.J.S. Osterne, C.F. Lossio, S. Serna, N.C. Reichardt, K. S. Nascimento, E.J.M. Van Damme, B.S. Cavada, Purification and characterization of a highly thermostable GlcNAc-binding lectin from *Collaea speciosa* seeds, Int. J. Biol. Macromol. 193 (2021) 1562–1571.

[29] N. Lannoo, W.J. Peumans, E. Van Pamel, R. Alvarez, T.C. Xiong, G. Hause, C. Mazars, E.J.M. Van Damme, Localization and *in vitro* binding studies suggest that the cytoplasmic/nuclear tobacco lectin can interact *in situ* with high-mannose and complex N-glycans, FEBS Lett 580 (2006) 6329–6337.

[30] J. Lundstrøm, E. Gillon, V. Chazalet, N. Kerekes, A. Di Maio, T. Feizi, Y. Liu, A. Varrot, D. Bojar, Elucidating the glycan-binding specificity and structure of agglutinin, a new R-type lectin, Beilstein J. Org. Chem. 20 (2024) 306–320.

[31] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589.

[32] J.I. Blanco Capurro, M. Di Paola, M.D. Gamarra, M.A. Martí, C.P. Modenutti, An efficient use of X-ray information, homology modeling, molecular dynamics and knowledge-based docking techniques to predict protein-monosaccharide complexes, Glycobiology 29 (2019) 124–136.

[33] A. Imberty, F. Casset, C.V. Gegg, M.E. Etzler, S. Pérez, Molecular modelling of the *Dolichos biflorus* seed lectin and its specific interactions with carbohydrates: alpha-D-N-acetyl-galactosamine, Forssman disaccharide and blood group A trisaccharide, Glycoconj. J. 11 (1994) 400–413.

[34] M. Agostino, T. Velkov, T. Dingjan, S.J. Williams, E. Yuriev, P.A. Ramsland, The carbohydrate-binding promiscuity of *Euonymus europaeus* lectin is predicted to involve a single binding site, Glycobiology 25 (2015) 101–114.

[35] Q. Wuyun, Y. Chen, Y. Shen, Y. Cao, G. Hu, W. Cui, J. Gao, W. Zheng, Recent progress of protein tertiary structure prediction, Molecules 29 (2024) 832.

[36] K.B. Bobbili, N. Sivaji, B. Priya, K. Suguna, A. Surolia, Structure and interactions of the phloem lectin (phloem protein 2) Cus17 from *Cucumis sativus*, Structure 31 (2023) 464–479, e5.

[37] L.M.G.G. de Camargos, D. de O. Lopes, A.H.F. Castro, Characterization of Lectin from *Bauhinia holophylla* using bioinformatics tools, Horticulturae 10 (2024) 682.

[38] F. Nieto-Fabregat, M.P. Lenza, A. Marseglia, C. Di Carluccio, A. Molinaro, A. Silipo, R. Marchetti, Computational toolbox for the analysis of protein-glycan interactions, Beilstein J. Org. Chem. 20 (2024) 2084–2107.

[39] S. Pérez, Computational modeling of protein-carbohydrate interactions: current trends and future challenges, Adv. Carbohydr. Chem. Biochem. 83 (2023) 133–149.

[40] L. Pol-Fachin, Insights into the effects of glycosylation and the monosaccharide-binding activity of the plant lectin CrataBL, Glycoconj. J. 34 (2017) 515–522.

[41] V.J.S. Osterne, V.R. Pinto-Junior, M.V. Oliveira, K.S. Nascimento, E.J.M. Van Damme, B.S. Cavada, Computational insights into the circular permutation roles on ConA binding and structural stability, Curr. Res. Struct. Biol. 7 (2024) 100140.

[42] S.S. Hessel, F.M. Dwivany, I.M. Zainuddin, K. Wikantika, I. Celik, T.B. Emran, T. E. Tallei, A computational simulation appraisal of banana lectin as a potential anti-SARS-CoV-2 candidate by targeting the receptor-binding domain, J. Genet. Eng. Biotechnol. 21 (2023) 148.

[43] F. Bonnardel, S. Perez, F. Lisacek, A. Imberty, Structural Database for Lectins and the UniLectin Web Platform, Methods Mol. Biol. 2132 (2020) 1–14.

[44] M. Tiemeyer, K. Aoki, J. Paulson, R.D. Cummings, W.S. York, N.G. Karlsson, F. Lisacek, N.H. Packer, M.P. Campbell, N.P. Aoki, A. Fujita, M. Matsubara, D. Shinmachi, S. Tsuchiya, I. Yamada, M. Pierce, R. Ranzinger, H. Narimatsu, K. F. Aoki-Kinoshita, GlyTouCan: an accessible glycan structure repository, Glycobiology 27 (2017) 915–919.

[45] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics, Nucleic Acids Res 37 (2009) D233–D238.

[46] J. Hirabayashi, H. Tateno, T. Shikanai, K.F. Aoki-Kinoshita, H. Narimatsu, The Lectin Frontier Database (LfDB), and data generation based on frontal affinity chromatography, Molecules 20 (2015) 951–973.

[47] I. Yamada, M. Shiota, D. Shinmachi, T. Ono, S. Tsuchiya, M. Hosoda, A. Fujita, N. P. Aoki, Y. Watanabe, N. Fujita, K. Angata, H. Kaji, H. Narimatsu, S. Okuda, K. F. Aoki-Kinoshita, The GlyCosmos Portal: a unified and comprehensive web resource for the glycosciences, Nat. Methods 17 (2020) 649–650.

[48] Z.L. Klamer, C.M. Harris, J.M. Beirne, J.E. Kelly, J. Zhang, B.B. Haab, CarboGrove: a resource of glycan-binding specificities through analyzed glycan-array datasets from all platforms, Glycobiology 32 (2022) 679–690.

[49] A. Varki, Biological roles of glycans, Glycobiology 27 (2017) 3–49.

[50] J. Lundstrøm, E. Korhonen, F. Lisacek, D. Bojar, LectinOracle: a generalizable deep learning model for lectin-glycan binding prediction, Adv. Sci. (Weinh). 9 (2022) e2103807.

[51] D. Bojar, L. Meche, G. Meng, W. Eng, D.F. Smith, R.D. Cummings, L.K. Mahal, A useful guide to Lectin binding: machine-learning directed annotation of 57 unique lectin specificities, ACS Chem. Biol. 17 (2022) 2993–3012.

[52] X. He, L. Zhao, Y. Tian, R. Li, Q. Chu, Z. Gu, M. Zheng, Y. Wang, S. Li, H. Jiang, Y. Jiang, L. Wen, D. Wang, X. Cheng, Highly accurate carbohydrate-binding site prediction with DeepGlycanSite, Nat. Commun 15 (2024) 5163.

[53] R.C.C. da Silva, R.S. Roldan-Filho, M.A. de Luna-Aragão, R.L. de Oliveira Silva, J.R. C. Ferreira-Neto, M.D. da Silva, A.M. Benko-Iseppon, Omics-driven bioinformatics for plant lectins discovery and functional annotation - A comprehensive review, Int. J. Biol. Macromol. 279 (2024) 135511.

[54] O. Oyelaran, J.C. Gildersleeve, Glycan arrays: recent advances and future challenges, Curr. Opin. Chem. Biol. 13 (2009) 406–413.

[55] O.C. Grant, R.J. Woods, Recent advances in employing molecular modelling to determine the specificity of glycan-binding proteins, Curr. Opin. Struct. Biol. 28 (2014) 47–55.

[56] A. Son, W. Kim, J. Park, W. Lee, Y. Lee, S. Choi, H. Kim, Utilizing molecular dynamics simulations, machine learning, cryo-EM, and NMR spectroscopy to predict and validate protein dynamics, Int. J. Mol. Sci. 25 (2024) 9725.

[57] D. Neumann, C.-M. Lehr, H.-P. Lenhof, O. Kohlbacher, Computational modeling of the sugar-lectin interaction, Adv. Drug Deliv. Rev. 56 (2004) 437–457.

[58] J. Lin, F. Du, M. Long, P. Li, Limitations of phage therapy and corresponding optimization strategies: a review, Molecules 27 (2022) 1857.

[59] J.S. Temme, C.T. Campbell, J.C. Gildersleeve, Factors contributing to variability of glycan microarray binding profiles, Faraday Discuss 219 (2019) 90–111.

[60] B. Donczo, J. Kerekgyarto, Z. Szurmai, A. Guttman, Glycan microarrays: new angles and new strategies, Analyst 139 (2014) 2650–2657.

[61] P. Peluso, B. Chankvetadze, Recent developments in molecular modeling tools and applications related to pharmaceutical and biomedical research, J. Pharm. Biomed. Anal. 238 (2024) 115836.

[62] R.J. Woods, Predicting the structures of glycans, glycoproteins, and their complexes, Chem. Rev. 118 (2018) 8005–8024.

[63] R.J. Woods, M.B. Tessier, Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes, Curr. Opin. Struct. Biol 20 (2010) 575–583.

[64] S.H. Wang, T.J. Wu, C.W. Lee, J. Yu, Dissecting the conformation of glycans and their interactions with proteins, J. Biomed. Sci. 27 (2020) 93.

[65] E. Barbierato, A. Gatti, The challenges of machine learning: a critical review, Electronics (Basel) 13 (2024) 416.

[66] D. Sidak, J. Schwarzerová, W. Weckwerth, S. Waldherr, Interpretable machine learning methods for predictions in systems biology from omics data, Front. Mol. Biosci. 9 (2022) 926623.

[67] D. Bojar, F. Lisacek, Glycoinformatics in the artificial intelligence era, Chem. Rev. 122 (2022) 15971–15988.

[68] H. Li, A.W.T. Chiang, N.E. Lewis, Artificial intelligence in the analysis of glycosylation data, Biotechnol. Adv. 60 (2022) 108008.

[69] A. Yom, A. Chiang, N.E. Lewis, Boltzmann model predicts glycan structures from Lectin binding, Anal. Chem 96 (2024) 8332–8341.

[70] R. Warkentin, D.H. Kwan, Resources and methods for engineering "designer" glycan-binding proteins, Molecules 26 (2021) 380.

[71] M.L.R. Macedo, C.F.R. Oliveira, C.T. Oliveira, Insecticidal activity of plant lectins and potential application in crop protection, Molecules 20 (2015) 2014–2033.

[72] M. Mazalovska, J.C. Kouokam, Plant-derived lectins as potential cancer therapeutics and diagnostic tools, Biomed. Res. Int. 2020 (2020) 1631394.

[73] E.H.E. Konozy, M.E.F.M. Osman, Plant lectin: a promising future anti-tumor drug, Biochimie 202 (2022) 136–145.

[74] M. Nabi-Afjadi, M. Heydari, H. Zalpoor, I. Arman, A. Sadoughi, P. Sahami, S. Aghazadeh, Lectins and lectibodies: potential promising antiviral agents, Cell. Mol. Biol. Lett. 27 (2022) 37.

[75] S.G. De-Simone, D.W. Provance Jr, Lectin affinity chromatography: an efficient method to purify horse IgG3, Methods Mol. Biol 2178 (2021) 301–310.

[76] S. Hu, D.T. Wong, Lectin microarray, Proteomics Clin. Appl 3 (2009) 148–154.