ORIGINAL ARTICLE

Gene sequencing and *in silico* characterization of mannose-binding lectin in South African chicken breeds

Peter Ayodeji Idowu[1] (ID), Takalani J. Mpofu[1] (ID), Oliver T. Zishiri[2] (ID), Ogun Joshua Oluwamayowa[3] (ID), Khathutshelo A. Nephawe[1] (ID), Bohani Mtileni[1] (ID)

[1]Department of Animal Sciences, Faculty of Science, Tshwane University of Technology, Pretoria, South Africa
[2]Department of Genetics, Westville Campus, University of KwaZulu-Natal, Durban, South Africa
[3]Institute of Animal Breeding and Husbandry, University of Kiel, Kiel, Germany

ABSTRACT

Objective: To sequence and characterize the mannose-binding lectin gene of three South African chicken breeds, namely, Potchefstroom Koekoek (PK), Venda (VN), and Ovambo (OV), to ascertain their genetic and immunologic diversity.

Materials and Methods: Total RNA was isolated from hepatic samples, quantified, and reverse-transcribed to generate cDNA. The *MBL* gene was amplified by PCR, confirmed by gel electrophoresis, purified, and sequenced using Sanger sequencing. Sequences were analyzed with FinchTV and submitted to GenBank. Comparative sequences were retrieved from National Centre for Biotechnology Information for multiple sequence alignment and phylogenetic analysis using MEGA. ProtParam and ExPASy were used for physicochemical analysis. Secondary structures were predicted using PDBsum, while tertiary structures were modeled with Swiss-Model, refined by GalaxyWEB, and validated by ProSA. Functional domain analysis, binding site prediction, and ligand interaction studies were also performed.

Results: MBL sequences showed breed-specific differences in protein length, isoelectric points, and thermostability. PK and VN MBLs had acidic *pI* values (< 7), while OV displayed a higher, alkaline *pI*. Conserved Glu–Pro–Asn (EPN) and Trp–Asn–Asp (WND) residues, linked by calcium ions, were identified for mannose-binding. Phylogenetic analysis revealed that PK breeds clustered closely with the White Leghorn (95%), OV clustered nearby PK breeds and Leghorn (82%), while VN clustered more closely with Indian Assel breeds (96%). Two conserved motifs (IPR033990 and IPR001304) were detected. Secondary and tertiary structures revealed predominant random coils in PK and OV, and more alpha-helices in VN. Binding site analysis identified key regions likely involved in immune modulation.

Conclusion: This research reveals variation in *MBL* genes and their immune relevance in South African chicken breeds, offering a basis for breeding strategies.

Introduction

South African chicken breeds are known to have unique characteristics such as resistance to disease and preservation of genetic resources traits suited to local conditions [1]. Among these breeds, Potchefstroom Koekoek (PK), Ovambo (OV), and Venda (VN) are the most predominant breeds used by poultry farmers in South Africa due to their adaptability, resilience, and significance in smallholder farming systems [2,3].

Despite their widespread usage, little or no molecular study has been done on their immunological traits, such as the mannose-binding lectin (*MBL*) protein. Gaining insights into the genetic variability, structural characteristics, and functional properties of *MBL* in these breeds is important for advancing both scientific knowledge and

---

Contact Peter Ayodeji Idowu ✉ ayodejiidowuolu@gmail.com ⬛ Department of Animal Science, Tshwane University of Technology, Pretoria, South Africa.

How to cite this article: Idowu PA, Mpofu TJ, Zishiri OT, Oluwamayowa OJ, Nephawe KA, Mtileni B. Gene sequencing and *in silico* characterization of mannose-binding lectin in South African chicken breeds. J Adv Vet Anim Res 2025; 12(3):941–955.

breeding strategies. Most especially given their exposure to diverse pathogens under free-range and low-input production systems [1,3]. Therefore, characterizing their *MBL* gene could provide insights into the diversity and capacity of South African chicken breeds' innate immune system.

Generally, chickens are known to consist of animal lectins [4,5]. These lectins contribute to antimicrobial defense, cell communication, and pathogen recognition [6,7]. Based on their carbohydrate recognition domains (CRDs), animal lectins can be categorized into three distinct types. This includes p-type, I-type, and c-type [8]. The p-type is found mainly in tissues with a preference to bind to the galactose or N-acetyl galactosamine region of pathogens and is the main factor involved in cell adherence, development, and immune defense [9]. The I-type lectin (galectin) binds primarily to beta-galactoside sugars such as lactose and N-acetyl lactosamine, with a major role in cancer progression, tissue development, and wound healing [9]. C-type lectins, mainly expressed in liver cells, preferentially bind carbohydrates like fucose, mannose, and N-acetylglucosamine. Mannose-binding lectin (*MBL*), a soluble Ca^{2+}-dependent member of this family, plays a key role in pathogen recognition, immune regulation, and clearance of apoptotic cells [10–12]. It specifically binds terminal mannose and other carbohydrate-rich residues and is found on the surface of pathogens like Gram-positive and Gram-negative bacteria and, in some cases, parasites, viruses, yeast, parasites, and mycobacteria [12,13]. Upon binding to these sugar residues, *MBL* neutralizes pathogens by marking them for immune recognition and further initiates the lectin pathway of the complement system [10,14]. This process leads to opsonization and enhanced phagocytosis, which ultimately results in the clearance of the pathogens from the host cell [4,15]. Finally, *MBL* can differentiate between self-cells, non-self-cells, and apoptotic cells to ensure targeted immune responses without harming host tissues [10,14,15].

Bodi et al. [16] reported that *MBL* exists as a trimeric structure, having a molecular weight of approximately 96 kilodaltons, composed of three identical 32 kilo Dalton subunits. Each subunit contains a collagen-like domain, an N-terminal nitrogen-rich cross-linking segment, and a C-terminal CRD [9,10,15]. These domains assemble into a classical triple-helical structure [10]. In chickens, *MBL* is capable of forming multiple oligomeric states, ranging from dimers to hexamers [15,17], with the homotrimer identified as the fundamental building block [18]. The trimer consists of three identical polypeptide chains, forming a collagen-like triple helix with globular lectin domains at the C-terminal end [19]. By binding a calcium ion, each lectin domain can specifically interact with sugars like N-acetyl-D-glucosamine, mannose, N-acetyl-mannosamine, fucose, and glucose [20].

Each chicken *MBL* (c*MBL*) subunit is organized into several structural regions: a collagen-like region, a neck domain, a CRD, and a linker region [9,10,21]. The collagen-like region, characterized by Gly-XY repeats, is critical for trimer assembly and interaction with *MBL*-associated serine proteases to induce complement pathway activation [20,22,23]. Connecting the collagen-like domain and the CRD, the neck region contributes to the stability of the trimer [10,15], while the CRD contains a conserved carbohydrate-binding site for sugar recognition [24]. Oligomerization of *MBL* increases the number of CRDs, enhancing multivalent ligand binding [25]. Also, most *MBL* variants possess a linker region between the neck and CRD, providing flexibility for optimal carbohydrate binding [26].

Chicken *MBL* contains an EPN (Glu-Pro-Asn) motif in its CRD, enabling binding to D-mannose, L-fucose, and GlcNAc [12,27]. Ligand-binding motifs are highly diverse across species; for example, saltwater clam (*Glycymeris yessoensis*) lectins exhibit motifs such as Glu–Pro–Asp (EPD), Gln–Pro–Gly (QPG), Gln–Pro–Ser (QPS), Tyr–Pro–Gly (YPG), and Tyr -Pro -Thr (YPT) [28]. *MBL* also interacts with other immune molecules like dendritic cells, pentraxins, and the serum amyloid p component, reinforcing its role in bridging innate and adaptive immunity [29,30].

Therefore, this study aimed to sequence and characterize the *MBL* gene in selected South African indigenous chicken breeds using gene sequencing and *in silico* approaches. Through computational analyses, the study explored the physicochemical properties, subcellular localization, functional domains, evolutionary relationships, secondary and tertiary protein structures, and potential ligand-binding sites of the cMBL protein. The objective was to gain insights into the immunological diversity of MBL and its role in breed-specific indigenous South African chicken breeds. By promoting sustainable approaches to poultry farming and reducing antibiotic dependence, these findings not only enhance poultry health and productivity but also contribute to achieving the UN Sustainable Development Goals of zero hunger and good health and well-being.

Materials and Methods

Ethical approval

The study was approved by the Animal Research Ethics Committee of the Faculty of Science, Tshwane University of Technology (AREC2021/10/002; date: October 18, 2021), and conducted in accordance with ARRIVE guidelines and local animal welfare regulations.

Animal tissue and total RNA extraction

Liver samples from three South African chicken breeds were collected and stored at –80°C in liquid nitrogen. Total RNA was extracted, quantified, and quality-checked before synthesizing cDNA using a first-strand synthesis kit. qPCR was conducted using two primer sets targeting fragments of 544 bp (5'-GAT AAG CCG GAA AAC CCT GAA-3' / 5'-GTT ACA ACA ATT CCA CGT TCT CCT-3') [31] and 835 bp (5'-GGT AAA GGT GCT GAT CTG TGG-3' / 5'-TGA GAG AAG AAA GTT GGA TTT-3') [32].

PCR amplification and sequencing

The *MBL* gene was sequenced following PCR amplification, product purification, Sanger sequencing, and sequence analysis. Genomic DNA was extracted and amplified using NEB OneTaq 2X Master Mix with Standard Buffer (New England Biolabs, M0482S). Each 20 µl PCR reaction contained 10 µl of Master Mix, 1 µl of genomic DNA (20 µg/µl), 1 µl each of forward and reverse primers (10 µM), and 7 µl of nuclease-free water (E476). Thermal cycling included an initial denaturation at 94°C for 5 min, 35 cycles of 94°C for 30 sec, 50°C for 30 sec, 68°C for 60 sec, and a final extension at 68°C for 10 min, followed by storage at 4°C until further use.

Gel electrophoresis and visualization of PCR products

PCR product integrity and size were evaluated on a 1% agarose gel stained with EZ-Vision® Bluelight DNA Dye and visualized using a gel documentation system to confirm successful amplification.

PCR product purification using the ExoSAP method

PCR products were purified enzymatically using the ExoSAP method. For each reaction, 10 µl of amplified DNA was combined with 2.5 µl of an ExoSAP mixture containing Exonuclease I (20 U/µl; NEB M0293L) and Shrimp Alkaline Phosphatase (1 U/µl; NEB M0371). The mixture was incubated at 37°C for 15 min to remove residual primers and dephosphorylate unused nucleotides, followed by enzyme inactivation at 80°C for 15 min.

Sanger sequencing reaction and post-sequencing cleanup

Purified PCR fragments were sequenced using the BrilliantDye™ Terminator Cycle Sequencing Kit v3.1 (Nimagen, BRD3-100/1000) following the manufacturer's protocol. Sequencing products were then cleaned using the ZR-96 DNA Sequencing Clean-up Kit (Zymo Research, D4053) to eliminate unincorporated dye terminators and salts.

Capillary electrophoresis and sequence data analysis

Sequencing analysis was performed on an Applied Biosystems ABI 3500XL Genetic Analyzer equipped with a 50 cm capillary array and POP-7 polymer. The resulting chromatograms were visualized and interpreted using FinchTV, a freely available software designed for high-quality electropherogram viewing. To support comparative analysis, the corresponding coding sequences of the *MBL* gene were retrieved from other chicken *MBL* sequences available in the National Centre for Biotechnology Information (NCBI) Protein Database (https://www.ncbi.nlm.nih.gov/protein) [33]. The derived sequence data from this study have been submitted to GenBank and are available with the following accession numbers: PP782170 (PK), PP782171 (OV), and PP782172 (VN).

Sequence analysis

The Expasy server (https://web.expasy.org/translate) [34] was used to convert the nucleotide sequences to amino acid sequences (proteins), and the longest open reading frame (ORF), which is highlighted in red, was selected for this study. This ORF with the highest red is known to contain the full protein-coding sequence, with the longest length and position, making it suitable for further protein analysis [34].

Prediction of amino acid sequences and functional characterization

Amino acid sequences were inferred from nucleotide sequences using the ExPASy Translate tool. The resulting protein sequences were analyzed with ExPASy Protein tools to determine their properties, including consensus motifs for chicken *MBL*. Functional domains were mapped, and the protein's ontology and classification were predicted using the InterPro server, which catalogs homologous protein domain families (https://www.ebi.ac.uk/interpro/) [35,36].

Multiple sequence alignment and phylogenetic analysis

For evolutionary analysis, seven *Gallus gallus* chicken mannose-binding lectins *cMBL* with one *MBL* from a plant as an outgroup were retrieved from the NCBI with accession numbers AF231714.1, KF469209.1, KU378610.1, KU378616.1, KF469210.1, KF469208.1, JF717877.1, and KC329532.1 (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), with three generated protein sequences from this study, PP782170 (PK), PP782171 (OV) and PP782172 (VN), making a total of ten [10] c*MBL*. Phylogenetic trees were produced in MEGA software 11.0.21 using the maximum likelihood test with a bootstrap test of 1,000 replicates [37].

Physicochemical property analysis

The ProtParam tool (http://web.expasy.org/protparam/ [34]) was used to determine physicochemical characteristics of *MBL* from South African chicken breeds, including molecular weight (MW), aliphatic index (AI), grand average of hydropathy (GRAVY), isoelectric point (*pI*), and conserved signal peptide sites (SPCS).

Subcellular location and solubility prediction

The probable subcellular localization of the protein was assessed using CELLO (http://cello.life.nctu.edu.tw/[38]). Protein solubility and hydrophobic regions were evaluated with SOSUI (http://harrier.nagahama-i-bio.ac.jp/sosui/ [39]), where hydrophobic segments were annotated as potential transmembrane regions. Signal peptide cleavage sites were predicted using the TOPCONS server (http://topcons.cbr.su.se/pred [40]).

Predictions and validations of the secondary and tertiary structures of proteins

To gain insights into protein function, the secondary structures were predicted using the PDBsum database (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) [41], while tertiary structures were modeled using the SwissModel tool for homology study (https://swissmodel.expasy.org/) [42]. These models were further refined with the GalaxyWEB refiner tool [43], which uses ab initio methods to refine loop and terminal regions. The predicted structures were validated with PDBsum, which identified the amino acid sequences involved in forming secondary structures, such as α-helices, β-sheets, coils, and loops. This provides insights into both structural and functional aspects of *MBL*. Lastly, ProSA (Protein Structure Analysis) was used to predict potential errors in the 3D models provided in PDB format [43,44].

Prediction of binding sites

The Galaxysite tool [45] was employed to predict possible ligand-binding sites within the protein's tertiary structure. This tool does not only predict the most likely binding pockets but also suggests potential ligand molecules that may interact with these sites. These predicted ligand interactions provide valuable insights for future therapeutic research and functional characterization of the *MBL* studies.

Results

Primary structure and subcellular location of selected South African chicken breeds

Table 1 shows the MW, which was 27, 27 and 33 kDa, and the SPCS were 29–30, 28–29 and 29–30 for the PK, VN and OV breeds. The number of amino acids was 251, 254 and 313 and the aliphatic index observed was 77.73, 79.49 and 79.74 for the PK, VN and OV breeds. The instability index observed were 30.88, 29.57 and 24.18 and the GRAVY of the protein sequences predicted were −0.520, −0.468 and −0.415 for PK, VN and OV breeds. This study further observed that both the PK and VN breeds are located in the periplasmic subcellular with an isoelectric *pI* of 5.85. In contrast, the OV breed is located in the cytoplasmic inner membrane, with a *pI* of 8.90.

Multiple sequence alignment of cMBL of selected South African chicken breeds

Figure 1 shows the multiple sequence alignment of the *MBL* gene. The contoured area shows the site of the signal peptide. All the chicken breeds sampled in this study have gaps at the 10th, 165th to 168th, and 274th positions. Also, PK, VN, and OV have 6, 3, and 3 insertions in all 317 protein sequences with respect to other chicken breeds. This shows that these three breeds are closely related to each other. In the three breeds studied, the MBL protein sequences were highly conserved.

Prediction of the functional domain of the South African cMBL

Two functional domains were identified using the InterPro server (https://www.ebi.ac.uk/interpro/), IPR033990 and IPR001304, in the amino acid sequence of the chicken *MBL* (Fig. 2). The two functional domains, IPR033990 (collectin domain), also called c-type lectin domain (CTLD), and the second, IPR001304 (c-type lectin domain), were observed in all the chicken breeds sampled. The

Table 1. Analysis of the primary structure and physicochemical properties of MBL from South African chicken breeds using the ProtParam server.

Chicken Breeds	SPCS	No aa	MW	*pI*	GRAVY	Instability Index	Aliphatic -Index
PK	29–30	251	27092.81	5.85	−0.520	30.88	77.73
VN	28–29	254	27362.16	5.85	−0.468	29.57	79.49
OV	29–30	313	33572.29	8.90	−0.415	24.18	79.74

SPCS: Signal Peptide-Conserved Sites, No. aa: Number of amino acids, GRAVY: Grand Average of Hydropathicity: *pI*: Isoelectric point. MW: Molecular weight. PK=Potchefstroom Koekoek breed, VN= Venda breed, and OV=Ovambo breeds. Chromosome number: 6
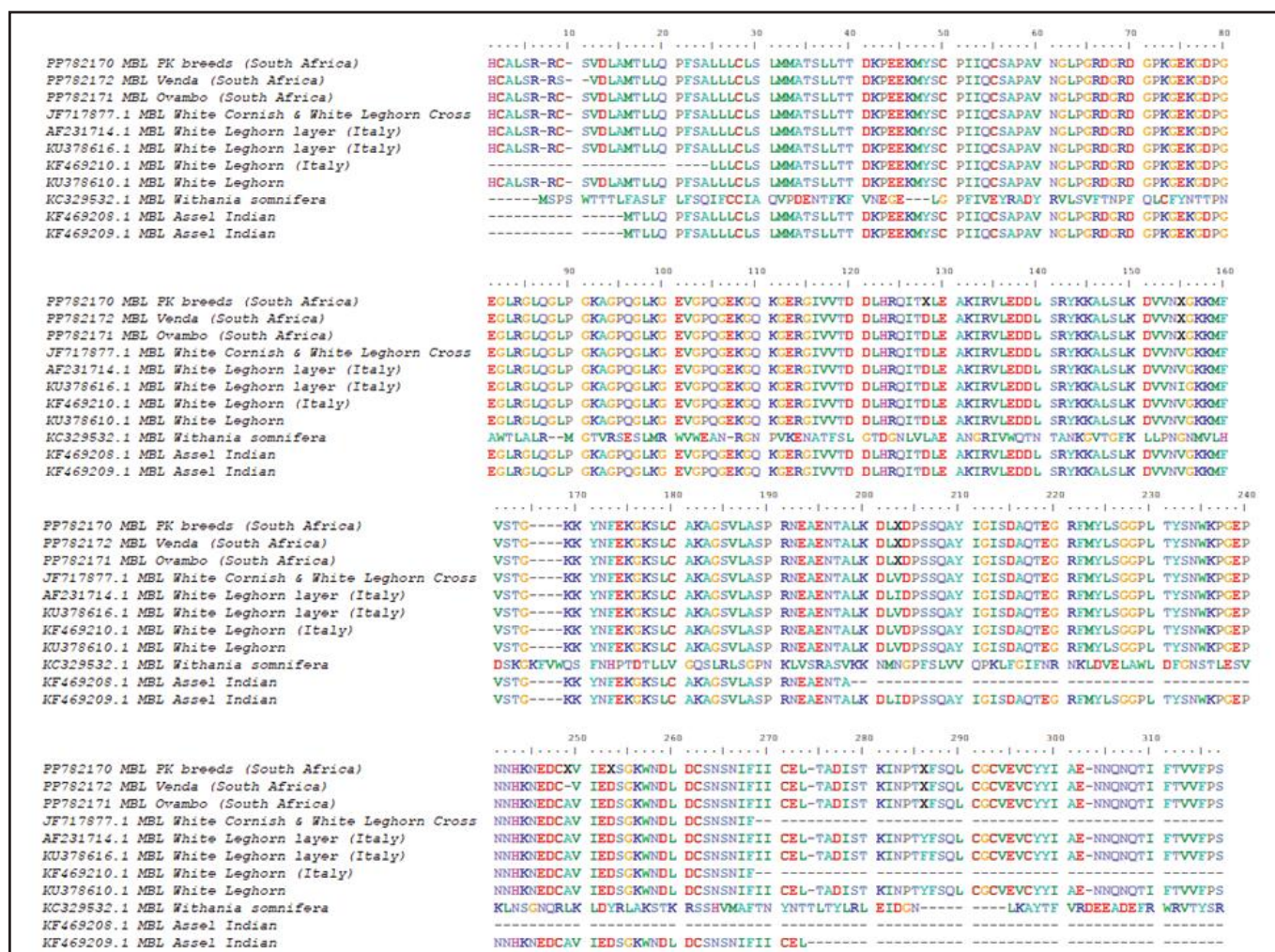
Figure 1. Alignment of mannose-binding lectin protein sequences from chicken breeds. Sequences were aligned using MEGA 11 by Clustal W; Residues that are identical or share similar properties are highlighted using consistent color.

representative domains are collagen and C-LECT-2 and the domain observed in this study is c-type lectin-like.

Phylogenetic analysis of the MBL protein sequence of South African chicken breeds

To investigate the relationships and variation among c*MBL* genes, a phylogenetic tree was generated using MEGA 11. The maximum likelihood method was applied to infer evolutionary relationships. The computed data indicated that there are nine different clusters, including the outgroup. South African breeds PK and Ovambo clustered closely with European lines (White Leghorn and White Cornish White Leghorn Crossbred). This suggested shared genetic ancestry despite geographic separation. The Indian breeds Assel and White Leghorn layer formed a distinct regional cluster, suggesting localized evolutionary lines. The VN breed from South Africa occupied a separate cluster, highlighting significant divergence from other African and European breeds. This may reflect unique selection pressures or genetic isolation. The plant species *Withania somnifera MBL* was included as a reference for rooting the phylogenetic tree.

Predictions and validations of the secondary and tertiary structures of proteins

The PROCHECK tool in PDBsum was used to evaluate the predicted tertiary structures of MBL proteins from three chicken breeds (Table 2, Fig. 1). The PK breed cMBL contains 15.5% β-strands, 27.9% α-helices, and 56.6% other structural elements. The OV breed has a similar composition with 15.4% β-strands, 28.3% α-helices, and 56.3% remaining components. The VN breed differs slightly, comprising 17.6% β-strands, 20.8% α-helices, a short helix of 1.3%, and 60.4% other elements.

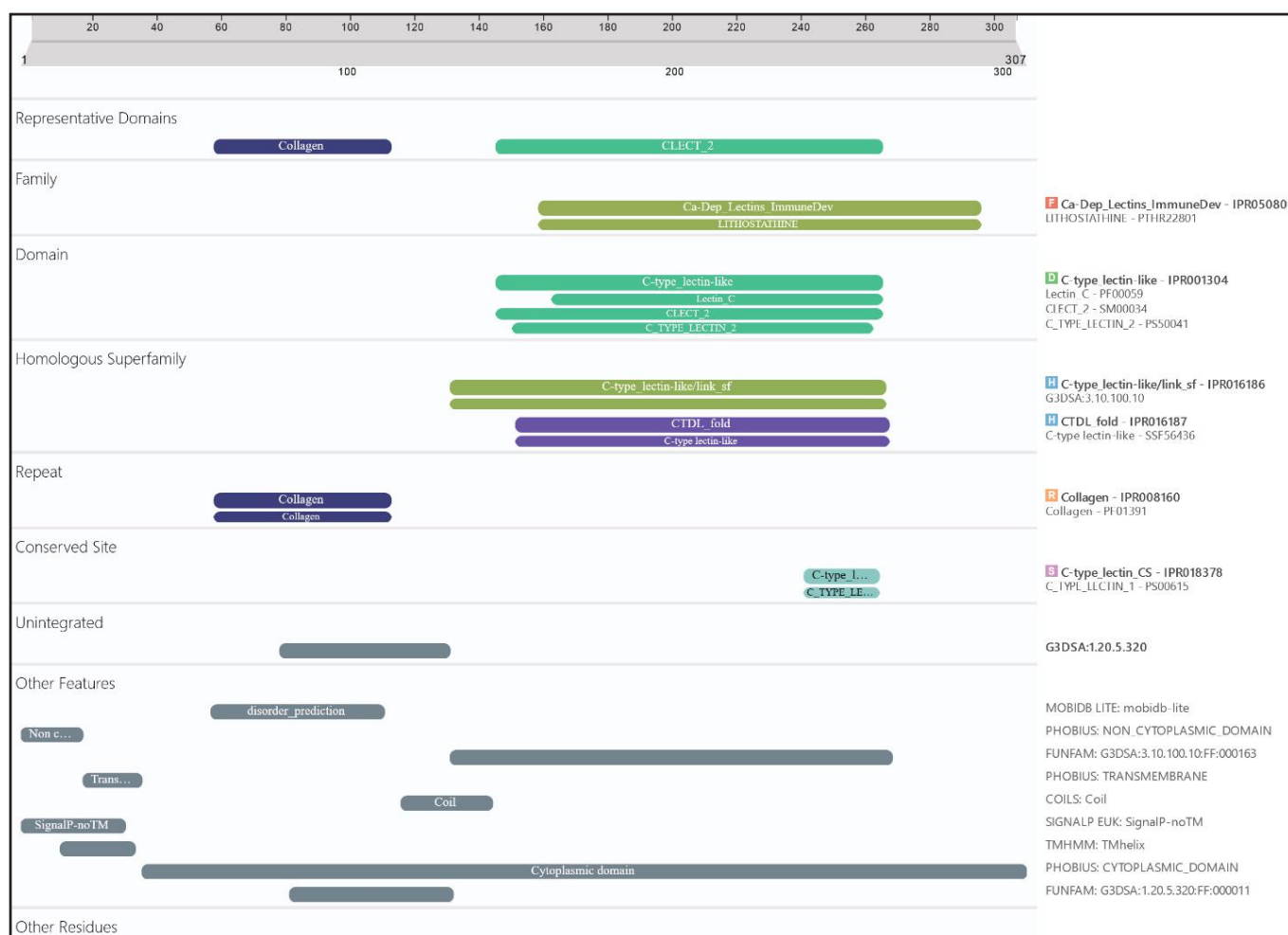Based on the amino acid sequence of the protein, PROCHECK predicts its secondary structure and compares

Figure 2. Prediction of the functional domain of South African chicken mannose-binding lectin observed by the InterPro server.

Table 2. Prediction of secondary structure and possible compositions of the three indigenous chicken breeds from South Africa.

Breeds	Strand	Alpha helix	Others	Total residues
PK	39 (15.5%)	70 (27.9%)	142 (56.6%)	251
OV	39 (15.4%)	72 (28.3%)	143 (56.3%)	254
VN	55 (17.6%)	65 (20.8%)	193 (60.4% + 1.3% Helix)	331

PK: Potchefstroom Koekoek; OV: Ovambo; VN: Venda.

it with the secondary structure shown in the crystallo-graphic data. The secondary structure elements that are projected to occur are often beta strands, loops, and alpha helices. The procheck result shows the overall quality of the protein structure (Fig. 4).

The stereochemical quality of the predicted MBL protein structures was assessed using Ramachandran plot analysis, which showed that 92.6%, 93.7%, and 95.3% of residues fell within the most favored regions for the PK, OV, and VN breeds, respectively. Additionally, ProSA, a web-based tool for detecting errors in protein 3D structures, was employed. The proteins yielded z-scores of −5.91, −6.36, and −6.16, indicating high-quality models (Fig. 5). The high proportion of residues in favored regions confirms strong agreement between predicted and observed secondary structures, and the z-scores indicate the models are reliable and acceptable [46].

Prediction of binding sites

The galaxy site prediction tool revealed possible binding sites and potential interacting ligands (Table 3). Figures 6, 7, and 8 show the possible noncovalent interactions (such as hydrogen bonds, salt bridges, and hydrophobic bonds) for the predicted ligands of the MBL proteins.
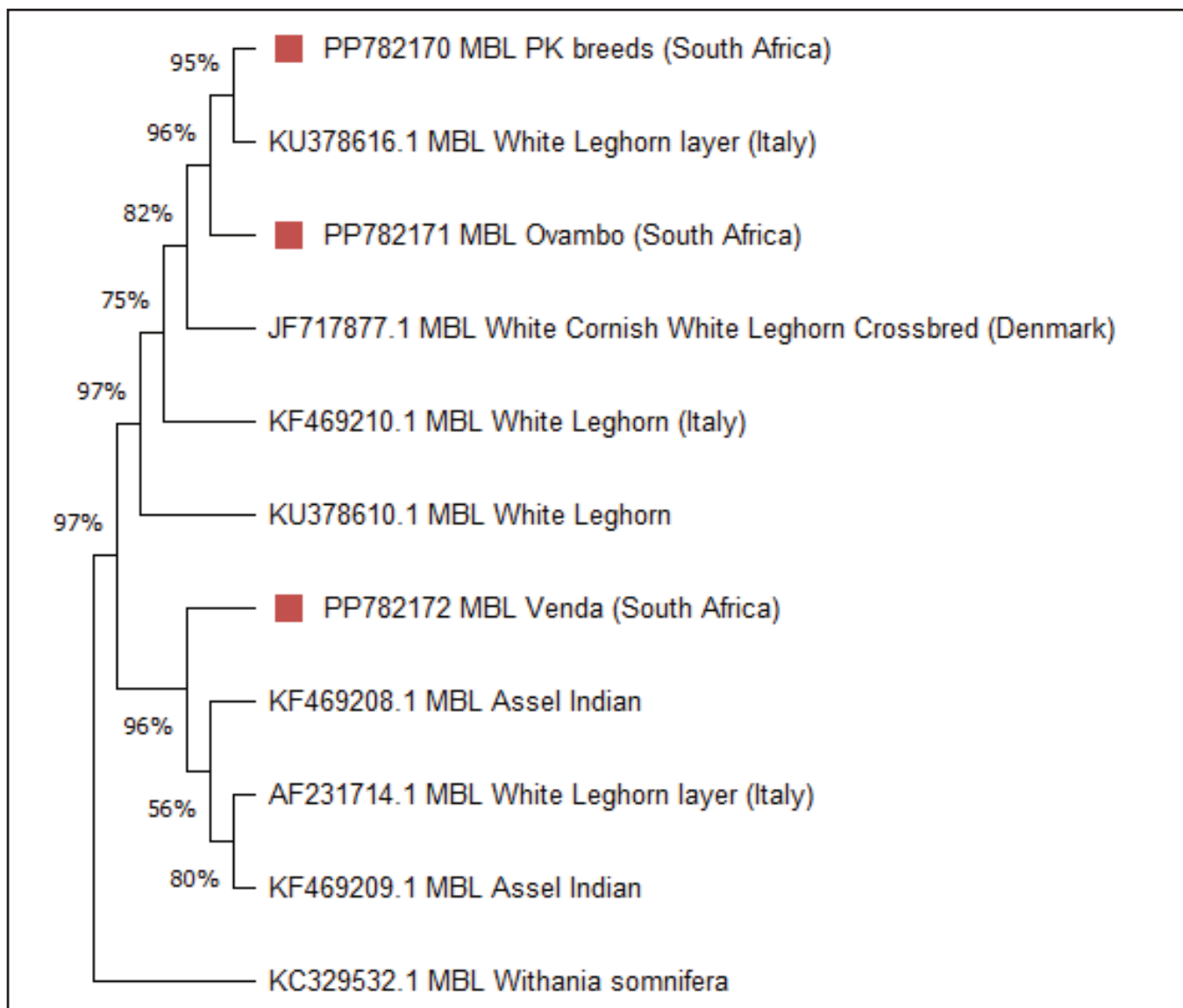
Figure 3. Maximum-likelihood phylogenetic tree based on protein sequences of chicken (*Gallus gallus*) from various breeds, with 1000 bootstrap replicates. The protein sequences of different breeds were retrieved from the NCBI database, alongside three breeds from this study with accession numbers PP782170 (PK), PP782171 (OV), and PP782172 (VN). MBL *Withania somnifera* was used as the reference/outgroup.

Discussion

The isoelectric *pI* of a protein reflects the pH at which it carries no net electrical charge [47]. In this study, the MBL protein of the OV breed had an alkaline *pI* (8.90), while the PK and VN breeds had acidic *pI* (5.85). This variation indicates differences in their net charge under physiological conditions [48]. These differences suggest that the OV MBL may be less mature or structurally distinct compared to the PK and VN MBL. Also, OV and *pI* are negatively charged in acidic environments and positively charged in alkaline environments, while PK and VN are positively charged in acidic environments and negatively charged in alkaline

environments [49]. Previous study has reported that most premature proteins tend toward alkalinity, whereas mature proteins generally exhibit more acidic *pI* distributions [50]. This could be attributed to post-translational modifications or differences in amino acid composition [51]. The relatively higher *pI* of the OV MBL protein indicates a lower proportion of acidic residues (aspartic acid and glutamic acid) or a higher content of basic residues (lysine and arginine) [52]. In contrast, the lower *pI* values in PK and VN suggest these proteins are more negatively charged [48]. These physicochemical differences may influence how the MBL proteins interact with pathogens, ligands, or immune components across the chicken breeds.
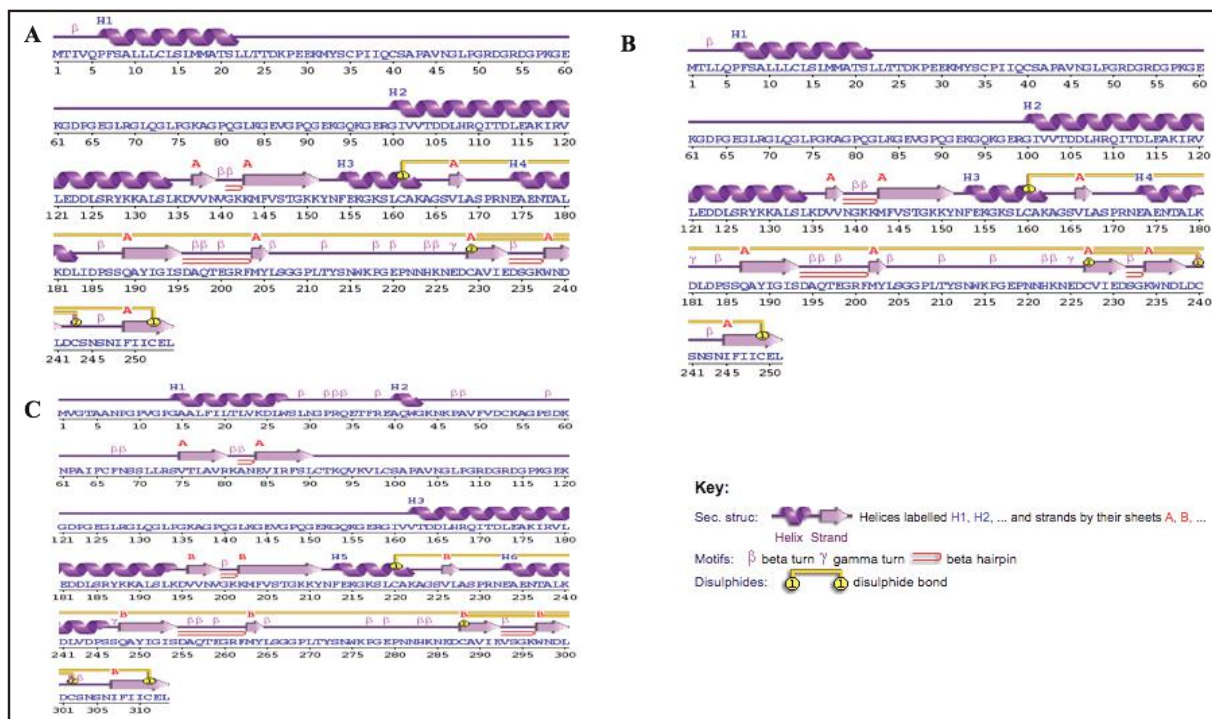
Figure 4. The PROCHECK tool from PDBsum predicted the secondary structures of the proteins. PK (A), OV (B), and VN (C).
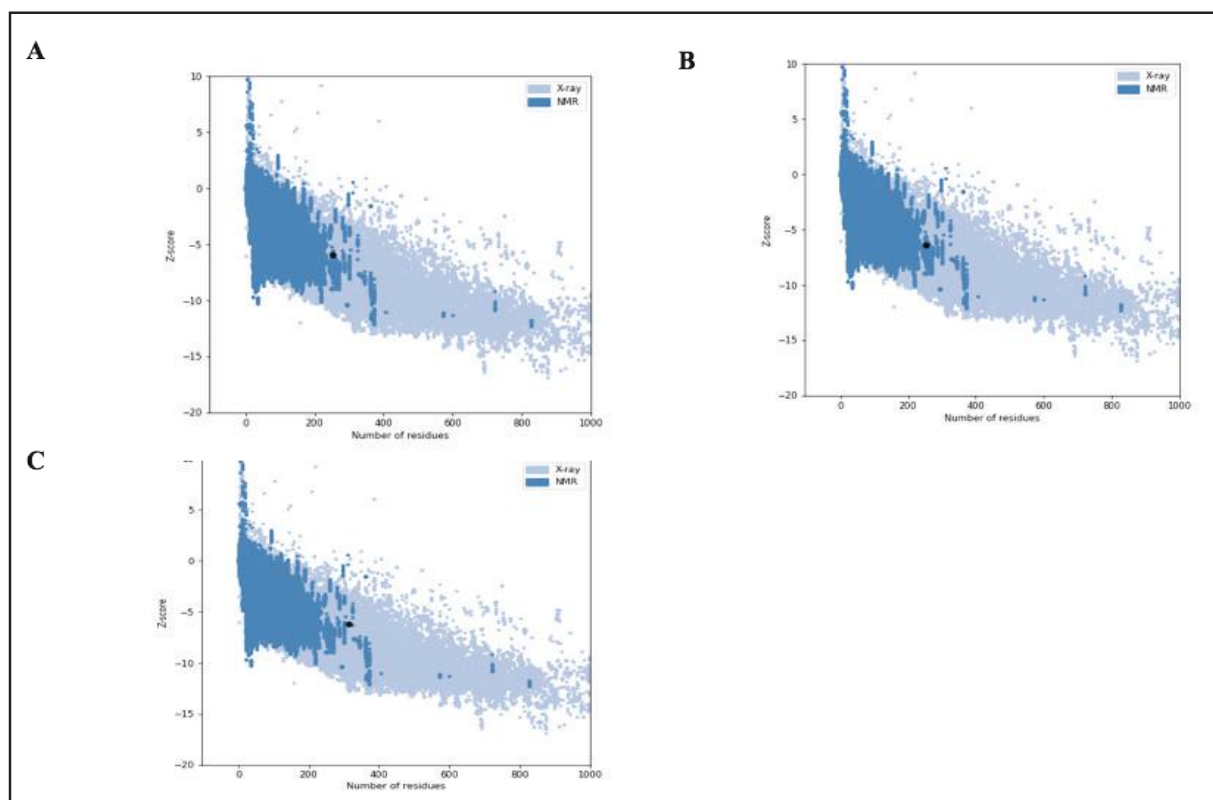

Figure 5. ProSA plot with different *z* scores of PK (A), OV (B), and VN (C). The black dot represents the position of the protein structure in comparison with the standard parameters for proteins of an equivalent size.

A significant correlation has been observed between a protein's isoelectric point (*pI*) and its subcellular localization [50]. The acidic *pI* values of PK and VN MBLs align with their predicted localization in inner membrane compartments [53]. This environment typically favors hydrophilic and acidic proteins, which support solubility and interactions with other biomolecules [54]. On the other hand, the alkaline *pI* of the OV MBL protein suggests greater hydrophobicity, which potentially facilitates stronger interactions with membrane lipids [55] and enhances signaling stability and structural anchoring [53,56]. These variations are likely driven by breed-specific evolutionary changes, such as amino acid substitutions, insertions, or deletions within the *MBL* gene [33].

Table 3. Binding sites and potential ligands of the MBL protein predicted by the Galaxy tool for the three chicken breeds.

Ligand Name	Binding sites
Alpha-D-glucopyranose (GLC)	187Q 215K 218E 222H 225E 236N 237D 238L
2-acetamido-2-deoxy-beta-D-galactopyranose (NGA)	187Q 215K 218E 220N 222H 225E 236N 238L
alpha-L-fucopyranose (FUC)	218E 222H 225E 236N 237D 238L

The MW of the cMBL protein observed in the present study for the PK, VN, and OV were 27, 27 and 33.5 kDa, respectively. These values are close to the range of the theoretical MW, which ranges from 25 to 27 kDa for processed MBL using native mass spectrometry [57]. Also, OV MW is very close to the MW of 32 kDa reported by Ulrich-Lynge et al. [12]. The study of Zhang et al. [57] also reported that the molecular weight of Ross Broiler chicken breeds was 26 kDa. The differences observed in South African chicken breeds may suggest breed-specific modifications. In other species, the molecular weights of human MBL, pumpkin [*Cucurbita pepo* MBL and wild garlic (*Allium ursinum*)] MBL were 31 kDa, 22.6 *kDa,* and 14.83 kDa, respectively [33,58,59]. Variations in the MW of the MBL protein in this study could be attributed to species specificity and glycosylation [54]. Nevertheless, differences in the MW of MBL in comparison to other chicken breeds sampled could reflect the underlying genetic diversity that influences immune system functionality and overall health [4,16]. Also, this variation could be a result of a single-nucleotide polymorphism, which affects the protein's structure and function [60,61]. In detail, variations in the promoter region and exon 1 of the *MBL* gene cause differences in *MBL* oligomerization, which subsequently impact the molecular weight of the protein. It is important to know that the higher the
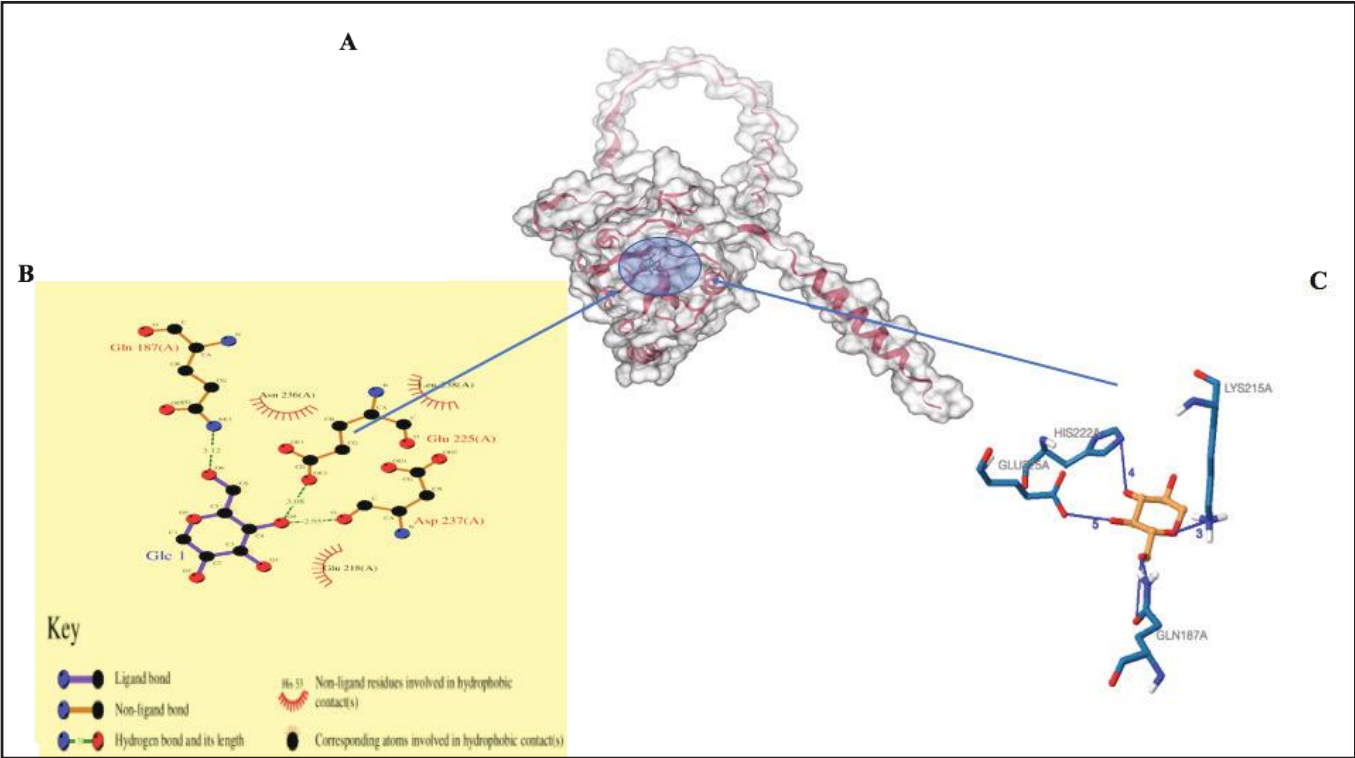

Figure 6. Three-dimensional structure of MBL protein predicted by Swiss Model. The interaction within the binding site with Alpha-D-glucopyranose (GLC) is predicted by the galaxy prediction tool (A). The ligand interaction plots (B). The interaction chains predicted by the Protein Interaction Ligand Profile tool (C).

Figure 7. Three-dimensional structure of MBL protein predicted by Swiss Model. The interaction within the binding site with 2-acetamido-2-deoxy-beta-D-galactopyranose (NGA) is predicted by the galaxy prediction tool (A). The ligand interaction plots (B). The interaction chains predicted by the Protein Interaction Ligand Profile tool (C).
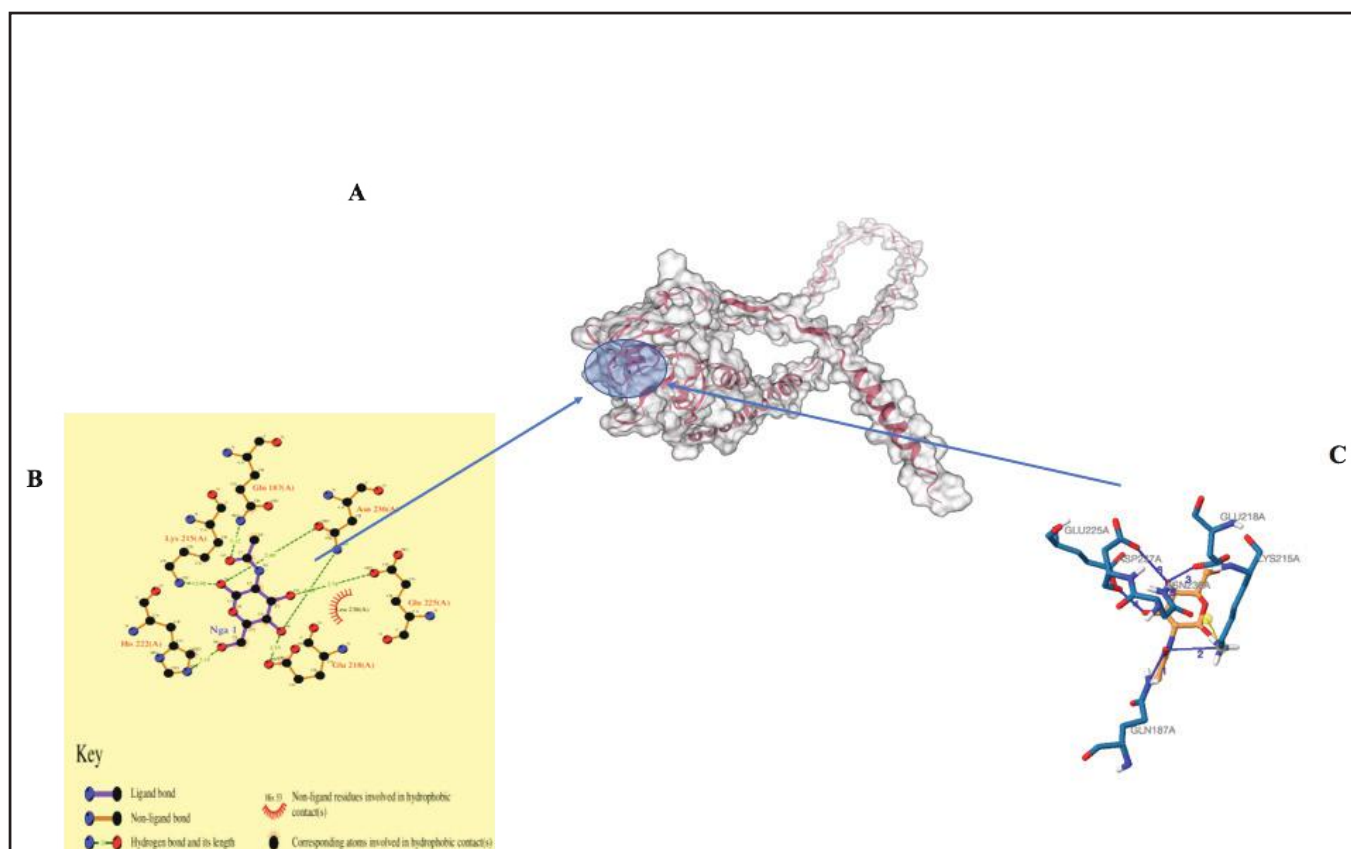
molecular weight, the greater the degree of polymerization [61]. This will, in turn, influence the functional capacity of *MBL* in pathogen recognition and complement system activation [62].

The phylogenetic analysis of the *MBL* gene in South African chicken breeds depicts some level of variation of the *MBL* gene. The PK and Ovambo breeds show close genetic affinity with White Leghorn and White Cornish crossbreds (European lines). This clustering could reflect historical gene flow or a similar *MBL* gene introduced during breeding programs [37]. Such genetic conservation suggests that these South African indigenous breeds could possess immune characteristics related to the commercial chicken breeds. In contrast, the VN breed exhibits clear genetic separation from PK and OV, clustering instead with Indian native breeds (Assel). This divergence may reflect unique local adaptations, limited introgression from commercial breeds, and the preservation of ancestral genetic signatures [63]. Also, adaptation to similar environmental situations and the influence of uncontrolled mating systems in indigenous chicken populations may have influenced the c*MBL* gene variation observed in VN

[30]. Consequently, each protein could have evolved due to differences in breed response to diseases, thus leading to antigenic variation [64]. These findings align with broader studies of South African indigenous chickens, which report moderate to high genetic differentiation among local lines and emphasize the presence of multiple maternal lineages, including those tracing back to the Indian subcontinent [65]. This study suggests that OV and VN breeds are genetically distant from each other, explaining the diversity within South Africa's chicken genetic resources.

The functional domain called collectin domain IPR033990 (CTLD) was found at positions 141-251, 142-253, and 201-313 in the MBL protein sequence of PK, VN and OV chicken breeds. Similar collectins are mostly found in human collectin, lung surfactant protein, liver collectin, and MBL [59]. They can bind carbohydrates on the surface of the pathogen, necrotic or apoptotic cells, and allergens [10,59]. Also, they mediate activities such as phagocytosis [61], identifying the high-rich mannose region and triggering the killing of cells [59].

The functional domain IPR001304 (c-type lectin) was found in the regions 133–250, 132–253 and 191–312 in
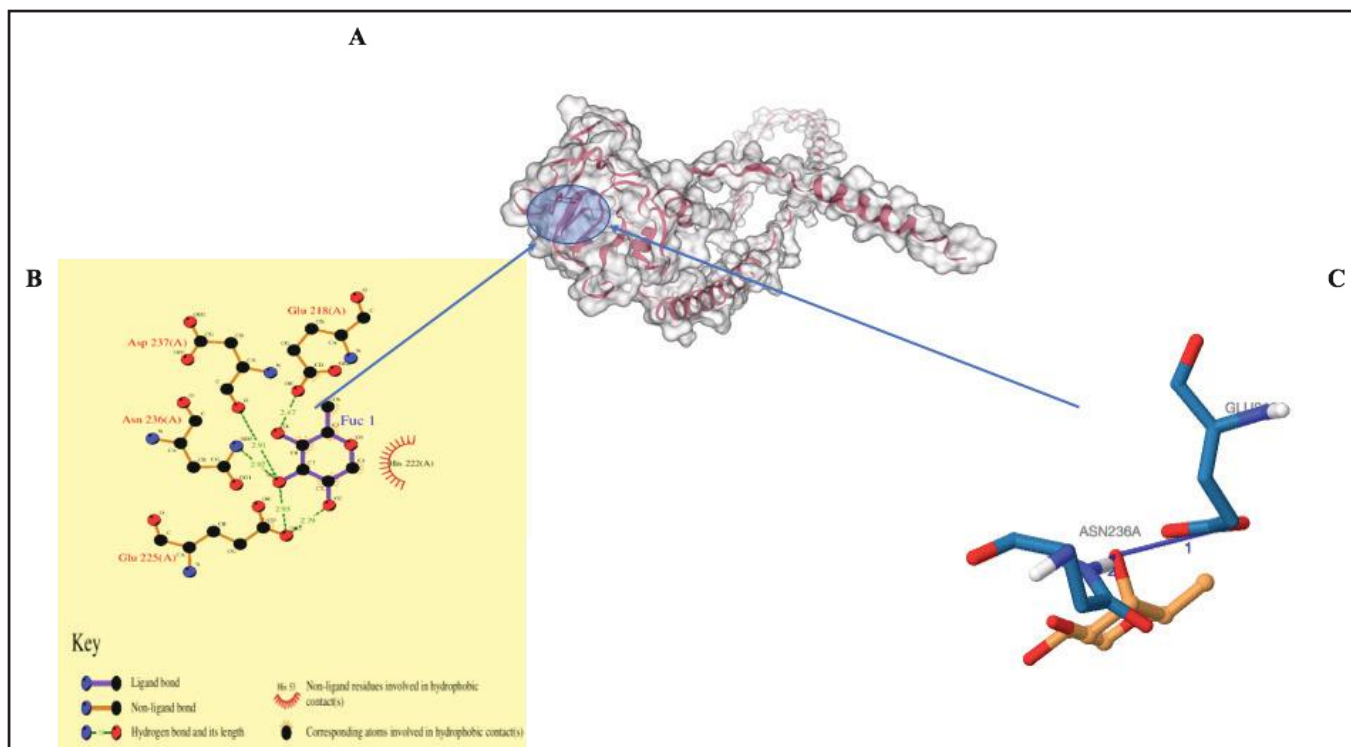
Figure 8. Three-dimensional structure of MBL protein predicted by Swiss Model. The interaction within the binding site with alpha-L-fucopyranose (FUC) is predicted by the galaxy prediction tool (A). The ligand interaction plots (B). The interaction chains predicted by the Protein Interaction Ligand Profile tool (C).

the cMBL protein sequence of the PK, OV and VN chicken breeds, respectively. The c-type lectin is first characterized in some animal lectins, known to be a Ca^{2+}-dependent recognition domain, also known as the CTLD, consisting of about 110–130 residues. It consists of four perfectly conserved cysteines and two disulfide bonds. Both domains have two representative domains, namely collagen and C-LECT-2. The conserved regions in several animal lectins mostly consist of a Ca^{2+}-dependent carbohydrate recognition domain that is shared by several distinct protein families [6,8,10]. The two functional domains are later divided into subdomains. The collectin domain has both c-lect and collectin-like, while the c-type has lectin 2, lectin c, and c-lect 2. The homologous family is the CTDL fold IPR016187 and the lectin-like c-type IPR016186. The MBL is among the few c-type lectin families that have the c-type lectin domain and the conserved domain of Ca^{2+}, which binds to the sugar or mannose region [25].

The CRD observed in the present study shows that the cMBL protein contains all the sequence characteristics of C-type lectin with great homology with other cMBL from a previous study [58]. The four cysteines (Cys) that maintain the distinctive double-loop structure are the two most significant structural components of C-type lectins [19,66]. The c-type lectins consist of two structural elements. The

first is the four cysteines with the aim of stabilizing the double-loop nature (Cys170-Cys184, Cys193-Cys213) and the 161-EPN-163, 171-WND-181 motifs essential for ligand binding under the supervision of a calcium ion present in the binding pocket region of MBL, which is important for the specificity of MBL [67]. This finding aligns with several studies [5,27] regarding the cysteine region present in cMBL under the influence of Ca^{2+}-dependent binding affinity.

In this study, the protein secondary structure was categorized into four types: α-helix, extended strand, β-turn, and random coil. For the PK breed, α-helices were the most prevalent, followed by random coils, extended strands, and β-turns. In contrast, for the OV and VN breeds, random coils were dominant, with α-helices, extended strands, and β-turns occurring in decreasing order.

The impact of the electrostatic bulkiness of the R group could result in coil formation [67]. Jimenez et al. [33] observed similar phases in a study on MBL structure in *Allium* species.

Understanding protein function requires accurate prediction of both secondary and tertiary structures, which is essential for identifying functional sites and protein–ligand interactions. In this study, homology modeling was used due to its effectiveness in predicting three-dimensional

protein structures from available templates [68]. Model quality was assessed using the *z*-score, which reflects the overall structural reliability and conformational energy. All three cMBL protein models exhibited negative *z*-scores, with their positions in the blue region of the plot, indicating high-quality structural models [69]. Regarding the protein model, the structural component of a protein is directly related to its function and the internal underlying force, depending on its interaction with other molecules [70]. Identifying binding sites on proteins is essential in understanding their function, interactions, and potential roles in various biological processes [71]. Glycosidic bonding is of two types: [1] the N-linkage, which occurs when the asparagine side chain attaches nitrogen atoms to the sugar surface, and [2] the O-linkage, which occurs when the serine side chain attaches oxygen atoms to the sugar surface. The MBL region proposed that Asn236, Leu225, Glu218, Gln187, and Asp237, with the ligand name Glc, are the major residues involved in the binding of the mannose region. The major proteins that occupy the glycosylation region that are involved in binding are the Glutamic acid Glu187, Asp237, and Gln187. Also, there is a projection that Glu218, Asp236, Gln187, His222, Lys215, and Leu236, with the ligand name Nga, are the residues observed to be involved in the binding of the mannose region to the MBL region and occupy the glycosylation region except for Leu236. Lastly, in the MBL region, Asp237, Asn236, Glu225, and His222 with ligand name Fuc1 were the main residues involved in carbohydrate binding and are the proteins found in the glycosylation region, except for His222. The Asn and Asp have been reported to be important residual proteins for carbohydrate recognition [62]. Therefore, three ligand names were predicted as the site where mannose binds to lectin (glycosylation) in this study. These ligand areas were also observed in a previous study [27,28].

Conclusion

This study successfully combined gene sequencing and *in silico* characterization to investigate the *MBL* gene in three indigenous South African chicken breeds. Using computational tools. The physicochemical properties, functional domains, subcellular localization, secondary and tertiary structures, and evolutionary relationships of the cMBL protein were analyzed. The results revealed breed-specific structural features and conserved functional motifs, revealing the immunological significance of MBL in these local breeds. Understanding the genetic and structural diversity of *MBL* contributes valuable insights into breed-specific traits with implications for improving poultry health and productivity. Furthermore, this study supports sustainable livestock development goals by highlighting the potential of indigenous breeds in enhancing disease resilience without over-reliance on antibiotics. To understand the interaction between the protein and sugar region (mannose), protein modeling will be recommended for future study to elucidate the immunomodulatory roles of MBL across different chicken populations.

List of abbreviations

AI, aliphatic index; AREC, Animal Research Ethics Committee; ARRIVE, Animal Research: Reporting of In Vivo Experiments; cMBL, chicken mannose-binding lectin; CRD, carbohydrate recognition domain; CTLD, C-type lectin domain; GlcNAc, N-acetyl-D-glucosamine; GRAVY, grand average of hydropathicity; ExPASy, Expert Protein Analysis System; IPR, InterPro protein domain/family identifier; kDA, kilo Dalton; MASP, MBL-associated serine proteases; MBL, mannose-binding lectin; MEGA, Molecular Evolutionary Genetics Analysis software; MW, molecular weight; NCBI, National Center for Biotechnology Information; ORF, open reading frame; OV, Ovambo; PDB, Protein Data Bank; *pI*, isoelectric point; PK, Potchefstroom Koekoek; ProSA, Protein Structure Analysis; SAP, serum amyloid P component; SPCS, signal peptide-conserved sites; VN, Venda.

Acknowledgment

Conflicts of interest

The author declares that there is no conflict of interest.

Authors' contributions

Peter A. Idowu conceptualized the study, designed the methodology, collected the sample, sequenced the data, validated the nucleotide sequencing, performed the *in silico* analyses, interpreted the data, drafted the manuscript, and did the final editing. Takalani J. Mpofu assisted with data collection, funding, review of the manuscript, and supervision. Ogun J. Mayowa assisted with the *in silico* analysis, data visualization, and manuscript write-up. Oliver T. Zishiri provided supervision, critical revision of the manuscript, and guidance on bioinformatics analyses. Nephawe K.A. assisted with funding, supervision, and manuscript editing. Mtileni B. assisted with funding, supervision, data validation, and manuscript editing. contributed to sample collection, gene sequencing, and laboratory validation of

nucleotide sequences. All authors read and approved the final manuscript.

References

[1] Idowu PA, Zishiri O, Nephawe KA, Mtileni B. Current status and intervention of South Africa chicken production: a review. World Poult Sci J 2021; 77(1):115–33; https://doi.org/10.1080/00439339.2020.1866965

[2] Hlokoe VR, Tyasi TL. Quantitative and qualitative traits characterisation of indigenous chickens in Southern African countries. Online J Anim Feed Res 2022; 12(6):333–40; https://doi.org/10.51227/ojafr.2022.44

[3] Idowu PA, Mpayipheli M, Muchenje V. Practices, housing and diseases within indigenous poultry production in Eastern Cape, South Africa. J Agric Sci 2018; 10(11):111–21; https://doi.org/10.5539/jas.v10n11p111

[4] Laursen SB, Nielsen OL. Mannan-binding lectin (MBL) in chickens: molecular and functional aspects. Dev Comp Immunol 2000; 24(2–3):85–101; https://doi.org/10.1016/S0145-305X(99)00066-X

[5] Zhang W, Bouwman KM, Van Beurden SJ, Ordonez SR, Van Eijk Haagsman HP, et al. Chicken mannose-binding lectin has antiviral activity towards infectious bronchitis virus. Virology 2017; 509:252–9; https://doi.org/10.1016/j.virol.2017.06.028

[6] Chettri D, Boro M, Sarkar L, Verma AK. Lectins: biological significance to biotechnological application. Carbohydr Res 2021; 506:108367; https://doi.org/10.1016/j.carres.2021.108367

[7] Kilpatrick DC. Animal lectins: a historical introduction and overview. Biochim Biophys Acta Gen Subj 2002; 1572(2–3):187–97; https://doi.org/10.1016/S0304-4165(02)00308-2

[8] Dias R, Machado L, Migliolo L, Franco O. Insights into animal and plant lectins with antimicrobial activities. Molecules 2015; 20(1):519–41; https://doi.org/10.3390/molecules20010519

[9] Loh SH, Park JY, Cho EH, Nah SY, Kang YS. Animal lectins: potential receptors for ginseng polysaccharides. J Ginseng Res 2017; 41(1):1–9; https://doi.org/10.1016/j.jgr.2015.12.006

[10] Idowu PA, Idowu AP, Zishiri OT, Mpofu TJ, Veldhuizen EJA, Nephawe KA, et al. Activity of mannose-binding lectin on bacterial-infected chickens: a review. Animals 2021; 11(3):787; https://doi.org/10.3390/ani11030787

[11] Coelho LCBB, PMdS S, De Oliveira WF, De Moura MC, Pontual EV, Gomes FS, et al. Lectins as antimicrobial agents. J Appl Microbiol 2018; 125(5):1238–52; https://doi.org/10.1111/jam.14055

[12] Ulrich-Lynge SL, Dalgaard TS, Norup LR, Kjærup RM, Olsen JE, Sørensen P, et al. The consequence of low mannose-binding lectin plasma concentration in relation to susceptibility to *Salmonella* Infantis in chickens. Vet Immunol Immunopathol 2015; 163(1–2):23–32; https://doi.org/10.1016/j.vetimm.2014.11.003

[13] Jogi J, Sharma V, Nayak A, Gupta V, Shakya P, Bordoloi S. Pattern recognition receptors (PRRs). Int J Curr Microbiol App Sci 2018; 7(9):3400–6; https://doi.org/10.20546/ijcmas.2018.709.422

[14] Miyake Y. Classification of C-type lectins and recognition of pathogens. Microbiol Immunol 2025; 69(5):257–69; https://doi.org/10.1111/1348-0421.13211

[15] Dos Santos Silva PM, De Oliveira WF, Albuquerque PBS, Dos Santos Correia MT, Coelho LCBB. Insights into anti-pathogenic activities of mannose lectins. Int J Biol Macromol 2019; 140:234–44; https://doi.org/10.1016/j.ijbiomac.2019.08.059

[16] Bodi J, Nsibu C, Hirayama K. Immunogenetic mechanisms of black water fever: article review. Gene Technol 2021; 10(161):1–8; Available via https://www.walshmedicalmedia.com/open-access/immunogenetic-mechanisms-of-black-water-fever-article-review.pdf (Accessed January 17, 2025)

[17] Mori K, Ohtani K, Jang S, Kim Y, Hwang I, Roy N, et al. Scavenger receptor CL-P1 mainly utilizes a collagen-like domain to uptake microbes and modified LDL. Biochimica Et Biophysica Acta (BBA) - Gen Subjects 2014; 1840(12):3345–56; https://doi.org/10.1016/j.bbagen.2014.08.015

[18] Hammad NM, El Badawy NE, Ghramh HA, Al Kady LM. Mannose-binding lectin: a potential therapeutic candidate against Candida infection. Biomed Res Int 2018; 2018:2813737; https://doi.org/10.1155/2018/2813737

[19] Gupta A, Gupta GS. Applications of mannose-binding lectins and mannan glycoconjugates in nanomedicine. J Nanopart Res 2022; 24:228; https://doi.org/10.1007/s11051-022-05594-1

[20] Zhang H, Peatman E, Liu H, Niu D, Feng T, Kucuktas H, et al. Characterization of a mannose-binding lectin from channel catfish (*Ictalurus punctatus*). Res Vet Sci 2012; 92(3):408–13; https://doi.org/10.1016/j.rvsc.2011.03.024

[21] Mason C, Tarr A. Human lectins and their roles in viral infections. Molecules 2015; 20(2):2229–471; https://doi.org/10.3390/molecules20022229

[22] Iqbal H, Fung KW, Gor J, Bishop AC, Makhatadze GI, Brodsky B, et al. A solution structure analysis reveals a bent collagen triple helix in the complement activation recognition molecule mannan-binding lectin. J Biol Chem 2023; 299(2):102799; https://doi.org/10.1016/j.jbc.2022.102799

[23] Degn SE, Thiel S. Humoral pattern recognition and the complement system. Scand J Immunol 2013; 78(2):181–93; https://doi.org/10.1111/sji.12070

[24] Veldhuizen EJ, Van Eijk M, Haagsman HP. The carbohydrate recognition domain of collectins. FEBS J 2011; 278(20):3930–41; https://doi.org/10.1111/j.1742-4658.2011.08206.x

[25] Drickamer K, Taylor ME. Recent insights into structures and functions of C-type lectins in the immune system. Curr OpIn Struct Biol 2015; 34:26–34; https://doi.org/10.1016/j.sbi.2015.06.003

[26] Balabanova L, Golotin V, Kovalchuk S, Bulgakov A, Likhatskaya G, Son O, et al. A novel bifunctional hybrid with marine bacterium alkaline phosphatase and Far Eastern holothurian mannan-binding lectin activities. PLoS One 2014; 9(11):e112729; https://doi.org/10.1371/journal.pone.0112729

[27] Zhang M, Liu Y, Song C, Ning J, Cui Z. Characterization and functional analysis of a novel mannose-binding lectin from the swimming crab Portunus trituberculatus. Fish Shellfish Immunol 2019; 89:448–57; https://doi.org/10.1016/j.fsi.2019.04.007

[28] Mizgina TO, Baldaev SN, Likhatskaya GN, Molchanova VI, Kokoulin MS, Filshtein AP, et al. Molecular cloning and characteristics of a lectin from the bivalve *Glycymeris yessoensis*. Mar Drugs 2023; 21(2):55; https://doi.org/10.3390/md21020055

[29] Schou TW, Permin A, Christensen JP, Cu HP, Juul-Madsen HR. Mannan-binding lectin (MBL) in two chicken breeds and the correlation with experimental *Pasteurella multocida* infection. Comp Immunol Microbiol Infect Dis 2010; 33(3):183–95; https://doi.org/10.1016/j.cimid.2008.08.010

[30] Yu JC, Hu CC, Chang WH, Chen PC, Lee MS, Peng KT, et al. An integrated microfluidic system using mannose-binding lectin for bacteria isolation and biofilm-related gene detection. Microfluid Nanofluidics 2018; 22:31; https://doi.org/10.1007/s10404-017-2031-3

[31] Kjærup RM, Norup LR, Skjødt K, Dalgaard TS, Juul-Madsen HR. Chicken mannose-binding lectin (*MBL*) gene variants with influence on MBL serum concentrations. Immunogenetics 2013; 65:461–71; https://doi.org/10.1007/s00251-013-0689-6

[32] Mamu G, Brah GS, Mukhopadhyay CS, Deka D, Ramneek R. Molecular characterization of mannose-binding lectin protein in chickens. Indian J Anim Sci 2014; 84(1):34–6; https://doi.org/10.56093/ijans.v84i1.37299

[33] Jimenez HJ, Martins LS, Carvalho R, Montarroyos AV, Filho RMM. *In silico* characterization and phylogenetic analysis of a

mannose-specific lectin in Allium species. Genet Mol Res 2019; 18(2):18187.

[34] Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Humana Press, Totowa, USA, pp 571–607, 2005. https://doi.org/10.1385/1-59259-890-0:571

[35] Ibtehaz N, Kagaya Y, Kihara D. Domain-PFP allows protein function prediction using function-aware domain embedding representations. Commun Biol 2023; 6:1103; https://doi.org/10.1038/s42003-023-05476-9

[36] Joshi JS, Badgujar NV, Patel HV, Raval AP, Tarapara BV, Shah FD. Identifying the impact of structurally and functionally high-risk nonsynonymous SNPs on human patched protein using in-silico approach. Gene Rep 2021; 23:101097; https://doi.org/10.1016/j.genrep.2021.101097

[37] Kumar S, Stecher G, Suleski M, Sunderland M, Sharma S, Tamura K. MEGA12: molecular evolutionary genetics analysis version 12 for adaptive and green computing. Mol Biol Evol 2024; 41(12):msae263; https://doi.org/10.1093/molbev/msae263

[38] Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res 2015; 43(1):W401–7; https://doi.org/10.1093/nar/gkv485

[39] Thakur CJ, Saini S, Notra A, Chauhan B, Arya S, Gupta R, et al.Deciphering the functional role of hypothetical proteins from Chloroflexus aurantiacus J-10-f1 using bioinformatics approach. Mol Biol Res Commun 2020; 9(3):129–39; https://doi.org/10.22099/mbrc.2020.36894.1495

[40] Da Silva FG, Maia RT, Jimenez HJ, de Moraes Filho RM, Silva GC. In silico characterization, phylogeny and molecular docking of the Bowman–Birk type protease inhibitor in species from Phaseoleae tribe. Sci Plena 2024; 20(10):100202; https://doi.org/10.14808/sci.plena.2024.100202

[41] El Khoury G, Azzam W, Rebehmed J. PyProtif: a PyMol plugin to retrieve and visualize protein motifs for structural studies. Amino Acids 2023; 55:1429–36; https://doi.org/10.1007/s00726-023-03323-z

[42] Waterhouse AM, Studer G, Robin X, Bienert S, Tauriello G, Schwede T. The structure assessment web server: for proteins, complexes and more. Nucleic Acids Res 2024; 52(W1):W318–323; https://doi.org/10.1093/nar/gkae270

[43] Ramya S, Prabhu P. In-silico approaches for molecular characterization, structural function prediction and peptide toxicity analysis of the matrix protein of vesicular stomatitis Indiana virus (VSIV). BIO Web Conf 2025; 172:02007; https://doi.org/10.1051/bioconf/202517202007

[44] Singh RK, Chaurasiya AK, Kumar A. Ab initio modeling of human IRS1 protein to find novel target to dock with drug MH to mitigate T2DM diabetes by insulin signaling. 3 Biotech 2024; 14:108; https://doi.org/10.1007/s13205-024-03955-2

[45] Rehman HM, Naz M, Ghous AG, Malik M, Ahmad S, Bashir H. Computational design and evaluation of a novel temporin 1CEa–IL24 fusion protein for anti-tumor potential. Biomed Res Ther 2025; 12(2):7138–52; https://doi.org/10.15419/bmrat.v12i2.959

[46] Rad M, Ebrahimipour G, Bandehpour M, Akhavan O, Yarian F. Neisseria meningitidis detection by coupling bacterial factor H onto Au/scFv antibody nanohybrids. Appl Phys A 2023; 129:401; https://doi.org/10.1007/s00339-023-06620-2

[47] Mohanta TK, Khan A, Hashem A, Abd_Allah EF, Al-Harrasi A. The molecular mass and isoelectric point of plant proteomes. BMC Genomics 2019; 20:631; https://doi.org/10.1186/s12864-019-5983-8

[48] Kozlowski LP. Proteome-pI: proteome isoelectric point database. Nucleic Acids Res 2017; 45(D1):D1112–6; https://doi.org/10.1093/nar/gkw978

[49] Bing SJ, Li YQ, Sun GJ, Wang CY, Liang Y, Hua DL, et al. Effect of different acidic or alkaline environments on structural characteristics, functional and physicochemical properties of Lentinus edodes protein. Process Biochem 2023; 130:15–25; https://doi.org/10.1016/j.procbio.2023.04.001

[50] Kurotani A, Tokmakov AA, Sato KI, Stefanov VE, Yamada Y, Sakurai T. Localization-specific distributions of protein pI in human proteome are governed by local pH and membrane charge. BMC Mol Cell Biol 2019; 20:36; https://doi.org/10.1186/s12860-019-0221-4

[51] PIeroni L, Iavarone F, Olianas A, Greco V, Desiderio C, Martelli C, et al. Enrichments of post-translational modifications in proteomic studies. J Sep Sci 2020; 43(1):313–6; https://doi.org/10.1002/jssc.201900804

[52] Meyer NH, Dellago H, Tam-Amersdorfer C, Merle DA, Parlato R, Gesslbauer B, et al. Structural fuzziness of the RNA-organizing protein Serf determines a toxic gain-of-interaction. J Mol Biol 2020; 432(4):930–51; https://doi.org/10.1016/j.jmb.2019.11.014

[53] Muller MP, Jiang T, Sun C, Lihan M, Pant S, Mahinthichaichan P, et al. Characterization of lipId–protein interactions and lipId-mediated modulation of membrane protein function through molecular simulation. Chem Rev 2019; 119(9):6086–161; https://doi.org/10.1021/acs.chemrev.8b00608

[54] Mapiour M, Abdelrasoul A. Critical influences of plasma pH on human protein properties for modeling considerations: size, charge, conformation, hydrophobicity, and denaturation. J Compos Sci 2023; 7(1):28; https://doi.org/10.3390/jcs7010028

[55] Phadungsil W, Grams R. Agglutination activity of Fasciola gigantica DM9-1, a mannose-binding lectin. Korean J Parasitol 2021; 59(2):173–80; https://doi.org/10.3347/kjp.2021.59.2.173

[56] Abeyratne-Perera HK, Chandran PL. Mannose surfaces exhibit self-latching, water structuring, and resilience to chaotropes: implications for pathogen virulence. Langmuir 2017; 33(36):9178–89; https://doi.org/10.1021/acs.langmuir.7b01006

[57] Zhang W, Van Eijk M, Guo H, Van Dijk A, Bleijerveld OB, Verheije MH, et al. Expression and characterization of recombinant chicken mannose-binding lectin. Immunobiology 2017; 222(3):518–28; https://doi.org/10.1016/j.imbio.2016.10.019

[58] Milutinovic B, Mitic N, Roncevic J, Goc S, Jankovic M. Glycome complexity of human seminal plasma high molecular mass components: evaluation of the contribution of acid-soluble glycoproteins/mucins and extracellular vesicles. Arch Biochem Biophys 2016; 609:20–30; https://doi.org/10.1016/j.abb.2016.09.005

[59] Zhao J, Zhang F, Liu X, St. Ange K, Zhang A, Li Q, et al. Isolation of a lectin binding rhamnogalacturonan-I containing pectic polysaccharide from pumpkin. Carbohydr Polym 2017; 163:330–6; https://doi.org/10.1016/j.carbpol.2017.01.067

[60] Kalia N, Singh J, Kaur M. The ambiguous role of mannose-binding lectin (MBL) in human immunity. Open Med 2021; 16(1):299–310; https://doi.org/10.1515/med-2021-0239

[61] Heindel DW, Koppolu S, Zhang Y, Kasper B, Meche L, Vaiana CA, et al. Glycomic analysis of host response reveals high mannose as a key mediator of influenza severity. Proc Natl Acad Sci USA 2020; 117(43):26926–35; https://doi.org/10.1073/pnas.2008203117

[62] Gupta A, Gupta GS. Status of mannose-binding lectin (MBL) and complement system in COVID-19 patients and therapeutic applications of antiviral plant MBLs. Mol Cell Biochem 2021; 476:2917–42; https://doi.org/10.1007/s11010-021-04107-3

[63] Lawal RA, Martin SH, Vanmechelen K, Vereijken A, Silva P, Al-Atiyat RM, et al. The wild species genome ancestry of domestic chickens. BMC Biol 2020; 18(13):13; https://doi.org/10.1186/s12915-020-0738-1

[64] Trzilova D, Tamayo R. Site-specific recombination: how simple DNA inversions produce complex phenotypic heterogeneity in bacterial populations. Trends Genet 2021; 37(1):59–72; https://doi.org/10.1016/j.tig.2020.09.004

[65] Lyimo CM, Weigend A, Msoffe PL, Eding H, Simianer H, Weigend S. Global diversity and genetic contributions of chicken populations from African, Asian and European regions. Anim Genet 2014; 45(6):836–48; https://doi.org/10.1111/age.12230;

[66] Larsen JB, Laursen MA, Hvas CL, Larsen KM, Thiel S, Hvas AM. Reduced mannose-binding lectin-associated serine protease (MASP)-1 is associated with disturbed coagulation in septic shock. Thromb Haemost 2019; 119(6):952–61; https://doi.org/10.1055/s-0039-1685140

[67] Cummings RD, Chiffoleau E, Van Kooyk Y, McEver RP. C-type lectins. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al. (eds). Essentials of glycobiology 4th edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 435–50, 2022. Available via https://www.ncbi.nlm.nih.gov/books/NBK579916/

[68] Nivetha R, Meenakumari M, Bhuvaragavan S, Hilda K, Janarthanan S. *In silico* analysis of carbohydrate-binding pockets in the lectin genes from various species of Canavalia. Comput Biol Chem 2021; 92:107477; https://doi.org/10.1016/j.compbiolchem.2021.107477

[69] Ogun OJ, Thaller G, Becker D. Molecular structural analysis of porcine CMAH–native ligand complex and high throughput virtual screening to identify novel inhibitors. Pathogens 2023; 12(5):684; https://doi.org/10.3390/pathogens12050684

[70] Dorn M, E Silva MB, Buriol LS, Lamb LC. Three-dimensional protein structure prediction: methods and computational strategies. Comput Biol Chem 2014; 53(B):251–76; https://doi.org/10.1016/j.compbiolchem.2014.10.001

[71] Sobolev OV, Afonine PV, Moriarty NW, Hekkelman ML, Joosten RP, Perrakis A, et al. A global Ramachandran score identifies protein structures with unlikely stereochemistry. Structure 2020; 28(11):1249–58; https://doi.org/10.1016/j.str.2020.08.005;