



Article

Directed Mutagenesis for Arginine Substitution of a *Phaseolus acutifolius* Recombinant Lectin Disrupts Its Cytotoxic Activity

Dania Martínez-Alarcón ^{1,†}, José Luis Castro-Guillén ^{2,†}, Elaine Fitches ³, John A. Gatehouse ^{3,‡}, Stefan Przyborski ³, Ulisses Moreno-Celis ⁴, Alejandro Blanco-Labra ¹ and Teresa García-Gasca ^{4,*}

¹ Centro de Investigación y de Estudios Avanzados Unidad Irapuato, Departamento de Biotecnología y Bioquímica, Irapuato 36821, Guanajuato, Mexico; dania.martinez.alarcon@gmail.com (D.M.-A.); alejandroblancolabra@gmail.com (A.B.-L.)

² Tecnológico Nacional de México/Instituto Tecnológico Superior de Irapuato (ITESI), Km. 12.5, Carretera Irapuato-Silao, El Copal, Irapuato 36821, Guanajuato, Mexico; jose.cg1@irapuato.tecnm.mx

³ Department of Biosciences, Durham University, Durham DH1 3LE, UK; e.c.fitches@durham.ac.uk (E.F.); stefan.przyborski@durham.ac.uk (S.P.)

⁴ Facultad de Ciencias Naturales, Universidad Autónoma de Querétaro, Querétaro 76230, Querétaro, Mexico; ulisses.moreno@uaq.mx

* Correspondence: tggasca@uaq.edu.mx,

† These authors contributed equally to this work.

‡ This paper is dedicated to the memory of Professor John Gatehouse (1951–2022).

Abstract: Recently, we reported that a recombinant Tepary bean (*Phaseolus acutifolius*) lectin (rTBL-1) induces apoptosis in colon cancer cell lines and that cytotoxicity was related to differential recognition of β 1-6 branched *N*-glycans. Sequencing analysis and resolution of the rTBL-1 3D structure suggest that glycan specificity could be strongly influenced by two arginine residues, R103 and R130, located in the carbohydrate binding pocket. The aim of this work was to determine the contribution of these residues towards cytotoxic activity. Two rTBL-1 mutants were produced in *Pichia pastoris*, biochemically characterized, and cytotoxic effects were evaluated on human colorectal cancer cells (HT-29). Substitution of either of the arginine residues with glutamines resulted in significant reductions in cytotoxic activity, with losses of 1.5 and 3 times for R103 and R130, respectively. Docking analysis showed that the mutations decreased lectin affinity binding to some Epidermal Growth Factor Receptor (EGFR)-related *N*-glycans. Together, these findings confirm that both of the selected arginine residues (R103 and R130) play a key role in the recognition of tumor cell glycoconjugates by rTBL-1.

Keywords: recombinant lectins; tepary bean; branched *N*-glycans; cancer



Citation: Martínez-Alarcón, D.; Castro-Guillén, J.L.; Fitches, E.; Gatehouse, J.A.; Przyborski, S.; Moreno-Celis, U.; Blanco-Labra, A.; García-Gasca, T. Directed Mutagenesis for Arginine Substitution of a *Phaseolus acutifolius* Recombinant Lectin Disrupts Its Cytotoxic Activity. *Int. J. Mol. Sci.* **2024**, *25*, 13258. <https://doi.org/10.3390/ijms252413258>

Academic Editor: Nam Deuk Kim

Received: 30 September 2024

Revised: 5 December 2024

Accepted: 6 December 2024

Published: 10 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Anticancer lectins are a specialized group of proteins that are able to recognize and interact with specific carbohydrate structures of cancer cells [1]. By specifically targeting saccharide structures, these lectins can disrupt crucial cellular processes, interfere with signaling pathways, and ultimately exert cytotoxic and antiproliferative effects on cancer cells. As a consequence, these proteins have gained considerable attention in cancer research due to their potential applications in cancer therapies [2]. One notable example of an anticancer lectin is the Phytohemagglutinin (PHA) derived from kidney bean (*Phaseolus vulgaris*) seeds. PHA has demonstrated cytotoxic effects on cancer cells, leading to cell death and reduced tumor growth [3]. PHA's anticancer effect stems from its ability to recognize and bind to specific carbohydrate molecules displayed on the surface of cancer cells. By binding to these structures, PHA can interfere with vital cellular processes, disrupt signaling pathways, and induce various anti-tumor effects, including the induction of

apoptosis, inhibition of cell proliferation, modulation of immune responses, and inhibition of angiogenesis [4–7].

In 2012, we reported that a group of PHA-homologous native lectins derived from Tepary bean (*Phaseolus acutifolius*) exhibited cytotoxic effects on various types of cancer cell lines [8] by inducing apoptosis and cell cycle arrest [9]. It was found that the Tepary Bean Lectin Fraction (TBLF) contained at least two cytotoxic lectins that were characterized at molecular and bioactive levels [10,11]; furthermore, when administered orally, TBLF exhibited the ability to inhibit early colon tumorigenesis in rats [12]. As Tepary Bean Lectin-1 was identified as the main lectin responsible for this cytotoxic effect, we subsequently produced it as a recombinant protein by bench-top fermentation of transformed *Pichia pastoris* and determined its three-dimensional structure [13,14]. The recombinant lectin-1 (rTBL-1) was found to recognize β 1-6 branched *N*-glycans, which are overrepresented in several cancer types, including colon cancer. Recognition of β 1-6 branched *N*-glycans does not appear to be affected by elongation with more units as sialic acid [13]. We also found that rTBL-1 retains its cytotoxic apoptosis-inducing effect on colon cancer cells via influencing epidermal growth factor cell receptor (EGFR) signaling pathways and principally by causing partial receptor degradation and activation of p38 MAPK signaling [15].

The structural characterization of rTBL-1 enabled us to compare its putative binding pocket and residues responsible for binding to the trisaccharide unit Gal(β 1-4)GlcNAc(β 1-2)Man with that previously reported for PHA [16]. We identified that, unlike other closely related legume agglutinins with anticancer activity, rTBL-1 has an arginine residue instead of lysine at position 103, which could be involved in the formation of hydrogen bonding to carbohydrates. Additionally, we identified the presence of an additional arginine within the binding pocket and adjacent to other residues involved in the recognition of carbohydrates reported for PHA. Therefore, as a highly conserved arginine residue is crucial for carbohydrate recognition in other families of lectins, such as the siglektins (sialic acid immunoglobulin-superfamily lectins) [17,18], we hypothesize that the sequence discrepancy in rTBL-1 could explain its cytotoxic effect as compared to the legume leucoagglutinins. To investigate this further, we have carried out a conservative substitution of two arginine residues with the uncharged amino acid glutamine in order to evaluate the impact of these modifications on the cytotoxic effect of rTBL-1 lectin.

2. Results and Discussions

2.1. Production of rTBL-1 Mutants

Multiple alignments of leucoagglutinin sequences, along with the findings of Nagae et al. [16], revealed that all sequences have identical carbohydrate binding pockets (CBP) (Figure 1A, blue triangles), except *P. acutifolius* and rTBL-1, where an arginine (R) instead of lysine (K) is located at position 103. This change is particularly interesting because, despite the similarity between these amino acids, conservative substitutions such as replacing R with K in other lectins, such as the animal I-type lectins (siglektins), have shown a conserved arginine residue to be essential for sialic-binding [17,19]. Sialic acid is particularly abundant in tumor cells where abnormal sialylation is considered as a hallmark of cancer [20], so it has been studied for cancer-involved membrane receptors such as EGFR [21].

An additional arginine residue located at position 130 (R130) in TBL-1 differs from the presence of a lysine in the *P. vulgaris* lectin or tryptophan in *P. costaricensis* or *P. coccineus* lectins. R130 does not appear to have direct interaction with the ligand but is located adjacent to two of the residues involved in carbohydrate recognition and also for metal-bonding cations. These bonds are essential for interactions between monomers and, therefore, for the recognition of complex carbohydrates. It is thus possible that R130 may contribute to enhancing the cytotoxic effects of TBL-1. It was also found that the sequon, unique consensus sequence for *N*-glycosylation (N-X-S/T, where X represents any amino acid except P) [22] was located between residues 13 and 15 (N-E-T).

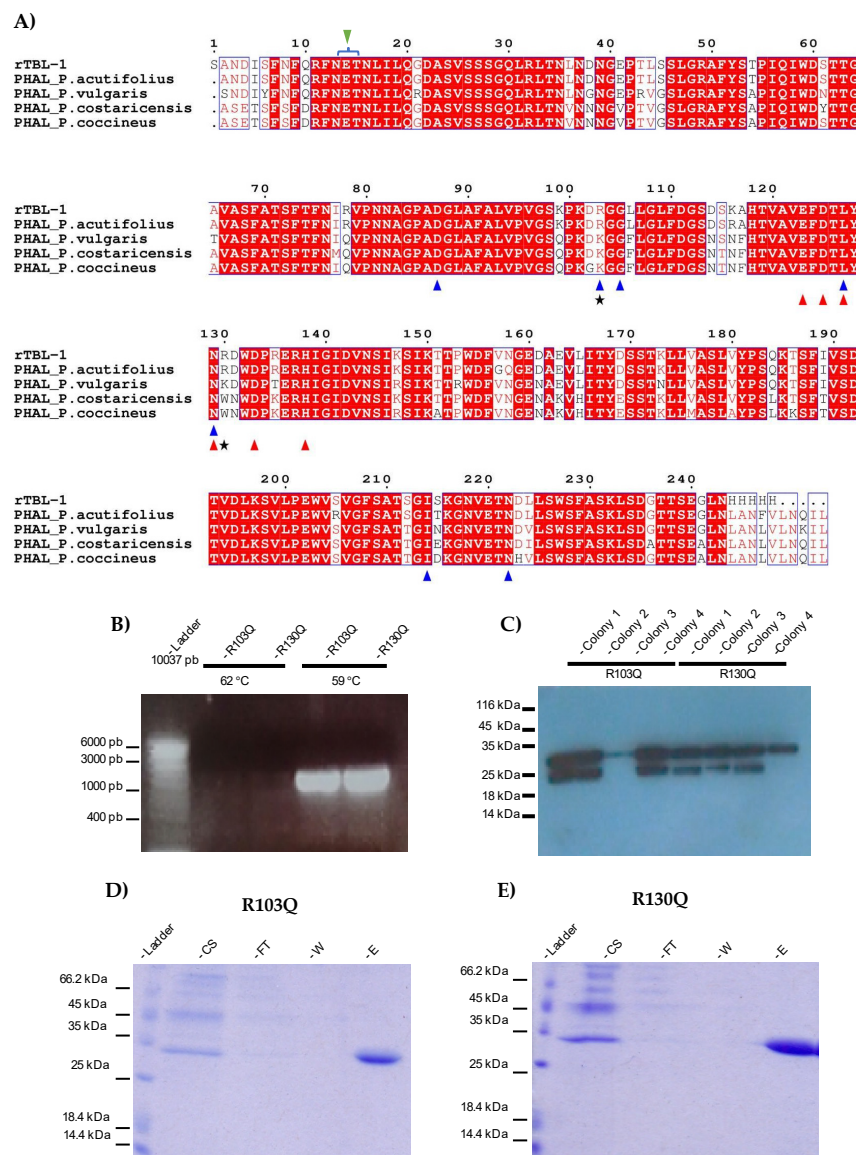


Figure 1. Mutants' design and production. (A) Sequence alignment of rTBL-1 homologous lectins. PHAL_*Phaseolus costaricensis* (Q5ZF34_9FABA), PHAL_*Phaseolus coccineus* (Q84RP8_PHACN), PHAL_*Phaseolus vulgaris* (PHAL_PHAVU), PHAL_*Phaseolus acutifolius* (Q40750_PHAAT), rTBL-1 sequence [14]. White letters represent identical residues, semiconserved residues are displayed in red enclosed into empty boxes and non-conserved residues are depicted in black. Blue triangles show residues that correspond to the CBP similarly to PHA-L from *P. Vulgaris*; red triangles show residues interacting with metal cations; green triangle depicts sequence for N-glycosylation marked with a blue bracket; black stars show residues presumably responsible for the cytotoxic effect of rTBL-1 (R103 and R130). (B) Agarose gel electrophoresis showing pGAP α ZB-rTBL-1 vector amplification. Lane 1, ladder; lanes 2–5, amplification products. (C) Western blot screening using anti-His antibodies for transformed colonies. (D,E) R103Q and R130Q SDS-PAGE electrophoretic profiles stained for total protein. Lane 1, ladder; lane 2, 25 μ L of culture supernatant (CS); lane 3, 25 μ L of flow through (FT); lane 4, 25 μ L wash with 10 mM imidazole buffer (W); and lane 5, 25 μ L elution with 200 mM imidazole buffer (E). PHAL, Leukocyte phytohemagglutinin.

Two rTBL-1 mutants were generated by substituting either R-103 or R-130 within the CBP with a non-charged glutamine (Q) residue. The resulting proteins were designated as R103Q and R130Q mutant lectins, respectively. Site-directed mutagenesis was performed on the original coding sequence of rTBL-1 by inverse amplification of the pGAPZ α B-rTBL-1

plasmid expression vector [13] to obtain amplicons of approximately 3000 bp (Figure 1B). Whilst amplification parameters were tested at two different temperatures (62 and 59 °C), viable products were only obtained at 59 °C (Figure 1B).

The vectors were gel purified, re-ligated, and subsequently cloned into Top10 *Escherichia coli* cells. Sequence verified plasmid DNAs were linearized and integrated into the GAPDH locus of *P. pastoris* through homologous recombination. Four colonies of each mutant were selected for expression screening via western analysis of supernatants derived from shake flask cultures. As shown in Figure 1C, all selected colonies secreted a recombinant protein of approximately 32 kDa consistent with the predicted mass for the mutants. The presence of an additional immunoreactive protein (ca. 25 kDa) observed in 6 of the 8 culture supernatants suggests degradation at the N-terminus of the mutants, since the His tag is at the C-terminus of the protein. Based on the performance of the small-scale cultures, colony number #1 for each mutant was selected to scale up production in a 7 L bench top fermenter.

The electrophoretic profile of the culture media obtained for both fermentations was very similar to the one previously reported for the rTBL-1 fermentation, where the recombinant product is presented as the most predominant band. The final yields per liter of culture supernatant following bench-top fermentation were 205 mg and 184 mg, respectively, for R103Q and R130Q. Proteins were purified by nickel affinity chromatography, and fractions were collected after elution with 200 mM imidazole. Analysis of purified proteins by SDS-PAGE electrophoresis showed the presence of a single band of ~30 kDa in Coomassie-stained gels (Figure 1D,E).

2.2. R103Q and R130Q Biochemical Characterization

rTBL-1 and the purified mutants were all separated as proteins of approx. 30 kDa under denaturing SDS-PAGE gel electrophoresis in line with predicted masses for the recombinant lectins (Figure 2A). However, it is well-known that rTBL-1 folds into a homo-tetramer [13]. To assess the oligomeric forms of R103Q and R130Q mutant lectins, their native molecular weights were determined using gel filtration chromatography using a high-resolution size exclusion column (ENrich™ SEC650, BioRad, Hercules, CA, USA). The elution profiles exhibited a single peak at approximately 120 kDa, indicating that both mutants are able to conform homo-tetramers in solution and that the K to Q substitutions did not disrupt their ability to oligomerize.

A single N-glycosylation site near their N-terminal region was found for rTBL-1; this sequence leads to the addition of an N-glycosidic antenna of approximately ~2.48 kDa on asparagine-13. All recombinant proteins were treated with various combinations of glycosidases in order to compare the type and estimated size of attached glycosidic antennas. Subsequently, the samples were subjected to SDS-PAGE, and the gels were stained using the Schiff-PASS technique, enabling the identification of protein-bound carbohydrates (Figure 2B).

Both R103Q and R130Q mutants showed similar molecular weights (Figure 2A). As rTBL-1 and the two mutants showed no changes in the protein sequon, an analysis of the glycosidic antennas was performed using several glycosidases. Upon treatment with N-glycosidase-F, both mutants and rTBL-1 exhibited a reduction in size of between 1.84 and 2.48 kDa, accompanied by the complete carbohydrate removal. No changes in intensity or molecular size were observed after treatment with O-glycosidases indicating that rTBL-1 and the mutants contain only N-linked glycans. Nano-LC-MS/MS analysis depending on standard data of tryptic peptides, was used to confirm the position of the carbohydrate antennas from samples treated with deglycosylating enzymes. This analysis confirmed the presence of an N-deamidated asparagine in the NETN peptide, which corresponds to positions 13–16 of the proteins. Deamination is known to be a byproduct of PNGaseF treatments by removing N-linked sugars.

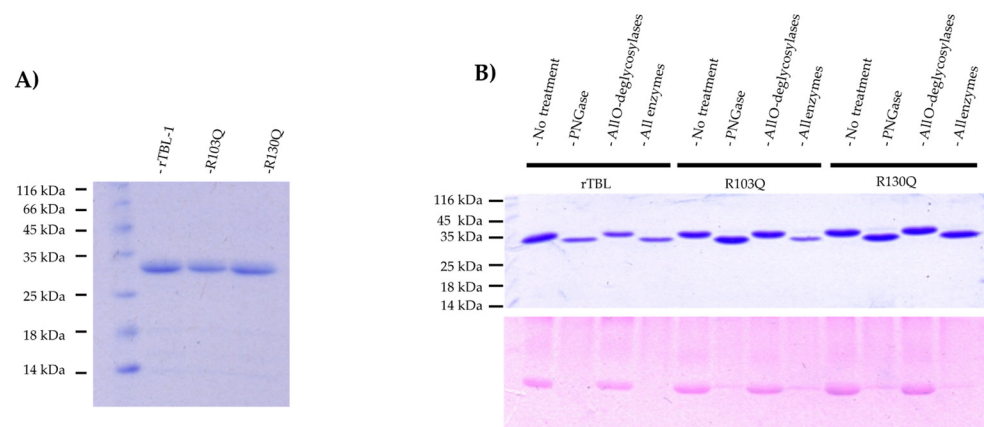


Figure 2. Characterization of the mutants. **(A)** Molecular size comparison between rTBL-1 and the mutants. SDS-PAGE of the electrophoretic profiles of rTBL-1, R103Q, and R130Q. Lane 1, ladder; lane 2, 5 μg of rTBL; lane 3, 5 μg of R103Q; lane 4, 5 μg of R130Q. **(B)** Analysis of glycosylation. Top panel shows a coomassie R-250-stained SDS-PAGE. N-Glycosidase F (PNGaseF), all O-glycosidases (β 1,4-Galactosidase, endo- α -N-acetylgalactosaminidase, α 2-3,6,8,9-Neuraminidase, and β -N-Acetylglucosaminidase), all N- and O-glycosidases (N-Glycosidase F, β 1,4-Galactosidase, endo- α -N-acetylgalactosaminidase, α 2-3,6,8,9-Neuraminidase, and β -N-Acetylglucosaminidase).

2.3. Cytotoxic Effect of R103Q and R130Q

The cytotoxicity of rTBL-1 and its mutants was assessed on HT-29 colorectal cancer cells through dose–response curves using concentrations ranging from 0 to 50 $\mu\text{g}/\text{mL}$ (Figure 3). The percentage of cell survival was normalized against the negative PBS-treated control treatment. The results show concentration-dependent responses for all proteins. However, rTBL-1 showed significantly greater cytotoxic effects as compared to either of the mutants, as evidenced by significant differences ($p < 0.05$) in lethal concentrations (LC_{50} values) derived from dose–response assays. LC_{50} values were 2.37 $\mu\text{g}/\text{mL}$ ($y = -16 - 317x + 88.654$) for rTBL-1 as compared to 4.94 $\mu\text{g}/\text{mL}$ ($y = -12.702x + 112.59$) and 7.22 $\mu\text{g}/\text{mL}$ ($y = -8.6725x + 112.7$) for R103Q and R130Q, respectively.

These results show that each arginine substitution plays a role in exerting the cytotoxic effects of rTBL-1. R130Q exhibited the greatest inhibition of cytotoxicity, being 3 times less cytotoxic as compared to rTBL-1 and 1.5 times lower than R103Q. No statistical differences ($p > 0.05$) were observed for all tested concentrations of R130 as compared to control cells. For R103, significant differences ($p < 0.05$) were observed for the 5 to 50 $\mu\text{g}/\text{mL}$ treatments as compared to the control cells. These results highlight the differential role of arginine residues 103 or 130 in the lectin’s carbohydrate binding activity.

2.4. Docking Analysis

In order to know more about the participation of R103 and R130 residues, a docking analysis was performed using the rTBL-1 sequence (PDB ID: 6tt9); this sequence has one extra alanine beside the first serine of the rTBL-1 sequence [14]. This change did not affect the tridimensional structure, but it was necessary to go through the numbering one position, changing R103Q to R104Q and R130Q to R131Q only for the docking analysis. On the other hand, all docking procedures were performed using the LZERD web server according to its specifications (<https://lzerd.kiharalab.org> accessed on 28 January 2023) [23].

2.4.1. Docking Analysis with Experimentally Recognized N-glycans

Previously we have shown that rTBL-1 recognized experimental complex N-glycans from the Mammalian Glycan Array (GA) 5.4; therefore, a docking analysis was performed considering only those whose relative signal was above 40% [13]. Figure 4 shows the interaction between rTBL-1 and experimental N-glycans, where GA-01, GA-02, GA-03, GA-04, GA-06, GA-07, and GA-09 exhibited a docking interaction in the upper zone between

monomers 1 and 2 of tetrameric rTBL-1 located near the CBP, confirming recognition and suitable interaction positions of the experimentally recognized *N*-glycans. Only those complex glycans having four sialylated branches (GA-08) or highly elongated branches (GA-05) interacted directly in the CBP with a large part of the glycan structure. These results provide the first two selection criteria for those *N*-glycans that had direct and considerable interactions with the lectin CBP, namely the presence of at least four-ended sialylated branches and/or elongated branches.

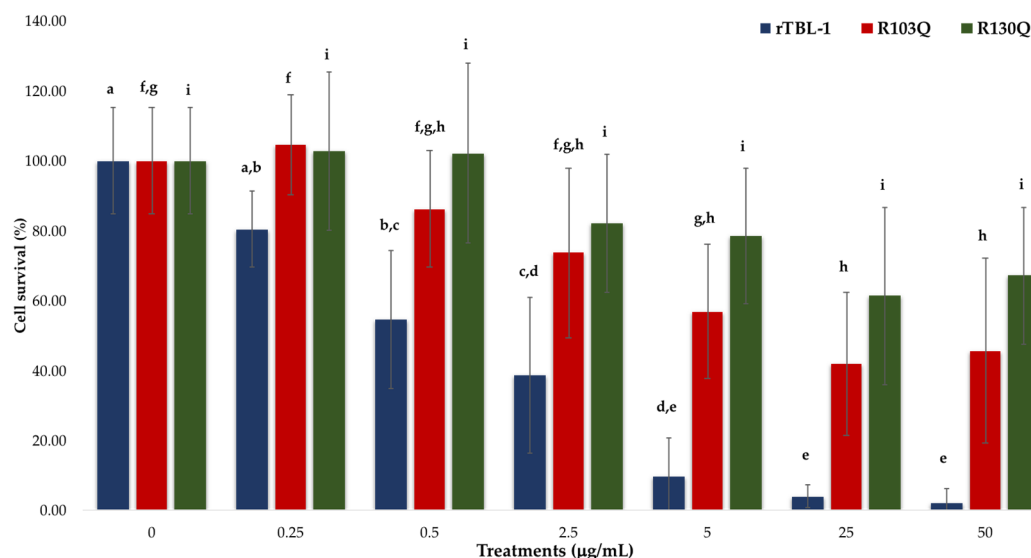


Figure 3. Cytotoxic effect of rTBL-1 and its mutants on HT-29 colon cancer cells. Cells were treated with rTBL-1, R103Q, or R130Q (0, 0.25, 0.5, 2.5, 5, 25, and 50 µg/mL) for 24 h and cell survival percentages were calculated from hemocytometer counts. Small letters show significant differences for each separate treatment at different concentrations (Tukey $p < 0.05$), where differences between treatments for rTBL-1 were pointed out with letters a–e, for R103Q with letters f–h, and for R130Q with letter i.

2.4.2. Docking Analysis with EGFR-Related *N*-glycans

EGFR is an important glycoprotein for cancer biology, being highly expressed and considered as a proteomic marker in different types of cancers [24–26]. EGFR-specific *N*-glycosylations are correlated with disease state/progression of colorectal cancer [27], and, for this reason, EGFR-related *N*-glycans were considered in this work for docking analysis using the structures of rTBL-1 and the in silico mutated versions R104Q and R131Q.

Complexity in the structure of *N*-glycans was also considered to select the EGFR-related glycans for the docking analysis, since rTBL-1 previously showed high affinity for complex glycans [13]. Furthermore, the protein–protein interaction of rTBL-1 with the EGFR has been observed [15]. On the other hand, due to the possible positions of its flexible structure, it was necessary to use several EGFR-related *N*-glycan conformers generated on the GLYCAM web server [29] in each docking analysis. This allowed us to screen a wide variety of docking results and select those within the CBP, and specifically those whose interactions were related to the mutated residues. Thirty-two of the 52 EGFR-related *N*-glycans (GlyConnect-<https://glyconnect.expasy.org> accessed on 8 February 2023) [30] were used for docking analysis with rTBL-1.

Only eight *N*-glycans were not recognized by rTBL-1 (considering that the resulting interactions were located between monomers [11]), and 24 EGFR-related *N*-glycans had reliable docking positions, including six interacting directly with CBP (Figure 5). Interestingly, when the reported molecular mass of each selected *N*-glycan (GlyConnect platform, <https://glyconnect.expasy.org/> accessed on 8 February 2023) [30] was compared, a molecular mass-lectin recognition correlation was found (Supplementary Material, Table S1).

The unrecognized *N*-glycans (those *N*-glycans with unsuitable docking interactions) had an average mass of 1.77 ± 0.33 kDa and included mainly complex two-branched glycans and high-mannose *N*-glycans. The *N*-glycans whose docking interaction was appropriate but without considerable interactions with CBP had 2.26 ± 0.25 kDa as average mass. Some of these *N*-glycans have three branches, two of which are located directly on the mannose (α 1-6) of the *N*-acetyl-glucosamine-trimannoside core. Most of these *N*-glycans were fucosylated. Finally, the EGFR-related *N*-glycans that had multiple interaction points with the CBP of rTBL-1 showed an average mass of 4.05 ± 0.25 kDa, having four branches containing sialic acid (*N*-acetylneuraminic acid [NeuAc(α 2-3)]) at the end of each branch and presenting a *N*-glycan core fucosylation at the innermost *N*-acetylglucosamine residue of the chitobiose (GlcNAc₂) core. An interesting feature of these *N*-glycans is that their structure was formed by four sialylated antennae and a core-fucosylated structure.

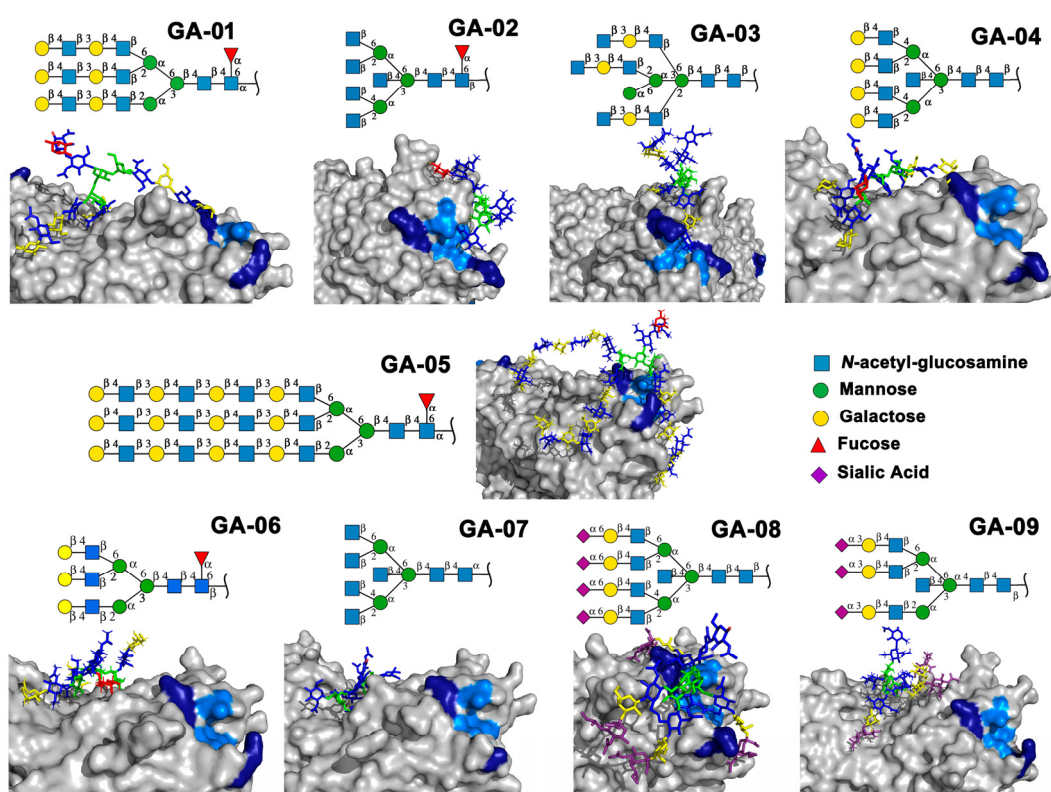


Figure 4. Molecular docking of the experimental *N*-glycans from the Mammalian Glycan Array 5.4 with rTBL-1 structures. The graphical 2D representation of *N*-glycans is shown according to the Symbol Nomenclature for Glycans (SNFG) [28]. The *N*-glycans from the mammalian glycan array “Version 5.4” are labeled as GA-##, where ## is the number described by Martínez-Alarcón et al. (2020) [13] and is related to the signal’s relative intensity of lectin recognition. Docking results are shown as topological representations of interactions at the same surface location between the *N*-glycans with the recombinant lectins: rTBL-1, R104Q, and with the R131Q lectin. The *N*-glycan in topological representations is presented as sticks colored according to the SNFG (as indicated by geometric codes), the protein surface in gray, highlighting the CBP in light blue and the arginine residues in dark blue. GA, glycan array.

It is worth mentioning that all those EGFR-related *N*-glycans whose interaction with Tepary bean lectin-related structures (rTBL-1, R104Q, and R131Q) have multiple contact points on the CBP region, especially for core-fucosylated and highly sialylated *N*-glycans. This fact is important since core fucosylation has been associated with a tumorigenic phenotype, and direct inhibition is able to inhibit cellular signaling for cell proliferation, migration, and tumor growth [31,32]. On the other hand, aberrant sialylation in cancer cells plays

an important role in tumor progression, favoring the evasion of cell death and promoting metastasis, immune evasion, and drug resistance. Aberrant sialylation, specifically α 2-3- and α 2-6-linked sialic acids, has been reported in highly metastatic and drug-resistant colorectal cancer and other tumor cells [33–38]. Highly sialylated *N*-glycans of EGFR have also been shown to disrupt ligand interactions by avoiding dimerization and subsequent phosphorylation [39], thereby protecting tumors against drug-induced cell death [40]; therefore, sialylation is an important target against cancer progression [41]. In consequence, the high affinity of rTBL-1 for highly sialylated EGFR-related *N*-glycans, especially those having α 2-3-linked sialic acids, suggest that its interaction directly promotes apoptosis of colon cancer cells, and this preference could explain the anticancer effect of this lectin and potential uses for the treatment of other types of cancer.

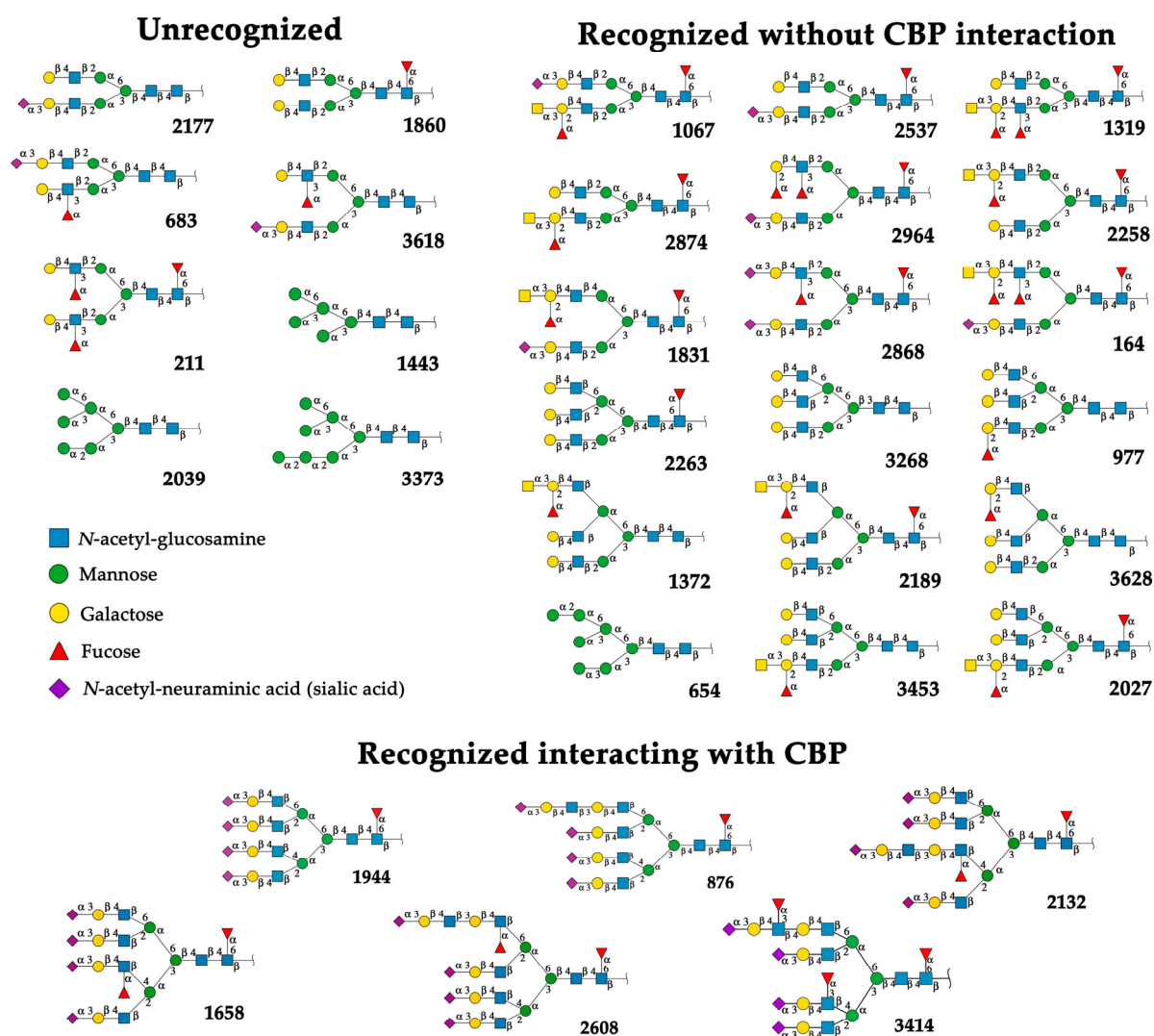


Figure 5. EGFR-related *N*-glycans used in docking process with rTBL-1. Graphical 2D representation of the *N*-glycans was created according to SNFG [28]. All EGFR-related *N*-glycans were named according to the GlyConnect platform [30]. CBP, carbohydrate-binding pocket.

2.4.3. The Role of R104 and R131 Residues in the Affinity rTBL-1 Toward EGFR-Related *N*-glycans by Docking Analysis

To determine the effect of punctual mutations on the lectin’s recognition of EGFR-related *N*-glycans, the high-molecular-mass *N*-glycans that interacted directly with a large part of their structure on the lectin CBP (forming a preliminary docking process using the structures shown in Figure 5) were chosen for docking analysis with in silico-mutated lectin

structures (R104Q and R131Q). Considering that all *N*-glycans have a $\text{Man}_3\text{GlcNAc}_2$ core, in this work the branches or “antennae” of complex *N*-glycans were numbered according to their relative position with respect to the mannose residue to which they are attached, either to $\text{Man}(\alpha 1-6)$ or $\text{Man}(\alpha 1-3)$ to simplify the nomenclature of the EGFR-related *N*-glycans (Supplementary Material, Figure S1). For instance, branch 1 corresponded to the antennae attached to carbon 6 of $\text{Man}(\alpha 1-6)$ of the $\text{Man}_3\text{GlcNAc}_2$ core, and branch 2 corresponded to the antennae attached to carbon 2 of the same mannose residue. In the same way, branch 3 corresponded to carbon 4 of $\text{Man}(\alpha 1-3)$ of the $\text{Man}_3\text{GlcNAc}_2$ core, and branch 4 corresponded to the antennae attached to carbon 2 of the same mannose residue.

Some of the EGFR-related *N*-glycans have elongated antennae. Generally, branches of complex *N*-glycans are composed of $\text{NeuAc}(\alpha 2-3)\text{Gal}(\beta 1-4)\text{GlcNAc}(\beta 1)$. However, in elongated antennae, two additional monosaccharides are present, which are *N*-acetylglucosamine with a glycosidic bond $\beta 1-3$ [$\text{GlcNAc}(\beta 1-3)$] and galactose attached to the *N*-acetylneuraminic acid (sialic acid) [$\text{NeuAc}(\alpha 2-3)$] found at the end of the branch, as follows: $\text{NeuAc}(\alpha 2-3)\text{Gal}(\beta 1-4)\text{GlcNAc}(\beta 1-3)\text{Gal}(\beta 1-4)\text{GlcNAc}(\beta 1)$.

Figure 6 shows the topological representation of the docking interactions of the selected EGFR-related *N*-glycans and the three lectin structures: rTBL-1, R104Q, or R131Q. The lectin structure was colored as a gray surface, and the CBP zone was highlighted as a light-blue zone, with residues R104 and R131 as dark-blue zones and mutated residues (Q104 and Q131) as dark-red surfaces, in order to compare the different ways in which the glycans interacted with the CBP of the lectins. When *N*-glycans interacted with rTBL-1, all except *N*-glycan 876 had a considerable interaction area on the CBP surface. Then, when comparing the docking results of the R104Q lectin with the EGFR-related *N*-glycans, *N*-glycans such as 876 and 2132 had no interaction surface in the CBP. In the case of R131Q lectin, the interaction positions were slightly changed compared to those of the rTBL-1 lectin, except for glycan 876, which had a very different location.

When comparing the 3D-interaction maps of all docking interactions (Figure 7), where protein-interacting residues are highlighted as orange sticks with their respective surface and chain position number, it is noteworthy that in the interactions of the *N*-glycans with the lectin rTBL-1, the R131 residue participates in most of them (1944, 1658, 2132, and 3414), and the R104 residue only in the interactions between rTBL-1 and 2132, 2608, and 3414 glycans. Both residues interact by a hydrogen bond (Supplementary Material, Tables S2–S19) and this could explain how the change of arginine to glutamine allowed hydrogen bond interactions by the amino group of the glutamine residue. Additionally, the Q104 residue participated in most of the interactions with *N*-glycans, such as 876, 1944, 1658, 2608, and 3414; in the R104Q lectin—2608 *N*-glycan interaction, it was possible to observe the participation of the R131 residue. Finally, the mutated Q131 residue participated only in the interactions of 1944, 1658, and 2608 *N*-glycans, having the additional participation of the R104 residue in the case of the interactions with 1658 and 2608 *N*-glycans.

At this point, it is remarkable that the arginine to glutamine mutation ($R \rightarrow Q$) allows, in most cases, the interaction of the lectin with *N*-glycans. Nevertheless, the R104Q mutation completely changed the interaction profile with the EGFR-related 2132 *N*-glycan, which has an elongated antenna with a second fucosylation at branch 3. This change is evident in the respective docking analysis, as there were no interactions in the CBP zone. This may be an indication that R104 plays a crucial role in the specific recognition of the lectin to certain *N*-glycan structures, and although the mutated residues may interact with *N*-glycans in most cases, they affect the three-dimensional configuration of the CBP and its surrounding regions and ultimately influence the specific recognition of the lectin and its target *N*-glycans.

In order to characterize the interaction interfaces between the EGFR-related *N*-glycans and the CBP of the recombinant and mutant Tepary bean lectins, the solvation free energy gain upon formation of the interface (Δ^iG) was determined for all docking interactions (Table 1). The Δ^iG was calculated as a difference in solvation energies of all residues between interfacing structures, and a negative value corresponds to hydrophobic interfaces

or is indicative of positive protein affinity according to the docking methodology of the PDBePISA web server [42].

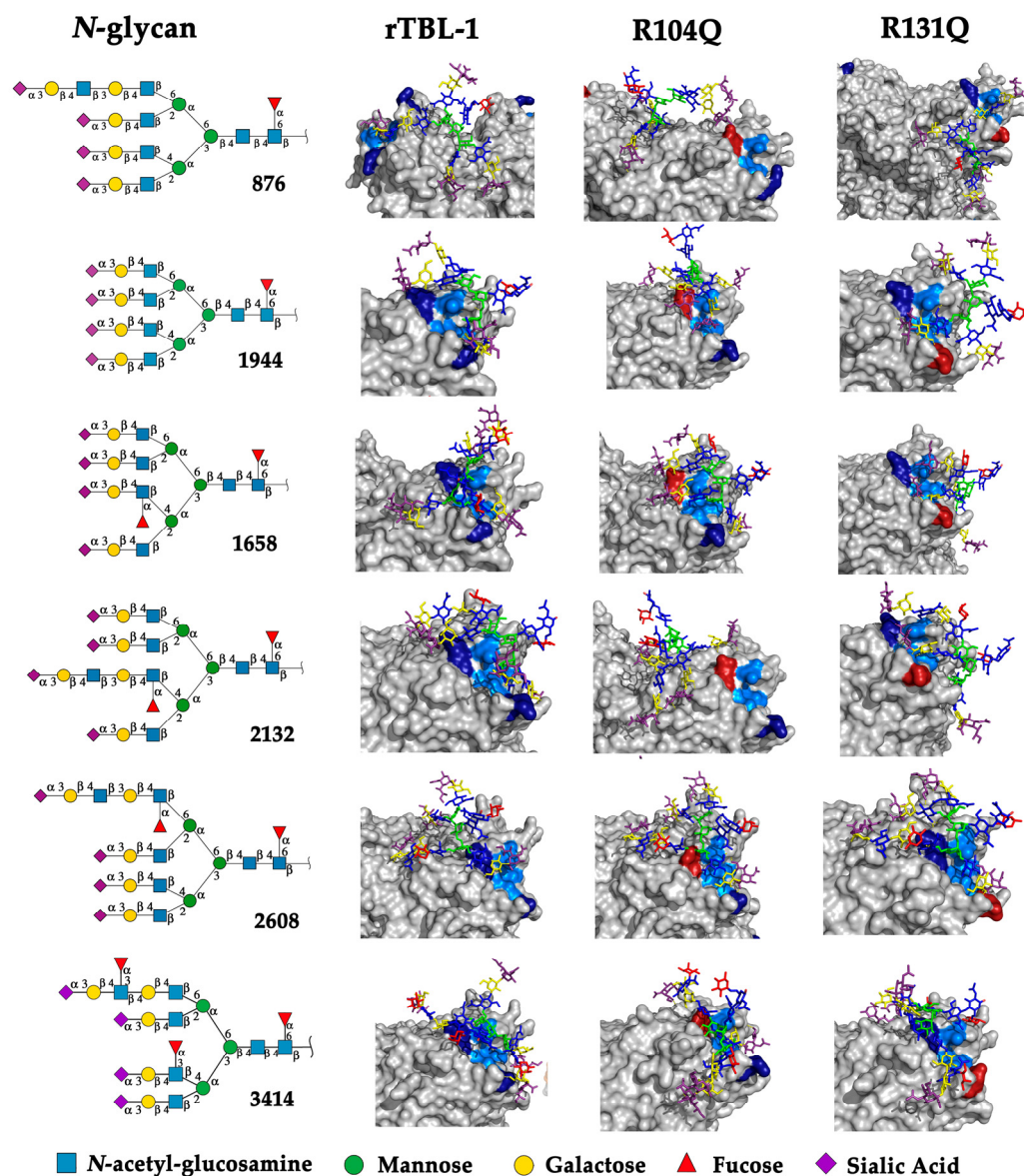


Figure 6. Molecular docking results for EGFR-related *N*-glycans with rTBL-1 and mutants R104Q and R131Q. For each *N*-glycan a graphical 2D representation according to SNFG [28] is shown. Docking results are shown as topological representations of interactions at the same surface location between the *N*-glycans, and the recombinant lectins: rTBL-1, R104Q, and R131Q. The *N*-glycan in topological representations is presented as sticks colored according to the SNFG, the protein surface in gray, highlighting the CBP in light blue, the mutated residues in red, and the original arginine residues in dark blue.

An important finding regarding the role of residues R104 and R131 in the recognition of complex *N*-glycans, considering the Δ^iG , is the appreciable loss of affinity in the recognition of EGFR-related *N*-glycan 2132 when R104 is mutated. Although the glycan is recognized and interacts with the lectin, it no longer interacts directly with the CBP and its solvation free energy upon formation of interfaces (Δ^iG) decreases considerably, indicating a significant loss of affinity (compare the docking results of glycan 2132 (Figure 5) with the energies of Table 1). This phenomenon also occurs with the interaction between lectins with EGFR-related *N*-glycan 876, which has an elongated antenna at branch 1 without a second

glycosylation, but it is not as representative since this glycan interacts with CBP with few residues with the rTBL-1. Additionally, this could also provide insight into the importance of N-glycan 2132 as a target molecule for tepary lectin and the effect it has on colon cancer.

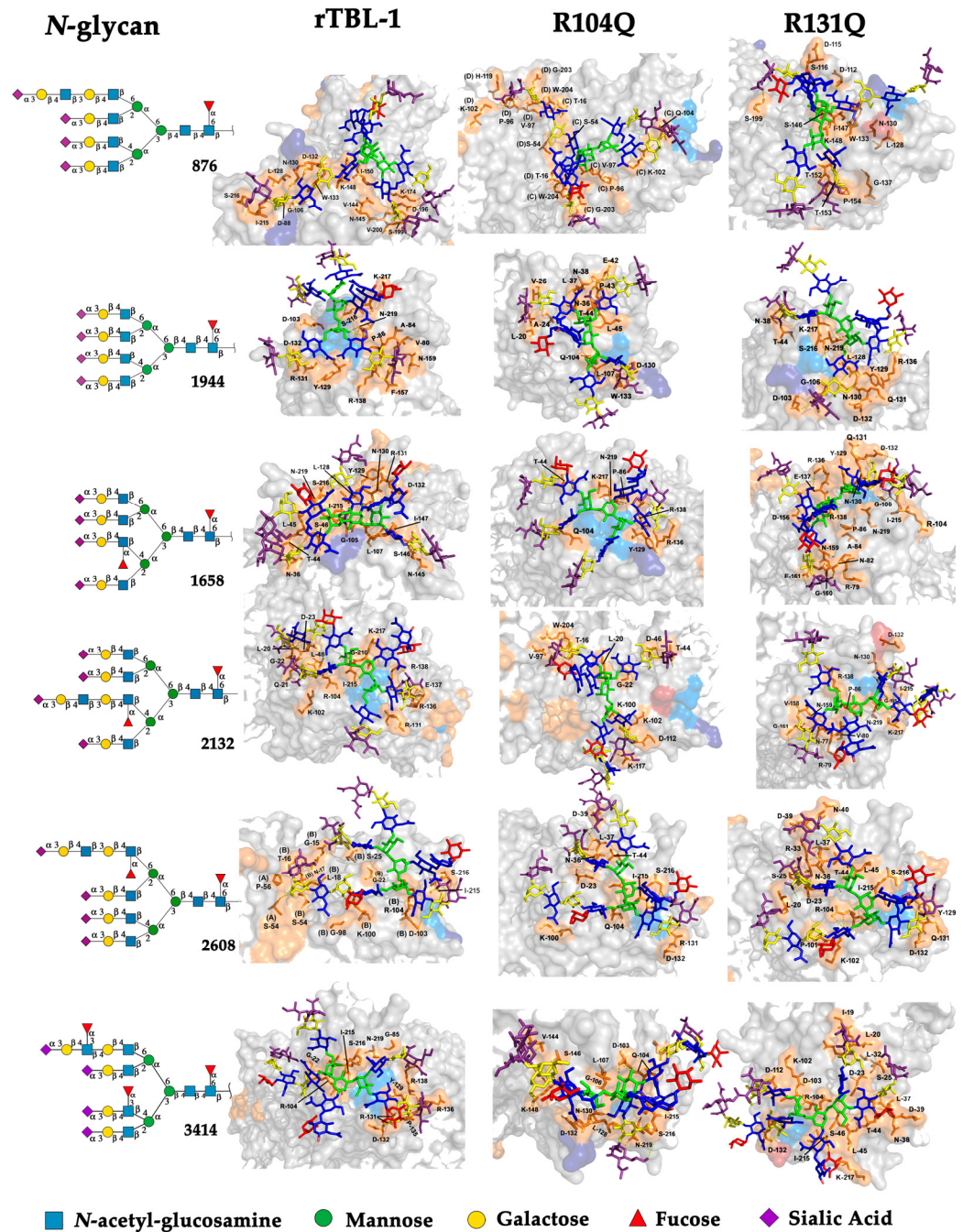


Figure 7. Three-dimensional maps for the interactions between EGFR-related N-glycans to rTBL-1, and mutants R104Q and R131Q. For each N-glycan, a graphical 2D representation according to SNFG [28] is shown. The interaction maps are shown as topological representations of the interactions of the N-glycans and the recombinant lectins: rTBL-1, R104Q, and R131Q. The target N-glycans are shown as sticks, and the structure of N-glycan residues is colored according to SNFG. The docking-interacting residues of lectins are shown as orange sticks with their respective surface and relative sequence number as labels.

Table 1. Comparison of the solvation free energy gain upon formation of the interface (Δ^iG) of the complex EGFR-related *N*-glycans with the rTBL-1 and mutated Tepary bean lectins (R104Q, and R131Q).

2D-Graphical <i>N</i> -glycan Structure According to (SNFG) [28]	<i>N</i> -glycan ID [30]	Reported Molecular Mass (kDa) [30]	Δ^iG (Kcal/mol)		
			rTBL-1	R104Q	R131Q
	876	4.05	-12.79	-6.57	-8.03
	1944	3.68	-7.99	-6.44	-8.20
	1658	3.83	-10.39	-13.80	-15.09
	2132	4.19	-10.99	-3.18	-11.75
	2608	4.19	-7.86	-7.50	-10.99
	3414	4.34	-12.59	-8.91	-5.83

■ *N*-acetyl-glucosamine
 ● Mannose
 ● Galactose
 ▲ Fucose
 ◆ Sialic Acid

rTBL-1, Recombinant lectin; R104Q, Recombinant R104Q mutant lectin; R131Q, Recombinant R131Q mutant lectin.

Another interesting finding is the probable effect of fucosylations on the recognition of *N*-glycans by tepary bean lectins. While the *N*-glycan 876 with elongated branch 1 did not have much interaction with CBP and even lost it when the residues were mutated, the *N*-glycan 2608, which had a similar structure but with a fucosylation on elongated branch 1, did have considerable interaction with CBP, which was maintained (and even

apparently enhanced in the R131Q mutant), suggesting that this fucosylation could be important to recognize this structure with an elongated branch 1 (compare Figure 5 and Table 1). Fucosylations in non-elongated branches apparently have no effect (compare interactions with 1944 and 1658 EGFR-related glycans in Figure 5); however, when a fucosylation in branch 3 is present, the values of the Δ^iG increased apparently even more in the interactions with mutated residues, indicating an apparent improvement in affinity between the interfaces of the interaction (compare Δ^iG values of 1944 and 1658 *N*-glycans in Table 1). In addition, the presence of a third fucosylation in *N*-glycans (i.e., 3414 EGFR-related *N*-glycan) also showed enhanced affinity in the case of the interaction with rTBL-1, but a considerable decrease in the interactions with R104Q and R131Q mutated lectins.

The docking results suggested that the observed loss of cytotoxicity of the mutants was related to the loss of affinity of the studied lectin-*N*-glycan interactions. In fact, each interaction depends on the glycan type suggesting a specific effect. Collectively, the results show that the interactions with *N*-glycans, such as 876, 1658, 2132, and 3414, could promote cytotoxicity since important changes in solvation free energy gain were determined for one or both mutants for each *N*-glycan. In particular, the substitution of R104 with Q104 decreased the affinity and (in some cases) the location of the interaction for three of the six tested *N*-glycans, and an important increase was only observed for one of them. On the other hand, for the substitution of R131 for Q131, an increase in affinity was observed for three *N*-glycans and a decrease for two. Both the decrease and increase in interactions compared to rTBL-1 could influence the properties of the lectin, where a decrease is primarily associated with a loss of affinity, while an increase could be linked to binding dissociation or a regulatory function.

3. Materials and Methods

3.1. Site-Directed Mutagenesis

To generate the mutants, two sets of primers containing the specific mutations were employed to amplify the recombinant plasmid pGAPZ α B-rTBL-1, previously generated by Martinez-Alarcon et al. [13]. The modifications were introduced near the 5' end of the reverse primers (underlined content). R103Q was generated using primers 126_F (5'-GGTCTTCTAGGTCTGTTCCGACGGC-3') and 126_R (5'-GCCTTGGTCTTTGGGCTTAGAGC-3'), while R130Q was generated using primers 152_F (5'-CAGAGAGCGTCATATTGGCATCG-3') and 152_R (5'-GGGTCCCAGTCCCTGGTTGTACAAG-3'). The amplified products were gel purified using a commercial BioLabs kit (Monarch DNA Gel Extraction. Hitchin, UK). Purified plasmids were subsequently ligated at 25 °C for 1 h using T4 ligase. The ligation products were transformed into *Escherichia coli* TOP10 cells (Invitrogen Life Technology, Cat. C4040-06. Waltham, MA, USA) via electroporation (resistance of 100 Ohms, capacitance of 125 μ Fa, and 1.8 Volts), and cells were plated on Luria–Bertani (LB) medium containing 25 μ g/mL zeocin as a selection agent, followed by incubation at 37 °C. After 24 h, transformants were screened by colony PCR (using gene-specific primers or pGAP α B primers), and positive clones were verified by digestion and sequencing.

3.2. Transformation of *Pichia pastoris*

The procedure was followed based on the patent MX/a/2018/009442. Sequence-verified plasmids were digested with *Avr*II at 37 °C for 16 h, and complete linearization of the DNA was confirmed through nucleic acid electrophoresis. Subsequently, the DNA was ethanol precipitated, and the pellet was resuspended in 20 μ L of nuclease-free sterile water. The concentration, determined by NanoDrop spectrophotometry (NanoDrop 2000c from Thermo Scientific; Hercules, CA, USA) and electrophoresis, was confirmed to be >5 μ g/ μ L. This DNA material was used to transform *Pichia pastoris* protease-deficient strain SMD1168H (Invitrogen Life Technology, Cat. C18400. Waltham, MA, USA) using the *Pichia* EasyCompTM Transformation Kit (Thermo Fisher Scientific; Pittsburgh, PA, USA) following the manufacturer's protocol. Transformed cells were recovered and plated on Yeast Extract–Peptone–Glycerol (YPG) medium supplemented with 50 μ g/mL zeocin and incubated at 30 °C for 48 h.

Ten colonies of *P. pastoris* were selected, and each one was inoculated into 10 mL of liquid YPG medium supplemented with 25 µg/mL zeocin. The cultures were incubated at 30 °C with shaking for 48 h, after which the supernatants were collected by centrifugation at 11,800× *g* for 10 min. For SDS-PAGE electrophoresis, 25 µL aliquots of each supernatant were loaded and transferred onto nitrocellulose membranes using a semi-dry transfer system (ATTO blotter, Tokyo, Japan) at a constant voltage of 10 V for 1 h in 1X TBS (Tris-Buffered saline, Tris-HCl 15.2 mM, Tris base 46.2 mM, and NaCl 1500 mM). Ponceau red staining was performed to visualize the positions of molecular size markers, followed by blocking the membranes with 5% (*w/v*) defatted milk powder in 0.1% (*v/v*) Tween TBS for 1 h. The membranes were incubated overnight with anti-histidine monoclonal antibodies at a 1:1000 dilution in blocking solution. After 16 h, the primary antibodies were removed by several washes with blocking solution, and the membranes were further incubated for 2 h in blocking solution containing secondary anti-goat antibodies coupled to alkaline phosphatase. Finally, chemiluminescence was used to visualize the immunoreactivity.

3.3. Heterologous Expression and Purification of R103Q and R130Q Mutants

P. pastoris cells expressing recombinant rTBL-1 were cultivated in a 7.5 L vessel Applikon ez-control laboratory fermenter, as described previously [13], with the exception of maintaining the pH at 5.0. After cultivation, the culture was centrifuged at 7500× *g* for 30 min at 4 °C to separate the secreted proteins from the cells. The resulting supernatant was clarified by sequential filtration using 2.7 µm and 0.7 µm glass fiber filters (Whatman, Kent, UK). Purification of the recombinant proteins from the supernatant was performed using nickel-affinity chromatography, followed by dialysis and freeze-drying, following the established protocol [14]. The protein content in the lyophilized samples was determined by SDS-PAGE gels, where the bands corresponding to intact proteins were compared to GNA (Sigma Aldrich, Burlington, MA, USA) standards through visual inspection and by capturing an image of the destained gel using a commercial flat-bed scanner. Image analysis was conducted using a custom-written software “ProQuantify”, version 1.0, supplied by Rodrigo Guerrero from Universidad Autónoma de Queretaro, Mexico.

3.4. Size Exclusion Chromatography

Size exclusion chromatography (SEC) was carried out using a High-Resolution ENrich™ SEC 650 column (Bio-Rad, Hercules, CA, USA) on an NGC™ chromatography system (Bio-Rad). Prior to the assay, the column was calibrated with the gel filtration standard #15119001 (Bio-Rad), following the supplier’s instructions. Protein samples at a concentration of 10 mg mL⁻¹ were centrifuged at 8000× *g* for 30 min. The column was equilibrated with 50 mL of buffer D (20 mM MES, 100 mM NaCl, pH 6.5), and then 200 µL of the sample were injected into the system, followed by a 40 mL isocratic elution using buffer D. Fractions were monitored at 280 nm for absorbance and collected.

3.5. Identification of the Glycosylation Present in the Recombinant Lectin

To determine the size and composition of the glycosidic antennas on Mut126 and Mut153, the proteins underwent glycan digestion using a combination of five distinct glycosidases that target O and N-glycosylations. This process was performed utilizing the Glycoprotein Deglycosylation kit (EMD Millipore, Darmstadt, Germany) according to the manufacturer’s instructions.

The glycosidases used were: (i) N-Glycosidase F: this enzyme cleaves all N-linked oligosaccharides, unless their core is α1-3 fucosylated; (ii) Endo-α-N-acetylgalactosaminidase that separates common nucleus (Galβ1,3GalNAc) from all O-glycosylations (it only works when no additional sugars are adhered to the main nucleus); (iii) α2-3,6,8,9-Neuraminidase that cleaves all sialic acids that may be attached to the common core of the O-glycosylations; (iv) β1,4-Galactosidase that releases only non-reducing terminal galactose, bound in β1-4 to O-glycosylations; (v) β-N-Acetylglucosaminidase that cleaves all GlcNAc residues bound to β-non-reducing terminal O-glycosylations. Following digestion, samples were subjected

to SDS-PAGE and subsequently stained with Coomassie blue. Additionally, the periodic acid-Schiff (PAS) staining technique (Sigma Chemical, St. Louis, MO, USA) was employed in accordance with the prescribed procedure based on the method originally proposed by Hotchkiss [43].

3.6. Cytotoxic Activity

To assess the cytotoxic activity of the mutants, HT-29 human colon cancer cells (ATCC, HTB-38, Manassas, VA, USA) were employed. Initially, 10,000 cells were seeded per well in 48-well plates containing 0.5 mL of medium A, which consisted of DMEM (Sigma-Aldrich, D-5796, Burlington, MA, USA) supplemented with 10% *v/v* fetal bovine serum (FBS) (Sigma-Aldrich, 12003C) and 1% *v/v* L-glutamine-antibiotics (Sigma-Aldrich, G6784). The plates were incubated at 37 °C with 5% CO₂ and 95% relative humidity until cells reached 80% confluence. After 48 h, the culture medium was replaced with medium B, composed of DMEM supplemented with 2% FBS *v/v* and 1% L-glutamine *v/v*, to synchronize the cells. Following an additional 24 h incubation, the medium was removed, and the cells were treated with DMEM supplemented with 0.5% BSA *w/v* (AlbuMax, Life Technologies, 11020021, Carlsbad, CA, USA) 1% L-glutamine *v/v*, and various concentrations of the respective treatments. PBS was used as the negative control, while rTBL served as the positive control. After 24 h incubation, the cells were harvested using 0.15% trypsin *w/v* and counted using a hemocytometer; triplicate wells were counted per treatment to determine cell numbers. Cell survival was calculated considering the negative control to be 100%. Lethal concentration 50 (LC₅₀) was calculated by plotting the log₁₀ of the treatment's concentration vs. cell survival % and taking the value of the ordered pair (*y*) as 50%.

Statistical analysis was determined by comparing the mean values for each treatment using a one-way ANOVA with a Tukey post hoc analysis (*p* < 0.05) to determine which means were significantly different.

3.7. Docking Analysis Procedure

EGFR-related *N*-glycans were identified using the GlyConnect platform (<https://glyconnect.expasy.org/> accessed on 8 February 2023) [30] from the ExPasy portal of the SIB Swiss Institute of Bioinformatics. The representative structures of the 52 EGFR-related glycans were chosen. Structures of *N*-glycan models were created using a "Build via Text" tool from the GLYCAM web server (<https://glycam.org> accessed on 8 February 2023) [29], based on the monosaccharide sequence on the GlyConnect platform. Structures from the main conformers were selected according to differences in the position of their glycan branches and were considered for the docking process.

To test the effect of mutations on lectin-glycan docking, two rTBL mutant structures were generated using the PyMOL software version 2.5.1 [44], considering the R103Q and R130Q experimentally obtained mutants. As the available rTBL-1 sequence (PDB ID: 6tt9) has one extra alanine beside the first serine of the rTBL-1 sequence [14], it was necessary to go through the numbering one position. Therefore, only for the docking analysis, R103 changes to R104, and R130 changes to R131.

Docking analysis of *N*-glycans with rTBL-1 and its two *in silico*-mutated structures was performed on the LZERD platform [23] according to its requirements. A list of 500 ranked models with their respective structure files was obtained for each docking, and those with higher Ranksum values were selected. However, docking models whose interactions were located in the tetrameric zone of the quaternary structure of rTBL-1 were discarded because of steric hindrance present between monomers [11]. Also, docking results were compared with the glycosylated recombinant Tepary bean lectin structure (PDB ID: 6tt9) in order to discard those whose interactions were located in the natural glycosylation site. Interactions near the mutation sites (R104Q and R131Q) and near the CBP were preferentially selected.

Interaction residues for each docking result were analyzed using the Protein-Ligand Interaction Profiler server [45,46] (PLIP-server, <https://plip-tool.biotech.tu-dresden.de/plip-web/plip/index> accessed on 20 February 2023) and the software Discovery Stu-

dio Visualizer version 21.1.0.20298 (Dassault Systèmes Biovia Corp. Velizy-Villacoublay, France) [47]. In addition, the types of interactions and the solvation free energy gain of the interfaces formed in these interactions were analyzed. using the PDBePISA web server [42] (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html accessed on 20 February 2023). Lectin-interacting residues of each docking result were identified and corroborated by PLIP-server [45]. Discovery Studio software [47] and the PDBePISA web server [42]. The total solvation energy gain upon formation of the interface of the interactions was obtained as the sum of the solvation energy gain corresponding to each of the protein-glycan interacting residues, expressed in kCal/mol, upon formation of the interface (Supplementary Material, Tables S2–S19).

4. Conclusions

The analysis of lectin sequences revealed that the *P. acutifolius* lectin, and consequently rTBL-1 have arginine residues at positions 103 and 130 (104 and 131 for docking analysis, respectively). The conservative substitution of either of these residues by glutamine in recombinant rTBL-1 mutants led to significantly decreased cytotoxicity towards colon cancer cells, suggesting that both residues are important for the establishment of lectin-*N*-glycan interactions. Moreover, R130Q inhibited cytotoxicity towards HT-29 human colon cancer cells by two-fold more than R103Q, which is indicative of distinct roles for each of the arginine residues.

Docking analysis for previously studied experimental *N*-glycans recognized by rTBL-1, showed correlation for the interaction in the CBP (close to R104 and R131). Just two of the experimentally recognized *N*-glycans interacted with the CBP, and these were highly sialylated and core-fucosylated *N*-glycans. The same pattern was observed for 32 of 52 EGFR-related *N*-glycans, where the low molecular mass (average mass: 1.77 ± 0.33 kDa) bisected *N*-glycans have no interaction at all, and 18 EGFR-related *N*-glycans with an average mass of 2.26 ± 0.25 kDa had similar interaction zones with rTBL-1 without interacting with CBP, such as those shown by the experimentally recognized *N*-glycans. Only the high molecular mass *N*-glycans (average mass: 4.05 ± 0.25 kDa) with sialylated four antennae and core-fucosylated structure had a direct docking interaction in the CBP of rTBL-1.

The interaction between the recombinant lectin and certain EGFR *N*-glycans occurs within a complex biological context, where the final effect depends on multifactorial parameters. However, this mutant study provides a better understanding of the molecular mechanism of action. It is also likely that EGFR is one of the target molecules for the rTBL-1, and many other membrane glycoproteins could be interacting, as aberrant sialylation and fucosylation are common in other glycoproteins involved in the development of many types of cancers, including colon cancer.

This research has expanded our understanding of how the Tepary bean lectin affects the microenvironment within which cancerous cells develop, creating new opportunities to uncover the mechanisms by which this lectin exerts its anticancer effects. Further studies are needed to gain a deeper understanding of the molecular interactions related to the lectin's cytotoxic effects to develop new strategies against EGFR-dependent cancers, particularly colon cancer.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms252413258/s1>.

Author Contributions: Conceptualization, D.M.-A., E.F., A.B.-L. and T.G.-G.; methodology, D.M.-A., J.L.C.-G., E.F., J.A.G. and S.P.; formal analysis, D.M.-A., E.F., S.P., A.B.-L. and T.G.-G.; investigation, D.M.-A., J.L.C.-G., E.F., A.B.-L. and T.G.-G.; data curation, D.M.-A., J.L.C.-G. and T.G.-G.; writing—original draft preparation, D.M.-A., J.L.C.-G. and T.G.-G.; writing—review and editing, J.L.C.-G., U.M.-C., A.B.-L. and T.G.-G.; visualization, J.L.C.-G., U.M.-C. and T.G.-G.; supervision, E.F., J.A.G., S.P., A.B.-L. and T.G.-G.; project administration, T.G.-G.; funding acquisition, T.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CONAHCYT Ciencia de Frontera (171873). D.M.-A. received a Master's scholar fellowship from CONAHCYT. Partial publication fee support was provided by the Universidad Autónoma de Querétaro.

Institutional Review Board Statement: The study was approved by the Ethics Committee of Facultad de Ciencias Naturales, Universidad Autónoma de Querétaro (protocol code 29FCN2021 on 29 April 2021).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data of mutants generation can be found in the repository link <https://repositorio.cinvestav.mx/handle/cinvestav/1634>.

Acknowledgments: We thank the technical support of Josué López-Martínez and the administrative support of Juan Trejo-Vega.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yau, T.; Dan, X.; Ng, C.C.; Ng, T.B. Lectins with potential for anti-cancer therapy. *Molecules* **2015**, *20*, 3791–3810. [CrossRef]
2. Mazalovska, M.; Kouokam, J.C. Plant-Derived Lectins as Potential Cancer Therapeutics and Diagnostic Tools. *Biomed. Res. Int.* **2020**, *2020*, 1631394. [CrossRef]
3. Wang, P.; Leng, X.; Duan, J.; Zhu, Y.; Wang, J.; Yan, Z.; Min, S.; Wei, D.; Wang, X. Functional Component Isolated from *Phaseolus vulgaris* Lectin Exerts In Vitro and In Vivo Anti-Tumor Activity Through Potentiation of Apoptosis and Immunomodulation. *Molecules* **2021**, *26*, 498. [CrossRef]
4. Kiss, R.; Camby, I.; Duckworth, C.; De Decker, R.; Salmon, I.; Pasteels, J.L.; Danguy, A.; Yeaton, P. In vitro influence of *Phaseolus vulgaris*, *Griffonia simplicifolia*, concanavalin A, wheat germ, and peanut agglutinins on HCT-15, LoVo, and SW837 human colorectal cancer cell growth. *Gut* **1997**, *40*, 253–261. [CrossRef]
5. Bardocz, S.; Grant, G.; Duguid, T.J.; Brown, D.S.; Pusztai, A.; Pryme, I.F. Intracellular levels of polyamines in Krebs II lymphosarcoma cells in mice fed phytohaemagglutinin-containing diets are coupled with altered tumour growth. *Cancer Lett.* **1997**, *121*, 25–29. [CrossRef]
6. Pryme, I.F.; Pusztai, A.; Bardocz, S.; Ewen, S.W. A combination of dietary protein depletion and PHA-induced gut growth reduce the mass of a murine non-Hodgkin lymphoma. *Cancer Lett.* **1999**, *139*, 145–152. [CrossRef]
7. Pryme, I.F.; Bardocz, S.; Pusztai, A.; Ewen, S.W. The growth of an established murine non-Hodgkin lymphoma tumour is limited by switching to a phytohaemagglutinin-containing diet. *Cancer Lett.* **1999**, *146*, 87–91. [CrossRef]
8. Valadez-Vega, C.; Morales-González, J.A.; Sumaya-Martínez, M.T.; Delgado-Olivares, L.; Cruz-Castañeda, A.; Bautista, M.; Sánchez-Gutiérrez, M.; Zuñiga-Pérez, C. Cytotoxic and antiproliferative effect of tepary bean lectins on C33-A, MCF-7, SKNSH, and SW480 cell lines. *Molecules* **2014**, *19*, 9610–9627. [CrossRef]
9. Moreno-Celis, U.; López-Martínez, F.J.; Cervantes-Jiménez, R.; Ferríz-Martínez, R.A.; Blanco-Labra, A.; García-Gasca, T. Tepary Bean (*Phaseolus acutifolius*) Lectins Induce Apoptosis and Cell Arrest in G0/G1 by P53(Ser46) Phosphorylation in Colon Cancer Cells. *Molecules* **2020**, *25*, 1021. [CrossRef]
10. Torres-Arteaga, I.; Castro-Guillén, J.L.; Mendiola-Olaya, E.; García-Gasca, T.; Ángeles-Zaragoza, M.V.; García-Santoyo, V.; Torres-Castillo, J.A.; Aguirre, C.; Phinney, B.; Blanco-Labra, A. Characterization of Two Non-fetuin-binding Lectins from Tepary Bean (*Phaseolus acutifolius*) Seeds with Differential Cytotoxicity on Colon Cancer Cells. *J. Glycobiol.* **2016**, *5*, 1000117. [CrossRef]
11. Torres-Arteaga, I.; Blanco-Labra, A.; Mendiola-Olaya, E.; García-Gasca, T.; Aguirre-Mancilla, C.; Ortega-de-Santiago, A.L.; Barboza, M.; Lebrilla, C.B.; Castro-Guillén, J.L. Comparative study, homology modelling and molecular docking with cancer associated glycans of two non-fetuin-binding Tepary bean lectins. *Glycoconj. J.* **2022**, *40*, 69–84. [CrossRef]
12. Moreno-Celis, U.; López-Martínez, J.; Blanco-Labra, A.; Cervantes-Jiménez, R.; Estrada-Martínez, L.A.; García-Pascalín, A.E.; Guerrero-Carrillo, M.J.; Rodríguez-Méndez, A.J.; Mejía, C.; Ferríz-Martínez, R.A.; et al. *Phaseolus acutifolius* Lectin Fractions Exhibit Apoptotic Effects on Colon Cancer: Preclinical Studies Using Dimethylhydrazine or Azoxi-Methane as Cancer Induction Agents. *Molecules* **2017**, *22*, 1670. [CrossRef]
13. Martínez-Alarcón, D.; Varrot, A.; Fitches, E.; Gatehouse, J.A.; Cao, M.; Pyati, P.; Blanco-Labra, A.; Garcia-Gasca, T. Recombinant Lectin from Tepary Bean (*Phaseolus acutifolius*) with Specific Recognition for Cancer-Associated Glycans: Production, Structural Characterization, and Target Identification. *Biomolecules* **2020**, *10*, 654. [CrossRef]
14. Palmerín-Carreño, D.; Martínez-Alarcón, D.; Dena-Beltrán, J.L.; Vega-Rojas, L.J.; Blanco-Labra, A.; Escobedo-Reyes, A.; García-Gasca, T. Optimization of a Recombinant Lectin Production in *Pichia pastoris* Using Crude Glycerol in a Fed-Batch System. *Processes* **2021**, *9*, 876. [CrossRef]
15. Dena-Beltrán, J.L.; Nava-Domínguez, P.; Palmerín-Carreño, D.; Martínez-Alarcón, D.; Moreno-Celis, U.; Valle-Pacheco, M.; Castro-Guillén, J.L.; Blanco-Labra, A.; García-Gasca, T. EGFR and p38MAPK Contribute to the Apoptotic Effect of the Recombinant Lectin from Tepary Bean (*Phaseolus acutifolius*) in Colon Cancer Cells. *Pharmaceuticals* **2023**, *16*, 290. [CrossRef]

16. Nagae, M.; Soga, K.; Morita-Matsumoto, K.; Hanashima, S.; Ikeda, A.; Yamamoto, K.; Yamaguchi, Y. Phytohemagglutinin from *Phaseolus vulgaris* (PHA-E) displays a novel glycan recognition mode using a common legume lectin fold. *Glycobiology* **2014**, *24*, 368–378. [CrossRef]
17. Radhakrishnan, A.; Chellapandian, H.; Ramasamy, P.; Jeyachandran, S. Back2Basics: Animal lectins: An insight into a highly versatile recognition protein. *J. Proteins Proteom.* **2023**, *14*, 43–59. [CrossRef]
18. Bianchet, M.A.; Ahmed, H.; Vasta, G.R.; Amzel, L.M. Structural aspects of lectin-ligand interactions. In *Animal Lectins: A Functional View*; Vasta, G.R., Ahmed, H., Eds.; CRC Press Taylor & Francis Group: Boca Raton, FL, USA, 2008; pp. 25–26. [CrossRef]
19. Yoshimura, A.; Hatanaka, R.; Tanaka, H.; Kitajima, K.; Sato, C. The conserved arginine residue in all siglecs is essential for Siglec-7 binding to sialic acid. *Biochem. Biophys. Res. Commun.* **2021**, *534*, 1069–1075. [CrossRef]
20. Ren, X.; Lin, S.; Guan, F.; Kang, H. Glycosylation Targeting: A Paradigm Shift in Cancer Immunotherapy. *Int. J. Biol. Sci.* **2024**, *20*, 2607–2621. [CrossRef]
21. Stroop, C.J.; Weber, W.; Gerwig, G.J.; Nimtz, M.; Kamerling, J.P.; Vliegenthart, J.F. Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor. *Glycobiology* **2000**, *10*, 901–917. [CrossRef]
22. Luke, B. *Beyond the Sequon: Sites of N-Glycosylation*; IntechOpen: London, UK, 2012. [CrossRef]
23. Venkatraman, V.; Yang, Y.D.; Sael, L.; Kihara, D. Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinform.* **2009**, *10*, 407. [CrossRef]
24. Khorana, A.A.; Ryan, C.K.; Cox, C.; Eberly, S.; Sahasrabudhe, D.M. Vascular endothelial growth factor, CD68, and epidermal growth factor receptor expression and survival in patients with Stage II and Stage III colon carcinoma: A role for the host response in prognosis. *Cancer* **2003**, *97*, 960–968. [CrossRef]
25. Luque-Garcia, J.L.; Martinez-Torrecuadrada, J.L.; Epifano, C.; Canamero, M.; Babel, I.; Casal, J.I. Differential protein expression on the cell surface of colorectal cancer cells associated to tumor metastasis. *Proteomics* **2010**, *10*, 940–952. [CrossRef]
26. Gan, Y.; Chen, D.; Li, X. Proteomic analysis reveals novel proteins associated with progression and differentiation of colorectal carcinoma. *J. Cancer Res. Ther.* **2014**, *10*, 89–96. [CrossRef]
27. Sethi, M.K.; Fanayan, S. Mass Spectrometry-Based N-Glycomics of Colorectal Cancer. *Int. J. Mol. Sci.* **2015**, *16*, 29278–29304. [CrossRef]
28. Varki, A.; Cummings, R.D.; Aebi, M.; Packer, N.H.; Seeberger, P.H.; Esko, J.D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25*, 1323–1324. [CrossRef]
29. Woods Group. GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA. 2005–2024. Available online: <http://legacy.glycam.org> (accessed on 8 February 2023).
30. Duvaud, S.; Gabella, C.; Lisacek, F.; Stockinger, H.; Ioannidis, V.; Durinx, C. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res.* **2021**, *49*, W216–W227. [CrossRef]
31. Zhou, Y.; Fukuda, T.; Hang, Q.; Hou, S.; Isaji, T.; Kameyama, A.; Gu, J. Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation. *Sci. Rep.* **2017**, *7*, 11563. [CrossRef]
32. De Freitas Junior, J.C.M.; Morgado-Diaz, J.A. The role of N-glycans in colorectal cancer progression: Potential biomarkers and therapeutic applications. *Oncotarget* **2016**, *7*, 19395–19413. [CrossRef]
33. Kim, H.J.; Kim, S.C.; Ju, W.; Kim, Y.H.; Yin, S.Y.; Kim, H.J. Aberrant sialylation and fucosylation of intracellular proteins in cervical tissue are critical markers of cervical carcinogenesis. *Oncol. Rep.* **2014**, *31*, 1417–1422. [CrossRef]
34. Yang, S.T.; Liu, C.H.; Chao, W.T.; Liu, H.H.; Lee, W.L.; Wang, P.H. The role of sialylation in gynecologic cancers. *Taiwan J. Obstet. Gynecol.* **2023**, *62*, 651–654. [CrossRef]
35. Park, J.-J.; Lee, M. Increasing the α 2,6 Sialylation of Glycoproteins May Contribute to Metastatic Spread and Therapeutic Resistance in Colorectal Cancer. *Gut Liver* **2013**, *7*, 629–641. [CrossRef]
36. Pearce, O.M.T.; Läubli, H. Sialic acids in cancer biology and immunity. *Glycobiology* **2016**, *26*, 111–128. [CrossRef]
37. Dorsett, K.A.; Marciel, M.P.; Hwang, J.; Ankenbauer, K.E.; Bhalerao, N.; Bellis, S.L. Regulation of ST6GAL1 sialyltransferase expression in cancer cells. *Glycobiology* **2021**, *31*, 530–539. [CrossRef]
38. Munkley, J. Aberrant Sialylation in Cancer: Therapeutic Opportunities. *Cancers* **2022**, *14*, 4248. [CrossRef]
39. Yen, H.Y.; Liu, Y.C.; Chen, N.Y.; Tsai, C.F.; Wang, Y.T.; Chen, Y.J.; Hsu, T.L.; Yang, P.C.; Wong, C.H. Effect of sialylation on EGFR phosphorylation and resistance to tyrosine kinase inhibition. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6955–6960. [CrossRef]
40. Britain, C.M.; Holdbrooks, A.T.; Anderson, J.C.; Willey, C.D.; Bellis, S.L. Sialylation of EGFR by the ST6Gal-I sialyltransferase promotes EGFR activation and resistance to gefitinib-mediated cell death. *J. Ovarian Res.* **2018**, *11*, 12. [CrossRef]
41. Dobie, C.; Skropeta, D. Insights into the role of sialylation in cancer progression and metastasis. *Br. J. Cancer* **2021**, *124*, 76–90. [CrossRef]
42. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [CrossRef]
43. Hotchkiss, R.D. A microchemical reaction resulting in the staining of polysaccharide structures in fixed tissue preparations. *Arch. Biochem.* **1948**, *16*, 131–141. Available online: <https://pubmed.ncbi.nlm.nih.gov/18921848/> (accessed on 22 January 2021).
44. DeLano, W.L. *PyMOL*, version 2.3.4; The PyMOL Molecular Graphics System; DeLano Scientific LLC: San Carlos, CA, USA, 2022. Available online: <https://pymol.org/2/> (accessed on 20 December 2022).
45. Adasme, M.F.; Linnemann, K.L.; Bolz, S.N.; Kaiser, F.; Salentin, S.; Haupt, V.J.; Schroeder, M. PLIP 2021: Expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucl. Acids Res.* **2021**, *49*, W530–W534. [CrossRef]

46. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucl. Acids Res.* **2015**, *43*, W443–W447. [[CrossRef](#)]
47. Dassault Systèmes. *BIOVIA Discovery Studio Visualizer*, version 21.1.0.20298; Dassault Systèmes: San Diego, CA, USA, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.