Lectin-carbohydrate interactions by protein bioinformatics: Parkia lectins case study

Benildo Sousa Cavada (bscavada@gmail.com)
Universidade Federal do Ceará (UFC)
Vinicius Jose Silva Osterne
Ghent University
Jorge Luiz Coelho Domingos
Universidade Federal do Ceará (UFC)
Messias Vital Oliveira
Universidade Federal do Ceará (UFC)
Francisco William Viana Martins
Universidade Federal do Ceará (UFC)
Francisco Vinicius Rodrigues Cruz
Universidade Federal do Ceará (UFC)
Francisco Edilcarlos Oliveira Lima
Universidade Federal do Ceará (UFC)
Henrique Sousa Oliveira
Universidade Federal do Ceará (UFC)
Jeanlex Soares Sousa
Universidade Federal do Ceará (UFC)
Wandemberg Paiva Ferreira
Universidade Federal do Ceará (UFC)
Kyria Santiago Nascimento
Universidade Federal do Ceará (UFC)
Vanir Reis Pinto-Junior
Universidade Federal do Ceará (UFC)

Research Article

Keywords: Lectins, Parkia, molecular dynamics

Posted Date: September 2nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1987770/v1

Abstract

Lectins are proteins that reversibly bind to carbohydrates without altering their structures. These proteins are present in practically all living beings and exert different functions. Understanding the molecular basis underlying the interaction between lectins and carbohydrates can help elucidate many biological activities of lectins. Parkia lectins have unique structural features within the legume family. They have protomers that associate as dimers, each with 3 β-prism domains, very similar to Moraceae lectins. This pattern is not conserved in any other Leguminoseae lectins. Each domain is unique in composition, but all have specificity for D-mannose and derivatives. This work aims to use docking and molecular dynamics approaches to characterize the interaction between *Parkia platycephala* (PPL) and *Parkia biglobosa* (PBL) lectins and D-mannose, building, as a result, a model to study lectin-carbohydrate interactions in general. MD trajectories demonstrate the stability of the lectins, whether in their native state or interacting with D-mannose. In addition, both molecular mechanics with generalized Born solvation and surface area (MM/GBSA) and molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) were used. When combined with the Interaction Entropy (IE) method, the binding energy of each domain with D-mannose was calculated to determine the participation of each amino acid in each domain during carbohydrate interaction. Trajectory analysis, as performed herein, has allowed for the expansion of knowledge about lectin-carbohydrate interactions based on our model, as well as the residues responsible for the binding with monosaccharides, thus contributing to future studies of Parkia lectins.

1. Introduction

Lectins comprise a group of proteins able to reversibly bind carbohydrates in a specific way through interaction with glycoconjugates in solution or on the cell surface. In particular, legume lectins stand out as a group with several purified and characterized proteins with closely related physicochemical properties and biological activities, as reported in the literature [1–3]. At the structural level, legume lectins are usually composed of a monomer of ~ 230–240 amino acids with around 25–30 kDa of molecular weight. The monomer folds into a jellyroll motif or β-sandwich and has two characteristic domains. The metal-binding site (MBS) binds divalent cations, such as calcium and manganese. It is essential for the formation of the second domain, the carbohydrate recognition domain (CRD). It is responsible for lectin activity [4]. These monomers can organize themselves into dimers or tetramers, depending on the lectin [5–7]. In general, this profile is found in most legume lectins with some exceptions. For example, lectins from the Parkia genus display a completely different structural organization in their tertiary and quaternary structures [2].

The Parkia genus is part of the Mimosoideae subfamily (Fabaceae family).The literature contains seven studies reporting on Parkia hololectins, as follows: *Parkia javanica* (PJL) [8], *P. speciosa* (PSL) [9], *P. discolor* (PDL) [10], *P. biglandulosa* (PbiL) [11], *P. pendula (PpeL)* [12], *P. roxburghii* (PRL) [11], *P. panurensis* (PpaL) [13], *P. biglobosa* (PBL) [14, 15] and *P. platycephala* (PPL-1) [16]. These lectins are specific to mannosides and glycosides with optimal activity in a pH range between 6 and 8 and stable

above 60°C. Structurally, these proteins share a high degree of similarity at primary structure level. Parkia lectins present molecular mass ranging between 45 and 49 kDa and are composed of ~ 447 amino acids [2]. Only two lectins have had their three-dimensional structure experimentally determined. PPL structure complexed with Xman (5-bromo-4-chloro-3-indolyl-α-D-mannose) was elucidated to a resolution of 2.5 Å, and PBL complexed with α-methyl-mannoside was resolved at a resolution of 2.12 Å, both by X-ray crystallography [15]. PBL and PPL monomers consist of three β-prism domains tandemly arranged, each one presenting a different CRD, but all with the same specificity. Furthermore, two monomers are able to associate to form a dimer, totaling six CRDs in the native structure [15, 17]. Parkia lectins present antinociceptive and anti-inflammatory activities in mice, as well as antiproliferation of lymphoma and hybridoma murine cell lines [11, 14, 15]. Therefore, understanding the molecular behavior of these proteins when interacting with carbohydrates is important to elucidate the biological mechanism(s) that underlie(s) such biological activities.

Several works have demonstrated the behavior of legume lectins that present β-sandwich folding, but no work has, so far, demonstrated simulations of proteins with folding related to Parkia lectins [18–20]. Overall, therefore, this work aims to investigate the dynamic behavior of PPL and PBL, as well as properties of their respective carbohydrate recognition domains (CRDs) and relevant residues, analyzing in detail the interactions between the lectins and D-mannose by molecular docking and dynamics.

2. Methods

2.1 Ligand-binding predictions

The initial coordinates of PBL and PPL in complex with D-mannose (MAN) applied in molecular dynamics simulations were obtained by molecular docking. Inputs for the simulations were obtained according to the following procedures. Ideal MAN coordinates were downloaded from PubChem (MAN code), and both PPL and PBL crystallographic coordinates were obtained from the Protein Data Bank (PDB) with access codes 1ZGS [17] and 4MQ0 [15], respectively. GOLD protein-ligand docking software, v. 2.0 [21, 22], was chosen for simulations with the following settings: number of operations, 10,000; number of poses, 30; population size, 100; selection pressure, 1.1; number of islands, 5; niche size, 2; crossover frequency, 95, and ChemPLP as the scoring function [23]. The simulation was performed at the center of each carbohydrate-recognition domain (CRD) and all atoms in a 6 Å radius. Each CRD was evaluated individually, and the best poses were selected, considering the score, hydrogen bonds, nonpolar interactions, and ligand geometry penalties [24, 25]. The algorithm's validity was assessed by redocking the crystallographic ligands with the respective lectins, 5-bromo-6-chloro-3-indolyl-α-D-mannopyranoside (X-man) for PPL and α-methyl-D-mannopyranoside for PBL.

2.2 Molecular dynamics simulations

Molecular dynamics simulations were performed for PPL and PPL dimers, both unliganded and bound to MAN, to compare the differences in structure and dynamic behavior between the complexed and uncomplexed forms of the lectins. Four different systems were set as inputs for the simulations: PPL-

only, PBL-only, PPL-MAN and PBL-MAN. Files for the simulations were prepared in the CHARMM-GUI web server [26, 27] and applied on the AMBER21 suite [28] with the parameters of CHARMM36m force field [29]. All systems were solvated with the TIP3P water model and neutralized with counter ions (Na^+). 90,679, 90,642, 90,647 and 90,599 water molecules were added to PPL-only, PBL-only, PPL-MAN and PBL-MAN systems, respectively. All systems were subjected to energy minimization steps using the combination of steepest descent and conjugate gradient algorithms with 10 kJ/mol as convergence criteria, and the minimized systems were equilibrated with 600,000 steps (500 ps) of NVT followed by the same number of steps under the NPT ensemble. Temperature and pressure of systems were set to 300 K and 1 bar maintained by a Langevin thermostat and an isotropic Monte-Carlo barostat [30], respectively. Time step for the simulations was set to 2 fs, and the SHAKE algorithm [31] was applied to constrain covalent bonds involving hydrogen atoms. Particle Mesh Ewald (PME) [32] was applied to calculate long-distance electrostatic interactions with a threshold of 10 Å.

2.3 Trajectory analysis

The following analyses were performed with the generated trajectory: root mean square deviation (RMSD), root mean square fluctuations (RMSF), radius of gyration (RoG), intermolecular hydrogen bonds (HB), contact frequency (CF) and solvent accessible surface area (SASA). Cpptraj [33], Xmgrace and VMD [34] were used to extract and analyze the trajectory data.

2.4 Binding free energy calculation

Binding free energies of the complexes between PPL or PBL with D-mannose ligands were estimated by molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) and molecular mechanics/Generalized Born (GB) solvent accessible surface area (MM/GBSA) [35, 36]. The entire simulation was applied according to Eq. (1), estimated using Eq. (2), as follows:

$$\Delta G_{binding} = G_{complex} - (G_{protein} + G_{ligand}), \quad (1)$$

where $G_{complex}$ is the free energy of the complex, $G_{protein}$ is the free energy of protein, and G_{ligand} is the free energy of ligand, and

$$G = E_{gas} + G_{solvation} - TS, \quad (2)$$

where E_{gas} refers to molecular mechanics free energy in gas phase, including the contributions of electrostatic, van der Waals, and internal strain energy. Solvation free energy ΔG_{solv} includes polar (ΔG_{PB}) and nonpolar (ΔG_{SA}) contributions (Eq. 3), as

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{SA} \quad (3).$$

PB and GB (MM/GBSA approach) models were used to estimate the polar contribution. The interior and exterior dielectric constants were set to 1 and 80, respectively. Nonpolar solvation energy was estimated by solvent-accessible surface area (SASA); for this, the surface tension proportionality constant γ was set

to 0.00542 kcal× mol$Å^2$, and b, the free energy of nonpolar solvation for a point solute, was set to 0.92 kcal× mol. Sphere radius for SASA calculation was 1.4 Å. After simulation, the energy file was used to generate the binding free value of PPL and PPL with mannose ligands. The contribution for each relevant residue for binding was generated by VMD using the output of MM/PBSA as input for the software.

2.5 Interaction Entropy (IE) method

In this section, we briefly discuss the IE method used to obtain the entropic term. Protein-ligand binding free energy (ΔG) can be given by

$$\Delta G = \Delta G_{gas} + \Delta G_{solv}, \quad (1)$$

where ΔG_{gas} stands for gas-phase free energy, and ΔG_{solv}stands for solvation free energy. Rewriting the gas-phase component [37, 38], we have

$$\Delta G_{gas} = \frac{1}{\beta}\frac{\int dq_w dq_p dq_l e^{-\beta\left(E_p+E_l+E_{pl}^{int}+E_w+E_{pw}^{int}+E_{lw}^{int}\right)}}{\int dq_w dq_p dq_l e^{-\beta\left(E_p+E_l+E_w+E_{pw}^{int}+E_{lw}^{int}\right)}},$$

2

where β = 1/k_BT and k_B is the Boltzmann constant, T is the absolute temperature, E_p, $E_l$and E_w are internal energies of the protein, ligand and waters, respectively. E_{pl}^{int} and E_{pw}^{int} are interaction energies of protein-ligand, protein-water and ligand-water, respectively. The protein-ligand interaction incorporates electrostatic and van der Waals interactions. After integration and some mathematical manipulations of Eq. (2), we have

$$G_{gas} =< E_{pl}^{int} > + k_B T\ ln\left(< e^{\beta \Delta E_{pl}^{int}} >\right), G_{gas} =< E_{pl}^{int} > T\Delta S. \quad (3)$$

Therefore, the IE is defined as

$$-T\Delta S = k_B T ln\left(< e^{\beta \Delta E_{pl}^{int}} >\right), \quad (4)$$

where ΔE_{pl}^{int} is the fluctuation of protein-ligand interaction energy around the ensemble averaged energy, as defined by

$$\Delta E_{pl}^{int} = E_{pl}^{int} - < E_{pl}^{int} >$$

5

,

where $< E_{pl}^{int} >$ represents the ensemble average of protein-ligand interaction energy, which is given by

$$< E_{pl}^{int} >= \frac{1}{t}\int_0^t E\left(t'\right) dt' = \frac{1}{N}\sum_{j=1}^{N} E_{pl}^{int}\left(t_j\right),$$

6

where

$$< e^{\Delta E_{pl}^{int}} >= \frac{1}{N}\sum_{j=1}^{N} e^{\beta \Delta E_{pl}^{int}}$$

7

.

Therefore, taking G_{solv} from the MM/PBSA method and G_{gas} obtained by the IE method, we obtain ΔG_{final}.

3. Results And Discussion

3.1 Molecular docking

Molecular docking was employed to generate inputs for the molecular dynamics simulations and to generate data about the binding of these two lectins with their main monosaccharide binder. Chemscores for the binding with each site can be seen in Table 1.

Table 1
Docking results of PPL and PBL domains interacting with D-mannose (in arbitrary units).

PPL	**Final score**	**Hbond score**	**Nonpolar score**	**Buried score**	**Ligand torsion**
Domain 1	-22.41	-5.73	-15.28	-1.40	0.00
Domain 2	-24.81	-7.94	-21.17	4.30	0.00
Domain 3	-27.61	-6.43	-16.61	-4.74	0.17
PBL	**Final score**	**Hbond score**	**Nonpolar score**	**Buried score**	**Ligand torsion**
Domain 1	-29.86	-7.30	-18.13	-4.43	0.00
Domain 2	-36,93	-8.55	-25.35	-3.03	0.00
Domain 3	-32,69	-6.97	-19.71	-6.01	0.00

It can be seen from these scores that binding to mannose is not similar in the two lectins, with PBL showing higher score values than PPL in all domains. A previous work by Bari and colleagues (2016) [15] performed an in-depth analysis to understand the interaction of *Parkia platycephalla* and *Parkia biglobosa* lectins with monosaccharides by applying molecular docking, and the authors verified

differences between domains resulting from their different sequences. Such differences lead to different topologies that finally affect the number of interactions in each site. The interactions of the three sites of PPL and PBL with MAN can be seen in the ligplots of Figs. 1 and 2, and are in line with previous observations. These data were applied in subsequent experiments of molecular dynamics to expand information and generate more data about the binding and overall dynamics of Parkia lectins.

3.4 Molecular dynamics

Four systems were set in CHARMM-GUI, including PBL-only, PPL-only, PPL-MAN and PBL-MAN, which ran in the pmemd.cuda module of AMBER21 suite. For minimization, all systems were subjected to 7500 steps of the steepest descent method and 5,300, 5,800, 6,000 and 6,100 steps of conjugate gradient method for PBL, PBL-MAN, PPL and PPL-MAN, respectively. The energy of the systems was around −1.18e + 06 kcal/mol. The MD simulation was set to run until 100,000 frames with 2 fs time step having been generated, totaling 100 ns. Initial trajectory analysis indicated that the simulations ran without any glaring issues, and for the simulations with mannose, the ligands maintained themselves inside the CRD for the entirety of the run, which is an indication of specificity and correct ligand parameterization.

3.5 Trajectory analysis

3.5.1 Root-mean square deviation and root-mean square fluctuation

RMSD data show that each simulation equilibrated at different time points, except PPL-MAN, which equilibrated at around 40 ns, while the aforementioned PPL-MAN reached equilibrium quite fast at 10 ns, with values varying between 2.0 and 4.5 Å (Fig. 3A).

This suggests that the number of steps was adequate for the analysis and that the systems stably transitioned from an energy-minimized state to a fully solvated state. The average RMSD values for PBL, PBL-MAN, PPL and PPL-MAN are 3.56 Å, 3.08 Å, 3.13 Å, and 2.41 Å, respectively. These results demonstrate small non-significant deviations between bound and unbound PBL and PPL. They also indicate that the impact of binding to MAN is similar on both lectins in relation to the reduction of RMSD values. For PBL, presence of the ligand did not affect the time for the system to equilibrate, while for PPL, the complexed form equilibrated much faster than its unbound counterpart. Considering the similarity between these lectins, this is a surprising fact. Both lectins deviated less when complexed with MAN in comparison to their unbound form. It should be noted that this was only a slight change for PBL, whereas the variation was quite noticeable for PPL. Equilibration is highly variable protein- to-protein and simulation-to-simulation, but it is generally accepted that ligand binding usually helps in shifting the simulation to an equilibrated state in a smaller number of steps owing to CRD stabilization in a closed position, as observed by Nascimento and colleagues (2018) with lectins from *Arachis* [39] and Cavada et al. (2019) [40] for the lectins of *Canavalia ensiformis* and *Canavalia brasiliensis*.

Using theoretical B-factors, RMSF data (Fig. 3B) were applied to verify the flexibility of the proteins per residue during all simulations. As shown in the RMSF plot, fluctuation was dramatically different between PPL and PBL, both in unliganded and complexed forms. Overall, PBL residues fluctuated more than PPL residues in all cases. The binding with MAN also differentially affected both lectins. For PPL, the binding had very little effect on fluctuation; meanwhile, for PBL, the binding was very stabilizing, in particular the regions from 150–300 and 600–900, which include residues from PBL's CRD.

The binding of lectins with their specific respective sugars tends to cause local stabilization with consequent reduction of fluctuation. This was observed for other legume lectins, such as *Canavalia bonariensis* [41], although exceptions can be pointed out, such as *Arachis* lectins for which specific ligand binding had no effect or increased fluctuation on specific residue ranges [39]. These results clearly show that very similar lectins can present dramatically different properties and that these differences are especially evident when dynamics properties are considered. The structures of Parkia lectins are dramatically different from those of other legume lectins, but their overall binding follows the same general rules.

3.5.2 Radius of gyration

Radius of gyration analysis evaluates the folding and compaction of the lectin. This analysis was applied in the current work to evaluate the differences caused by ligand binding. Results show that differences in RoG between lectins, both bound and unbound, were very small, suggesting that Parkia lectins do not change their compaction during binding. The average RoG values for PBL, PBL-MAN, PPL and PPL-MAN are 32.55 Å, 32.60 Å, 32.39 Å and 32.17 Å, respectively. The overall shallowness of Parkia lectin CRDs could explain these results. This property was also observed for the *Dioclea lasiophylla* lectin [19]. In this example, no significant changes in RoG could be observed after binding with monosaccharides or oligomannose *N*-glycans, suggesting that ligand binding is not necessarily a factor affecting the compaction of lectin structure. The energy x RoG graph (Fig. 4) demonstrates that PPL and PBL adopt conformations that vary between − 8500 and − 7250 kcal/mol, with little difference between their unbound and bound states.

3.5.3 H-bonds and contacts analysis

A few differences could be observed in the number of H-bonds between sites and lectins. Previous works have already reported differences in sites that could lead to variability in carbohydrate interactions [5] (Supplementary Fig. 1). Average H-bonds showed differences between lectins with PBL CRD 1 averaging 2.5 H-bonds with MAN, PBL CRD 2 averaging 4.5, and PBL CRD 3 averaging 3, while for PPL, averages were 2.5, 4 and 3, respectively. However, for all sites, one can see an overall stability with very few frames with 0 bonds. Also, at no point during the trajectory did any MAN leave the binding sites, which can happen, even when the ligand is specific to the lectin, as verified for Arachis lectins. **S**tability in interactions suggests a more favorable outcome.

Contact frequency analysis for PPL and PBL is depicted in Fig. 5. At first glance, it can be seen that the contacts with MAN are divided between very high contact residues (> = 80%) and very low contact residues (< = 20%), which are probably random interactions.

High contact residues include the following: For PBL CRD 1, Gly15, Gly16, Ile92, Gly136, Tyr137, Tyr138 and Asp140; for PBL CRD 2, Ala163, Gly164, Asp165, Phe238, Val240, Gly280, Gly281, Tyr282, Tyr283 and Asp285; for PBL CRD 3, Gly309, Gly310, Phe387, Gly430, Asp431, Tyr432 and Asp434. For PPL CRD 1, Gly15, Gly16, Val92, Gly136, Tyr137, Tyr138 and Asp140; for PPL CRD 2, Ala163, Gly164, Asp165, Phe239, Val241, Ser281, Gly282, Tyr283, Tyr284 and Asp286; for PPL CRD 3, Gly310, Gly311, Phe388, Gly431, Asp432, Tyr433 and Asp435. Despite small differences, the frequencies are very similar between the two lectins. When comparing this with the static analysis, qualitatively, almost all residues have been previously reported, but not their frequencies; we have now done this, and as a result, the CRDs of both Parkia lectins tested in the current work could be completely defined.

3.5.4 SASA analysis

SASA analysis of the systems was performed to understand the behavior of the protein solvent and evaluate the impact of interaction with carbohydrates on this behavior (Supplementary Fig. 2A-B). The solvent can interact with the protein by polar and nonpolar interactions, but the interaction with carbohydrates can change the conformation of the protein in a way that affects solvent accessibility. As expected, PBL and PPL undergo a reduction in SASA after interacting with ligands since carbohydrates interact with the surface of the molecule in the CRD cavity.

The average SASA values for the dimers of PBL and PBL-MAN are 35340 Å and 34005 Å, respectively (Supplementary Fig. 2C). According to the evaluation of each domain separately in the PBL system (Supplementary Fig. 3A-C), domain 1 presented an average value of 5327/5479 Å (monomer1/monomer2), domain 2 presented 6399/6193 Å, and domain 3 presented 6021/5919 Å. In the PBL-MAN system, the domains presented the following average values: domain 1 is 5291/5300 Å, domain 2 is 5968/5853 Å and domain 3 is 5761/5630 Å. As shown in Supplementary Fig. 4A-C, CRD1 (domain 1) had an average SASA value of 462/481 Å for PBL and 356/404 Å for PBL-MAN, CRD2 (domain 2) was 827/933 Å for PBL and 455/571 Å for PBL-MAN, and CRD3 (domain 3) was 748/472 Å for PBL and 547/281 Å for PBL-MAN.

For the PPL and PPL-MAN systems, the average SASA values for the dimers were 33493 Å and 32041 Å (Supplementary Fig. 2C), respectively. The average values of domains 1, 2 and 3 from the PPL system were 5428/5410 Å, 5432/5452 Å and 5809/5960 Å (Supplementary Fig. 3D-F), and their respective CRDs were 525/420 Å, 671/626 Å and 688/686 Å (Supplementary Fig. 4D-F). Regarding the PPL-MAN system, the average SASA value of domain 1 was 5191/5248 Å, domain 2 was 5123/5230 Å and domain 3 was 5565/5682 Å, while their respective CRDs were 361/262 Å, 461/462 Å and 551/544 Å.

PPL has a slightly smaller SASA area than PBL, as reflected at the dimer and domain level, but in domain 2, a more significant difference could be observed, even at the CRD level. In the unbound state, CRD2 of

PBL shows a higher SASA area, followed by CRD3 and, with greater difference, CRD1. PPL, on the other hand, has less discrepant SASA values among its CRDs and relatively lower values than those of PBL. In the bound state with MAN, reductions in SASA values can be seen for both lectins, being more intense in PBL, specifically in CRD2. SASA values of both lectins in the bound state tend toward similarity.

Despite the high similarity between PBL and PPL, it was possible to observe that both molecules have structural differences in their unbound and MAN-bound state in relation to their global structure, which tend to decrease in their bound state at the CRD level. Some small differences in SASA values were observed between monomer 1 and monomer 2, which was more intense in CRD3 of PBL. These differences were not random since they were observed in the PBL and PBL-MAN systems, as well as in repetitions performed. Thus, the interaction between PBL monomers does not have the same impact on the corresponding domains of each protomer. This was not observed in PPL or in other evaluated lectins.

3.5.5 Binding free energy

The binding free energy of PPL and PBL with mannose molecules was calculated using MM/GBSA and MM/PBSA combined with Entropy Interaction (EI) for entropy calculation. The results obtained are presented in Table 2. The binding free energy of PBL is similar to that of PPL, but domain 3 of PBL binds more favorably than domain 3 of PPL. Despite this difference, both have a similar pattern of binding in that domain 1 has higher free energy, followed by domains 2 and 3. The energy decomposition analysis demonstrated the participation of each amino acid residue in the binding free energy (Figs. 6 and 7). For PPL, domain 1 residues Gly15, Gly16, Gly55, Gly56, Lys57, Val92, Val94, Arg134, Ala135, Gly136, Tyr137, Tyr138 and Leu139 contributed favorably to the binding energy, while Leu93 and Asp140 interfere energetically to the binding free energy. Domain 2 residues Ala163, Gly164, Gly204, Gly205, Gln206, Phe239, Val241, Lys280, Ser281, Gly282, Tyr283 and Tyr284 contribute favorably, while Glu240 and Asp286 interfere with the interaction. Domain 3 residues Gly310, Gly311, Gly350, Gly351, Val352, Phe388, Thr390, Arg429, Ala430, Gly431 and Tyr433 contribute to the interaction, while Asp312, Thr389 and Asp435 caused interference. It is interesting to note that some residues in contact with the ligand do not favor interaction, including Asp140, Asp286 and Asp435, all of which have a contact frequency of 100%. Other residues that have high frequency of contact with the ligand, but still do not favor interaction, include Gly55, Gly56, Val94, Lys57, Ala135, Leu139, Gly204, Gly205, Gln206, Lys280, Gly350, Gly351, Val352, Thr390, Arg429 and Ala430. Instead, these residues favor the creation of a hydrophobic or electrostatic environment that does favor interaction.

Table 2
Energy contributions for the interaction of PPL and PBL with D-mannose in kcal/mol.

MM/GBSA						
PPL				PBL		
Energy component	MN1	MN2	MN3	MN1	MN2	MN3
ΔG_{gas}	-97.70 ± 5.23	-73.17 ± 4.98	-65.00 ± 4.45	-97.71 ± 5.22	-71.10 ± 2.45	-72.90 ± 5.02
ΔG_{solv}	57.73 ± 2.73	43.52 ± 3.72	40.64 ± 4.14	57.73 ± 2.72	44.31 ± 2.27	43.38 ± 3.83
ΔG_{total}	-39.97 ± 2.70	-29.65 ± 2.73	-24.35 ± 2.50	-39.98 ± 2.70	-26.79 ± 2.29	-29.52 ± 2.70
Entropy	9.94	8.45	6.43	9.89	3.08	8.22
ΔG_{final}	-30.03 ± 2.70	-21.20 ± 2.73	-17.93 ± 2.50	-30.09 ± 2.70	-23.71 ± 2.29	-21.30 ± 2.70
MM/PBSA						
Energy component	MN1	MN2	MN3	MN1	MN2	MN3
ΔG_{gas}	-97.70 ± 5.23	-73.17 ± 4.98	-65.00 ± 4.45	-97.71 ± 5.22	-71.10 ± 2.45	-72.90 ± 5.02
ΔG_{solv}	67.24 ± 2.77	53.93 ± 4.18	56.49 ± 4.81	67.26 ± 2.76	53.18 ± 3.07	53.84 ± 4.26
ΔG_{total}	-30.46 ± 3.49	-19.24 ± 3.73	-8.51 ± 3.29	-30.45 ± 3.49	-17.92 ± 2.79	-19.06 ± 3.67
Entropy	9.94	8.45	6.43	9.89	3.08	8.22
ΔG_{final}	-20.52 ± 3.49	-10.79 ± 3.73	-2.08 ± 3.29	-20.56 ± 3.49	-14.84 ± 2.79	-10.84 ± 3.67

For PBL, domain 1 residues that contributed favorably to binding were Gly15, Gly16, Gly55, Gly56, Lys57, Ile92, Val94, Arg134, Ala135, Gly136, Tyr137, Tyr138 and Asp140, while Asp17 and Leu93 interfere with binding. In domain 2. Ala 163, Gly164, Asp165, Gly203, Gly204, Phe238, Val240, Lys279, Gly280, Gly281, Tyr282 and Tyr283 favor binding, while Glu239 and Asp285 interfere with the interaction. Finally, in domain 3, Gly309, Gly310, Gly349, Gly350, Phe387, Thr389, Arg428, Ala429, Gly430, Asp431 and Tyr432 contribute to free binding energy, whereas Asp311, Thr388 and Asp434 interfere with interaction. Residues Gly56, Lys57, Arg134, Ala135, Gly203, Gly204, Lys279, Gly349, Gly350, Thr389, Arg428 and Ala429 lack high frequency of contact, but they still participate in the interaction. However, residues Asp285 and Asp434 also have a high frequency of contact, but they do not favor the binding energy.

Although the binding energies of PPL and PBL are similar in domain 1 and domain 2, it is possible to notice some differences in the contribution of some residues, mainly in the interference level of Asp140 (domain 1) and Asp286 (domain 2) in PPL, which is greater, but balanced by the level of contributions from favorable interactions to a more negative free energy. Domain 3 binding energy is much lower for PPL since Asp435 interference is 9.5x greater than that of Asp434 in PBL. Furthermore, the energetic contribution of Gly309, Gly310, Gly430, Asp431 and Tyr432 is higher in PBL than their equivalents in PPL.

The interference of Asp residues present in the domains is greater in PPL, possibly resulting from amino acid differences between their structures and the additional amino acid residue present in PPL structure, which changes the position and behavior of Asp residues, which disfavor interactions of PPL with mannose.

The information obtained from dynamics trajectories provided more details of the interaction of Parkia lectins with their specific ligands. As presented in this work, the level of structural data on the interaction with carbohydrates expanded the work of Bari et al. [15], who only evaluated static structures, resulting in an increase in the number of residues that compose, in part, the CRDs of these two lectins.

4. Conclusion

Trajectory analysis allowed the determination of molecular stability and behavior of Parkia lectins in solution, as well as the expansion of knowledge about interaction with carbohydrates and the residues responsible for the binding with monosaccharides for this group of lectins. We will continue to add experimental data in future studies to confirm the conclusions of the present work.

Declarations

Acknowledgements

This study was supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP). B.S.C, J.S.S, W.P.F and K.S.N are senior investigators of CNPq.

Disclosure statement

No potential conflict of interest was reported by the authors.

Availability of data

Data available on request from the authors.

Competing interests

The authors declare no competing interests.

Code Availability

N/A.

Ethics approval

N/A.

Consent to participate

N/A.

Consent for publication

N/A.

Author Contributions

Benildo Sousa Cavada and Vanir Reis Pinto-Junior built the project and coordinated the team to execute the docking and molecular dynamics simulation processes and were responsible for writing the manuscript. Kyria Santiago do Nascimento, Jeanlex Soares Sousa and Wandemberg Paiva Ferreira reviewed the docking and molecular dynamics results and optimized simulation conditions during the execution of the methodologies. In addition to providing infrastructure for project execution. Vinicius Jose Silva Osterne., Messias Vital Oliveira. and Francisco William Viana Martins performed docking and molecular dynamics of the PPL lectin, created figures, supplementary figures and tables. Francisco Vinicius Rodrigues Cruz, Francisco Edilcarlos Oliveira Lima, Henrique Sousa Oliveira performed docking and molecular dynamics of the PBL lectin, created figures, supplementary figures and tables. Jorge Luiz Coelho Domingos and Vanir Reis Pinto-Junior were responsible for the binding free energy calculations by MMGBSA, MMPBSA and for the entropy calculation using the interaction entropy method.

References

1. Van Damme E (1998) Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles. Critical Reviews in Plant Sciences 17:575–692
2. Cavada BS, Osterne VJS, Oliveira MV, et al (2020) Reviewing Mimosoideae lectins: A group of under explored legume lectins. Int J Biol Macromol 154:159–165
3. Barre A, Van Damme EJM, Simplicien M, et al (2021) Man-Specific Lectins from Plants, Fungi, Algae and Cyanobacteria, as Potential Blockers for SARS-CoV, MERS-CoV and SARS-CoV-2 (COVID-19) Coronaviruses: Biomedical Perspectives. Cells 10.: https://doi.org/10.3390/cells10071619

4. Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) Legume lectin structure. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1383:9–36
5. Cavada B, Pinto-Junior V, Osterne V, Nascimento K (2018) ConA-Like Lectins: High Similarity Proteins as Models to Study Structure/Biological Activities Relationships. International Journal of Molecular Sciences 20:30
6. Nascimento KS, Silva MTL, Oliveira MV, et al (2020) Dalbergieae lectins: A review of lectins from species of a primitive Papilionoideae (leguminous) tribe. Int J Biol Macromol 144:509–526
7. Cavada BS, Pinto-Junior VR, Osterne VJS, et al (2020) Comprehensive review on Caelsalpinioideae lectins: From purification to biological activities. Int J Biol Macromol 162:333–348
8. Utarabhand P, Akkayanont P (1995) Purification of a lectin from Parkia javanica beans. Phytochemistry 38:281–285
9. Suvachittanont W, Peutpaiboon A (1992) Lectin from Parkia speciosa seeds. Phytochemistry 31:4065–4070
10. Cavada BS, Madeira SVF, Calvete JJ, et al (2000) Purification, chemical, and immunochemical properties of a new lectin from Mimosoideae (Parkia discolor). Prep Biochem Biotechnol 30:271–280
11. Kaur N, Singh J, Kamboj SS, et al (2005) Two novel lectins from Parkia biglandulosa and Parkia roxburghii: isolation, physicochemical characterization, mitogenicity and anti-proliferative activity. Protein Pept Lett 12:585–595
12. Coriolano MC, de Melo CML, Silva F de O, et al (2014) Parkia pendula seed lectin: potential use to treat cutaneous wounds in healthy and immunocompromised mice. Appl Biochem Biotechnol 172:2682–2693
13. Cavada BS, Bari AU, Pinto-Junior VR, et al (2020) Purification and partial characterization of a new lectin from Parkia panurensis Benth. ex H.C. Hopkins seeds (Leguminosae family, Mimosoideae subfamily) and evaluation of its biological effects. International Journal of Biological Macromolecules 145:845–855
14. Silva HC, Bari AU, Rocha BAM, et al (2013) Purification and primary structure of a mannose/glucose-binding lectin from Parkia biglobosa Jacq. seeds with antinociceptive and anti-inflammatory properties. J Mol Recognit 26:470–478
15. Bari AU, Santiago MQ, Osterne VJS, et al (2016) Lectins from Parkia biglobosa and Parkia platycephala: A comparative study of structure and biological effects. Int J Biol Macromol 92:194–201
16. Mann K, Farias CMS, Del Sol FG, et al (2001) The amino-acid sequence of the glucose/mannose-specific lectin isolated from Parkia platycephala seeds reveals three tandemly arranged jacalin-related domains. European Journal of Biochemistry 268:4414–4422
17. Gallego del Sol F, Nagano C, Cavada BS, Calvete JJ (2005) The first crystal structure of a Mimosoideae lectin reveals a novel quaternary arrangement of a widespread domain. J Mol Biol 353:574–583

18. Neco AHB, Pinto-Junior VR, Araripe DA, et al (2018) Structural analysis, molecular docking and molecular dynamics of an edematogenic lectin from Centrolobium microchaete seeds. Int J Biol Macromol 117:124–133
19. Cavada BS, Pinto-Junior VR, Osterne VJS, et al (2021) In depth analysis on the carbohydrate-binding properties of a vasorelaxant lectin from Dioclea lasiophylla Mart Ex. Benth seeds. Journal of Biomolecular Structure and Dynamics 1–14
20. Cavada BS, Osterne VJS, Pinto-Junior VR, et al (2020) Molecular dynamics and binding energy analysis of Vatairea guianensis lectin: a new tool for cancer studies. J Mol Model 26:22
21. Verdonk ML, Cole JC, Hartshorn MJ, et al (2003) Improved protein-ligand docking using GOLD. Proteins 52:609–623
22. Liebeschuetz JW, Cole JC, Korb O (2012) Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Mol Des 26:737–748
23. Korb O, Stützle T, Exner TE (2009) Empirical scoring functions for advanced protein-ligand docking with PLANTS. J Chem Inf Model 49:84–96
24. de Ávila MB, Xavier MM, Pintro VO, de Azevedo WF Jr (2017) Supervised machine learning techniques to predict binding affinity. A study for cyclin-dependent kinase 2. Biochem Biophys Res Commun 494:305–310
25. Heck GS, Pintro VO, Pereira RR, et al (2017) Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem 24:2459–2470
26. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29:1859–1865
27. Lee J, Hitzenberger M, Rieger M, et al (2020) CHARMM-GUI supports the Amber force fields. J Chem Phys 153:035103
28. Case DA, Metin Aktulga H, Belfon K, et al (2021) Amber 2021. University of California, San Francisco
29. Huang J, Rauscher S, Nawrocki G, et al (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 14:71–73
30. Berendsen HJC, Postma JPM, van Gunsteren WF, et al (1984) Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics 81:3684–3690
31. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
32. Essmann U, Perera L, Berkowitz ML, et al (1995) A smooth particle mesh Ewald method. The Journal of Chemical Physics 103:8577–8593
33. Roe DR, Cheatham TE 3rd (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput 9:3084–3095
34. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–8, 27–8

35. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10:449–461
36. Miller BR 3rd, McGee TD Jr, Swails JM, et al (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. J Chem Theory Comput 8:3314–3321
37. Duan L, Liu X, Zhang JZH (2016) Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein-Ligand Binding Free Energy. J Am Chem Soc 138:5722–5728
38. Sun Z, Yan YN, Yang M, Zhang JZH (2017) Interaction entropy for protein-protein binding. The Journal of Chemical Physics 146:124124
39. Nascimento KS, Araripe DA, Pinto-Junior VR, et al (2018) Homology modeling, molecular docking, and dynamics of two α-methyl-D-mannoside-specific lectins from Arachis genus. J Mol Model 24:251
40. Cavada BS, Osterne VJS, Pinto-Junior VR, Nascimento KS (2019) ConBr, the Lectin from Canavalia brasiliensis Mart. Seeds: Forty Years of Research. Current Protein & Peptide Science 20:600–613
41. Cavada BS, Silva MTL, Osterne VJS, et al (2020) Exploring the carbohydrate-binding ability of Canavalia bonariensis lectin in inflammation models. J Mol Recognit 33:e2870

Figures

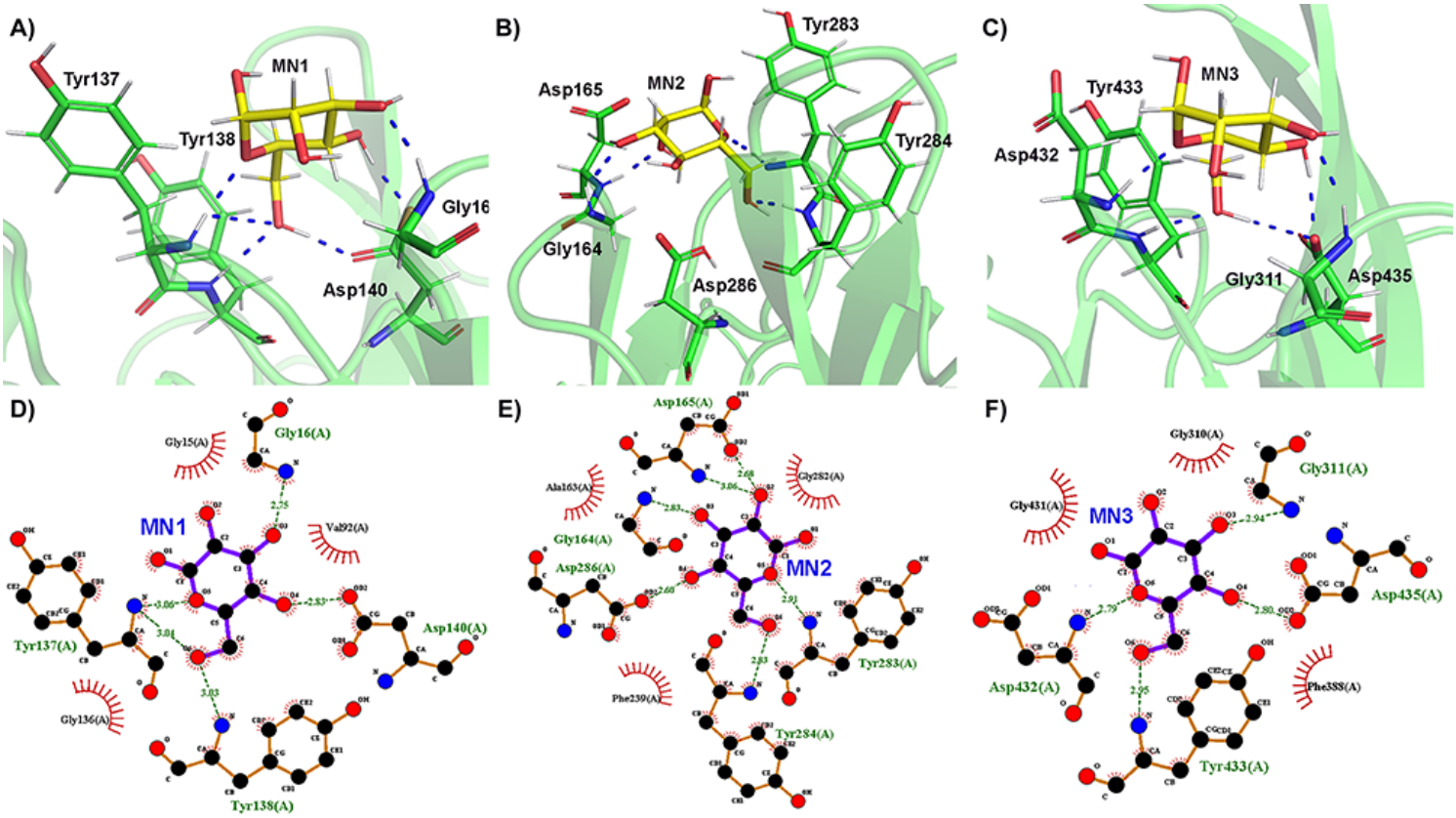

Figure 1

PPL domains complexed with D-mannose (MAN) from docking results. Polypeptide chains are represented in cartoon representation in green. MAN is in stick representation. A) Domain 1, B) Domain 2, C) Domain 3. LIGPLOT representations of hydrogen bonds and hydrophobic interactions around MAN with all CRD residues from D) Domain 1, E) Domain 2 and F) Domain 3.

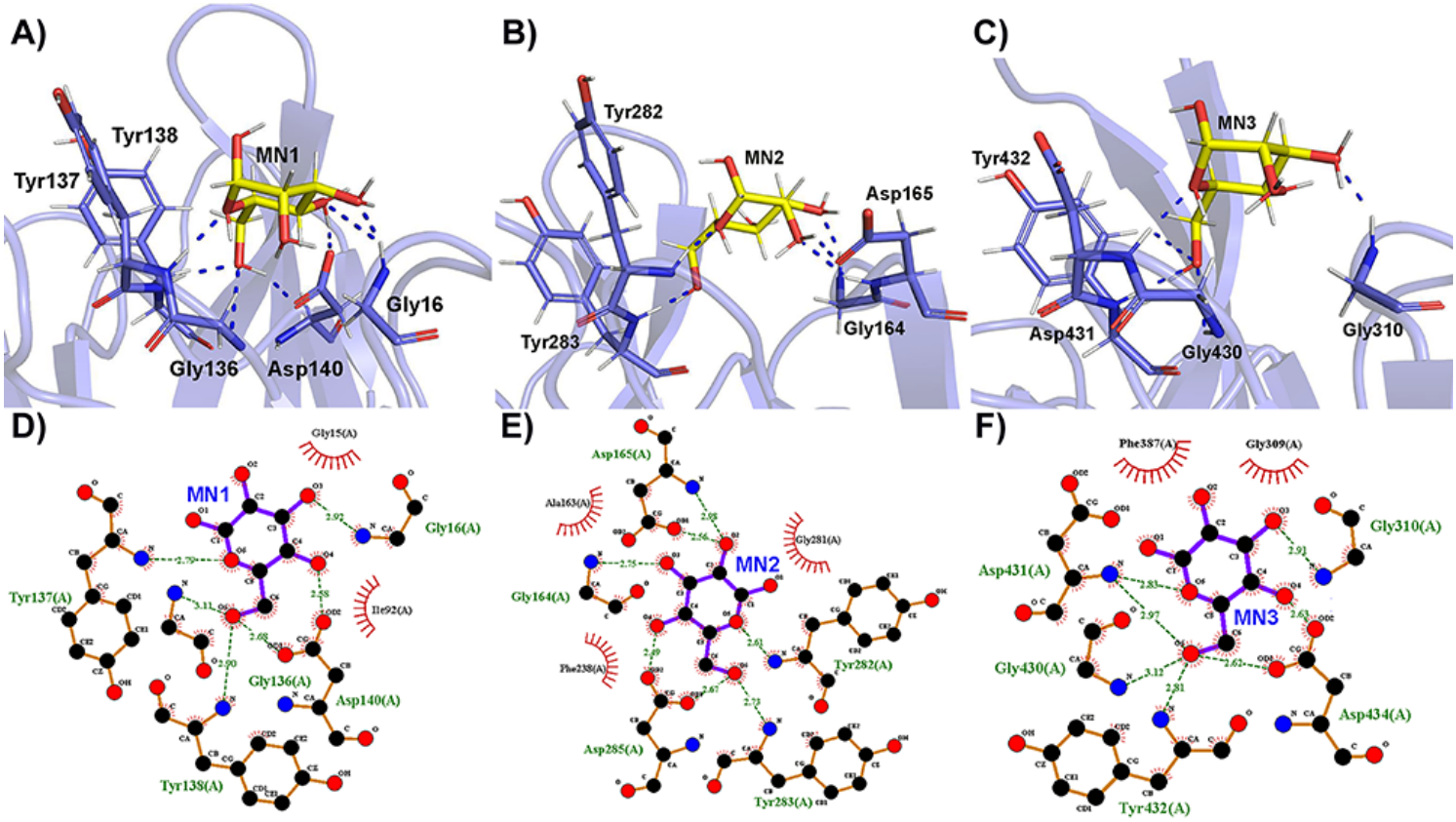

Figure 2

PBL domains complexed with D-mannose (MAN) from docking results. Polypeptide chains are represented in cartoon representation in purple. MAN is in stick representation: A) Domain 1, B) Domain 2, C) Domain 3. LIGPLOT representations of hydrogen bonds and hydrophobic interactions around MAN with all CRD residues from D) Domain 1, E) Domain 2 and F) Domain 3.

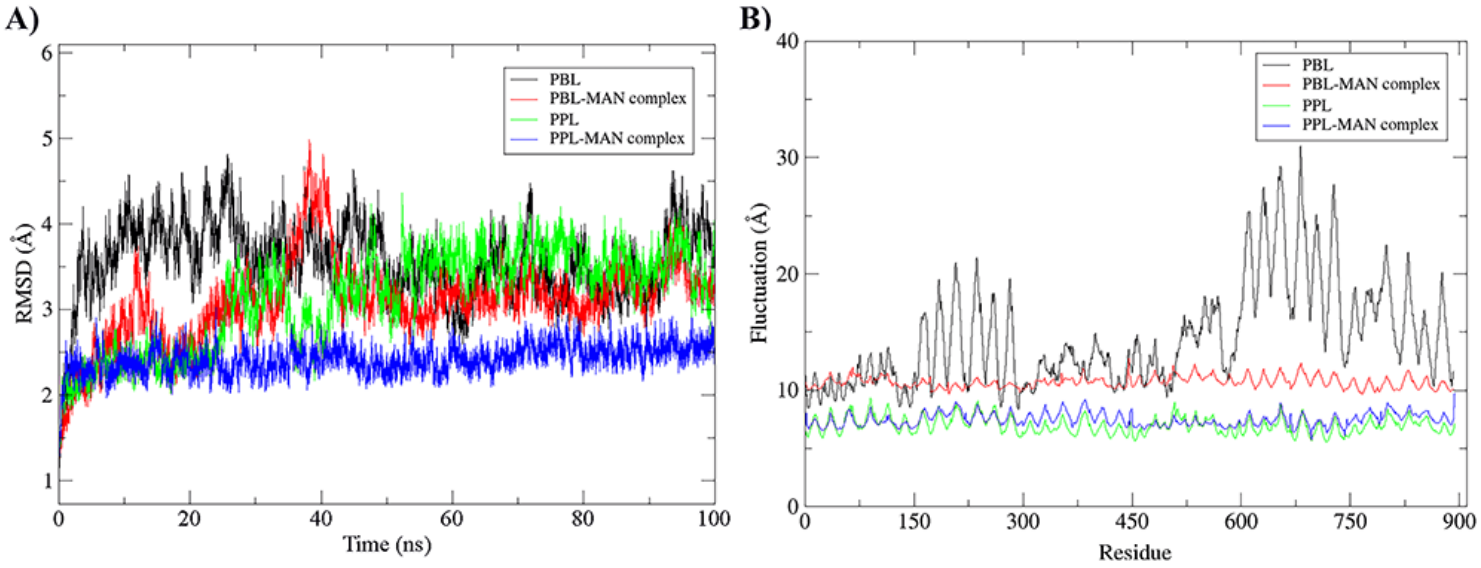

Figure 3

Molecular dynamics trajectories of PPL and PBL in their native and complexed with D-mannose (MAN) states. A) Mean square deviation graph (RMSD) and B) Mean quadratic fluctuation graph (RMSF).

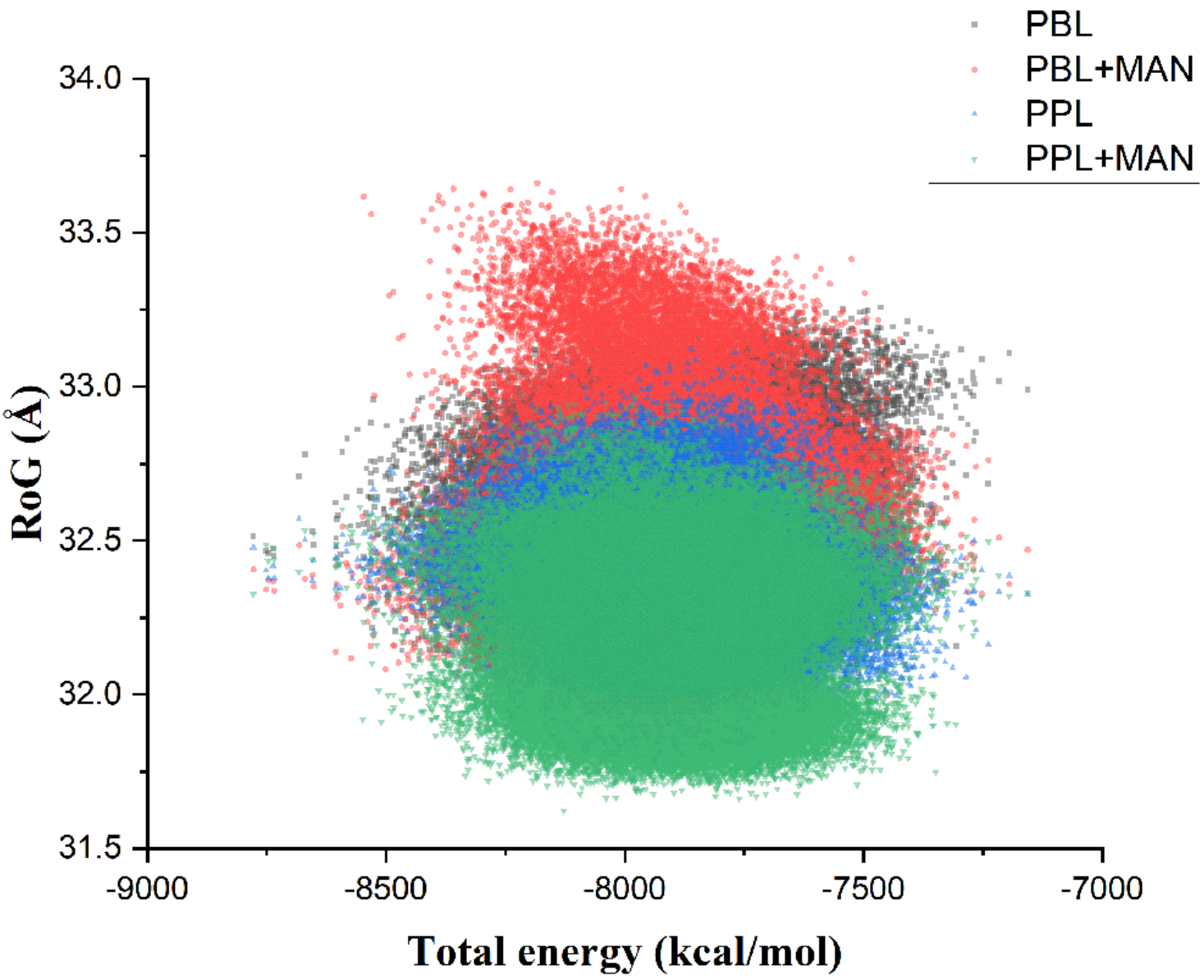

Figure 4

Molecular compaction results from molecular dynamics trajectories correlating Radius of Gyration (RoG) and total energy of PPL and PBL structures in their native and D-mannose (MAN) complexed states.

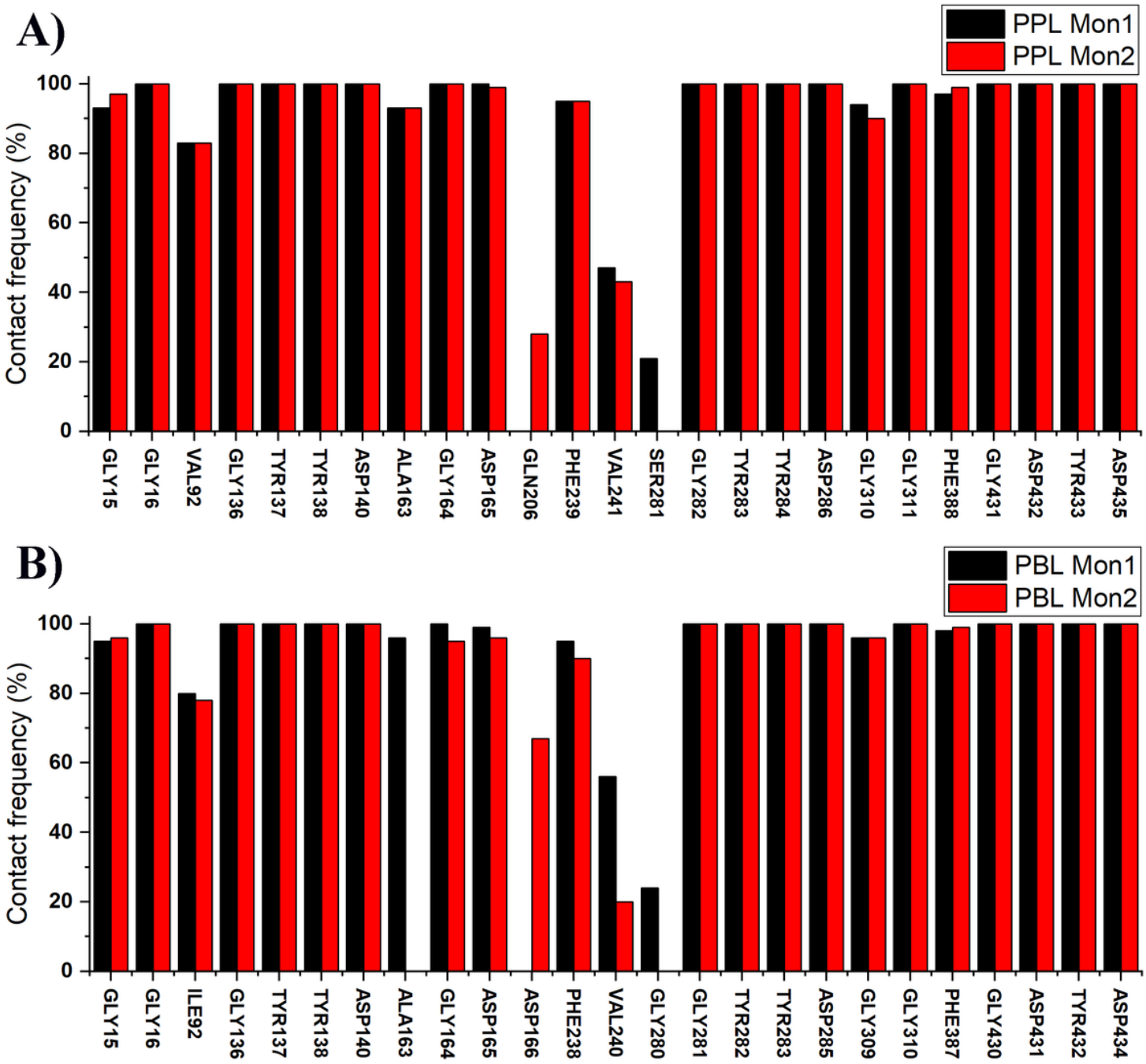

Figure 5

Plot showing the contact frequency of CRD residues from PPL and PBL complexed with D-mannose during molecular dynamics trajectories. A) Residues of the two PPL monomers, B) residues of the two PBL monomers.

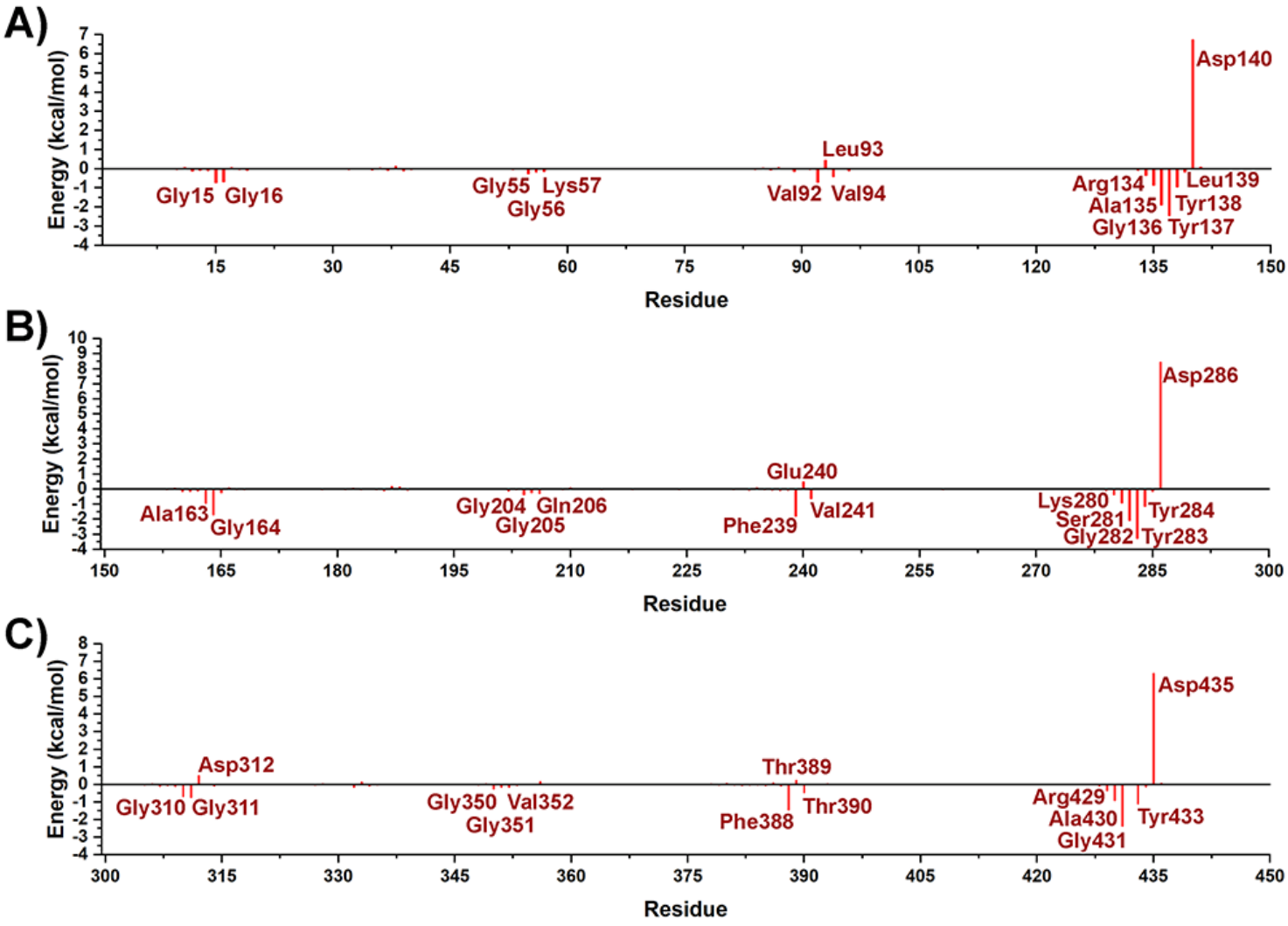

Figure 6

Decomposition of binding free energy per residue from molecular dynamics simulations of PPL in complex with D-mannose: A) domain 1, B) domain 2 and C) domain 3.

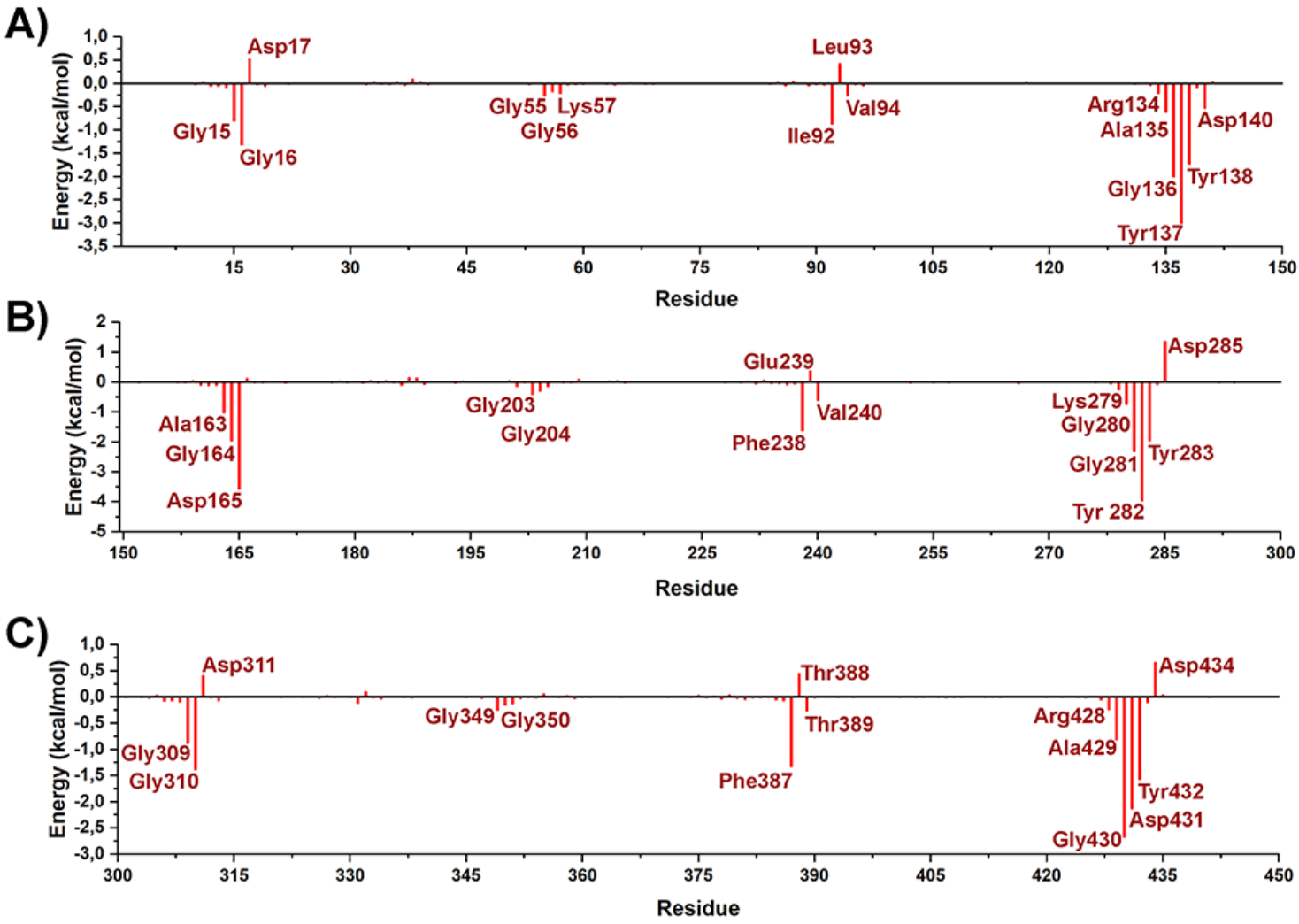

Figure 7

Decomposition of binding free energy per residue from molecular dynamics simulations of PBL in complex with D-mannose: A) domain 1, B) domain 2 and C) domain 3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf