

Mutation of the Flo1 flocculation protein for enhancing the oligomannose binding

Chenhong Zhang¹ and Shijun Zhong¹

¹Dalian University of Technology

May 5, 2020

Abstract

Yeast cell flocculation is a common cellular adhesion process which increases the efficiency of yeast cell harvest and recovery in the beer industry and biofuel production. The flocculation involves the binding of exopolysaccharides to the flocculation protein of neighboring cell. It is of great interest to examine the roles of key amino acid residues in modulating polysaccharide binding which regulates the flocculation. In this work, proper mutation sites were identified via alanine scanning mutagenesis incorporated with molecular dynamics simulations of protein-mannobiose complexes based on the crystal structure of flocculation protein Flo1p. Two mutants Q117N and Q117R have been selected out of nine examined mutants, with stronger binding free energies for the mannobiose binding, comparing to the wild type protein. The two mutants also show enhanced binding in the protein-mannotriose complexes which consist of three mannose residues, fulfilling the capacity of the binding site, and can be used to evaluate the binding strength of oligomannose. Hydrogen bonding, conformational stability and shape change were analyzed to help understand the protein-mannotriose interactions. This study also shows a way to alternate oligosaccharide binding to proteins functioning in other biological processes.

KEYWORDS

alanine scanning mutagenesis, binding free energy, flocculation protein mutant, molecular dynamics simulations, oligosaccharide binding

INTRODUCTION

The binding of oligosaccharides on protein is essential for many important biological processes such as recognition, signaling and adhesion (Kaltner, Abad-Rodriguez, Corfield, Kopitz, & Gabius, 2019; A. Mishra et al., 2019; Varki, 2017). Cellular adhesion between yeast cells is a natural phenomenon called flocculation which can improve the efficiency of *Saccharomyces cerevisiae* and allows them to survive under adverse conditions (Cheng, Zhang, Xue, Bai, & Zhao, 2017), for example, in the beer industry (Sariki, Kumawat, Singh, & Tomar, 2019) and the production of bioethanol (He, Zhao, & Bai, 2012; Li, Zhao, Chang, Zhang, & Bai, 2012). The flocculation mechanism has been illustrated in several similar models which present two kinds of interactions, i.e. the adhesion between exopolysaccharides (EPS) and the binding of EPS mannose residues on the N-terminal domain of flocculation protein (Bruckner & Mosch, 2012; K. V. Goossens et al., 2015; Soares, 2011; Verstrepen & Klis, 2006).

A flocculation protein consists of three domains, i.e. the N-terminal domain, the middle domain and the C-terminal domain. The N-terminal domain is a carbohydrate binding domain (Groes, Teilmann, Olesen, Poulsen, & Henriksen, 2002) required for flocculation (Veelders et al., 2010). It can be extended beyond the cell surface (Kobayashi, Hayashi, Kuroki, & Sone, 1998) to recognize and bind the N-oligosaccharides of neighboring cell (Helenius & Aebi, 2004). These N-oligosaccharides have an α -1,6-mannose chain of up to 50 mannose residues, attached with short α -1,2-mannose branches which are terminated in an α -1,3-linked mannose residue (Ballou, 1990; Dean, 1999; Lesage & Bussey, 2006). The middle domain is a highly

repetitive sequence rich in Ser and Thr, to which some short O-oligosaccharides are linked, like tree branches to adhere the N-oligosaccharides in the outer layer of cell wall and thus to support the structure of cell wall (Levin, 2011). The Ser/Thr-rich sequence is a flexible chain of 800-1260 amino acids, long enough to deliver the N-terminal domain out of the cell wall for interacting with the neighboring cell (K. Goossens & Willaert, 2010). The C-terminal domain contains a hydrophobic domain and binds, in Golgi body, to glycosylphosphatidylinositol (GPI) which anchors the C-terminal on the cell membrane after transporting through the secretory pathway (Kondo & Ueda, 2004). The GPI anchor can be cleaved off on the plasma membrane, but leaving the GPI remnant on the C-terminal for linking to the glucans in the inner layer of cell wall (Verstrepen & Klis, 2006).

The family of flocculins contains up to 30 members, including the proteins Flo1, Flo5 and Flo11, which are encoded by the correlated genes FLO1, FLO5 and FLO11, respectively (Teunissen & Steensma, 1995). Among them, Flo1 can form stronger flocculation (Verstrepen, Jansen, Lewitter, & Fink, 2005) due to its longer middle sequence of tandem repeats which deliver the N-terminal carbohydrate binding domain out of the cell surface to capture the N-oligosaccharides of neighboring cell. The flocculation can be enhanced in the presence of Ca^{2+} (Soares, 2011; Veelders et al., 2010), presenting in the carbohydrate binding site (K. V. Goossens et al., 2015), by increasing the electrostatic interaction with the negatively charged hydroxyl oxygen atoms of N-oligosaccharides. In contrast, the flocculation can be inhibited by a mannose molecule which binds in the carbohydrate binding site and blocks the binding of N-oligosaccharides (K. V. Goossens et al., 2015; Veelders et al., 2010).

Mutation is a conventional way to alternate the interaction between a ligand and its receptor (Sahu, Moharana, Sahu, & Pattanayak, 2019), i.e. the binding of N-oligosaccharides on the N-terminal domain. Mutation of binding site residues will directly affect the binding affinity, for examples, via changing a nonpolar sidechain to a polar one to strengthen the electrostatic interaction by forming additional hydrogen bond, or by replacing a longer sidechain to a shorter one, thus creating more space for a larger ligand. Proper mutation positions can be determined using a mutagenesis technique called alanine scanning which replaces binding site residues with alanine one by one to scan the variation of biochemical and biophysical properties, especially the binding free energy (Coskuner-Weber & Uversky, 2019; Massova & Kollman, 1999). Here, the ligand binding free energy is of special interest. Molecular dynamics (MD) simulation provides a convenient approach to investigate the binding free energies of protein-ligand systems (Huggins et al., 2019; S. K. Mishra, Sund, Aqvist, & Koca, 2012; Wang et al., 2019) to compare wild-type protein with mutants (Bhayye, Roy, & Saha, 2018; Kulczycka-Mierzejewska, Sadlej, & Trylska, 2018; Martinez-Archundia, Hernandez Mojica, Correa-Basurto, Montano, & Camacho-Molina, 2019; Saladino & Gervasio, 2016) based on experimentally determined protein crystal or NMR structures.

Crystal structures for the N-terminal domains of Flo1 flocculins from both *S. cerevisiae* (N-Flo1p) and *S. pastorianus* (N-Lg-Flo1p), have been solved (PDB IDs: 4LHL (K. V. Goossens et al., 2015), 4LHN (K. V. Goossens et al., 2015), 4LHK (K. V. Goossens et al., 2015) and 4GQ7 (Sim, Groes, Olesen, & Henriksen, 2013)). 4LHL is the apo form of N-Flo1p in which three amino acid residues in the L3 loop are missing (W196, G197 and G198). 4LHN is an intact N-Flo1p complexed with a mannose molecule and a calcium cation Ca^{2+} . In addition, 4LHK and 4GQ7 are homologues of N-Flo1p and can be inhibited by mannose and glucose (K. V. Goossens et al., 2015; Sim et al., 2013). 4LHK is an N-Lg-Flo1p associated with α -1,2-mannobiose. 4GQ7 is the apo-N-Lg-Flo1p, which is similar for 74% in sequence and can be well superimposed in three dimensional structure with the apo-N-Flo1p (4LHL) (K. V. Goossens et al., 2015).

In this work, the alanine scanning mutagenesis was introduced to determine suitable mutation positions which affect the oligomannose binding on the Flo1p protein based on the 4LHN crystal structure. Using Amber 14 (Case et al., 2014), MD simulations and binding free energy calculations were performed on the wild type (WT) protein and several selected mutants to investigate the possibility of enhancing the binding of oligomannoses to N-Flo1p. Detailed analyses were derived from the MD trajectories to provide an atomistic understanding of the interactions between the oligosaccharides and the protein.

RESULTS AND DISCUSSION

During the binding of oligomannose in flocculation, two or three mannose residues can simultaneously be bound in the binding site (Bruckner & Mosch, 2012). The binding site in the crystal structure 4LHN is larger enough to accommodate up to three sugar residues (K. V. Goossens et al., 2015). The binding diagram is illustrated in Figure 1. In the crystal structure of 4LHN, one mannose molecule is bound in the binding site of N-Flo1p (Figure 1A) (K. V. Goossens et al., 2015). We assume that the crystal structure provides the lowest energy pose of the single mannose. However, it is expected that certain mutants might have stronger binding free energy for mannobiose (FIGURE 1B) and mannotriose (FFigure 1C), comparing with the wild type protein. The cases of mannobiose and mannotriose bindings (Figure 1B and 1C) can be used to estimate the binding strength of the whole oligomannose (Figure 1D). According to the connections of N-linked oligosaccharides (Ballou, 1990; Dean, 1999), mannobiose (manb) has the (α 1-3) linkage, and mannotriose (mant) the (α 1-3) - (α 1-2) linkages, with all mannose residues in the form of α -D-mannose enantiomer. There is no need to compute larger oligomannose since a sugar residue of mannotetraose already spreads out of the active site in our test dockings.

2.1 Alanine scanning mutagenesis

Amino acid residues with any atom within 5 Å around the ligand mannobiose in the WT-manb complex were tested one by one using the alanine scanning algorithm built in Amber 14. The free energy changes of mannobiose binding for mutants comparing with the WT protein ($G = G_{\text{wildtype}} - G_{\text{mutant}}$) are shown in Figure 2. Undoubtedly, a big loss in binding free energy indicates an important residue which contributes a lot to the binding. A very small change or zero change in binding free energy means that the residue does not contribute much to the binding. Therefore, mutation should be applied to residues which cause middle changes in binding free energy. A middle amount indicates both the participation in binding and the opportunity for improving.

Our calculations clearly show some residues have a greater impact on the binding than others, that is, the mutants D160A and D161A which reduce the ligand binding free energy by about 25 kcal/mol, indicating that the side chains of the two aspartic acid residues play a very large part in the binding free energy. This means that this conserved aspartic acid motif is important and should not be mutated, consistent with the observation throughout the entire epithelial adhesin family (Maestre-Reyna et al., 2012). The residues N224, V226 and W228 have strong interactions with Ca^{2+} which is located in the binding site (K. V. Goossens et al., 2015) and involved in the binding of carbohydrates (Kerr, McDiarmid, Fraser, & Schulz, 2018). Therefore, these residues will not be considered for mutation in order to avoid disrupting the oligomannose binding. The residues S227, P195 and S101 contribute less than 1 kcal/mol to the overall binding free energy, indicating that these residues might not be important for the binding of mannobiose.

Each of the residues K194, W196, Q98 and Q117 contributes a middle amount to the binding free energy, i.e. 4~6 kcal/mol, and might be improved via mutation. Carefully viewing the shape of the active site (Figure 3), it can be seen that tryptophan residue W196 should not be mutated since it forms an edge of the binding pocket to prevent the ligand from unbinding movements. In addition, lysine residue K194 and glutamine residue Q98 are closer to the ligand, with possible hydrogen bonds to assist binding. They can be mutated to similar residues to form mutants like K194R, Q98D, Q98E and Q98N. Only Q117 is located a little far from the center of binding site and offer a wide range of options for mutants including Q117D, Q117E, Q117K, Q117N and Q117R which all can form hydrogen bond.

Flo1p-mannobiose complexes

MD simulations of 50 ns were performed on the mannobiose-binding complexes of the wild type N-Flo1p protein, as well as nine mutants mentioned above, followed by the binding free energy calculations using the MM/GBSA (molecular mechanics/generalized Born model and solvent accessible surface area) method (Gohlke, Kiel, & Case, 2003; Onufriev, Bashford, & Case, 2004). The total binding free energy G_{bind} is divided into its components in the formula, $G_{\text{bind}} = E_{\text{ele}} + E_{\text{vdw}} + G_{\text{GB}} + G_{\text{np}}$, where E_{ele} is the electrostatic interaction, E_{vdw} the van der Waals interaction, G_{GB} the polar solvation energy calculated via the generalized Born model, and G_{np} the non-polar solvation energy calculated using the solvent accessible surface area. The

entropy change is neglected since the ligand is same for all complexes.

Comparison of binding free energy values

The values of all the above energy terms are shown in Table 1 for the ten complexes. Seven complexes, including Q98E, Q98N, Q117E, Q117D, Q117K, Q117N and Q117R, receive more favorable binding free energy than the wild type protein, while two mutants Q98D and K194R reduce the binding free energy. Especially, two mutants Q117N and Q117R have much stronger bindings than the wild type protein, since the total binding free energy is -9.81 and -13.23 kcal/mol for Q117N and Q117R, respectively, more than the value for the WT protein. Other five mutants Q98E, Q98N, Q117E, Q117D and Q117K have slightly better binding free energy values from -1.54 to -5.37 kcal/mol which are not very significantly larger and might be reduced if other computational corrections will be added such as the entropy contribution. Considering the two poor mutants, Q98D has a significant solvation energy penalty of $G_{GB} + G_{np}$ =100.31 kcal/mol larger than 94.21 kcal/mol for the WT, while K194R has a smaller electrostatic energy -87.05 kcal/mol comparing to -95.07 kcal/mol for the WT. This result indicates that the sidechain of K194 plays an important role in the oligomannose binding which is consistent with the experimental report (K. V. Goossens et al., 2015). From the above analysis, it can be seen that the binding can be greatly improved in Q117N and Q117R which thus are proper for further validation.

Conformational change

For better understanding the binding difference between the WT protein and mutant Q117R, the root mean square deviation (RMSD) values were calculated for ligand, Ca^{2+} and residues within 5 Å around the ligand. FIGURE 4 shows that the RMSD for WT-manb is stable around 1.25 Å, while the RMSD for Q117R-manb is frequently changing in the period of time <17 ns and stabilizes after 20 ns. It indicates that the ligand was reorienting its poses, due to the incorporation of the significantly larger arginine.

The structure on the upper left of Figure 4 is a snapshot at 17 ns which shows the mannobiose binding pose with its second sugar structure flexible and exposed to solvent water molecules. The upper right shows a well bound pose at 20 ns which forms two hydrogen bonds with S101 and R117 respectively, while the L3 loop moves closer to its first sugar ring to form hydrogen bonding interactions at residues D161 and K194. In order to check the changes of mannobiose binding, the binding free energy of Q117R-manb was averaged for each 10 ns of the MD trajectory, resulting at -24.98, -30.94, -33.63, -32.58 and -35.86 kcal/mol, showing that the mannobiose binding became stronger after 20 ns.

Flo1p-manotriose complexes

Based on the above analysis of the binding free energy values, it is clear that two mutants Q117N and Q117R can enhance the binding of mannobiose which contains two sugar residues. Consequently, it is expected that the two mutants can also enhance the binding of manotriose which contains three sugar residues, which would more closely resemble to the situation of oligosaccharide binding. Hence, two Flo1p-manotriose complexes Q117N-mant and Q117R-mant were subjected to estimation via MD simulations of 50 ns followed by binding free energy calculations.

Binding free energy and binding stability

The binding free energy results are shown in Table 2. It can be seen that the binding free energy is -40.37 and -42.03 kcal/mol for Q117N and Q117R, respectively, which are at least 15 kcal/mol stronger than that for WT-mant. In principle, the mutants Q117N and Q117R can strengthen the oligomannose binding and thus improve the flocculation. Especially, Q117R presents the strongest binding free energy. Checking the variation of binding strength with the increment of sugar residue number, via comparing the manotriose binding (Table 2) with the mannobiose binding (Table 1), it can be encouraged that one more sugar residue increases the binding free energy about 2.32 kcal/mol for the WT, 7.87 kcal/mol for Q117N and 6.97 kcal/mol for Q117R. This suggests that our proposed mutations will further increase the binding free energy as function of the length of an oligomannos chains.

For checking the binding stability of mannotriose in the binding pocket. Figure 5 shows the trajectory fluctuation of the RMSD values of mannotriose plus Ca^{2+} and residues within 5 Å around the mannotriose. Here, we have to point out that the absolute value of RMSD is not so important since the initial reference conformations are different. Of more importance is that a smaller fluctuation suggests a more stable binding. It can be seen that the RMSD is rather stable in the cases of the two mutants Q117N and Q117R, but changes frequently in the WT protein. This indicates that the two mutants can tightly hold the ligand, while the WT protein does not keep the ligand very steadily although the ligand still lies in the binding site. This result has proven that proper mutation can stabilize the oligomannose binding and thus the flocculation.

Binding modes and hydrogen bonds

To show the changes of the binding mode and the pocket shape after the mutations have been introduced, representative conformations were extracted from the last 10 ns trajectories of MD simulations which tend to be proceeding steadily. The conformation with the lowest energy is used as the representative here in order to express the best binding in the concerned complex. For the sake of comparison, the representative conformations from Q117N-mant and Q117R are respectively superimposed with the representative of WT-mant. It can be seen that the binding poses of mannotriose molecule in Q117N-mant and WT-mant are significantly different even if the binding pocket residues are overlapped closer and the two pocket shapes are similar to each other (Figure 6A). Similarly, the superposition of Q117R-mant and WT-mant shows different mannotriose binding poses and poor overlaps of the pocket residues which might be caused by the larger sidechain of arginine (Figure 6B).

The binding mode of mannotriose in Q117N is shown in Figure 6C, in which mannotriose lies along to the valley of binding pocket, stabilized by four hydrogen bonds which can be expressed in the donor-acceptor notation residue:atom...residue:atom (Kulczycka-Mierzejewska et al., 2018) as 1:O4...Asp161:OD1, 2:O4...Asn117:OD1, Ser227:OG...2:O4 and 3:O4...Gly103:O, where the numbers 1, 2 and 3 indicate the sugar residues in mannotriose, while OD1 and OG are the oxygen atoms of the amino acid residues. In the case of Q117R-mant (Figure 6D), the L3 loop moves closer to hold the ligand tightly together with R117 so that the binding pose looks like a seedling growing from the bottom where the first sugar residue is firmly grasped by three stable hydrogen bonds 1:O4...Asp161:OD1, 1:O6...Asp161:OD1 and 1:O2...Gln98:OE1.

Some hydrogen bonds are stable and some hydrogen bonds are considerably flexible. Here, a hydrogen bond is judged to be linked if the donor-acceptor distance is not more than 3.0 Å and the donor-H-acceptor angle is not smaller than 135°. Either a longer distance or a smaller angle means the breakdown of the hydrogen bond. Thus, a hydrogen bond has two states, i.e. the linked and broken states during a period of time. The percentage of linked states over all snapshots in the concerned period of time is called the occupancy of hydrogen bond. A larger occupancy means a stable hydrogen bond.

Table 3 gives the occupancies of intermolecular hydrogen bonds in the three complexes based on the trajectories of the last 10 ns. Meanwhile, the averaged donor-acceptor distances and donor-H-acceptor angles are provided for showing the linked state. In the WT-mant complex, there are only two stable hydrogen bonds with higher occupancy, i.e. 1:O4...Asp161:OD1 (92.60%) and 1:O2...Gln 98:OE1 (66.30%) which are related to the first sugar residue, while other four hydrogen bonds have very low occupancy of less than 18.00%. This suggests that the mannotriose molecule only binds its first sugar residue well, but leaves other two sugar residues flexible in the WT-mant complex. In the case of Q117N-mant, there are four stable hydrogen bonds with high occupancy from 84.30% to 99.10%, one for the first sugar residue, two for the second sugar residue and one for the third sugar residue. This indicates that mannotriose molecule can bind all the three sugar residues well in the Q117N pocket. Similarly, Q117R can hold the mannotriose molecule firmly via three stable hydrogen bonds for the first sugar residue with high occupancies from 86.70% to 97.10%. Meanwhile, one more hydrogen bond between Arg117 and the third sugar residue can help stabilize the mannotriose binding in the Q117R pocket.

CONCLUSIONS

It is believed that mutation can provide opportunity to improve the oligosaccharide binding on flocculin

protein. Therefore, residues in the binding site of N-Flo1p protein were tested one by one via alanine scanning mutagenesis technique implemented in the MD simulation software Amber 14, based on the 50 ns MD simulations of WT-mannobiose complex. Here, mannobiose was used since it is a key step to grow single mannose to oligomannose to investigate the oligosaccharide binding. The changes in binding free energy during the alanine scanning were used to determine proper mutation sites. A big loss in binding free energy indicates that this residue contributes a lot and should not be mutated. A very small change or zero change in binding free energy means that the residue does not participate much in the binding and should not be considered for mutation too. A residue causing a middle change indicates its participation in binding and will be suitable for mutation toward improving the binding free energy. Several sites including Q98, Q117 and K194 were found to be suitable for mutation.

Two mutants Q117N and Q117R out of the nine investigated mutants have been identified to enhance the mannobiose binding, via the MM/GBSA binding free energy calculations applied to the last 10 ns trajectories of 50 ns MD simulations. The binding free energy values for Q117N and Q117R are about 10 kcal/mol more than the value for the WT protein. Further calculations were performed on the complexes with mannotriose which contains three sugar molecules and can be used to evaluate the binding strength of oligomannose since the binding pocket cannot hold more than three mannose residues. Q117N and Q117R both present stronger binding than the wild type protein. The above energy analysis for the bindings of both mannobiose and mannotriose, shows that mutation can improve the N-Flo1p-oligomannose binding. A stronger binding of oligosaccharide on the N-terminal domain of Flo1 protein means a more favorable flocculation between yeast cells. Thus, the mutants Q117N and Q117R can be considered as candidates for further experimental validation.

1. COMPUTATIONAL DETAILS

2. Structure preparation

The N-Flo1p structure is taken from 4LHN (K. V. Goossens et al., 2015). The protonation was automatically done since there is no histidine residue. Hydrogen atoms were automatically added and easily minimized using Amber 14 (Case et al., 2014). Mutants were generated by manually changing the residue names, followed by quick minimization for the replaced sidechain. The oligomannoses were constructed using the dataset GLYCAM-06j (Kirschner et al., 2008) which is implemented in Amber14. The poses of mannobiose and mannotriose were initially generated using AutoDock Vina (Trott & Olson, 2010) with the docking center located at the centroid of the crystal ligand coordinates and further minimized using Amber 14. Minimizations were performed in two steps. The first step was done in vacuum by a cycle of 2500 steps of the steepest descent method followed by 2500 steps of the conjugate gradient method. The second step was done in the water solvents of TIP3P model (Neria, Fischer, & Karplus, 1996) neutralized with sodium ions by several cycles of the above combined methods, applied first to water molecules via restricting the protein-ligand complex plus Ca^{2+} with a restraint of 1000 kcal/(mol[?]Å²), and then to the whole system. The Amber ff14SB force field (Maier et al., 2015) was used in this work.

Molecular dynamics simulation

For each protein-ligand complex, the periodic boundary condition was applied based on the water box with distances of 12 Å from the outer atoms of protein-ligand complex to the box edges. Long-range electrostatic interactions were calculated using the Particle mesh Ewald algorithm (Darden, York, & Pedersen, 1993) and the positions of hydrogen atoms were constrained using the SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1977). The whole system was heated up from 0 K to 300 K at a step of 50 K per 10 ps, with NVT ensemble, using the Langevin thermostat to control the temperature (Izaguirre, Catarella, Wozniak, & Skeel, 2001). During the heating process, a weak restraint of 10 kcal/(mol[?]Å²) is applied to the heavy atoms of the complex to allow the hydrogen atoms to move freely. The equilibration was done with NPT ensemble, via three steps, each lasting 100 ps with a restraint on the heavy atoms gradually reduced from 10 to 5 and 0 kcal/(mol[?]Å²). The time step was set to be 1 fs for heating and equilibration. Finally, the production of 50 ns was performed, with the aids of GPU acceleration, on the well-balanced system at a time step of 2 fs. Alanine scanning mutagenesis (Coskun-Weber & Uversky, 2019) was performed based

on the trajectories of WT-manb complex. The MM/GBSA binding free energies were calculated from the final 10 ns trajectories, using the SANDER program implemented in Amber14.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 21536006) and the Supercomputing Center of Dalian University of Technology. Zhang thanks Lu Sun and Dan Zhao for their assistance in computations. Zhong thanks Christoffer Lind, Xinqing Zhao and Fengwu Bai for helpful discussions.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

ORCID

Shijun Zhong <http://orcid.org/0000-0002-4213-2808>

REFERENCES

- Ballou, C. E. (1990). Isolation, characterization, and properties of *Saccharomyces cerevisiae* mnn mutants with nonconditional protein glycosylation defects. *Methods Enzymol*, 185, 440-470. doi:10.1016/0076-6879(90)85038-p
- Bhayye, S. S., Roy, K., & Saha, A. (2018). Molecular dynamics simulation study reveals polar nature of pathogenic mutations responsible for stabilizing active conformation of kinase domain in leucine-rich repeat kinase II. *Structural Chemistry*, 29 (3), 657-666. doi:10.1007/s11224-017-1059-z
- Bruckner, S., & Mosch, H. U. (2012). Choosing the right lifestyle: adhesion and development in *Saccharomyces cerevisiae*. *FEMS Microbiology Reviews*, 36 (1), 25-58. doi:10.1111/j.1574-6976.2011.00275.x
- Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., . . . Kollman, P. A. (2014). AMBER 14. University of California, San Francisco.
- Cheng, C., Zhang, M., Xue, C., Bai, F., & Zhao, X. (2017). Development of stress tolerant *Saccharomyces cerevisiae* strains by metabolic engineering: New aspects from cell flocculation and zinc supplementation. *Journal of Bioscience and Bioengineering*, 123 (2), 141-146. doi:10.1016/j.jbiosc.2016.07.021
- Coskuner-Weber, O., & Uversky, V. N. (2019). Alanine Scanning Effects on the Biochemical and Biophysical Properties of Intrinsically Disordered Proteins: A Case Study of the Histidine to Alanine Mutations in Amyloid-beta42. *Journal of Chemical Information and Modeling*, 59 (2), 871-884. doi:10.1021/acs.jcim.8b00926
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N[\log(N)]$ method for Ewald sums in large systems. *Journal of Chemical Physics*, 98 (12), 10089-10092. doi:10.1063/1.464397
- Dean, N. (1999). Asparagine-linked glycosylation in the yeast Golgi. *Biochimica et Biophysica Acta*, 1426 (2), 309-322. doi:10.1016/s0304-4165(98)00132-9
- Gohlke, H., Kiel, C., & Case, D. A. (2003). Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *Journal of Molecular Biology*, 330 (4), 891-913. doi:10.1016/s0022-2836(03)00610-7
- Goossens, K., & Willaert, R. (2010). Flocculation protein structure and cell-cell adhesion mechanism in *Saccharomyces cerevisiae*. *Biotechnology letters*, 32 (11), 1571-1585. doi:10.1007/s10529-010-0352-3
- Goossens, K. V., Ielasi, F. S., Nookaew, I., Stals, I., Alonso-Sarduy, L., Daenen, L., . . . Willaert, R. G. (2015). Molecular mechanism of flocculation self-recognition in yeast and its role in mating and survival. *mBio*, 6 (2), e00427-00415. doi:10.1128/mBio.00427-15

- Groes, M., Teilum, K., Olesen, K., Poulsen, F. M., & Henriksen, A. (2002). Purification, crystallization and preliminary X-ray diffraction analysis of the carbohydrate-binding domain of flocculin, a cell-adhesion molecule from *Saccharomyces carlsbergensis*. *Acta Crystallographica Section D-Structural Biology*, *58*, 2135-2137. doi:10.1107/s0907444902015494
- He, L. Y., Zhao, X.-Q., & Bai, F.-W. (2012). Engineering industrial *Saccharomyces cerevisiae* strain with the FLO1-derivative gene isolated from the flocculating yeast SPSC01 for constitutive flocculation and fuel ethanol production. *Applied Energy*, *100* (none), 33-40.
- Helenius, A., & Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annual Review of Biochemistry*, *73*, 1019-1049. doi:10.1146/annurev.biochem.73.011303.073752
- Huggins, D. J., Biggin, P. C., Damgen, M. A., Essex, J. W., Harris, S. A., Henchman, R. H., . . . van der Kamp, M. W. (2019). Biomolecular simulations: From dynamics and mechanisms to computational assays of biological activity. *Wiley Interdisciplinary Reviews-Computational Molecular Science*, *9* (3), e1393. doi:10.1002/wcms.1393
- Izaguirre, J. A., Catarello, D. P., Wozniak, J. M., & Skeel, R. D. (2001). Langevin stabilization of molecular dynamics. *Journal of Chemical Physics*, *114* (5), 2090-2098. doi:10.1063/1.1332996
- Kaltner, H., Abad-Rodriguez, J., Corfield, A. P., Kopitz, J., & Gabius, H. J. (2019). The sugar code: letters and vocabulary, writers, editors and readers and biosignificance of functional glycan-lectin pairing. *Biochemical Journal*, *476* (18), 2623-2655. doi:10.1042/BCJ20170853
- Kerr, E. D., McDiarmid, D. C., Fraser, J. A., & Schulz, B. L. (2018). Cell Wall and Whole Cell Proteomes Define Flocculation and Fermentation Behavior of Yeast. *Fermentation*, *4* (3), 80-90. doi:10.3390/fermentation4030080
- Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., & Woods, R. J. (2008). GLYCAM06: A generalizable Biomolecular force field. Carbohydrates. *Journal of Computational Chemistry*, *29* (4), 622-655. doi:10.1002/jcc.20820
- Kobayashi, O., Hayashi, N., Kuroki, R., & Sone, H. (1998). Region of Flo1 proteins responsible for sugar recognition. *Journal of Bacteriology*, *180* (24), 6503-6510. doi:0021-9193/98/\$04.0010
- Kondo, A., & Ueda, M. (2004). Yeast cell-surface display - applications of molecular display. *Applied Microbiology and Biotechnology*, *64* (1), 28-40. doi:10.1007/s00253-003-1492-3
- Kulczycka-Mierzejewska, K., Sadlej, J., & Trylska, J. (2018). Molecular dynamics simulations suggest why the A2058G mutation in 23S RNA results in bacterial resistance against clindamycin. *Journal of Molecular Modeling*, *24* (8), 191. doi:10.1007/s00894-018-3689-5
- Lesage, G., & Bussey, H. (2006). Cell wall assembly in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews*, *70* (2), 317-343. doi:10.1128/MMBR.00038-05
- Levin, D. E. (2011). Regulation of Cell Wall Biogenesis in *Saccharomyces cerevisiae*: The Cell Wall Integrity Signaling Pathway. *Genetics*, *189* (4), 1145-1175. doi:10.1534/genetics.111.128264
- Li, Q., Zhao, X. Q., Chang, A. K., Zhang, Q. M., & Bai, F. W. (2012). Ethanol-induced yeast flocculation directed by the promoter of TPS1 encoding trehalose-6-phosphate synthase 1 for efficient ethanol production. *Metabolic engineering*, *14* (1), 1-8. doi:10.1016/j.ymben.2011.12.003
- Maestre-Reyna, M., Diderrich, R., Veelders, M. S., Eulenburg, G., Kalugin, V., Bruckner, S., . . . Essen, L. O. (2012). Structural basis for promiscuity and specificity during *Candida glabrata* invasion of host epithelia. *Proceedings of the National Academy of Sciences of the United States of America*, *109* (42), 16864-16869. doi:10.1073/pnas.1207653109
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical*

Theory and Computation, 11 (8), 3696-3713. doi:10.1021/acs.jctc.5b00255

Martinez-Archundia, M., Hernandez Mojica, T. G., Correa-Basurto, J., Montano, S., & Camacho-Molina, A. (2019). Molecular dynamics simulations reveal structural differences among wild-type NPC1 protein and its mutant forms. *Journal of Biomolecular Structure & Dynamics*, 1-6. doi:10.1080/07391102.2019.1664324

Massova, I., & Kollman, P. A. (1999). Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *Journal of the American Chemical Society*, 121 (36), 8133-8143. doi:10.1021/ja990935j

Mishra, A., Behura, A., Mawatwal, S., Kumar, A., Naik, L., Mohanty, S. S., . . . Dhiman, R. (2019). Structure-function and application of plant lectins in disease biology and immunity. *Food and Chemical Toxicology*, 134, 110827. doi:10.1016/j.fct.2019.110827

Mishra, S. K., Sund, J., Aqvist, J., & Koca, J. (2012). Computational prediction of monosaccharide binding free energies to lectins with linear interaction energy models. *Journal of Computational Chemistry*, 33 (29), 2340-2350. doi:10.1002/jcc.23081

Neria, E., Fischer, S., & Karplus, M. (1996). Simulation of activation free energies in molecular systems. *Journal of Chemical Physics*, 105 (5), 1902-1921. doi:10.1063/1.472061

Onufriev, A., Bashford, D., & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, 55 (2), 383-394. doi:10.1002/prot.20033

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23 (3), 327-343. doi:10.1016/0021-9991(77)90098-5

Sahu, S. N., Moharana, M., Sahu, R., & Pattanayak, S. K. (2019). Impact of mutation on podocin protein involved in type 2 nephrotic syndrome: Insights into docking and molecular dynamics simulation study. *Journal of Molecular Liquids*, 281, 549-562. doi:10.1016/j.molliq.2019.02.120

Saladino, G., & Gervasio, F. L. (2016). Modeling the effect of pathogenic mutations on the conformational landscape of protein kinases. *Current Opinion in Structural Biology*, 37, 108-114. doi:10.1016/j.sbi.2016.01.005

Sariki, S. K., Kumawat, R., Singh, V., & Tomar, R. S. (2019). Flocculation of *Saccharomyces cerevisiae* is dependent on activation of Slt2 and Rlm1 regulated by the cell wall integrity pathway. *Molecular Microbiology*, 112 (4), 1350-1369. doi:10.1111/mmi.14375

Sim, L., Groes, M., Olesen, K., & Henriksen, A. (2013). Structural and biochemical characterization of the N-terminal domain of flocculin Lg-Flo1p from *Saccharomyces pastorianus* reveals a unique specificity for phosphorylated mannose. *The FEBS Journal*, 280 (4), 1073-1083. doi:10.1111/febs.12102

Soares, E. V. (2011). Flocculation in *Saccharomyces cerevisiae*: a review. *Journal of Applied Microbiology*, 110 (1), 1-18. doi:10.1111/j.1365-2672.2010.04897.x

Teunissen, A. W., & Steensma, H. Y. (1995). Review: the dominant flocculation genes of *Saccharomyces cerevisiae* constitute a new subtelomeric gene family. *Yeast*, 11 (11), 1001-1013. doi:10.1002/yea.320111102

Trott, O., & Olson, A. J. (2010). Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry*, 31 (2), 455-461. doi:10.1002/jcc.21334

Varki, A. (2017). Biological roles of glycans. *Glycobiology*, 27 (1), 3-49. doi:10.1093/glycob/cww086

Veelders, M., Bruckner, S., Ott, D., Unverzagt, C., Mosch, H. U., & Essen, L. O. (2010). Structural basis of flocculin-mediated social behavior in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (52), 22511-22516. doi:10.1073/pnas.1013210108

Verstrepen, K. J., Jansen, A., Lewitter, F., & Fink, G. R. (2005). Intragenic tandem repeats generate functional variability. *Nature genetics*, 37 (9), 986-990. doi:10.1038/ng1618

Verstrepen, K. J., & Klis, F. M. (2006). Flocculation, adhesion and biofilm formation in yeasts. *Molecular Microbiology*, 60 (1), 5-15. doi:10.1111/j.1365-2958.2006.05072.x

Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews*, 119 (16), 9478-9508. doi:10.1021/acs.chemrev.9b00055

TABLE 1 Binding free energy values (kcal/mol) of Flo1p-mannobiose complexes

system	[?] E_{ele}	[?] E_{VDW}	[?] G_{GB}	[?] G_{np}	[?] G_{bind}	[?][?] G_{bind}
WT-mant	-111.80	-27.67	120.22	-5.67	-24.95	0.00
Q117N-mant	-115.14	-35.43	117.02	-6.81	-40.37	-15.42
Q117R-mant	-127.29	-34.22	125.82	-6.35	-42.03	-17.08

TABLE 2 Binding free energy values (kcal/mol) of Flo1p-mannotriose complexes

system	[?] E_{ele}	[?] E_{VDW}	[?] G_{GB}	[?] G_{np}	[?] G_{bind}	[?][?] G_{bind}
WT-mant	-111.80	-27.67	120.22	-5.67	-24.95	0.00
Q117N-mant	-115.14	-35.43	117.02	-6.81	-40.37	-15.42
Q117R-mant	-127.29	-34.22	125.82	-6.35	-42.03	-17.08

TABLE 3 Occupancies of intermolecular hydrogen bonds in the three complexes

complex	donor	acceptor	occupancy (%)	distance (\AA)	angle ($^{\circ}$)	angle ($^{\circ}$)
WT-mant	1:O4	Asp161:OD1	92.60	2.62	154.13	154.13
	1:O2	Gln98:OE1	66.30	2.67	161.35	161.35
	Gln117:NE2	2:O2	18.00	2.88	153.92	153.92
	Ser227:OG	2:O2	17.00	2.81	162.82	162.82
	2:O6	Gly103:O	14.10	2.75	159.96	159.96
Q117N-mant	2:O4	Gln117:OE1	9.20	2.77	158.34	158.34
	1:O4	Asp161:OD1	96.70	2.60	154.88	154.88
	2:O4	Asn117:OD1	99.10	2.67	165.55	165.55
	Ser227:OG	2:O4	85.00	2.78	159.75	159.75
Q117R-mant	3:O4	Gly103:O	84.30	2.71	162.85	162.85
	1:O4	Asp161:OD1	97.10	2.62	158.10	
	1:O6	Asp161:OD1	92.80	2.71	157.79	
	1:O2	Gln98:OE1	86.70	2.69	160.32	
	Arg117:NH1	3:O6	31.50	2.87	152.62	

(Table of Content)

Yeast cell flocculation involves the interaction between the exopolysaccharides and the flocculation protein of neighboring cell. The binding site residues of the flocculation protein N-Flo1p were examined herein using computational alanine scanning mutagenesis. Two mutants Q117N and Q117R show stronger binding free energies for the protein-mannotriose binding. This study provides a way to modulate the yeast cell harvest and recovery in the beer industry and biofuel production.

Figure Legends

FIGURE 1 Diagram of oligomannose binding on the N-Flo1p

FIGURE 2 Binding free energy changes in alanine scanning for mannobiose binding

FIGURE 3 Binding pose of mannobiose on the wild type N-Flo1p protein. The manb is the mannobiose.

FIGURE 4 Variation of RMSDs of mannobiose plus Ca^{2+} and residues in the active site. The numbers 1 and 2 denote the mannose residues.

FIGURE 5 Fluctuation of RMSDs of mannotriose plus Ca^{2+} and residues in the active site

FIGURE 6 Binding modes of mannotriose in (A) Q117N-mant (red) superimposed with WT-mant (cyan), (B) Q117R-mant (red) superimposed with WT-mant (cyan), (C) Q117N-mant and (D) Q117R-mant, where the numbers 1, 2 and 3 denote the mannose residues; the green ball is the calcium cation.









