# Supporting Information

# Delineating the tryptophan-galactosylamine conjugate mediated structural distortions in $A\beta_{42}$ protofibril

Arushi Dabas and Bhupesh Goyal*

Department of Chemistry & Biochemistry, Thapar Institute of Engineering & Technology, Patiala–147004, Punjab, India

*Corresponding author

Email: bhupesh@iitbombay.org; bhupesh@thapar.edu

## Table of contents

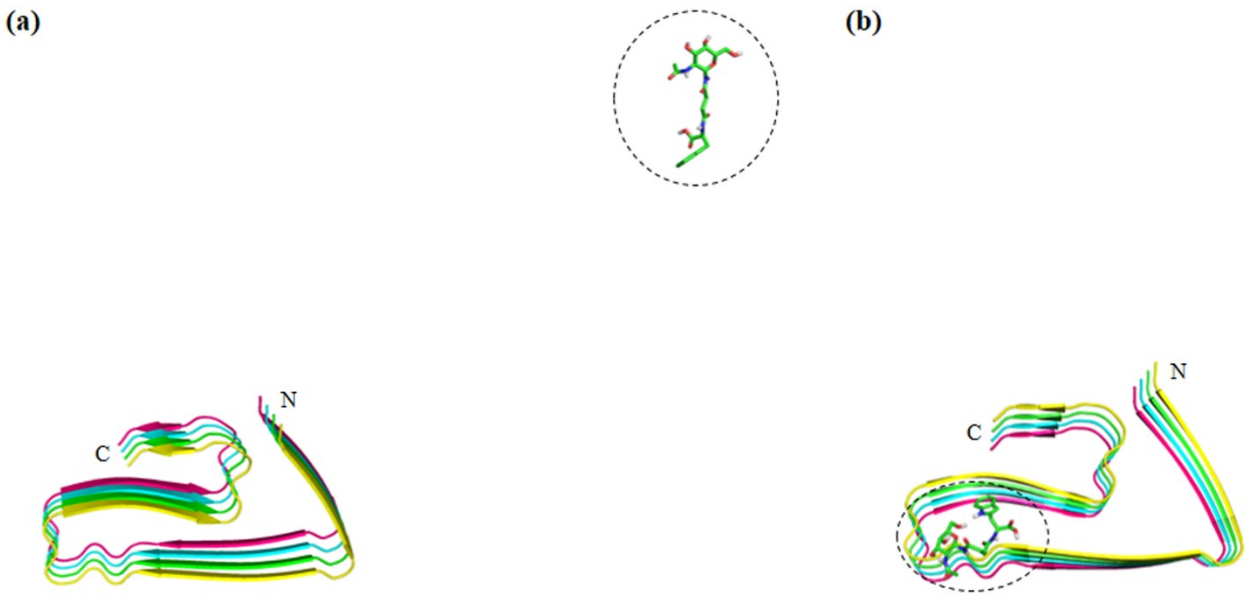



**Figure S1**: Input conformation of 5OQV protofibril and WGalNAc for molecular docking (panel a). The best-docked pose of WGalNAc with 5OQV protofibril (panel b).

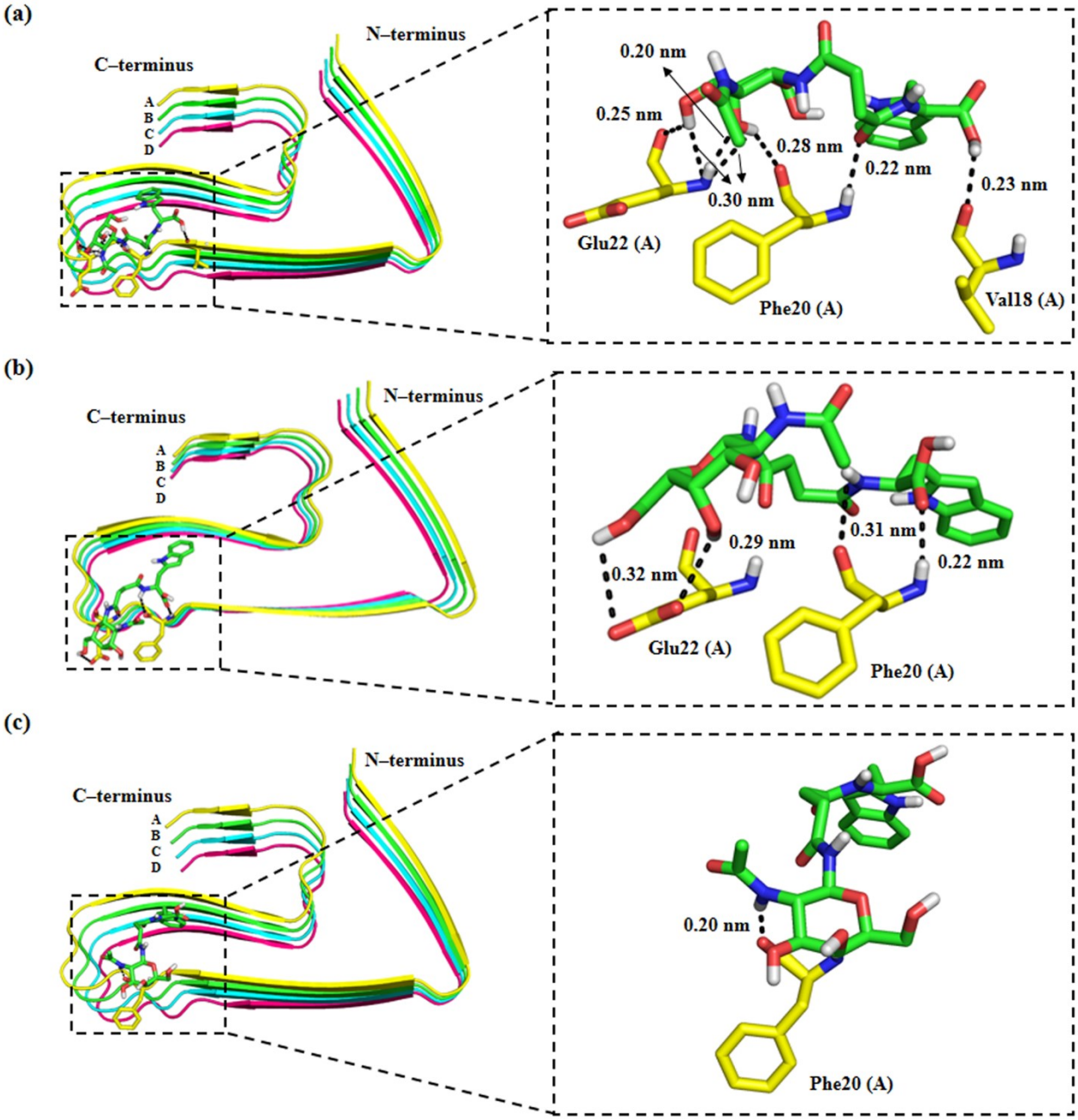



**Figure S2**: Illustration of the docked poses of 5OQV protofibril (cartoon) with WGalNAc (stick) with variable distance (panel b) and position (panel c). WGalNAc displayed hydrogen bonds with chain A residues of 5OQV protofibril (enlarged view).

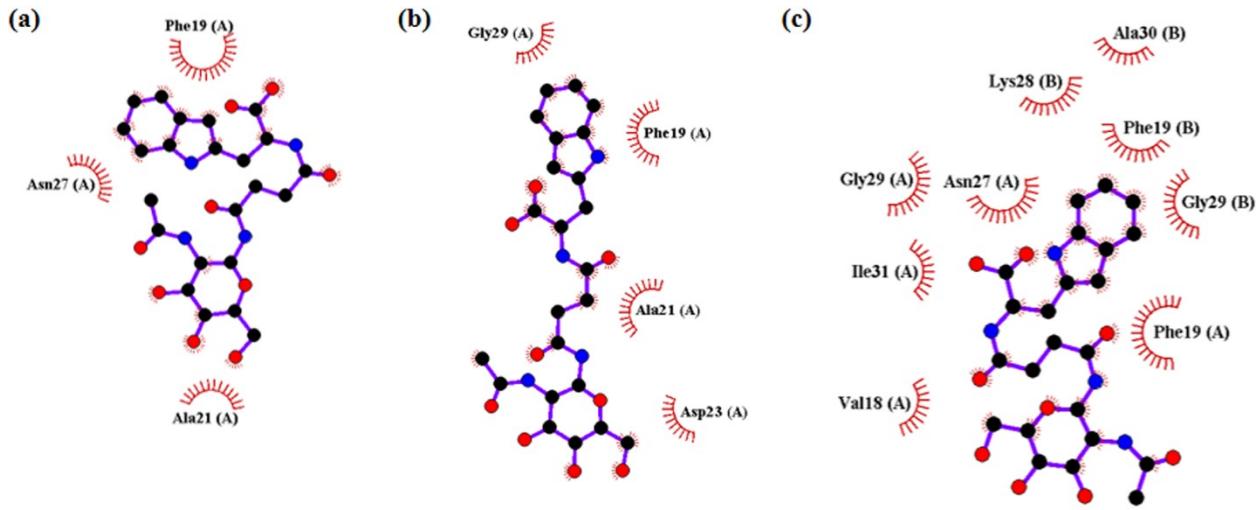



**Figure S3**: 2D interaction map generated using LigPlot+ depicts the hydrophobic contacts between 5OQV protofibril and WGalNAc with variable distance (panel b) and position (panel c).

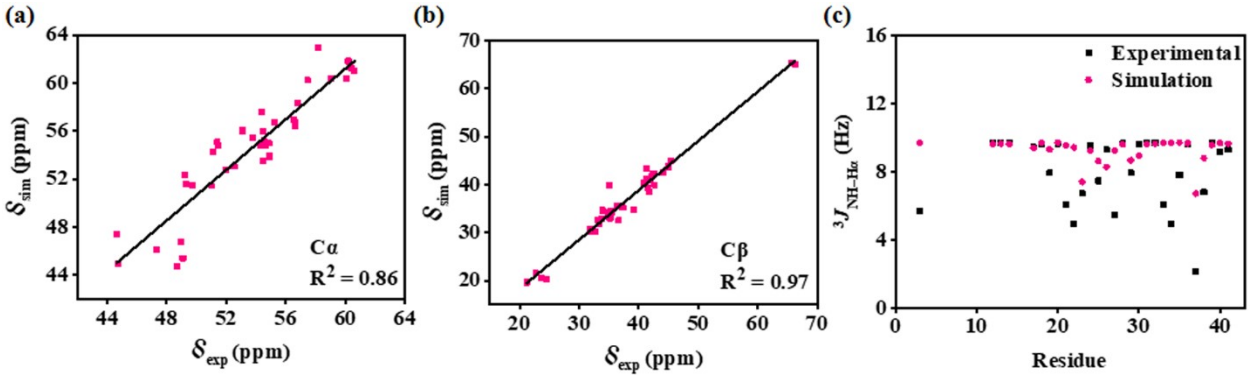



**Figure S4**: Correlation between experimental and simulated NMR chemical shift data calculated for Cα (panel a) and Cβ (panel b) atoms of the 5OQV protofibril. The values of the *J*-coupling ($^3J_{NH\text{-}H\alpha}$) constants for $A\beta_{42}$ residues computed for the experimental structure and representative member of the most-populated conformational cluster of 5OQV protofibril.

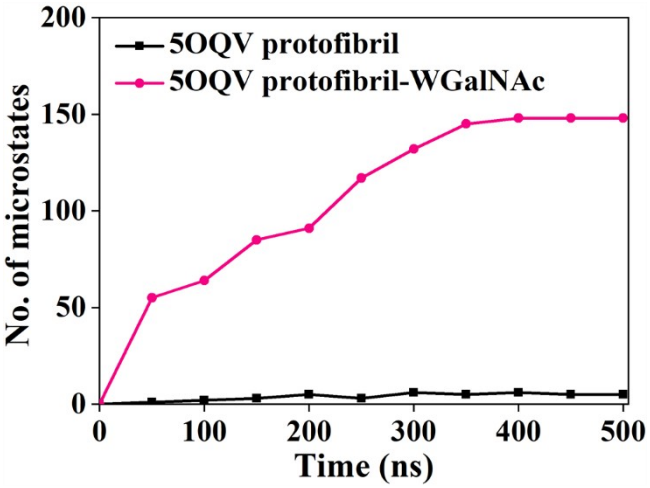



**Figure S5**: Evolution of microstates (equivalent to conformational clusters) in 5OQV protofibril and 5OQV protofibril-WGalNAc during simulation.

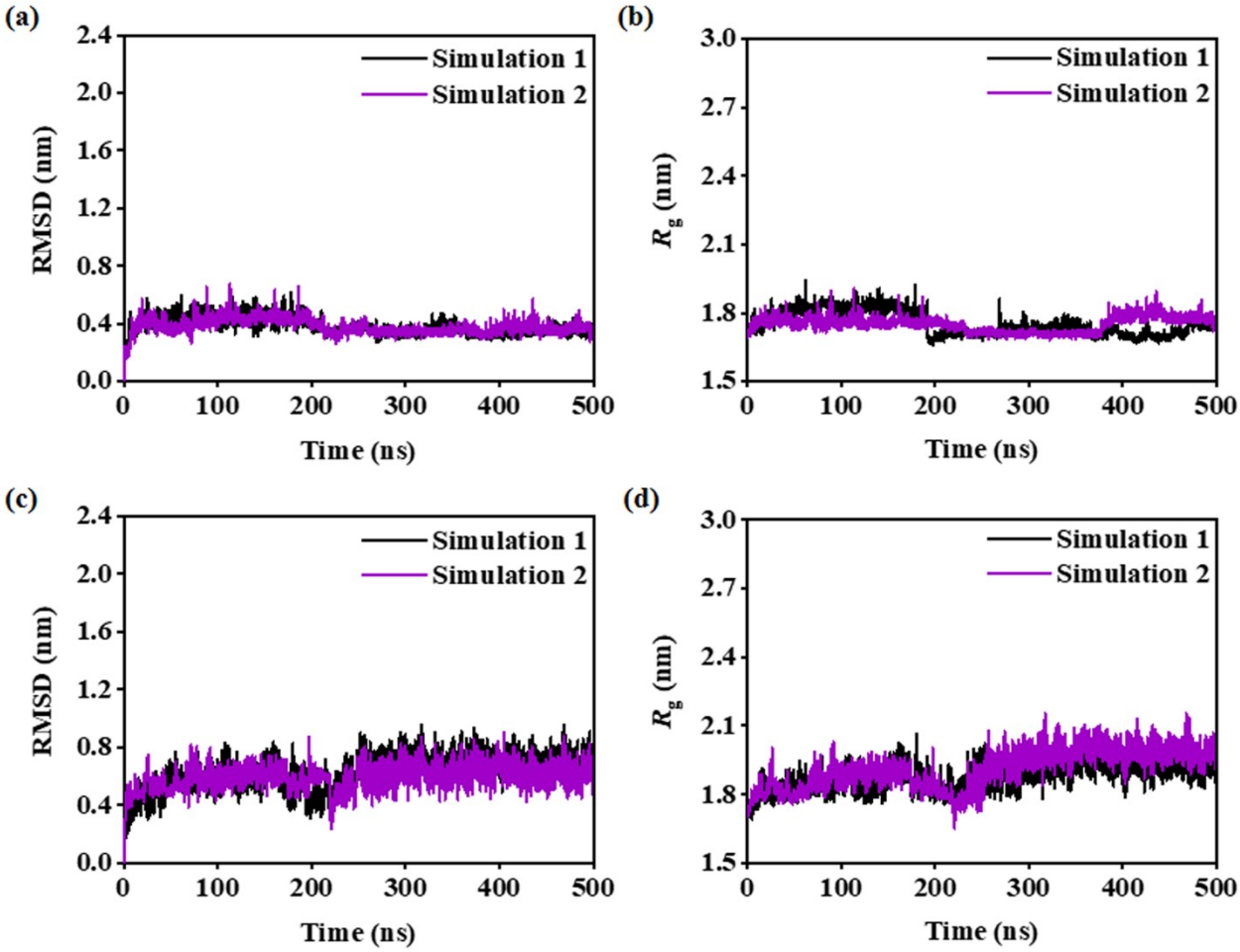



**Figure S6**: Variations in the RMSD and $R_g$ of 5OQV protofibril (panels a-b) and 5OQV protofibril-WGalNAc (panels c-d) for the repeat simulations.

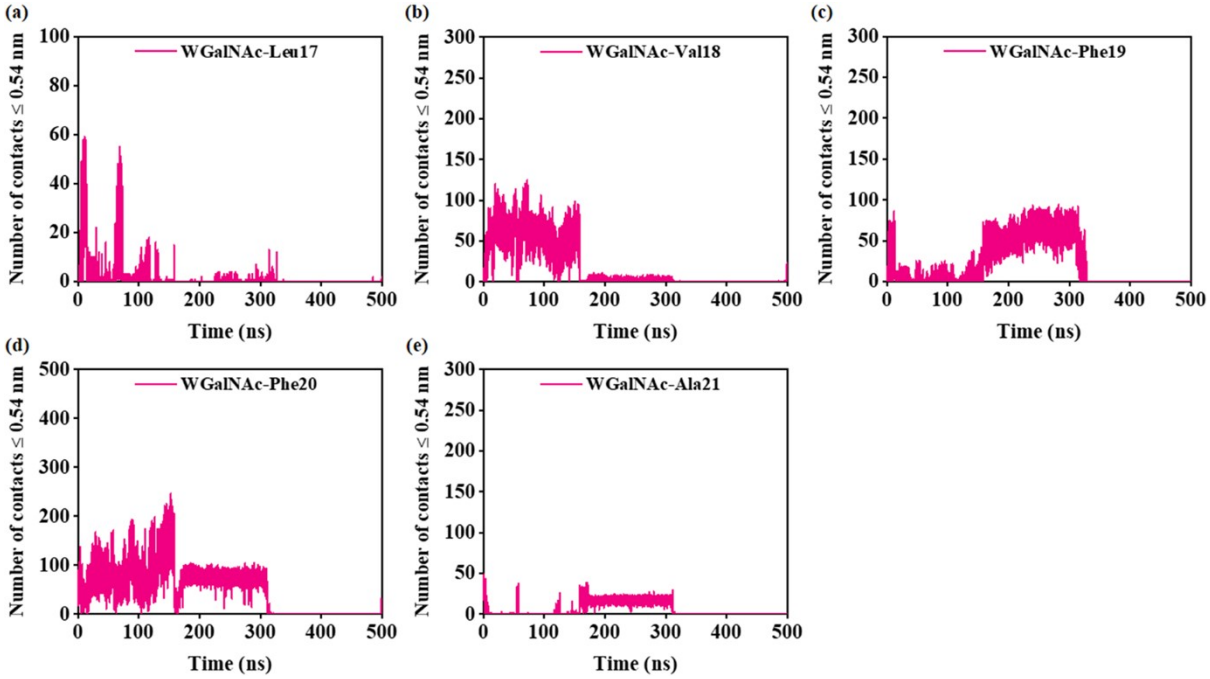



**Figure S7**: Time evolution of the contacts between WGalNAc and residues of the CHC region of 5OQV protofibril.

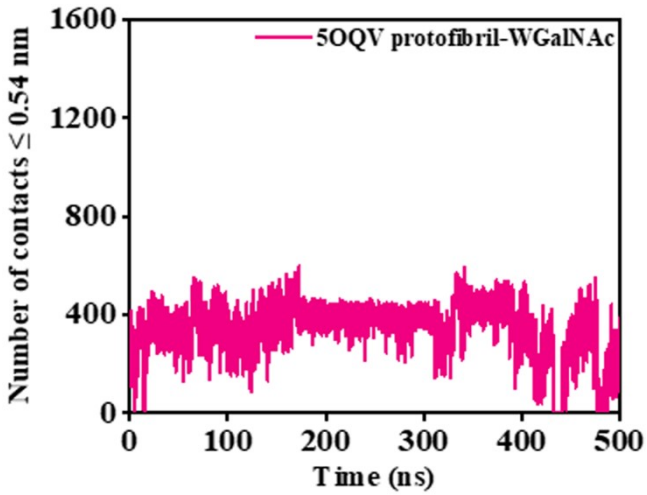



**Figure S8**: Number of contacts of WGalNAc with 5OQV protofibril during simulation.

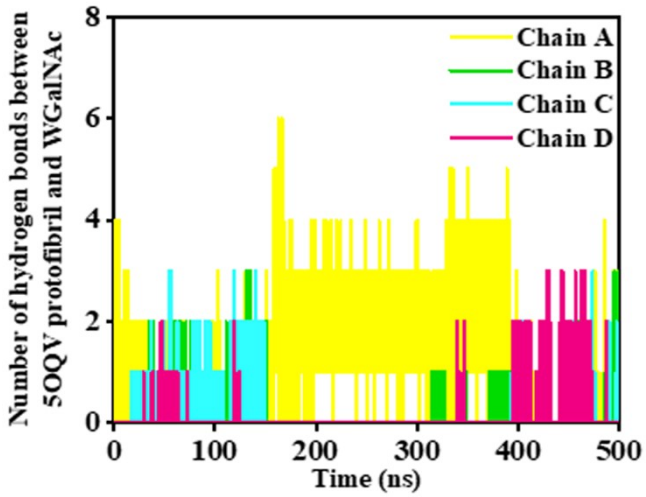



**Figure S9**: Variations in the number of hydrogen bonds of WGalNAc calculated chain-wise of 5OQV protofibril during simulation.

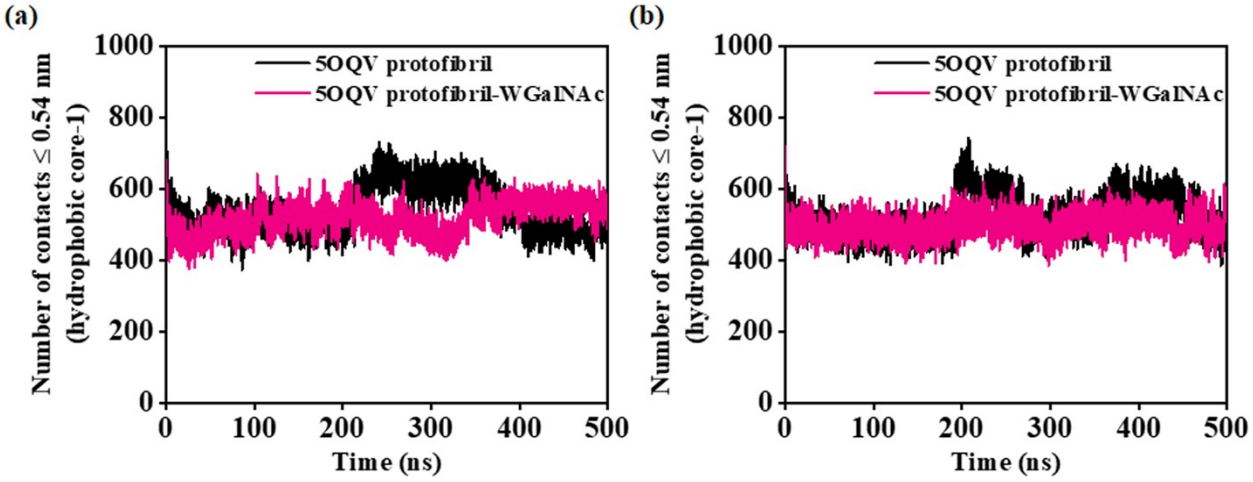



**Figure S10**: WGalNAc influences side chain contacts within hydrophobic core-1 of 5OQV protofibril as noted in the repeat simulations 2 (panel a), and 3 (panel b).

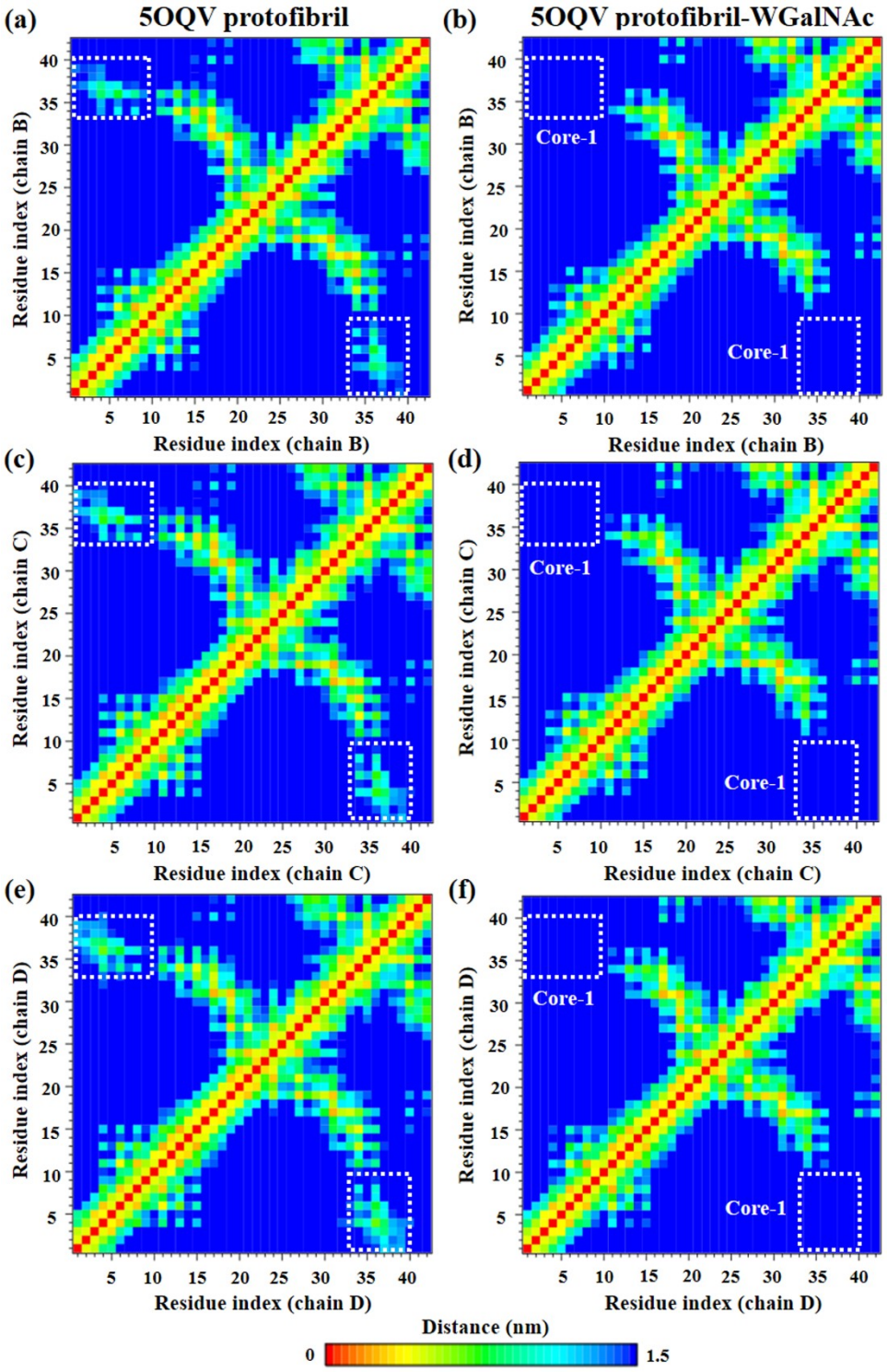



**Figure S11**: Intrachain side-chain interaction map for chains B, C, and D of 5OQV protofibril in absence and presence of WGalNAc. The cut-off distance between atoms used to define contact is 1.5 nm. The tertiary contacts between residues in 5OQV protofibril chains and hydrophobic core 1 were disrupted (shown in dotted rectangular boxes) on the incorporation of WGalNAc.

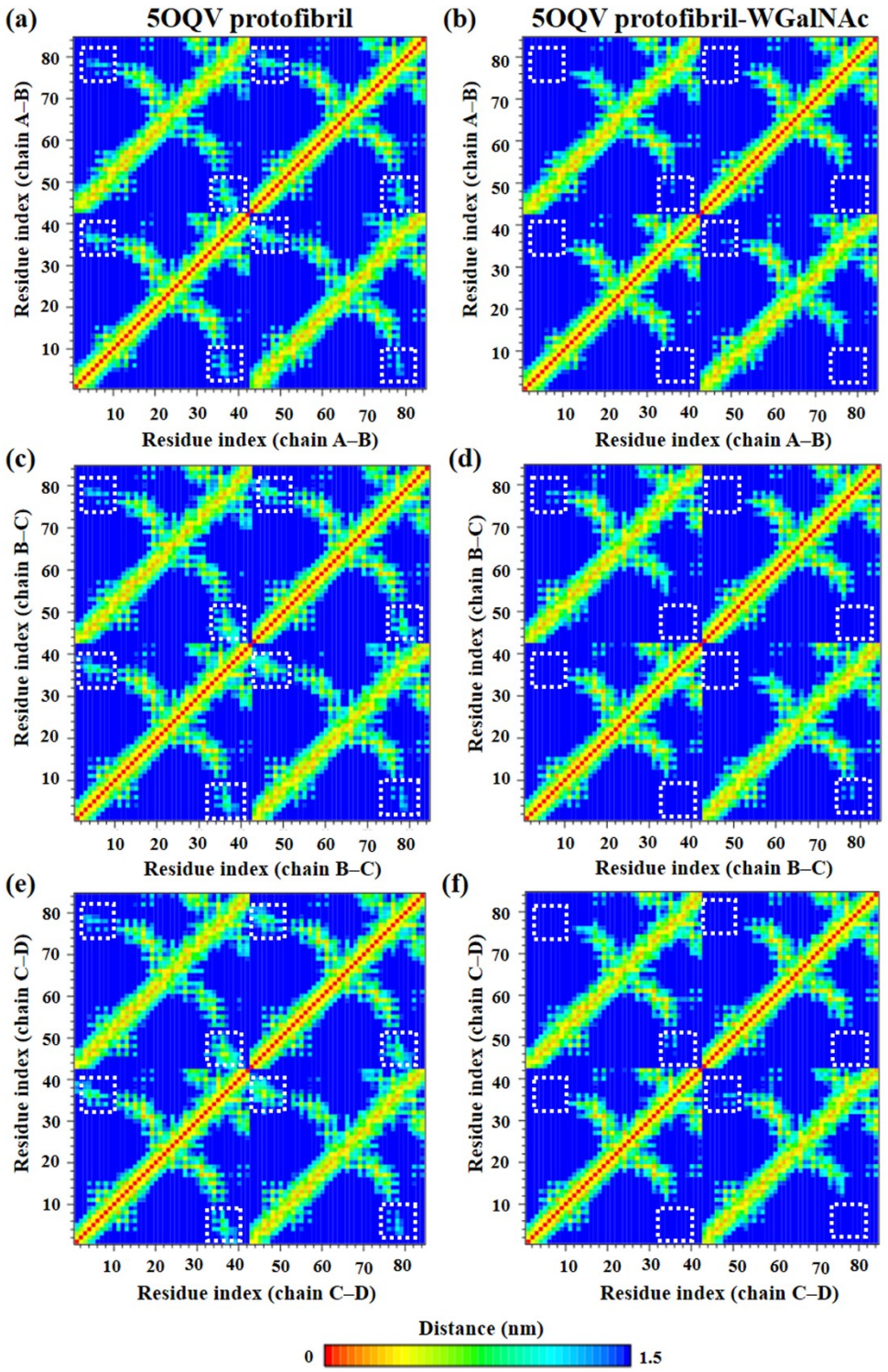



**Figure S12**: Interchain side-chain interaction map of 5OQV protofibril in absence and presence of WGalNAc. The cut-off distance between atoms used to define contact is 1.5 nm. In the presence of WGalNAc, the interchain tertiary contacts between the residues of 5OQV protofibril and hydrophobic core 1 were disrupted as shown in dotted rectangular boxes.

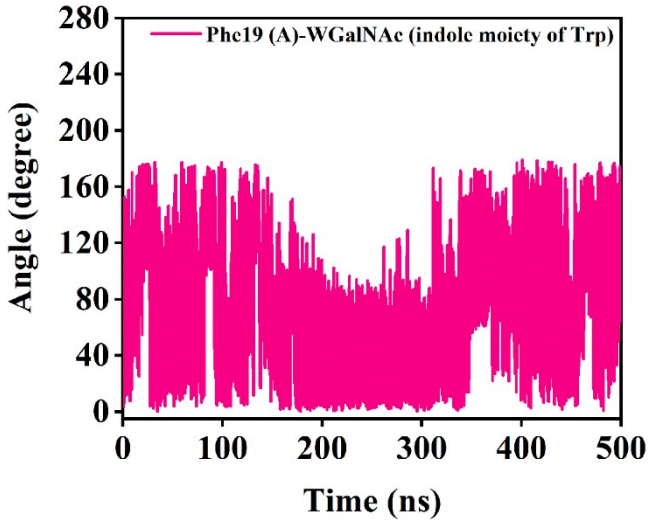



**Figure S13**: Variations in the angle between planes of Phe19 (chain A) of 5OQV protofibril and indole moiety of Trp of WGalNAc during simulation.

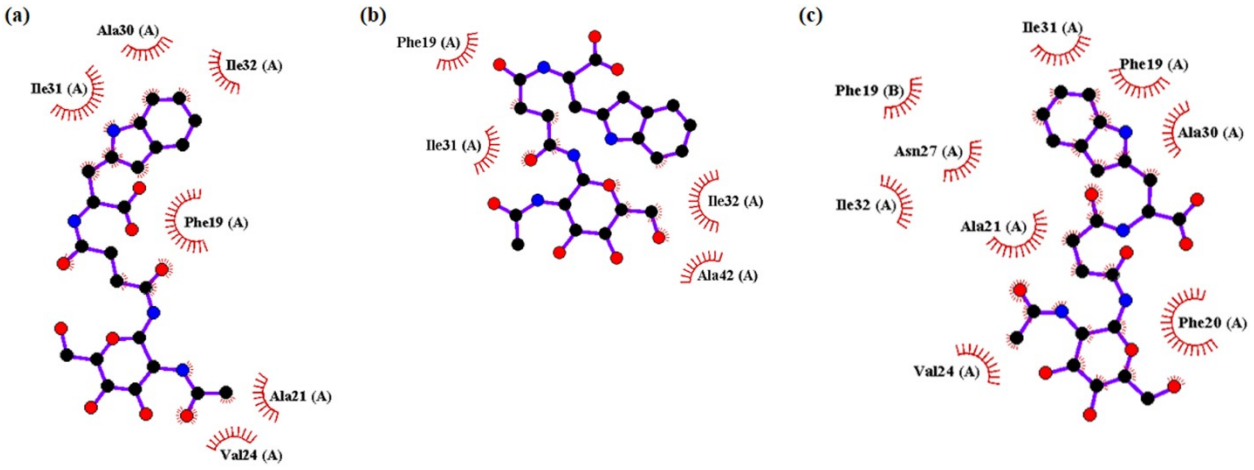



**Figure S14**: 2D interaction maps displaying the hydrophobic contacts in the representative conformations of the three most-populated microstates $m_1$, $m_2$, and $m_3$ of 5OQV protofibril in panels a, b, and c, respectively.

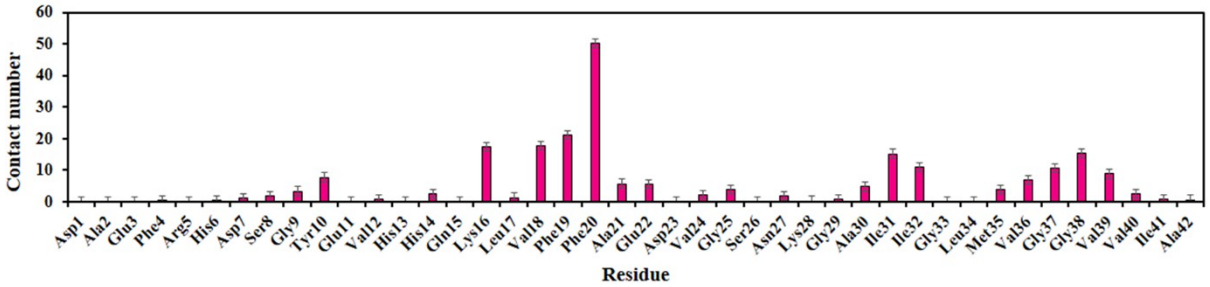



**Figure S15**: Number of contacts of WGalNAc with the residues of 5OQV protofibril during simulation.

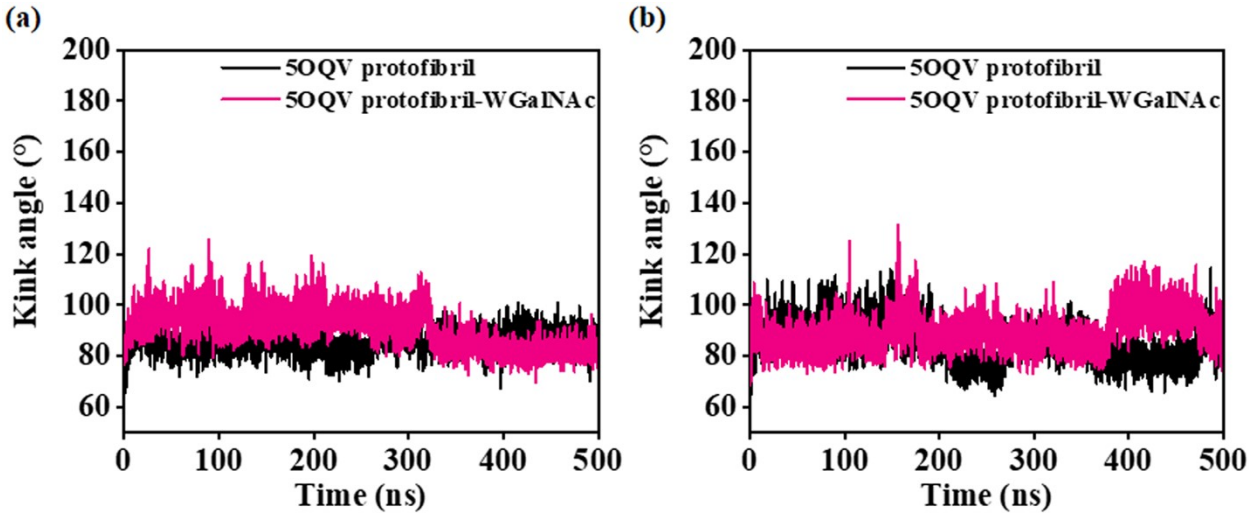



**Figure S16**: Variation of kink angle in 5OQV protofibril and 5OQV protofibril-WGalNAc in the repeat simulations 2 (panel a), and 3 (panel b).

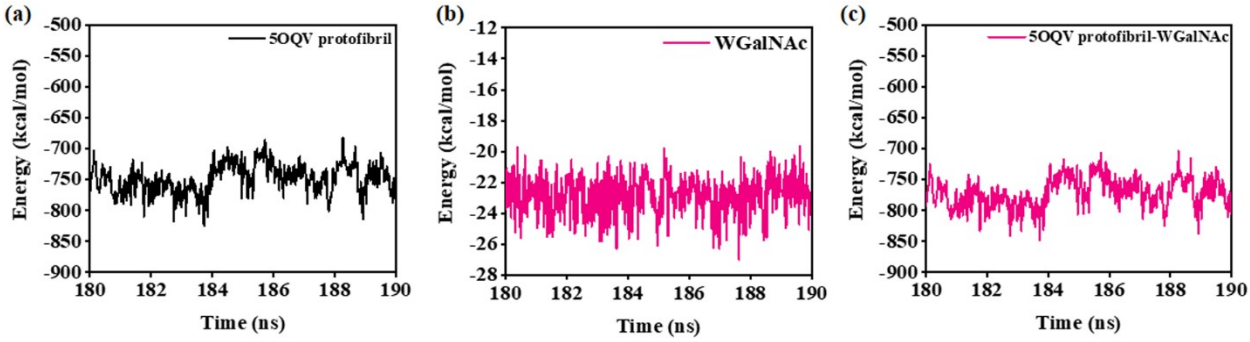



**Figure S17**: Variations in the polar solvation energies for 5OQV protofibril, WGalNAc, and 5OQV protofibril-WGalNAc during simulation.

**Table S1**: AutoDock Vina binding energy (kcal/mol) and the key residues of 5OQV protofibril involved in hydrogen bonds and hydrophobic contacts with WGalNAc.

| Compound | AutoDock Vina binding energy (kcal/mol) | 5OQV protofibril residues participating in hydrogen bonds with WGalNAc | | | 5OQV protofibril residues involved in hydrophobic contacts with WGalNAc |
|---|---|---|---|---|---|
| | | Residue | Atom[a] | Distance (nm) | |
| WGalNAc | –6.60 | Val18 (A) | CO: HXT | 0.23 | Phe19 (A), Ala21 (A) Asn27 (A) |
| | | Phe20 (A) | NH: OC | 0.22 | |
| | | | CO: HO | 0.28 | |
| | | Glu22 (A) | HN: HO | 0.20 | |
| | | | CO: HO | 0.25 | |
| | | | N: HO | 0.30 | |
| | | | N: HO | 0.30 | |

[a]Atom on the left correspond to main chain of 5OQV protofibril, whereas on the right correspond to WGalNAc.

**Table S2**: Top three docked conformations of WGalNAc with 5OQV protofibril.

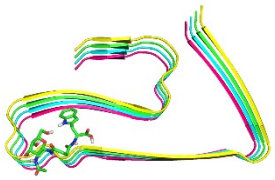
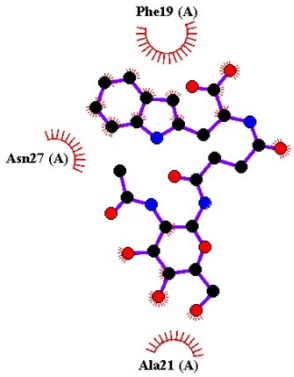
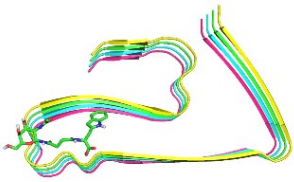
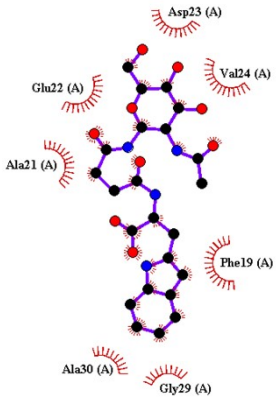
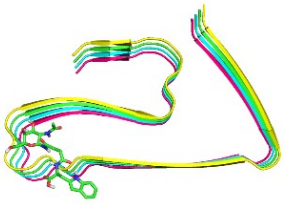
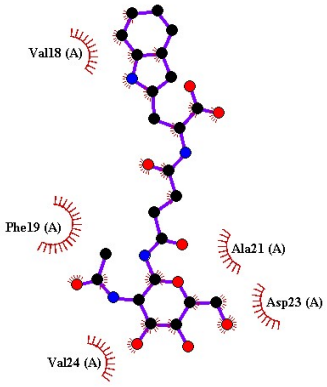
| Docked pose | Binding energy (kcal/mol) | Docked conformation | Hydrophobic contacts | Residues involved in hydrophobic contacts |
|---|---|---|---|---|
| 1 | –6.60 |  |  | Phe19 (A), Ala21 (A), Asn27 (A) |
| 2 | –6.10 |  |  | Phe19 (A), Ala21 (A), Glu22 (A), Asp23 (A), Val24 (A), Gly29 (A), Ala30 (A) |
| 3 | –5.80 |  |  | Val18 (A), Phe19 (A), Ala21 (A), Asp23 (A), Val24 (A) |

**Table S3**: AutoDock Vina binding energy (kcal/mol) and the key residues of 5OQV protofibril involved in hydrogen bonds and hydrophobic contacts with WGalNAc in the docked poses obtained with variable distance and position.

| Compound | AutoDock Vina binding energy (kcal/mol) | 5OQV protofibril residues participating in hydrogen bonds with WGalNAc | | | 5OQV protofibril residues involved in hydrophobic contacts with WGalNAc |
|---|---|---|---|---|---|
| | | Residue | Atom[a] | Distance (nm) | |
| WGalNAc | –6.60 | Val18 (A) | CO: HXT | 0.23 | Phe19 (A), Ala21 (A), Asn27 (A) |
| | | Phe20 (A) | NH: OC | 0.22 | |
| | | | CO: HO | 0.28 | |
| | | Glu22 (A) | HN: HO | 0.20 | |
| | | | CO: HO | 0.25 | |
| | | | N: HO | 0.30 | |
| | | | N: HO | 0.30 | |
| Varied relative distance | | | | | |
| WGalNAc | –6.70 | Phe20 (A) | NH: OC | 0.22 | Phe19 (A), Ala21 (A), Asp23 (A), Gly29 (A) |
| | | | CO: HN | 0.31 | |
| | | Glu22 (A) | OE1: HO | 0.29 | |
| | | | OE2: HO | 0.32 | |
| Varied relative position | | | | | |
| WGalNAc | –7.10 | Phe20 (A) | CO: HN | 0.20 | Val18 (A), Phe19 (A), Asn27 (A), Gly29 (A), Ile31 (A), Phe19 (B), Lys28 (B), Gly29 (B), Ala30 (B) |

**Table S4**: The molecular docking analyses for 5OQV protofibril-WGalNAc using AutoDock Vina and MVD.

| Compound | Binding energy | | Hydrogen bonds | | Hydrophobic contacts | |
|---|---|---|---|---|---|---|
| | AutoDock Vina (kcal/mol) | MVD (MolDock score) | AutoDock Vina | MVD | AutoDock Vina | MVD |
| WGalNAc | –6.60 | –175.65 | Val18 (A), Phe20 (A), Glu22 (A) | Val18 (A), Ala30 (A), Lys28 (C) | Phe19 (A), Ala21 (A), Asn27 (A) | Phe19 (A), Gly29 (A), Ile31 (A), Phe19 (B), Asn27 (B), Gly29 (B), Ala30 (B), Ile31 (B), Phe19 (C), Gly29 (C), Ala30 (C), Ile31 (C), Gly29 (D), Ala30 (D) |

**Table S5**: Variations in average RMSD of chains A, B, C, and D regions for 5OQV protofibril and 5OQV protofibril-WGalNAc.

| System | Average RMSD (nm) | |
|---|---|---|
| | 5OQV protofibril | 5OQV protofibril-WGalNAc |
| Chain A | 0.42 ± 0.02 | 0.58 ± 0.03 |
| Chain B | 0.34 ± 0.02 | 0.56 ± 0.03 |
| Chain C | 0.31 ± 0.02 | 0.57 ± 0.03 |
| Chain D | 0.31 ± 0.02 | 0.60 ± 0.03 |

**Table S6**: Conformational clustering analysis with population of the three most-populated clusters of 5OQV protofibril and 5OQV protofibril-WGalNAc.

| System | Number of clusters | Population of three most-populated clusters (%) | | |
|---|---|---|---|---|
| | | $m_1$ | $m_2$ | $m_3$ |
| 5OQV protofibril | 7 | 78.71 | 15.55 | 5.39 |
| 5OQV protofibril-WGalNAc | 145 | 34.04 | 28.45 | 9.43 |

**Table S7**: Detailed binding free energy calculated for 5OQV protofibril–WGalNAc.

| Energy terms (kcal/mol) | 5OQV protofibril–WGalNAc |
|---|---|
| $\Delta E_{vdW}$ | –27.16 ± 1.90 |
| $\Delta E_{elec}$ | 8.17 ± 0.40 |
| $\Delta E_{MM}{}^{a}$ | –18.99 ± 1.50 |
| $\Delta G_{ps}$ | 0.35 ± 1.10 |
| $\Delta G_{nps}$ | –3.12 ± 0.20 |
| $\Delta G_{solv}{}^{b}$ | –2.77 ± 0.90 |
| $\Delta G_{binding}{}^{c}$ | –21.76 ± 2.40 |

$^{a}\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{elec}$ ; $^{b}\Delta G_{solv} = \Delta G_{ps} + \Delta G_{nps}$ ; $^{c}\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solv}$

**Table S8**: Interchain binding free energy averaged over the three neighbouring chains (*i.e.*, chain A–B, chain B–C, and chain C–D) for 5OQV protofibril and 5OQV protofibril-WGalNAc.

| Energy terms (kcal/mol) | 5OQV protofibril | 5OQV protofibril–WGalNAc |
|---|---|---|
| $\Delta E_{vdW}$ | –195.92 ± 7.18 | –166.76 ± 21.24 |
| $\Delta E_{elec}$ | 2.79 ± 0.77 | 3.73 ± 1.89 |
| $\Delta E_{MM}{}^{a}$ | –193.13 ± 6.41 | –163.03 ± 19.35 |
| $\Delta G_{ps}$ | 79.22 ± 18.07 | 58.56 ± 21.99 |
| $\Delta G_{nps}$ | –20.40 ± 0.54 | –17.41 ± 4.59 |
| $\Delta G_{solv}{}^{b}$ | 58.82 ± 17.53 | 41.15 ± 17.40 |
| $\Delta G_{binding}{}^{c}$ | –134.31 ± 11.12 | –121.88 ± 1.95 |

$^{a}\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{elec}$ ; $^{b}\Delta G_{solv} = \Delta G_{ps} + \Delta G_{nps}$ ; $^{c}\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solv}$

**Table S9**: ADMET properties of WGalNAc.

| Category | Property (unit) | WGalNAc | References range |
|---|---|---|---|
| Absorption | Caco-2 permeability log (cm/s) | –5.81 | Optimal: Higher than –5.15 |
| | HIA (Human Intestinal Absorption) (%) | 55.91 | Optimal:>=80%<br>Medium:30-80%<br>Poor:<=30% |
| | log $D_{7.4}$ (Lipophilicity) | 1.46 | Optimal:1~3 |
| Distribution | Blood-brain barrier penetration (%) | 26.73 | Optimal:<=30% Medium:30%-70% Poor:>=70% |
| | PPB (Plasma protein binding) (%) | 56.69 | Optimal:<=90% Poor:>90% |
| Metabolism | CYP2C9 Substrate | Low | >0.5: A substrate;<br><0.5: Non substrate |
| | CYP2C9 Inhibition (%) | 43.66 | |
| | CYP3A4 Substrate | Low | >0.5: A substrate;<br><0.5: Non substrate |
| | CYP3A4 Inhibition (%) | 41.45 | |
| Excretion | Clearance microsome (ml/ min/Kg) | 0.05 | High: >15 ml/min/kg; Moderate: 5-15 ml/min/kg; Low: <5 ml/min/kg |
| | $T_{1/2}$ (Half-life) (h) | 113.01 | Long half-life: >3 h; Short half-life: <3 h |
| Toxicity | hERG blockers (%) | 43.49 | Optimal:<=30% Medium:30%-70% Poor:>=70% |
| | AMES (Ames mutagenicity) (%) | 41.24 | Optimal:<=30% Medium:30%-70% Poor:>=70% |
| | DILI (Drug Induced Liver Injury) (%) | 53.82 | Optimal:<=30% Medium:30%-70% Poor:>=70% |