Supplementary data

Atomistic scale simulation study of structural properties in the silk-fibrohexamerin complex

Wooboum Park[a,*], Taeyoung Yoon[a,*], Hyun Joon Chang[b], Juneseok You[a,†] and Sungsoo Na[a,†]

[a]Department of Mechanical Engineering, Korea University, 02841, Seoul, Republic of Korea

[b]HITS Inc., 124, Teheran-ro, Gangnam-gu, Seoul, 06234, Republic of Korea

*Equal contribution to the paper

[†]Corresponding Author: Sungsoo Na, Juneseok You

E-mail: nass@korea.ac.kr, protrossko1@korea.ac.kr

1. Supplementary Methods

1.1 Molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) Calculation

To investigate the structural/interfacial properties of fibrohexamerin (P25), which contains N-glycan, and the *Bombyx mori* silkworm, the binding energy for each simulation model was calculated using MM/PBSA calculations.[1] The binding energy is expressed as:

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

where $G_{complex}$ represents the total free energy of the fibrohexamerin and silkworm silk proteins, $G_{protein}$ and G_{ligand} denote the total free energy of fibrohexamerin and silkworm silk, respectively. The total free energy can be expressed as follows:

$$G_x = \langle E_{MM} \rangle + \langle G_{solvation} \rangle \quad \textbf{(2)}$$

where G_x represents the total free energy of the protein, ligand, or complex. The molecular mechanics potential energy E_{MM} is

$$E_{MM} = E_{bonded} + E_{nonbonded} = E_{bonded} + (E_{VdW} + E_{elec.}) \quad \textbf{(3)}$$

where E_{bonded} represents bonded interactions consisting of a bond, angle, dihedral, and improper interactions. Nonbonded interactions include van der Waals interaction $(E_{VdW})$and electrostatic forces (E_{Elec}). The free energy of solvation is expressed as follows

$$G_{solvation} = G_{polar} + G_{nonpolar} \quad \textbf{(4)}$$

where G_{polar} and $G_{nonpolar}$ represent the electrostatic and non-electrostatic contributions to the solvation-free energy, respectively. The polar solvation energy can be approximated by solving the Poisson–Boltzmann equation:

$$\nabla \cdot [\varepsilon(r)\nabla \cdot \varphi(r)] - \varepsilon(r)\kappa(r)^2 \sinh[\varphi(r)] + \frac{4\pi\rho^f(r)}{kT} = 0 \quad \textbf{(5)}$$

where $\varphi(r)$, $\varepsilon(r)$, and $\rho^f(r)$ represent the electrostatic potential, the dielectric constant, and the fixed charge density, respectively. The term $-\kappa^2$ refers to the reciprocal of the Debye

length, which is affected by the ionic strength of the solution. Non-polar solvation energy includes repulsive and attractive forces generated by cavity formation and van der Waals interactions between the solute and solvent, respectively. The energy is expressed as follows

$$G_{nonpolar} = G_{cavity} + G_{VdW} \quad \textbf{(6)}$$

where G_{cavity} and G_{VdW} represent the work done by the solute to generate a cavity in the solvent and depends on the geometry of the solute and

$$G_{nonpolar} = \gamma A + pV + G_{VdW} \quad \textbf{(7)}$$

G_{VdW}, γ, A, p, and V represent the van der Waals energy between the solvent and solute, a coefficient related to the surface tension of the solvent, the SASA, a coefficient related to the pressure of the solvent, and the SAV, respectively. The first and second terms account for the energy of cavity formation.

2. Supplementary Figures and Table

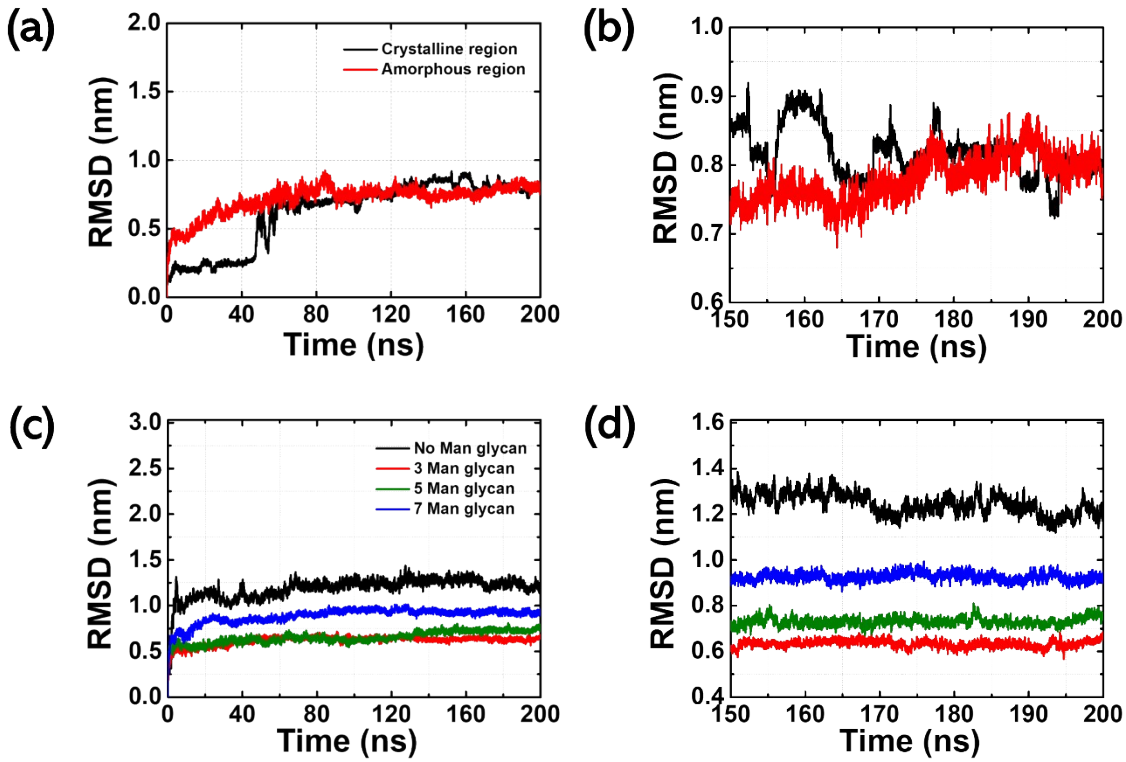

Figure S1. RMSD graph for crystalline/amorphous domain(a). RMSD graph for the crystalline/amorphous domain for the 150-200 ns range(b). RMSD for each N-glycan according to Mannose order(c) and RMSD for N-glycan over the 150-200 ns range(d).

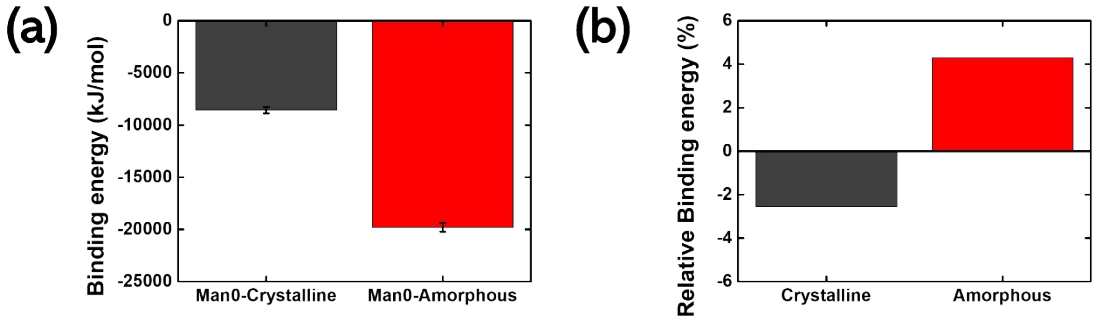

Figure S2. (a) The Binding energy of crystalline (-8569.85 kJ/mol, Black) and amorphous (-19788.18 kJ/mol, Red) domain according to without N-glycan (0-Crystalline and 0-Amorphous) and (b) the relative binding energy of crystalline (-2.56 %, Black) and amorphous (4.30 %, Red) domain according to N-glycan without/with relative binding energy for N-glycan (Man0/Man3). All data were analyzed by extracting 1000 snapshots at 10 ps in the 190-200 ns range.

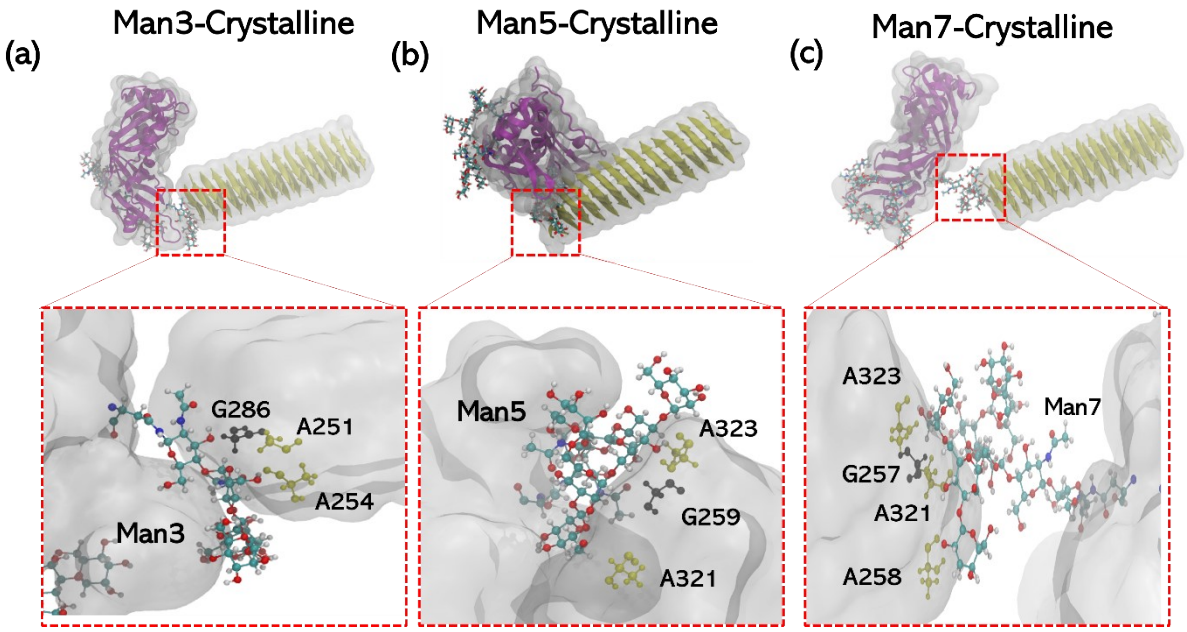

Figure S3. MD simulation result of crystalline domain according to the mannose molecule type. (a) 3-type (Man3), (b) 5-type (Man5), and (c) 7-type (Man7) crystalline domain-P25 protein complex when it is a mannose molecule. P25 protein is expressed as magenta color, crystalline domain is expressed as yellow color, and N-glycan is expressed as licorice molecule. In the additional figures, interacting amino acids are represented by one letter code and index number.

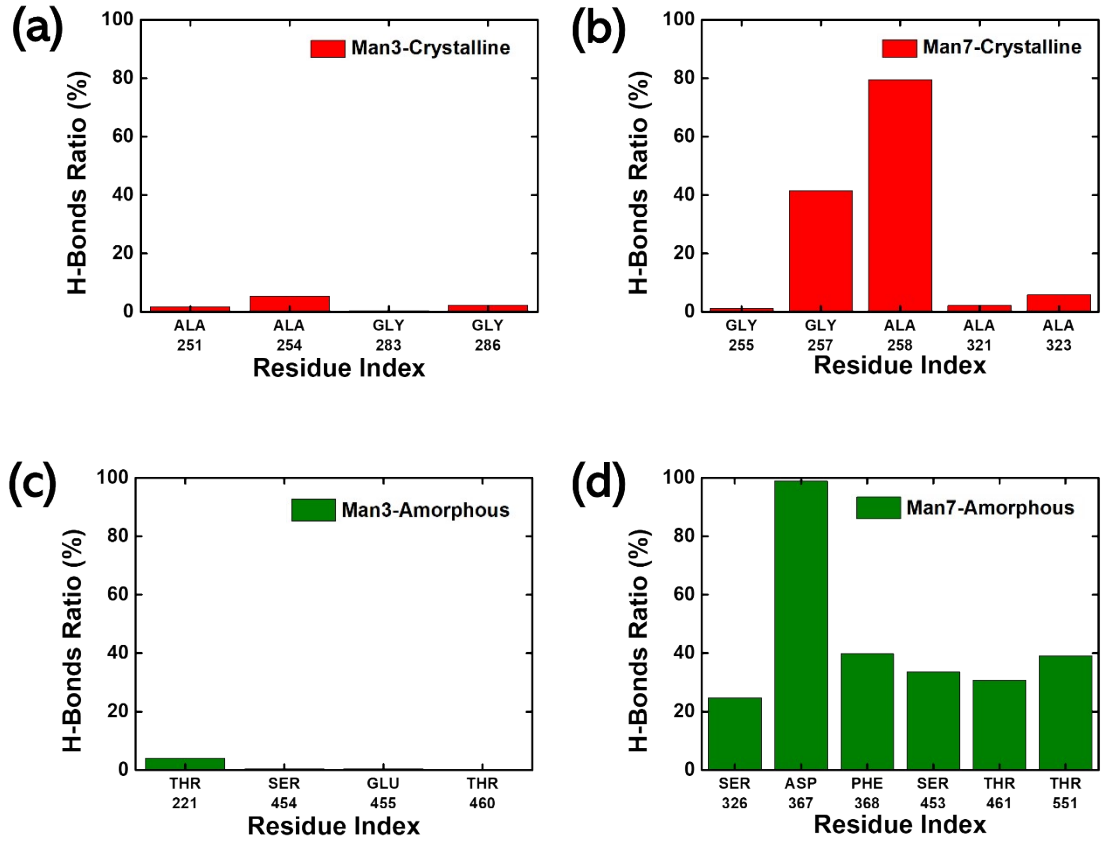

Figure S4. Analysis of amino acids contributing to hydrogen bonding with mannose molecule in (a) 3-type high mannose (Man3-Crystalline), (b) 7-type high mannose (Man7-Crystalline) N-glycan and crystalline domain; (c) 3-type high mannose (Man3-Amorphous), (b) 7-type high mannose (Man7-Amorphous) N-glycan and amorphous domain.

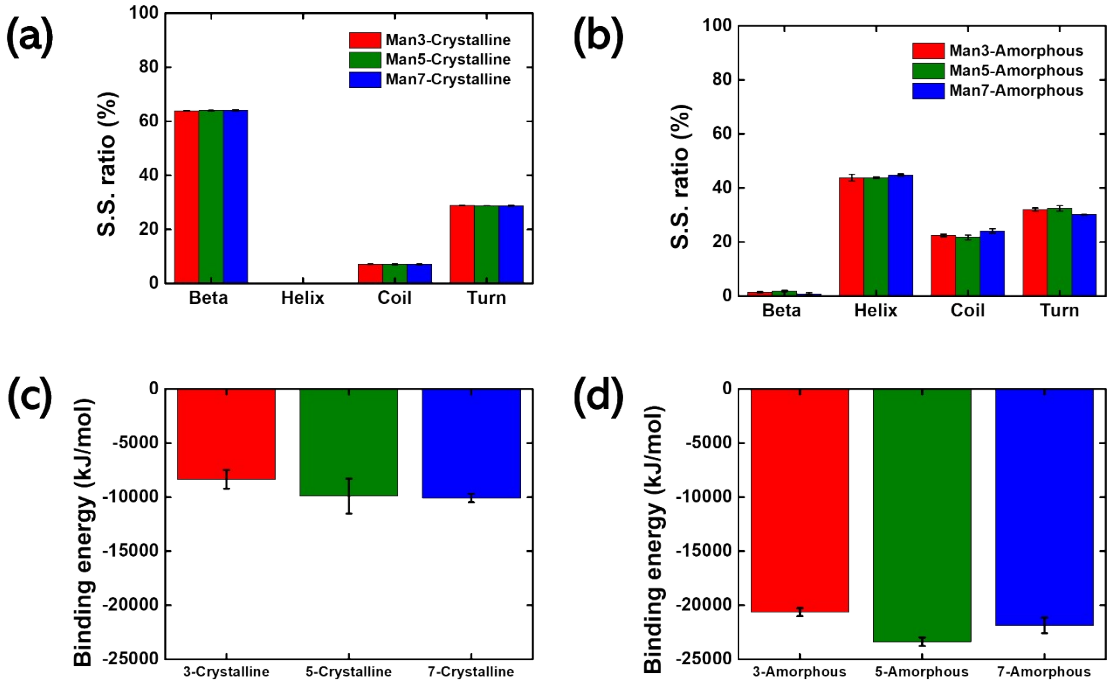

Figure S5. Secondary structure distribution of (a) crystalline and (b) amorphous domain according to the mannose molecule type of N-glycan; (c) The Binding energy of crystalline and (d) amorphous domain according to the mannose molecule type (Man3-Crystalline : -8356.20 kJ/mol, Man5-Crystalline : -9906.95 kJ/mol, Man7-Crystalline : -10087.85 kJ/mol, Man3-Amorphous : -20634.51 kJ/mol, Man5-Amorphous : -23369.38 kJ/mol, Man7-Amorphous : -21858.17 kJ/mol)

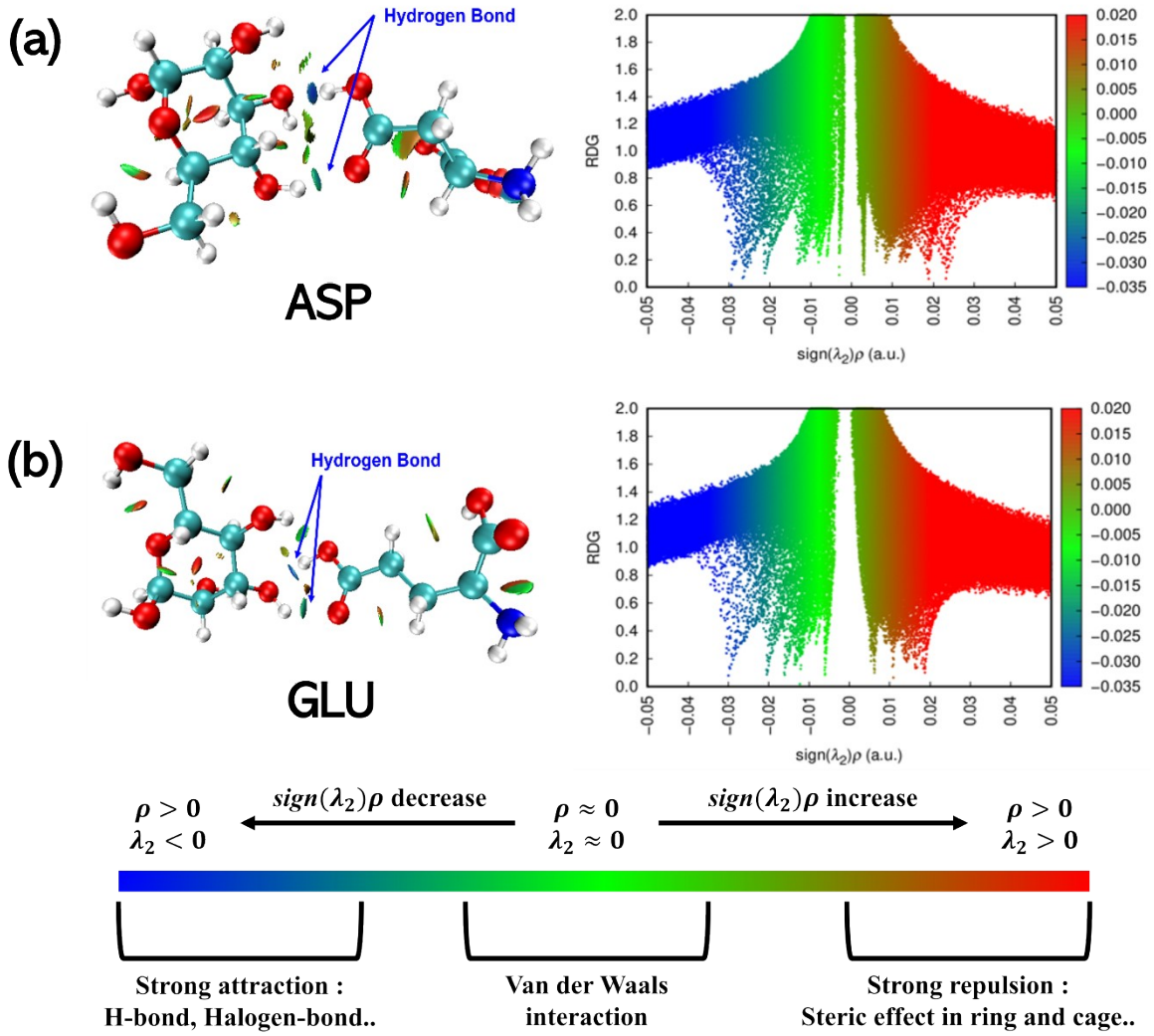

Figure S6. NCI analysis of negative charged amino acids in amorphous domain (Aspartic acid(a), Glutamic acid(b)) with mannose molecule. The blue, green, and red areas in the snapshot represent the hydrogen bond interaction, the van der Waals interaction, and the steric effect, respectively.

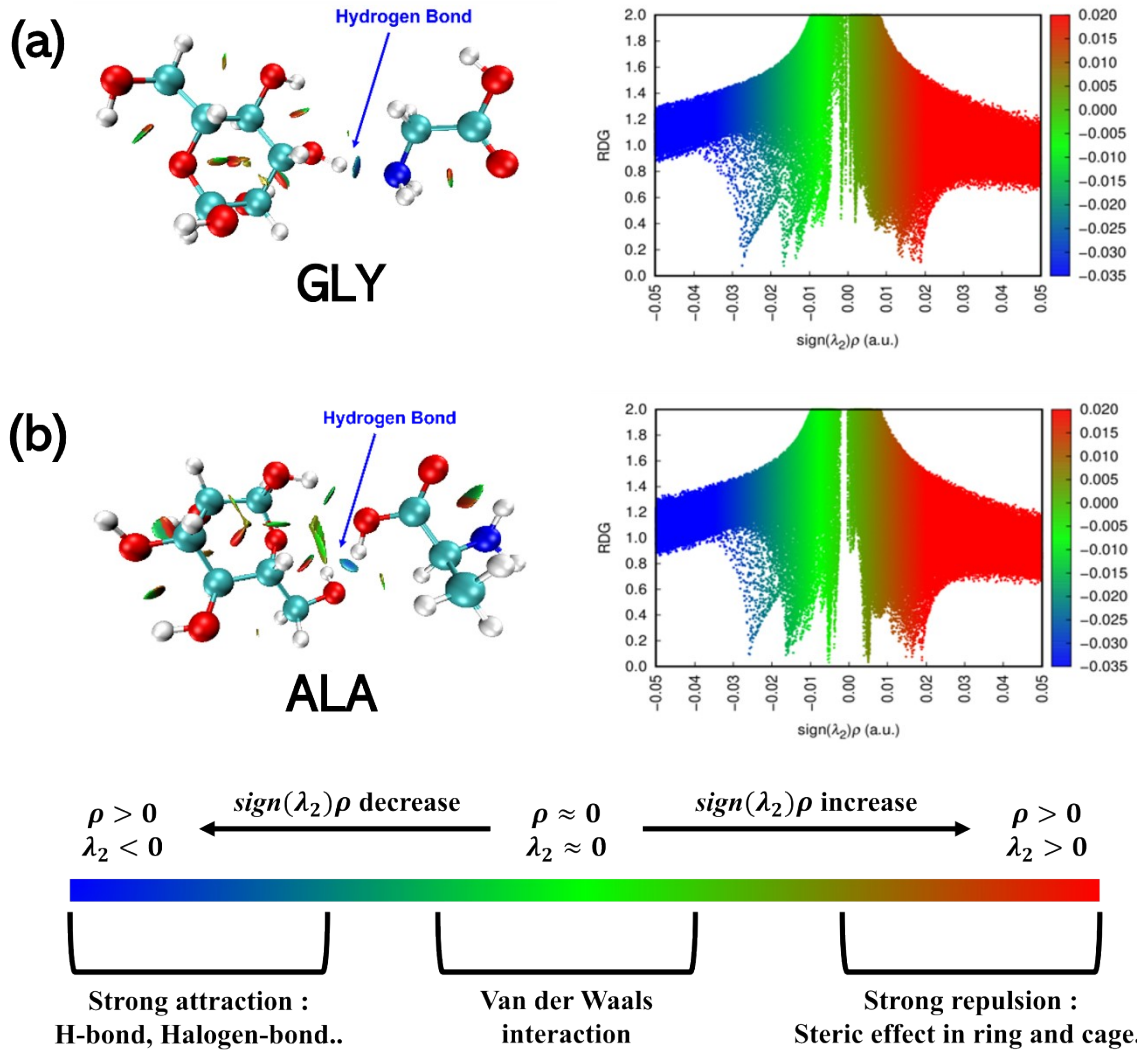

Figure S7. NCI analysis of amino acids in crystalline domain (Glycine(a), Alanine(b)) with mannose molecule. The blue, green, and red areas in the snapshot represent the hydrogen bond interaction, the van der Waals interaction, and the steric effect, respectively.

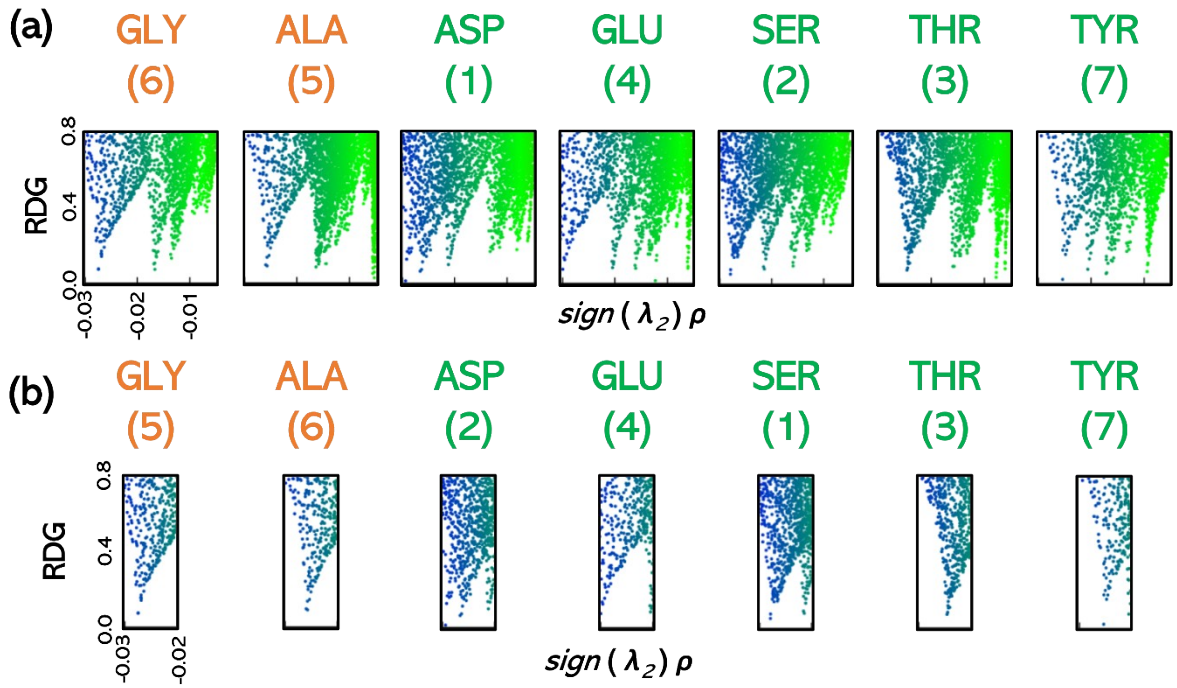

Figure S8. NCI analysis of amino acids in crystalline (Orange text) and amorphous (Light green text) domain (Glycine, Alanine, Aspartic acid, Glutamic acid, Serine, Threonine, Tyrosine) with mannose molecule (a) considering $sign(\lambda_2)\rho$ from -0.03 to -0.005 (blue and green area) and RGD from 0.0 to 0.8; (b) considering $sign(\lambda_2)\rho$ from -0.03 to -0.02 (blue area) and RGD from 0.0 to 0.8. The blue, green, and red areas in the snapshot represent the hydrogen bond interaction, the van der Waals interaction, and the steric effect, respectively.

Amino acid	Energy gap (eV)	Hardness (η)	Softness (σ)	Chemical potential (μ)	Electrophilicity (ω)
GLY	6.3394 (1)	3.1697 (1)	0.1577 (7)	-6.6484 (1)	6.9725 (7)
ALA	6.0858 (3)	3.0429 (3)	0.1643 (5)	-6.7483 (3)	7.4839 (5)
ASP	5.9976 (5)	2.9988 (5)	0.1667 (3)	-6.7853 (6)	7.6765 (2)
GLU	5.9911 (6)	2.9956 (6)	0.1669 (2)	-6.7731 (4)	7.6571 (3)
SER	6.1707 (2)	3.0584 (2)	0.1621 (6)	-6.6732 (2)	7.2165 (6)
THR	6.0319 (4)	3.0160 (4)	0.1658 (4)	-6.7845 (5)	7.6309 (4)
TYR	4.1203 (7)	2.0600 (7)	0.2427 (1)	-7.0875 (7)	12.1913 (1)

Table S1. HOMO–LUMO energy gap and quantum chemical descriptors of crystalline amino acids (glycine, alanine) and amorphous and mannose amino acids (aspartic acid, glutamic acid, serine, threonine, and tyrosine).

3. Reference

(1) Rashmi Kumari, R. K., Open Source Drug Discovery Consortium, Andrew Lynn. G_mmpbsa-A GROMACS tool for high-throughput MM-PBSA calculation. *J. Chem. Inf. Model.* **2014**, *54*, 1951-1962.