Advancing Glycoprotein Insights: Bridging Glycobiology and Structural Biology

Protein glycosylation is integral in biology; however, quantitative annotation of the structure-function relationship of glycosylation is challenging. This article reports the findings of an interdisciplinary and international team formed to tackle this challenge.

The surfaces of cells and viruses are often coated with a thick layer of sugar (also known as glycans), presenting a challenge for identifying specific proteins or biomolecules within this complex environment. In nature, however, efficient biomolecular interactions take place within this forest of glycans. Most extracellular proteins are post-translationally modified by glycans with varying sizes and complexity. The masking effect of these glycans allows bacteria and viruses to evade recognition and attack by the host immune system.

Structural biology focuses on providing pivotal structural information at an atomic resolution to help to understand the mechanisms governing biological functions. Such an underlying principle is also used by the immune system, where antibodies recognize the structural characteristics of foreign substances, such as antigens, to achieve immune protection. Nevertheless, structural biology encounters two technical challenges when examining the intricate structures of protein glycans. First, glycan molecules have a complex chemical composition, and their biosynthesis lacks a one-to-one template. Second, glycans are highly dynamic, making it challenging to accurately define three-dimensional spatial distributions. Therefore, the intrinsic microscopic compositional and conformational heterogeneity presents a fundamental challenge to the accurate modeling of glycan structures.

The research team led by Shang-Te Danny Hsu (Academia Sinica) tried to solve this problem from the structural understanding of protein glycosylation started in 2018 through a collaboration with Hui-Wen Chang at the National Taiwan University and Kay-Hooi Khoo at Academia Sinica. They used cryo-electron microscopy (cryo-EM) to determine the near-atomic structure of the spike protein of a type I feline infectious peritonitis virus (FIPV).[1] They identified an exceptionally well-defined cryo-EM map enabling the modeling of one of the longest N-glycan structures in the literature. The atomic model of the fully glycosylated FIPV spike protein was built manually using the information derived from mass spectrometry (MS) analysis. Nonetheless, the data analysis and model building are laborious and not very quantitative.

To address these issues, they collaborated with Khoo to develop a robust workflow to quantitatively characterize the N-glycosylation profiles of the spike proteins of SARS-CoV-2 variants.[2] They also determined the cryo-EM structures of the same spike proteins, with an emphasis on recovering structural information of glycans. They further collaborated with Cyril Hanus at the Inserm, France, and Mateusz Sikora at the Max Planck Institute of Biophysics, Germany, to develop the computational tool called GlycoSHIELD[3] to integrate the experimental observables with computational modeling to automate the process of model building of glycoproteins (**Fig. 1**). Such a procedure can be accomplished using a personal computer in minutes rather than using a high-performance computing center for full-blown all-atom molecular dynamics simulations of fully glycosylated spike proteins solvated in an explicit solvent model. Additionally, GlycoSHIELD offers predictive powers for identifying potential receptor binding sites and antigenic targets for therapeutic purposes based on the common characteristic that the less-shielded protein surfaces are generally more likely to be used for host receptor binding and targeted by neutralizing antibodies.

Small-angle X-ray scattering (SAXS) is exceptionally versatile in describing protein structures and dynamics in solution. The research team has benefited tremendously from the support of the biological SAXS **TPS 13A** beamline research team at the NSRRC. In the context of GlycoSHIELD, they used SAXS to characterize two highly glycosylated proteins, namely, the extracellular domains (ECDs) of N-cadherin pertinent to cell–cell adhesion and

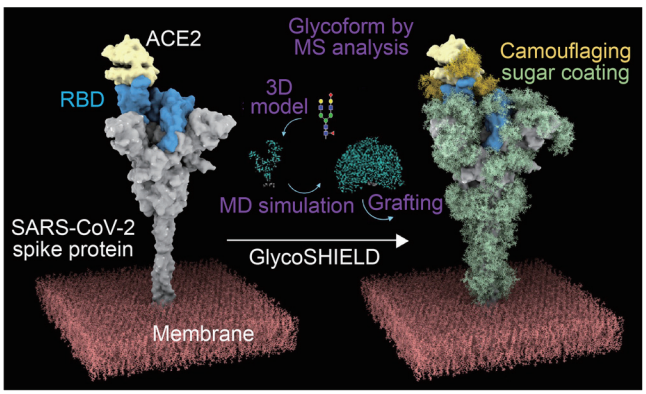

Fig. 1: Schematic overview of the principle of GlycoSHIELD.

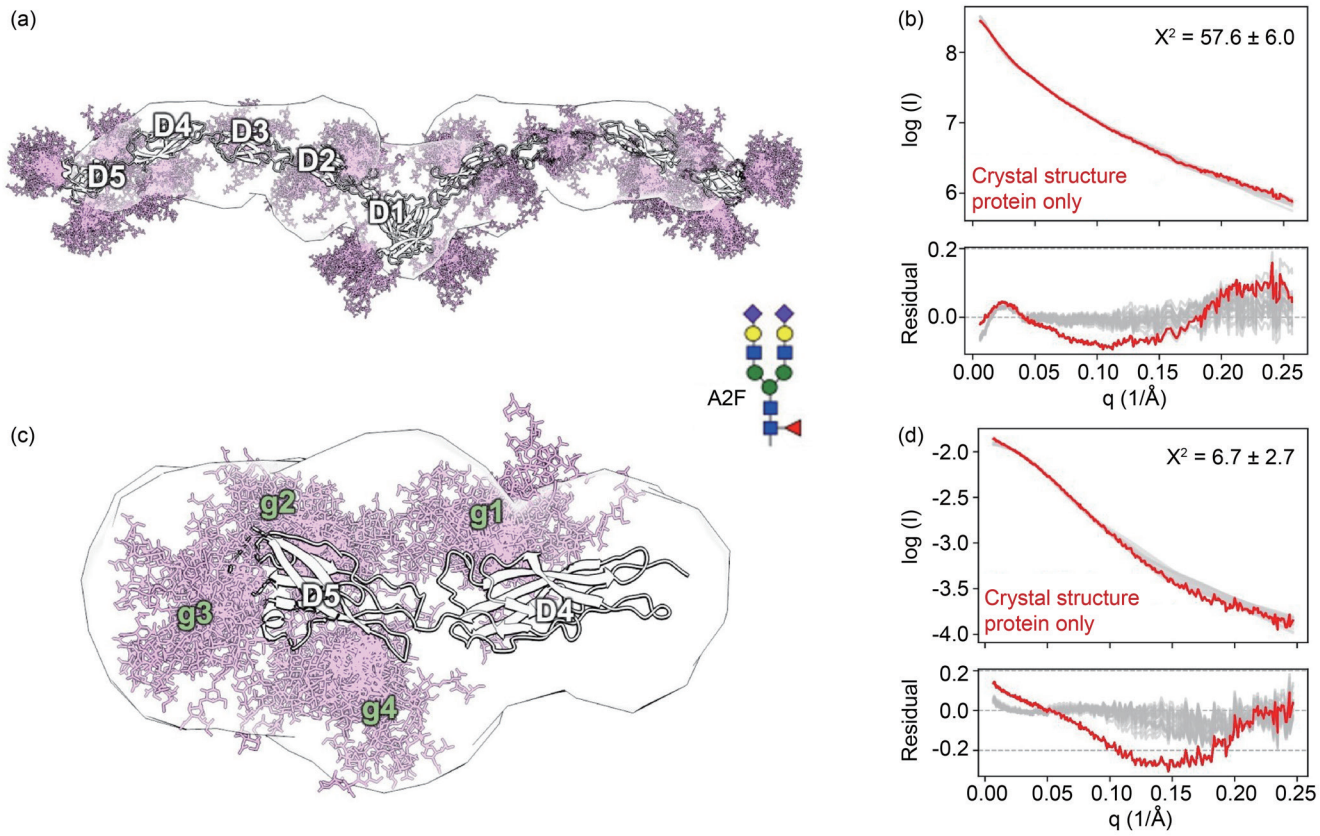

Fig. 2: SAXS analysis of N-cadherin variants. (a) The molecular envelope of N-cadherin ECD 1-5 (D1-D5) in a seagull-like homodimeric assembly. The crystal structure, indicated by a white outline, is modified with a biantennary complex type (A2F) N-glycan in its putative N-glycosylation sites, which are indicated by pink sticks. (b) The experimental SAXS profile is superimposed with the back-calculated SAXS profiles based on the crystal structure (red) and the additional N-glycan structures in 20 different conformations (gray). The residuals are shown below the SAXS profiles. Similar structural representations of the D4 and D5 of N-cadherin are shown in (c), and the corresponding SAXS data analyses are shown in (d). [Reproduced from Ref. 3]

the ECD of the receptor-like protein tyrosine phosphatase (PTPRA), which is crucial for transmitting extracellular signals to trigger intracellular chemical reactions.

N-cadherin is a calcium-binding glycoprotein comprising five folded ECDs (**Fig. 2(a)**). Molecular dynamics simulations have shown that glycosylation is important for maintaining the extension of N-cadherin. However, most of the glycans are removed to solve the crystal structure of N-cadherin, which is an intrinsic issue in protein crystallography. The team used SAXS to demonstrate that the fully glycosylated N-cadherin variants indeed adopt an extended seagull-like dimer structure, which is consistent with the crystallographic finding (**Fig. 2(a)**). By modeling different glycan ensembles onto the putative N-glycosylation sites, they could further improve the agreement between the experimental and back-calculated SAXS profiles based on the atomic ensemble structures. This illustrates how the combination of GlycoSHIELD modeling and experimental SAXS inputs can generate a more realistic glycoprotein conformational ensemble (**Figs. 2(a)–2(d)**).[3]

Structural analysis of highly glycosylated intrinsically disordered proteins represents an even more significant challenge to structural biology. In collaboration with Khoo and his colleagues, they comprehensively profiled the glycosylation of the PTPRA-ECD, which comprises only 120 amino acids with four N-glycosylation sites and over 30 O-glycosylation sites. Using the SAXS data derived from the PTPRA-ECD fused to a dimeric Fc scaffold and GlycoSHIELD, the team generated an ensemble structure with the N- and O-glycans decorating PTPRA-ECD to account for the spatial occupancy by the N- and O-glycans (**Fig. 3**). They further collaborated with Takayuki Uchihashi at Nagoya University in Japan to directly visualize the opening and closing states of the PTPRA-ECD through high-speed molecular force microscopy to confirm the bottlebrush-like molecular envelope of PTPRA-ECD inferred by SAXS. This interdisciplinary exercise underscores the importance of an integrated structural-biology approach to characterize such a highly complex and dynamic system.[4]

The current modeling algorithm of GlycoSHIELD is limited to the steric effect of glycan ensemble structures on the protein of interest. An improved algorithm is required to account for the non-covalent interactions. Furthermore, the compositional heterogeneity of the individual glycosylation sites presents a combinatorial problem of a large number

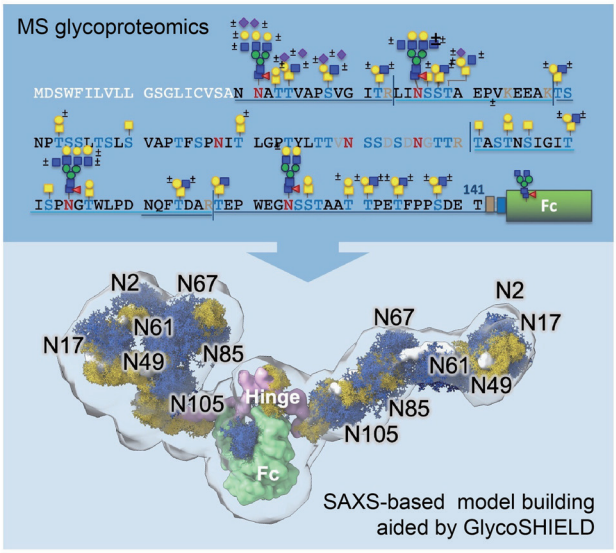

Fig. 3: Integrated structural analysis of the densely N- and O-glycosylated and intrinsically disordered PTPRA-ECD fused to Fc. The representative N- and O-glycoforms derived from MS analysis are shown on the top, and the atomic model of the fully glycosylated protein structure is shown below. The N- and O-glycans are shown in blue and yellow, respectively. The molecular envelope of the Fc-fused PTPRA-ECD is shown on a transparent surface. [Reproduced from Ref. 4]

of overall glycosylation patterns that one should consider when describing the physicochemical properties of the protein of interest at an atomic level. The optimization of GlycoSHIELD's molecular models and experimental observations is necessary to achieve a balance between the computational performance and the completeness of experimental observations while ensuring that the molecular models of glycoproteins comply with the basic laws of physics and chemistry, an area that will require further explorations. (Reported by Shang-Te Danny Hsu, Academia Sinica)

This report features the work of Shang-Te Danny Hsu and his collaborators published in Cell ***187****, 1296 (2024), and the work of Kay-Hooi Khoo and his collaborators published in JACS Au* ***3****, 1864 (2023).*

TPS 13A Biological Small-angle X-ray Scattering

- Biological Small/Wide Angle X-ray Scattering
- Biological Macromolecules, Protein Solution, Life Science

References

1. T. J. Yang, Y. C. Chang, T. P. Ko, P. Draczkowski, Y. C. Chien, Y. C. Chang, K. P. Wu, K. H. Khoo, H. W. Chang, S.-T. D. Hsu, PNAS **117**, 1438 (2020).
2 C. W. Kuo, T. J. Yang, Y. C. Chang, P. Y. Yu, S.-T. D. Hsu, K. H. Khoo, Glycobiology. **32**, 60 (2022).
3. Y. X. Tsai, N. E. Chang, K. Reuter, H. T. Chang, T. J. Yang, S. von Bülow, V. Sehrawat, N. Zerrouki, M. Tuffery, M. Gecht, I. L. Grothaus, L. C. Ciacchi, Y. S. Wang, M. F. Hsu, K. H. Khoo, G. Hummer, S.-T. D. Hsu, C. Hanus, M. Sikora, Cell **187**, 1296 (2024).
4. Y. C. Chien, Y. S. Wang, D. Sridharan, C. W. Kuo, C. T. Chien, T. Uchihashi, K. Kato, T. Angata, T. C. Meng, S.-T. D. Hsu, K. H. Khoo, JACS Au **3**, 1864 (2023).

The Key to Accurate Cell Division: Centrosome Regulation

Cep57 regulates centrosome maturation and microtubule nucleation via phase separation, ensuring accurate cell division.

Liquid–liquid phase separation (LLPS) is a process driven by multivalent interactions among biomolecules, such as proteins and nucleic acids, forming dense, dynamic structures known as biomolecular condensates. Recent evidence supports LLPS as a fundamental organizing principle for membrane-less organelles. The centrosome, consisting of a pair of centrioles surrounded by pericentriolar material (PCM), plays a crucial role in regulating cell division. As cells prepare for mitosis, the centrosome undergoes a maturation process characterized by the expansion of the PCM and an increase in its microtubule nucleation capacity by recruiting centrosomal scaffolding proteins and microtubule nucleation factors. This facilitates the assembly of the mitotic spindle and ensures the precise segregation of chromosomes during mitosis.

Human Cep57 is a coiled-coil scaffold protein located in the inner layer of the PCM and controls the process of centriole duplication and centrosome maturation for faithful cell division. Cep57 truncation mutations are genetically linked to mosaic-variegated aneuploidy syndrome, which features centrosome amplification and aberrant spindle formation. However, the molecular mechanisms by which Cep57 regulates PCM organization and microtubule assembly remain unknown. Phase separation has been implicated in some key centrosomal scaffolds, such as SPD-5 (a human CDK5RAP2 functional homolog) and the Cep63/Cep152 complex. During interphase, Cep57 forms a complex with the middle layer scaffold Cep63 and Cep152, serving as regulators for centrosome maturation. However, it is unclear whether Cep57 undergoes phase separation and how these centrosomal scaffolding components coordinate to organize the centrosome.