Recent Advances in Carbohydrate Chemistry and Application

Chi-Huey Wong

Department of Chemistry, The Scripps Research Institute
La Jolla, California 92037, USA

Abstract: Recent advances in carbohydrate chemistry and glycobiology have led to a better understanding of carbohydrates and the associated biological glycosylation in biology. This report describes new advances in carbohydrate synthesis and its application to the understanding of glycosylation in protein folding, cancer progression, influenza and SARS-CoV-2 infection and the development of carbohydrate-based medicines.

Key Words: glycan array, programmable synthesis, vaccine, cancer, SARS-CoV-2

1. Introduction

Carbohydrates are one of the four major classes of biomolecules that make up cells. They are often linked to lipids as glycolipids or to proteins as glycoproteins through glycosylation. These glycoconjugates are often expressed on cell surface and are involved in almost everything, including protein folding, viral infection, spread of cancer cells, differentiation of stem cells, egg-sperm interaction, and many other biological recognition events. Despite their importance, the roles of carbohydrates and the associated glycosylation reactions in biology and disease progression have not been well understood, mainly due to the lack of tools and methods and the availability of homo-geneous substances for the study of this class of molecules. So, over the years, we have been actively involved in the development of new methods for the study of carbohydrates and the associated glycosylation reactions [1]. These include the development of programmable one-pot chemical synthesis and chemo-enzymatic synthesis of oligosaccharides and glycoproteins, design of glyco-sylation probes for identification of cancer-specific glycans on cell surface, and development of glycan microarrays for the high throughput analysis of protein-sugar interaction. These methods have led to a better understanding of glycosylation in biology and the

development of new medicines, such as carbohydrate-based vaccines against cancer and viral infection, homogeneous anti-bodies with well-defined Fc glycan to improve effector functions and inhibitors of carbohydrate enzymes associated with disease progression (Figure 1).
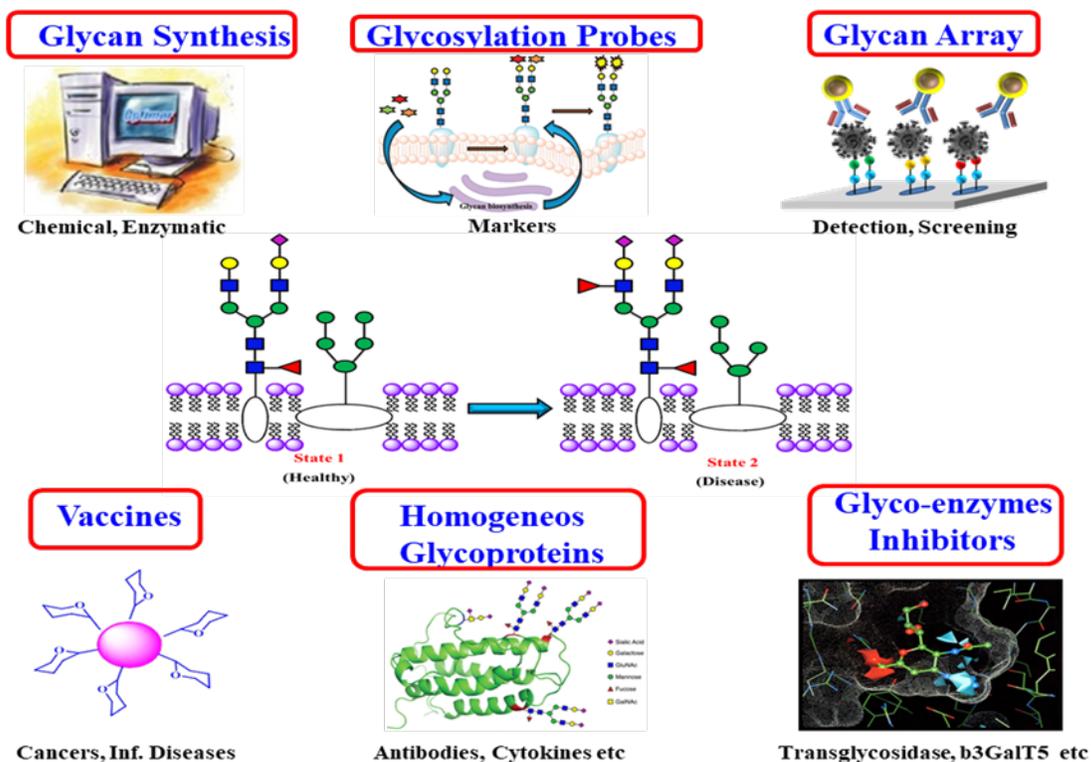

Figure 1. Development of new tools and methods for the study of cell-surface glycans: chemo-enzymatic and programmable one-pot synthesis of oligosaccharides, design of glycosylation probes to monitor biological glycosylation, and glycan microarray to study protein-glycan interaction. These methods have been used for development of vaccines against cancer and microbial infection, homogeneous antibodies to improve effector functions and glyco-enzyme inhibitors.

2. Programmable One-Pot and Chemo-enzymatic Synthesis of Oligosaccharides

In 1999, my laboratory developed the first programmable one-pot synthesis of oligosaccharides using a library of about 50 building blocks with defined relative reactivity value (RRV) for sequential glycosylation reaction [2]. Since then, we have expanded the library to more than 50,000 building blocks with predicted RRV using machine learning [3], and also included the reactivity of hydroxyl acceptors in the presence of protecting groups [4]. We have also updated the software to

guide the selection of building blocks for one-pot reaction. To use the program, one can simply enter the glycan sequence of interest to the computer, and it will show the building blocks to be added sequentially to a reaction mixture, starting from the most reactive to the least reactive one. After all the building blocks are added, the product with protecting groups is formed quickly, usually in minutes, and after global deprotection and purification, the oligosaccharide is obtained. This rapid programmable synthesis method has been used for the synthesis of various glycans including heparin pentasaccharides (Figure 2, [5]),
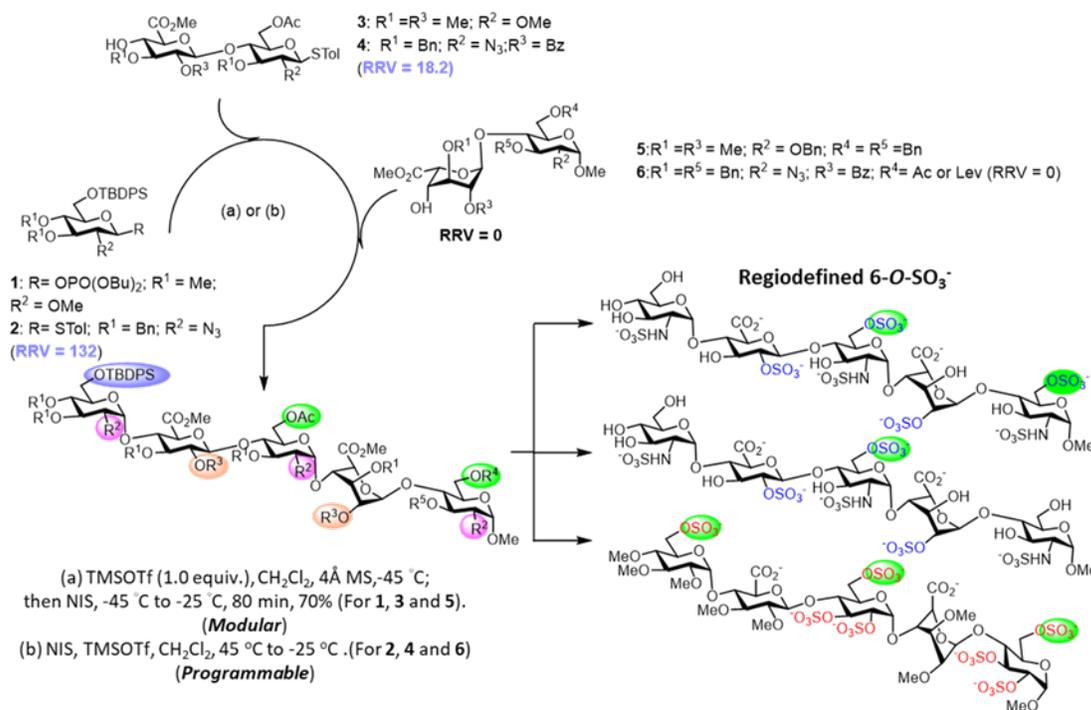

Figure 2. Programmable and Modular Synthesis of Heparin Pentasaccharides

and it is expected that this method and the automated solid-phase synthesis [6], as well as the enzymatic method, will be used to further expand the structural diversity of glycans [1].

The development of programmable one-pot method was inspired by the way an oligosaccharide is made by enzymes in our body. As shown in the enzymatic synthesis of cancer-specific Globo-H and SSEA4 (Figure 3, [7]), the first enzyme product becomes the substrate of the second enzyme, and its product becomes the substrate of the third enzyme. So, the final product can be obtained in a one-pot manner with all the enzymes mixed together, including the sugar nucleotide regeneration system that was designed to reduce the cost and eliminate the problem of feed-back inhibition caused by the released nucleoside phosphate in the glycosylation reaction.
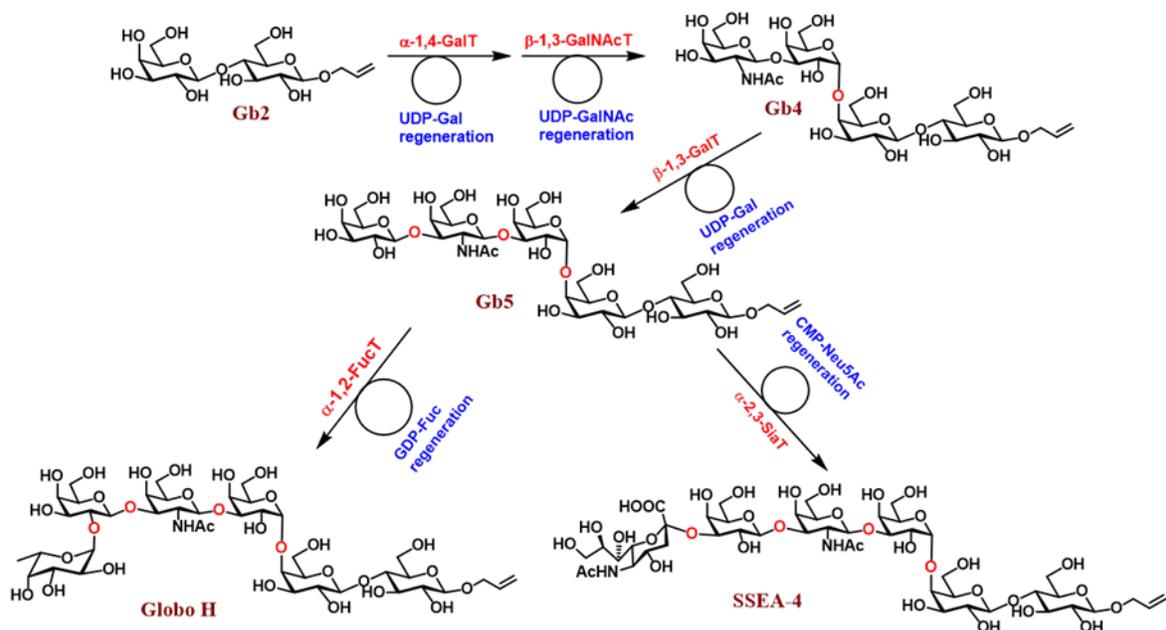
Figure 3. Chemo-enzymatic Synthesis of the Globo-series Glycans of SSEA3, SSEA4 and Globo-H

3. Glycan Microarray

With various glycans available, we can use them to create glycan arrays for the study of protein-glycan interaction [8].
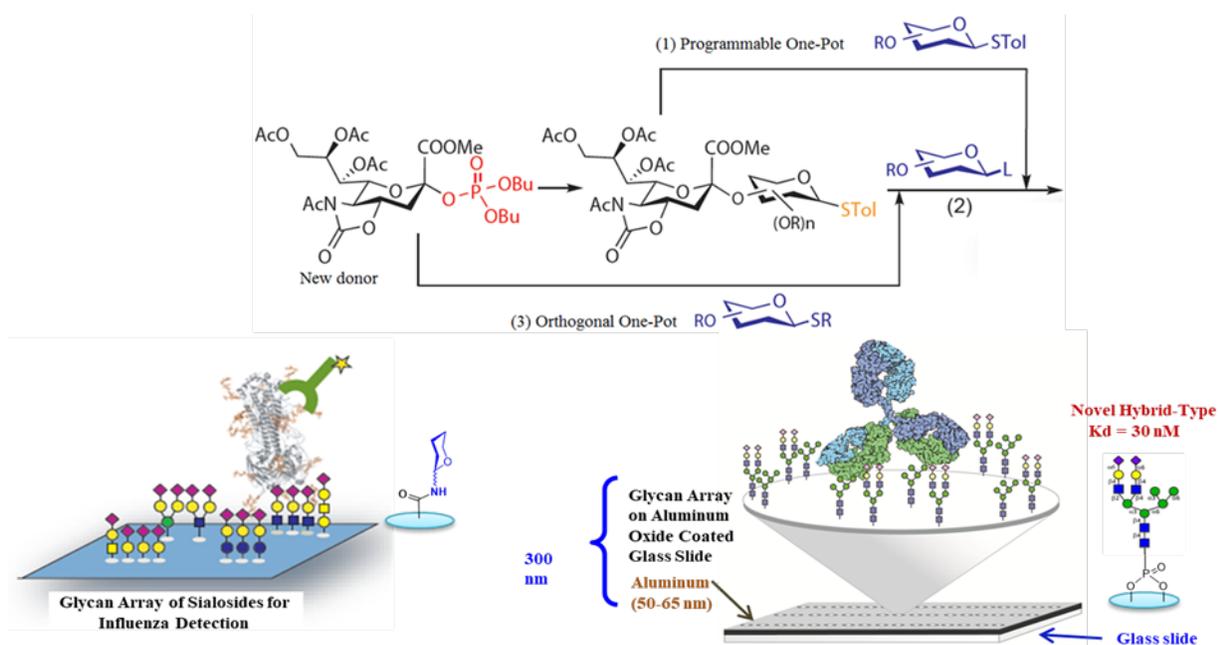

Figure 4. Synthesis of Sialosides and N-glycans for the Preparation of Glycan Array

Figure 4 shows the profiling of influenza hemagglutinins interacting with an array of sialosides on a glass slide. We know human influenza virus recognizes alpha-2,6-linked sialosides and avian virus recognizes 2,3-linked sialosides on the surface of epithelial cells in our upper and lower airway tracks. Using the glycan array, we can further understand the nature and structure of the internal glycans recognized by different hemagglutinin subtypes and define their binding specificity. We also developed the more homogeneous aluminum oxide-coated glycan array to profile some of the broadly neutralizing antibodies isolated from HIV positive patients who did not develop into AIDS [9]. These antibodies are known to target some unusual N-glycans on gp120 and neutralize over 70% of the more than 2000 HIV variants known to date. As an example, the unusual hybrid-type glycan as shown in Figure 4 was identified as a nanomolar ligand for the antibody PG9 [10] and work is in progress to use this unusual glycan as epitope for vaccine design, hopefully to elicit PG9-like antibody against HIV.

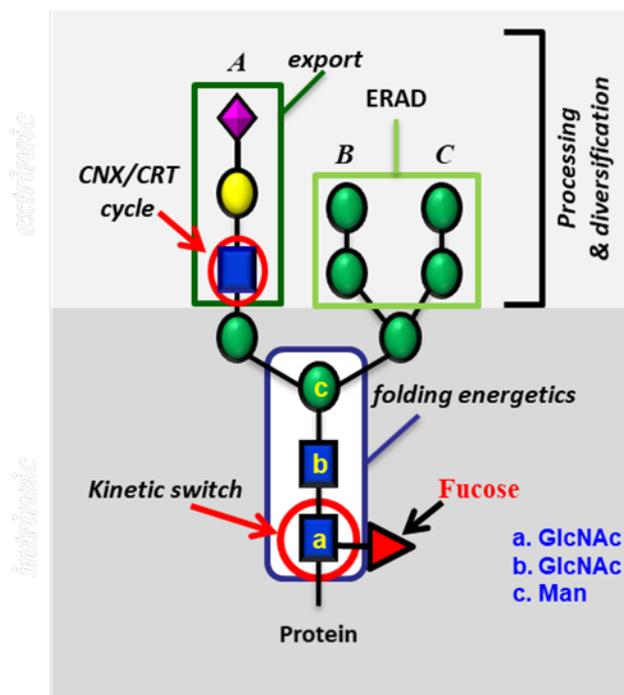## 4. Impact of Glycosylation on Protein Folding

Figure 5. Impact of glycosylation on protein folding and stability using a model protein (hCD2ad) without proline and disulfide bond to interfere. The core trisaccharide is essential for protein folding and stability. GlcNAc (a) speeds up the folding by 0.8 kcal/mol and stabilizes the folded structure by 2 kcal/mol. Additional stabilization (1.1 kcal/mol) is from Man-GlcNAc (c-b). The rest of glycans are involved in function.

To understand how glycosylation affects protein folding, we first studied the O-glycosylation where N-acetylglucosamine is attached to the Ser or Thr residue. We then studied the N-glycosylation and used hCD2ad as a model (Figure 5) because it does not have the proline residue or disulfide bond to interfere with folding. We found that the first sugar attached to the Asn residue, i.e., GlcNAc, is the most important, as it speeds up the protein folding by 0.8 kcal/mol and stabilized the folded structure by 2 kcal/mol, and another one kcal is from the second and the third sugars together. These three sugars constitute the core trisaccharide which is highly conserved from yeast to human, and the rest of sugars are not associated with folding but may involve in different recognition events [11,12].

To further understand the origin of the 2 kcal contributed by the GlcNAc residue to stabilize the protein, we performed an NMR study using another model without proline or disulfide. It showed that the GlcNAc residue has a hydrophobic interaction with the nearby Phe group that accounts for about one kcal/mol, and this interaction changes the orientation of Phe to trigger the following hydrophobic interactions in the beta sheet region, very much like a domino effect that accounts for another one kcal/mol.

5. Low-Sugar Universal Vaccines Against Influenza and SARS-CoV-2

This finding has led to the development of universal vaccines against influenza and SARS-CoV-2 through glyco-engineering. Both human viruses are RNA viruses that are easy to mutate. In addition, the virus surface immunogen, either hemagglutinin or spike (S) protein, is highly glycosylated to shield the conserved epitopes from immune response. Our proposition is to remove the shielded glycans on S protein to better expose the conserved epitopes and use the low-sugar S protein as vaccine.

The SARS-Cov-2 spike protein is a trimer and each monomer contains about 1300 amino acids, 22 N-glycosites and 2 O-glycosites. There are more than 10 million S protein sequences reported by GISAID today though most of them have insignificant mutation rate. Currently, more than 85% of confirmed cases are infected by omicron BA.2 subvariant (March 20, 2022) while 3 months ago, the delta variant accounted for more than 98% of the confirmed cases. Although current vaccines have proven highly effective against severe infection, breakthrough infection occurs relatively often, suggesting the need for a broadly protective vaccine.

We carried out the sequence alignment of more than 6 million S proteins from GISAID and identified 12 highly conserved epitopes, 7 in the receptor binding domain (RBD) and 5 in the stem domain (HR2); and 10 of these 12 epitopes are shielded by glycans.
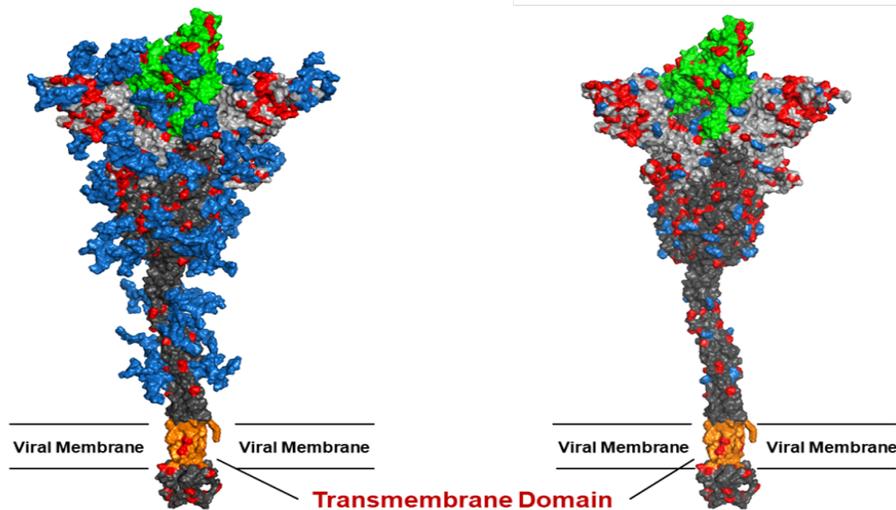
Figure 6. (Left) Fully Glycosylated Spike Protein; (Right) Mono-GlcNAc Decorated (or Low-Sugar) Spike Protein. Green RBD; gray S1; dark gray S2; blue glycans; red mutation sites (mutation rate >0.05%).

We found that when the N-glycans of the wild-type spike protein are trimmed down to a mono-GlcNAc decorated glycoform (Figure 6), it induced enhanced and broadly protective antibody, CD4 and CD8 T-cell responses against all variants of concern, including alpha, beta, gamma, delta and omicron variants (Figure 7).
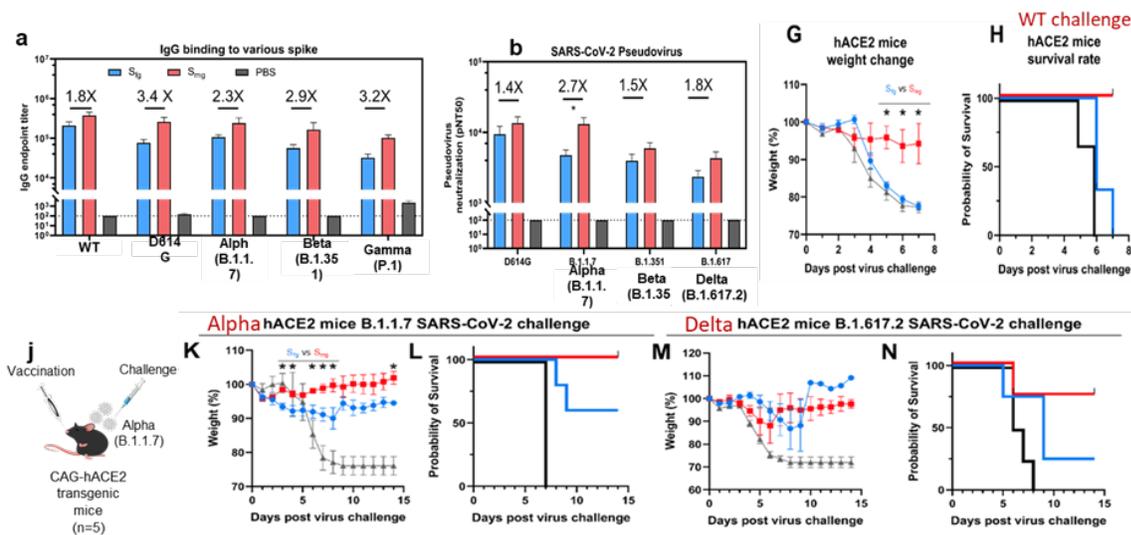

Figure 7. Mono-GlcNAc decorated (or low-sugar) spike vaccine exhibited enhanced binding, neutralization breadth and potency against SARS-CoV-2 and variants.

We also found that vaccination in mice with the mRNA of S protein with deletion of specific glycosites, especially the N-glycosites in S2, induced more broadly protective antibody and CD8 T-cell responses against the variants of concern compared to the unmodified mRNA vaccine.

Overall, compared to the fully glyco-sylated S protein, the mono-GlcNAc decorated (or low-sugar) S protein vaccine generated stronger antibody and CD8 T-cell responses to S protein and stronger protective activity against variants of concern as shown in the neutralization and challenge study [13].

Using the single B cell technology, we isolated the splenocytes of mice immunized with the low-sugar S vaccine and used the fully glycosylated S protein to sort B cells for analysis, and we have identified several broadly neutral-izing antibodies, and one of these anti-bodies recognized a conserved region in RBD

(417-482) with picomolar affinity and neutralizing activity. This antibody could not be generated from vaccination of the fully glycosylated S protein, probably due to the shielding of N-glycans, especially at N-165. The broadly neutralizing antibody identified from immunization of mono-GlcNAc décor-ated S further demonstrated that removal of shielded glycans from S protein is an effective strategy for development of a broadly protective vaccine against SARS-CoV-2 and variants.

The development of low-sugar S protein as a broadly protective vaccine is based on the strategy used in our previous development of universal vaccines against influenza virus. Like S protein, the cell surface hemagglutinin (HA) of influenza is also a trimer and each mono-mer contains 6 N-glycosites. We also found that the regions covered by glycans are highly conserved. We then treated the H5N1 HA protein with endoglycosidase H to make low-sugar HA (Figure 8).
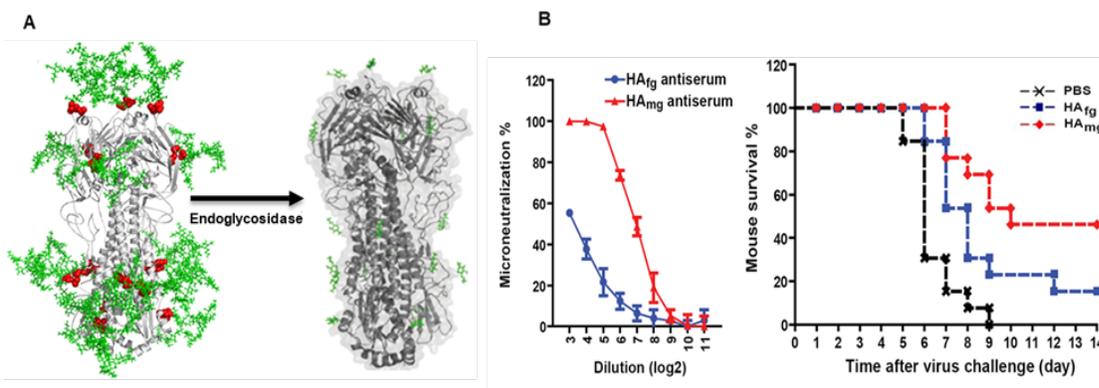

Figure 8. Development of broadly protective influenza vaccine using trimeric hemagglutinin with all 18 N-glycans trimmed down to the mono-GlcNAc state as immunogen. Such a low-sugar vaccine induced more anti-bodies, and CD4 (as well as CD8) T-cell responses to the conserved stem of hemagglutinin showed improved neutralization in the cell-based assay and better protection in the challenge study.

Compared to the fully glycosylated HA, immunization of the low-sugar HA in mice generated stronger antibody and T-cell responses against H5N1 and subtypes with cross reaction against H1N1, as shown in the microneutral-ization and challenge study. Most recently, we have constructed a chimeric HA vaccine with a consensus H5 sequence as head and a consensus H1 as stem and trim the glycans of the chimeric HA to the low-sugar state. Immunization of this chimeric HA vac-cine generated a strong immune response against various H1 and H5 subtypes, and neutralized H3, H7 and H9 subtypes [14]. This chimeric HA vaccine is also highly protective against various H1N1 and H5N1 subtypes in the challenge study.

To further improve the efficacy of antibody against influenza virus, we modified the glycan of antibody FI6 which is active against most type A influenza viruses. It is known that the glycosylation of antibody at Asn-297 will affect its binding to Fc receptors on immune cells and subsequently the effector functions such as antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cellular phargo-cytosis (ADCP), complement dependent cytotoxicity (CDC) and vaccinal effect. We first trim the antibody glycan to GlcNAc followed by enzymatic trans-glycosylation to prepare a series of homo-geneous antibodies with well-defined glycan in Fc and measure their interaction with FcγIIIA receptor *in vitro* with biacore assay. This interaction is a measure of ADCC and through analysis of various homogeneous glycoforms, we found that the 2,6-siayl complex-type biantennary glycan (SCT) is the best glycan for FcγIIIA binding and is also the best for ADCC as shown in the cell-based and challenge study in mice (Figure 9).
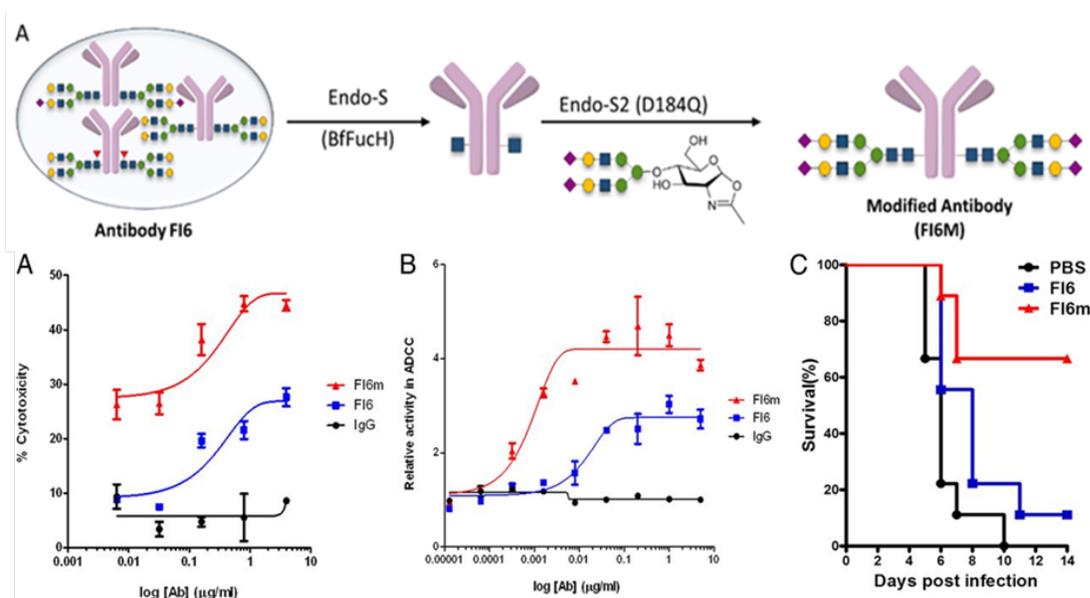

Figure 9. Optimization of broadly neutralizing anti-Influenza antibody Fl6 through glycoengineering of the Fc glycan to SCT to improve the cytotoxicity, ADCC and survival rate.

So, any therapeutic antibody can be optimized through glycoengineering as shown here. Recently, we solved the structure of the Fc moiety containing the SCT glycan, showing that the glycan is more attached to the protein to make a better interaction with the receptor FcγIIIA (Figure 10).
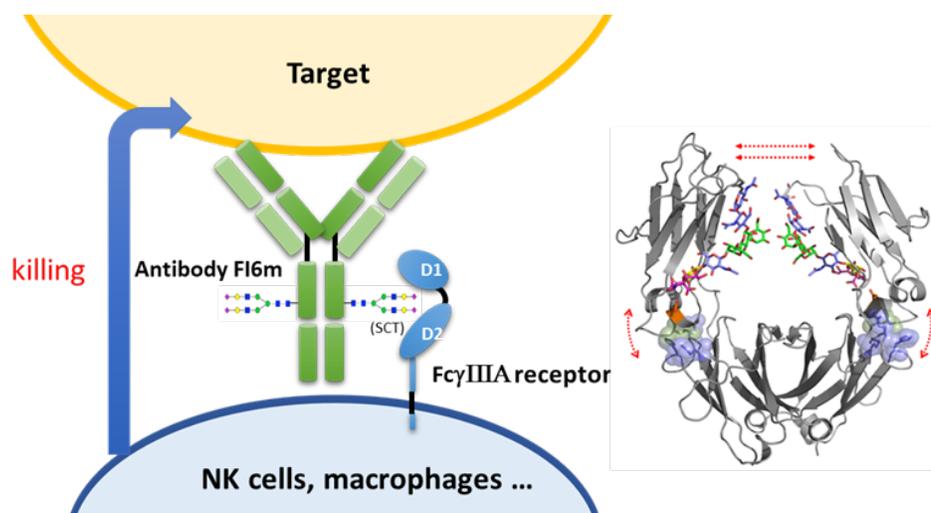

Figure 10. (Left) Homogeneous antibody with well-defined Fc-glycan (2,6-sialyl biantennary glycan) optimized for antibody dependent cellular cytotoxicity; (Right) X-ray structure of Fc-SCT showed that SCT is more attached to the antibody backbone to facilitate FcγIIIA-receptor binding.

6. Broadly Effective Carbohydrate Vaccine and Antibody against Cancer

Identifying a non-self protein antigen for cancer vaccine development is a real challenge; because cancer cells are derived from normal cells, it is difficult to find such unique proteins, if they exist. So, we decided to go after unique glycans on the cancer cell surface. The reason is that glycans are smaller in size compared to proteins and they are assembled by multiple enzymes. If the assembly lines go wrong, the glycan structure will change, and we may have a better chance to find cancer-specific glycan structures that are not present in normal cells. Since cancer cells are usually highly fucosylated and sialylated, we used the alkyne-containing glycosylation probes to monitor the fucosylation and sialylation process in normal and cancer cells (Figure 11, [15]).
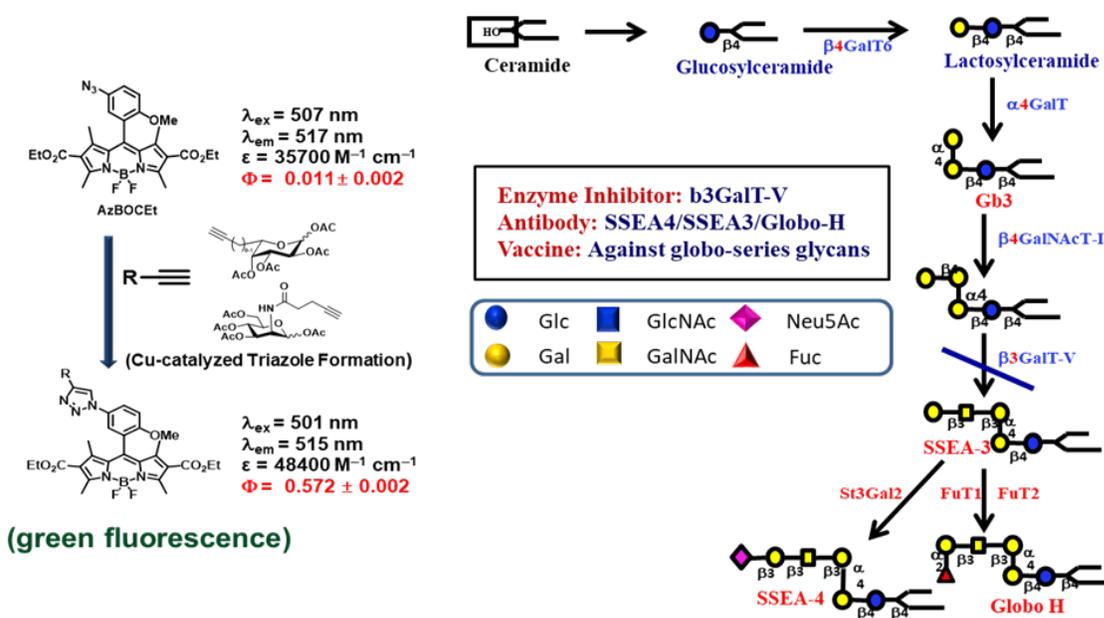
Figure 11. Identification of cancer cell-specific glycans using glycosylation probes of fucose and sialic acid containing an alkyne group. After incorporation of the probes into glycoconjugates, the cells are fixed and reacted with azido-BODIPY to form a green fluorescent triazole for LC-MS analysis to confirm that the globo-series glycans are cancer specific.

We first feed the cells with a probe to let the sugar probe incorporate into glycoproteins or glycolipids. We then fix the cell and mixed the fixed cell with BODIPY-azide for click reaction to form the fluorescent triazole adduct, that is then separated by gel and the glycan structure is determined by LC/MS analysis.

Using these probes, the globo-series glycolipids are often identified from different kinds of cancer cells, and they are absent in normal cells. These glycolipids include SSEA3, Globo-H and SSEA4, which are originally expressed on embryonic stem cells but disappear after differentiation. For some reason, they show up again in cancer cells. So far, we found that at least 15 types of cancer have expressed globo-series glycolipids, and the key enzyme that drives the expression of these glycolipids is a galactosyltransferase called B3GalT5, which catalyzes the galactosylation of Gb4 to SSEA3 and is then converted to SSEA4 and Globo-H. Using the pull-down experiment, we found that CAV1 and FAK interact with SSEA4 and Globo-H, respectively, and then AKT and RIP interact with CAV1 and FAK to form a complex. Knockdown of B3GALT5 in breast cancer cell stopped the synthesis of globo-series glycolipids and caused dissociation of the protein complex; RIP then interacts with FADD and activates caspase leading to cell apoptosis. However, the knockdown has no effect on normal cells. We also found that SSEA4 is increasingly expressed in tyrosine kinase inhibitor (TKI)-resistant non-small cell lung cancer (NSCLC) with T790M mutation; without T790M mutation, there is no significant change in the SSEA4 level. This study suggests that the TKI-resistant non-small lung cancer

cells with T790M mutation have high expression of SSEA4 which may be a good target for therapeutic development.

Previous study showed that SSEA4 is a good target for antibody therapy against pancreatic cancer that has high expression of SSEA4 [16]. The chimeric SSEA4 antibody chMC813-70 was engineered to have the SCT glycan or the more stable 3FSCT derivative in the Fc region to enhance ADCC (Figure 12).
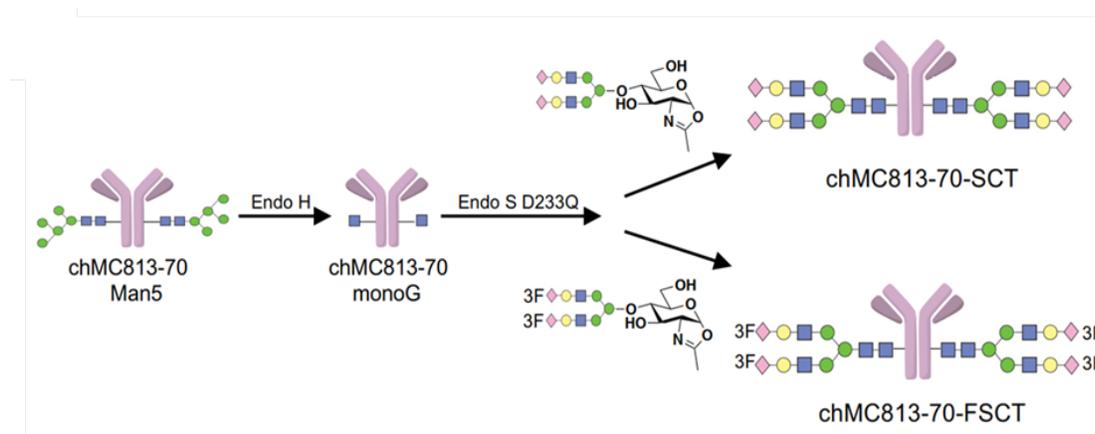

Figure 12. Homogeneous Anti-SSEA-4 Antibodies with Maximal ADCC Activity Through Fc Glycoengineering

As shown in the cell-based assay, the homogeneous antibody is more effective than the unmodified antibody against the three pancreatic cancer cell lines expressing different levels of SSEA4, and it is more effective against HPAC cells with higher expression of SSEA4. The homogeneous antibody can also be used to isolate a subpopulation of NK cells enriched with FcγIIIA receptor for ADCC. The isolated NK cells indeed showed 6-fold more effective than the unseparated NK cells in killing the target cells.

This work demonstrated that the Fc glycan of an antibody can be engineered to optimize its interaction with a specific Fc receptor to improve the corresponding specific effector function, and the homogeneous antibody can also be used to isolate a specific population of immune cells enriched with specific Fc receptor for cell-based therapy.

Globo-H was also used to develop cancer vaccines through conjugation to KLH or DT, and adjuvanted with QS21 or C34. The Globo-H/KLH/QS21 vaccine was developed by the Danishefsky group and synthesized by our method and now is in phase 3 global trial for the treatment of Globo-H positive triple negative breast (TNB) cancer patients. The Globo-H/DT/C34 vaccine developed by us with C34 adjuvant designed to increase IgG response is in phase 2 trial for the treatment of different cancers with Globo-H expression. It is interesting that both vaccines can induce immune responses to target not only Globo-H, but also the other two globo-series glycans, namely SSEA3 and SSEA4 [17].

To understand why the GH-DT vaccine induces immune response against the three globo-series glycans, we mixed dendritic cells with the vaccine and investigated how the vaccine is processed

and presented using antibody staining and LC-MS analysis. We found that Globo-H is first processed by fucosidase 1 in early endosome to generate SSEA3, which is the common epitope of these three globo-series glycans. We also found that the antigen presentation is mainly through MHCII to produce anti-Globo-H antibody and anti-SSEA3 antibody that cross-reacted with SSEA4. When SSEA4-DT was used as a vaccine, the glycan was very slowly processed, and the antibody generated is mainly specific for SSEA4, while in the case of SSEA3-DT, it was quickly processed to the less immunogenic glycans Gb4, Gb3 and Gb2 to induce a weak and non-specific antibody response.

7. Conclusion

In summary, glycosylation is an important biological modification for normal and disease processes. Understanding the role of glycans and the associated glycolsylation reaction in biology will facilitate the development of better strategies for treatment or prevention of diseases associated with carbohydrate recognition. Over the years, advances in carbohydrate chemistry and glycobiology have made important contributions to help reach this goal.

8. Acknowledgment

9. References

1. Krasnova L, Wong CH. *J. Am. Chem. Soc.*, 2019, 141, 9, 3735–3754.

2. Zhang Z, Ollmann IR, Ye XS, Wischnat R, Baasov T, Wong C. *J. Am. Chem. Soc.*, 1999, 121, 734-53.

3. Cheng CW, Zhou Y, Pan WH, Dey S, Wu CY, Hsu WL, Wong CH. *Nat. Commun.*, 2018, 9, 5202.

4. Chang CW, Lin MH, Chan CK, Su KY, Wu CH, Lo WC, Lam S, Cheng YT, Liao PH, Wong CH, Wang CC. *Angew. Chem. Int. Ed.*, 2021, 60, 12413-12423.

5. Dey S, Lo HJ, Wong CH. *J. Am. Chem. Soc.*, 2019, 141, 26, 10309–10314.

6. Pardo-Vargas A, Delbianco M, Seeberger PH. *Curr. Opin. Chem. Biol.*, 2018, 46, 48–55.

7. Tsai TI, Lee HY, Chang SH, Wang CH, Tu YC, Lin YC, Hwang DR, Wu CY, Wong

CH. *J. Am. Chem. Soc.*, 2013, 135, 14831-14839.

8. Liao HY, Hsu CH, Wang SC, Liang CH, Yen HY, Su CY, Jan JT, Ren CT, Cheng TJ, Wu CY, Wong CH. *J. Am. Chem. Soc.*, 2010, 132, 14849-14856.

9. Shivatare SS, Chang SH, Tsai TI, Tseng SY, Lin YS, Cheng YY, Ren CT, Pawar S, Tsai CS, Shi HW, Zeng YF, Liang CH, Kwong PD, Burton DR, Wu CY, Wong CH. *Nat. Chem.*, 2016, 8, 338-346.

10. Shivatare VS, Shivatare SS, Lee CC, Liang CH, Liao KS, Cheng YY, Saidachary G, Wu CY, Lin NH, Kwong PD, Burton D, Wu CY, Wong CH. *J. Am. Chem. Soc.*, 2018, 140, 5202-5210.

11. Hanson SR, Culyba EK, Hsu TL, Wong CH, Kelly JW, Powers ET. *Proc. Nat. Acad. Sci. U.S.A.*, 2009, 106, 3131-3136.

12. Culyba EK, Price JL, Hanson SR, Dhar A, Wong CH, Gruebele M, Powers ET, Kelly JW. *Science*, 2011, 331, 571-575.

13. Huang HY, Liao HY, Chen X, Cheng CW, Wang SW, Shahed-Al-Mahmud M, Chen TH, Lo JM, Liu YM, Ma HH, Chang, YH, Tsai CY, Huang PY, Chang SY, Chao TL, Kao HC, Tsai YM, Chen YH, Chen CY, Lee KC, Wu CY, Jan JT, Lin KI, Cheng TJR, Ma C, Wong CH *Sci. Transl. Med.*, 2022, 14, *eabm0899.*

14. Chen JR, Yu YH, Tseng YC, Chiang WL, Chiang MF, Ko YA, Chiu YK, Ma SH, Wu CY, Jan JT, Lin KI, Ma C, Wong CH. *Proc. Nat. Acad. Sci. U.S.A.*, 2014, 111, 2476-2481.

15. Shie JJ, Liu YC, Lee YM, Lim C, Fang JM, Wong CH. *J. Am. Chem. Soc.*, 2014, 136, 9953-9961.

16. Lin CW, Wang YJ, Lai TY, Hsu TL, Han SY, Wu HC, Shen CN, Dang V, Chen MW, Chen LB, Wong CH. *Proc. Nat. Acad. Sci. U.S.A.*, 2021, 118, e2114774118.

17. Samuel J. Danishefsky SJ, Shue YK, Chang MN, Wong CH. *Acc. Chem. Res.*, 2015, 48, 3, 643–652.