nature portfolio

Peer Review File

Mammalian Cell-Based Production of Glycans, Glycopeptides and Glycomodules - Expanding the Glycoscience Toolbox

Corresponding Author: Dr Yoshiki Narimatsu

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have responded to many of the original critique points in the manuscript but in many cases they believe the reviewers have misunderstood the point or nature of their submission. In any case, this review here will not go back point-by-point over the issues. In this reviewers opinion, the study does clearly provide additional tools and expanded approaches to help identify the various types of glycan species generated by cells using these specific unique biosynthetic platforms. However, whether these platforms provide close similarities to the actual glycans the cells produce is not really clear, and their relative abundance of such materials also is not clear. If one is using engineered cell lines that have highly limited diversities of materials, this approach might be comparable or used in parallel to just growing more cells for harvesting to obtain the limited diversity of glycans, but it is clear that the 'recombinant' strategy proposed here could be superior. Also, as noted in earlier points, the products generated by these approaches are heterogeneous (are incompletely synthesized?), some more than others, and would require extensive purification for characterization. Also, the quantities of material generated by the approaches put forward here are not clear to this reviewer, and while it might be scalable it is not clear to what degree this is reasonably accomplished. In addition, using a limited set of protein carriers for example, might actually bias the nature of glycans produced on the individual platforms. Overall, it is clear that the studies provide another, specialized set of techniques to prepare glycans for use, it is not clear that this methods paper is at present providing sufficiently novel sets of approaches beyond the use of natural glycans produced by the cells on their endogenous glycoproteins.

Reviewer #2

(Remarks to the Author)
In previous work, the authors produced mucin repeat proteins that were surface displayed, and in this work they produce them as secreted proteins. They show that secreted glycoproteins can be stripped to produce glycans in mg scales (with same glycans also displayed on proteins and peptides). A similar approach is also followed for N-glycans and GAGs. So, now we can produce isolated glycans and also the same glycan on a repeat scaffold. This may represent an alternate approach to produce glycans in laboratories lacking chemoenzymatic synthesis capabilities. Another strength of the paper is the volume of data that is on display.

On the flip side, the work to some degree feels incremental as the technology for CRISPR KO/KI exists and is well established for the last decade. Some of the isolated glycans are not homogeneous due to incomplete reactions within cells, and it is argued that this can be improved in the future as necessary, or down-stream separation is possible. There are many functional studies in the submission focused on demonstrating the utility of the produced glycans/glycoproteins and they are all well-performed, but no strong example of novel biology.

Point-by-point list of reviewer comments and responses/actions:

Reviewer #1 (Remarks to the Author):
Comment #1: The authors here have attempted to address a need in the field, namely the generation of glycoconjugates in a stable, recombinant format for analyzing the molecular interactions of the glycans in the context of a protein linkage. The approach is a recombinant chimeric library of proteins designed to generate O-, N-, and GAG-linked glycans. The authors have shown that this strategy does allow the production of recombinant glycoconjugates with selected features using engineered cell lines, which they produced earlier for studies of glycan recognition. There are some features of this approach that are interesting, but the major biological insights using this approach are very limited, and it is not clear that they provide information beyond that of using the engineered cell lines themselves.

Response #1: This summary may reflect a conceptional misunderstanding of what our study is about – the Glycocarrier technology is not an extension of our previously reported cell-based glycan arrays that champions the concept of presenting glycans in the right concept of their glycoconjugate nature and the cell surface (Narimatsu et al. Mol Cell 2019, PMID: 31227230). Instead, the Glycocarrier technology provides a simple, scalable, and sustainable process for producing glycans in many formats - and notably without their natural context, e.g. as glycolipids or specific proteins. We believe that the cell-based platform holds promise to eliminate most of the chemical and chemoenzymatic processes with their supply, cost, and environmental challenges, and provide access to individual glycans in amounts and grades suitable for any biomedical use – just like how we today produce almost all recombinant glycoprotein therapeutics in CHO cells! We fully appreciate that our manuscript may have lacked clarity in presenting this concept and we have now extensively revised the text.

Action #1: We have extensive revised the manuscript to mitigate misunderstandings and stress the aims of our study.

Some of the major difficulties in this approach are below.
Query #1: The products have a degree of heterogeneity in the glycan attached, and this could complicate analyses, as well as a degree of heterogeneity in site utilization, especially for O-linked GalNAc type products. It is unclear as to the total heterogeneity as the authors state that "… MS was only possible after removal of sialic acid (Sia) residues by enzymatic desialylation."

Response #1: We regret our use of language ("heterogeneity" instead of "incomplete synthesis") in the text. Heterogeneity may not be a major difficulty of cell-based production of glycans. Clearly, all types of syntheses, chemical and/or enzymatic, involves steps that reach different levels of completion, and there are simple ways to separate and isolate final homogenous products. While our wording was unfortunate, we clearly demonstrate that many different types of glycans can be produced in homogeneous and/or quite simple mixtures that (1) could be optimized by further genetic engineering for desirable glycans and (2) used to purify any desirable glycan to homogeneity. We have chosen not to include experiments and illustrations to support this, but we have data and if deemed needed we are happy to include an example as a new Supplementary Figure.

We must disagree with the comment on MS analysis of sialylated glycans. We provide analysis of sialylation with isolated N- and O-glycans and short glycopeptides, and it is only our analysis of the intact

glycomodules with 10/18 sialylated glycans that required removal of sialic acids. We hope that the reviewer can appreciate that even analysis of these large complex glycomodules without sialic acids by native/intact MS is already close to state-of-the-art in the field.

We have clarified the text accordingly. We would like to stress that this is a method paper, and we do not intend to or perceive to have produced a comprehensive library of glycans. We provide the framework and validation for how defined glycans can be produced in cell by the Glycocarrier technology. We also want to stress that this method is geared to produce individual desirable glycans in amounts often impossible or impractical by chemical and chemoenzymatic processes.

Action #1: We have clarified the text throughout, and Figures 2-4 and Supplementary Figures 5-6 have been revised.

Query #2: The relatively simple glycans produced by these approaches are not expanding the available glycans for the community, rather displaying them on a single peptide format. Such a display has no particular biological presentation in the context of a 3-4 dimensional glycoprotein, but as the authors state, it is a scaffold. As such, it provides more of a 'multivalent' presentation on a single molecule that standard type glycan arrays.

Response #2: We refer to response to query#1. The Glycocarriers are only carriers of glycans. Glycocarriers are designed to allow isolation and scalable production of glycans in many formats to eliminate the challenges with chemical and chemoenzymatic synthesis. Our previously described cell-based glycan array (Narimatsu et al. Mol Cell 2019, PMID: 31227230) provides what the reviewer is looking for in this comment. Mammalian cells do generally not produce free glycans and the Glycocarriers are only used as carriers to produce glycans. The objective was to enable use of these glycans in many formats, as free released glycans, as short glycopeptides without protein identity, and on glycomodules in multivalent presentation again without protein identity.

The misunderstanding here is likely caused by our selection of studies of clusters of GalNAc-type O-glycans, which is a unique feature of this type of protein glycosylation. Our intent was to demonstrate that simple clusters O-glycans without protein identity can also be produced and studied, but this clearly has distracted from the point that the method is to produce glycans without their natural protein context.

We have revised the text to stress that we are not trying to supply glycans for glycan arrays – this need has largely been overcome by our cell-based glycan array technology – but we are trying to supply glycans in quantities and forms for wider uses. We hope the reviewer will appreciate that most of the glycans produced and used for glycan arrays are in great scarcity and consumed to maintain these arrays.

Action #2: We have clarified the text throughout.

Query #3: The products in regard to complex N-glycans consisted of mixtures including bi-, tri-, and tetra-antennary N-glycans, with or without fucose and sialic acid.

Response #3: We previously demonstrated how far genetic glycoengineering has come in terms of dissecting and controlling N-glycan biosynthesis (please see our papers: Yang et al. Nat Biotechnol 2015, PMID: 26192319; Tian et al. Nat Commun 2019, PMID: 31040271). In the present study, we present data showing that we can produce N-Glycocarriers that obtain (1) stoichiometric attachment of N-glycans at all 10 glycosites and (2) that we can control distinct features to produce homogenous N-glycans (high Man and biantennary) or N-glycans e.g. with/without core fucose or sialic acids (**Figure 3**). Since this is not a resource paper, we consider the chosen examples as sufficiently representative to validate the strategy. However, as discussed in the text further engineering can obviously be used to obtain homogenous bi and/or triantennary glycans of most N-glycan structures. The order of N-glycoengineering examples presented in Figure 3 may have been misleading, and we therefore include a new Figure 3 to improve clarity. Note, that we are not capable of driving complete synthesis of tetraantennary N-glycans (>70%), but simple purification steps could be employed here.

Action #3: We have clarified the text throughout and included a revised Figure 3.

Query #4: For the GAGs, the authors observe the products with complex GAG structures are “high MW smears”.

Response #4: This may be an oversight as we did provide the state-of-the-art disaccharide analysis of the produced GAGs in **Suppl. Figures 10** and **11**. This type of data is unfortunately too large and complex to present in a main Figure. We did not perform extensive engineering of GAGs primarily because these are still difficult to analyse, as also revealed by the Reviewers comment on “high MW smears”. We have already demonstrated that the GAG biosynthesis can be dissected to produce GAGs with/without particular features (Chen et al. Nature Methods 2018, PMID: 30104636; Karlsson et al. Sci Adv 2021, PMID: 34936441) just like other types of glycosylation. In this method paper, we clearly demonstrate that we can (1) produce Glycocarriers with GAG chains attached and (2) engineer the feature of GAG chains (HS or CS) on these. Clearly, for production of a comprehensive library of GAG chains our continuously expanding large library of engineered cells (Chen et al. Nature Methods 2018, PMID: 30104636) could be employed, but this would be more suitable for a resource paper. The key novelty of our method is the identification of a short peptide sequence motif that obtains stoichiometric attachment of GAGs. Please note that we do discuss the challenge left with driving efficient HS synthesis, so this is not all straightforward but conceptually conceivable.

Action #4: We have clarified the text in the GAG section to stress this.

Query #5: Perhaps the authors could separate the Glycocarrier products to produce more homogeneous material for studying biological interactions.

Response #5: We agree that it is fairly straigthforward, when/if needed, to use standard separation protocols to obtain more homogenous glycans, especially if combined with further genetic engineering. However, we clearly illustrate that glycans can be produced in relatively pure states, and we do not believe that it is not within the scope of the study to produce comprehensive libraries of glycans.

We may also further point out that while homogenous glycans clearly are an advantage with printed glycan arrays, the genetic engineering approach and cell-based production also provide the options to produce and use more heterogenous mixtures of glycans (or glycoconjugates displayed on cells or secreted molecules) with/without distinct features (e.g. fucose, sialic acids, sulfate groups). This provides unique opportunities to take advantage of the step-by-step biosynthesis of glycans to narrow down biological interactions with different features of glycans, for example specific interactions with a sialylated epitope on N-glycans can follow dissection of sialic acid linkage, core LacNAc/LacDiNAc dissacharides or polysaccharides, N-glycan branching, and core fucose. However, this is best performed with the cell-based glycan array, as previously demonstrated (Narimatsu et al. Mol Cell 2019, PMID: 31227230), but the present method allows for production of such glycans in diverse formats with ease in quantities to support structural and other studies requiring greater amounts.

Action #5: None, apart from the general text editing already referred to.

Query #6: The authors use a very limited 'peptide' scaffolds, e.g., EPO, Serglycin, etc. Do the scaffolds themselves give different presentations that have different biological recognition? Are the scaffold-based presentations superior in some ways compared to natural expression on cell surfaces or cellular glycoproteins from engineered cells?

Response #6: This comment reemphasizes our concern that the concept for our method to produce glycans was not clearly presented. Our aim was to identify only one single carrier design for each type of glycans, and these designs should be inert to and not impact the glycosylation outcome in cells as well as the biological context. Thus, the method provides quite the opposite of what the Reviewer is querying and requesting.

Action #6: None, beyond Actions #1-4

Query #7: One advantage of having such engineered glycopeptide products is their use as chemical carriers for cell-specific delivery as well as their use in semi-synthetic glycoprotein production. The potential of these approaches was demonstrated in supplementary data, but it is not clear that this approach, given the issues about the heterogeneity of the products above, is superior to existing strategies of generating neoglycoproteins and chemo-synthetic glycopeptides. This aspect of the work is key to demonstrating its importance.

Response #7: We fully agree that demonstrating the value of the diverse formats of glycans the Glycocarriers provide is most important. We used a graphic summary Figure (**Figure. 5**) to illustrate this and with the experimental support provided in **Supplementary Figures 12-18**, as these are too complex

to present in simple formats as main figures. We are happy to move one or more of these into main figures of course.

However, the comment that the method is not superior to existing ways of producing glycans may rely on a misunderstanding. We addressed heterogeneity and glycoconjugate context in previous Response/Actions above, and further we need to clarify that once stable expression of a Glycocarrier in an appropriately glycoengineered cell for a particular glycan is established – this cell provides (1) unlimited access to the glycan in any quantity and (2) facile access to multiple formats of this glycan from a single Glycocarrier for use in diverse applications. These aspects are not provided by current chemical and chemoenzymatic processes.

Action #7: We have clarified the text throughout to avoid misunderstanding.

Query #8: The authors analyzed only a few proteins for their recognition of these engineered glycopeptide products, and provided no specific evidence that the insights from these materials are superior to those using the engineered cell lines. This aspect of the work is key to demonstrating its importance.

Response #8: We refer to our previous Responses and Actions. The cell-based glycan arrays are indeed available to dissect specificities in interactions (Narimatsu et al. Mol Cell 2019, PMID: 31227230), however, it is important to appreciate that many interactions with glycans are not guided by specific presentation of these on distinct glycoproteins or other glycoconjugates. But perhaps more importantly, once binding specificities have been characterized there is a need to have access to glycans in quantities suitable for further studies and biomedical applications. The aim with our study is to provide a new method to produce glycans in cells without particular context (i.e., except clusters and multivalency) with an approach that is much simpler, cost-effective, and sustainable compared to traditional synthetic methods. We do provide several examples (**Figure. 5** and **Suppl. Fig. 12-18**) to demonstrate and validate what the Gycocarriers can support.

Reviewer #2 (Remarks to the Author):

Comment #1: The paper comes from a leading research group that presents a tour de force approach to synthesize designer glycopeptides and glycans using recombinant engineered cells. They propose that this approach is scalable for the production of glycoconjugates that can be used for diverse applications. From a technical perspective, the study is well done, with detailed mass spectrometry characterization of produced products, and indeed the products are as expected based on the cell type they are produced in. Another strength of the paper is the ability to produce diverse (O-, N- and GAG) glycan-types on designer peptide backbones. Data analysis and statistics are appropriate.

Response #1: We appreciate the praise and recognition of the efforts used to derive this method. We only want to point out that the method is not to produce glycoconjugates, but glycans in different formats that also include short glycopeptides and multimeric glycomodules without protein identity. Our manuscript must have been misleading in this regard and we have extensively revised the text.

Action #1: The revised manuscript has been extensively modified to clarify this.

Major comments:

Query #1: If the goal is to produce defined complex sugars using genetically engineered cells, the target glycans may be obtained by simply stripping the carbohydrates from knock-out/knock-in cell surface, rather than taking trouble to produce specific secreted glycopeptides. The yield if this were done would likely be orders-of-magnitude higher than the approach presented in this paper, and then they could be separated for downstream applications. This is the approach taken by Cummings and colleagues for producing glycan microarrays. Thus, there is a need to better justify this approach that relies on secreted glycoproteins—in what scenario would the presence of the underlying peptide scaffold be critical for generating new biological understanding.

Response #1: We disagree – stripping glycans from cell lysates suffers from major challenges even if cells are engineered for homogenous glycans due to intracellular biosynthetic intermediates and difficulties with isolation as glycans have to be released before purification. Most importantly though, this does not provide yields close to what can be obtained with a secreted product using standard bioprocessing processes, and today not a single pharmaceutical product is produced from the cell biomass of mammalian cells and all biologics are exclusive produced from secreted glycoproteins. The Cummings strategy is another elegant way to circumvent chemical synthesis, but this process is designed to probe glycosylation capacities of cells and to screen complex mixtures of glycans each in minute amounts for use in printed glycan arrays and structural analysis. Thus, not suitable for production of individual glycans in any significant quantities. The use of glycosides to prime synthesis of oligosaccharides in cells has only been shown for O-glycans and GAGs, and considerable challenges with degree of glycosylation and hence yields are found with both (Kudelka et al. Nat Methods 2016, PMID: 26619014; Chen et al. Nature Methods 2018, PMID: 30104636). Finally, stripping glycans of cells or any animal tissue source clearly does not provide the versatile platform provided here with defined glycans in different formats.

Action #1: None, apart from the general text editing already referred to.

Query #2: Among the carbohydrates produced, only a few of the cells can produce homogenous glycans, e.g. COSMC-KO, MGAT1-KO and B4GALT7-KO. When more complex glycans are synthesized in the target cells, the products are commonly heterogeneous. This is noted in Figure 2 for O-glycans, particularly in the case of B3GNT6-KI, FUT2-KI and GAL3ST4-KI. This is also noted for N-glycans in Fig. 3C and Supplementary Fig 8. Indeed, groups of related N-glycans are produced but there does not seem to be any good way to control the branching pattern of these carbohydrates. The same is true for the GAGs that contain varying types of disaccharide units and heterogenous HS/CS masses. These results somewhat run contrary to the stated motivation of the paper, to produce defined glycan products for structure-function studies. Wouldn't this be a limitation since it is not possible to produce arbitrary glycoconjugates, by design, using this method.

Response #2: We refer to our Responses and Actions to Reviewer #1. Most glycans can be produced close to homogenous by use of combinatorial genetic engineering, while further postproduction purification may be needed with some complex structures like for example tetraantennary N-glycans. However, the branching status of N-glycans can be controlled to a large extend with homogenous biantennary glycans illustrated as the prime example (**Figure 3c**). Triantennary and tetraantennary N-glycans may contain variable amounts of incomplete converted bi or triantennay glycans and postpurification may be needed in these cases, not dissimilar to the situation with chemical/chemoenzymatic synthesis.

Cell-based produced glycans is indeed very useful for structure-function analysis. Printed glycan arrays and the cell-based glycan array are extremely useful for discovery and dissection of specificities of interactions, but for detailed analysis of interactions with glycans by structural methods (x-ray, CryoEM, etc) access to defined glycans are clearly needed, and this is where the presented method can provide compounds in required quantities and formats.

We are uncertain what the Reviewer refers to regarding arbitrary glycoconjugates, but the presented method is based on Glycocarriers that use a single short peptide motif to prime attachment of glycans. These acceptor sequence motifs were designed and selected to have no or minimal impact on the glycans and their flexible use as short glycopeptides or multimeric glycomodules.

Action #2: None, beyond the previous Actions listed above.

Query #3: The previous work by this group used surface expressed glycoconjugates, and the current work extends the toolbox to produce many of these products in soluble form. There is concern that this would be incremental unless a definitive novel, significant result is presented. The results of binding studies in Supplementary Fig. 7, 13 and 14 measure molecular binding but are unable to distinguish between affinity vs. rebinding vs. avidity effects. The results with Gal3 condensates are interesting but a bit preliminary in nature (Fig. S15). The azido derivatization of glycopeptides is as expected (Fig. S16), and the glycoproteins in Fig. S18 could also be synthesized by other means, for example by simply expressing the protein and glycomodule in a single plasmid. Overall, a number of applications are presented, but how they advance the current science is not clearly evident.

Response #3: We disagree with this comment. We previously used genetic glycoengineering to develop libraries of isogenic cells that display distinct glycan features on the surface in the context of their natural glycoconjugates as a major advance in glycan array technologies. Our new method for sustainable production of glycans relies on a novel Glycocarrier technology and the purpose is entirely different. We do employ genetic engineering of glycosylation capacities of cells and the knowledge developed from our previous studies as design matrix for optimal engineering, but the concept of sustainable production of glycans in multiple formats without protein identity by Glycocarriers is not incremental.

Our use of application examples (**Figure. 5** and **Suppl Figs. 12-18**) made possible by our new method to produce glycans were selected only to illustrate the diversity of opportunities provided. Clearly, other methods can be used to produce the glycans we used in these examples, but a Glycocarrier first of all provides cost-effective and sustainable supply of glycans in different formats. The Reviewer points to our binding studies as incomplete but does not acknowledge that such studies with defined biomolecules (1) have not been done before, (2) chemical synthesis of the multimeric glycomodules would be challenging and costly, and (3) that further studies with the Glycocarrier produced glycans, glycopeptides and glycomodules could answer the questions raised. However, we believe that this is beyond the scope of this method paper. The scope in contrary is to highlight the novel method to access glycans and the opportunities to advance glycoscience this method provides.

Action #3: None, beyond the previous Actions listed above.

Minor comments:
Query #4: Expression of tandem repeat proteins in mammalian cells, has been shown to result in recombination that often results in truncated proteins with varying numbers of tandem repeats. Thus the number of repeats in the output protein may be different from the input plasmid. It would be helpful to clarify that such recombination is absent (or present) in CHO and HEK293 cell experiments performed by the authors. There seem to be data that would allow this evaluation in Supplemental Figures 6 and 7, but there is no discussion of this analysis in the manuscript.

Response 4: We agree that this is a major concern with mucin tandem repeat constructs, but this recombination has for long been solved by use of degenerate sequence variations in synthesis of such constructs as employed with our constructs.

Action #4: None.

Reviewer #3:

Comment #1: The concept of developing defined glycoproteins using gene KO or KI CHO and HEK293 cells was reported by the authors a few years ago (PMID: 26192319, PMID: 30104636, PMID: 31227230, PMID: 35247390). The current study is an extension of their previous work by using glycoengineered cells to produce secreted glycocarriers for presenting O-glycans, N-glycans and GAGs. This technology is a useful approach for producing short, homogeneous glycopeptides to overcome the challenge of obtaining recombinant mammalian glycosyltransferases for in vitro chemoenzymatic glycopeptide synthesis.

Response #1: We need to point out that the method is to produce glycans in diverse forms, not only glycopeptides.

Action #1: None, apart from the general text editing already referred to.

Comment #2: Although it is of great importance to produce homogeneous glycoforms, I'm not quite sure why the authors claim that the applications presented in Figure 5 are only achievable with glycocarriers developed in the current study, especially the examples presented in Figure 5A and 5C.

Response #2: To our knowledge only Glycocarriers can without unreasonable efforts produce the paired samples of free glycans, single glycan glycopeptides, and multimeric glycomodules.

Action #1: The manuscript text has been edited to clarify this throughout.

Specific comments and suggestions:

Query #1: What does "metabolic labeling refer to in Fig. 1a?

Response #1: We simply want to illustrate that the cell-based production of Glycoarriers can be combined with incorporation of unnatural sugar such as azido- or propagyl alkyne Neu5Ac by supplementing these metabolic feeds in the growth media, and we illustrated this concept in **Suppl. Fig. 16** in the revised manuscript.

Action #1: The manuscript text has been edited to clarify this.

Query #2: For "Production of N-glycans by N-Glycocarriers", What is the glycosylation efficiency for longer glycomodules with two or more glycosylation sites? Only N-glycomodules with one glycosylation site was described in the manuscript.

Response #2: This is incorrect. We demonstrate and validate that N-Glycocariers with glycomodules containing up to 10 repeats have stochiometric incorporation of N-glycans (**Fig. 3b**). The same was demonstrated for O-Glycocariers and we only found challenges with GAG-Glycocariers containing >5 repeats.

Action #2: None

Query #3: In the second paragraph of this section, you incorrectly cited Supplementary Fig5, which is the MS analysis of O-glycans not N-glycans. It should be Fig 3c.

Action #3: Fixed, thank you!

Query #4: Could authors present an application of glycosaminoglycans enabled by their new production technique?

Response #4: The GAG-Glycocarriers were indeed used to demonstrate binding specifies of the novel ScFv2 nanobodies to cancer-associated CS epitopes by mass photometry (**Suppl. Fig. 14**). Our paper on these new interesting nanobodies will appear in Nat Commun within the next few weeks.

Action #4: None