



## OPEN ACCESS

## EDITED BY

Jose De La Fuente,  
Spanish National Research Council (CSIC),  
Spain

## REVIEWED BY

Ruben Marrero Diaz De Villegas,  
CONICET Institute of Agrobiotechnology and  
Molecular Biology (IABIMO), Argentina  
Sina Sarsarshahi,  
SUNY Buffalo State University, United States  
Farong Ye,  
Shanghai Jiao Tong University, China

## \*CORRESPONDENCE

Cristina E. Martina

✉ cristina.e.martina@vanderbilt.edu

†These authors have contributed equally to  
this work

RECEIVED 16 October 2025

REVISED 09 December 2025

ACCEPTED 12 December 2025

PUBLISHED 07 January 2026

## CITATION

Araiza G, Morel J, Tran MH, Jiang S, Murray S,  
Miyamoto E, Vogts H, Brown KL, Meiler J and  
Martina CE (2026) Glycan masking in  
immunogen design: computational and  
experimental methods.  
*Front. Immunol.* 16:1726810.  
doi: 10.3389/fimmu.2025.1726810

## COPYRIGHT

© 2026 Araiza, Morel, Tran, Jiang, Murray,  
Miyamoto, Vogts, Brown, Meiler and Martina.  
This is an open-access article distributed under  
the terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Glycan masking in immunogen design: computational and experimental methods

Gustavo Araiza<sup>1,2†</sup>, Josuel Morel<sup>1,3†</sup>, Minh H. Tran<sup>1,4</sup>,  
Shan Jiang<sup>1,4</sup>, Sean Murray<sup>5</sup>, Emika Miyamoto<sup>4</sup>, Hannah Vogts<sup>6</sup>,  
Kyle L. Brown<sup>1,4,7</sup>, Jens Meiler<sup>1,2,3,4,8,9,10</sup> and Cristina E. Martina<sup>1,4\*</sup>

<sup>1</sup>Center for Structural Biology, Vanderbilt University, Nashville, TN, United States, <sup>2</sup>Chemical and Physical Biology Program, Vanderbilt University, Nashville, TN, United States, <sup>3</sup>Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, United States, <sup>4</sup>Department of Chemistry, Vanderbilt University, Nashville, TN, United States, <sup>5</sup>Department of Chemistry and Physics, Belmont University, Nashville, TN, United States, <sup>6</sup>Department of Biomedical Engineering, University of Iowa City, Iowa, IA, United States, <sup>7</sup>Department of Medicine, Division of Nephrology, Vanderbilt University Medical Center, Nashville, TN, United States, <sup>8</sup>Department of Pharmacology, Institute of Chemical Biology, Center for Applied Artificial Intelligence in Protein Dynamics, Vanderbilt University, Nashville, TN, United States, <sup>9</sup>Institute for Drug Discovery, Institute for Computer Science, Wilhelm Ostwald Institute for Physical and Theoretical Chemistry, University Leipzig, Leipzig, Germany, <sup>10</sup>Center for Scalable Data Analytics and Artificial Intelligence ScaDS.AI and School of Embedded Composite Artificial Intelligence SECAI, Leipzig, Germany

Epitope-focused vaccine design aims to improve upon existing immunization strategies by eliciting immune responses against specific epitopes targeted by known therapeutic antibodies. One of the techniques in epitope-focused immunogen design is glycan masking, in which sugars are used to hide epitopes on the protein of interest that are associated with low therapeutic potency. Here, we provide a detailed overview of the computational and experimental techniques associated with glycan masking for immunogen design at a biochemical and biophysical level. We will cover well-established and emerging *in silico* methods for predicting and engineering glycosylation sites. Additionally, we will discuss expression and validation of glycosylated immunogens *in vitro*. We hope this review will be a useful overview for scientists interested in performing glycan masking in their field of research.

## KEYWORDS

epitope-focused vaccine design, glycan masking, glycan shield, glycoengineering, immunogen design, N-linked glycosylation, protein-glycan modeling, reverse vaccinology

## 1 Introduction

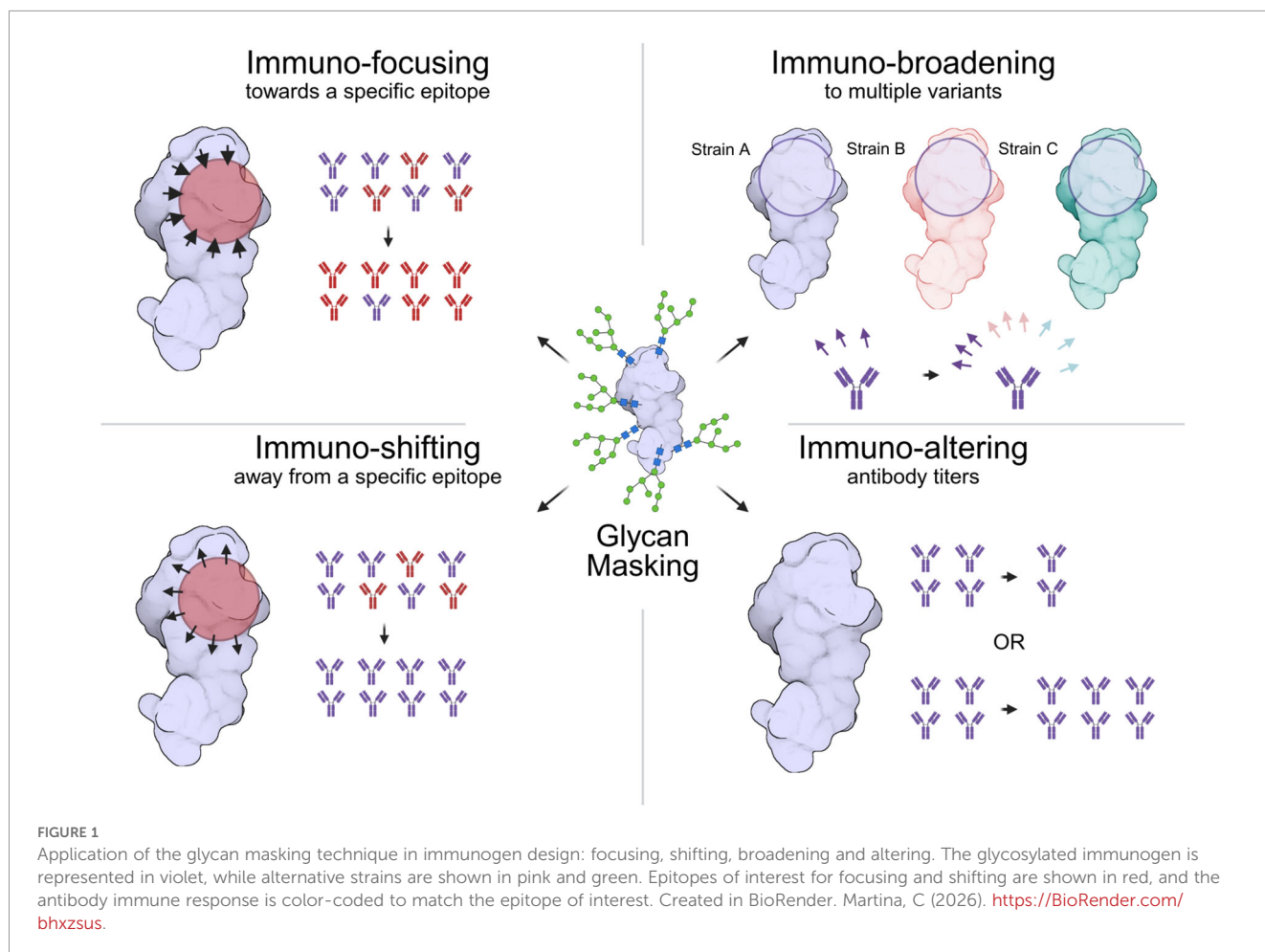
Glycan masking is a natural phenomenon in which viruses escape the immune response by masking therapeutic epitopes with poorly reactive sugars (1–3). This phenomenon is very common among viruses and their envelope glycoproteins, which are the first target of an antibody-mediated immune response. As B-cells mature and their

elicited antibodies' specificities are improved, the newfound selective pressure causes viruses to evolve towards escaping recognition by immune cells. In particular, influenza hemagglutinin (HA), HIV envelope (Env) and coronavirus Spike (S), are often heavily glycosylated and can acquire additional glycosylation sites through evolution (4–7). Glycans can prevent antibody binding through steric hindrance, by physically blocking access to the target epitope. Based on natural viral evasion mechanisms, researchers have developed glycan masking techniques as novel methods in vaccinology, redirecting the immune response aimed at specific outcomes (8).

In epitope-focused vaccine design, glycan masking can, for example, be used to immuno-focus the antibody-mediated immune response to epitopes that elicit therapeutically relevant antibodies (Abs), such as neutralizing antibodies (nAbs), or broadly neutralizing antibodies (bnAbs, see Figure 1). nAbs are Abs that can neutralize pathogenic infection by binding, and can do so through a variety of mechanisms, such as blocking viral receptor binding sites (9). bnAbs are nAbs that can neutralize a wide range of pathogenic strains, generally by binding to epitopes with high sequence and/or structural similarity between pathogenic strains. It is important to note that neutralization is not always necessary for an antibody to confer protection against a pathogen-of-interest (10, 11). In glycan masking, glycosylation sites are engineered such that irrelevant

epitopes are blocked by glycans, allowing, e.g. bnAb epitopes, to be preferentially recognized. Glycan masking has been applied to multiple pathogens, including HIV (12–17), influenza virus (18–20), Zika and Dengue virus (21) (for more information and examples, see Martina et al. (8)). Alternatively, glycan masking can be used to immuno-shift the response away from hypervariable, immunodominant epitopes and/or artificial regions (e.g. scaffold proteins like nanoparticle platforms, non-native residues introduced during antigen engineering, etc.) which would not elicit a broad or potent response (see Figure 1). This has also been performed for HIV (22, 23), influenza virus (18, 20, 24–28), malaria (29), and some artificial scaffolds (30–32). Finally, among the most common applications, glycan masking can also be used to immuno-broaden the response by eliciting cross-reactive antibodies (12–16, 18–20, 24, 25, 28, 33, 34) and to immuno-alter the response by lowering or increasing the overall antibody titers (17, 29, 31, 32) (see Figure 1).

Glycan masking is a very efficient way to redirect the immune response because the *N*-glycans presented by viruses (i.e. GlcNAc) are poorly reactive (35), with only a few classes of antibodies known to bind to them (36–38). Given that glycosylation is one of the most common post-translational modifications in human proteins (39), antibodies may have evolved to not recognize “self” sugars produced by the host cells. By replicating within human cells,



viruses adopt a human-like pattern of glycosylation and sugar propensity, thus being ignored by the immune system antibodies (35).

Glycan masking uses *N*-linked glycosylation, in which a nitrogen atom on an asparagine connects the protein and the sugar moieties, as opposed to *O*-linked glycosylation, which uses an oxygen atom (40, 41). In *N*-linked glycosylation, the amino acid sequon (NxS/T) recognized by glycosyltransferase enzymes is necessary for the addition of the first sugar ring (42) to the asparagine, while the second residue can be any amino acid except proline and the third residue can be either serine or threonine.

Engineering novel glycans into proteins is simple from a theoretical standpoint, but problems in protein expression and folding might arise (43, 44). However, well-placed glycosylation can help in increasing the solubility and overall stability of the proteins of interest (45, 46). These different outcomes are often protein-specific and position-specific, and it is challenging to identify viable, novel glycosylation sites on a protein. Various *in silico* tools have been developed over the years to help researchers with glycan modeling and predictions. Similarly, experimental methods to express and validate glycoproteins are being refined and innovated. Given the importance of glycan masking in the development of biotherapeutics such as vaccines, understanding the array of available tools is of extreme importance.

In this review, we will highlight established procedures, alternative approaches, and emerging technologies for glycan masking in immunogen design. We report currently available tools for predicting glycosylation propensity at the residue level, and for modeling glycans onto the protein of interest. We will include both recent methods that leverage artificial intelligence (AI) and classical non-AI methods. We also provide an overview of standard and novel methods that are used in glycan masking of immunogens, including protein expression, glycan presence and validation of protein folding. We hope this review will work as a guide for researchers working with glycans and glycan masking in immunogen design and beyond.

## 2 Computational techniques: predictions

Initial efforts in immunogen design through glycan masking aimed to predict positions in antigens that were most suitable for glycan insertion. Two main techniques have been used: alignment to homologous sequences, and sequence predictors. Each technique will be described in the following sections, but a summary of publications using these techniques is reported in Table 1. Additionally, a detailed overview of the historical development of AI tools in glycobiology, known as glycoinformatics, has been published by Bojar and Lisacek (47). In particular, datasets in glycomics, glycoproteomics, and glycan-binding have been used for training of various tools, from glycan classifiers to predictors for glycosylation.

## 2.1 Alignment to homologous sequences

Early efforts in identifying putative glycosylation sites for immunogen design were done by analyzing homologous sequences of viral proteins. For example, Lin SC. et al. (24) analyzed the sequences of 163 highly pathogenic influenza A H5N1 strains identified in human samples to identify the most variable regions, which were consequently masked through glycosylation. The rationale behind masking hyper-variable regions is to expand the breadth of action of the vaccine and to make it suitable for preventing infection from many H5N1 strains simultaneously. A total of nine glycosylation sites were identified from five hypervariable regions, but only six of them presented the engineered glycosylation upon SDS-PAGE inspection.

Similarly, Bajic et al. (18) aligned all available sequences of influenza H3 from the 1968 Hong Kong outbreak (H3 HK68) and saw that during viral evolution the number of glycosylation sites in HA doubled over time from 7 to 14. They then designed three

TABLE 1 List of publications using computational techniques to predict sites for glycosylation and to model glycans.

Technique	Publication	Details
Alignment of homologous sequences	Lin SC. et al., 2012 (24)	163 different sequences of influenza H5 HA
	Bajic G. et al., 2019 (18)	Available sequences for influenza H3 HA
	Boyoglu-Barnum S. et al., 2020 (19)	Available sequences for influenza group 1 and 2 HA
Predictors	Duan H. et al., 2018 (17)	NetNGlyc
Sequon modeling	Sampath S. et al., 2013 (29)	Rosetta
	Ingale J. et al., 2014 (16)	Rosetta
	Duan H. et al., 2018 (17)	Pymol
	Adolf-Bryfogle J. et al., 2024 (32)	Rosetta
Glycan modeling	Ingale J. et al., 2014 (16)	-
	Duan H. et al., 2018 (17)	Glycosylate
	Adolf-Bryfogle J. et al., 2024 (32)	Rosetta
	Lemmin et al., 2019 (55)	Python
	Krishna et al., 2024 (54)	RoseTTA Fold All Atom
	Abramson et al., 2024 (53)	AlphaFold3
	Ives et al., 2024 (63)	GlycoShape
Dialpuri et al., 2024 (84)	Privateer	
Molecular Dynamics	Ingale J. et al., 2014 (16)	-

The first column reports the technique name, the second column reports the publication references and the third column reports additional details.

hyperglycosylated immunogens starting from H3 HK68. In the first mutant, five additional glycosylation sites were added to mask everything except for the receptor binding site (RBS) epitope, totaling 12 glycans. The second presented the same five additional glycosylation sites as the first mutant, in addition to a glycosylation to conceal the RBS, for a total of 13. The third presented four additional glycans on top of the second, for a total of 17 glycans potentially able to mask the full surface of the protein. Following mouse immunizations, the authors were able to show that the additional engineered glycans do not affect the overall antibody titers and that the sera from each sample was able to recognize the WT protein H3 HK68. However, the different glycovariants elicited different classes of antibodies, enriching in specific germline  $V_H$  genes 1-69, 1-9, and 5-9-1 (18). This shows that glycan masking can enhance frequencies of preferred  $V_H$ -genes and that glycosylated immunogens can contribute to germline targeting.

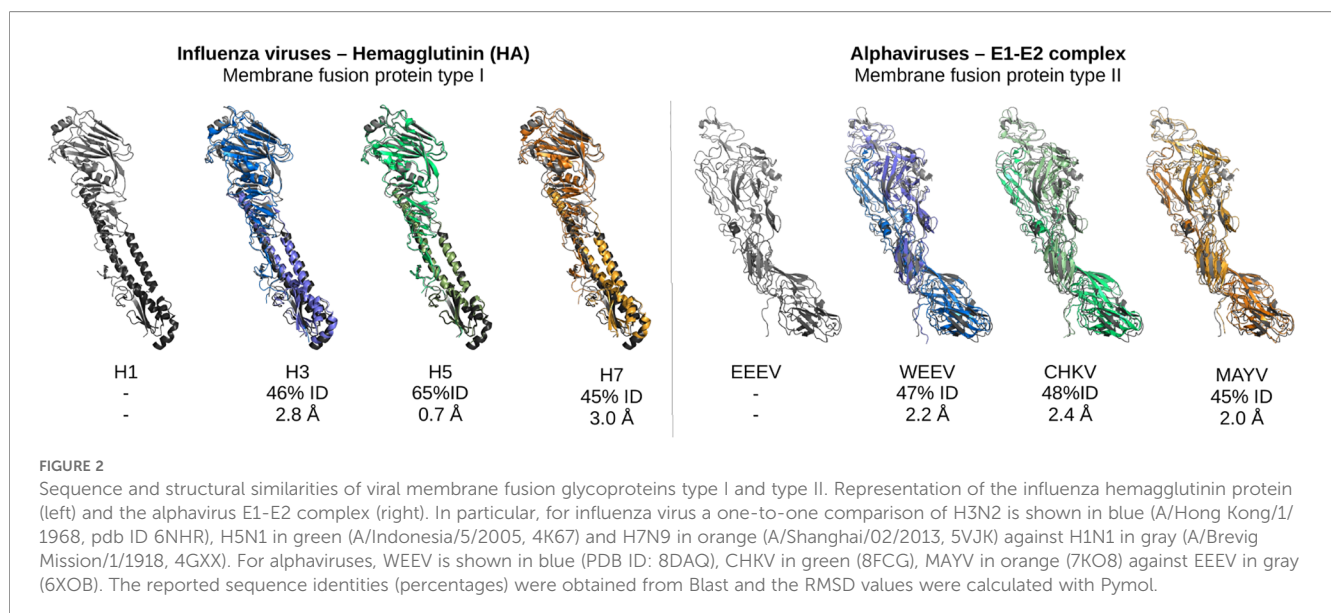
The identification of potential glycosylation sites through sequence alignment is a successful strategy due to the high structural similarity of the proteins within viral families. For example, HA of influenza virus H1 from 1918, which is a type I membrane fusion protein, bears a sequence identity of 46% to H3 HK68, 65% to an H5 strain from 2005, and 45% to an H7 strain from 2013. However, the structural similarity of the HA1-HA2 complex is 2.8 Å to H3, 0.7 Å to H5 and 3.0 Å to H7. Similarly, type II membrane fusion glycoproteins such as those of alphaviruses have low sequence similarity between different species (below 50%), but high structural similarity (see Figure 2). Structural conservation is generally more relevant than sequence similarity since glycans need to be solvent accessible and fully exposed on the surface of the protein. Sequences might vary substantially but still encode for the same fold, and may thus be misleading. The high structural similarity of viral proteins suggests that a potential glycosylation site present in one strain may be well-tolerated in the same structural region of another strain. Thus, this simpler method of selection for glycosylation sites could be highly effective.

## 2.2 Glycosylation predictors

Glycosylation predictors are computational tools to identify and predict sites on the protein of interest that are most likely to undergo natural or engineered glycosylation. An overview of the development and principles behind various glycosylation predictors are given below.

Different generations of glycosylation predictors have been developed over time (47). The predictors from various generations mainly differ from two aspects, one is the feature extraction input, which can be sequence based only or sequence/structure hybrid. The other is the algorithm underneath the tool, traditional machine learning algorithms such as tree forest, support vector machine (SVM) and neural networks are popular options at early generations, but the usage of more advanced algorithms such as protein language model (PLM) becomes more popular since its release around 2023. The first generation, available before 2005, includes bioinformatic tools to identify the presence of the NxS/T sequon within the amino acid sequence of the protein of interest and to evaluate its propensity for glycosylation. An example of such sequence-based tools is NetNGlyc (48), which was used by Duan et al. (17) to identify 13 novel glycosylation sites to insert in an anti-HIV nanoparticle immunogen named eOD-GT8 60-mer, which already presents ten native glycosylation sites. Duan and team (17) designed 13 glycovariants by adding a single glycosylation to the wild-type eOD-GT8 60-mer, and 35 additional variants presenting two glycans simultaneously. All 48 constructs successfully expressed in mammalian cells, and most retained binding to the epitope of interest (CD4 binding site) by target antibodies. The single variants, however, performed worse in masking non-CD4 epitopes, showing that the presence of multiple glycans was more efficient in masking irrelevant epitopes.

Later generations of predictors included additional information such as structural, functional and evolutionary features. Tools like GlycoMine (49, 50) and N-GlyDE (51) have been widely used to



glycosylate viral proteins, though they have rarely been used to identify potential sites for immunogen design through glycan masking. Potential reasons include the fact that the tools were trained only on natural sequences, and that datasets of engineered glycovariants, which could improve the model training, are currently missing or too small. In order to utilize these tools in an immunogen design context, one can apply them to screen engineered immunogens for glycosylation sites. By evaluating the probability of glycosylation at a given site or introduction of glycosylation in the design process, researchers can filter designs that remain unoccupied to increase the chances of successfully masking the target area on the immunogen and ensuring that unwanted epitopes are properly shielded.

### 3 Computational techniques: modeling

With recent advances in structural biology and protein modeling, multiple *in silico* tools have allowed for the accurate modeling of glycans on viral glycoproteins and on immunogens with varying accuracy and speeds. Different tools now allow glycosylation of proteins via sequon introduction, modeling of linked glycans, and evaluation of overall stability. In this section, physics-based and AI-based methods of glycoprotein modeling will be discussed, as well as their applications to the field of glycan masking and immunogen design. The most common techniques in epitope-focused immunogen design are summarized in Table 1.

In general terms, the modeling process typically begins with preparation of the target protein, ensuring sequons are present. Then, relevant glycans are selected for addition to the target protein, and finally the model is built. These tools for glycan modeling have been often utilized for immuno-focusing, increasing chances of success before *in vitro* assessment and playing an important role in vaccine design. Tools used for this application include Rosetta (52)

and other downloadable software suites, in addition to several web-based platforms. AI-based tools like AlphaFold3 (53) or RoseTTAFold All-Atom (RFAA) (54), each with unique algorithms and user interface, have been developed to also model glycans and other small molecules. All tools introduced in this section are summarized in Table 2, although some of them are no longer actively maintained or accessible. Still, their contributions to the field are worth mentioning.

### 3.1 Physics-based glycan modeling

Four commonly used software platforms perform physics-based glycan modeling: Rosetta Suite (32, 52), Glycosylator (55) GlycoShape and Molecular Dynamics (MD) simulation. The Rosetta software suite has played a significant role in general protein modeling and predictions, with the carbohydrate framework integration allowing computational glycan modeling and structural assessment (32, 56, 57). Specifically, the GlycanTreeModeler tool allows for the *de novo* glycosylation of input constructs at designated glycosylation sites; this is not, however, without limitations. Due to the extensive sampling needed to assess the glycans' rotational degrees of freedom; the approach can be computationally expensive. Additionally, current protocols use defined glycoforms, and this can lead to difficulty in predicting glycan microheterogeneity. Finally Rosetta can be sensitive to starting structures potentially affecting the final result, where certain backbone coordinates or glycan orientations may lead to traps in local energy minima. The protocol behind GlycanTreeModeler mimics the growth of trees, and the sugar root linked to the protein is modeled first and energy minimized, then the second layer of sugars are modeled and energy minimized and so on until the last sugar. The protocol uses Monte Carlo optimization and sampling of the glycans' freedom of movement (GlycanSampler). To do this, the algorithm assesses torsion angles,

TABLE 2 Comparison of computational tools for glycan-protein modeling.

Name	Cost	User interface	Algorithm	Evaluation metrics	Input file	Time efficiency	Type
Rosetta	Free for academia	Terminal	Monte Carlo optimization	Rosetta energy score (REU)	pdb	Slow	Physics-based
Glycosylator	Free	Terminal or GUI	CHARMM	N/A	pdb	fast	Physics-based
Glycoshape	Free	GUI	Steric optimized placement algorithm	N/A	pdb	fast	Physics-based
CHARMM	Free for academia	Terminal or GUI	MD	N/A	pdb	slow	Physics-based
RF All-Atom	Free for academia	Terminal	Deep learning	Confidence score, pLDDT and others	Fasta + SMILES	Normally faster than Rosetta	AI-based
AlphaFold3	Free for non-commercial	Terminal or GUI	Deep learning	Confidence score, pLDDT and others	Fasta + CCD codes	Similar to RF all-atom	AI-based
Privateer	Free	GUI	Cremer-Pople algorithm	Conformations, Torsions and others	pdb	fast	Other

Six aspects are taken into account: cost, user interface, underlying algorithm, evaluation metrics, input file type, time efficiency and type. N/A, Not applicable.

side chain optimization, and evaluates the optimization of the glycosidic linkages as it builds, and scores the resulting structures through a Rosetta-specific scoring function. The result is a glycosylated protein with a conformation similar to that of experimentally-determined structures (32). While it is true that the resulting glycoproteins are in agreement with experiments, we must note that the conformations predicted by Rosetta are approximations. It is also important to note that *in vivo* dynamics are significantly more flexible and present a more stochastic nature than those predicted or sampled by Rosetta.

Recently, GlycanTreeModeler has seen use in the glycan masking space to lower the overall immunogenicity of a scaffold protein. In a 2022 study by Kraft et al. (58), protein nanoparticles presenting HIV-based immunogens were glycosylated to shield them from the immune system and allowing the antigen to remain immunodominant over the nanoparticle scaffold. This allowed for the antigen to remain immunodominant over the nanoparticle scaffold. In that study, they evaluated the tertiary structure of the original nanoparticle, looking for outward-facing residues which were optimal for sequon placement. They then used CreateGlycanSequon mover for sequon insertion, and GlycanTreeModeler for glycosylation at the desired site. Finally, the designs were evaluated *in vivo*. There, they found that if the epitope is immuno-dominant to the scaffold being used, which is accomplished via glycosylation, there is no hindrance to the immunogenic effect of the antigen (58). Additionally, in a 2023 study by Dosey et al. (59) it was found that hyperglycosylation of stabilized immunogen candidates both increases and focuses antibody-mediated response to the target epitope in rabbits. In the creation of these glycosylated constructs, Rosetta was used to both guide sequon placement and in the modeling of the glycans before moving to *in vitro* and *in vivo* experiments.

Glycosylator is a Python-based program designed to model glycosylation via flexible interaction modes, including direct execution at the command line, in a Python script, or via the affiliated graphical user interface (GUI) (55). Contributing to the platform's ease of use and flexibility, Glycosylator has also been added to the BuildAMol toolkit which combines multiple programs made in Python for the purpose of molecular modeling (60). Glycosylator allows users to specify their protein of interest and the positions in which they would like glycosylation; there is also a sequon finder that can glycosylate automatically. Functionally, the platform focuses on the structural placement of glycans with users being able to determine glycan placement based on the integrated sequon finder. The user can either provide a PDB file of the glycan and the protein, or they can create the glycan directly using the glycan's IUPAC name. When written into the script, Glycosylator outputs the glycan that can then be used for glycosylation. The program subsequently samples multiple conformations of the sugars and their subsequent glycosylations, resolving any structural conflicts or clashes that are found and sampling geometrically viable conformations. While the program does resolve clashes, before using the structure in platforms such as Rosetta for assessment, it is recommended that the structure be refined first. Glycosylator's functionality can be useful in the field of

immunogen design, where modeling glycan shielding can aid in predicting blocked sites. Additionally, the program can be used to inform vaccine design by aiding in the understanding of conformational states of the target protein, and the effect that this can have on antibody binding. More specifically, the conformational sampling can allow for the prediction of potential steric clashes between the introduced glycan shield and the effect this may have on interaction between the antibody and the immunogen. This can be useful in confirming that engineered glycans effectively shield other potential epitopes without accidentally occluding the desired interaction site. Additionally, Glycosylator can be used in the addition of glycans for locking viral proteins in a prefusion state, as is often done in immunogen design against pathogens like HIV. For example, Bennett and colleagues (61) show, via microsecond-long MD simulation, that glycosylation at areas of interest can be used to lock the envelope protein of HIV-1 in the prefusion state, which has implications relevant to vaccine design due to the conformational changes that are needed prior to host cell entry. Additionally, Glycosylator has seen use in increasing the understanding of antibody antigen interactions through ascertaining the binding distance between antibody and antigen (62). In a 2021 study by Lee et al. (62), the authors note that the glycosylation of epitopes targeted by antibodies against HIV-1 increases the antibody-antigen distance in binding, while the creation of glycan holes decreases this distance. Glycosylator was utilized in the creation of glycosylated structures for subsequent MD evaluation, displaying that the program can be used to study natural phenomena. However, Glycosylator can also be used for engineered constructs, which can be studied through MD or other methods like protein docking. Although the examples do not directly address the design of immunogens, Lee et al. provide an example of how glycan addition or absence can alter the dynamics of protein-protein interactions, specifically antibody-antigen interactions, which is of great interest to those designing immunogens. Additionally, in the work by Bennett et al., Glycosylator is used to alter the protein to study the prefusion state of the envelope protein of HIV-1. Here, glycan alteration displays changes in the conformational state of a viral protein by biasing one state over another. In immunogen design, especially when there is a state of interest, like a prefusion state, design may be done to favor this state over another using programs such as Glycosylator.

GlycoShape is a free computational platform designed to build and visualize glycoprotein structures. The input structures can be derived from existing PDB files or AlphaFold predictions (63). This platform addresses the challenge of missing or incomplete glycan representations in experimental or predicted protein models by adding realistic glycan conformers using a steric-optimized placement algorithm. This algorithm evaluates whether glycans are compatible with protein surface by selecting glycan conformations from a curated library that minimizes steric clashes while maintaining biologically plausible orientations. The platform has a user-friendly web interface. For input, the user uploads a target protein structure in PDB format, then selects a glycan from its dropdown glycan list. After submitting the job, users

can directly visualize the resulting glycoprotein structure within the browser, without installing another protein structure visualization software. This visualization is powered by WebGL (Web Graphics Library) and supports output download in standard formats (e.g., PDB) for use in downstream applications or simulations. GlycoShape also supports an “Advanced Glycosylation” mode, by applying it users are able to model multiple glycosylation sites simultaneously (typically 5–10 glycans per run). However, modeling many glycan sites at once can increase processing time and potentially introduce steric clashes, depending on protein surface topology. Although the platform is very user-friendly, a notable limitation of GlycoShape is its absence of quantitative scoring metrics to evaluate and compare the energy landscape or structure quality of the generated glycoprotein models. Therefore, users who would like to have evaluation metrics of the models may need to feed GlycoShape output structures into other modeling tools like Rosetta or run MD simulations for further assessment. This platform has seen use in stabilization of HIV-1 Env, specifically being used to predict the areas which may be occluded by glycosylation and therefore masked from immune response (64). Kumar and colleagues (64) used multiple tools for glycosylation including GlycoShape emphasizing an approach in which multiple prediction tools can be used to see if a consensus is reached amongst them.

MD simulations allow for the investigation of molecular systems at the atomic level (65). While static structural models have the limitation of oftentimes underestimating the coverage of glycan shielding and potentially missing transient occlusion events, MD allows for the capture of the full dynamic range of the glycan shield and an understanding of these events, if present. Due to the great conformational diversity that glycans display, MD simulations can aid in the representation of the conformational landscape of glycoproteins, thereby allowing researchers to better understand the scope of glycan masking. For example, in a publication by Re and Mizuguchi (66), MD was employed to assess the glycan shield dynamics of Lassa virus to assess where glycan shielding may affect binding of antibodies, stating that this work may aid and influence vaccine design efforts in deciding which antibody response or epitope target sites on the virus are optimal for targeting based on occlusion caused by glycan shield dynamics (66). While MD provides these benefits, some of these phenomena occur on microsecond to millisecond timescales, which can be prohibitively computationally expensive to simulate, depending on the available resources. This idea of increasing our understanding of glycan conformational heterogeneity through computation has aided in our understanding of how viruses use glycans dynamically, and in the design of immunogens. In selecting a glycoprotein modeling platform, researchers may be choosing between Rosetta and Glycosylator. It is important to mention that these two platforms do not necessarily have the same depth and usability, and are not necessarily used for the same purpose. For example, Rosetta offers a comprehensive scoring function for assessing structures and more rigorous structural refinement, while Glycosylator offers a more accessible interface and rapid structure generation. Rosetta also offers more options for assessment methods and potentially for

structure filtering when used in a protein design context, whereas in Glycosylator, multiple platforms might be needed for design and filtering of immunogen designs. Both programs are useful in the field of immunogen design, and the decision of which to use is ultimately up to the user and largely depends on the computational rigor of their experiments and the computational resources available.

Accurate MD simulations involve the consideration of several simulation conditions, including the forcefield (67). Force fields are used to calculate the potential energy landscape of the system as a function of particle coordinates. There are force fields that have been made specifically with consideration paid to the dynamics of glycans, such as GLYCAM/AMBER force fields (68–70), and there others, such as CHARMM have added the functionality to assess glycoproteins over time (71). Hybrid approaches combining protein and glycan simulation methods, including AMBER and GLYCAM, are commonly used and widely validated for viral glycoproteins. It is important to emphasize that oftentimes the reliability and accuracy of MD simulations can depend on the force field chosen, and there has been much effort paid to comparing force fields for both glycoproteins and protein-carbohydrate complexes (72). In addition to the force field, great attention is paid to many other factors such as the simulation box, solvent, temperature, pressure, equilibration steps and other simulation conditions based on the question being answered.

In addition to viral immune escape, glycans can play a large role in the host cell entry of viruses. It has been found via MD simulation and subsequent *in vitro* experiments, that hinge glycans on the spike protein of some coronaviruses increase infectivity and aid in cell entry (73). Studies investigating SARS-CoV-2 and its glycan shield have found regions that could be susceptible to human leukocyte antigen detection due to gaps in glycosylation (74). In addition to the study of naturally occurring glycan shields, MD simulations can aid in the design of immunogens. For example, the glycoprotein prediction platform GlycoShape, utilizes individual MD trajectories, that they congregate into a single dataset allowing for the prediction of a conformational landscape (63). Platforms such as GlycoShape allow for the accurate prediction of not only single structures but also consider the structural diversity of glycans, which plays a large functional role in entry, protection from the immune system, and now in the design of immunogens utilizing glycans to expose epitopes of immunological relevance. For example, Nuqui et al. (75) successfully created simulation-informed immunogen designs for SARS-CoV-2 using information on dynamics, specifically the breathing of the S2 trimer, that was not directly visible in Cryo-EM determined structures alone. The MD simulations conducted here informed the substitution of tryptophans used to lock the S2 trimer in the perfusion state, creating a protein that was more thermostable and elicited an antibody response in mice (75).

### 3.2 AI-based approaches to glycosylation

In addition to the aforementioned downloadable software packages for the modeling of glycoproteins and the assessment of

glycan conformations there are also several web-based and AI platforms that seek to improve glycan modeling, such as RFAA (54), AlphaFold (53, 76), GlycoShape, and others described below. These types of tools, broadly speaking, have pros and cons that should be considered before modeling or creating a pipeline for the modeling of glycoproteins. For example, these tools often have differences in their time efficiency, as the generation time per structure, when compared. Rosetta typically takes the longest, especially if the defined protocol involves an energy minimization step, which is typical for many modeling pipelines (52). For example, for a 2300 aa amino acid long sequence modeled with mannose-5, the time needed can be around one hour (32). The AI based platforms, such as AlphaFold (53, 76) or ColabFold (77), are typically quicker and can take as little as a few minutes for the generation of a structure (78, 79). For context regarding the general needs of each platform type and their pros and cons, refer to Table 2.

RFAA is a structural prediction network that allows for the *de novo* structural modeling of proteins as well as small molecule, nucleic acid, and glycan spatial prediction with an input FASTA sequence and IUPAC glycan name for glycoprotein modeling (54). For the accurate prediction of structures, RFAA uses a three-track neural network utilizing a 1D sequence alignment track, a 2D distance matrix, and a 3D coordinate matrix, while taking into account the sequence, residue distances, and positions, respectively. The use of deep learning platforms for the prediction and/or modeling of glycans in immunogen design, is fairly new, with RFAA's publication in March of 2024. Specifically in the glycan masking space for vaccine design, biophysical methods still outperform AI-based methods. Because of this, there have not been many applications to immunogen design and glycan masking. However, given the promise of RFAA and other similar deep learning platforms, the accuracy of glycan masking predictions will likely be increased, and the protocols for the modeling of a glycoprotein will be simplified over time, allowing a larger range of researchers to adopt the technology for this purpose. It is worth noting that most ML-based structure prediction platforms, like RFAA and AlphaFold2, were designed to yield 'the best' structure given a protein sequence. Confidence metrics enable these algorithms to rank the quality of each structure. This is useful for traditional applications, but is not as useful when an ensemble of protein structures is needed, as in determining the potential space occupied by a glycan chain. These tools can be modified to yield ensembles of protein structures, but may be more computationally expensive than, for example, traditional approaches like Rosetta or MD (80). Ensemble representation of glycans can be beneficial as they possess significant rotational freedom and therefore, do not occupy a fixed location. As a consequence, the use of ensemble representation is important in the determination of the occupancy of the glycan shield, in terms of probabilistic location of the glycans. This allows researchers to better distinguish between constantly shielded epitopes and those which may be transiently revealed. On the other hand, while the ensemble-based approach to glycoprotein modeling offers a realistic representation of biological reality, it can introduce computational cost, as previously mentioned.

AlphaFold3 (AF3) is an AI-based protein structure prediction tool developed by DeepMind (53). Compared to its previous version AlphaFold2 (AF2) (76), and AlphaFold Multimer (AFm) (81), key improvements were made to support a broader range of molecular species. While AF2 and AFm can only predict protein structures or complexes, AF3 can also predict interactions of proteins with other molecule types, including DNA, RNA, small-molecule ligands, and glycans. This expanded functionality allows AF3 to model glycoprotein structures. To use AF3 to model glycoproteins, users provide the protein's amino acid sequence along with Chemical Component Dictionary codes (CCDs) (82) representation of the glycan, rather than providing pre-existing structural files like PDBs. AF3 provides scoring metrics to evaluate the reliability of its predictions. These metrics include the confidence score (ranging from 0 to 100) and pLDDT (predicted Local Distance Difference Test, ranging from 0 to 100). Higher scores indicate higher confidence in the reliability of the predicted structure. Confidence values above 70 often are considered reliable, and values above 90 indicate a very high confidence in structural accuracy (83).

### 3.3 Others - tools for assessing glycoprotein models

Unlike other tools described in previous sections, Privateer is not a tool for generating glycoprotein models, but a tool that provides validation metrics for existing glycoprotein structures (84, 85). According to its website, Privateer doesn't model glycoproteins directly but provides a quality assessment of modeled or experimentally determined glycans. It extracts sugar molecules from the input PDB structure by user, then calculates sugar molecules' parameters according to the Cremer-Pople puckering metrics (86) to confirm their conformations, then compares them to statistically derived distributions from high-quality PDB structures to evaluate how plausible the conformations of the extracted sugar molecules are. Finally, it gives a validation report to inform the user of any potential issues the glycans may have in the given glycoprotein. For example, the report includes fields like number of conformation issues, number of torsion issues and so on.

Finally, GLYCO (87) is another program used for the assessment of glycan masking, while it does not glycosylate directly, it does assess glycan sterics of previously predicted structures. This can inform immunogen design in ascertaining the needed number of glycosylation sites to enhance immunofocusing, in addition to ensuring that glycans do not occlude the needed epitope.

### 3.4 Glycan masking with physics-based versus AI platforms

When considering the AI and web-based protocols for the modeling of glycoproteins (54, 63, 76, 77, 81), it is important to consider the differences in how they model when compared to

platforms such as Rosetta that utilize a more physics-based approach comparatively (32, 52, 56). Rosetta utilizes predefined energy functions and scoring terms to assess glycan conformations and generally is useful for the generation of high-resolution structures and quantitative scoring metrics. Rosetta also allows the exploration of the conformational space via the creation of an ensemble of structures rather than single best conformation, as is typical for AI-based platforms. This sampling of the conformational space can aid in distinguishing areas of the immunogen accessible to antibodies and those that are not. Additionally, Rosetta and other platforms like it, allow for customizability in the glycans being added, so the more niche parts of the field using novel or less common glycan structures may be supported in Rosetta but not on AI platforms, although this is changing as the field progresses. Due to this high resolution assessment of the conformational space however, Rosetta, as previously mentioned, takes more time and has a steeper learning curve when compared to tools like AF3, so if quicker results are needed then AI-based platforms should be considered. Additionally, AI-based platforms allow for the folding of proteins *de novo* with the user providing a sequence and glycosylation site. This allows for the aforementioned combined approach, which is common in the field of using tools like AlphaFold3 for protein structural prediction and Rosetta for glycosylation as a part of a glycoprotein modeling pipeline.

The modeling of glycans onto proteins is complex and requires intensive biophysical analysis, and the rise of AI and web-based protocols in addition to Rosetta and other physics-based systems is becoming more common also in specific fields such as glycan masking for immunogen design. The use of these programs in conjunction, in particular performing protein modeling with AlphaFold and glycan design with Rosetta, is currently the preferred method in glycan masking for immunogen design, and it combines the best of both fields.

## 4 Experimental techniques - glycoprotein biosynthesis

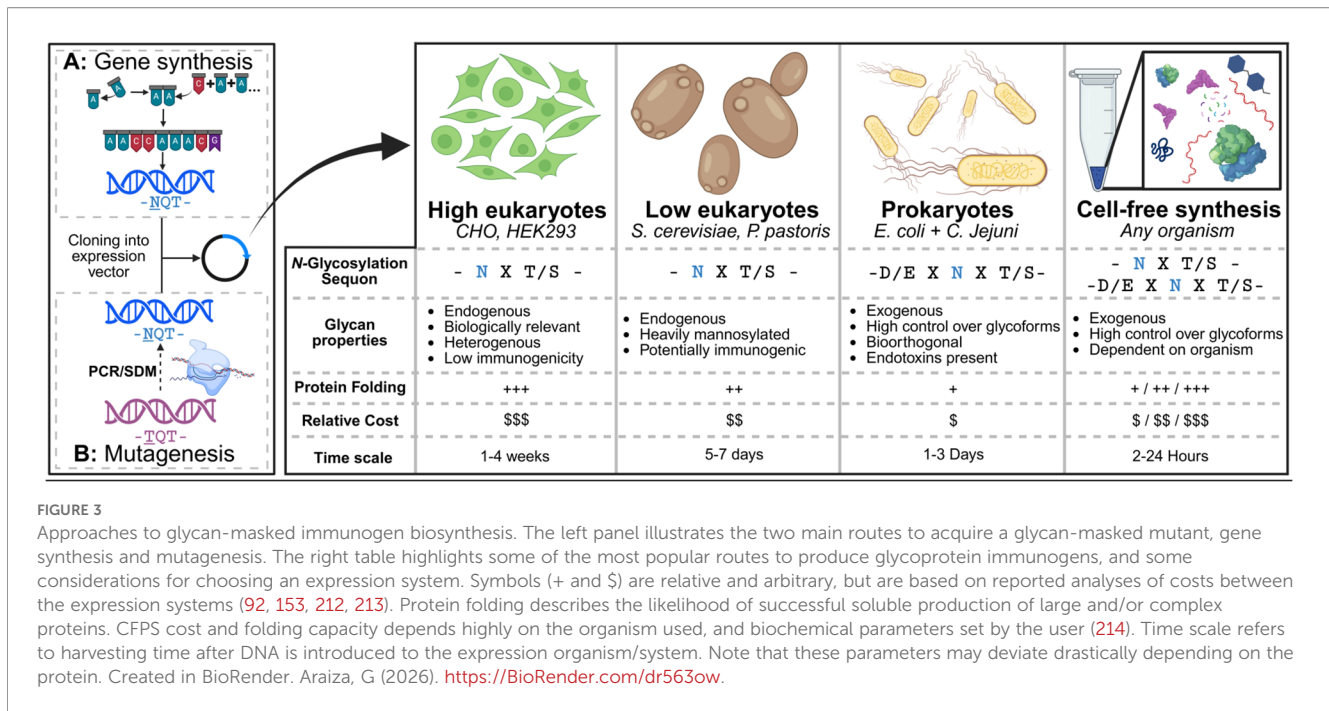
In many cases, the presence of N-linked glycans can improve biophysical and biochemical characteristics such as solubility, proteolytic stability, thermostability, half-life, and overall yields, provided the glycan does not interfere with the protein's folding or function (41, 44). Virus-associated immunogens in particular are usually robust to glycosylation along solvent-exposed regions and loops (59, 88, 89). Some viral proteins, such as HIV Env, are even hyperglycosylated and can contain 30 or more N-linked glycans on its gp120 subunit alone. However, not all pathogen-associated proteins can readily support artificially introduced glycans, and addition of glycans may even result in adverse immunogenic reactions (90). For synthesis of glycoprotein immunogens, the preferred route in the relevant literature is overwhelmingly in favor of eukaryotic cellular expression systems, primarily those of human cell lineages (91). These systems are robust and usually do not require highly specialized cell lines or host genome modification, though this may be a case-by-case consideration.

Here, we discuss general strategies for the biosynthesis of designed glycoprotein immunogens. We will highlight systems that are commonly used and offer alternatives that may be more accessible or convenient to researchers (Figure 3).

### 4.1 DNA, genes, and vectors

The first consideration scientists must take in validating their designs is how they will acquire the genotypes encoding their designed glycoimmunogens. DNA expression vectors are circular DNA that encode the information necessary for an organism to produce a protein exogenously. DNA vectors are typically between 3,000 to 10,000 base pairs (bp), though some vectors can be larger than 30,000 bp (92). The variance in size stems from the modularity of vectors and their features, as they encode the biological information needed for the organism to replicate the DNA, express the protein, traffic the protein, and so on. In addition, vectors can also encode features desirable to researchers, such as selection markers, purification tags, and so on (92, 93). The most direct approach for acquiring artificial genes is through oligonucleotide synthesis, for which many third-party services exist (94). These genes can then each be cloned into an expression vector. Advances in oligonucleotide synthesis have made acquisition of novel DNA sequences highly affordable and convenient for many laboratories throughout the world (95–97). When testing a small number of unique designs, scientists typically order the DNA sequence that encodes the protein-of-interest from a commercial vendor, which must be cloned into an expression vector that harbors signal peptides, affinity tags, among other features. Ordering full-length genes of glycan-masked constructs may be prohibitively expensive when ordering dozens or even hundreds of variants (95). However, depending on the complexity between constructs, it may be substantially more efficient in terms of time to order full-length genes of, say, hyperglycosylated constructs. Additionally, other mutations that accompany the glycoprotein may be necessary. Thornlow and colleagues (20) cloned five distinct genes when designing hyperglycosylated hemagglutinin constructs with up to ten engineered glycosylation sites, totaling 16 N-glycosylations, plus four cysteine mutations that occluded undesired trimer interface epitopes by 'locking' the trimeric structure through disulfide bonds between monomers. Similarly, unique strains of viruses may differ drastically from one another, but common epitopes may be desirable for immunofocusing. Brinkkemper and colleagues (98) demonstrated an alternative to chemical synthesis of template DNA. They isolated five strains of HIV, each widely deviating in phylogeny, directly from patient samples. Brinkkemper et al. then supplemented Env proteins' native glycan shields with engineered glycosylation sites to block immunodominant but nonprotective epitopes. In the study, several mutations were further introduced to stabilize the engineered Env variants, and were also integrated into a two-component nanoparticle system.

An alternative to purchasing full-length genes of glycosylated proteins and immunogens is site-directed mutagenesis (SDM) in



which the wild-type gene is mutated to present sequons for glycosylation (Figure 3). Polymerase chain reaction (PCR) allows specific amplification or modification to template DNA *in vitro*, and it has made SDM convenient and affordable (25, 99). Short (20 to 60 nucleotides), single-stranded DNA primers can be used to amplify site-specific mutants through PCR (100). Xu et al. (101) screened 16 glycosylation mutants and found that seven expressed. From this initial screen, Xu and colleagues (101) combined the four most desirable glycan mutants into constructs with up to four glycosylation sites that expressed well, without the need to order new, full-length genes. Alternatively, larger glycoprotein mutant libraries can be generated and screened at once through cell-surface display (102). These methods, however, may be more suited for *in vitro* stabilization of glycoproteins through directed evolution rather than engineering of novel glycosylation sites (103). This is because of the relatively low frequency of mutation during error-prone mutagenesis that is highly unlikely to yield the specific NxT/S motif if it was absent in the native sequence. Regardless, SDM makes introduction of the glycosylation sequon into wild-type sequences relatively trivial at the genetic level, but does not guarantee successful glycosylation when expressed. It is worth noting the potential utility of *in vitro* engineering techniques, as further stabilization or functional modification of glycan-masked immunogens may be necessary for some projects (104).

## 4.2 High eukaryote expression systems

High eukaryote-based expression systems, particularly those of mammalian cell origins, prevail in the literature for production of designed glycoprotein immunogens, especially for viral immunogens. Mammalian cells efficiently *N*-glycosylate and fold

complex proteins that simpler expression systems, such as bacteria, are not equipped to handle (91, 105). Additionally, the environment of mammalian cells is biologically relevant to the environment in which wild, pathogenic human viruses replicate. Thus, mammalian expression cell lines enable recombinantly expressed viral proteins to adopt native-like trimerization, cysteine pairing, host-like glycosylation, and are subject to the protein folding quality-control mechanisms of the endoplasmic reticulum, and secretion into the extracellular matrix. The recombinant expression of glycoproteins *in vivo* is essentially identical to expression of non-glycosylated proteins (106). The major distinction between non-glycosylated and *N*-glycosylated protein production in eukaryotes is the presence and class of the signal peptide encoded in the plasmid that facilitates trafficking of the translated protein into the endoplasmic reticulum, where the initial glycan is ligated, and then to the golgi apparatus for glycan maturation (91, 98). Care must be taken when selecting the secretion tag, as trafficking is dependent on the organism used for protein expression (93). Additionally, each propeptide will have unique trafficking properties regarding where the fused protein is localized, and what other, if any, post-translational modifications are performed (107–109).

Several human cell lines have been optimized for recombinant protein expression. The most ubiquitous of these, in the context of glycan masked immunogens, are human embryonic kidney (HEK) cells of the 293 lineage (91, 106). HEK293 clonal lineages, such as Gibco Expi293F™, have been engineered for faster growth rates, greater recombinant protein yields, and can be non-adherent to facilitate increased growth densities and scales (52, 106). Though these features are attractive on-paper, in practice they may adversely affect glycoprotein homogeneity and stability. High growth densities and prolonged expression times may influence glycan

identity and protein stability due to nutrient and/or oxidative stress, and other unforeseen biological phenomena (110). It is thus prudent for researchers to optimize and investigate the effects of expression time on their proteins' quality. Non-human mammalian cells, such as Chinese Hamster Ovary cells, are more commonly used for high-yield production of proteins due to their high suspension growth densities, and being less susceptible to contamination from human pathogens (106). CHO cells do vary in glycosylation patterns, including several non-human glycan residues and even some glycans with immunogenic potential, such as  $\alpha$ -galactose (111). Despite this, nearly 70% of available biopharmaceuticals are produced in CHO cells, mostly due to favorable large-scale production properties such as high productivity, processivity, and being less susceptible to contamination (105). However, with regards to vaccine production, cells of human origins are generally favored due to the biological relevance of the immunogens produced. It is important to note that bacterial immunogens may not be suitable for production in mammalian cells, as these tend to be toxic. In addition, most pathogenic bacterial proteins are not natively *N*-glycosylated, and mutation of native, incidental NxT/S sequons that were not intended for glycosylation may be required. Regardless, the versatility of mammalian cell expression systems has enabled scientists to establish both high-throughput expression and screening workflows (112), and high-yielding protocols for recombinant glycoprotein production and characterization (113). Insect (25, 114) and plant cells (115–117) have also been reported in the use of producing glycoprotein immunogens, though these are sparse in comparison to mammalian cell lines regarding pharmaceuticals, mostly due to the presence of harmfully immunogenic glycan residues in these organisms. While they currently remain relatively unpopular in the realm of vaccine production, we speculate that successful humanization of glycosylation patterns within some plants, namely *Nicotiana benthamiana*, and/or other organisms may facilitate large-scale expression workflows more conducive to biopharmaceutical production.

### 4.3 Alternative glycoprotein expression systems

Mammalian cell facilities can be prohibitively expensive and, in most cases, demand specialized equipment, training, and often demand increased biosafety clearance for laboratories to operate (118). Thus, some laboratories may require more accessible organisms to cultivate such as fungi or prokaryotes (119). Additionally, transfection into mammalian cells requires that the plasmid DNA be of high quality and quantity, and may require specialized reagents (120–122). The preparation of sufficient DNA for transfection, although well established, can be a bottleneck in throughput for experiments (123). Simpler organisms such as yeast and prokaryotes offer less rigorous methods of transformation, such as heat-shock, that require only basic laboratory chemicals and equipment (124–126). Additionally, these organisms offer faster

growth rates and cheaper reagents to facilitate growth compared to high-eukaryotes. Thus, although organisms such as yeast and bacteria deviate from humans regarding glycosylation patterns, or even glycocompetence, their convenience and accessibility alone may offset their disadvantages in some experiments, particularly in the prototyping phase. We must emphasize that yeast and prokaryotes are best utilized in the design phases, since their glycans vary drastically in composition, branching, occupancy, and especially immunogenicity (127). Thus, mammalian cells are still recommended for the final stages of glycoimmunogen characterization. Such an approach has been used to screen immunogenic libraries through yeast display followed by production of hits in mammalian cells (128). This allows for a time and cost-efficient pipeline that still allows for experimentation on vaccine candidates produced in biologically relevant environments. Ultimately, it is at the discretion of the scientist to weigh such advantages and disadvantages when designing their experiments (see Figure 3).

Among the many species of yeast, two stand out as the most common in laboratories: *Saccharomyces cerevisiae*, and *Pichia pastoris* (129). Laboratory yeast strains can grow densely with affordable reagents and reliably with minimal equipment. However, yeast glycosylation is known to be mannose-favored, which deviates greatly from human glycosylation patterns (130, 131). In addition, induction of yeast in protein production may require methanol, which is expensive, toxic, and may require additional chemical safety clearance (129). Thus, while yeast can be a useful production medium for both the prototyping phase and for large-scale production, yeast's hypermannosylation is known to alter immunogenicity, which can confound the design of glycan-masked vaccine candidates. *P. pastoris* has been used to produce a glycosylated immunogen based on the stem domain of Influenza hemagglutinin (132). In this work, Wang and coworkers (132) demonstrated the use of *P. pastoris* expression through both shake-flask culturing and bioreactor cultivation. They reported protein yields of 10 mg/L through shake-flask and 100 mg/L in fed-batch cultivation. Additionally, Burnap and colleagues (133) used *P. pastoris* to produce the Abdala vaccine for SARS-CoV-2, which itself is based on the receptor-binding domain (RBD) of the spike protein. The scalability of yeast systems are showcased here, as Burnap and coworkers (133) acquired their samples from 75L and 3,000L lots of the Abdala vaccine. We note that the SARS-CoV-2 Spike RBD construct used by Burnap and colleagues is relatively small at about 26 kDa without glycans, and highly stable. Therefore, other constructs may vary in success with respect to production scalability. *P. pastoris* has shown use for immunogen production in other SARS-CoV-2 vaccine constructs (134–136). *S. cerevisiae* tends to see more use in immunogen engineering through yeast display (137, 138), but is also popular for production of other biopharmaceuticals (139). An extensive review showcasing the versatility of yeast in production of vaccines and pathogen-associated proteins up to 2018 has been reported elsewhere (140).

Bacterial organisms, namely *Escherichia coli*, offer some of the fastest growth rates among model laboratory organisms, high growth densities, and are highly affordable and accessible (92).

Additionally, they are relatively easily genetically engineered and many unique strains have been made commercially available for a wide variety of niche applications (141). However, most prokaryotes, including *E. coli*, lack the complex quality control of protein folding and *N*-glycosylation offered by Eukarya. Prokaryotic organisms may, thus, seem to be poor choices for production of glycoproteins at first glance. However, as Keys and Aebi (142) pointed out in their review on this subject, this is not an inherent weakness, and has presented an opportunity to engineer glycosylation pathways that can *N*-glycosylate exogenously expressed proteins in a bio-orthogonal manner. For this to occur, machinery from an *N*-glycosylating organism must be produced alongside the protein-of-interest (143, 144). One such example is the oligosaccharyltransferase PglB from the bacteria *Campylobacter jejuni*. While *C. jejuni* glycosylates with GalNAc-based heptasaccharides that are known to be immunogenic, work has been done on modifying the glycans into more eukaryote-like GlcNAc glycans (143). Furthermore, the glycoforms produced by *C. jejuni* are usually homogenous, which may be advantageous for, e.g. biophysical experiments but do not simulate the heterogeneity of native eukaryotic glycosylation. It must be noted that the *N*-glycosylation sequon in *C. jejuni* deviates from that of eukaryotes, as it is D/E-X-N-X-S/T, where X is any amino acid besides proline, with DQNAT being the optimal acceptor sequence for PglB glycosylation (145). The orthogonality of engineered *E. coli* glycosylation has been heavily exploited in the production of glycoconjugate vaccines against bacterial pathogens, wherein carrier proteins are *N*-linked to pathogen-associated bacterial glycans (146). These carrier proteins, however, are typically established and stable, and are not themselves the immunogenic target. The limiting factor in bacterial production of glycoimmunogens may instead be the complexity of the antigens of interest themselves, as opposed to the glycocompetency of *E. coli* (105). The primary challenge in producing redesigned glycoprotein immunogens in bacteria would most likely be solubility, as many viral fusion protein targets are metastable (147). Incompatibilities with the *C. jejuni* sequon may also conflict with a researcher's designs.

Cell-free protein synthesis (CFPS) is a rapidly expanding technology in both popularity and utility. A review on vaccines produced through CFPS has been made available by Hu and Kamat (148), and others are available for glycoproteins (149, 150). Briefly, CFPS utilizes the biosynthetic machinery of an organism of choice and is supplemented with a mixture of biochemical reagents (151). A genetic template is then introduced to the mixture, wherein the desired protein is produced *in vitro*. The speed, modularity, and throughput of CFPS systems are what make CFPS an attractive option when screening designed protein sequences. CFPS components are relatively affordable and easy to acquire, and reactions as small as 10  $\mu$ L can yield sufficient protein for simple assays such as enzyme-linked immunosorbance assays (ELISA) (152, 153). Adiga and colleagues reported a portable system for the rapid production and purification of therapeutic proteins using a CHO-based lysate system which was supplemented by purified microsomes to facilitate glycosylation (154). Through this, they

produced the glycosylated therapeutic protein erythropoietin, the vaccine immunogen diphtheria toxoid DT5, and other therapeutic proteins. It must be noted that eukaryote-based CFPS systems require supplementation with microsomes to glycosylate efficiently (155). Counterintuitively, CFPS with prokaryotic lysates for glycoproteins may offer the highest degree of control in terms of glycan identity and homogeneity. DeWinter and colleagues recently reported the development of cell-free glycoprotein synthesis (CFGpS) through mutant PglB-expressing *E. coli*-derived extracts that are capable of exclusively linking recombinant proteins to GlcNAc (156). This study demonstrates promise in terms of homogeneity of glycoforms for target glycoproteins, but suffers from similar shortcomings to that of *in vivo* bacterial expression. Given the novelty of this study, it is unclear how robust this system, or CFPS in general, will be for immunogens and overall vaccine production, though other, similar platforms have been used for various therapeutic glycoproteins (112, 140, 144, 149). The inherent stability of the protein-of-interest is likely to be the limiting factor in terms of expressibility in CFPS.

## 5 Experimental techniques: glycan verification

Recombinant glycoprotein purification is a vast field in itself and has been extensively covered elsewhere (157). The next step from acquiring pure glycoprotein samples is to evaluate whether or not the engineered sequon is properly glycosylated. The presence of the sequon for *N*-glycosylation (N<sub>x</sub>S/T) does not inherently guarantee glycosylation. Thus, it is essential to experimentally verify the presence of glycans, which can be done through various techniques: SDS-PAGE, immunoblotting, enzymatic cleavage of the sugar and mass spectrometry. A summary of the publications using these techniques is shown in Table 3.

### 5.1 SDS-PAGE and western blotting

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) is commonly used to differentiate proteins based on their size (158). Denaturing agents (such as SDS) and reducing agents (such as DTT or  $\beta$ -mercaptoethanol) lead to complete unfolding of the proteins in the mixture, which are then loaded into a polyacrylamide matrix. The denatured proteins migrate at different speeds based on their molecular weight. However, it should be noted that glycosylation often produces diffuse, smeared bands due to glycan heterogeneity, which can complicate interpretation. In combination with SDS-PAGE, Western blotting is used to validate the presence of a protein of interest—for example, via detection of a polyhistidine tag (159). Antibodies specific to the tag and/or to the protein of interest are used to target and identify the corresponding bands within the mixture. Nevertheless, it should be noted that epitope tags may be masked or altered by glycosylation, potentially reducing antibody binding. These techniques are also widely used to identify glycosylated variants

**TABLE 3** List of publications using experimental techniques to verify the presence of glycans on the immunogens.

Technique	Publication	Details
SDS-PAGE/Western	Garrity RR. et al., 1997 (33)	Both
	Pantophlet R. et al., 2003 (12)	Both
	Selvarajah S. et al., 2008 (22)	Both
	Lin SC. et al., 2012 (24)	Both
	Lin SC. et al., 2014 (25)	Both
	Chen TH. et al., 2019 (34)	Both
	Sampath S. et al., 2013 (29)	Both
	Forsell MNE. et al., 2013 (23)	SDS-PAGE only
	Eggink D. et al., 2014 (28)	Both
	Ingale J. et al., 2014 (16)	SDS-PAGE only
	Du L. et al., 2016 (185)	SDS-PAGE only
	Tai W. et al., 2019 (31)	Both
	Boyoglu-Barnum S. et al., 2020 (19)	SDS-PAGE only
	Chen TS. et al., 2021 (26)	Both
Adolf-Bryfogle J. et al., 2024 (32)	SDS-PAGE only	
PNGaseF/EndoH	Garrity RR. et al., 1997 (33)	PNGaseF only
	Lin SC. et al., 2012 (24)	Both
	Lin SC. et al., 2014 (25)	PNGaseF only
	Chen TH. et al., 2019 (34)	Both
	Sampath S. et al., 2013 (29)	PNGaseF only
	Forsell MNE. et al., 2013 (23)	PNGaseF only
	Ingale J. et al., 2014 (16)	PNGaseF only
	Boyoglu-Barnum S. et al., 2020 (19)	PNGaseF only
	Chen TS. et al., 2021 (26)	PNGaseF only
	Adolf-Bryfogle J. et al., 2024 (32)	PNGaseF only
Mass Spectrometry	Ingale J. et al., 2014 (16)	MALDI-TOF
	Du L. et al., 2016 (185)	MALDI-TOF
	Duan H. et al., 2018 (17)	LC-MS

The first column reports the technique name, the second column reports the publication references and the third column reports additional details.

in natural glycoprotein and in glycan masking. Since glycosylation is a covalent modification, denaturing and reducing agents do not remove the glycans from the protein, but the increase in weight is detectable via SDS-PAGE. Since protein glycosylation in the endoplasmic reticulum and Golgi is a stochastic process, glycan heterogeneity is obtained and the additional molecular weight ranges from 0.5 kDa to 2.5 kDa (160). This shift is typically visible on 10%, 12%, or 15% acrylamide SDS-PAGE gels. Glycan-specific antibodies are available, but they are often very specific for sugar type, modification, and overall chemistry (161). Given that

these antibodies are specific to glycans and not the POI, non-specific detection of glycans should be expected in unpurified protein samples.

Testing glycosylated immunogens by SDS-PAGE is a simple and cost-effective method to indirectly demonstrate glycan presence. It is often used in glycan masking experiments either as a standalone technique or in combination with Western blotting (see Table 3). However, this method is low resolution, and when multiple glycans are present on the same protein, additional techniques may be required to identify glycoforms in samples. SDS-PAGE cannot distinguish between glycan occupancy vs. glycan composition, nor can it resolve glycoforms differing by <1 kDa. Alternative techniques may be required to identify glycoforms in samples.

## 5.2 Enzymatic cleavage

Enzymatic cleavage of glycans is a very common technique to verify that glycans are present on the immunogen when performing glycan masking. In particular, the protein of interest is treated with specific glycosidases, enzymes which cleave glycans at points along the glycan chain (162). Following the deglycosylation of the protein, SDS-PAGE and Western blotting enable verification of the deglycosylation by comparing molecular weights and protein-specific identification via epitope tags (163, 164). These experimental methods enable rapid and inexpensive qualitative verification of protein glycosylation, though are not robust when identification of glycans and homogeneity is desired.

Different glycosidases are key for the cleavage of the glycosidic bond to separate glycans from their proteins. The most common enzyme used to validate glycan presence on the protein in glycan masking for immunogen design is Peptide:N-glycosidase F (PNGase-F) (165). This enzyme is very versatile, cleaving the glycosidic bonds between the innermost N-acetylglucosamine and asparagine residue of N-linked glycans as shown in Figure 4. PNGaseF has the advantage of leaving the separated oligosaccharide intact, which can then be analyzed and identified through mass spectrometry. To note, after cleavage of the sugar, PNGaseF will change the asparagine (N) to aspartic acid (D), as shown in Figure 4. This is very important to take into account if performing antibody binding experiments with the sequon but without the sugar itself. Binding to the antibody might be affected by the change from N to D due to PNGaseF cleavage.

Though PNGaseF is widely used, it is unable to cleave N-linked glycans with  $\alpha$ -1,3 core fucosylation, which is commonly found in N-linked glycans of plants (166). PNGaseA exhibits the same enzymatic activity as PNGaseF, but it is not inhibited by the presence of an  $\alpha$  1,3 core fucosylation and it can cleave more complex glycans. Endoglycosidase H (Endo H) is an additional glycosidase with higher specificity compared to PNGaseF, as it only removes high-mannose or certain hybrid N-linked glycosylations (167). It cleaves the beta-linkage in between two N-

acetylglucosamine residues, leaving one sugar attached to the protein (see Figure 4).

Small molecule glycosidase inhibitors such as Kifunensine and Swainsonine can be used during transfection to increase sensitivity to Endo H without compromising protein yield (168). The use of chemical inhibitors in glycoprotein studies is suggested as a cost effective and experimentally simple way to address certain issues that are inherent to studying glycoproteins such as the need to effectively remove glycans for crystallization or create greater homogeneity of glycoforms. Glycosidase inhibitors have also been used to optimize monoclonal antibody production and function. Li et al. discusses cases where these inhibitors have been used to glycoengineer mAbs as a way to enhance their therapeutic activity. They also suggest their potential use in addressing species dependent variations in N-glycan processing of mAbs (160). Kifunensine in particular has also been used to study the role of glycans in regulating viral entry (169). However, it should be noted that Kifunensine can induce high-mannose glycosylation and may artificially increase epitope shielding, potentially altering immunogenicity readouts. On the other hand, glycosidase inhibitors can be used strategically to control the glycan forms presented by the immunogens, for example through mannose enrichment. Finally, cheaper chemical methods to remove glycans from proteins (i.e. sodium hypochlorite, NaOCl) exist, though they are less efficient than enzymes (170), and the harsh conditions might damage and unfold the protein. Glycosidases are experimentally useful, but they are often paired with other experimental methods which help verify the successful removal of the glycans, maintenance of the protein's structure, and identification of the removed glycans.

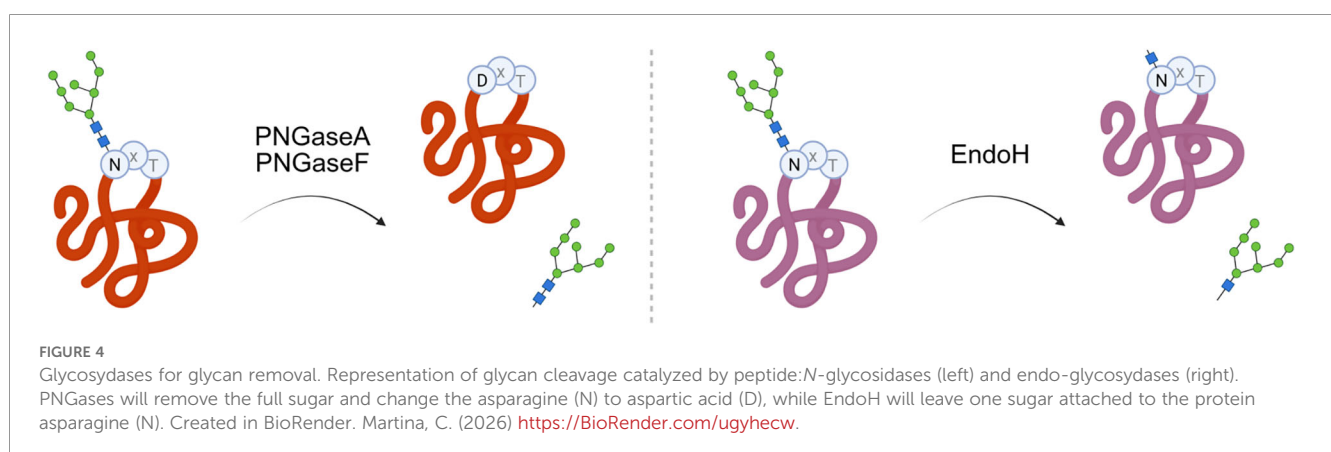
### 5.3 Mass spectrometry

Mass spectrometry (MS) is a powerful analytical tool for evaluating glycosylation sites and determining the structures of the glycans (171). The analysis of N-glycopeptides through MS presents several obstacles. First, the structural heterogeneity of glycans, in terms of composition, branching, and linkage, complicates MS

analysis, as each glycosylation site on a peptide can be occupied by various oligosaccharide structures. Additionally, the increased acidity and negative charge of glycosylated peptides reduce their ionization efficiency, particularly when their MS signals compete with a much larger number of non-glycosylated peptides that produce stronger ion currents (172). Consequently, their MS signals appear considerably weaker than those from non-glycosylated peptides. Enrichment of N-glycopeptides before MS analysis becomes important to enhance their detection sensitivity (173). Lectin-based affinity enrichment utilizes the selective binding between lectins and sugar moieties present on glycopeptides (174). Hydrophilic interaction chromatography relies on polar interactions occurring between glycan hydroxy groups and the stationary phase; this method has gained widespread adoption for its effectiveness in uniformly enriching N-glycopeptides (175). Chemical covalent interaction relies on forming covalent bonds between specific functional groups on the glycan structures and chemically modified solid supports or beads (176). Given the structural diversity among glycoforms and the heterogeneous nature of N-glycosylation sites, there is no single universal N-glycopeptide enrichment strategy, and depending on each study, the appropriate method of choice is selected with further optimization (177).

Regarding MS analysis following glycopeptide sample preparation, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is the most commonly used technique. The workflow involves enriching glycopeptides, treating them with PNGase-F for deglycosylation, then analyzing the resulting free peptides. This enables identification of N-glycosylation site locations and the individual peptides. The masses of released glycans can also be used to deduce their structural information (178–180).

Glycopeptides can also be characterized using electrospray ionization (ESI)-based LC-MS/MS strategies. In this approach, glycopeptides are first separated by liquid chromatography (LC) based on their hydrophobicity and glycan composition, then ionized under gentle electrospray conditions to preserve labile glycan structures. The resulting ions are analyzed in the mass spectrometer, where tandem MS (MS/MS) fragmentation provides information about both the peptide backbone and



attached glycans. This method enables identification of glycosylation sites and more precise glycan structural information, including stereochemistry of glycosidic linkages and branching patterns (181, 182).

Overall, with the improvement and evolution of MS technology and sample preparation techniques, MS has made great strides in glycoprotein analysis. Contemporary applications now encompass glycopeptide sequence determination, glycosylation site occupancy, and comprehensive glycan identification.

## 6 Experimental techniques: functional, and conformational studies

Protein glycosylation and hyper-glycosylation may lead to either global or local misfolding. In the former case, the protein's structural integrity is compromised by the presence of the glycan(s) and protein misfolding occurs in the ER and/or Golgi (183). In these situations, the protein itself is unlikely to mature and thus be extracted in high quantities. In the case of local misfolding, glycosylation may partially alter folds proximal in space to the glycosylation site. This may pose a risk in the development of vaccine candidates, since altered conformations could elicit antibodies able to recognize the designed immunogen but not the native viral protein itself (184). One avenue to verify that engineered glycans do not affect target immunogen folding is by assessing functional activity. In some cases, however, immunogenic properties may be more important than maintaining the immunogen's native function and may even demand inactivation of the function while maintaining the overall structure. Determination of the molecular structure of engineered glycan-masked immunogens reveals the conformational integrity of the antigen, usually unambiguously (6). An overview of examples for verification of structural and functional integrity of glycan masked immunogens is presented in Table 4.

### 6.1 Structural validation

In most cases, it can be assumed that a glycan-masked protein has folded correctly if target antibodies bind (Table 4). This is only the case for conformational epitopes, where epitope residues are proximal in space but not in sequence, thus binding depends on the immunogen's folding. However, certain caveats must be considered: 1. Glycan masking may sterically prevent antibody access even if correctly folded. 2. Some antibodies tolerate partial misfolding. For some projects, known antibodies against desirable conformational epitopes are not available, thus proper folding cannot be confidently deduced from antibody binding, necessitating molecular-level structural determination. However, these methods are generally non-trivial, as they can be highly time-intensive, demand specialized equipment and training, and typically require protein samples that are of superb quality. Still, performing structural validation may be critical to evaluating novel, engineered glycan-masked immunogens.

Low-resolution (with respect to structure) but high-throughput validation of structural integrity is commonplace among designed glycan masked immunogens. ELISAs are biochemical, colorimetric assays that use antibodies for capture and detection of target

TABLE 4 List of publications for structural and functional studies of glycosylated immunogens.

Technique	Publication	Method
Structural validation	Garrity RR et al., 1997 (33)	Binding to antibodies and receptors
	Pantophlet R. et al., 2003 (12)	Binding to antibodies
	Pantophlet R. et al., 2004 (13)	Binding to antibodies
	Selvarajah S. et al., 2005 (14)	Binding to antibodies
	Ahmed FK. et al., 2012 (15)	Binding to antibodies
	Selvarajah S. et al., 2008 (22)	Binding to antibodies, electron microscopy
	Sampath S. et al., 2013 (29)	Binding to receptors
	Forsell MNE. et al., 2013 (23)	Binding to antibodies and receptors
	Eggink D. et al., 2014 (28)	Binding to antibodies
	Ingale J. et al., 2014 (16)	Binding to antibodies and receptors
	Du L. et al., 2016 (185)	Binding to antibodies and receptors
	Duan H. et al., 2018 (17)	Binding to antibodies, electron microscopy
	Bajic G. et al., 2019 (18)	Luminex assay
	Thornlow DN. et al., 2021 (20)	Binding to antibodies
	Tai W. et al., 2019 (31)	Binding to antibodies
	Functional activity	Boyoglu-Barnum S. et al., 2020 (19)
Chen TS. et al., 2021 (26)		Binding to antibodies
Adolf-Bryfogle J. et al., 2024 (32)		Electron microscopy
Lin WS. et al., 2021 (99)		Binding to antibodies
Garrity RR et al., 1997 (33)		Viral replication
Wanzeck K. et al., 2011 (206)		Viral replication
Lin SC. et al., 2012 (24)	Hemagglutination activity	
Sampath S. et al., 2013 (29)	Cytoadherence assay	
Lin HH. et al., 2019 (21)	Flagellin activity	
Chen TS. et al., 2021 (26)	Hemagglutination activity	

The first column reports the technique name, the second column reports the publication references and the third column reports additional details.

proteins. They are among the most common antibody-antigen detection methods due to their affordability, convenience, and versatility, and do not require advanced laboratory equipment. ELISA has been used extensively for preliminary validation of designed glycan-masked immunogens such as HIV Env (12–15, 17, 22, 23, 33, 64), influenza HA (19, 20, 26, 28), Zika virus envelope protein (31), Ebola virus GP (101), and MERS/SARS Spike protein (73, 185). Biolayer interferometry (BLI) is a biophysical assay that grants real-time kinetic data on binding, can detect affinity reductions due to glycan occlusion, even when structure remains intact, and has also been used for a variety of immunogenic targets (16, 64, 186–188). Both methods can be label-free and can also support qualitative collection of binding data with crude protein in clarified cell lysates or media, therefore supporting high-throughput workflows (189, 190), yet they cannot verify epitope geometry. Flow cytometry has also been used and is popular for cell-based assays (17, 18, 29, 185, 186), though this route may demand covalent fluorescent labeling to the protein-of-interest which may adversely affect protein integrity (191). This obstacle may be overcome via site-specific labeling methods such as sortase tagging (192). Still, flow-cytometry provides an extensive range of data when characterizing cellular responses to immunogens *in vitro* (193–196). Other methods such as luminex assays (18), differential scanning fluorimetry (64, 101, 187), AlphaScreen (185, 196), surface plasmon resonance (17), and circular dichroism spectroscopy (36), among others, have been used towards the same end but are not as frequently reported.

Molecular structure determination of glycoproteins is a paradoxical issue with regard to known techniques. The elucidation of protein structures is most popularly done through cryogenic electron microscopy (Cryo-EM), x-ray crystallography (XRC), or nuclear magnetic resonance (NMR). Put simply, the resolution obtained from the former two rely on the rigidity of the analyte, as they are measurements of electron densities (197). Resolution is lost in more flexible regions, such as glycans and flexible loops, where electron densities may not conform. NMR is the opposite, as its signal relies on the resonance, i.e. free movement, of nuclei, but sample preparation demands isotope labeling (198). In our review, we have found that, of the studies that experimentally validate the glycan-masked immunogens' molecular structures, Cryo-EM is the most popular route. This is likely due to the robustness of Cryo-EM in analyzing heterogenous samples, such as those of glycoproteins, in addition to quick data collection compared to other methods like XRC (197, 199). Cryo-EM tends to perform best on large proteins and complexes, ideally exceeding 100 kDa, but has been shown for complexes down to 50 kDa (200, 201). Berndsen and colleagues (5) demonstrated the use of Cryo-EM in evaluating the molecular topology of glycan-masked HIV-1 Env, lending insight into potentially accessible pockets of the Env protein. In addition, Cryo-EM can be amenable to relatively high-throughput, low-resolution structural evaluation of constructs through negative-stain cryo-EM (98, 101, 199, 202). Cryo-EM is thus particularly useful when structurally characterizing protein nanoparticle constructs, including designed hyperglycosylated

constructs (17, 19, 32, 58, 59). XRC has shown use in glycan-masked immunogen design (18) and for many viral glycoproteins, though heterogenous samples may interfere with crystallization (203, 204). NMR may be more suited for the characterization of glycan residues (205).

## 6.2 Functional activity

The simplest way to verify whether engineered glycosylation affects viral integrity is to assess viral replication with the mutants. This has been done by Garrity and team with HIV (33), who generated nine viral variants inserting from one to four glycosylation in the full length genome of HIV-1. They then compared the replication and infectivity rates of each variant, discovering that only one mono-glycosylated variant had rates comparable to the ones of the WT virus. Three other monoglycosylated variants and one biglycosylated variant show a slower growth but similar infectivity, and the remaining bi- tri- and tetra- glycosylated variants were highly affected by the glycan presence for both growth and infectivity. Similarly, Wanzeck and coworkers (206) tested their two variants of influenza virus (bearing two and four additional glycans) for replication efficiency, confirming that viral replication wasn't affected by glycosylation.

Another method to test functional activity and protein integrity after glycosylation is to use known functions of the antigens of interest. For example, influenza HA binds sialic acid exposed on blood cells, which clump together, resulting in hemagglutination (207). Thus, a hemagglutination assay can be used to verify the structural and functional integrity of glycan masked versions of HA. This has been done by Lin and team who tested eight single glycan mutants of HA H5N1 for hemagglutination activity, of which six variants retained functional activity, and two failed to agglutinate red blood cells (24). Similarly, Chen and team, tested for hemagglutination activity three adenovirus vectors encoding for HA glycosylated variants, showing that the mutants were functionally similar to the wild-type (26).

Similarly to the HA of influenza virus, the malaria *P. vivax* antigen Duffy binding protein (PvDBP) binds to the Duffy receptor also exposed on the surface of red blood cells. Its functional activity can be tested *in vitro* with an established cytoadherence assay: the antigen is considered fully functional when five or more red blood cells bind to the cells overexpressing PvDBP and GFP (208). Sampath and coworkers tested three of their PvDBP glycan mutants presenting one, three and eight additional glycans and saw that, compared to the wild-type protein, binding was reduced to 4% in one candidate and to 75% and 88% in the remaining two (29).

Functional assays for glycan masked immunogens are important, since single glycosylations can disrupt the antigen activity and/or viral replicability. However, this approach is not always possible, since many pathogenic antigens might lack specific activities that could be used in parallel to structural validation. Additionally, unlike many structural approaches, functional assays are not generalizable and depend on the pathogen in question.

## 7 Discussion

Glycan masking is a powerful method in epitope-focused vaccine design and is gaining considerable interest in the field of reverse vaccinology. Among others, its main applications in the field of immunogen design are immuno-focusing, immuno-shifting, immuno-broadening, and immuno-altering (8). Here, we focus on computational and experimental methods to design immunogens through the addition of N-linked glycosylations (42).

Many computational methods are available to predict glycosylation sites as well as insert sequons for subsequent glycoprotein modeling. Tools that analyze sequences for potential glycosylation sites based on homologous sequences are available (18, 24) in addition to glycosylation predictors that identify sequons from the protein sequence (48) and structure (49–51). More advanced tools utilize machine learning and allow for the prediction of other post-translational modifications (209). In addition to sequon predictors and tools used for sequon insertion, recent advances in both structural biology and AI have allowed for the creation of tools for glycoprotein modeling, which can typically insert and subsequently model glycosylation computationally. These tools can be broadly separated into physics-based and AI-based. Physics-based glycan modeling examples include the GlycanTreeModeler pipeline within the RosettaSuite and Glycosylator (32, 52, 55). Additionally, computational tools based on MD pay great attention to the development of forcefields that accurately predict glycosylation and glycan pose over the course of a simulation (68–71). Lastly, AI approaches have played a large role in recent years in the field of glycoprotein modeling as it relates to immunogen design. Tools such as AlphaFold (53, 76), ColabFold (77), RFAA (54), and others utilize AI to allow for accurate, easy-to-use, and rapid glycoprotein prediction utilizing sequences as input. The trajectory of this field appears to be following the general trend of machine learning being incorporated into workflows. However, one of the major gaps in training glycan-based AI models is the relatively small and sparse datasets of glycoproteins and glycan conformations. Additionally glycans are flexible and often poorly resolved experimentally, affecting the quality of the available datasets. Sequon and glycosylation predictors can help in the computational design, but none can predict whether the engineered site will actually be glycosylated in a given expression host and whether the glycosylation will disrupt folding, trafficking, or immunogen stability. Physics-based methods remain valuable with regard to modeling glycoprotein structures and interactions, and with a more extensive collection of *in silico* and *in vitro* glycan masking data, new AI model could be specifically trained on and for glycoprotein and glycan dynamics.

Once the glycoprotein structure has been predicted and assessed computationally, there are a number of techniques that can be used experimentally to synthesize these glycoproteins. For viral targets especially, mammalian cells are the most favored due to the biological relevance of the produced glycoproteins (91). Yeast make for well-rounded expression hosts as they are cost-effective, glycosylate natively, and are robust in small and large scale growths.

The primary downside to yeast as expression hosts in glycan masking is the hypermannosylation of proteins. The large size of yeast mannose glycans may lead to false negatives when assessing target antibody binding. Bacterial immunogens may be best produced in glycocompetent/glycoengineered prokaryotes, as these are, in many cases, toxic to mammalian cells. Human-derived cell lines, such as HEK293, are the most representative expression hosts of how engineered glycans will affect target antibody binding. Beyond this, it is difficult to speculate which non-human expression hosts are best in mimicking the effects of human glycosylation patterns on target antibody binding. Each antibody/antigen complex will be unique and the natural heterogeneity of glycoforms already convolutes rational inference of the effects of glycan masking. This is also difficult to predict computationally, as running all possible iterations of designed glycan-masked immunogens in complex with every known glycoform of all common laboratory organisms is not feasible.

CFPS systems are also rapidly growing in popularity with recent use in vaccine production as reviewed by Hu (148). CFPS has several limitations, however, with one being the limited scalability of reactions. Increasing the CFPS reaction scale is complicated due to the need for pure biochemical components. CFPS systems of prokaryotic origins show some promise in the production of homogenous, human-like glycoforms, but have not yet been tested in the realm of immunogen design. These systems are still relatively early in their development, and we anticipate that the improvement and diversification of these systems may impact future production workflows based on cost and time effectiveness.

The explosion in available technologies have also enabled high-throughput approaches and applications for nearly all expression platforms (112, 129, 149), shifting the throughput bottleneck towards gene synthesis. While individual genes are affordable, screening a high number of genes can be costly. Several high-throughput commercial gene synthesis services (such as IDT eBlock or Twist BioScience) offer custom DNA synthesis at roughly \$0.06–\$0.07 per base. These costs rapidly add up, especially since most immunogens are far larger than 100 amino acids. The accessibility of artificial gene synthesis has seen drastic improvements throughout the decades but is now the limiting factor in high-throughput screening of proteins. Streamlining workflows in the chemical synthesis and purification of nucleoside reagents used in gene synthesis may improve downstream affordability.

After glycoprotein synthesis, experimental techniques are available for glycan verification with SDS-PAGE/Western Blot (160, 210, 211), PNGaseF/EndoH (165), and MS (171, 177, 179) being commonly used for this purpose. Blotting methods can be used to differentiate glycosylated and non-glycosylated variants based on their size, while enzyme based methods can be used to remove present glycans for evaluation. MS can be used in sequence determination as well as glycan identification and site localization.

Lastly, there are numerous techniques for functional and conformational study of glycoproteins. Structural validation can be done via lower resolution (with respect to structure) methods such as ELISA, BLI, and flow cytometry which are commonplace in

the field of glycan-masked immunogen design. At residue-level resolution, Cryo-EM is perhaps the most robust choice when characterizing structures of engineered glycoproteins (201). However, the flexibility of glycans limits the usefulness of Cryo-EM in characterizing the potential mechanisms of antibody perturbation from engineered glycans. XRC can be used if atomic-level resolution is necessary, provided the glycoprotein samples are pure and glycoforms are homogenous. Additionally, NMR can be useful when used in the characterization of glycan residues (205). Comprehensive studies combining high-resolution structure data with binding assay datasets and computational simulations seem to be preferred when characterizing the effects of glycan masking. To our knowledge, there are no individual techniques or models that can accurately predict the extent to which glycans within and around epitopes-of-interest can obstruct antibody binding. AI-based methods may also be limited in use here, as datasets of glycoproteins are often limited in accuracy, size, and/or are incomplete. Even so, systematic placement of glycosylation sequons at rationally defined sites has led to extensive success in the realm of glycan masked immunogen design. This can be at least partially attributed to the relative simplicity of the NxT/S sequon, which enables scientists to infer the feasibility of glycosylation on solvent-accessible areas on the surfaces of antigens-of-interest. Precise identification of sites exposed on immunogen surfaces can be sufficient to narrow down the number of constructs tested. Subsequent filtering through other parameters, *e.g.* distance from the epitope, non-neutralizing epitopes, can also lessen the screening burden. Overall, glycan masking is a versatile and accessible strategy in immunofocusing, and readily compatible with nucleotide vaccine delivery platforms. We aim for this review to assist researchers from both computational and experimental backgrounds to enable the use of glycan masking in their research.

## Author contributions

GA: Supervision, Methodology, Data curation, Investigation, Writing – review & editing, Conceptualization, Writing – original draft. JoM: Writing – review & editing, Writing – original draft, Supervision, Investigation, Data curation, Methodology, Conceptualization. MT: Writing – review & editing, Conceptualization, Investigation, Methodology, Writing – original draft, Data curation. SJ: Methodology, Data curation, Writing – review & editing, Investigation, Writing – original draft. SM: Writing – review & editing, Writing – original draft. EM: Writing – original draft, Writing – review & editing. HV: Writing – review & editing, Writing – original draft. KB: Writing – review & editing, Writing – original draft. JeM: Writing – review & editing, Funding acquisition, Writing – original draft. CM: Methodology, Writing – review & editing, Software, Investigation, Writing – original draft, Conceptualization, Visualization, Validation, Data curation, Project administration, Supervision.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by grants from the National Institutes of Health (NIH) U01 AI150739 and U01 AI187059, and cooperative agreements from the Advanced Research Projects Agency for Health (ARPA-H) AY1AX000039.

## Acknowledgments

The authors acknowledge BioRender for the creation of the following figures: **Figure 1** (license LD28VR5AFE), **Figure 3** (SQ28VR65N9) and **Figure 4** (JQ28VR4QHP). **Figure 2** was created using Pymol (The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC). Artificial intelligence tools (ChatGPT, OpenAI) were used only for proofreading and grammatical corrections of the manuscript.

## Conflict of interest

The authors declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. Artificial intelligence tools (ChatGPT, OpenAI) were used exclusively for language editing, including proofreading and minor grammatical corrections. No generative AI tools were employed for the generation of scientific content, data interpretation, or conclusions. The authors take full responsibility for the integrity and accuracy of all scientific aspects of the manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Sommerstein R, Flatz L, Remy MM, Malinge P, Magistrelli G, Fischer N, et al. Arenavirus glycan shield promotes neutralizing antibody evasion and protracted infection. *PLoS Pathog.* (2015) 11:e1005276. doi: 10.1371/journal.ppat.1005276
- Francica JR, Varela-Rohena A, Medvec A, Plesa G, Riley JL, Bates P, et al. Steric shielding of surface epitopes and impaired immune recognition induced by the ebola virus glycoprotein. *PLoS Pathog.* (2010) 6:e1001098. doi: 10.1371/journal.ppat.1001098
- Walls AC, Tortorici MA, Frenz B, Sniijder J, Li W, Rey FA, et al. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol.* (2016) 23:899–905. doi: 10.1038/nsmb.3293
- Gong Y, Qin S, Dai L, Tian Z. The glycosylation in SARS-CoV-2 and its receptor ACE2. *Signal Transduct Target Ther.* (2021) 6:396. doi: 10.1038/s41392-021-00809-8
- Berndsen ZT, Chakraborty S, Wang X, Cottrell CA, Torres JL, Diedrich JK, et al. Visualization of the HIV-1 Env glycan shield across scales. *Proc Natl Acad Sci.* (2020) 117:28014–25. doi: 10.1073/pnas.2000260117
- Morano NC, Guo Y, Becker JE, Li Z, Yu J, Ho DD, et al. Structure of a zoonotic H5N1 hemagglutinin reveals a receptor-binding site occupied by an auto-glycan. *Structure.* (2025) 33:228–33. doi: 10.1016/j.str.2025.01.001
- Cao L, Diedrich JK, Kulp DW, Pauthner M, He L, Park SKR, et al. Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. *Nat Commun.* (2017) 8:14954. doi: 10.1038/ncomms14954
- Martina CE, Crowe JE, Meiler J. Glycan masking in vaccine design: Targets, immunogens and applications. *Front Immunol.* (2023) 14:1126034. doi: 10.3389/fimmu.2023.1126034
- Burton DR. Antiviral neutralizing antibodies: from *in vitro* to *in vivo* activity. *Nat Rev Immunol.* (2023) 23:720–34. doi: 10.1038/s41577-023-00858-w
- Dong J, Gilchuk I, Li S, Irving R, Goff MT, Turner HL, et al. Anti-influenza H7 human antibody targets antigenic site in hemagglutinin head domain interface. *J Clin Invest.* (2020) 130:4734–9. doi: 10.1172/JCI136032
- Bangaru S, Lang S, Schotsaert M, Vandervan HA, Zhu X, Kose N, et al. A site of vulnerability on the influenza virus hemagglutinin head domain trimer interface. *Cell.* (2019) 177:1136–1152.e18. doi: 10.1016/j.cell.2019.04.011
- Pantophlet R, Wilson IA, Burton DR. Hyperglycosylated mutants of human immunodeficiency virus (HIV) type 1 monomeric gp120 as novel antigens for HIV vaccine design. *J Virol.* (2003) 77:5889–901. doi: 10.1128/JVI.77.10.5889-5901.2003
- Pantophlet R, Wilson IA, Burton DR. Improved design of an antigen with enhanced specificity for the broadly HIV-neutralizing antibody b12. *Protein Eng Des Sel.* (2004) 17:749–58. doi: 10.1093/protein/gzh085
- Selvarajah S, Puffer B, Pantophlet R, Law M, Doms RW, Burton DR. Comparing antigenicity and immunogenicity of engineered gp120. *J Virol.* (2005) 79:12148–63. doi: 10.1128/JVI.79.19.12148-12163.2005
- Ahmed FK, Clark BE, Burton DR, Pantophlet R. An engineered mutant of HIV-1 gp120 formulated with adjuvant Quil A promotes elicitation of antibody responses overlapping the CD4-binding site. *Vaccine.* (2012) 30:922–30. doi: 10.1016/j.vaccine.2011.11.089
- Ingale J, Tran K, Kong L, Dey B, McKee K, Schief W, et al. Hyperglycosylated stable core immunogens designed to present the CD4 binding site are preferentially recognized by broadly neutralizing antibodies. *J Virol.* (2014) 88:14002–16. doi: 10.1128/JVI.02614-14
- Duan H, Chen X, Boyington JC, Cheng C, Zhang Y, Jafari AJ, et al. Glycan masking focuses immune responses to the HIV-1 CD4-binding site and enhances elicitation of VRC01-class precursor antibodies. *Immunity.* (2018) 49:301–311.e5. doi: 10.1016/j.immuni.2018.07.005
- Bajic G, Maron MJ, Adachi Y, Onodera T, McCarthy KR, McGee CE, et al. Influenza antigen engineering focuses immune responses to a subdominant but broadly protective viral epitope. *Cell Host Microbe.* (2019) 25:827–835.e6. doi: 10.1016/j.chom.2019.04.003
- Boyoglu-Barnum S, Hutchinson GB, Boyington JC, Moin SM, Gillespie RA, Tsybovsky Y, et al. Glycan repositioning of influenza hemagglutinin stem facilitates the elicitation of protective cross-group antibody responses. *Nat Commun.* (2020) 11:791. doi: 10.1038/s41467-020-14579-4
- Thornlow DN, Macintyre AN, Oguin TH, Karlsson AB, Stover EL, Lynch HE, et al. Altering the immunogenicity of hemagglutinin immunogens by hyperglycosylation and disulfide stabilization. *Front Immunol.* (2021) 7:12. doi: 10.3389/fimmu.2021.737973
- Lin HH, Yang SP, Tsai MJ, Lin GC, Wu HC, Wu SC. Dengue and zika virus domain III-flagellin fusion and glycan-masking E antigen for prime-boost immunization. *Theranostics.* (2019) 9:4811–26. doi: 10.7150/thno.35919
- Selvarajah S, Puffer BA, Lee FH, Zhu P, Li Y, Wyatt R, et al. Focused dampening of antibody response to the immunodominant variable loops by engineered soluble gp140. *AIDS Res Hum Retroviruses.* (2008) 24:301–14. doi: 10.1089/aid.2007.0158
- Forsell MNE, Soldemo M, Dosenovic P, Wyatt RT, Karlsson MCI, Karlsson Hedestam GB. Independent expansion of epitope-specific plasma cell responses upon HIV-1 envelope glycoprotein immunization. *J Immunol.* (2013) 191:44–51. doi: 10.4049/jimmunol.1203087
- Lin SC, Lin YF, Chong P, Wu SC, Pekosz A. Broader neutralizing antibodies against H5N1 viruses using prime-boost immunization of hyperglycosylated hemagglutinin DNA and virus-like particles. *PLoS One.* (2012) 7:e39075. doi: 10.1371/journal.pone.0039075
- Lin SC, Liu WC, Jan JT, Wu SC, Kang SM. Glycan masking of hemagglutinin for adenovirus vector and recombinant protein immunizations elicits broadly neutralizing antibodies against H5N1 avian influenza viruses. *PLoS One.* (2014) 9:e92822. doi: 10.1371/journal.pone.0092822
- Chen TH, Yang YL, Jan JT, Chen CC, Wu SC. Site-specific glycan-masking/unmasking hemagglutinin antigen design to elicit broadly neutralizing and stem-binding antibodies against highly pathogenic avian influenza H5N1 virus infections. *Front Immunol.* (2021) 12:692700. doi: 10.3389/fimmu.2021.692700
- Hershewe JM, Warfel KF, Iyer SM, Peruzzi JA, Sullivan CJ, Roth EW, et al. Improving cell-free glycoprotein synthesis by characterizing and enriching native membrane vesicles. *Nat Commun.* (2021) 12:2363. doi: 10.1038/s41467-021-22329-3
- Eggink D, Goff PH, Palese P. Guiding the Immune Response against Influenza Virus Hemagglutinin toward the Conserved Stalk Domain by Hyperglycosylation of the Globular Head Domain. *J Virol.* (2014) 88:699–704. doi: 10.1128/JVI.02608-13
- Sampath S, Carrico C, Janes J, Gurumoorthy S, Gibson C, Melcher M, et al. Glycan masking of plasmodium vivax Duffy binding protein for probing protein binding function and vaccine development. *PLoS Pathog.* (2013) 9. doi: 10.1371/journal.ppat.1003420
- Kulp DW, Steichen JM, Pauthner M, Hu X, Schiffner T, Liguori A, et al. Structure-based design of native-like HIV-1 envelope trimers to silence non-neutralizing epitopes and eliminate CD4 binding. *Nat Commun.* (2017) 8:1655. doi: 10.1038/s41467-017-01549-6
- Tai W, Chen J, Zhao G, Geng Q, He L, Chen Y, et al. Rational design of zika virus subunit vaccine with enhanced efficacy. *J Virol.* (2019) 93:e02187–18. doi: 10.1128/JVI.02187-18
- Adolf-Bryfogle J, Labonte JW, Kraft JC, Shapovalov M, Raemisch S, Lütke T, et al. Growing Glycans in Rosetta: Accurate *de novo* glycan modeling, density fitting, and rational sequon design. *PLoS Comput Biol.* (2024) 20:e1011895. doi: 10.1371/journal.pcbi.1011895
- Garrity RR, Rimmelzwaan G, Minassian A, Tsai WP, Lin G, De Jong JJ, et al. Refocusing neutralizing antibody response by targeted dampening of an immunodominant epitope. *J Immunol.* (1997) 159:279–89. doi: 10.4049/jimmunol.159.1.279
- Chen T, Liu W, Lin C, Liu C, Jan J, Spearman M, et al. Glycan-masking hemagglutinin antigens from stable CHO cell clones for H5N1 avian influenza vaccine development. *Biotechnol Bioeng.* (2019) 116:598–609. doi: 10.1002/bit.26810
- Newby ML, Allen JD, Crispin M. Influence of glycosylation on the immunogenicity and antigenicity of viral immunogens. *Biotechnol Adv.* (2024) 70:108283. doi: 10.1016/j.biotechadv.2023.108283
- Liang Z, Li C, Gong X, Ye G, Jiang Y, Shi H, et al. Development of Glycan-masked SARS-CoV-2 RBD vaccines against SARS-related coronaviruses. *PLoS Pathog.* (2024) 20:e1012599. doi: 10.1371/journal.ppat.1012599
- Sok D, Pauthner M, Briney B, Lee JH, Saye-Francisco KL, Hsueh J, et al. A prominent site of antibody vulnerability on HIV envelope incorporates a motif associated with CCR5 binding and its camouflaging glycans. *Immunity.* (2016) 45:31–45. doi: 10.1016/j.immuni.2016.06.026
- Garces F, Sok D, Kong L, McBride R, Kim HJ, Saye-Francisco KF, et al. Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell.* (2014) 159:69–79. doi: 10.1016/j.cell.2014.09.009
- Dutta H, Jain N. Post-translational modifications and their implications in cancer. *Front Oncol.* (2023) 13:1240115. doi: 10.3389/fonc.2023.1240115
- Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al. *Essentials of glycobiology, 4th ed.* Cold Spring Harbor (NY: Cold Spring Harbor Laboratory Press (2022).
- He M, Zhou X, Wang X. Glycosylation: mechanisms, biological functions and clinical implications. *Signal Transduct Target Ther.* (2024) 9:194. doi: 10.1038/s41392-024-01886-1
- Rini JM, Esko JD. Glycosyltransferases and glycan-processing enzymes. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, editors. *Essentials of glycobiology, 3rd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY (2015).
- Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc Natl Acad Sci.* (2008) 105:8256–61. doi: 10.1073/pnas.0801340105
- Mitra N, Sinha S, Ramya TNC, Surolia A. N-linked oligosaccharides as outitters for glycoprotein folding, form and function. *Trends Biochem Sci.* (2006) 31:156–63. doi: 10.1016/j.tibs.2006.01.003
- Somnuk P, Hauhart RE, Atkinson JP, Diamond MS, Avirutnan P. N-linked glycosylation of dengue virus NS1 protein modulates secretion, cell-surface expression,

- hexamer stability, and interactions with human complement. *Virology*. (2011) 413:253–64. doi: 10.1016/j.virol.2011.02.022
46. Saedi S, Wan H, Kang H, Gao J, Wu WW, Malik T, et al. N-linked glycans on the stalk of influenza virus neuraminidase promote functional tetramer formation by compensating for local hydrophobicity. *J Virol*. (2025) 99:e00879–25. doi: 10.1128/jvi.00879-25
47. Bojar D, Lisacek F. Glycoinformatics in the artificial intelligence era. *Chem Rev*. (2022) 122:15971–88. doi: 10.1021/acs.chemrev.2c00110
48. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. In: *Biocomputing 2002*. World Scientific, Kauai, Hawaii, USA (2001). p. 310–22.
49. Li FF, Li C, Wang M, Webb GI, Zhang Y, Whisstock JC, et al. GlycoMine: a machine learning-based approach for predicting N-, C- and O-linked glycosylation in the human proteome. *Bioinformatics*. (2015) 31:1411–9. doi: 10.1093/bioinformatics/btu852
50. Li FF, Li C, Revote J, Zhang Y, Webb GI, Li J, et al. GlycoMinestruct: a new bioinformatics tool for highly accurate mapping of the human N-linked and O-linked glycoproteomes by incorporating structural features. *Sci Rep*. (2016) 6:34595. doi: 10.1038/srep34595
51. Pitti T, Chen CT, Lin HN, Choong WK, Hsu WL, Sung TY. N-GlyDE: a two-stage N-linked glycosylation site prediction incorporating gapped dipeptides and pattern-based encoding. *Sci Rep*. (2019) 9:15975. doi: 10.1038/s41598-019-52341-z
52. Koehler Leman J, Weitzner BD, Lewis SM, Adolf-Bryfogle J, Alam N, Alford RF, et al. Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat Methods*. (2020) 17:665–80. doi: 10.1038/s41592-020-0848-2
53. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. (2024) 630:493–500. doi: 10.1038/s41586-024-07487-w
54. Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*. (2024) 384:eadd2528. doi: 10.1126/science.adl2528
55. Lemmin T, Soto C. Glycosylator: a Python framework for the rapid modeling of glycans. *BMC Bioinf*. (2019) 20:513. doi: 10.1186/s12859-019-3097-6
56. Labonte JW, Adolf-Bryfogle J, Schief WR, Gray JJ. Residue-centric modeling and design of saccharide and glycoconjugate structures. *J Comput Chem*. (2017) 38:276–87. doi: 10.1002/jcc.24679
57. Frenz B, Rämisch S, Borst AJ, Walls AC, Adolf-Bryfogle J, Schief WR, et al. Automatically fixing errors in glycoprotein structures with Rosetta. *Structure*. (2019) 27:134–139.e3. doi: 10.1016/j.str.2018.09.006
58. Kraft JC, Pham MN, Shehata L, Brinkkemper M, Boyoglu-Barnum S, Sprouse KR, et al. Antigen- and scaffold-specific antibody responses to protein nanoparticle immunogens. *Cell Rep Med*. (2022) 3:100780. doi: 10.1016/j.xcrm.2022.100780
59. Dosey A, Ellis D, Boyoglu-Barnum S, Syeda H, Saunders M, Watson MJ, et al. Combinatorial immune refocusing within the influenza hemagglutinin RBD improves cross-neutralizing antibody responses. *Cell Rep*. (2023) 42:113553. doi: 10.1016/j.celrep.2023.113553
60. Kleinschmidt N, Lemmin T. BuildAMol: a versatile Python toolkit for fragment-based molecular design. *J Cheminformatics*. (2024) 16:104. doi: 10.1186/s13321-024-00900-6
61. Bennett AL, Edwards RJ, Kosheleva I, Saunders C, Bililign Y, Williams A, et al. Microsecond dynamics control the HIV-1 envelope conformation. *BioRxiv Prepr Serv Biol*. (2023) 2023.05.17.541130. doi: 10.1101/2023.05.17.541130
62. Lee M, Changela A, Gorman J, Rawi R, Bylund T, Chao CW, et al. Extended antibody-framework-to-antigen distance observed exclusively with broad HIV-1 neutralizing antibodies recognizing glycan-dense surfaces. *Nat Commun*. (2021) 12:6470. doi: 10.1038/s41467-021-26579-z
63. Ives CM, Singh O, D'Andrea S, Fogarty CA, Harbison AM, Satheesan A, et al. Restoring protein glycosylation with GlycoShape. *Nat Methods*. (2024) 21:2117–27. doi: 10.1038/s41592-024-02464-7
64. Kumar S, Moral-Sánchez ID, Singh S, Newby ML, Allen JD, Bijl TPL, et al. The design and immunogenicity of an HIV-1 clade C pediatric envelope glycoprotein stabilized by multiple platforms. *Vaccines*. (2025) 13:110. doi: 10.1101/2024.09.14.613016
65. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron*. (2018) 99:1129–43. doi: 10.1016/j.neuron.2018.08.011
66. Re S, Mizuguchi K. Glycan cluster shielding and antibody epitopes on lassa virus envelop protein. *J Phys Chem B*. (2021) 125:2089–97. doi: 10.1021/acs.jpcc.0c11516
67. Lopes PEM, Guvench O, MacKerell AD. Current status of protein force fields for molecular dynamics. *Methods Mol Biol Clifton NJ*. (2015) 1215:47–71. doi: 10.1007/978-1-4939-1465-4\_3
68. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, et al. GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. J Comput Chem*. (2008) 29:622–55. doi: 10.1002/jcc.20820
69. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. (2004) 25:1157–74. doi: 10.1002/jcc.20035
70. Mallajosyula SS, Jo S, Im W, MacKerell AD. Molecular dynamics simulations of glycoproteins using CHARMM. *Methods Mol Biol Clifton NJ*. (2015) 1273:407–29. doi: 10.1007/978-1-4939-2343-4\_25
71. Guvench O, Hatcher E, Venable RM, Pastor RW, MacKerell AD. CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J Chem Theory Comput*. (2009) 5:2353–70. doi: 10.1021/ct900242e
72. Plazinska A, Plazinski W. Comparison of carbohydrate force fields in molecular dynamics simulations of protein-carbohydrate complexes. *J Chem Theory Comput*. (2021) 17:2575–85. doi: 10.1021/acs.jctc.1c00071
73. Chmielewski D, Wilson EA, Pintilie G, Zhao P, Chen M, Schmid MF, et al. Structural insights into the modulation of coronavirus spike tilting and infectivity by hinge glycans. *Nat Commun*. (2023) 14:7175. doi: 10.1038/s41467-023-42836-9
74. Grant OC, Montgomery D, Ito K, Woods RJ. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci Rep*. (2020) 10:14991. doi: 10.1038/s41598-020-71748-7
75. Nuqui X, Casalino L, Zhou L, Shehata M, Wang A, Tse AL, et al. Simulation-driven design of stabilized SARS-CoV-2 spike S2 immunogens. *Nat Commun*. (2024) 15:7370. doi: 10.1038/s41467-024-50976-9
76. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. (2021) 596:583–9. doi: 10.1038/s41586-021-03819-2
77. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. (2022) 19:679–82. doi: 10.1038/s41592-022-01488-1
78. Yin R, Feng BY, Varshney A, Pierce BG. Benchmarking ALPHAFOLD for protein complex modeling reveals accuracy determinants. *Protein Sci*. (2022) 31:e4379. doi: 10.1002/pro.4379
79. Ruffolo JA, Chu LS, Mahajan SP, Gray JJ. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nat Commun*. (2023) 14:2389. doi: 10.1038/s41467-023-38063-x
80. Monteiro Da Silva G, Cui JY, Dalgarno DC, Lisi GP, Rubenstein BM. High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. *Nat Commun*. (2024) 15:2464. doi: 10.1038/s41467-024-46715-9
81. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. *BioRxiv Prepr Serv*. (2021) 10:04.463034. doi: 10.1101/2021.10.04.463034
82. Westbrook JD, Shao C, Feng Z, Zhuravleva M, Velankar S, Young J. The chemical component dictionary: complete descriptions of constituent molecules in experimentally determined 3D macromolecules in the Protein Data Bank. *Bioinformatics*. (2015) 31:1274–8. doi: 10.1093/bioinformatics/btu789
83. Edmunds NS, Genc AG, McGuffin LJ, Cowen L. Benchmarking of AlphaFold2 accuracy self-estimates as indicators of empirical model quality and ranking: a comparison with independent model quality assessment programmes. *Bioinformatics*. (2024) 40:btac491. doi: 10.1093/bioinformatics/btac491
84. Dialpuri JS, Bagdonas H, Schofield LC, Pham PT, Holland L, Bond PS, et al. Online carbohydrate 3D structure validation with the Privateer web app. *Acta Crystallogr Sect F Struct Biol Commun*. (2024) 80:30–5. doi: 10.1107/S2053230X24000359
85. Dialpuri JS, Bagdonas H, Schofield LC, Pham PT, Holland L, Agirre J. Monitoring carbohydrate 3D structure quality with the Privateer database. *Beilstein J Org Chem*. (2024) 20:931–9. doi: 10.3762/bjoc.20.83
86. Cremer D, Pople JA. General definition of ring puckering coordinates. *J Am Chem Soc*. (1975) 97:1354–8. doi: 10.1021/ja00839a011
87. Lee M, Reveiz M, Rawi R, Kwong PD, Chuang GY. GLYCO: a tool to quantify glycan shielding of glycosylated proteins. *Bioinforma Oxf Engl*. (2022) 38:1152–4. doi: 10.1093/bioinformatics/btab791
88. D'Addabbo A, Tong T, Crooks ET, Osawa K, Xu J, Thomas A, et al. Impact of glycan depletion, glycan debranching and increased glycan charge on HIV-1 neutralization sensitivity and immunogenicity. *Glycobiology*. (2024) 34:cwae063. doi: 10.1093/glycob/cwae063
89. Hung HC, Tan BF, Lin WS, Wu SC. Glycan masking of NTD loops with a chimeric RBD of the spike protein as a vaccine design strategy against emerging SARS-CoV-2 Omicron variants. *J Med Virol*. (2024) 96:e29893. doi: 10.1002/jmv.29893
90. Lembo A, Molinaro A, De Castro C, Berti F, Biagini M. Impact of glycosylation on viral vaccines. *Carbohydr Polym*. (2024) 342:122402. doi: 10.1016/j.carbpol.2024.122402
91. Swiech K, De Freitas MCC, Covas DT, Picanço-Castro V. Recombinant glycoprotein production in human cell lines. In: Garcia-Fruitós E, editor. *Insoluble proteins*. Springer New York, New York, NY (2015). p. 223–40. (Methods in Molecular Biology; vol. 1258).
92. Schütz A, Bernhard F, Berrow N, Buyel JF, Ferreira-da-Silva F, Haustaete J, et al. A concise guide to choosing suitable gene expression systems for recombinant protein production. *STAR Protoc*. (2023) 4:102572. doi: 10.1016/j.xpro.2023.102572
93. O'Neill P, Mistry RK, Brown AJ, James DC. Protein-specific signal peptides for mammalian vector engineering. *ACS Synth Biol*. (2023) 12:2339–52. doi: 10.1021/acssynbio.3c00157

94. Catani M, De Luca C, Medeiros Garcia Alcântara J, Manfredini N, Perrone D, Marchesi E, et al. Oligonucleotides: current trends and innovative applications in the synthesis, characterization, and purification. *Biotechnol J.* (2020) 15:e1900226. doi: 10.1002/biot.201900226
95. Hughes RA, Ellington AD. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol.* (2017) 9:a023812. doi: 10.1101/cshperspect.a023812
96. Kosuri S, Church GM. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat Methods.* (2014) 11:499–507. doi: 10.1038/nmeth.2918
97. Weng Z, Li J, Wu Y, Xiu X, Wang F, Zuo X, et al. Massively parallel homogeneous amplification of chip-scale DNA for DNA information storage (MPHAC-DIS). *Nat Commun.* (2025) 16:667. doi: 10.1038/s41467-025-55986-9
98. Brinkkemper M, Kerster G, Brouwer PJM, Tran AS, Torres JL, Ettema RA, et al. Mosaic and mixed HIV-1 glycoprotein nanoparticles elicit antibody responses to broadly neutralizing epitopes. *PLoS Pathog.* (2024) 20:e1012558. doi: 10.1371/journal.ppat.1012558
99. Lin WS, Chen IC, Chen HC, Lee YC, Wu SC. Glycan masking of epitopes in the NTD and RBD of the spike protein elicits broadly neutralizing antibodies against SARS-CoV-2 variants. *Front Immunol.* (2021) 12:795741. doi: 10.3389/fimmu.2021.795741
100. Weiner MP, Costa GL, Schoettlin W, Cline J, Mathur E, Bauer JC. Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene.* (1994) 151:119–23. doi: 10.1016/0378-1119(94)90641-6
101. Xu D, Powell AE, Utz A, Sanyal M, Do J, Patten JJ, et al. Design of universal Ebola virus vaccine candidates via immunofocusing. *Proc Natl Acad Sci.* (2024) 121:e2316960121. doi: 10.1073/pnas.2316960121
102. Packer MS, Liu DR. Methods for the directed evolution of proteins. *Nat Rev Genet.* (2015) 16:379–94. doi: 10.1038/nrg3927
103. Zimmermann A, Prieto-Vivas JE, Voordeckers K, Bi C, Verstrepen KJ. Mutagenesis techniques for evolutionary engineering of microbes – exploiting CRISPR-Cas, oligonucleotides, recombinases, and polymerases. *Trends Microbiol.* (2024) 32:884–901. doi: 10.1016/j.tim.2024.02.006
104. Ye J, Wen F, Xu Y, Zhao N, Long L, Sun H, et al. Error-prone PCR-based mutagenesis strategy for rapidly generating high-yield influenza vaccine candidates. *Virology.* (2015) 482:234–43. doi: 10.1016/j.virol.2015.03.051
105. Lalonde ME, Durocher Y. Therapeutic glycoprotein production in mammalian cells. *J Biotechnol.* (2017) 251:128–40. doi: 10.1016/j.jbiotec.2017.04.028
106. Clarke EC. Considerations for glycoprotein production. In: Bradfute SB, editor. *Recombinant glycoproteins.* Springer US, New York, NY (2024). p. 329–51. *Methods in Molecular Biology*; vol. 2762.
107. Carnell GW, Billmeier M, Vishwanath S, Suau Sans M, Wein H, George CL, et al. Glycan masking of a non-neutralising epitope enhances neutralising antibodies targeting the RBD of SARS-CoV-2 and its variants. *Front Immunol.* (2023) 14:1118523. doi: 10.3389/fimmu.2023.1118523
108. Kou Y, Xu Y, Zhao Z, Liu J, Wu Y, You Q, et al. Tissue plasminogen activator (tPA) signal sequence enhances immunogenicity of MVA-based vaccine against tuberculosis. *Immunol Lett.* (2017) 190:51–7. doi: 10.1016/j.imlet.2017.07.007
109. Cai H, Yu DH, Hu XD, Li SX, Zhu YX. A combined DNA vaccine-prime, BCG-boost strategy results in better protection against *Mycobacterium bovis* challenge. *DNA Cell Biol.* (2006) 25:438–47. doi: 10.1089/dna.2006.25.438
110. Morgan RG, Molnár E, Jones RF, Collard TJ, Lane JD, Greenhough A, et al. Nutrient stress alters the glycosylation status of LGR5 resulting in reduced protein stability and membrane localisation in colorectal tumour cells: implications for targeting cancer stem cells. *Br J Cancer.* (2015) 112:714–9. doi: 10.1038/bjc.2015.4
111. Mizukami A, Caron AL, Picanço-Castro V, Swiech K. Platforms for recombinant therapeutic glycoprotein production. In: Picanço-Castro V, Swiech K, editors. *Recombinant glycoprotein production.* Springer New York, New York, NY (2018). p. 1–14. *Methods in Molecular Biology*; vol. 1674.
112. Chapple SD, Dyson MR. High-throughput expression screening in mammalian suspension cells. In: Chen YW, Yiu CPB, editors. *Structural genomics.* Springer US, New York, NY (2021). p. 117–25. *Methods in Molecular Biology*; vol. 2199.
113. Hacker DL, Durrer L, Quinche S. CHO and HEK293 cultivation and transfection in single-use orbitally shaken bioreactors. In: Hacker DL, editor. *Recombinant protein expression in mammalian cells.* Springer New York, New York, NY (2018). p. 123–31. *Methods in Molecular Biology*; vol. 1850.
114. Clarke EC, Collar AL, Ye C, Cai Y, Anaya E, Rinaldi D, et al. Production and purification of filovirus glycoproteins in insect and mammalian cell lines. *Sci Rep.* (2017) 7:15091. doi: 10.1038/s41598-017-15416-3
115. Margolin E, Chapman R, Williamson A, Rybicki EP, Meyers AE. Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnol J.* (2018) 16:1531–45. doi: 10.1111/pbi.12963
116. Margolin E, Allen JD, Verbeek M, Van Diepen M, Ximpa P, Chapman R, et al. Site-specific glycosylation of recombinant viral glycoproteins produced in *Nicotiana benthamiana*. *Front Plant Sci.* (2021) 12:709344. doi: 10.3389/fpls.2021.709344
117. Srisangsung T, Phetphong T, Manopwisedjaroen S, Rattanapit K, Bulaon CJI, Thitithanyanont A, et al. The impact of N-glycans on the immune response of plant-produced SARS-CoV-2 RBD-Fc proteins. *Biotechnol Rep Amst Neth.* (2024) 43:e00847. doi: 10.1016/j.btre.2024.e00847
118. Stark JC, Jaroentomeechai T, Warfel KF, Hershewe JM, DeLisa MP, Jewett MC. Rapid biosynthesis of glycoprotein therapeutics and vaccines from freeze-dried bacterial cell lysates. *Nat Protoc.* (2023) 18:2374–98. doi: 10.1038/s41596-022-00799-z
119. Ashok A, Brison M, LeTallec Y. Improving cold chain systems: Challenges and solutions. *Vaccine.* (2017) 35:2217–23. doi: 10.1016/j.vaccine.2016.08.045
120. Sherba JJ, Hogquist S, Lin H, Shan JW, Shreiber DI, Zahn JD. The effects of electroporation buffer composition on cell viability and electro-transfection efficiency. *Sci Rep.* (2020) 10:3053. doi: 10.1038/s41598-020-59790-x
121. Tan S, Tao Z, Loo S, Su L, Chen X, Ye L. Non-viral vector based gene transfection with human induced pluripotent stem cells derived cardiomyocytes. *Sci Rep.* (2019) 9:14404. doi: 10.1038/s41598-019-50980-w
122. Fus-Kujawa A, Prus P, Bajdak-Rusinek K, Teper P, Gawron K, Kowalcuk A, et al. An overview of methods and tools for transfection of eukaryotic cells *in vitro.* *Front Bioeng Biotechnol.* (2021) 9. doi: 10.3389/fbioe.2021.701031
123. Cohen M, Randolph DE, Lozano ME, Anderson PW, Crissman J, Triana FJ, et al. A fully automated high-throughput plasmid purification workstation for the generation of mammalian cell expression-quality DNA. *SLAS Technol.* (2022) 27:227–36. doi: 10.1016/j.slast.2022.01.005
124. Pope B, Kent HM. High efficiency 5 min transformation of *Escherichia coli.* *Nucleic Acids Res.* (1996) 24:536–7. doi: 10.1093/nar/24.3.536
125. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc.* (2007) 2:31–4. doi: 10.1038/nprot.2007.13
126. Kawai S, Hashimoto W, Murata K. Transformation of *Saccharomyces cerevisiae* and other fungi: Methods and possible underlying mechanism. *Bioeng Bugs.* (2010) 1:395–403. doi: 10.4161/bbug.1.6.13257
127. Malladi SK, Singh R, Pandey S, Gayathri S, Kanjo K, Ahmed S, et al. Design of a highly thermostolerant, immunogenic SARS-CoV-2 spike fragment. *J Biol Chem.* (2021) 296:100025. doi: 10.1074/jbc.RA120.016284
128. Rathore U, Saha P, Kesavardhana S, Kumar AA, Datta R, Devanarayanan S, et al. Glycosylation of the core of the HIV-1 envelope subunit protein gp120 is not required for native trimer formation or viral infectivity. *J Biol Chem.* (2017) 292:10197–219. doi: 10.1074/jbc.M117.788919
129. Fernández FJ, Gómez S, Vega MC. High-throughput protein production in yeast. In: Vincetelli R, editor. *High-throughput protein production and purification.* Springer New York, New York, NY (2019). p. 69–91. *Methods in Molecular Biology*; vol. 2025.
130. Tang H, Wang S, Wang J, Song M, Xu M, Zhang M, et al. N-Hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in *Saccharomyces cerevisiae.* *Sci Rep.* (2016) 6:25654. doi: 10.1038/srep25654
131. Vervecken W, Kaigorodov V, Callewaert N, Geysens S, De Vusser K, Contreras R. *In vivo* synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris.* *Appl Environ Microbiol.* (2004) 70:2639–46. doi: 10.1128/AEM.70.5.2639-2646.2004
132. Wang SC, Liao HY, Zhang JY, Cheng TJR, Wong CH. Development of a universal influenza vaccine using hemagglutinin stem protein produced from *Pichia pastoris.* *Virology.* (2019) 526:125–37. doi: 10.1016/j.virol.2018.10.005
133. Burnap SA, Calvaresi V, Cabrera G, Pousa S, Limonta M, Ramos Y, et al. Structural and functional glycosylation of the abdal COVID-19 vaccine. *Glycobiology.* (2025) 35:cwaf001. doi: 10.1093/glycob/cwaf001
134. Chen WH WH, Tao X, Agrawal AS, Algaissi A, Peng BH, Pollet J, et al. Yeast-expressed SARS-CoV recombinant receptor-binding domain (RBD219-N1) formulated with aluminum hydroxide induces protective immunity and reduces immune enhancement. *Vaccine.* (2020) 38:7533–41. doi: 10.1016/j.vaccine.2020.09.061
135. Chen WH WH, Pollet J, Strych U, Lee J, Liu Z, Kundu RT, et al. Yeast-expressed recombinant SARS-CoV-2 receptor binding domain RBD203-N1 as a COVID-19 protein vaccine candidate. *Protein Expr Purif.* (2022) 190:106003. doi: 10.1016/j.pep.2021.106003
136. Hotez PJ, Adhikari R, Chen WH, Chen YL, Gillespie P, Islam NY, et al. From concept to delivery: a yeast-expressed recombinant protein-based COVID-19 vaccine technology suit able for global access. *Expert Rev Vaccines.* (2023) 22:495–500. doi: 10.1080/14760584.2023.2217917
137. Nguyen NL, Kim JM, Park JA, Park SM, Jang YS, Yang MS, et al. Expression and purification of an immunogenic dengue virus epitope using a synthetic consensus sequence of envelope domain III and *Saccharomyces cerevisiae.* *Protein Expr Purif.* (2013) 88:235–42. doi: 10.1016/j.pep.2013.01.009
138. Kim YS, Bhandari R, Cochran JR, Kuriyan J, Wittrup KD. Directed evolution of the epidermal growth factor receptor extracellular domain for expression in yeast. *Proteins.* (2006) 62:1026–35. doi: 10.1002/prot.20618
139. Maneira C, Chamas A, Lackner G. Engineering *Saccharomyces cerevisiae* for medical applications. *Microb Cell Factories.* (2025) 24:12. doi: 10.1186/s12934-024-02625-5
140. Kumar R, Kumar P. Yeast-based vaccines: New perspective in vaccine development and application. *FEMS Yeast Res.* (2019) 19. doi: 10.1093/femsyr/foz007

141. Morão IG, Manzine LR, Clementino LOD, Wrenger C, Nascimento AS, Beddoe T. A scalable screening of *E. coli* strains for recombinant protein expression. *PLoS One*. (2022) 17:e0271403. doi: 10.1371/journal.pone.0271403
142. Keys TG, Aebi M. Engineering protein glycosylation in prokaryotes. *Curr Opin Syst Biol*. (2017) 5:23–31. doi: 10.1016/j.coisb.2017.05.016
143. Schwarz F, Huang W, Li C, Schulz BL, Lizak C, Palumbo A, et al. A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. *Nat Chem Biol*. (2010) 6:264–6. doi: 10.1038/nchembio.314
144. Ding N, Ruan Y, Fu X, Lin Y, Yu H, Han L, et al. Improving production of N-glycosylated recombinant proteins by leaky *Escherichia coli*. *3 Biotech*. (2019) 9:302. doi: 10.1007/s13205-019-1830-5
145. Chen MM, Glover KJ, Imperiali B. From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochem*. (2007) 46:5579–85. doi: 10.1021/bi602633n
146. Dow JM, Mauri M, Scott TA, Wren BW. Improving protein glycan coupling technology (PGCT) for glycoconjugate vaccine production. *Expert Rev Vaccines*. (2020) 19:507–27. doi: 10.1080/14760584.2020.1775077
147. White JM, Delos SE, Brecher M, Schornberg K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol*. (2008) 43:189–219. doi: 10.1080/10409230802058320
148. Hu VT, Kamat NP. Cell-free protein synthesis systems for vaccine design and production. *Curr Opin Biotechnol*. (2023) 79:102888. doi: 10.1016/j.copbio.2022.102888
149. Palma JA, Bunyatov MI, Hulbert SW, Jewett MC, DeLisa MP. Bacterial glycoengineering: Cell-based and cell-free routes for producing biopharmaceuticals with customized glycosylation. *Curr Opin Chem Biol*. (2024) 81:102500. doi: 10.1016/j.cbpa.2024.102500
150. Bidstrup EJ, Kwon YH, Kim K, Bandi CK, Aw R, Jewett MC, et al. Cell-free systems for the production of glycoproteins. *Methods Mol Biol Clifton NJ*. (2024) 2762:309–28. doi: 10.1007/978-1-0716-3666-4\_19
151. Sawasaki T, Ogasawara T, Morishita R, Endo Y. A cell-free protein synthesis system for high-throughput proteomics. *Proc Natl Acad Sci*. (2002) 99:14652–7. doi: 10.1073/pnas.232580399
152. Hunt AC, Vögeli B, Hassan AO, Guerrero L, Kightlinger W, Yoeseop DJ, et al. A rapid cell-free expression and screening platform for antibody discovery. *Nat Commun*. (2023) 14:3897. doi: 10.1038/s41467-023-38965-w
153. Guzman-Chavez F, Arce A, Adhikari A, Vadhin S, Pedroza-Garcia JA, Gandini C, et al. Constructing cell-free expression systems for low-cost access. *ACS Synth Biol*. (2022) 11:1114–28. doi: 10.1021/acssynbio.1c00342
154. Adiga R, Al-adhami M, Andar A, Borhani S, Brown S, Burgenson D, et al. Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat BioMed Eng*. (2018) 2:675–86. doi: 10.1038/s41551-018-0259-1
155. Gurramkonda C, Rao A, Borhani S, Pilli M, Deldari S, Ge X, et al. Improving the recombinant human erythropoietin glycosylation using microsome supplementation in CHO cell-free system. *Biotechnol Bioeng*. (2018) 115:1253–64. doi: 10.1002/bit.26554
156. DeWinter MA, Wong DA, Fernandez R, Kightlinger W, Thames AH, DeLisa MP, et al. Establishing a cell-free glycoprotein synthesis system for enzymatic N-glycanylation. *ACS Chem Biol*. (2024) 19:1570–82. doi: 10.1021/acscchembio.4c00228
157. Wingfield PT. Overview of the purification of recombinant proteins. *Curr Protoc Protein Sci*. (2015) 80:6.1.1–6.1.35. doi: 10.1002/0471140864.ps0601s80
158. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. (1970) 227:680–5. doi: 10.1038/227680a0
159. Sule R, Rivera G, Gomes AV. Western blotting (immunoblotting): history, theory, uses, protocol and problems. *BioTechniques*. (2023) 75:99–114. doi: 10.2144/btn-2022-0034
160. Scheller C, Krebs F, Wiesner R, Wätzig H, Oltmann-Norden I. A comparative study of CE-SDS, SDS-PAGE, and Simple Western—Precision, repeatability, and apparent molecular mass shifts by glycosylation. *Electrophoresis*. (2021) 42:1521–31. doi: 10.1002/elps.202100068
161. Gillmann KM, Temme JS, Marglous S, Brown CE, Gildersleeve JC. Anti-glycan monoclonal antibodies: Basic research and clinical applications. *Curr Opin Chem Biol*. (2023) 74:102281. doi: 10.1016/j.cbpa.2023.102281
162. Stanley P. Golgi glycosylation. *Cold Spring Harb Perspect Biol*. (2011) 3:a005199. doi: 10.1101/cshperspect.a005199
163. Crabtree MB, Kinney RM, Miller BR. Deglycosylation of the NS1 protein of dengue 2 virus, strain 16681: construction and characterization of mutant viruses. *Arch Virol*. (2005) 150:771–86. doi: 10.1007/s00705-004-0430-8
164. Vieyres G, Thomas X, Descamps V, Duverlie G, Patel AH, Dubuisson J. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus. *J Virol*. (2010) 84:10159–68. doi: 10.1128/JVI.01180-10
165. Wang T, Voglmeir J. PNGases as valuable tools in glycoprotein analysis. *Protein Pept Lett*. (2014) 21:976–85. doi: 10.2174/0929866521666140626111237
166. Tretter V, Altmann F, März L. Peptide- $N^4$ -( $N$ -acetyl- $\beta$ -glucosaminyl) asparagine amidase F cannot release glycans with fucose attached  $\alpha 1 \rightarrow 3$  to the asparagine-linked  $N$ -acetylglucosamine residue. *Eur J Biochem*. (1991) 199:647–52. doi: 10.1111/j.1432-1033.1991.tb16166.x
167. Yamamoto K. Microbial endoglycosidases for analyses of oligosaccharide chains in glycoproteins. *J Biochem (Tokyo)*. (1994) 116:229–35. doi: 10.1093/oxfordjournals.jbchem.a124510
168. Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, et al. Glycoprotein structural genomics: solving the glycosylation problem. *Struct Lond Engl*. (2007) 15:267–73. doi: 10.1016/j.str.2007.01.011
169. Yang Q, Hughes TA, Kelkar A, Yu X, Cheng K, Park S, et al. Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration. *eLife*. (2020) 9:e61552. doi: 10.7554/eLife.61552
170. Fischler DA, Orlando R. N-linked glycan release efficiency: A quantitative comparison between NaOCl and PNGase F release protocols. *J Biomol Tech JBT*. (2019) 30:58–63. doi: 10.7171/jbt.19-3004-001
171. Xu S, Wu R. Glycobiology and proteomics: has mass spectrometry moved the field forward? *Expert Rev Proteomics*. (2023) 20:303–7. doi: 10.1080/14789450.2023.2255748
172. Cao L, Qu Y, Zhang Z, Wang Z, Prytkova I, Wu S. Intact glycopeptide characterization using mass spectrometry. *Expert Rev Proteomics*. (2016) 13:513–22. doi: 10.1586/14789450.2016.1172965
173. Palaniappan KK, Bertozzi CR. Chemical glycoproteomics. *Chem Rev*. (2016) 116:14277–306. doi: 10.1021/acs.chemrev.6b00023
174. Santorelli L, Capitoli G, Chinello C, Piga I, Clerici F, Denti V, et al. In-depth mapping of the urinary N-glycoproteome: distinct signatures of ccRCC-related progression. *Cancers*. (2020) 12:239. doi: 10.3390/cancers12010239
175. Qing G, Yan J, He X, Li X, Liang X. Recent advances in hydrophilic interaction liquid chromatography materials for glycopeptide enrichment and glycan separation. *TRAC Trends Anal Chem*. (2020) 124:115570. doi: 10.1016/j.trac.2019.06.020
176. Xiao H, Suttapitugsakul S, Sun F, Wu R. Mass spectrometry-based chemical and enzymatic methods for global analysis of protein glycosylation. *Acc Chem Res*. (2018) 51:1796–806. doi: 10.1021/acs.accounts.8b00200
177. Riley NM, Bertozzi CR, Pitteri SJ. A pragmatic guide to enrichment strategies for mass spectrometry-based glycoproteomics. *Mol Cell Proteomics MCP*. (2021) 20:100029. doi: 10.1074/mcp.R120.002277
178. Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. *Nat Protoc*. (2012) 7:1285–98. doi: 10.1038/nprot.2012.062
179. Devakumar A, Mechref Y, Kang P, Novotny MV, Reilly JP. Identification of isomeric N-glycan structures by mass spectrometry with 157 nm laser-induced photofragmentation. *J Am Soc Mass Spectrom*. (2008) 19:1027–40. doi: 10.1016/j.jasms.2008.03.005
180. Both P, Green AP, Gray CJ, Sardžik R, Voglmeir J, Fontana C, et al. Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. *Nat Chem*. (2014) 6:65–74. doi: 10.1038/nchem.1817
181. Bagdonaitė I, Malaker SA, Polasky DA, Riley NM, Schjoldager K, Vakhrushev SY, et al. Glycoproteomics. *Nat Rev Methods Primer*. (2022) 2:48. doi: 10.1038/s43586-022-00128-4
182. Illiano A, Pinto G, Melchiorre C, Carpentieri A, Faraco V, Amoresano A. Protein glycosylation investigated by mass spectrometry: an overview. *Cells*. (2020) 9:1986. doi: 10.3390/cells9091986
183. Xu C, Ng DTW. Glycosylation-directed quality control of protein folding. *Nat Rev Mol Cell Biol*. (2015) 16:742–52. doi: 10.1038/nrm4073
184. Graham BS, Gilman MSA, McLellan JS. Structure-based vaccine antigen design. *Annu Rev Med*. (2019) 70:91–104. doi: 10.1146/annurev-med-121217-094234
185. Du L, Tai W, Yang Y, Zhao G, Zhu Q, Sun S, et al. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. *Nat Commun*. (2016) 7:13473. doi: 10.1038/ncomms13473
186. Hauser BM, Sangesland M, St. Denis KJ, Lam EC, Case JB, Windsor IW, et al. Rationally designed immunogens enable immune focusing following SARS-CoV-2 spike imprinting. *Cell Rep*. (2022) 38:110561. doi: 10.1016/j.celrep.2022.110561
187. Bruun TUJ, Do J, Weidenbacher PAB, Utz A, Kim PS. Engineering a SARS-CoV-2 vaccine targeting the receptor-binding domain cryptic-face via immunofocusing. *ACS Cent Sci*. (2024) 10:1871–84. doi: 10.1021/acscentsci.4c00722
188. Pierce BG, Keck ZY, Wang R, Lau P, Garagusi K, Elkholy K, et al. Structure-based design of hepatitis C virus E2 glycoprotein improves serum binding and cross-neutralization. *James Ou JH editor. J Virol*. (2020) 94:e00704–20. doi: 10.1128/JVI.00704-20
189. Machen AJ, O'Neil PT, Pentelute BL, Villar MT, Artigues A, Fisher MT. Analyzing dynamic protein complexes assembled on and released from biolayer interferometry biosensor using mass spectrometry and electron microscopy. *J Vis Exp*. (2018) 138:57902. doi: 10.3791/57902
190. Aydin S, Emre E, Ugur K, Aydin MA, Sahin İ, Cinar V, et al. An overview of ELISA: a review and update on best laboratory practices for quantifying peptides and proteins in biological fluids. *J Int Med Res*. (2025) 53:03000605251315913. doi: 10.1177/03000605251315913

191. Toseland CP. Fluorescent labeling and modification of proteins. *J Chem Biol.* (2013) 6:85–95. doi: 10.1007/s12154-013-0094-5
192. Ta HT, Peter K, Hagemeyer CE. Enzymatic antibody tagging: toward a universal biocompatible targeting tool. *Trends Cardiovasc Med.* (2012) 22:105–11. doi: 10.1016/j.tcm.2012.07.004
193. Wang HL, Kajbaf K, Gau BC, Dawdy AW, Edwards R, Bare B, et al. A novel *in vitro* expression assay by LC/MS/MS enables multi-antigen mRNA vaccine characterization. *Sci Rep.* (2025) 15:10336. doi: 10.1038/s41598-025-94616-8
194. Sanjuan Nandin I, Fong C, Deantonio C, Torreno-Pina JA, Pecetta S, Maldonado P, et al. Novel *in vitro* booster vaccination to rapidly generate antigen-specific human monoclonal antibodies. *J Exp Med.* (2017) 214:2471–90. doi: 10.1084/jem.20170633
195. Gong S, Beukema M, De Vries-Idema J, Huckriede A. Assessing human B cell responses to influenza virus vaccines and adjuvants in a PBMC-derived *in vitro* culture system. *Vaccine.* (2025) 44:126563. doi: 10.1016/j.vaccine.2024.126563
196. Du L, Zhao G, Yang Y, Qiu H, Wang L, Kou Z, et al. A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in middle east respiratory syndrome coronavirus spike protein. *Perlman S editor. J Virol.* (2014) 88:7045–53. doi: 10.1128/JVI.00433-14
197. Atanasova M, Bagdonas H, Agirre J. Structural glycobiology in the age of electron cryo-microscopy. *Curr Opin Struct Biol.* (2020) 62:70–8. doi: 10.1016/j.sbi.2019.12.003
198. Valverde P, Quintana JI, Santos JI, Ardá A, Jiménez-Barbero J. Novel NMR avenues to explore the conformation and interactions of glycans. *ACS Omega.* (2019) 4:13618–30. doi: 10.1021/acsomega.9b01901
199. Chua EYD, Mendez JH, Rapp M, Ilca SL, Tan YZ, Maruthi K, et al. Better, faster, cheaper: recent advances in cryo-electron microscopy. *Annu Rev Biochem.* (2022) 91:1–32. doi: 10.1146/annurev-biochem-032620-110705
200. Drulyte I, Hurdiss DL. Cryo-electron microscopy-based epitope mapping. In: Rockberg J, Nilvebrant J, Malm M, Thalén NB, editors. *Epitope mapping protocols.* Springer US, New York, NY (2025). p. 61–77. (Methods in Molecular Biology; vol. 2937).
201. Wu M, Lander GC. How low can we go? Structure determination of small biological complexes using single-particle cryo-EM. *Curr Opin Struct Biol.* (2020) 64:9–16. doi: 10.1016/j.sbi.2020.05.007
202. Pratap PP, Cottrell CA, Quinn J, Carnathan DG, Bader DLV, Tran AS, et al. Immunofocusing on the conserved fusion peptide of HIV envelope glycoprotein in rhesus macaques. *BioRxiv Prepr Serv Biol.* (2024) 2024:11.27.625755. doi: 10.1101/2024.11.27.625755
203. Verdager N, Ferrer-Orta C, Garriga D. X-ray crystallography of viruses. In: Mateu MG, editor. *Structure and physics of viruses.* Springer Nature Switzerland, Cham (2024). p. 135–69. Subcellular Biochemistry; vol. 105.
204. Domsic J, McKenna R. Structural studies of viral proteins – X-ray crystallography. In: Agbandje-McKenna M, McKenna R, editors. *Structural virology.* London, United Kingdom: The Royal Society of Chemistry (2010). p. 121–34.
205. Fernández de Toro B, Peng W, Thompson AJ, Dominguez G, Cañada FJ, Pérez-Castells J, et al. Avenues to characterize the interactions of extended N-glycans with proteins by NMR spectroscopy: the influenza hemagglutinin case. *Angew Chem Int Ed.* (2018) 57:15051–5. doi: 10.1002/anie.201807162
206. Wanzeck K, Boyd KL, McCullers JA. Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. *Am J Respir Crit Care Med.* (2011) 183:767–73. doi: 10.1164/rccm.201007-1184OC
207. Hierholzer JC, Suggs MT, Hall EC. Standardized viral hemagglutination and hemagglutination-inhibition tests. II. Description and statistical evaluation. *Appl Microbiol.* (1969) 18:824–33. doi: 10.1128/am.18.5.824-833.1969
208. Chitnis CE, Miller LH. Identification of the erythrocyte binding domains of *Plasmodium vivax* and *Plasmodium knowlesi* proteins involved in erythrocyte invasion. *J Exp Med.* (1994) 180:497–506. doi: 10.1084/jem.180.2.497
209. Ertelt M, Mulligan VK, Maguire JB, Lyskov S, Moretti R, Schiffrer T, et al. Combining machine learning with structure-based protein design to predict and engineer post-translational modifications of proteins. *PLoS Comput Biol.* (2024) 20:e1011939. doi: 10.1371/journal.pcbi.1011939
210. Eckard AD, Dupont DR, Young JK. Development of two analytical methods based on reverse phase chromatographic and SDS-PAGE gel for assessment of deglycosylation yield in N-glycan mapping. *BioMed Res Int.* (2018) 2018:3909674. doi: 10.1155/2018/3909674
211. He S, Liu X, Wu D, Zhou X. Detection of de-N-glycosylated EPO with SDS-PAGE: A complementary confirmation procedure for recombinant EPO in blood samples. *Drug Test Anal.* (2022) 14:1974–83. doi: 10.1002/dta.3324
212. Meyer HP RD. Microbial expression systems and manufacturing from a market and economic perspective. In: Agbo EC, editor. *Innovations in biotechnology.* London, United Kingdom: InTech (2012).
213. Lebozec K, Jandrot-Perrus M, Avenard G, Favre-Bulle O, Billiald P. Quality and cost assessment of a recombinant antibody fragment produced from mammalian, yeast and prokaryotic host cells: A case study prior to pharmaceutical development. *New Biotechnol.* (2018) 44:31–40. doi: 10.1016/j.nbt.2018.04.006
214. Batista AC, Soudier P, Kushwaha M, Faulon J. Optimising protein synthesis in cell-free systems, a review. *Eng Biol.* (2021) 5:10–9. doi: 10.1049/enb2.12004