Title: Molecular mechanisms of immune evasion by host protein glycosylation of a bacterial immunogen used in nucleic acid vaccines

Authors: Cinar, M.S., Adams, T.M., Ozdilek, A., and Avci, F.Y.*

Department of Biochemistry, Emory Vaccine Center, Emory University School of Medicine, Atlanta, GA, 30322 USA

* Correspondence to: Fikri Y. Avci, Email: favci@emory.edu

Keywords: mRNA vaccines, nucleic acid vaccines, mammalian N-glycosylation, glycomics, glycopeptidomics, molecular dynamics, *Mycobacterium tuberculosis*, antigenicity, immunogenicity

Author List + Contributions:

Mukaddes Sena Cinar – design, data acquisition, analysis, interpretation, writing, and reviewing

Trevor Adams – design, data acquisition, analysis, interpretation, writing, and reviewing

Ahmet Ozdilek – conceptualization and design

Fikri Y. Avci – conceptualization, design, interpretation, writing, reviewing, and correspondence

Abstract:

Nucleic acid vaccines, including mRNA platforms, represent a transformative approach by inducing adaptive immunity against *in situ*-expressed antigens. Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), remains a leading global health threat. The only licensed TB vaccine, Bacillus Calmette-Guérin (BCG), offers limited protection against pulmonary TB in adults and adolescents. Underscoring the urgent need for more effective vaccination strategies. However, clinical trials targeting the *Mtb* Ag85 complex proteins as nucleic acid vaccines have failed to confer protective immunity. We hypothesized that the glycosylation of *Mtb* immunogens expressed in mammalian hosts as a post-translational modification impairs their immunogenicity and antigenicity. Here, we show that Ag85B expressed in human cells undergoes aberrant N-glycosylation, leading to significantly dampened antigenicity compared to its native, non-glycosylated bacterial counterpart. Site-directed mutagenesis to eliminate N-glycosylation motifs restored immune recognition and response. We then elucidated the molecular and cellular mechanisms by which host-derived glycosylation can impair the efficacy of nucleic acid vaccines, particularly targeting non-viral pathogens. This insight provides a critical foundation for the rational design of next-generation nucleic acid vaccines, where controlling or circumventing undesirable post-translational modifications may be essential to elicit robust protective immunity.

32 **Introduction:**

33 Nucleic acid vaccines, including DNA, recombinant vector, and mRNA platforms, represent a next-
34 generation immunization strategy that diverges from conventional vaccine modalities(*1*). Instead of delivering
35 whole pathogens or purified protein antigens, these vaccines introduce the genetic sequences encoding
36 pathogen-associated proteins directly into host cells, expressing the target antigens *in situ*. This platform offers
37 significant advantages in manufacturing scalability, flexibility, and rapid adaptability, particularly in response to
38 emerging infectious diseases(*2*).

39 While nucleic acid vaccines have long been studied in both human and veterinary contexts, they were
40 not approved for human use until the global urgency of the COVID-19 pandemic catalyzed the clinical
41 deployment of several mRNA-based vaccines(*3-6*). This landmark success has revitalized interest in using
42 nucleic acid vaccine platforms for prokaryotic and eukaryotic pathogens.

43 However, an essential immunological distinction emerges when applying nucleic acid vaccine technology
44 to non-viral pathogens. Unlike viral proteins, which are naturally synthesized within host cells during infection,
45 non-viral antigens are not expressed in host cells upon infection. We postulated that employing nucleic acid
46 vaccines against non-viral pathogens would cause the proteins expressed by host cells to be structurally and
47 immunologically different from native pathogen proteins. This mismatch arises from the host's post-translational
48 modifications (PTMs), especially glycosylation for targets entering the secretory pathway. Glycans may obscure
49 or alter immunogenic epitopes and/or engage immune regulatory pathways that suppress effective responses(*7,*
50 *8*). Such host-derived glycans could therefore compromise vaccine efficacy by reducing the quality and
51 magnitude of immune recognition and function.

52 This concept is particularly relevant for tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*),
53 which remains one of the deadliest infectious diseases worldwide(*9, 10*). With an estimated one-third of the
54 world's population infected by *Mtb,* 10.8 million new TB cases, and 1.25 million deaths annually, TB presents an
55 urgent global health challenge exacerbated by the emergence of multidrug- and extensively drug-resistant *Mtb*
56 strains(*9, 11*). Current treatments are long, costly, and often not curative, highlighting the need for more effective
57 vaccines(*9, 12, 13*). The only approved vaccine, Bacillus Calmette-Guérin (BCG), is an attenuated
58 *Mycobacterium bovis* strain introduced nearly a century ago and confers limited protection(*14*). It particularly fails
59 to prevent adult pulmonary TB, necessitating the development of protective, new-generation vaccines(*15-18*).

60 Among the most promising TB vaccine targets employed in multiple vaccine clinical trials are antigen 85
61 (Ag85) proteins, Ag85A and Ag85B—fibronectin-binding mycolyltransferases critical to *Mtb* cell wall assembly
62 and survival(*19*). These proteins are abundantly expressed on the outermost surface of *Mtb*, highly
63 immunogenic, and essential for intracellular replication, making them high-priority targets in multiple clinical
64 vaccine trials(*20*). However, despite their promise, Ag85A and Ag85B-based vaccines delivered as nucleic acid
65 vaccines failed to induce protective immunity in several clinical trials(*21-26*).

66 Our prior work revealed that Ag85A produced in mammalian expression systems undergoes N-
67 glycosylation, which significantly dampens its immunogenicity compared to the native *Mtb*-expressed form(*27*).
68 Building on this foundational work, we examined Ag85B, the other highly targeted *Mtb* antigen in nucleic acid
69 vaccine development, to elucidate the molecular mechanisms of immune evasion caused by host glycosylation
70 and to inform the rational design of future nucleic acid vaccines for non-viral pathogens.

Results:

Ag85B is microheterogeneously glycosylated at four canonical N-linked glycosylation sequons when expressed in human Expi293 cells.

Mammalian N-glycosylation takes place canonically at N-x-S/T(C) sequons, where "x" may be any amino acid except proline. *Mycobacterium tuberculosis* Ag85B (*Mtb* Ag85B) is not N-glycosylated when expressed in its native form; however, the protein sequence predicts that four canonical N-glycosylation sites may be glycosylated if expressed in mammalian hosts **(Fig. 1a)**. To mimic the protein expressed by nucleic acid vaccines targeting Ag85B, we expressed the Ag85B protein (293-F Ag85B) in Expi293 cells, a variant of a human embryonic kidney cell line (HEK293-F), and investigated its potential glycosylation.

To assess the N-glycosylation of 293-F Ag85B, we first conducted N-glycan site-occupancy analysis using nLC-MS/MS mass spectrometry and an EndoH/PNGase F glycosidase with heavy water ($H_2{}^{18}O$) labeling scheme(*28*). This method simplifies N-glycan analysis by binning N-glycan sites into one of three potential groups: unoccupied (+0 mass shift), high-mannose/hybrid (+3 mass shift), and complex (+203 mass shift). The differentiation between high-mannose/hybrid and complex sites is based on EndoH's recognition of terminal mannose groups. Our study confirmed N-glycan presence in the four canonical sequons: N52, N224, N234, and N280 **(Fig. 1a)**. Our analysis revealed that N-glycans heavily occupy these four sites, with the lowest occupancy being 68.1% at site N52 (sequon N_{52}NS) and the highest occupancies being 95.8% at both sites N224 (sequon N_{224}NT) and N280 (sequon N_{280}GT) **(Fig. 1b)**. This also aligns with the known phenomenon that NxT sites are favored by oligosaccharyltransferase (OST) over NxS sites(*29*).

Glycomics analysis using LC-MS of 293-F Ag85B revealed the heterogeneous population of glycoforms on the protein. A total of 75 different glycan compositions were detected, 18 of which are present with greater than 1% of total glycan abundance **(Fig. 1c)**. Additionally, 72.0% of all compositions were identified as complex glycoforms **(Fig. 1d)**. The most abundant composition identified is the complex type N-glycan HexNAc(4)Hex(3)Fuc(1), accounting for about 14% of all glycoforms present. Moreover, several structures with greater than 1% abundance have varying amounts of sialylation and core fucosylation **(Fig. 1c)**. This data is generally in agreement with the site occupancy results, with the discrepancies between the two experiments likely due to the hybrid glycan classification scheme used in our glycomics approach.

98 In addition to the classical roles of N-glycans in contributing to glycoprotein structure and stability,
99 Sialylation and fucosylation of N-glycans play a part in immunoregulation, including self-recognition, in the human
100 immune system(*8, 30, 31*) , with particular importance placed on the interaction of Siglecs with sialic acids.
101 Furthermore, bacterial species utilize terminal sialic acid and fucose moieties for “molecular mimicry” to evade
102 immune responses(*8, 32-34*). To assess the presence of these immunoregulatory terminal sugars, we analyzed
103 the extent of sialylation and fucosylation these proteins underwent during post-translational modifications. Of all
104 mammalian-expressed Ag85B glycans, 46% are sialylated and 91% are fucosylated. The prevalence of such
105 moieties on the surface of the mammalian-expressed Ag85B may alter its antigenicity **(Fig. 1e-1f)**.

106 To better understand the specific glycan profile at each identified sequon, we performed glycomics-
107 informed glycopeptidomics analysis using mass spectrometry(*35*). A representative MS2 spectrum of the N52
108 glycan site demonstrates the rich fragmentation generated from our stepped high-energy collisional dissociation
109 (sHCD), which provides functional fragments both of the peptide backbone and also of the glycan itself in a single
110 MS2, which are used by the search engine pGlyco3 to identify glycopeptides **(Supplementary Fig. 1)**.

111 N-glycan microheterogeneity (site-specific diversity and distribution) was observed at all sequons. The
112 most abundant glycan at site N52 was the high-mannose structure HexNAc(2)Hex(5) **(Supplementary Fig. 2a)**,
113 which agrees with site occupancy data showing an abundance of terminal mannose structures at that site. Site
114 N224 contrasts this with abundant complex glycosylation and complete fucosylation of non-oligomannose
115 structures **(Supplementary Fig. 2b)**. At site N234, a mixture of oligomannose, hybrid, and complex structures
116 coexist with the more processed glycans appearing to be biantennary fucosylated strcutures **(Supplementary**
117 **Fig. 2c)**. Finally, site N280’s glycopeptide analysis data contradicts the site occupancy data with the most
118 common glycoform being the oligomannose HexNAc(2)Hex(5) N-glycan, which commonly forms a bottleneck at
119 MGAT1 processing **(Supplementary Fig. 2d)**. One explanation for this is the higher branching at this site (as
120 evidenced by more HexNAc groups) ‘spreads’ the glycoform distribution across a broad range of potential
121 structures, diluting the signal. Ion suppression may also play a role, mainly since this site differs from the others
122 as it requires a non-tryptic peptide, which is not ideal for analysis. However, the previously observed patterns of
123 high fucosylation levels and abundant sialylation among the more processed glycans remain consistent.
124

Molecular dynamics simulations reveal the detrimental effect of N-glycosylation on the surface accessibility of Ag85B.

To better illuminate the significant impact that N-glycosylation may have on the Ag85B structure and surface accessibility, we performed molecular dynamics (MD) simulations. The glycosylated protein (293-F Ag85B) was generated using GLYCAM(*36, 37*) to add site-specific, abundant glycan structures, identified through glycomics-informed glycopeptidomics mass spectrometry data **(Supplementary Fig. 2)**, to a previously published *Mtb* Ag85B X-ray diffraction structure, 1F0N(*38*). As a control and for comparison, we also simulated the native, nonglycosylated protein structure. After confirming the sampling quality of all three production replicates **(Supplementary Fig. 3)**, we evaluated surface accessibility to water and larger molecules, such as antibodies, using various probe sizes and software tools.

First, solvent-accessible surface areas (SASAs) were assessed using a 1.4 Å probe corresponding to the radius of a water molecule, with CPPTRAJ(*39*). SASA values were calculated for each frame across the 1-microsecond production replicates and plotted over time **(Fig. 2a)**. Throughout the MD simulation, the glycosylated protein consistently exhibited a significant reduction in water accessibility across nearly all conformations sampled when compared with the nonglycosylated protein. Although one replicate from each condition showed some overlap, these points correspond to the lowest accessibility values in the nonglycosylated protein and the highest in the glycosylated protein. When re-examined as distributions in violin plots, the reduction in solvent accessibility in the glycosylated protein becomes even more apparent **(Fig. 2b)**. The mean SASA for the nonglycosylated protein was 116 nm^2. In comparison, that of the glycosylated protein was 107 nm^2. Overall, as demonstrated by the mean SASA plots, the two proteins show distinct accessibility profiles **(Fig. 2a-2b)**.

Next, to assess antibody-accessible surface areas (ASAs), we used a larger 10 Å probe with NACCESS, which provides residue-specific accessibility calculations(*40, 41*). To simplify NACCESS analysis, we used hierarchical clustering to reduce the glycosylated protein trajectory to ten representative structural conformations **(Supplementary Fig. 4a-4b)**. Figure 2c illustrates various angles of each glycan's conformational flexibility across these ten structures, highlighting the vast three-dimensional space that glycans can occupy on the protein surface over time. When assessing residue-specific ASAs, regions of the protein backbone closest to glycans consistently showed decreased accessibility. In contrast, residues distant from glycosylation sites either exhibited

153 no change or increased accessibility **(Fig. 2c)**. Together, our SASA and ASA results demonstrate that N-glycans
154 occupy substantial spatial regions on the surface of Ag85B, leading to measurable losses in solvent and antibody
155 accessibility. These structural alterations may influence how epitopes are exposed or shielded, potentially
156 affecting recognition by humoral immune components. This may result in the shielding of vulnerable antigenic
157 targets and the exposure of non-productive epitopes.
158

159 **Molecular dynamics simulations illustrate the detrimental effect of Ag85B N-glycosylation on**
160 **established antibody binding sites.**

161 Humoral immunity is the key mediator by which vaccines achieve “sterilizing" immunity against
162 pathogens(*42*). One common immune evasion strategy employed by pathogens is the use of “glycan shields,”
163 where glycans sterically obstruct antibody binding by masking surface-exposed epitopes(*43, 44*). Hence, to
164 investigate whether glycosylation similarly impairs antibody access, we utilized previously identified antibody
165 binding sites against Ag85B(*45*). These sites were identified using proteolytic affinity-mass spectrometry and
166 biosensor analysis against human serum samples from healthy, vaccinated, and infected donors.

167 The identified antibody-binding sites were mapped onto the glycosylated Ag85B ensemble structure
168 generated from the molecular dynamics simulations **(Fig. 3a)**. When the antibody binding sites were located
169 distantly from glycosylation sites, little to no loss in accessibility to a 10 Å probe was observed, as demonstrated
170 by sites S23–Y31 and L59–W72 **(Fig. 3b-3c)**. However, as antibody sites approached glycosylated residues,
171 greater loss of accessibility was observed, with several residues within each site exhibiting a significant reduction
172 in accessibility to the 10 Å probe **(Fig. 3d-3h)**.

173 Notably, antibody binding site K260-F275, although not directly adjacent to a glycosylation site, contains
174 at least three residues with a significant loss of accessibility **(Fig. 3f)**. In addition, glycosylation site N280 is
175 located within antibody binding site N276–Y287, where many residues show a pronounced decrease in surface
176 accessibility. Compared to sites A157–F164 (21.6–97.3% loss of accessibility) and K260–F275 (4.3–63.6% loss),
177 both of which are not directly at a glycosylation site, site N276–Y287 exhibits a far more extensive loss, ranging
178 from 81.3 to 99.9% **(Fig. 3g)**. Antibody binding site W288–G306 is also significantly affected by the nearby N280
179 glycan even though it does not contain a glycosylation site within the epitope. Most exposed surface residues in
180 this site lose accessibility to a 10 Å probe, with reductions ranging from 45.8 to 99.8% **(Fig. 3h)**. The significant

181 loss of accessibility observed across multiple antibody binding sites, either in proximity to or directly overlapping
182 glycosylation sites, highlights the substantial role glycans play in shielding key epitopes from recognition by the
183 adaptive immune response.
184
185 **N-glycans mask known T cell epitopes of Ag85B.**
186 Similar to how unnatural mammalian host glycans can shield antibody binding regions on the native
187 protein, they can also impact cellular adaptive immunity through several mechanisms: (1) by masking protease
188 cleavage sites and preventing proper processing of MHC epitopes, (2) by interfering with or altering MHC
189 molecule binding in the microsomes, (3) through the presentation of glycosylated peptides to T cells, which can
190 modulate T cell responses differently from the native peptides, and (4) by inducing deamidation of asparagine to
191 aspartic acid during antigen processing, leading to the presentation of epitopes that differ from those derived
192 from the native protein(*46, 47*).
193 To assess how glycosylation affects cellular immunity, we systemically mapped previously identified MHC
194 binding epitopes onto the glycosylated Ag85B ensemble structure generated from molecular dynamics
195 simulations **(Fig. 4a**(*48*). We hypothesized that sites proximal to non-native N-glycosylation would exhibit
196 reduced accessibility. For example, the peptide A111–E125, identified as presented by DPB1*04:01 during
197 human latent TB infection(*49*), is not located near any glycosylation sites **(Fig. 4a)**. When examining its
198 accessible surface area using a 10 Å probe, most residues in this peptide exhibit minimal variation **(Fig. 4 b)**,
199 with two residues showing significant but mild changes in accessibility (C113 decreasing, A114 increasing). In
200 contrast, F121–P138, another T cell peptide identified during latent TB infection(*50, 51*), lies in close proximity
201 to two glycosylation sites (N52 and N224). Several of its surface-accessible residues exhibit a substantial
202 decrease in accessibility upon glycosylation, ranging from 20.6 to 99.5% **(Fig. 4c)**. Unlike A111–E125, no
203 residues within F121–P138 display increased accessibility. These data suggest that proximity to non-native N-
204 glycosylation may hinder the accessibility of epitopes.
205 Two additional Ag85B peptides, F164–S171 and K220–L228, were identified using human MHC
206 transgenic mice and are presented to T cells via HLA-A*02:01 in a DNA vaccination model **(Fig. 4d–4e)**[52].
207 Peptides presented by HLA-A*02:01 are of particular interest due to the high prevalence of this haplotype in the
208 global population, especially among individuals of Caucasian descent(*53, 54*). Glycosylation of Ag85B

209 significantly affects the accessibility of both peptides. For F164–S171, only one residue is surface-accessible,
210 and this accessibility is completely lost upon the addition of a glycan at N224 **(Fig. 4d)**. In comparison, K220–
211 L228 contains multiple accessible residues, many of which show a loss of accessibility ranging from 41.4 to
212 100% when N224 is glycosylated **(Fig. 4e)**.

213 Finally, using a reverse immunogenetics approach, the peptide I245–F253 was identified as an HLA-
214 B*35-restricted epitope in humans(*55*). Within this peptide, several residues, particularly those near the N234
215 glycan, exhibit a marked reduction in surface accessibility, ranging from 45.6-100% **(Fig. 4f)**. Together, the
216 substantial loss of accessibility observed in several T cell epitopes, whether adjacent to or directly overlapping
217 with glycosylation sites, demonstrates the potential for glycans to interfere with epitope processing and
218 presentation. In some cases, glycopeptides may be presented in place of the native epitope, with important
219 implications for T cell recognition and immunogenicity.

220

221 **Glycosylation of Ag85B negatively impacts antigenicity, which is rescued by the non-glycosylated**
222 **mutant variant.**

223 Using *in silico* approaches, we predict that glycosylation significantly impairs both antibody and T cell
224 epitopes of Ag85B. We evaluated the effect in vitro by using a monoclonal antibody targeting the Ag85 complex
225 to assess whether glycosylation impacts antibody binding. We performed a competitive ELISA in which plates
226 were coated with bacterially expressed Ag85B, and the ability of Ag85B variants to compete for antibody binding
227 was evaluated. We found that glycosylation of Ag85B significantly reduces its ability to inhibit antibody binding
228 to the bacterial protein. Conversely, when all glycosylation sites were removed via asparagine-to-glutamine
229 mutations within the N-x-S/T sequon (293-F dAg85B), this inhibitory capacity was successfully rescued, reaching
230 levels comparable to bacterial-expressed Ag85B **(Fig. 5a)**.

231 To further characterize these interactions, we employed biolayer interferometry to measure the binding
232 kinetics of the monoclonal antibody against the different Ag85B variants **(Fig. 5b-5e)**. The bacterially expressed
233 Ag85B, representing the wild type binding profile expected to be seen in natural infections, exhibited a robust
234 wavelength shift of approximately 0.30 nm during association, which returned to baseline upon dissociation **(Fig.**
235 **5b)**. In contrast, the glycosylated Ag85B showed a substantially reduced shift of approximately 0.05 nm,
236 indicating a diminished association rate **(Fig. 5c)**. The engineered 293F dAg85B mutant restored binding **(Fig.**

237 **5d)**. Bacterial Ag85B exhibited a dissociation constant of 53.61 nM, glycosylated Ag85B had three orders of
238 magnitude weakened binding affinity with a dissociation constant of 29.80 mM, and the deglycosylated mutant
239 demonstrated a rescued affinity of 219.0 nM comparable to the bacterial Ag85B **(Fig. 5e)**.

240 **Discussion:**

241 The success of COVID-19 mRNA vaccines has underscored the potential of nucleic acid vaccine
242 technologies as powerful prophylactic tools. This breakthrough has spurred widespread interest in applying
243 mRNA and DNA platforms to a broad spectrum of infectious diseases. However, while nucleic acid vaccines are
244 well-suited to viral antigens naturally synthesized in host cells, their application to non-viral pathogens, such as
245 bacteria and parasites, presents unique biological challenges. Notably, when these foreign antigens are
246 expressed in mammalian cells, they may acquire host-derived post-translational modifications (PTMs),
247 particularly N-linked glycosylation, that are absent in their native pathogen context. These glycan modifications
248 can mask critical epitopes or engage immune regulatory pathways, ultimately dampening the desired immune
249 response. Such unintended alterations could undermine the efficacy of nucleic acid vaccines targeting non-viral
250 pathogens. Understanding and accounting for these host-mediated modifications is therefore essential to
251 optimizing antigen expression and immunogenicity in next-generation nucleic acid vaccine platforms.

252 Importantly, the role of glycosylation in shaping immune responses is not limited to non-viral
253 pathogens(*56, 57*). In viral vaccine targets such as the HIV-1 envelope glycoprotein (Env), dense glycan
254 shielding on the gp120 subunit has been shown to hinder antibody access to conserved neutralizing epitopes(*58,*
255 *59*). Studies have demonstrated that selective removal of N-glycosylation sites can unmask these epitopes,
256 directing the immune system toward the generation of broadly neutralizing antibodies(*41, 60, 61*). Similarly, in
257 SARS-CoV-2, the spike protein's glycosylation landscape influences antigenicity and immune targeting(*62, 63*).
258 Recent work has shown that eliminating specific glycosylation sites on the spike can enhance the breadth and
259 potency of the antibody response(*64-66*). These findings across both viral and non-viral systems highlight that
260 glycosylation is not a peripheral concern but a central determinant of vaccine efficacy. As such, nucleic acid
261 vaccine design must move beyond sequence selection to consider how expression in mammalian hosts alters
262 antigen structure and immune presentation. Incorporating glycosylation-aware design strategies will be critical
263 to unlocking the full potential of nucleic acid vaccines across diverse pathogen classes.

264 Our investigation into the *Mtb* antigen, Ag85B, provides a direct and mechanistically informative example
265 of this principle. Ag85B is a secreted, highly immunogenic protein critical for *Mtb* cell wall biosynthesis and has
266 been widely used in TB vaccine candidates, including those delivered via nucleic acid platforms(*19, 21-23, 26*).
267 In its native bacterial context, Ag85B is not glycosylated. However, when expressed in mammalian host cells

268 during nucleic acid vaccination, Ag85B acquires N-linked glycans at specific asparagine residues. We
269 demonstrated that this aberrant glycosylation significantly reduces its antigenicity and dampens immune
270 activation. Importantly, removal of these glycosylation sites via site-directed mutagenesis restored immune
271 recognition and response, confirming that the glycan moieties themselves were directly responsible for immune
272 evasion.

273 At a structural level, these glycans likely occlude key B-cell or T-cell epitopes, hinder antigen processing
274 by antigen-presenting cells, or alter MHC loading and presentation, mechanisms well-documented in other
275 glycan-shielded antigens such as HIV Env(*56, 58, 61, 67*). Glycans can also engage inhibitory immune receptors
276 (e.g., DC-SIGN, Siglecs), actively suppressing immune responses(*31, 68, 69*). Understanding these molecular
277 mechanisms, including how host-derived glycosylation alters epitope exposure, antigen stability, and receptor
278 interactions, is critical not only for optimizing individual antigens but for establishing design principles that guide
279 antigen engineering across vaccine platforms. A previous study demonstrated the power of informed protein
280 engineering in glycoconjugate vaccine design by utilizing glycan conjugation on Ag85B via lysine residues.
281 Targeted mutations prevented conjugation at key epitopes, preserving T cell function while maintaining glycan
282 conjugation at non-critical sites(*70*).

283 While N-linked glycosylation is a well-characterized and structurally defined post-translational
284 modification with clear implications for antigen recognition and immune activation, O-linked glycosylation may
285 also modulate immune responses. O-glycans, typically added to serine or threonine residues, can influence
286 protein conformation, mucosal immunity, and receptor engagement(*31, 71*). However, O-linked glycosylation is
287 inherently more heterogeneous and less predictable than N-linked glycosylation, with no strict consensus
288 sequence and extensive variation in glycan composition and site occupancy(*71*). Moreover, the biosynthesis of
289 O-glycans is less stringently regulated, often varying with cell type, activation state, and even environmental
290 conditions. This lack of predictability makes both the identification of relevant O-glycosylation sites and their
291 targeted modulation technically challenging. Given these complexities, prioritizing O-glycosylation for
292 immunogen engineering may not be practical unless N-glycan modulation fails to restore immunogenicity.
293 However, characterization of potential O-glycosylation sites via approaches such as mass spectrometry or lectin
294 blots may be an important but overlooked step in characterizing engineered antigens. In this work, we specifically

295 focus on N-linked glycosylation due to its defined structural rules, clearer immunological impact, and greater
296 potential for rational design in nucleic acid vaccine contexts.

297 In this study, we provide mechanistic insight into how host-cell-mediated N-glycosylation impairs the
298 antigenicity of Ag85B, a major TB vaccine candidate antigen. When expressed in human Expi293 cells, Ag85B
299 is microheterogeneously glycosylated at four canonical N-linked glycosylation sequons, leading to altered
300 structural and immunological properties. We demonstrated, using molecular dynamics simulations, that these
301 glycans significantly impede the surface accessibility of Ag85B, including the occlusion of known B-cell and T-
302 cell epitopes, thereby hindering recognition by the immune system. Functional assays further confirmed that
303 glycosylated Ag85B shows reduced binding to an Ag85B-specific monoclonal antibody, as measured by ELISA
304 and biolayer interferometry (BLI). Importantly, this dampened antigenicity was effectively rescued by a non-
305 glycosylated mutant variant of Ag85B, highlighting that N-glycans are directly responsible for impaired immune
306 engagement. Thus, structural insights into host-imposed antigen remodeling underscore the potential need for
307 a paradigm shift in nucleic acid vaccine design, one that fully integrates biochemical and immunological
308 considerations to avoid glycan-induced immune silencing.

309 This study has wide-reaching implications for the rational design of nucleic acid vaccines primarily aimed
310 at non-viral pathogens, while also significantly impacting viral pathogens. As nucleic acid platforms drive antigen
311 expression in mammalian cells, unintended post-translational modifications, particularly N-linked glycosylation,
312 may structurally remodel microbial proteins in ways that diminish their immunogenic potential. Our work
313 emphasizes the importance of evaluating and, when necessary, modifying candidate antigens to preserve
314 epitope integrity and immune visibility. Taken together, our findings support the need for glycosylation-aware
315 design principles in future nucleic acid vaccine engineering and underscore the importance of structural and
316 molecular immunology in bridging antigen expression to immune activation.

317 **Methods:**

318 **Proteins and Bacteria.**

319 *Mycobacterium tuberculosis* H37Rv strain was acquired from the Biodefense and Emerging Infections
320 (BEI) Research Resources Repository of the National Institute of Allergy and Infectious Diseases at the National
321 Institutes of Health (NIH). Native Ag85B proteins were purified from the *Mtb* strain H37Rv.

322

323 **Expression and purification of *E. coli* BL21-expressed recombinant Ag85B.**

324 Plasmids with the Ag85B gene were generated using polymerase chain reactions (PCR) and DNA
325 assembly techniques (NEBuilder HiFi DNA Assembly, New England Biolabs) to insert the gene into a pET-28b(+)
326 vector backbone with an N-terminal His-tag and TEV cleavage site. All PCR reactions were confirmed with gel
327 electrophoresis techniques using 1% agarose in Tris-acetate-EDTA buffer gel (Addgene). PCR reactions were
328 purified using the PCR purification methods (Invitrogen PureLink Quick PCR Purification Kit).

329 Plasmids were transformed into *Escherichia coli* (*E. coli*) BL21 cells and incubated at 37ºC with an orbital
330 shaking speed of 200 rpm until an OD of 0.7-1.0. Protein expression was induced with 1 mM isopropyl β-D-
331 thiogalactopyranoside (IPTG) and incubated at 18ºC with an orbital shaking speed of 200 rpm for 24 hours before
332 harvesting cells for protein purification. All proteins were purified using IMAC columns (BioRad, Cat. # 12009287)
333 and then separated by a HiLoad Superdex 200 pg preparative size-exclusion column (Cytiva, Cat. # 28989335)
334 using a BioRad fast-protein liquid chromatography. Proteins were validated using mass spectrometry.

335

336 **Expression and purification of mammalian-expressed recombinant Ag85B.**

337 Plasmids with the Ag85B variant gene insertions were first generated for mammalian cell protein
338 expression. *Homo sapiens* codon optimized genes with the tissue plasminogen activator leader signal sequence
339 at the N-terminal and a TEV cleavage site with an 8xHis tag at the C-terminal were synthesized by Twist
340 BioSciences. 293-F Ag85B synthesized gene was inserted into a pGEc2-DEST vector with a CMV promoter
341 using PCR and DNA assembly technique (NEBuilder HiFi DNA Assembly, New England Biolabs). All PCR
342 reactions were confirmed with gel electrophoresis techniques using 1% agarose in Tris-acetate-EDTA buffer gel
343 (Addgene). PCR reactions were purified using the PCR purification methods (Invitrogen PureLink Quick PCR

344 Purification Kit). Mutant variant genes were cloned through Twist BioSciences. All plasmids were confirmed
345 through sequencing methods (Eurofins).
346 Recombinant Ag85B proteins were expressed in Expi293F™ cells cultured in Expi293 expression
347 medium (Gibco, Cat. # A14635). Cells were transiently transfected with the ExpiFectamine™ 293 Transfection
348 Kit (Gibco, Cat. # A14524) or Polyethylenimine "Max" (PEI MAX) (Polysciences, CAS # 49553-93-7) following
349 the producer's protocol or an established protocol by Dr. Dumoux, respectively. All proteins were purified using
350 IMAC columns (BioRad, Cat. # 12009287) and then separated by a HiLoad Superdex 200 pg preparative size-
351 exclusion column (Cytiva, Cat. # 28989335) using a BioRad fast-protein liquid chromatography. Proteins were
352 validated using mass spectrometry.
353
354 **Released Ag85B N-glycan analysis (glycomics) via LC-MS.**
355 Glycans from 20µg purified recombinant Ag85B were released and labeled with InstantPC dye using the
356 Agilent AdvanceBio Gly-X N-glycan prep with InstantPC kit (Agilent, Part #: GX96-201PC) as described in the
357 kit instructions. 2µL of the resulting 100µL sample was injected onto a 2.1x150mm 1.8µm-particle size
358 AdvanceBio Amide HILIC column (Cat. #: 859750-913). Glycans were separated with a two-buffer LC gradient
359 using 50 mM ammonium formate as buffer A and acetonitrile with 0.1% formic acid as buffer B with an Agilent
360 1290 Infinity II liquid chromatography system. The buffer gradient proceeded as follows: 0-45min 23 to 44% A,
361 45-46min 44 to 56% A, 46-47min 56 to 60% A, 47-49min 60 to 23% A, 49-60min 23% A. Eluting glycans were
362 analyzed with an Agilent 6545XT quadrupole time-of-flight (QTOF) mass spectrometer with a dual Agilent Jet
363 Stream (AJS) electrospray ionization (ESI) source operating in positive ion mode. Raw files were analyzed
364 utilizing Skyline with a transition list derived from the human N-glycan compositions described in the GlyGen(*72*)
365 Database (accessed August 2024). Glycan compositions were relatively quantified by summing the peak area
366 of all observed adducts and isomers using Skyline(*73*) and Python. Figures were generated using GraphPad
367 Prism 9 software.
368
369 **N-linked glycopeptide and site occupancy analysis – Sample Preparation**
370 Intact glycopeptides were generated utilizing purified Ag85B. 50µg of purified glycoprotein was reduced
371 for one hour at 56°C in 10 mM DTT. Ag85B cysteines were alkylated using 25 mM iodoacetamide in the dark at

372 room temperature for 45 minutes. Alkylation was quenched with an additional 10 mM DTT. Glycoprotein was
373 then digested with one of two separate protease conditions: 1) α-lytic protease (NEB, Part #: P8113S), or 2)
374 trypsin/LysC mix (Promega, Part #: V5071) followed by GluC (Promega, Part #: V1651). All protease digestions
375 were carried out at 37°C overnight (~16-18 hours) at a 1:20 µg:µg protease:protein ratio (e.g., protease condition
376 2 required two overnight digests). After each digest, the protease was heat-inactivated at 95°C for 5 minutes.

377 Glycopeptides were desalted via reverse-phase cleanup using a BioPureSPN MINI C18 column (The
378 Nest Group, Part #: HUM S18V). Desalted peptides were dried down with a vacuum centrifuge and then
379 resuspended in a solution of 95.9% water, 4% acetonitrile, and 0.1% formic acid. Resuspended peptides were
380 filtered through a conditioned 0.2µm filter (PALL, Cat. #: ODM02C34) before transferring to autosampler vials.

381 For site occupancy analysis(*28*), desalted glycopeptides were instead resuspended in 50µL of a 50 mM
382 sodium phosphate buffer with 5µL/500 units EndoH (NEB, Part #: P0702S) at 37°C overnight, followed by the
383 addition of 50µL 1% trifluoroacetic acid (TFA) to stop the reaction. Partially deglycosylated peptides were then
384 subjected to another round of reverse phase cleanup (see above), then resuspended in 50µL of a 20 mM
385 ammonium bicarbonate in heavy-labeled ^{18}O water (Cambridge Isotope Laboratories, Part #: DLM-4-100).
386 Glycopeptides were treated with 2µL/20 units PNGaseF (Promega, Part #: V4831) overnight at 37°C.
387 Deglycosylated peptides were then subjected to a final round of reverse-phase cleanup and 0.2µm filtering (see
388 above) before transfer to autosampler vials.

389

390 **N-linked glycopeptide and site occupancy analysis – LC-MS/MS Analysis**

391 For intact glycopeptide analysis, 1µg of Ag85B glycopeptides was injected onto a 75µmx15cm Acclaim
392 PepMap 100 C18 nanoViper column (Thermo Fisher Scientific, Part #: 164940) and eluted into the nanospray
393 source of an Orbitrap Eclipse Tribrid mass spectrometer with a field asymmetric ion mobility spectrometry
394 (FAIMS) frontend interface (Thermo Fisher Scientific). Glycopeptides were separated with a two-buffer LC
395 gradient using water with 0.1% formic acid as buffer A and 80% acetonitrile with 0.1% formic acid as buffer B
396 utilizing a Vanquish liquid chromatography system (Thermo Fisher Scientific) with a flow rate of 0.4µL/min. The
397 90-minute buffer gradient proceeded: 0-6min 2.5% B, 6-76min 2.5 to 32.5% B, 76-88min 32.5% to 100% B,
398 followed by a minute of column wash of 100% B at 1µL/min. The spray voltage was set to 2.35 kV, and the ion
399 transfer tube was heated to 305°C. FAIMS compensation voltages were set to cycle through -35V, -40V, and -

400 50V. Full MS scans were acquired from m/z 200 to 2000 at 60,000 resolution in the Orbitrap. MS/MS scans were
401 triggered with an intensity threshold of 2×10^4 and a mass tolerance of 20 ppm. MS/MS scans were acquired at
402 a resolution of 15,000 in the Orbitrap via stepped higher collisional dissociation (stepHCD), with HCD collision
403 energies of 15%, 25%, and 35% and an isolation window of 1.6 m/z.

404 For site occupancy analysis(*28, 57*), 10µg of EndoH/PNGaseF-treated peptides were injected onto a
405 2.1x150mm 2.7µm pore size Peptide Mapping C18 column (AdvanceBio, Part #: 653750-902) eluted into the
406 dual Agilent Jet Stream (AJS) electrospray source of an Agilent 6560 mass spectrometer (Agilent). Peptides
407 were separated with a two-buffer LC gradient using water with 0.1% formic acid as buffer A and acetonitrile with
408 0.1% formic acid as buffer B utilizing a 1290 Infinity II liquid chromatography system (Agilent) with a flow rate of
409 0.4 mL/min. The 45-minute buffer gradient proceeded as follows: 0-2.5min 2.5% B, 2.5-4min 2.5 to 6% B, 4-
410 12min 6% to 22% B, 12-33min 22 to 29% B, 33-38min 29 to 34% B, 38-40min 34 to 81% B, 40-43min constant
411 81% B wash, then equilibration from 43-45min 81 to 2.5% B. Eluting glycans were analyzed with an Agilent 6560
412 (QTOF) mass spectrometer equipped with a dual AJS ESI source operating in QTOF-only positive ion mode.
413 MS1 full scans were acquired from 100-3000 m/z, with an acquisition rate of 8 spectra/second and
414 125ms/spectrum. Data-dependent MS/MS scans were triggered with an isolation window of 1.3 m/z at a rate of
415 3 spectra/second and 333 ms/spectrum.

416

417 **N-linked glycopeptide and site occupancy analysis – Data Analysis**

418 For intact glycopeptide analysis, the resulting RAW files were searched for glycopeptide spectral matches
419 (GPSMs) using the search engine pGlyco3(*74*). RAW files were searched against a FASTA database containing
420 Ag85B and common contaminants and a custom glycan database derived from released glycan compositions
421 identified in glycomics experiments (see above). Search was conducted with a precursor tolerance of 10 ppm,
422 fragment tolerance of 20 ppm, and glycopeptide FDR set to 0.01. Resulting glycopeptide spectral matches
423 (GPSMs) were manually inspected to ensure accuracy. GPSMs were processed by the ppmFixer script(*75*) to
424 adjust for common errors due to monoisotopic peak misidentification, and site-specific glycan compositions were
425 relatively quantified using spectral counts(*76*).

426 Agilent .d file folders were analyzed for site occupancy analysis using MassHunter Qualitative Analysis
427 (Agilent). Extracted ion chromatograms for predicted doubly and triply charged tryptic and α-lytic peptides were

428 generated, and peak areas were obtained. The unmodified peptide denotes the non-glycosylated fraction, the
429 +204 modification denotes high-mannose and hybrid structures, and the +3 modification denotes complex
430 structures. Peak areas for all observable predicted peptides and charge states were summed to obtain the final
431 site-specific glyoprofile. Figures were generated using GraphPad Prism 9 software.
432

433 **3D Structural Modeling and Molecular Dynamics Simulations of Nonglycosylated and Glycosylated**
434 **Ag85B**

435 *Mtb Ag85B protein structure and glycoform generation* – An experimentally determined X-ray diffraction
436 structure of *Mtb* Ag85B (RCSB PDB: 1F0N) was obtained through the RCSB Protein Data Bank (*38, 77*). Glycans
437 were selected to install glycosylated Ag85B N-glycan sequons based on glycomics-informed glycoproteomics
438 analysis of site-specific abundant glycoforms. The glycoprotein builder available at GLYCAM-Web
439 (ww.glycam.org) was employed together with an in-house method that adjusts asparagine side chain torsion
440 angles and glycosidic linkages to relieve atomic overlaps with the core protein backbone, as described previously
441 (*35, 36*).

442 *Energy minimization and Molecular Dynamics (MD) Simulations* – Each protein structure was placed in
443 a periodic box of OPCBOX water molecules with a 1.4 Å buffer between the solute and solvent box. Additionally,
444 66 sodium and 60 chlorine ions were placed randomly within the box to mimic physiological salt concentrations
445 (~150 mM NaCl). Energy minimization for each production was completed following a previously established
446 ten-step protocol involving 4000 total steps of minimization and 40,000 steps of MD, which equate to a series of
447 energy minimizations and relaxation steps that allow the system to relax gradually under constant pressure of 1
448 atm and constant temperature of 300 K(*78*). All MD simulations were performed in the NVT canonical ensemble
449 with CUDA-enabled implementation of PMEMD as included in the Amber24 software suite(*79-82*). The ff19SB
450 force field and GLYCAM06j force field were used for the protein and carbohydrate moieties, respectively. The
451 Bussi thermostat was employed to maintain a constant temperature of 300 K. Periodic boundary conditions
452 under constant volume, no pressure regulation, and a 10 Å cutoff for non-bonded interactions were employed
453 throughout the simulations. The SHAKE algorithm was used to constrain bonds with hydrogen atoms, and the
454 integration time step was set to 2 fs. Each system was then equilibrated for 10 ns before initiating 3 independent
455 1 μs production MD simulations with random starting velocities for a total time of 3 μs per system.

456 *Surface Area Accessibility Analysis* – To determine solvent accessible surface areas, 3D structures of
457 each frame (every 1 ns) of each production run were later analyzed for their ability to interact with a 1.4 Å
458 spherical probe using the CPPTRAJ surf command(*82*). To determine the accessible surface area to a larger
459 probe on a residue level, ten hierarchical ensemble 3D structures were generated using all three productions of
460 each system, and then probed for their ability to interact with a 10 Å spherical probe using NACCESS(*40, 41,*
461 *82*). Data are displayed as violin plots, and statistical significance was calculated using either Welch's t-test or
462 Mann-Whitney U test, depending on distribution normality, using Python. All scripts are available on request.
463 Figures were generated using GraphPad Prism 9 software.
464
465 **Competitive Enzyme-Linked Immunosorbent Assay (ELISA)**.
466 To determine the needed concentration of the monoclonal antibody against the Ag85 Complex, a direct
467 ELISA was first conducted. Plates were coated with 2.0 µg/mL of *Mtb* Ag85B and incubated overnight at 4ºC.
468 Plates were blocked with 40 µl/well of 3% bovine serum albumin (BSA) in phosphate-buffered saline (PBS) for
469 2 hours on an orbital shaker at 100 rpm at room temperature. 20 µl/well of primary antibody (anti-*Mtb* Ag85
470 Complex IgG1, Clone P4B3G2, BEI Resources) was added using a 2-fold serial dilution and incubated for 2
471 hours on an orbital shaker at 100 rpm at room temperature. 20 µl/well of 1:2000 diluted anti-mouse IgG-Alkaline
472 phosphatase (AP) conjugated secondary antibody (Southern BioTech Cat. #1030-04) was added and incubated
473 for 1 hour on an orbital shaker at 100 rpm at room temperature. All serum dilutions were run in technical
474 replicates.
475 After determining the serum concentration needed, the competition ELISA was conducted. Plates were
476 coated with 2.0 µg/mL of *Mtb* Ag85B and incubated overnight at 4ºC. Plates were blocked with 40 µl/well of 3%
477 bovine serum albumin (BSA) in phosphate-buffered saline (PBS) for 2 hours on an orbital shaker at 100 rpm at
478 room temperature. While blocking was taking place, on a separate plate, competitors (*Mtb* Ag85B, BL21 Ag85B,
479 293-F Ag85B, 293-F dAg85B) with four different concentrations of 100 µg/mL, 50 µg/mL, 10 µg/mL, and 1.0
480 µg/mL were incubated with the determined 1:25600 antibody dilution for 2 hours on an orbital shaker at 100 rpm
481 at room temperature. Afterwards, 20 µl/well of the antibody-competitor solution was added to the blocked plate
482 in technical triplicate and incubated for 2 hours on an orbital shaker at 100 rpm at room temperature. Next, 20
483 µl/well of 1:2000 diluted anti-mouse IgG-AP conjugated secondary antibody (Southern BioTech Cat. #1030-04)

484 was added and incubated for 1 hour on an orbital shaker at 100 rpm at room temperature. Plates were imaged
485 using 20 µl/well of 2 mg/mL of AP substrate (Sigma Cat. #S0942-200TAB) diluted in 1M Tris Base, 0.3 mM $MgCl_2$
486 buffer, and OD readings were determined using the BioTek Synergy Neo with the Gen5 Microplate Reader and
487 Image Software. All samples were diluted in 1% BSA in PBS-T. All plates were washed with 1% BSA in PBS-T
488 between each incubation step. Figures and statistical analysis were completed using GraphPad Prism 9
489 software.
490

491 **Biolayer Interferometry (BLI) for Antibody-Binding Kinetic Assay.**

492 BLI assays were performed on a Sartorius Octet R8 instrument at 30ºC with shaking at 1000 rpm. All
493 measurements were adjusted by subtracting the background signal from blank loads. Anti-*Mtb* Ag85 Complex
494 IgG1 (Clone P4B3G2) was immobilized on anti-Mouse Fc Capture (AMC) biosensors (Sartorius, Cat. # 18-5088)
495 at 10 µg/ml in kinetic buffer (PBS, 0.02% Tween-20, 0.1% BSA, 0.05% sodium azide). After loading, biosensors
496 were incubated in kinetic buffer for 10 minutes to establish a baseline signal. Biosensors were then incubated in
497 various concentrations (500, 200, 100, 50, 25, and 10 nM) of the analytes (BL21 Ag85B, 293-F Ag85B, and 293-
498 F dAg85B) for kinetic analysis. Curve fitting to determine KD, k_{on}, and k_{off} for each analyte was performed using
499 a global 1:1 binding model with the Octet Analysis Software, with R^2 values greater than a 95% confidence level.
500 Biosensors were conditioned before loading the antibody and regenerated after the kinetic assay according to
501 the manufacturer's recommendations. Figures were generated using GraphPad Prism 9 software.
502

503 **Acknowledgments**

504 National Institutes of Health grants R01AI123383, R01AI152766, and R41AI157287 supported this work.

References:

1. W. R. Ellis, R. Rappuoli, S. Ahmed, in *Vaccines,* S. A. Plotkin, W. A. Orenstein, P. A. Offit, Eds. (Saunders, 2013), chap. 60, pp. 1182-1199.
2. I. Delany, R. Rappuoli, E. De Gregorio, Vaccines for the 21st century. *EMBO Mol Med* **6**, 708-720 (2014).
3. A. V. Paschall, F. Y. Avci, in *Vaccine Development: From Concept to Clinic,* A. K. Prasad, Ed. (The Royal Society of Chemistry, 2022), pp. 0.
4. D. S. Rajao, D. R. Perez, Universal Vaccines and Vaccine Platforms to Protect against Influenza Viruses in Humans and Agriculture. *Front Microbiol* **9**, 123 (2018).
5. G. Voss *et al.*, Progress and challenges in TB vaccine development. *F1000Res* **7**, 199 (2018).
6. S. J. Draper, J. L. Heeney, Viruses as vaccine vectors for infectious diseases and cancer. *Nature Reviews Microbiology* **8**, 62 (2009).
7. A. Ozdilek, F. Y. Avci, Glycosylation as a key parameter in the design of nucleic acid vaccines. *Curr Opin Struct Biol* **73**, 102348 (2022).
8. J. Y. Zhou, D. M. Oswald, K. D. Oliva, L. S. C. Kreisman, B. A. Cobb, The Glycoscience of Immunity. *Trends in Immunology* **39**, 523-535 (2018).
9. WHO, "Global tuberculosis report 2024," (World Health Organization, 2023).
10. T. Paulson, Epidemiology: A mortal foe. *Nature* **502**, S2-S3 (2013).
11. K. Dheda *et al.*, Multidrug-resistant tuberculosis. *Nat. Rev. Dis. Primers* **10**, 27 (2024).
12. K. J. Seung, S. Keshavjee, M. L. Rich, Multidrug-Resistant Tuberculosis and Extensively Drug-Resistant Tuberculosis. *Cold Spring Harb Perspect Med* **5**, a017863 (2015).
13. J. Manjelievskaia, D. Erck, S. Piracha, L. Schrager, Drug-resistant TB: deadly, costly and in need of a vaccine. *Transactions of The Royal Society of Tropical Medicine and Hygiene* **110**, 186-191 (2016).
14. S. Luca, T. Mihaescu, History of BCG Vaccine. *Maedica (Bucur)* **8**, 53-58 (2013).
15. G. A. Colditz *et al.*, Efficacy of BCG Vaccine in the Prevention of Tuberculosis: Meta-analysis of the Published Literature. *JAMA* **271**, 698-702 (1994).
16. B. B. Trunz, P. E. M. Fine, C. Dye, Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *The Lancet* **367**, 1173-1180 (2006).
17. C. L. Karp, C. B. Wilson, L. M. Stuart, Tuberculosis vaccines: barriers and prospects on the quest for a transformative tool. *Immunol Rev* **264**, 363-381 (2015).
18. R. Lai, A. F. Ogunsola, T. Rakib, S. M. Behar, Key advances in vaccine development for tuberculosis-success and challenges. *NPJ Vaccines* **8**, 158 (2023).
19. J. T. Belisle *et al.*, Role of the major antigen of Mycobacterium tuberculosis in cell wall biogenesis. *Science* **276**, 1420-1422 (1997).
20. M. K. Z. Babaki, S. Soleimanpour, S. A. Rezaee, Antigen 85 complex as a powerful Mycobacterium tuberculosis immunogene: Biology, immune-pathogenicity, applications in diagnosis, and vaccine design. *Microbial Pathogenesis* **112**, 20-29 (2017).
21. E. Lozes *et al.*, Immunogenicity and efficacy of a tuberculosis DNA vaccine encoding the components of the secreted antigen 85 complex. *Vaccine* **15**, 830-833 (1997).
22. K. Huygen *et al.*, Immunogenicity and protective efficacy of a tuberculosis DNA vaccine. *Nat Med* **2**, 893-898 (1996).
23. S. H. Kaufmann, Tuberculosis vaccines: time to think about the next generation. *Semin Immunol* **25**, 172-181 (2013).
24. M. D. Tameris *et al.*, Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet* **381**, 1021-1028 (2013).
25. B. P. Ndiaye *et al.*, Safety, immunogenicity, and efficacy of the candidate tuberculosis vaccine MVA85A in healthy adults infected with HIV-1: a randomised, placebo-controlled, phase 2 trial. *Lancet Respir Med* **3**, 190-200 (2015).
26. M. Tameris *et al.*, A double-blind, randomised, placebo-controlled, dose-finding trial of the novel tuberculosis vaccine AERAS-402, an adenovirus-vectored fusion protein, in healthy, BCG-vaccinated infants. *Vaccine* **33**, 2944-2954 (2015).
27. A. Ozdilek, A. V. Paschall, M. Dookwah, M. Tiemeyer, F. Y. Avci, Host protein glycosylation in nucleic acid vaccines as a potential hurdle in vaccine design for nonviral pathogens. *Proceedings of the National Academy of Sciences* **117**, 1280-1282 (2020).

28. P. M. Angel, J.-M. Lim, L. Wells, C. Bergmann, R. Orlando, A potential pitfall in18O-based N-linked glycosylation site mapping. *Rapid Communications in Mass Spectrometry* **21**, 674-682 (2007).
29. W. Breuer, R. A. Klein, B. Hardt, A. Bartoschek, E. Bause, Oligosaccharyltransferase is highly specific for the hydroxy amino acid in Asn-Xaa-Thr/Ser. *FEBS Lett* **501**, 106-110 (2001).
30. J. Li, H. C. Hsu, J. D. Mountz, J. G. Allen, Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. *Cell Chem Biol* **25**, 499-512 (2018).
31. S. S. Pinho, I. Alves, J. Gaifem, G. A. Rabinovich, Immune regulatory networks coordinated by glycans and glycan-binding proteins in autoimmunity and infection. *Cell Mol Immunol* **20**, 1101-1113 (2023).
32. M. Demirturk, M. S. Cinar, F. Y. Avci, The immune interactions of gut glycans and microbiota in health and disease. *Mol Microbiol* **122**, 313-330 (2024).
33. M. R. Kudelka, S. R. Stowell, R. D. Cummings, A. S. Neish, Intestinal epithelial glycosylation in homeostasis and gut microbiota interactions in IBD. *Nat Rev Gastroenterol Hepatol* **17**, 597-617 (2020).
34. M. P. Jennings, C. J. Day, J. M. Atack, How bacteria utilize sialic acid during interactions with the host: snip, snatch, dispatch, match and attach. *Microbiology (Reading)* **168**, (2022).
35. P. Zhao *et al.*, Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host Microbe* **28**, 586-601 e586 (2020).
36. R. Woods. (2005).
37. O. C. Grant *et al.*, Generating 3D Models of Carbohydrates with GLYCAM-Web. *bioRxiv*, 2025.2005.2008.652828 (2025).
38. D. H. Anderson, G. Harth, M. A. Horwitz, D. Eisenberg, An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the Mycobacterium tuberculosis 30 kda major secretory protein (antigen 85B), a mycolyl transferase11Edited by I. A. Wilson. *Journal of Molecular Biology* **307**, 671-681 (2001).
39. D. R. Roe, T. E. Cheatham, 3rd, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **9**, 3084-3095 (2013).
40. S. J. a. T. Hubbard, J.M. . (Computer Program, Department of Biochemistry and Molecular Biology, University College London, 1993).
41. T. Zhou *et al.*, Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation. *Cell Reports* **19**, 719-732 (2017).
42. A. J. Pollard, E. M. Bijker, A guide to vaccinology: from basic principles to new developments. *Nat Rev Immunol* **21**, 83-100 (2021).
43. M. Iraqi *et al.*, N-Glycans Mediate the Ebola Virus-GP1 Shielding of Ligands to Immune Receptors and Immune Evasion. *Front Cell Infect Microbiol* **10**, 48 (2020).
44. L. G. Baum, B. A. Cobb, The direct and indirect effects of glycans on immune function. *Glycobiology* **27**, 619-624 (2017).
45. F. Rinaldi *et al.*, Epitope and affinity determination of recombinant Mycobacterium tuberculosis Ag85B antigen towards anti-Ag85 antibodies using proteolytic affinity-mass spectrometry and biosensor analysis. *Analytical and Bioanalytical Chemistry* **411**, 439-448 (2019).
46. M. A. Wolfert, G.-J. Boons, Adaptive immune activation: glycosylation does matter. *Nature Chemical Biology* **9**, 776-784 (2013).
47. A. S. de Wit, F. Bianchi, G. van den Bogaart, Antigen presentation of post-translationally modified peptides in major histocompatibility complexes. *Immunology & Cell Biology* **103**, 161-177 (2025).
48. K. Huygen *et al.*, Immunogenicity and protective efficacy of a tuberculosis DNA vaccine. *Nature Medicine* **2**, 893-898 (1996).
49. C. S. Lindestam Arlehamn *et al.*, Memory T Cells in Latent Mycobacterium tuberculosis Infection Are Directed against Three Antigenic Islands and Largely Contained in a CXCR3+CCR6+ Th1 Subset. *PLOS Pathogens* **9**, e1003130 (2013).
50. A. S. Mustafa *et al.*, Identification and HLA Restriction of Naturally Derived Th1-Cell Epitopes from the Secreted Mycobacterium tuberculosisAntigen 85B Recognized by Antigen-Specific Human CD4⁺T-Cell Lines. *Infection and Immunity* **68**, 3933-3940 (2000).
51. M. T. Valle *et al.*, Epitope focus, clonal composition and Th1 phenotype of the human CD4 response to the secretory mycobacterial antigen Ag85. *Clinical and Experimental Immunology* **123**, 226-232 (2001).
52. A. Geluk *et al.*, Identification of HLA class II-restricted determinants of Mycobacterium tuberculosis-derived proteins by using HLA-transgenic, class II-deficient mice. *Proceedings of the National Academy of Sciences* **95**, 10797-10802 (1998).

53. E. Arrieta-Bolaños, D. I. Hernández-Zaragoza, R. Barquera, An HLA map of the world: A comparison of HLA frequencies in 200 worldwide populations reveals diverse patterns for class I and class II. *Frontiers in Genetics* **Volume 14 - 2023**, (2023).
54. D. S. Romaniuk, A. M. Postovskaya, A. A. Khmelevskaya, D. B. Malko, G. A. Efimov, Rapid Multiplex Genotyping of 20 HLA-A*02:01 Restricted Minor Histocompatibility Antigens. *Frontiers in Immunology* **Volume 10 - 2019**, (2019).
55. M. R. Klein *et al.*, HLA-B*35–Restricted CD8 T Cell Epitopes in the Antigen 85 Complex of Mycobacterium tuberculosis. *The Journal of Infectious Diseases* **183**, 928-934 (2001).
56. L. Sun *et al.*, Glycopeptide epitope facilitates HIV-1 envelope specific humoral immune responses by eliciting T cell help. *Nat Commun* **11**, 2550 (2020).
57. L. Cao *et al.*, Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. *Nature Communications* **8**, 14954 (2017).
58. L. P. Deimel, X. Xue, Q. J. Sattentau, Glycans in HIV-1 vaccine design – engaging the shield. *Trends in Microbiology* **30**, 866-881 (2022).
59. C. Upadhyay *et al.*, Signal peptide of HIV-1 envelope modulates glycosylation impacting exposure of V1V2 and other epitopes. *PLOS Pathogens* **16**, e1009185 (2021).
60. V. Dubrovskaya *et al.*, Targeted N-glycan deletion at the receptor-binding site retains HIV Env NFL trimer integrity and accelerates the elicited antibody response. *PLOS Pathogens* **13**, e1006614 (2017).
61. Q. Wei *et al.*, Glycan Positioning Impacts HIV-1 Env Glycan-Shield Density, Function, and Recognition by Antibodies. *iScience* **23**, (2020).
62. O. C. Grant, D. Montgomery, K. Ito, R. J. Woods, Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Scientific Reports* **10**, 14991 (2020).
63. L. Casalino *et al.*, Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Central Science* **6**, 1722-1734 (2020).
64. F. Zhang *et al.*, SARS-CoV-2 spike glycosylation affects function and neutralization sensitivity. *mBio* **15**, e01672-01623 (2024).
65. H.-Y. Huang *et al.*, Vaccination with SARS-CoV-2 spike protein lacking glycan shields elicits enhanced protective responses in animal models. *Science Translational Medicine* **14**, eabm0899 (2022).
66. C.-Y. Wu *et al.*, Glycosite-deleted mRNA of SARS-CoV-2 spike protein as a broad-spectrum vaccine. *Proceedings of the National Academy of Sciences* **119**, e2119995119 (2022).
67. H. Li *et al.*, Proximal Glycans Outside of the Epitopes Regulate the Presentation of HIV-1 Envelope gp120 Helper Epitopes1. *The Journal of Immunology* **182**, 6369-6378 (2009).
68. B. S. Bochner, N. Zimmermann, Role of siglecs and related glycan-binding proteins in immune responses and immunoregulation. *J Allergy Clin Immunol* **135**, 598-608 (2015).
69. P. Saini, O. S. Adeniji, M. Abdel-Mohsen, Inhibitory Siglec-sialic acid interactions in balancing immunological activation and tolerance during viral infections. *eBioMedicine* **86**, (2022).
70. F. Rinaldi *et al.*, Rational design, preparation and characterization of recombinant Ag85B variants and their glycoconjugates with T-cell antigenic activity against Mycobacterium tuberculosis. *RSC Advances* **8**, 23171-23180 (2018).
71. I. Brockhausen, H. H. Wandall, K. G. T. Hagen, P. Stanley, in *Essentials of Glycobiology,* A. Varki *et al.*, Eds. (Cold Spring Harbor Laboratory Press

Copyright © 2022 The Consortium of Glycobiology Editors, La Jolla, California; published by Cold Spring Harbor Laboratory Press; doi:10.1101/glycobiology.4e.10. All rights reserved., Cold Spring Harbor (NY), 2022), pp. 117-128.

72. W. S. York *et al.*, GlyGen: Computational and Informatics Resources for Glycoscience. *Glycobiology* **30**, 72-73 (2020).
73. L. K. Pino *et al.*, The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrom Rev* **39**, 229-244 (2020).
74. W.-F. Zeng, W.-Q. Cao, M.-Q. Liu, S.-M. He, P.-Y. Yang, Precise, fast and comprehensive analysis of intact glycopeptides and modified glycans with pGlyco3. *Nature Methods* **18**, 1515-1523 (2021).
75. T. M. Adams, P. Zhao, R. Kong, L. Wells, ppmFixer: a mass error adjustment for pGlyco3.0 to correct near-isobaric mismatches. *Glycobiology*, cwae006 (2024).
76. A. A. Dowle, J. Wilson, J. R. Thomas, Comparing the Diagnostic Classification Accuracy of iTRAQ, Peak-Area, Spectral-Counting, and emPAI Methods for Relative Quantification in Expression Proteomics. *J Proteome Res* **15**, 3550-3562 (2016).

77. H. Berman, K. Henrick, H. Nakamura, Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology* **10**, 980-980 (2003).
78. D. R. Roe, B. R. Brooks, A protocol for preparing explicitly solvated systems for stable molecular dynamics simulations. *J Chem Phys* **153**, 054123 (2020).
79. A. W. Götz *et al.*, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *Journal of Chemical Theory and Computation* **8**, 1542-1555 (2012).
80. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **9**, 3878-3888 (2013).
81. D. A. Case, Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Berryman, J. T., Brozell, S. R., Cerutti, D. S., Cheatham, III, T. E., Cisneros, G. A., Cruzeiro, V. W. D., Darden, T. A., Forouzesh, N., Ghazimirsaeed, M., Giambaşu, G., Giese, T., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, J., Huang, Z., Izadi, S., Izmailov, S. A., Kasavajhala, K., Kaymak, M. C., Kovalenko, A., Kurtzman, T., Lee, T. S., Li, P., Li, Z., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Manathunga, M., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O'Hearn, K. A., Onufriev, A., Pan, F., Pantano, S., Rahnamoun, A., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shajan, A., Shen, J., Simmerling, C. L., Skrynnikov, N. R., Smith, J., Swails, J., Walker, R. C., Wang, J., Wang, J., Wu, X., Wu, Y., Xiong, Y., Xue, Y., York, D. M., Zhao, C., Zhu, Q., and Kollman, P. A. . (2024).
82. D. A. Case *et al.*, AmberTools. *Journal of Chemical Information and Modeling* **63**, 6183-6191 (2023).

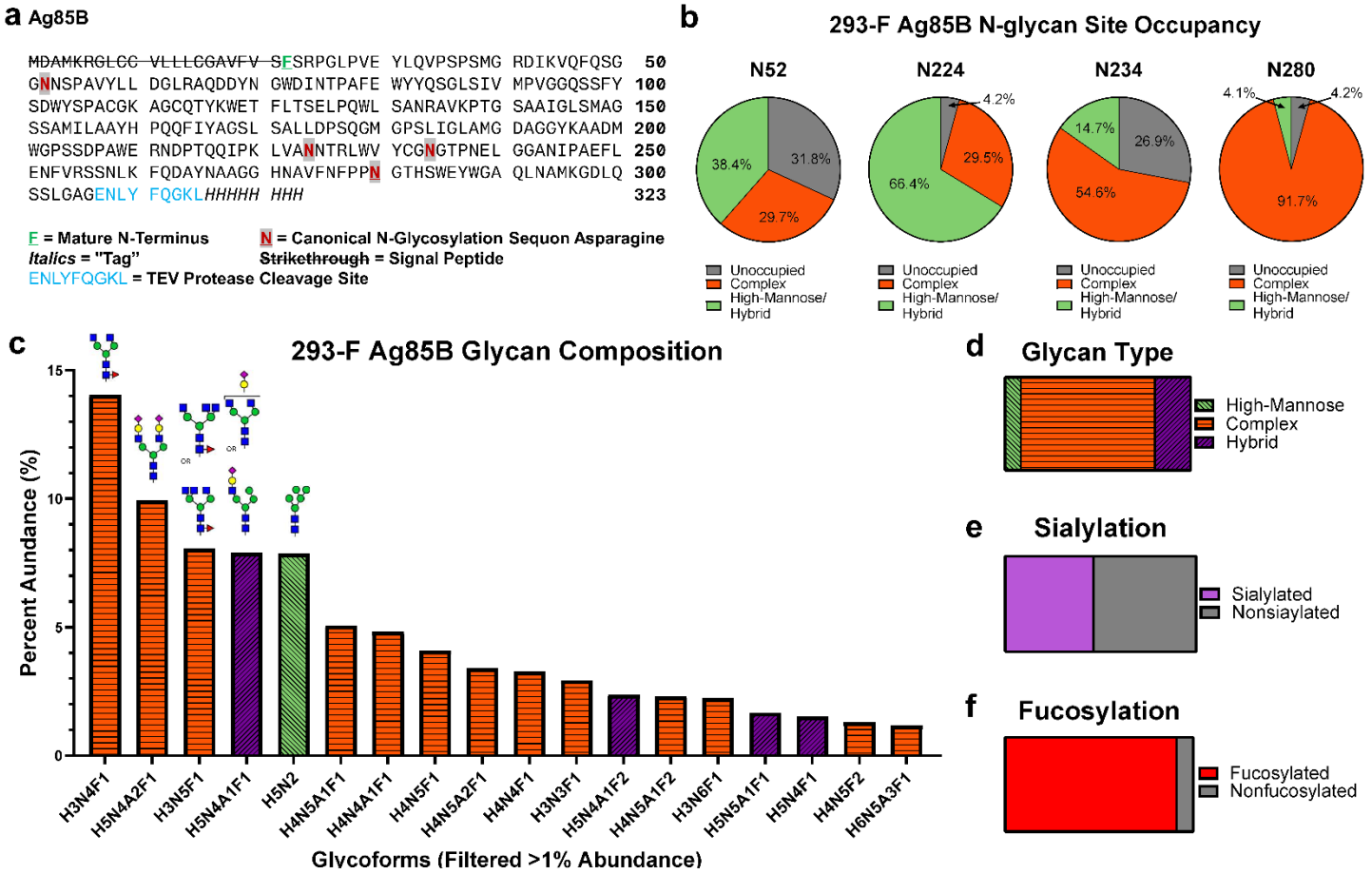
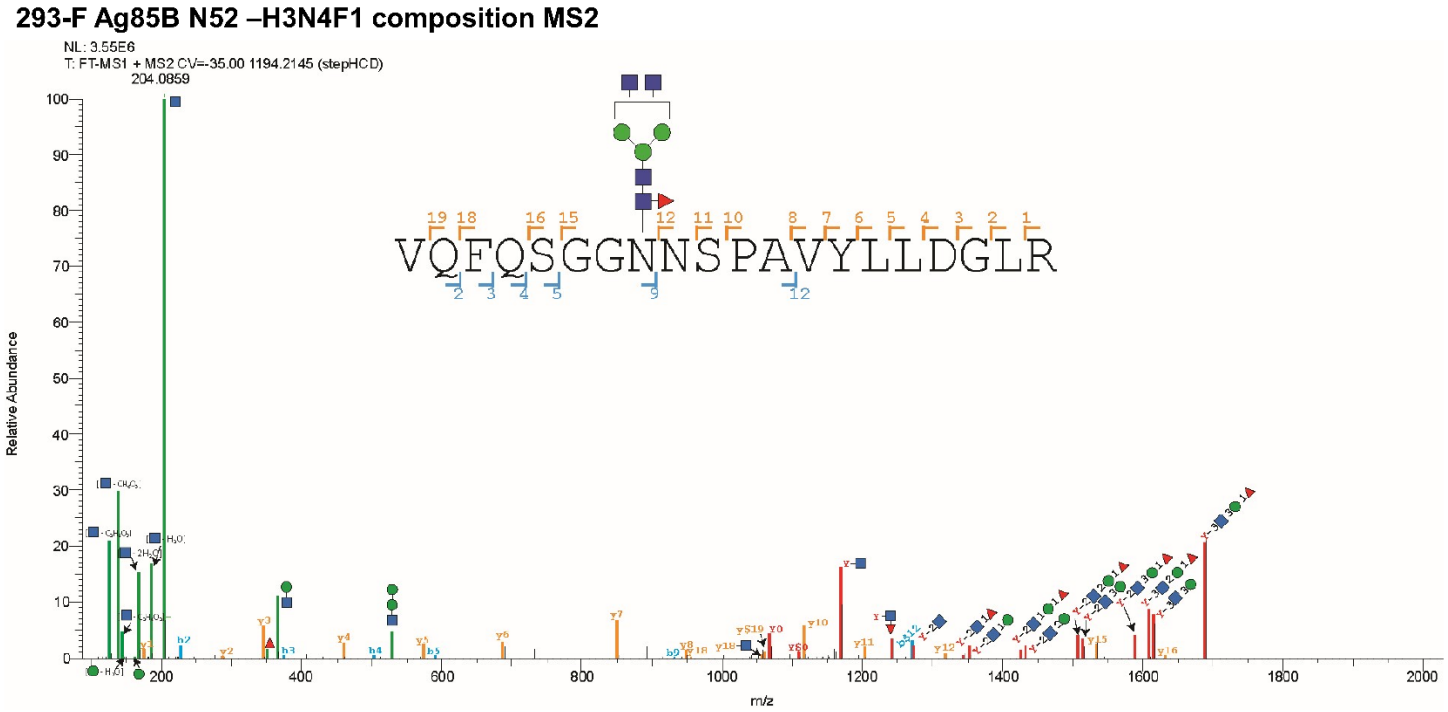
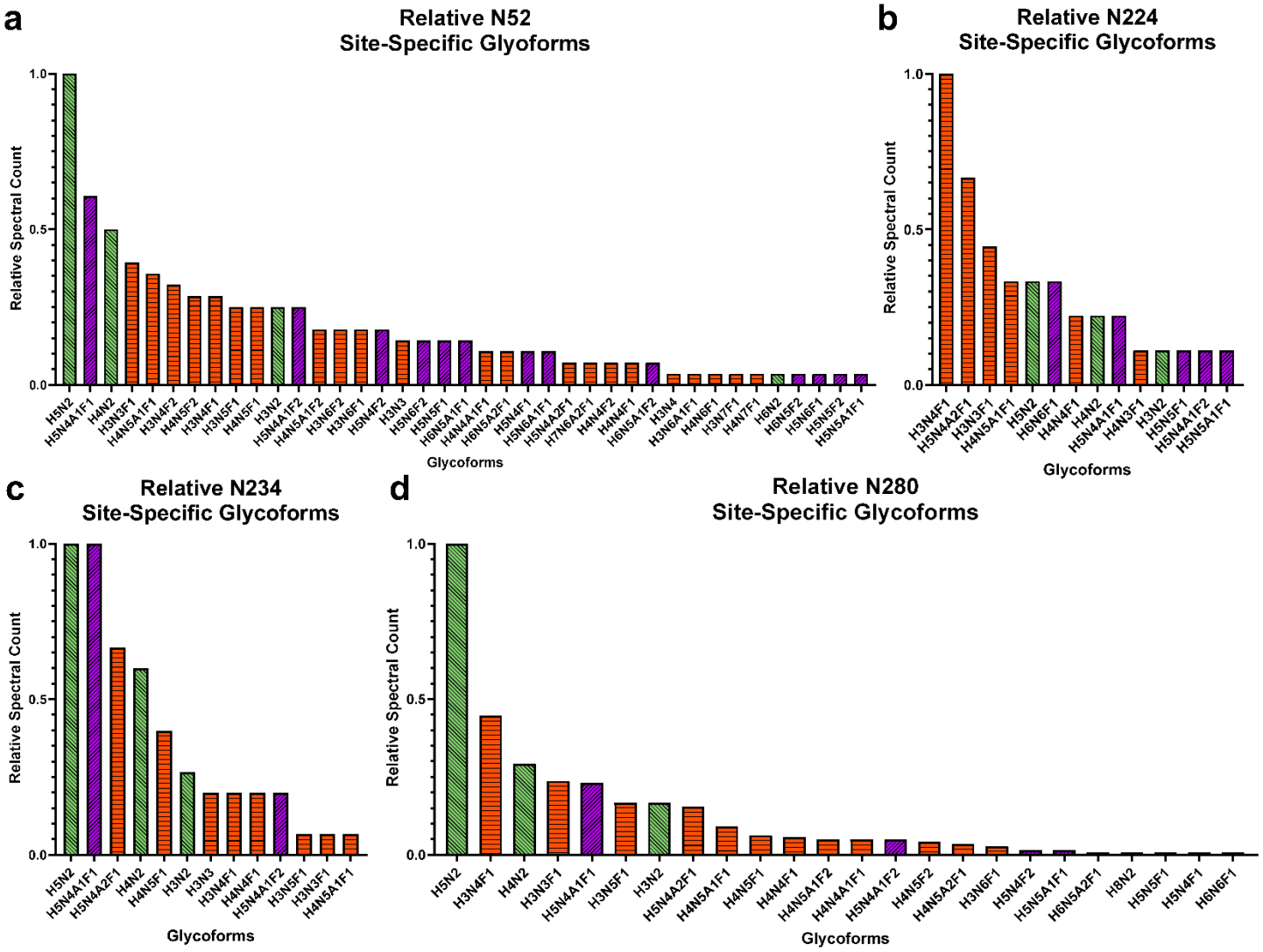


Figure 1: Ag85B is glycosylated at four canonical N-linked glycosylation sequons when expressed in mammalian Expi293 cells. (a) Protein sequence of Ag85B. The mature N-terminal is bolded, underlined, and shown in green. The canonical N-linked glycosylation sequon asparagine is bolded in red text, highlighted in gray, and underlined. The TEV protease cleavage site is displayed in blue text, and the 8xHIS tag is italicized. **(b)** Site occupancy of canonical N-linked glycosylation sites (N52, N224, N234, and N280) using EndoH and PNGase F digestion with heavy water, D_2O. **(c)** Glycomics data showing the percent abundance of released glycans from 293-F Ag85B using PNGase F digestion filtered to show greater than 1% abundance in the total identified glycoforms. Complex structures are represented in dotted orange, hybrid structures are represented in checkered purple, and high-mannose structures are represented in triangle light green. Putative structures shown for the most abundant glycan compositions; note that these compositions are a mixture of isomers. **(d)** Distribution of high-mannose (8.7%) and (72.0%) complex, and (19.3%) hybrid glycoform structures from the total identified glycoforms. **(e)** Distribution of sialylated (46.3%) glycoforms from the total identified glycoforms. **(f)** Distribution of fucosylated (91.1%) glycoforms from the total identified glycoforms.



Supplementary Figure 1: A representative spectrum of N52 glycan site demonstrates the rich fragmentation generated from the stepped high-energy collisional dissociation (sHCD), which provides functional fragments both of the peptide backbone and of the glycan itself in a single MS2, which are used by the search engine pGlyco3 to identify glycopeptides.



Supplemental Figure 2: Ag85B is microheterogeneously glycosylated when expressed in human Expi293 cells. Site-specific microheterogeneity of **(a)** N52, **(b)** N224, **(c)** N234, and **(d)** N280 using glycomics-informed glycopeptidomics. Relative spectral counts were normalized to the most abundant glycoforms present at each site. Complex structures are represented in dotted orange, hybrid structures are represented in checkered purple, and high-mannose structures are represented in triangle light green.

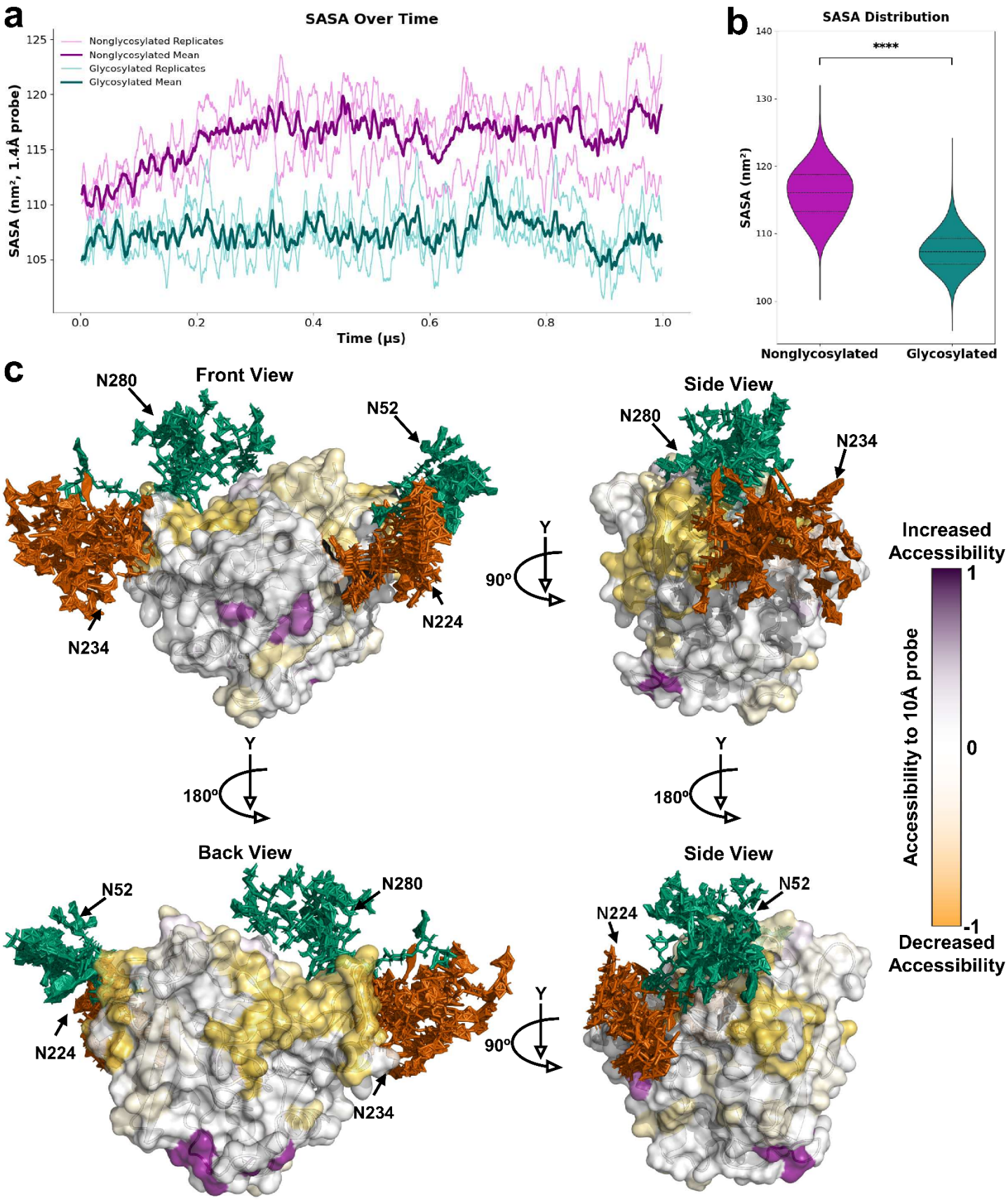
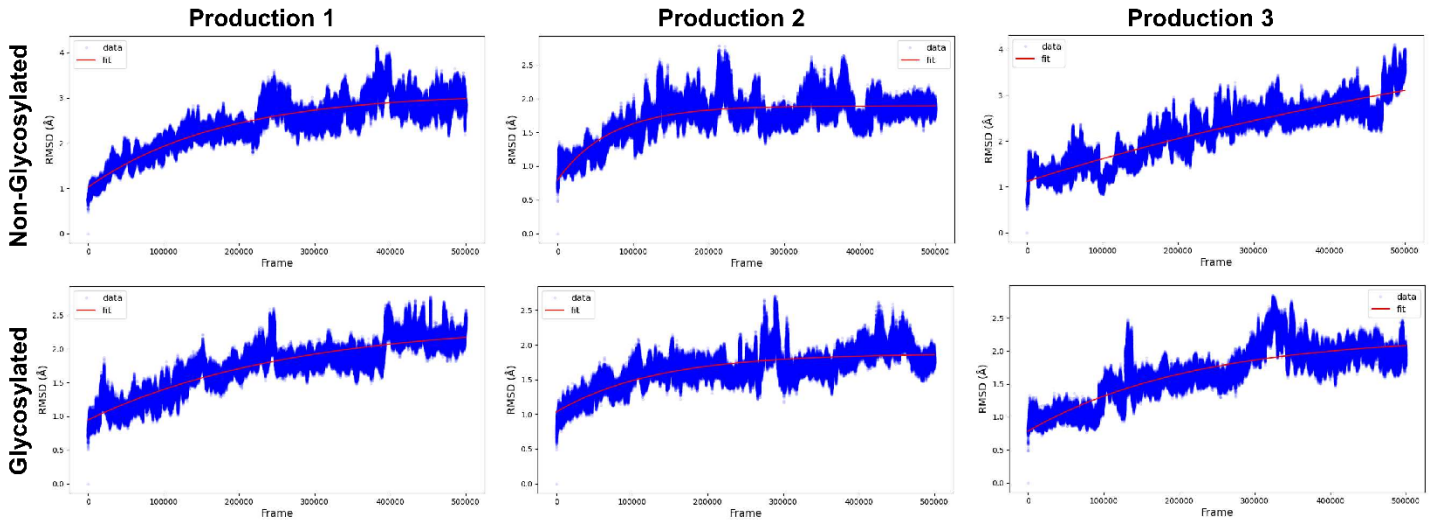
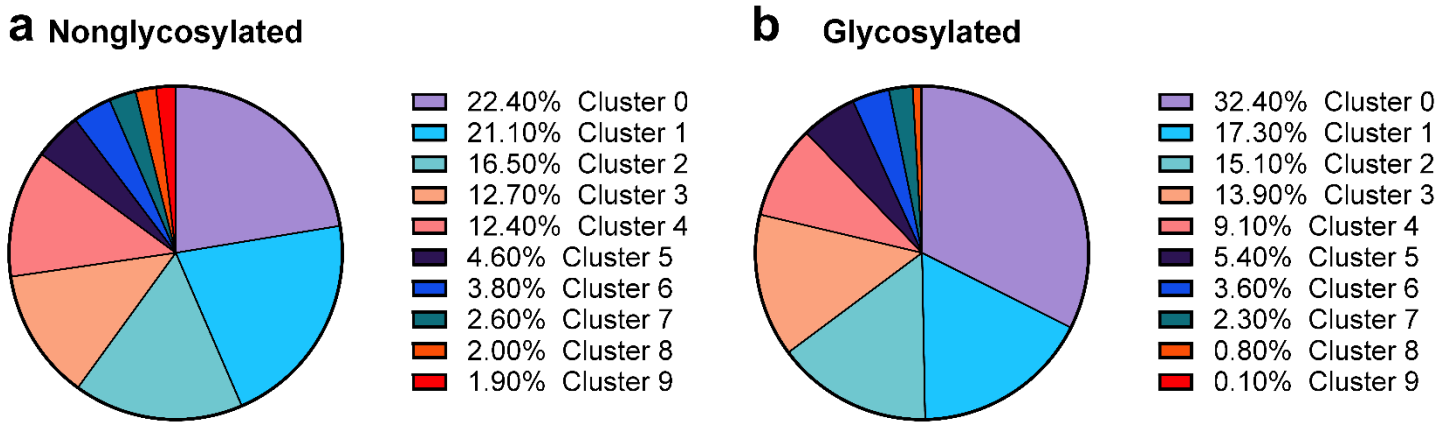

a
SASA Over Time
Nonglycosylated Replicates
Nonglycosylated Mean
Glycosylated Replicates
Glycosylated Mean
SASA (nm², 1.4Å probe)
Time (μs)
b
SASA Distribution
****
SASA (nm²)
Nonglycosylated
Glycosylated
c
Front View
N280
N52
N234
N224
Side View
N280
N234
90°
Y
180°
Back View
N52
N280
N224
N234
Side View
N52
N224
Increased Accessibility
Accessibility to 10Å probe
Decreased Accessibility

Figure 2: Molecular dynamic simulations indicate the impact of Ag85B N-glycosylation on the protein surface accessibility. (a) Water accessibility over three 1-microsecond-long molecular dynamic simulations of the nonglycosylated and glycosylated Ag85B with a probe size of 1.4 Angstroms. Replicates are displayed in a lighter shade of magenta and teal for the nonglycosylated and glycosylated MD runs, respectively. Mean of all three replicates displayed in a darker shade of magenta and teal for the nonglycosylated and glycosylated MD runs, respectively. **(b)** The water-accessible surface area from (a) is demonstrated as a distribution. $****p < 0.0001$ represents a significant difference between groups determined using the Mann-Whitney U test. **(c)** Hierarchal ensemble representation of the glycomics-informed glycopeptidomics site specific abundant glycan structures of Ag85B generated after three 1 microsecond long molecular dynamic simulations viewed from the front view (top left), back view (bottom left), side view when rotated 90 deg to the left in respect to the front view (top right), and side view when rotated 90 deg to the left in respect to the back view (bottom right). The gradient bar demonstrates the percent change in accessibility to a 10 Angstrom probe, representing the antibody-accessible surface area of the glycosylated Ag85B protein compared to the native, nonglycosylated protein. Residue-specific antibody-accessible surface areas were calculated using NACCESS. Positive/purple represents increased antibody accessibility, and negative/yellow represents decreased antibody accessibility. Complex/hybrid glycan structures are represented in orange. High-mannose structures are represented in green. The protein backbone is represented in white.



Supplemental Figure 3: Quality control of molecular dynamics simulations of the three replicate productions per system.



Supplemental Figure 4: Hierarchical clustering of molecular dynamics simulations for the **(a)** nonglycosylated and **(b)** glycosylated Ag85B in ten representative conformations.

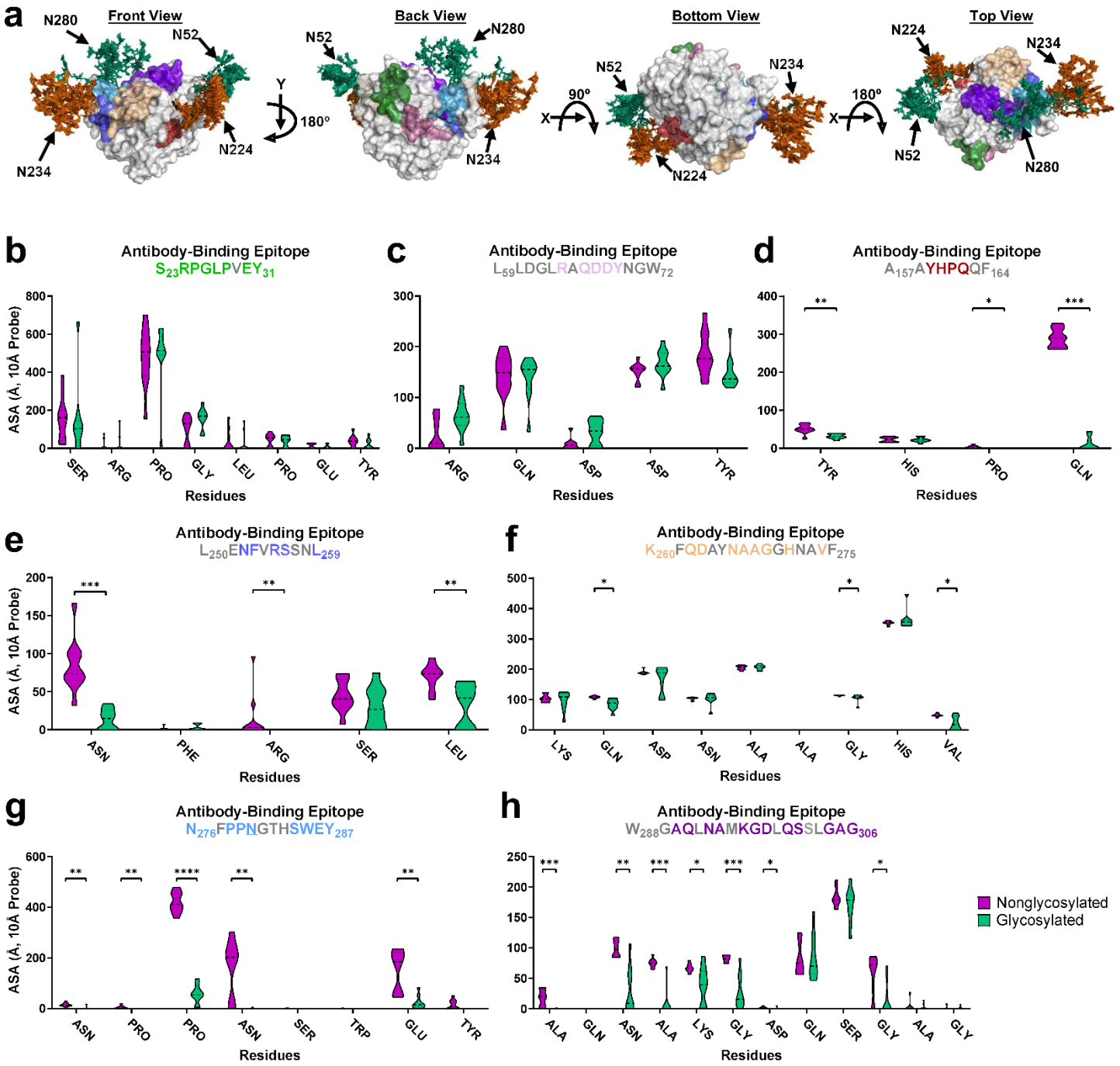


Figure 3: Molecular dynamic simulations illustrate the impact of Ag85B N-glycosylation on the accessibility of known antibody binding epitopes. (a) Hierarchical ensemble representation of the glycomics-informed glycopeptidomics site-specific abundant glycan structures of Ag85B generated after three 1-microsecond-long molecular dynamic simulations viewed from front, back, bottom, and top (from left to right), displaying specific antibody binding epitope locations distinguished through color coding. Antibody accessible surface areas of various antibody epitopes identified by Rinaldi et al. 2018 using proteolytic affinity-mass

753 spectrometry and biosensor analysis displayed through violin plots for **(b)** site S23-Y31 (green on panel a), **(c)**
754 site L59-W72 (pink on panel a), **(d)** site A157-F164 (red in panel a), **(e)** site L250-L259 (blue in panel a), **(f)** site
755 K260-F275 (tan in panel a), **(g)** site N276-Y287 (light-blue in panel a), and **(h)** site W288-G306 (purple in panel
756 a). Identified N-glycosylated asparagines are underlined in the graph title. For clarity, amino acids where there
757 was no accessibility before and after adding glycans at identified N-glycan sequons are colored in gray in the
758 graph titles and are removed from the shown data on the x-axis. $****p < 0.0001$, $***p < 0.001$, $**p < 0.01$, and
759 $*p < 0.05$ represent significant differences between groups determined using either Welch's t-test or the Mann-
760 Whitney U test, depending on distribution normality.

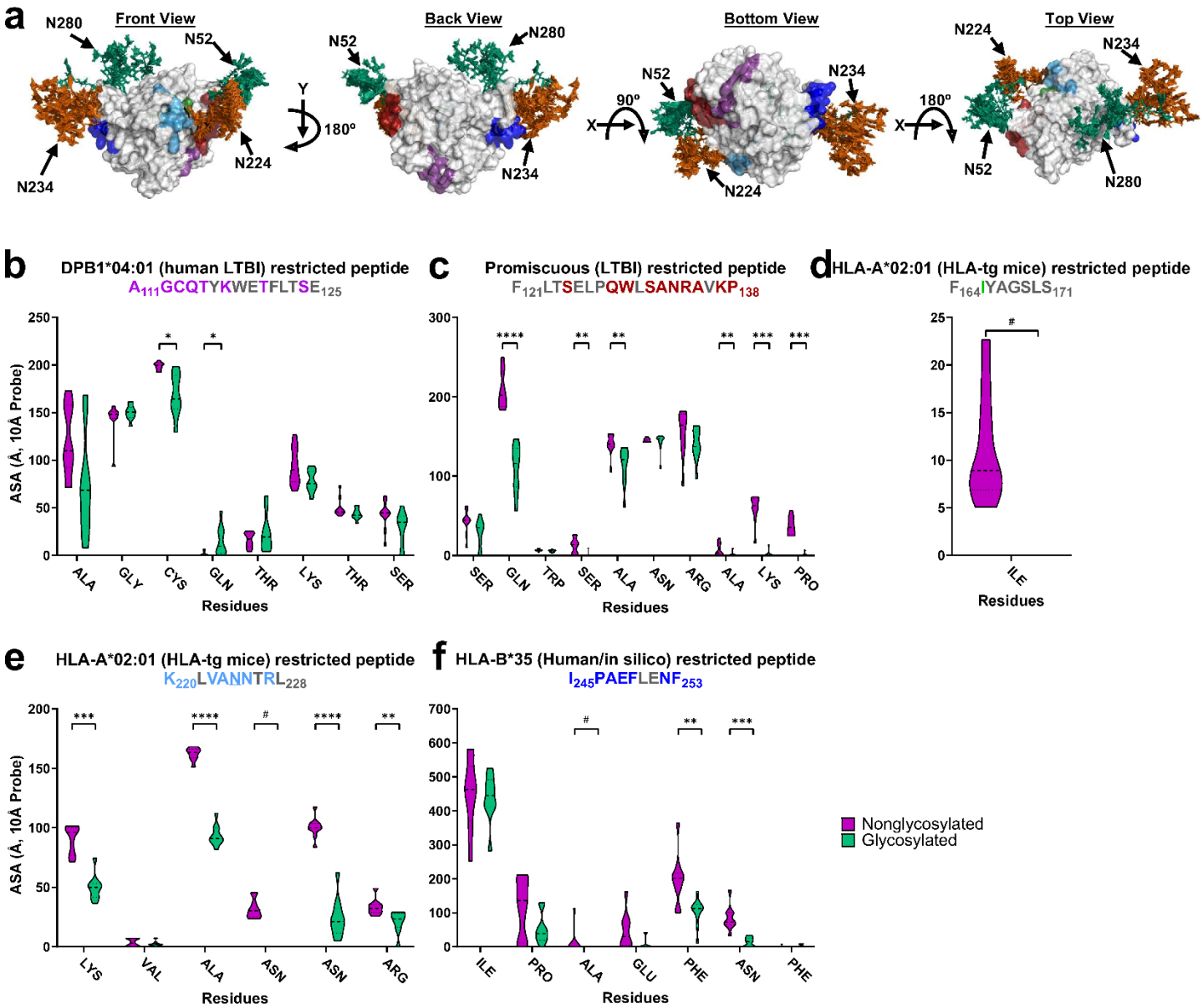
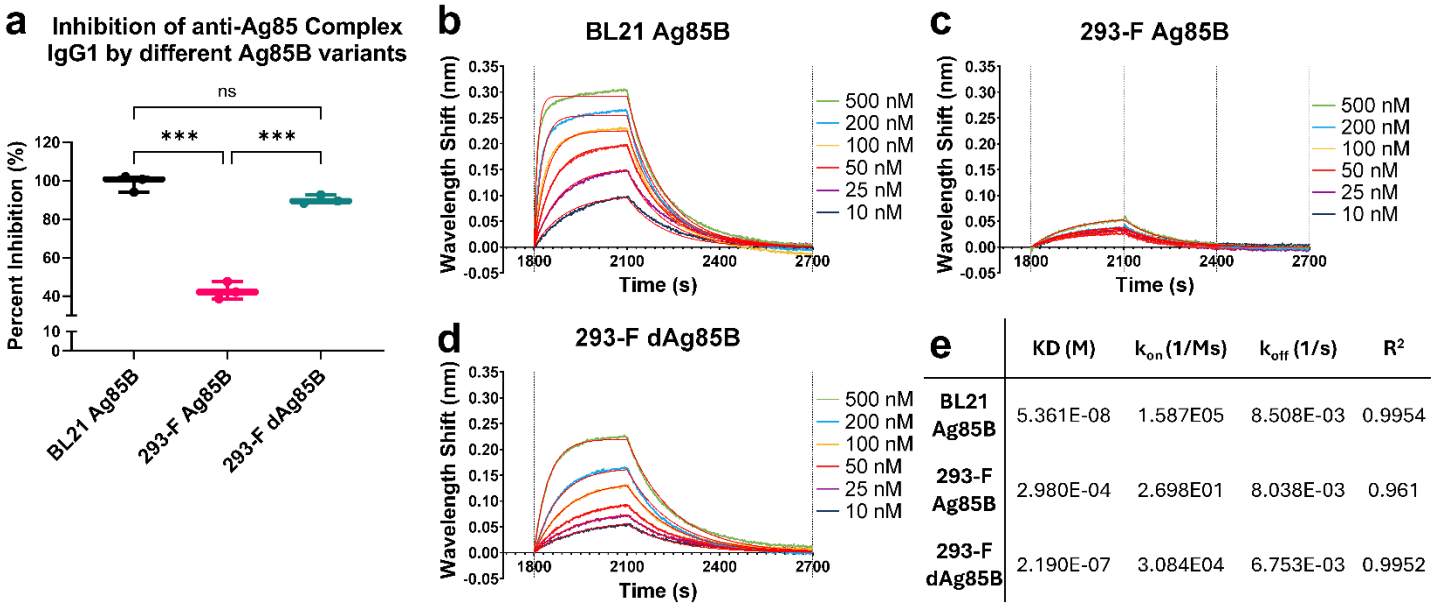


Figure 4: Molecular dynamic simulations demonstrate the impact of Ag85B N-glycosylation on the accessibility of known T-cell epitopes. (a) Hierarchical ensemble representation of the glycomics-informed glycopeptidomics site-specific abundant glycan structures of Ag85B generated after three 1-microsecond-long molecular dynamic simulations viewed from front, back, bottom, and top (from left to right), displaying specific T-cell epitope locations distinguished through color coding. Accessible surface areas of various T-cell epitopes summarized in Huygen et al. 2014 displayed through violin plots for **(b)** site A111-E125 (purple on panel a), **(c)** site F121-P138 (red on panel a), **(d)** site F164-S171 (green on panel a), **(e)** site K220-L228 (light-blue on panel a), and **(f)** site I245-F252 (blue on panel a). Identified N-glycosylated asparagines are underlined in the graph

771 title. For clarity, amino acids where there was no accessibility before and after adding glycans at identified N-
772 glycan sequons are colored in gray in the graph titles and are removed from the shown data on the x-axis. $****p$
773 < 0.0001, $***p < 0.001$, $**p < 0.01$, and $*p < 0.05$ represent significant differences between groups determined
774 using either Welch's t-test or the Mann-Whitney U test, depending on distribution normality. # represents
775 complete loss of accessibility after glycan addition.



e	KD (M)	k_{on} (1/Ms)	k_{off} (1/s)	R^2
BL21 Ag85B	5.361E-08	1.587E05	8.508E-03	0.9954
293-F Ag85B	2.980E-04	2.698E01	8.038E-03	0.961
293-F dAg85B	2.190E-07	3.084E04	6.753E-03	0.9952

Figure 5: Glycosylation of Ag85B impacts its antigenicity. (a) Glycosylation of Ag85B impacts the binding of a monoclonal antibody against Ag85 Complex (BEI Resources, Clone P4B3G2), and non-glycosylated mutant variants of Ag85B rescue binding, determined by competitive ELISA. Values represent the mean ± SD of the percentage of inhibition. $****p < 0.0001$, $**p < 0.01$, and ns represent significant differences between groups determined using two-way ANOVA with Tukey's multiple comparisons test. Biolayer interferometry using AMC anti-mouse Fc Biosensors loaded with a monoclonal antibody against Ag85 Complex (from panel A) binding association and dissociation to **(b)** BL21 Ag85B, **(c)** glycosylated Ag85B, and **(d)** non-glycosylated mutant variant of Ag85B. **(e)** KD, k_{on}, k_{off}, and R^2 were determined by global 1:1 fitting analysis using the Sartorius Octet Analysis Suite.