scientific reports

OPEN

Conserved topology of virus glycoepitopes presents novel targets for repurposing HIV antibody 2G12

Nathaniel L. Miller[1,2,3], Vidya Subramanian[1,2], Thomas Clark[1,2], Rahul Raman[1,2] & Ram Sasisekharan[1,2,4]✉

Complex glycans decorate viral surface proteins and play a critical role in virus–host interactions. Viral surface glycans shield vulnerable protein epitopes from host immunity yet can also present distinct "glycoepitopes" that can be targeted by host antibodies such as the potent anti-HIV antibody 2G12 that binds high-mannose glycans on gp120. Two recent publications demonstrate 2G12 binding to high mannose glycans on SARS-CoV-2 and select Influenza A (Flu) H3N2 viruses. Previously, our lab observed 2G12 binding and functional inhibition of a range of Flu viruses that include H3N2 and H1N1 lineages. In this manuscript, we present these data alongside structural analyses to offer an expanded picture of 2G12-Flu interactions. Further, based on the remarkable breadth of 2G12 N-glycan recognition and the structural factors promoting glycoprotein oligomannosylation, we hypothesize that 2G12 glycoepitopes can be defined from protein structure alone according to N-glycan *site* topology. We develop a model describing 2G12 glycoepitopes based on N-glycan site topology, and apply the model to identify viruses within the Protein Data Bank presenting putative 2G12 glycoepitopes for 2G12 repurposing toward analytical, diagnostic, and therapeutic applications.

The epitope surfaces on antigens that are targeted by an antibody have largely been defined based on amino acid contacts between the antibody and antigen. The surface of a virus presents epitopes of varying complexities ranging from a defined region on a single protein domain to interfaces involving several copies of the surface protein along different axes of symmetry in the entire viral assembly[1,2]. In this context, many viral surface antigens are also heavily glycosylated by the host's glycosylation machinery and glycans displayed on the viral surface play multifaceted roles in virus-host interaction and adaptation of the virus to the host[3–5].

Complex glycans are added to viral surface proteins as a part of the post-translational modification during viral replication in the host. The adaptation of a virus to its host can involve mutations that either add or remove glycosylation sites so as to evade host immune response by either masking or modifying a previously exposed epitope surface[6–8]. Although complex glycans on the viral surface are not likely to elicit an immune response, the clustering of these glycosylation sites on the viral surface antigens can generate distinct types of glycans that are different from the typical glycans present on host glycoproteins[9]. Distinct high-mannose type N-linked glycan structures are present on clustered glycosylation sites of the gp120 envelope protein of the human immunodeficiency virus (HIV)[10], primarily as a result of clustered N-linked glycans sterically-interfering with α-mannosidase function[11]. The presence of high-mannose N-linked glycosylation across diverse pathogen surfaces prompted previous studies to investigate if specific glycan-binding proteins or lectins can bind to these high mannose glycans and neutralize these viruses. Specifically, mannose-binding lectins including those from human C-type lectin family[12] and plant lectins such as banana lectins[13] have been demonstrated to bind high-mannose type glycan structures displayed on surfaces of diverse pathogens such as HIV and influenza viruses.

The distinct glycosylation patterns on viral surfaces also present "glycoepitopes" that are targets of antibodies generated by the host immune response to neutralize the virus. The evidence of such glycoepitopes is best established in the case of the HIV gp120[3,14]. Several antibodies isolated from HIV patients target these glycoepitopes

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Singapore-MIT Alliance in Research and Technology (SMART), Singapore 138602, Singapore. ✉email: rams@mit.edu

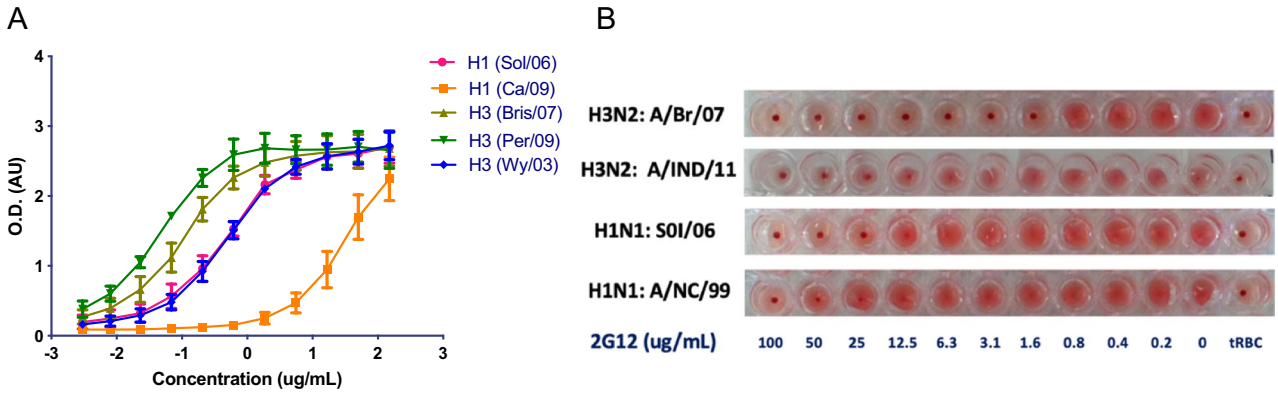


Figure 1. (**A**) Binding of 2G12 to full length recombinant HA from representative H1N1 and H3N2 strains in a quantitative ELISA (n = 3). Pink circles, HA from A/Solomon Island/2006 (H1N1); orange squares, HA from A/California/2009 (H1N1); dark green triangles, HA from A/Brisbane/2007 (H3N2); green triangles, HA from A/Perth/2009 (H3N2); blue diamond, HA from A/Wyoming/2003 (H3N2). (**B**) 2G12 inhibits agglutination of red blood cells by inactive influenza viruses. 4 HAU of inactive influenza viruses were incubated with serial dilution of 2G12 starting at 100 ug/mL for 45 min on ice followed by addition of 2% turkey red blood cells.

on gp120 and effectively neutralize the virus[15–17]. Among these antibodies the 2G12 antibody is unique in two aspects. First, instead of the conventional "Y" shape with 2 antigen binding Fab domains (VH–VL, VH′–VL′), 2G12 has an additional binding site formed by cross over of VH domains of each Fab (VH–VL′, VH′–VL and VH–VH′)[18]. Second, unlike other HIV antibodies that interact with both the glycan and the underlying protein surface ("glycan-protein glycoepitopes"[19]), 2G12 recognizes a solely glycan-based glycoepitope comprised of high-mannose type glycans at glycosylation sites including N295, N332, N339, and N392 on gp120[20,21]. Mutational studies to knockout these four glycosylation sites and nearby sites on gp120 have provided insights on the relative contributions of each glycan to the binding affinity of 2G12[22], revealing that 2G12 maintains high-affinity binding to a range of similar glycan configurations. Given this observation, in combination with evidence that 2G12 binds SARS-CoV-2 spike protein[23,24], the hemagglutinin (HA) of certain Influenza H3N2 viruses[25], and glycoproteins of *S. cerevisiae*[26], the anti-glycan antibody 2G12 is likely to target the glycoproteins of additional viruses. Yet, as 2G12-glycoprotein interactions do not depend directly on the protein surface itself, existing methods to identify novel repurposing targets for antibodies based on protein sequences and structures are not suited for predicting new targets for 2G12.

In this study, we first present experimental data describing interactions between 2G12 and a set of H1N1 and H3N2 viruses. Our results are consistent with recently published data[25] on 2G12-H3N2 interactions, yet extend these findings to certain H1N1 viruses and further demonstrate how certain N-glycan topologies on H3N2 enhance 2G12's binding, functional inhibition, and neutralization potency. Second, we present a structural analysis of 2G12 interactions across HIV[22], Flu[25], and SARS-CoV-2[23] to define a generalized model for 2G12 glycoepitopes based on N-glycan site topology at the protein surface. Finally, we hypothesized that any other viral surface glycoprotein that possesses a similar N-glycan site topology can potentially be recognized by 2G12. We therefore apply our model to survey the Protein Data Bank for glycoproteins that 2G12 putatively binds. The resulting "hits" serve as a list of viruses for which 2G12 repurposing toward analytical, diagnostic, and therapeutic applications may prove valuable.

Results

2G12 binds, functionally inhibits, and neutralizes select Influenza H1N1 and H3N2. It was recently reported than 2G12 binds and neutralizes select H3N2 Influenza A viruses[25]. At the time of this publication, our lab was evaluating unpublished data for interactions between 2G12 and a number of Influenza A viruses which we report now to expand the understanding of these interactions. Here we provide evidence indicating that 2G12 binds, functionally inhibits, and neutralizes a number of additional Influenza A viruses including H1N1 lineages (Fig. 1, Table 1). Briefly, we find that 2G12 binds H1N1 A/SoI/06 and H3N2 A/Per/09, A/Bris/07, A/Wy/03 at concentrations below 1 ug/mL, but only very weakly binds H1N1 Ca/09 (Fig. 1A).

Further, we evaluated 2G12's ability to functionally inhibit the HA of certain representative H1N1 and H3N2 viruses. Turkey red blood cells (tRBCs) predominantly express α2 → 6 sialylated glycan receptors that specifically bind to HA from human-adapted viruses. We thus measured the ability of 2G12 to inhibit agglutination of tRBCs by representative human-adapted H1N1 and H3N2 viruses (Fig. 1B). We find functional inhibition of HA from both H3N2 and H1N1 viruses, with the strongest effect identified for H3N2 A/Br/07 (inhibition as low as 1.6 ug/mL). Inhibition of H1N1 viruses was weaker, with the strongest effect identified for A/SoI/06 (25 ug/mL).

Finally, consistent with 2G12's effective inhibition of tRBC agglutination, we find that 2G12 neutralizes live H3N2 viruses. 2G12 potently inhibited the induction of CPE by Bris/07 with an EC50 of 0.1ug/ml. Inhibition of CPE induction by 2G12 was also observed for A/Sydney/97 (Syd/97) and A/Perth/11/2009 (Perth/09) H3N2 strains, albeit at a 10–20-fold lower concentration (Table 1). 2G12 failed to neutralize A/Vic/75. Importantly, the viruses surveyed here feature similar yet distinct N-glycans site topologies on the HA stalk, often in a ± 1 N-glycan

H3N2 influenza strains	Inoculum (PFUs)	EC50 (ug/mL)	CC50 (ug/mL)	SI
A/Brisbane/10/2007	40	0.1	> 100	> 10
A/Sydney/05/97	20	1.8	> 100	> 56
A/Perth/16/2009	32	1.5	> 100	> 67
A/Victoria/03/75	20	Not active	> 100	NA

Table 1. 2G12 neutralizes select H3N2 viruses in vitro. EC50, CC50, and selective index (SI) are reported according to inhibition of CPE observed in MDCK cells.

site format. We therefore sought to evaluate the totality of the binding, inhibition, and neutralization data in light of the N-glycan topology for these viruses to expand our understanding of the structural basis for 2G12 interactions.

A pair of N-glycan sites on the HA head is sufficient-but not optimal-for 2G12 interactions. Lee et al.[25] reported that a pair of N-glycans on the H3N2 HA head at sites N165 and N246 mediate overhead 2G12 binding poses and are critical for 2G12 neutralization of H3N2 viruses. Our data expand the understanding of N-glycan topology required for 2G12 interactions with HA, providing evidence that (1) homologous pairs of N-glycans on the H1N1 head are sufficient to mediate 2G12 binding and functional inhibition, and (2) glycans other than N165 and N246 on the H3N2 head significantly enhance 2G12 binding, functional inhibition, and neutralization.

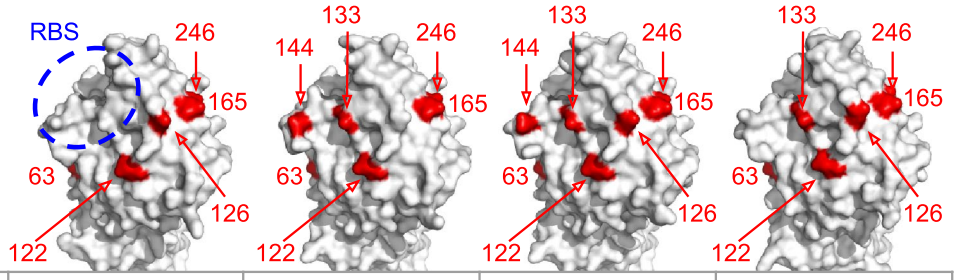
First, we find that H1N1 A/SoI/06 and A/NC/99 present N-glycans pairs on the HA head (N142, N177; Fig. 2B), and that 2G12 binds and functionally inhibits these strains (Fig. 1A,B). In contrast, H1N1 A/Ca/09 presents a more substantial stalk N-glycan cluster than A/SoI/06 and lacks a pair of head N-glycans, but is only weakly bound by 2G12 at saturating concentrations. Together, these data suggest that the pair of head N-glycans on A/SoI/06 and A/NC/99 is responsible for 2G12 binding and functional inhibition of these viruses. These data highlight the critical role that a single pair of clustered N-glycans can play at quaternary junctions, and we hypothesize that the mode of interaction between 2G12 and A/SoI/06 is likely analogous to the overhead quaternary pose reported for H3N2 viruses by Lee et al.[25].

Second, our data suggests that glycan pairs on the HA head alone are not optimal for 2G12 interactions. Examining the N-glycan site topology for H3N2 A/Ind/11 and H3N2 A/Vic/75 (Fig. 2A), which both feature N-glycan pairs in isolation, we find that these H3N2 viruses as well as the previously mentioned H1N1 viruses show weak functional inhibition (A/Ind/11 at 12.5 ug/mL, A/SoI/06 at 25 ug/mL, A/NC/99 at 25 ug/mL) and a lack of neutralization (no neutralization observed for A/Vic/75) relative to Flu viruses presenting additional head N-glycans. That is, H3N2 viruses bearing more substantial N-glycan clusters are the most potent binders (A/Per/09 and A/Br/07), inhibitors (A/Br/07 at 1.6 ug/mL), and neutralizers (A/Br/07 EC50 at 0.1 ug/mL). Note that we find 2G12 binding strength is not directly correlated with the number of head N-glycans, suggesting the specific N-glycan cluster topology influences 2G12 activity. While N-glycan pairs may be sufficient—and certain N-glycan pairs may be critical—these pairs alone do not appear optimal for 2G12 interactions.

The structural data on Flu-2G12 interactions presented by Lee et al.[25] suggests that non-overhead or partially-overhead binding poses may explain how additional head N-glycans optimize 2G12 binding and neutralization. The Lee et al.[25] negative-stain electron microscopy (nsEM) images indicate a variety of poses for 2G12 bound to HA that include poses in which 2G12 is not strictly overhead of HA and others in which 2G12 runs parallel to the vertical axis of HA[25]. These additional poses suggest that N-glycans other than N165 and N246 may be directly involved in the 2G12 binding event. We therefore next examine how different head N-glycan topologies contribute functionally to 2G12-Flu interactions.

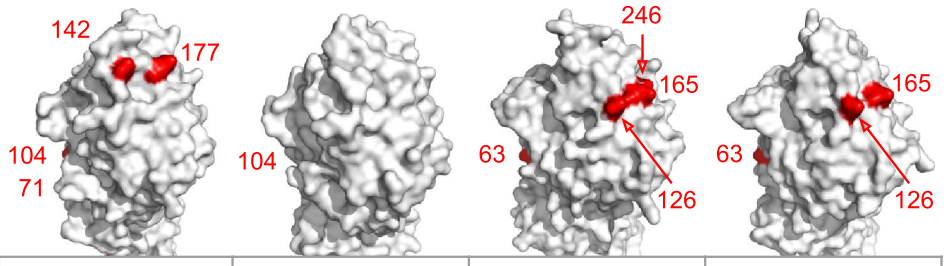
There is a binding vs. neutralization tradeoff dictated by N-glycan topology on the H3N2 HA head. While it is clear that additional N-glycan clustering on the HA head contributes to more effective 2G12 interactions, our data provide further evidence that the specific topology of head glycans drives a tradeoff between binding strength and neutralization potency. We observe that the strongest binder (A/Per/09, EC50 = 0.35 nM) is not the most potent neutralizer (EC50 = 1.5 ug/mL). Rather, A/Bris/07 neutralizes with more than an order of magnitude higher potency (EC50 = 0.1 ug/mL) despite binding with twofold lower strength (EC50 = 0.73 nM). The key difference between the N-glycan topologies for these two viruses is the additional N-glycan at site N144 for A/Bris/07. The N144 N-glycan is directly adjacent to the major functional site for HA—the receptor binding site (RBS)—offering a straightforward explanation for the significant increase in neutralization potency against A/Bris/07. That is, the N144 glycan may promote 2G12 binding poses that sterically interfere with RBS function to a greater extent than HA glycan clusters lacking an N144 glycan. Further, it is plausible that binding strength to A/Bris/07 is relatively reduced as a result of the glycan at site N133 disrupting binding poses spanning from N144 to N165/N246, as the N133 glycan would reside in the middle of the 2G12 paratope potentially creating a steric constraint for 2G12 binding. Such a mechanism has precedent in the literature on interactions between 2G12 and gp120 in which the glycan at site N411 resides in the middle of the 2G12 paratope, and it has been observed that knockout of this glycan enhances 2G12 binding strength beyond that of wild type gp120[22]. Indeed, consistent with these two hypotheses, the highest affinity monomeric glycoepitope (A/Per/09) is highly planar and can be visualized as presenting both monomeric and quaternary glycoepitopes according to the poses imaged in Lee at al.[25]. We next sought to integrate these functional observations across

A



Lineage	H3N2 Syd/97 (2YP2 + N133D, N144I)	H3N2 Wy/03 (6BKN)	H3N2 Bris/07 (6AOR)	H3N2 Per/09 (4KVN)
N-Glycan Sites	63, 122, 126, 165, 246	63, 122, 133, 144, 165, 246	63, 122, 126, 133, 144, 165, 246	63, 122, 126, 133, 165, 246

B



Lineage	H1N1 SOI/06 & NC/99 (5W6G)	H1N1 Ca/09 (3LZG)	H3N2 Ind/11 (5XRT + no Asn mutations)	H3N2 Vic/75 (4GMS)
N-Glycan Sites	71, 104, 142, 177	104	63, 126, 165, 246	63, 126, 165

Figure 2. HA head N-glycan topology of Influenza A viruses tested for 2G12 binding, functional inhibition, and neutralization. Surface representations of HA are shown in white with N-glycan sites highlighted in red and the receptor binding site (RBS) outlined in blue. (**A**) A/Syd/97, A/Wy/03, A/Bris/07, and A/Per/09 all feature the N-glycan pair at N165 and N246, plus two to four additional adjacent N-glycan sites. These viruses are more potently bound, inhibited, and neutralized by 2G12, and include N-glycans with varying degrees of proximity to the RBS. (**B**) A/SoI/06, A/NC/99, A/Ind/11, and A/Vic/75 present pairs of N-glycan sites on the HA head, while A/Ca/09 features a single N-glycan site on the HA head (reverse, not shown). H1N1 and H3N2 viruses solely presenting N-glycan pairs on the HA head are weakly bound and functionally inhibited by 2G12 (Table S1).

H1N1 and H3N2 viruses alongside existing structural models of 2G12 interactions with HIV and SARS-CoV-2 to build a topological model for 2G12 interactions with predictive potential.

Model of 2G12 topological glycoepitopes across HIV, Flu, and SARS-CoV-2. To define a generalizable model for topological glycoepitopes recognized by 2G12, we leveraged data across the three known viruses 2G12 binds to: HIV, Flu, and SARS-CoV-2. Specifically, we analyzed structural features from co-crystal complexes (6OZC[22]; 7L09[23]) and functional data from experimental binding and neutralization assays (38; Fig. 1). Further, we incorporated the theoretical basis for clustered N-glycans promoting high-mannose species[11] and performed computational glycan docking experiments to derive broad distributions of allowable 2G12 binding poses. By integrating these data, we learn from both the epitope and paratope sides of the 2G12 epitope-paratope interaction to define a generalizable topological model (Fig. 3).

As observed in the structure of 2G12 bound to HIV BG505 SOSIP.664 (6OZC[22]), 2G12 has two distinct primary binding sites with each site binding to a high mannose glycan at N295 and N392 with preference for the terminal α1–2 linked-mannose at the d1 arm of Man9[18]. Further, 2G12 presents a third binding "surface" formed at the VH–VH′ crossover interface, which interacts with additional mannose sugars in the d2 and d3 branches of two high-mannose glycans at N332 and N339[22]. Meanwhile, the 2G12-SARS-CoV-2 spike complex (7L09[23]) indicates that a single high-mannose N-glycan at N717 is capable of interacting with both the primary and secondary 2G12 binding sites simultaneously. Data from 2G12-H3N2/H1N1 interactions (38, Fig. 1) suggest that

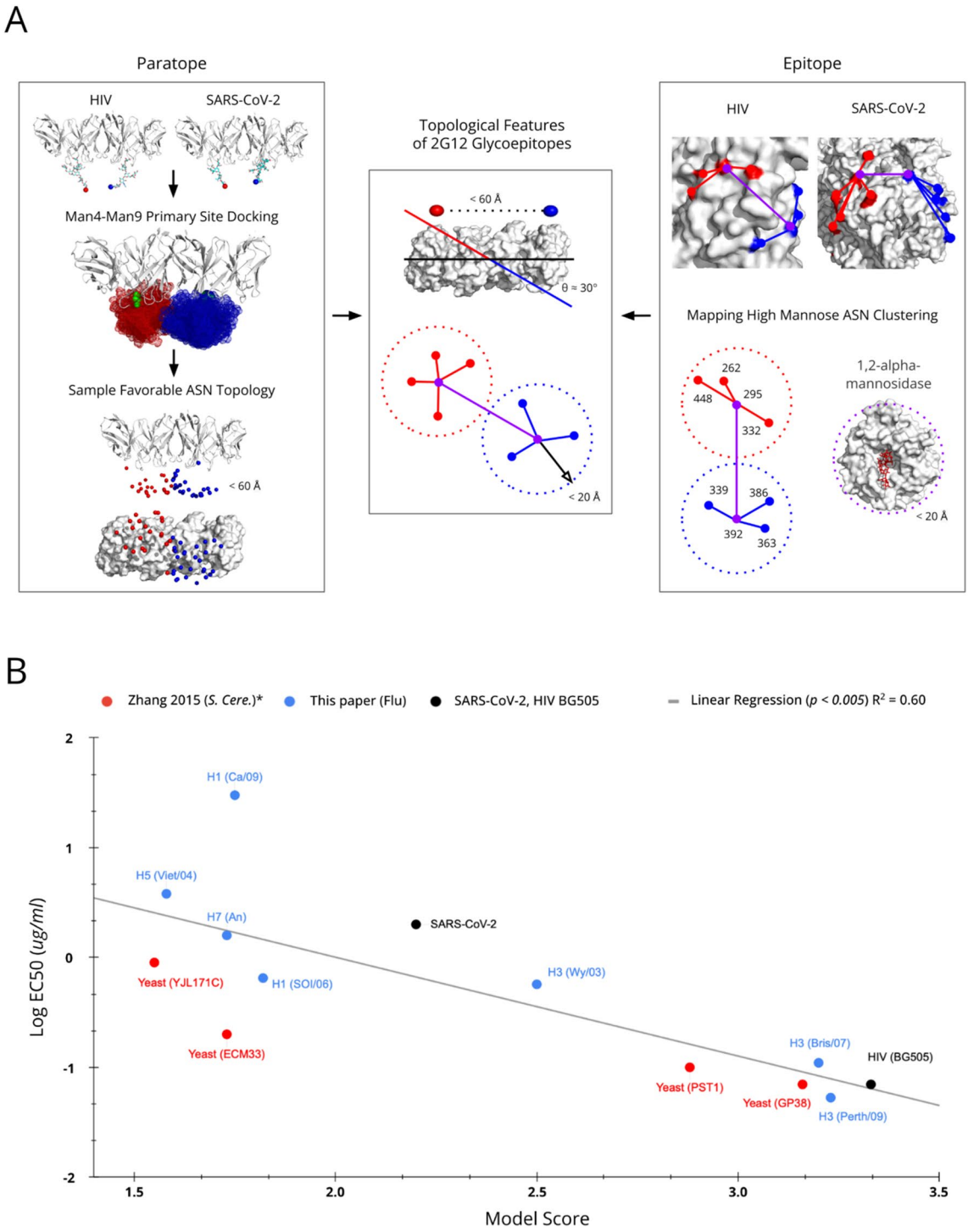


Figure 3. 2G12 Topological glycoepitope model and validation. (**A**) Left, Paratope: From the structures of 2G12 complexed to HIV BG505 (6OZC[22]) and SARS-CoV-2 spike (7L09[23]), we derive a distribution of canonical Man4–Man9 poses (shown as mesh) amenable to binding the 2G12 primary binding sites, and a corresponding distribution of N-glycan sites (shown as individual spheres). The most distant pair of N-glycan sites giving rise to a theoretically acceptable 2G12 binding pose are 60 Å apart. Right, Epitope: From the epitope side of the 2G12 interactions, we examine the clustering of primary (purple spheres) and secondary (red or blue spheres) N-glycan sites, and the steric constraints of ER α1,2-mannosidase[11] (5KIJ[29]). These features contribute to oligomannosylation at the primary Asn, secondary binding site interactions, and indirectly to the binding event via network effects (e.g., glycan support and stability). Middle: We integrate information from the epitope and paratope sides of the interaction to develop a simple feature set describing N-glycan site topologies likely to present 2G12 glycoepitopes. (**B**) 2G12 glycoepitope scores correlate with 2G12 binding strength. The apparent binding strength for 2G12 and 13 viral and non-viral glycoproteins measured in this study and elsewhere is plotted against the 2G12 glycoepitope score returned by the glycoepitope model for each glycoprotein. A statistically significant linear relationship is observed ($p < 0.005$), suggesting that the glycoepitope model provides a degree of predictive utility. *Note that for the glycoproteins measured in Zhang et al., 2015[26], solved structures do not exist and so high confidence Alpha Fold 2[28] predictions were employed (90%+ predicted accuracy for all sites). The *S. cerevisiae* glycoprotein predictions are limited to monomeric glycoepitopes, potentially explaining their systemic underscoring relative to other glycoprotein structures scored that include both monomeric and quaternary glycoepitopes.

Figure 4. PDB search for 2G12 repurposing targets. (**A**) Distribution of 2G12 glycoepitope scores for 1250 + viral glycoproteins in PDB. Each structure analyzed is plotted according to its top scoring 2G12 glycoepitope and ranked against all other glycoproteins in the set. Approximately 60% of viral glycoproteins did not present an N-glycan topology satisfying the minimum criteria of the model. The weakest known 2G12 binding glycoproteins score in the 80th to 90th percentiles and the top scoring glycoproteins are near exclusively HIV/SIV or HIV-derived glycoproteins (e.g., BG505). (**B**) Top scoring viral glycoproteins. Rows corresponding to known 2G12 binders are highlighted in light blue. *For *S. cerevisiae* glycoproteins, UniProt accession codes are shown. (**C**) Three hits with logical mechanisms of neutralization are shown, with the Asn residues for the highest scoring glycoepitope for each glycoprotein shown as red spheres and the putative 2G12 footprint shown as a red dotted oval. Left, the post-fusion glycoprotein B of HCMV features a 96th percentile scoring glycoepitope for which half of the glycoepitope footprint overlaps with a known neutralizing epitope (PDB: 5C6T[30], nAb 1G2 shown in blue). Note that the pre-fusion conformation also presents a putative, albeit lower scoring, 2G12 glycoepitope. Middle, the spike protein of PEDV features a number of high-scoring glycoepitopes (96th percentile) that suggest inter- and intra-chain locking (PDB: 6VV5[31]). Right, the receptor binding protein of Sosuga Virus (SosV) features a glycoepitope scoring in the 95th percentile that overlaps with the expected receptor binding site, which is shown in blue (PDB: 6SG8[32]).

a single pair of high-mannose N-glycans is capable of meditating 2G12 binding via the primary and secondary binding sites, though there is insufficient structural data to determine the nature of the secondary interactions.

Considering the totality of this structural and functional data it is appealing and logical to define a glycoepitope model based on N-linked glycans whose terminal mannose form a rhombus shape mirroring the 2G12 paratope, with the long axis defined by the primary binding sites (VH–VL′, VH′–VL) and the short axis defined by the secondary surface (VH–VH′). However, such a definition would limit application of the model as it would depend critically on N-glycan structures which tend not to be fully solved in most glycoprotein structures. Further, even structures that include full-length high-quality N-glycans (or computational approaches to populating N-glycans) would still not represent the range of conformations flexible glycans can adopt during 2G12 binding. Thus, we instead build our model based on the topology of Asn residues at N-glycan sites, as the minimal binding determinant for 2G12 appears to be a pair of N-glycan sites presenting high-mannose N-glycans capable of interacting with the two 2G12 primary binding sites, while nearby N-glycans sites clustered around these two primary N-glycan sites promote high-mannose species at the primary sites and interact with the secondary VH–VH′ binding surface. Toward not overfitting the limited structural information available (i.e., scoring glycoepitopes based on an "ideal" distance between primary binding N-glycan sites for gp120 of ~ 21.9 Å), we selected the simplest feature set that satisfies these objectives while also modeling distributions of possible N-glycan poses (Fig. 3), specifically:

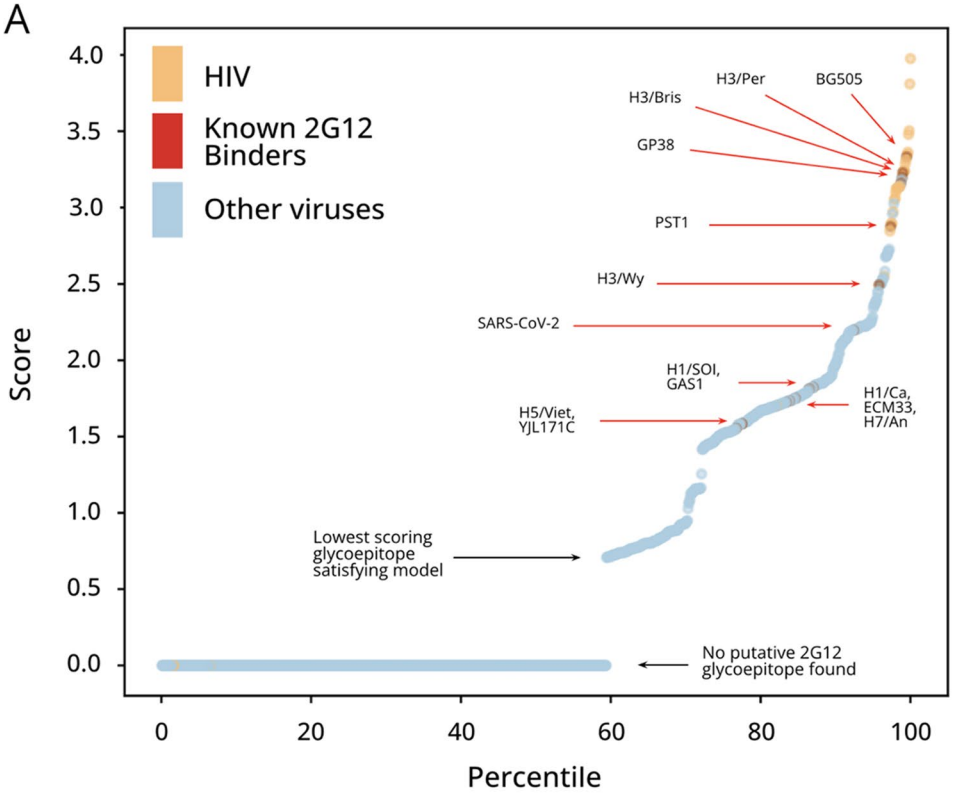
Feature 1 A surface exposed pair of N-glycan sites, wherein the two Asn residues (termed *primary Asn*) can be docked with canonical Man4-Man9 glycans such that there is a pose in which both glycans interact via a terminal mannose with the 2G12 primary binding sites.

Feature 2 Additional surface exposed N-glycan sites (termed *secondary Asn*) whose Asn residue is sufficiently close to a *primary Asn* to (1) sterically impair N-glycan processing by α1,2-mannosidase and (2) interact with the secondary binding surfaces of 2G12, where the number and proximity of these secondary sites is proportional to the likelihood of mediating these effects.

We derive primary Asn site distributions from a glycan docking exercise using 25 canonical Man4-Man9 glycans obtained from the Glycan Fragment Database[27], and clustering radii are computed based on steric constraints of ER α1,2-mannosidase[11] (Fig. 3A; see "Materials and methods" section). From this derivation, we find that satisfactory primary Asn pairs lie along a 60 Å axis that is rotated approximately 30° relative to the 2G12 paratope, and that secondary Asn tend to reside within 20 Å of each primary Asn. We subsequently wrote a lightweight algorithm to score all combinations of N-glycan sites on a given glycoprotein structure based on these two features (see "Materials and methods" section).

Toward validating the 2G12 glycoepitope model, we examined the relationship between model score and apparent 2G12 binding strength to 13 viral and non-viral glycoproteins (Fig. 3B). We identified 13 glycoproteins with measured 2G12 binding strength in the literature or in our lab, consisting of the following antigens: HIV-derived BG505[22], SARS-CoV-2 spike protein[23], Hemagglutinin from seven Influenza lineages spanning four sub-types, and four non-viral glycoproteins derived from *S. cerevisiae*[26]. Structures of the four *S. cerevisiae* glycoproteins are not solved and so were predicted via Alpha Fold 2[28], which returned predicted structures with high confidence (90%+) at all sites. We found a linear correlation between the apparent 2G12 binding strength and the modeled glycoepitope score ($R^2 = 0.60$, $p < 0.005$), suggesting that the model offers predictive value for identifying high-affinity 2G12 glycoepitopes. Notably, the *S. cerevisiae* glycoproteins appear to be systematically underscored by the algorithm. One potential explanation for this divergence from the trend is that the *S. cerevisiae* glycoepitopes are scored based solely on monomeric presentation. That is, the 2G12 binding experiments were performed under conditions in which quaternary interactions may occur, but the *S. cerevisiae* glycoproteins are scored by the model according to monomeric glycoepitope presentation because no structures of these glycoproteins exist in the proper quaternary biological context. Having defined a model for scoring 2G12 glycoepitopes based on N-glycan site topology, we next apply it to identify viruses with putative 2G12 glycoepitopes in the Protein Data Bank.

A



B

PDB*	Virus Hit	Score
6VY2, 6V0R, etc.	HIV/SIV or HIV/SIV-derived (CH505, BG505, etc.)	3.98, 3.33
4KVN, 6AOR, 6PDX, 6MZK etc.	Influenza A virus H3N2 A/Perth/16/2009; A/Brisbane/10/2007, etc.	3.22
P38616, Q12355	Yeast Proteins - Strongest Binders (GP38, PST1)	3.16, 2.88
6CV0	Infectious bronchitis coronavirus (IBV-γ) spike protein	2.58
5C6T, 2WY3, 3D2U	HCMV glycoprotein B (post-fusion)	2.54
3I1K, 3I1L	Porcine torovirus Hemagglutinin-Esterase	2.49
6SG8	Sosuga virus receptor binding protein	2.45
6N41	H1N1 A/Netherlands/002P1/1951	2.36
7CN9, 6M16, 7BYR, 6Z43, 6ZOW	SARS-CoV-2 Spike	2.25
6THB, 6THG	Cedar Virus attachement glycoprotein	2.24
6VV5, 6M15	Porcine epidemic diarrhea virus (PEDV) spike protein; Swine acute diarrhea syndrome coronavirus	2.21
6Q04, 6NB3, 6ZGF, 5N11, etc.	MERS-CoV S protein; RaTG13 Bat Coronavirus Spike Protein, OC43 NL/A/2005, etc.	2.22
2H6O	Epstein Barr Virus Major Envelope Glycoprotein gp350 (Human gammaherpesvirus 4 gp350)	2.2
3OQ3, 6BEB	Poxvirus Decoy Receptor	2.2
6PCI	Ebola Virus G Protein Trimer	2.2
6P91	Lassa Virus Glycoprotein	2.2
6M15, 6Y3Y	HKU spike, HKU1	2.17
6GH8	Lujo virus glycoprotein	1.94
6H70	GI.1 human norovirus protruding domain	1.89
4WST	H7N9 A/Taiwan/1/2013	1.83
4YY0	H6N1 A/chicken/Taiwan/A2837/2013	1.83
P22146, P38248, P46992	Yeast Proteins - Weaker Binders (GAS1, ECM33, YJL171C)	1.82, 1.73, 1.55
6UYM	Hepatitis C Virus Envelope Glycoprotein	1.71
3UBQ, 3AL4, 5IBL, etc.	H1N1 A/California/04/2009	1.63
4CQV, 3ZNM, 4KTH	Flu H5N1 A/turkey/Turkey/1/2005; Vietnam/1194/2004; Hubei/1/2010	1.58

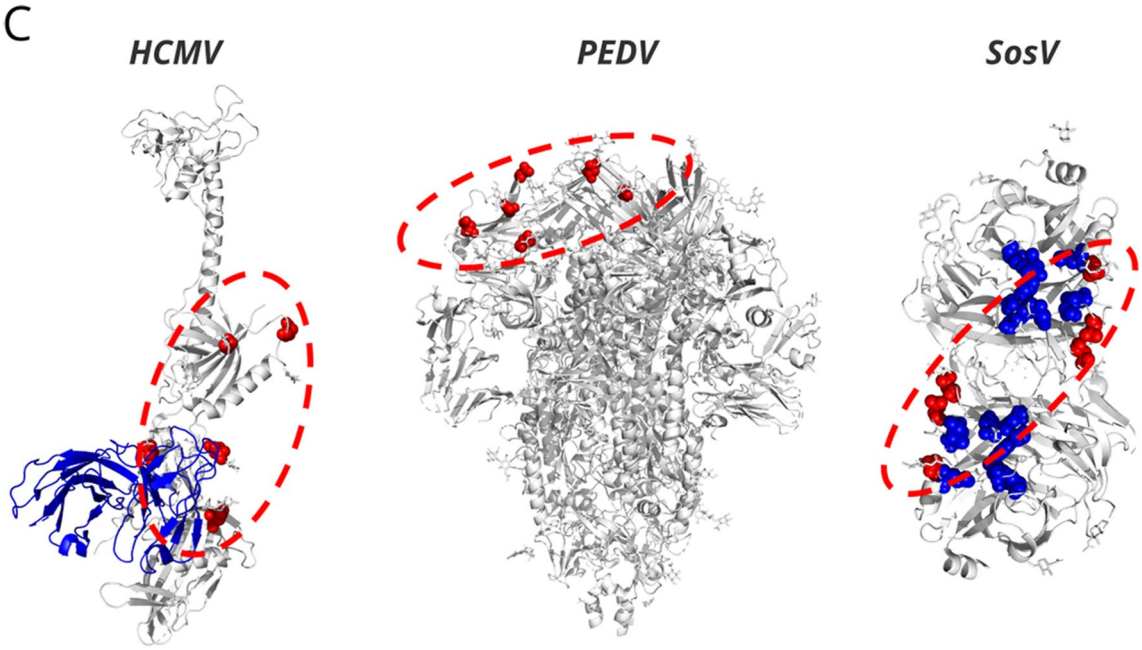


Figure 4. (continued)

Predicting putative 2G12 glycoepitopes on novel viruses. Finally, we sought to apply our 2G12 glycoepitope model to search for additional viruses that might be bound by 2G12 toward repurposing 2G12 for analytical, diagnostic, or therapeutic purposes. We identified over 1250 viral glycoproteins in the Protein Data Bank and scored them using our 2G12 glycoepitope model (Fig. 4). We obtained a score-percentile distribution in which approximately 40% of viral glycoproteins in the set display a pair of N-glycan sites satisfying the minimal topological constraints for presenting a potential 2G12 glycoepitope if both of these N-glycan sites were occupied by high-mannose N-glycans (Fig. 4A). However, it is unlikely that 2G12 binds these lowest scoring glycoproteins as the lowest scoring glycoproteins do not offer sufficient secondary N-glycan sites to promote high-mannose species at the primary sites or to interact with the secondary 2G12 binding surface. Rather, the lowest experimentally-verified 2G12 binders occur at the 80th score percentile (H5/Viet/04 and YJL171C) and the strongest known binders (EC50 ≤ 0.1 ug/mL) score in the 95th percentile or higher. Further, the 98th and above percentiles contain near exclusively HIV/SIV Env and HIV/SIV constructs (e.g., BG505 and CH505) reflecting the unique degree of N-glycan clustering within the HIV intrinsic mannose patch.

We therefore defined a set of "hits" as all glycoproteins with scores above the lowest scoring known 2G12 binder (H5/Viet/04) for manual investigation of structures presenting 2G12 glycoepitopes in a plausible biological context (Fig. 4B). During this curation process we eliminated enzymes, non-structural proteins, and glycoepitopes presented proximal to membranes or multimeric interfaces such that 2G12 is unlikely to bind in the context of the full virion. We further moved structures of viruses that do not infect eukaryotic cells and those containing multiple copies of glycoproteins in a non-biologically relevant conformation (e.g., in the crystal lattice). For a number of these hits, including HCMV glycoprotein B, porcine epidemic diarrhea virus (PEDV) spike protein, and Sosuga virus (SosV) receptor binding protein, the predicted 2G12 glycoepitope overlaps known neutralizing antibody epitopes or functionally critical domains and therefore presents a plausible basis for 2G12 neutralization (Fig. 4C). Future investigation is warranted to determine if 2G12 offers diagnostic, analytical, or therapeutic utility for these viruses presenting putative 2G12 glycoepitopes.

Discussion

2G12 was one of the first identified broadly neutralizing human antibodies directed against the high-mannose glycans of HIV gp120[33,34]. The unique structure of 2G12 presenting three distinct high-mannose glycan binding surfaces was subsequently described[15,18,35]. Following recent reports that 2G12 interacts with Influenza H3N2 and SARS-CoV-2, we report results demonstrating that 2G12 binds and functionally inhibits select Influenza A H1N1 viruses and additional H3N2 viruses. Through follow-up work to validate our algorithm, we further observed 2G12 binding to H5 and H7 strains. Further, our results indicate that pairs of N-glycans on the HA head are sufficient to mediate 2G12-HA interactions with H1 HA, which is consistent with the critical role identified by Lee et al., for N165 and N246[25] on H3 HA. Yet, our results also indicate that additional glycans beyond a sole pair substantially enhance 2G12 interactions and contribute to more potent neutralization.

Additional N-glycans may mediate enhanced effects through a variety of direct and indirect mechanisms. Potential direct effects include glycans proximal to the HA receptor binding site (RBS) interacting directly with the 2G12 paratope, which may favor 2G12 poses that sterically obstruct HA function. Indeed, a variety of 2G12 poses for a given HA are observed by Lee et al.[25], suggesting certain N-glycan topologies may more strongly favor

certain overhead vs. RBS-proximal poses. On the other hand, the additional N-glycans likely exert an indirect functional effect by modulating glycosylation state at the primary binding site glycans or contributing to glycan network effects such as stabilization. Indeed, Lee et al., present evidence that N165 and N246 are nearly entirely Man9 species, but it was previously unclear whether this effect is dependent on the other H3N2 head glycans or if the pair is sufficiently clustered to induce Man9 alone. Our results demonstrating 2G12 binding to a pair of N-glycans on the HA head of select H1N1 viruses suggest that a lone pair is indeed sufficient to induce high-mannose N-glycosylation and bind 2G12. Still, 2G12 is capable of binding a variety of Man4+ species[21], and so our H1N1 results alone do not indicate whether the pair promotes Man9 or other Man4+ species, leaving open the possibility that additional N-glycans on the H3N2 head affect the high-mannose species present at the critical binding pair. Our results further suggest that subtle differences in HA head glycan topology at sites beyond the N165/N246 pair produce substantial effects including a tradeoff between binding and neutralization for H3N2 viruses.

Our study has implications in two specific areas. *First*, our findings have direct relevance to influenza research and drug development efforts. Since inception into human populations, influenza viruses have acquired mutations that add or subtract glycosylation sites, with additions occurring in intervals of about 5 to 7 years[36]. Interestingly, the introduction of the swine origin 2009 H1N1 virus as the dominant seasonal strain reset the clock on glycosylation sites in the RBS of HA in H1N1 viruses. Compared to previous seasonal strains the 2009 H1N1 lacked glycosylation sites in the RBS of HA resulting in the unmasking of novel antigenic epitopes[37]. In contrast, despite the identification of triple reassortant H3N2 viruses lacking glycosylation sites in the RBS of HA, these strains haven't replaced the seasonal virus as the dominant strain and hence the glycosylation clock hasn't been reset for H3N2 viruses yet. Thus, the HA from seasonal H3N2 viruses since 1968 until 2011 has been accumulating glycosylation sites in the RBS. A recent glycomic analysis demonstrated that these high-mannose N-glycan sites contribute to pathogenicity[38]. Based on this work, 2G12 may have utility as a diagnostic and analytical tool to monitor the evolution of glycosylation sites in circulating H3N2 viruses. Additionally, given the high neutralization potency of 2G12 against select H3N2 strains, 2G12 might prove useful as a therapeutic. Further, 2G12 should be considered if there is ever a need for a rapid response against an emerging strain with novel N-glycosylation sites proximal to the RBS as 2G12 may potently neutralize such a strain.

Second, our approach to define glycoepitope similarity based on structural and geometric considerations goes beyond the conventional definition based on protein sequence similarity and permits identification of novel target epitope surfaces that can be successfully targeted by anti-glycan antibodies on diverse antigens. While until recently 2G12 was the only known antibody recognizing solely N-glycan-based epitopes, an additional structural class of antibodies known as Fab-dimerized glycan-reactive (FDG) antibodies was recently reported[39]. Further, the set of mutations required to mediate VH domain exchange on the 2G12 scaffold is known[40], suggesting the possibility of converting conventional antibodies with glycan-binding domains (e.g., the PGT series[16]) to domain-exchanged format, potentially mediating enhanced glycan-specific binding. Converted antibodies could be subsequently optimized for interactions with glycoepitopes bearing other glycan species or topologies than those preferred by 2G12 or the FDG antibodies documented to date. Therefore, the approach presented in this study may generalize to guide repurposing of additional anti-glycan antibodies to target topologically similar glycoepitope surfaces on diverse antigens. Further, given recent advances in predicting protein structure from sequence alone[41,28] and the application of Alpha Fold 2 to predict and score the structure of *S. cerevisiae* glycoproteins in this paper which correlated well with measured binding strength, the 2G12 topological model could also be applied to identify novel viral glycoproteins targets for 2G12 from the vast sequence space of emerging viruses.

Materials and methods

Antibody expression and purification. 2G12 variable heavy and light chain were cloned into full length IgG1 expression vector pcDNA3.3 HC and pcDNA3.3 LC. The antibodies were transiently expressed in Freestyle 293 expression system (Invitrogen, CA) and purified using Protein A Sepharose affinity chromatography (Pierce Biotech, GE, UK). The purified antibodies were concentrated, and buffer exchanged into 1× PBS using 30 kDa spin filter columns (Milipore, USA) and quantified using a nanodrop.

Enzyme linked immunoabsorbent assay (ELISA). For ELISA, Maxisorp high-binding 96 well microtiter plates were used to coat the antigen. All steps of the ELISA were performed at room temperature and all volumes used were 100 uL unless otherwise stated. Wells were coated with 50 ng of gp120 JR-FL or 200 ng/well of the various full length HA trimer overnight at 4 °C. Plates were washed three times with 1× PBS/0.05% (vol/vol) Tween 20, blocked for 1 h with 5% Blotto/PBS/0.05% Tween 20 and washed three times with PBS/0.05% Tween 20. Serial dilutions of all antibodies in PBS/0.05% Tween 20 were incubated for 2 h, and the unbound antibody was removed by washing three times as described above. The binding was detected via rabbit anti-human Fcγ HRP conjugate (1:10,000, Jackson Immuno Research, PA) and TMB substrate at 405 nm.

Viruses. The following inactive viruses were used in the study. A/Solomon Island/2006 (H1N1) (Meridien Life Sciences), A/New Caledonia/99 (H1N1) (Fitzgerald), and A/Brisbane/2007 (H3N2) (Meridien Life Sciences).

Hemagglutination inhibition assay. Hemagglutination inhibition assays were performed using turkey red blood cells (tRBC; Rockland, cat: R408-0050). Briefly, the pooled and washed 10% tRBC were centrifuged at 1000 rpm to pellet the cells. The supernatant was removed, and the pellet was resuspended in 1× PBS. This process was repeated three times and the cells were finally resuspended in 1× PBS to give a final concentration of 2% red blood cells. Twofold serial dilution of the different viruses was incubated with an equal volume of 2%

tRBC to determine the HA units for each of the viruses. For the hemagglutination inhibition assay, 4-HA units of each virus was pre-incubated on ice for 45 min with serial dilution of the various antibodies. At the end of the incubation, an equal volume of 2% tRBC was added and incubated on ice for an additional 1 h.

In vitro neutralization against select strains. 2G12 was evaluated against a select set of H3N2 virus strains via a standardized cytopathic effect (CPE) inhibition assay at The Institute for Antiviral Research (IAR) at Utah State University. Briefly, 8 serial dilutions of 2G12 were added to MDCK cells (IAR) infected with the specific H3N2 viruses. PFUs were measured according to cell culture infective dose (CCID50). The ability of 2G12 to prevent formation of CPE was measured 7 days post infection by neutral red uptake assay. In parallel, the toxicity was also determined and the selectivity index (SI) against each virus was calculated.

Statistical analyses. The following statistical analyses were used in this study. Figure 1: A nonlinear log(agonist) vs. response curve with four parameters was fit to the observations. The error bars represent standard error of the mean (SEM) across three replicates. The R^2 describing goodness of fit is > 0.90 for all curves. Figure 3: Linear regression was performed via the SciPy[42] linregress function in Python. The linear regression was evaluated for statistical significance using a Wald test with t-distribution. The null hypothesis was defined as a slope of zero.

2G12 glycoepitope-paratope analysis. The 2G12 glycoepitope epitope-paratope analysis was carried out using models of 2G12 bound to HIV (6OZC[22]) and SARS-CoV-2 (7L09[23]). Epitope**:** For each model, 5 canonical N-glycans for each of man4-man9 (from the Glycan Fragment Database[27]) were docked into each primary binding site using the terminal mannose for alignment. The distribution of possible N-glycan sites giving rise to the array of N-glycan poses was estimated from the position of the terminal GlcNAc. Paratope: For each model, the N-glycan sites interacting with the 2G12 paratope were visualized and annotated in PyMol[43] which were then used to construct network models toward building intuition for feature definition. To define the clustering radius of inclusion for secondary N-glycan sites, we analyzed the surface of ER α1,2-mannosidase adjacent to the enzymatic pocket as previously proposed[11] (5KIJ[29]). Validation: Apparent binding constants were estimated according to EC50 for HIV[22], SARS-CoV-2[23], Flu (this study), and *S. cerevisiae*[26] glycoproteins. For H1 and H3 lineages, EC50's from Fig. 1 were used. For H5 and H7 lineages, EC50s using ELISA as decribed above but with a single replicate (data not shown). For scoring *S. cerevisiae* glycoproteins, high confidence (90%+) Alpha Fold 2[28] predicted structures were employed according to Uniprot accession codes P38616, Q12355, P22146, P38248, and P46992. For the other glycoproteins, scores were computed on the following structures: 6AOR[44], 4KVN[45], 6BKN[46], 3LZG[47], 5UJZ[48], 6V0R[49], 6VXX[50], 3ZNM[51], 4LKI[52]. Note that the biological assembly was used to populate trimeric HA in cases where the given structure was monomeric. Linear regression was performed via the SciPy[42] linregress function in Python.

2G12 glycoepitope search. The Protein Data Bank was first searched for structures annotated as derived from viruses and containing at least three putative N-glycan sites, and all hits were downloaded. For each structure, surface exposure was calculated for all N-glycan sites using a relative surface area (RSA) of 0.2 via the FreeSASA library[53] and the theoretical MaxSASA values derived by Tien et al.[54]. For each surface exposed N-glycan site, the structure was searched for all other surface-exposed N-glycan sites within 60 Å (Feature 1). For each identified pair, all surface-exposed N-glycan sites whose ND2 atoms are within 20 Å of the ND2 atom of a pair-residue were identified and assigned to the nearest pair-residue (Feature 2). Secondary N-glycan sites were weighted from 0.5 to 1 depending on distance from the primary site (Feature 2). Finally, for each structure all glycoepitopes satisfying the minimal constraints were scored and ranked according to the sum of the square roots of the total weighting for both primary binding clusters. Top hits were manually curated for high relevance and feasible biological context.

Received: 21 July 2021; Accepted: 17 January 2022
Published online: 16 February 2022

References

1. De Filette, M. *et al.* An influenza A vaccine based on tetrameric ectodomain of matrix protein 2. *J. Biol. Chem.* **283**(17), 11382–11387. https://doi.org/10.1074/jbc.M800650200 (2008).
2. Robinson, J. E. *et al.* Quaternary epitope specificities of anti-HIV-1 neutralizing antibodies generated in rhesus macaques infected by the simian/human immunodeficiency virus SHIVSF162P4. *J. Virol.* **84**(7), 3443–3453. https://doi.org/10.1128/JVI.02617-09 (2010).
3. Crispin, M., Ward, A. B. & Wilson, I. A. Structure and immune recognition of the HIV glycan shield. *Annu. Rev. Biophys.* **47**, 499–523. https://doi.org/10.1146/annurev-biophys-060414-034156 (2018).
4. Grant, O. C., Montgomery, D., Ito, K. & Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **10**(14991), 2020. https://doi.org/10.1038/s41598-020-71748-7 (2020).
5. Tate, M. D. *et al.* Playing hide and seek: How glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection. *Viruses* **6**(3), 1294–1316. https://doi.org/10.3390/v6031294 (2014).
6. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**(6929), 307–312. https://doi.org/10.1038/nature01470 (2003).
7. Kobayashi, Y. & Suzuki, Y. Evidence for N-glycan shielding of antigenic sites during evolution of human influenza A virus hemagglutinin. *J. Virol.* **86**(7), 3446–3451. https://doi.org/10.1128/JVI.06147-11 (2012).

8. Wibmer, C. K. *et al.* Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. *PLoS Pathog.* **9**(10), e1003738. https://doi.org/10.1371/journal.ppat.1003738 (2013).
9. Crispin, M. & Doores, K. J. Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design. *Curr. Opin. Virol.* **11**, 63–69. https://doi.org/10.1016/j.coviro.2015.02.002 (2015).
10. Doores, K. J. *et al.* Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci. U.S.A.* **107**(31), 13800–13805. https://doi.org/10.1073/pnas.1006498107 (2010).
11. Behrens, A. J. & Crispin, M. Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* **44**, 125–133. https://doi.org/10.1016/j.sbi.2017.03.008 (2017).
12. Ji, X., Gewurz, H. & Spear, G. T. Mannose binding lectin (MBL) and HIV. *Mol. Immunol.* **42**(2), 145–152. https://doi.org/10.1016/j.molimm.2004.06.015 (2005).
13. Covés-Datson, E. M. *et al.* A molecularly engineered antiviral banana lectin inhibits fusion and is efficacious against influenza virus infection in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **117**(4), 2122–2132. https://doi.org/10.1073/pnas.1915152117 (2020).
14. Falkowska, E. *et al.* Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity* **40**(5), 657–668. https://doi.org/10.1016/j.immuni.2014.04.009 (2014).
15. Sanders, R. W. *et al.* The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. *J. Virol.* **76**(14), 7293–7305. https://doi.org/10.1128/jvi.76.14.7293-7305.2002 (2002).
16. Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**(5950), 285–289. https://doi.org/10.1126/science.1178746 (2009).
17. Pejchal, R. *et al.* A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science* **334**(6059), 1097–1103. https://doi.org/10.1126/science.1213256 (2011).
18. Calarese, D. A. *et al.* Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* **300**(5628), 2065–2071. https://doi.org/10.1126/science.1083182.75 (2003).
19. Miller, N. L., Clark, T., Raman, R. & Sasisekharan, R. Glycans in virus-host interactions: A structural perspective. *Front. Mol. Biosci.* **8**, 666756. https://doi.org/10.3389/fmolb.2021.666756 (2021).
20. Wang, D., Tang, J., Tang, J. & Wang, L. X. Targeting N-glycan cryptic sugar moieties for broad-spectrum virus neutralization: Progress in identifying conserved molecular targets in viruses of distinct phylogenetic origins. *Molecules* **20**(3), 4610–4622. https://doi.org/10.3390/molecules20034610 (2015).
21. Calarese, D. A. *et al.* Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. *Proc. Natl. Acad. Sci. U.S.A.* **102**(38), 13372–13377. https://doi.org/10.1073/pnas.0505763102 (2005).
22. Seabright, G. E. *et al.* Networks of HIV-1 envelope glycans maintain antibody epitopes in the face of glycan additions and deletions. *Structure* **28**(8), 897–909. https://doi.org/10.1016/j.str.2020.04.022 (2020).
23. Acharya, P. *et al.* A glycan cluster on the SARS-CoV-2 spike ectodomain is recognized by fab-dimerized glycan-reactive antibodies. *BioRxiv* **06**(30), 178897. https://doi.org/10.1101/2020.06.30.178897 (2020).
24. Mannar, D. *et al.* Glycan Reactive Anti-HIV-I Antibodies Bind the SARS-CoV-2 Spike Protein but Not Block Viral Entry. *bioRxiv*. https://doi.org/10.1101/2021.01.03.425141 (2021).
25. Lee, C. D. *et al.* A cross-neutralizing antibody between HIV-1 and influenza virus. *PLoS Pathog.* **17**(3), e1009407. https://doi.org/10.1371/journal.ppat.1009407 (2021).
26. Zhang, H. *et al.* Antibodies elicited by yeast glycoproteins recognize HIV-1 virions and potently neutralize virions with high mannose N-glycans. *Vaccine* **33**(39), 5140–5147. https://doi.org/10.1016/j.vaccine.2015.08.012 (2015).
27. Jo, S. & Im, W. Glycan fragment database: A database of PDB-based glycan 3D structures. *Nucleic Acids Res.* **41**, D470–D474. https://doi.org/10.1093/nar/gks987 (2013).
28. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2 (2021).
29. Xiang, Y., Karaveg, K. & Moremen, K. W. Substrate recognition and catalysis by GH47 α-mannosidases involved in Asn-linked glycan maturation in the mammalian secretory pathway. *Proc. Natl. Acad. Sci. U.S.A.* **113**(49), E7890–E7899. https://doi.org/10.1073/pnas.1611213113 (2016).
30. Chandramouli, S. *et al.* Structure of HCMV glycoprotein B in the postfusion conformation bound to a neutralizing human antibody. *Nat. Commun.* **6**, 8176. https://doi.org/10.1038/ncomms9176 (2015).
31. Kirchdoerfer, R. N. *et al.* Structure and immune recognition of the porcine epidemic diarrhea virus spike protein. *Structure* **29**(4), 385–392. https://doi.org/10.1016/j.str.2020.12.003 (2021).
32. Stelfox, A. J. & Bowden, T. A. A structure-based rationale for sialic acid independent host-cell entry of Sosuga virus. *Proc. Natl. Acad. Sci. U.S.A.* **116**(43), 21514–21520. https://doi.org/10.1073/pnas.1906717116 (2019).
33. Buchacher, A. *et al.* Generation of human monoclonal antibodies against HIV-1 proteins; electrofusion and Epstein-Barr virus transformation for peripheral blood lymphocyte immortalization. *AIDS Res. Hum. Retroviruses* **10**(4), 359–359. https://doi.org/10.1089/aid.1994.10.359 (1994).
34. Trkola, A. *et al.* Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **70**(2), 1100–1108. https://doi.org/10.1128/JVI.70.2.1100-1108.1996 (1996).
35. Scanlan, C. N. *et al.* The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1–>2 mannose residues on the outer face of gp120. *J. Virol.* **76**(14), 7306–7321. https://doi.org/10.1128/jvi.76.14.7306-7321.2002 (2002).
36. Altman, M. O. *et al.* Human influenza A virus hemagglutinin glycan evolution follows a temporal pattern to a glycan limit. *MBio* **10**(2), e00204–e00219. https://doi.org/10.1128/mBio.00204-19 (2019).
37. Wei, C. J. *et al.* Cross-neutralization of 1918 and 2009 influenza viruses: Role of glycans in viral evolution and vaccine design. *Sci. Transl. Med.* **2**(24), 24. https://doi.org/10.1126/scitranslmed.3000799 (2010).
38. Heindel, D. W. *et al.* Glycomic analysis of host response reveals high mannose as a key mediator of influenza severity. *Proc. Natl. Acad. Sci. U.S.A.* **117**(43), 26926–26935. https://doi.org/10.1073/pnas.2008203117 (2020).
39. Williams, W. B. *et al.* Fab-dimerized glycan-reactive antibodies are a structural category of natural antibodies. *Cell* **184**(11), 2955–2972. https://doi.org/10.1016/j.cell.2021.04.042 (2021).
40. Huber, M. *et al.* Very few substitutions in a germ line antibody are required to initiate significant domain exchange. *J. Virol.* **84**(20), 10700–10707. https://doi.org/10.1128/JVI.01111-10 (2010).
41. Senior, A. W. *et al.* Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710. https://doi.org/10.1038/s41586-019-1923-7 (2020).
42. Virtanen, P. *et al.* SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272. https://doi.org/10.1038/s41592-019-0686-2 (2020).
43. The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC.
44. Wu, N. C. *et al.* A structural explanation for the low effectiveness of the seasonal influenza H3N2 vaccine. *PLoS Pathog.* **13**(10), e1006682. https://doi.org/10.1371/journal.ppat.1006682 (2017).
45. Nakamura, G. *et al.* An in vivo human-plasmablast enrichment technique allows rapid identification of therapeutic influenza A antibodies. *Cell Host Microbe* **14**(1), 93–103. https://doi.org/10.1016/j.chom.2013.06.004 (2013).

46. Wu, N. C. *et al.* A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. *Nat. Commun.* **9**(1), 1264. https://doi.org/10.1038/s41467-018-03663-5 (2018).
47. Xu, R. *et al.* Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **328**(5976), 357–360. https://doi.org/10.1126/science.1186430 (2010).
48. Liu, Y. *et al.* CryoEM structure of an influenza virus receptor-binding site antibody-antigen interface. *J. Mol. Biol.* **429**(12), 1829–1839. https://doi.org/10.1016/j.jmb.2017.05.011 (2017).
49. Nogal, B. *et al.* Mapping polyclonal antibody responses in non-human primates vaccinated with HIV Env trimer subunit vaccines. *Cell Rep.* **30**(11), 3755–3765. https://doi.org/10.1016/j.celrep.2020.02.061 (2020).
50. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 181(2):281–292 (2020). https://doi.org/10.1016/j.cell.2020.02.058. Erratum in: Cell 183(6):1735 (2020).
51. Xiong, X. *et al.* Recognition of sulphated and fucosylated receptor sialosides by A/Vietnam/1194/2004 (H5N1) influenza virus. *Virus Res.* **178**(1), 12–14. https://doi.org/10.1016/j.virusres.2013.08.007 (2013).
52. Shi, Y. *et al.* Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. *Science* **342**(6155), 243–247. https://doi.org/10.1126/science.1242917 (2013).
53. Mitternacht, S. FreeSASA: An open source C library for solvent accessible surface area calculations. *F1000Research* **5**, 189. https://doi.org/10.12688/f1000research.7931.1 (2016).
54. Tien, M. Z., Meyer, A. G., Sydykova, D. K., Spielman, S. J. & Wilke, C. O. Maximum allowed solvent accessibilites of residues in proteins. *PLoS ONE* **8**(11), e80635. https://doi.org/10.1371/journal.pone.0080635 (2013).

Acknowledgements
Nathan Miller was supported in part by T32 ES007020/ES/NIEHS NIH. This work was supported in part by R37 GM057073 and by SMART, Singapore.

Author contributions
N.L.M. and V.S. contributed equally to the manuscript, with N.L.M. leading the structural and computational analyses and V.S. designing and performing the experiments. R.R. and N.L.M. conceived the concept and wrote the manuscript. T.C. assisted N.L.M. with the computational analysis. R.S. was involved in supervising the research and writing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-06157-z.

Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022