

RESEARCH ARTICLE

Effect of Glycosylation on an Immunodominant Region in the V1V2 Variable Domain of the HIV-1 Envelope gp120 Protein

Jianhui Tian^{1,2}✉, Cesar A. López¹✉, Cynthia A. Derdeyn³, Morris S. Jones⁴, Abraham Pinter⁵, Bette Korber¹, S. Gnanakaran^{1*}

1 Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America, **2** Center for Biomolecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States of America, **3** Department of Pathology and Laboratory Medicine and Emory Vaccine Center, Emory University, Atlanta, Georgia, United States of America, **4** University of California Berkeley, School of Public Health, Berkeley, California, United States of America, **5** New Jersey Medical School, Rutgers University, Newark, New Jersey, United States of America

✉ These authors contributed equally to this work.

* gnana@lanl.gov



CrossMark

[click for updates](#)

OPEN ACCESS

Citation: Tian J, López CA, Derdeyn CA, Jones MS, Pinter A, Korber B, et al. (2016) Effect of Glycosylation on an Immunodominant Region in the V1V2 Variable Domain of the HIV-1 Envelope gp120 Protein. *PLoS Comput Biol* 12(10): e1005094. doi:10.1371/journal.pcbi.1005094

Editor: Alan Rein, National Cancer Institute-Frederick, UNITED STATES

Received: December 1, 2015

Accepted: August 1, 2016

Published: October 7, 2016

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](#) public domain dedication.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by NIH grants P01AI088610, R01-AI-58706 (CAD), and the Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID; UM1-AI100645) of the National Institute of Allergy and Infectious Diseases; and the Los Alamos LDRD program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Heavy glycosylation of the envelope (Env) surface subunit, gp120, is a key adaptation of HIV-1; however, the precise effects of glycosylation on the folding, conformation and dynamics of this protein are poorly understood. Here we explore the patterns of HIV-1 Env gp120 glycosylation, and particularly the enrichment in glycosylation sites proximal to the disulfide linkages at the base of the surface-exposed variable domains. To dissect the influence of glycans on the conformation these regions, we focused on an antigenic peptide fragment from a disulfide bridge-bounded region spanning the V1 and V2 hyper-variable domains of HIV-1 gp120. We used replica exchange molecular dynamics (MD) simulations to investigate how glycosylation influences its conformation and stability. Simulations were performed with and without N-linked glycosylation at two sites that are highly conserved across HIV-1 isolates (N156 and N160); both are contacts for recognition by V1V2-targeted broadly neutralizing antibodies against HIV-1. Glycosylation stabilized the pre-existing conformations of this peptide construct, reduced its propensity to adopt other secondary structures, and provided resistance against thermal unfolding. Simulations performed in the context of the Env trimer also indicated that glycosylation reduces flexibility of the V1V2 region, and provided insight into glycan-glycan interactions in this region. These stabilizing effects were influenced by a combination of factors, including the presence of a disulfide bond between the Cysteines at 131 and 157, which increased the formation of beta-strands. Together, these results provide a mechanism for conservation of disulfide linkage proximal glycosylation adjacent to the variable domains of gp120 and begin to explain how this could be exploited to enhance the immunogenicity of those regions. These studies suggest that glycopeptide immunogens can be designed to stabilize the most relevant Env conformations to focus the immune response on key neutralizing epitopes.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

Heavy glycosylation of the envelope surface subunit, gp120, is a key adaptation of HIV-1, however, the precise effects of glycosylation on the folding, conformation and dynamics of this protein are poorly understood. The network of glycans on gp120 is of particular interest with regards to vaccine design, because the glycans both serve as targets for many classes of broadly neutralizing antibodies, and contribute to patterns of immune evasion and escape during HIV-1 infection. In this manuscript, we report on how glycosylation influences an immunogenic but disordered region of gp120. Glycosylation stabilizes the pre-existing conformation, and reduces its propensity to form other secondary structures. It also stabilizes preformed conformation against thermal unfolding. These complementary effects originate from a combination of multiple factors, including the observation that having a glycosylation site adjacent to the disulfide bond further promotes the formation of beta-strand structure in this peptide.

Introduction

Glycosylation, one of the most common intracellular modifications of proteins[1], is the covalent attachment of one or more carbohydrates (glycans) at specific amino acid sequence motifs. In N-linked glycosylation, the glycan is attached to an asparagine (Asn) residue in an Asn-Xaa-Ser/Thr motif, where Xaa can be any amino acid residue except proline. Based on secondary structure predictions of protein sequences, there appears to be a strong preference for N-linked glycosylation at beta-bends[2], where approximately 70% of N-linked glycan motifs occur, while 10% and 20% occur in alpha-helices and beta-sheets, respectively[1]. Lentiviral envelope proteins are among the most heavily glycosylated proteins in nature[3]. Carbohydrates constitute half of the HIV-1 Env gp120 mass, and cover much of its surface[4].

It has long been known that gp120 can accommodate a remarkable heterogeneity in terms of the number and location of glycosylation sites [5]. This variably glycosylated protein mediates the interactions with CD4 and coreceptor molecules that are critical for viral entry. However, the effects of glycosylation on the conformation and biology of gp120 are not well understood. In general, glycosylation can stabilize protein conformation[6], accelerate protein folding[7], promote secondary structure formation[8], reduce protein aggregation [6, 9], shield hydrophobic surfaces[10], promote disulfide pairing[11], and increase folding cooperativity [12]. Others have shown that glycosylation can stabilize a protein structure against thermal unfolding due to entropic effects[13, 14]. In some cases, glycosylation can slow down the folding process by stabilizing the on-pathway folding intermediates[15]. These varied effects of glycosylation on protein stability are sensitive to the number and location of glycans in the tertiary protein structure[16–20]. Furthermore, modeling approaches typically neglect the influence of non-specific and specific protein-protein and protein-glycan interactions, which play an important role in glycosylation effects[19, 21–25]. Despite recent computational studies of glycosylation[14, 23, 26–28], the effects of carbohydrate moieties on protein conformation and folding are incompletely understood, particularly when glycosylation occurs in or near a region with an unstructured conformation.

HIV-1 gp120 contains multiple highly immunogenic regions and serves as the major target for neutralizing antibodies. The network of glycans on gp120 is of particular interest with regards to HIV-1 vaccine design, because the glycans both serve as targets for many classes of broadly neutralizing antibodies[29–32], and contribute to patterns of immune evasion and

escape during HIV-1 infection[33–39]. Elucidating the relevant forms of glycans for neutralizing antibody epitope formation could aid in the design of glycopeptide-based vaccine immunogens for HIV (for examples, see: [33, 40–42]).

In this study, we investigated glycosylation patterns in the gp120 variable domains using an updated set of 4633 HIV-1 sequences from the 2014 Los Alamos HIV Database reference alignment. Our analysis highlights the enrichment of glycosylation at the base of the 4 disulfide bonded variable loops in gp120 (Fig 1A). This led us to focus on the effects of glycosylation at two conserved sites in the V1V2 domain that are proximal to the Cys at 157: the site at 156 which is immediately adjacent, and one at 160 that is nearby (Fig 1B). This region constitutes an important glycan-dependent target of broadly neutralizing antibodies (Fig 1B) [29–31, 43, 44]. This region of V1V2 contains both conserved and highly variable positions as captured in Fig 1C. It tends to maintain a region of high positive charge, and appears to have a flexible conformation, exhibiting a tendency to adopt different conformations. In the native trimer form of HIV-1, this region of V1V2 adopts a beta-strand conformation, which persists when it is bound to broadly neutralizing antibodies PG9 and PG16 [30, 44–46]. However, the V1V2

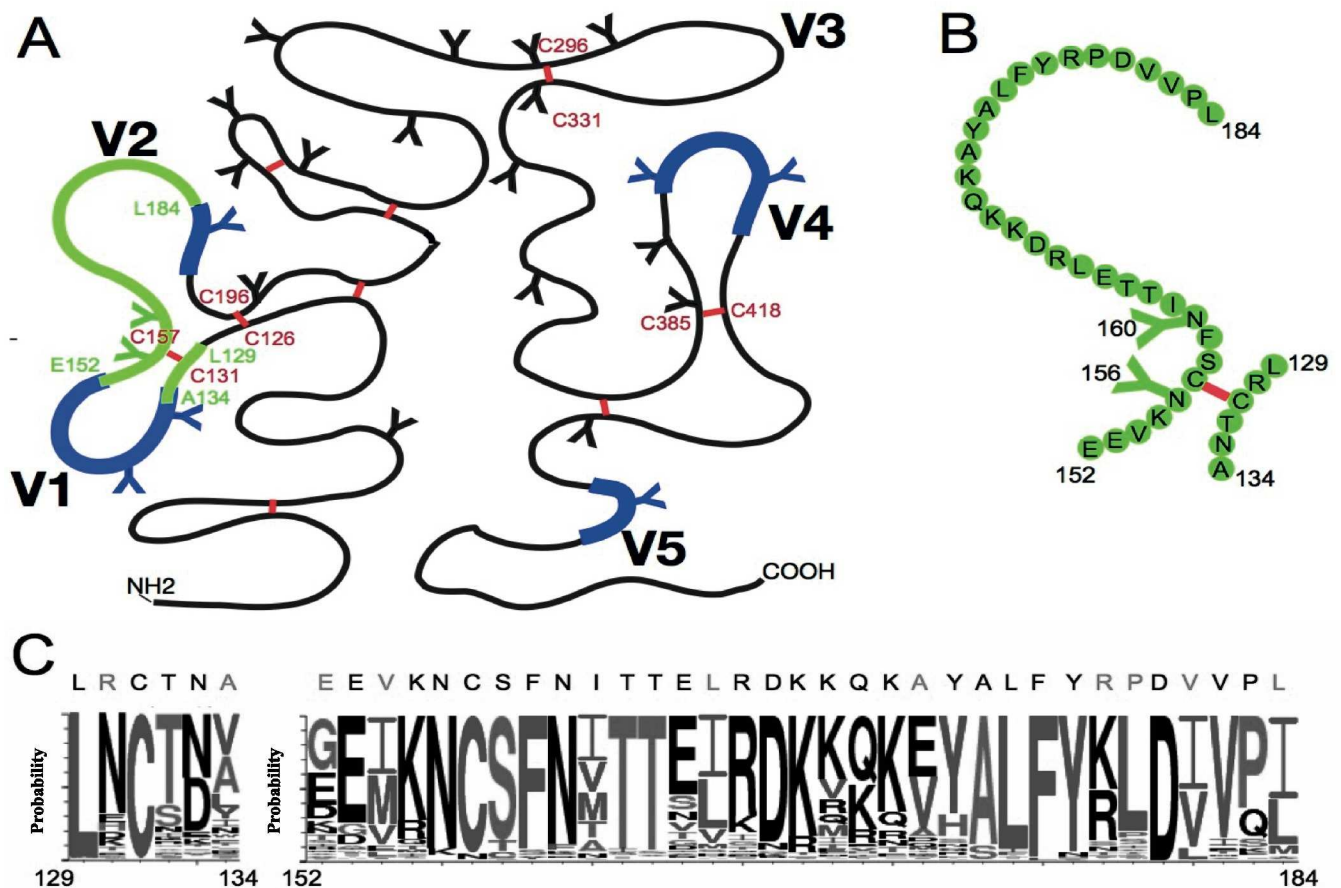


Fig 1. Map of gp120 variable domains and diagram of the V1V2 peptide construct considered in this study. (A) Cartoon diagram showing the variable regions of gp120 (V1–V5), disulfide bonds, and structural region corresponding to the peptide construct in the context of gp120. Blue—hyper-variable regions; Red—disulfide bonds; Green—V1V2 peptide region. This cartoon is modified from the original Leonard reference [49]. (B) The peptide construct (green) contains 6 amino acids from V1 and 33 amino acids from V2 connected by a disulfide bond (red). The sequence corresponds to the CAP45 strain (C) The sequence variability of the regions encompassing the peptide construct among HIV-1 isolates shown as a web logo (http://www.hiv.lanl.gov/content/sequence/ANALYZEALIGN/analyze_align.html). All residues are numbered according to the HXB2 reference sequence.

doi:10.1371/journal.pcbi.1005094.g001

region is found in several conformations such as beta-strand, helical and random coil when bound to monoclonal antibodies (mAbs) elicited by RV144 vaccination that have limited neutralization breadth [45, 47]. These findings suggest that certain conformations of V1V2 could preferentially elicit antibodies with neutralization breadth, and several groups are actively working towards this goal by developing V1V2-derived glycopeptide immunogens [40, 48]. However, the conformational variability of this region, further complicated by glycosylation, could present substantial obstacles to this strategy.

To better understand how glycosylation influences this immunogenic but disordered region of V1V2, we utilized an unbiased all atom MD simulations approach. A peptide construct (Fig 1B) was generated to mimic this immunologically important region of V1V2; 6 amino acids from V1 (129–134 HXB2 numbering) and 33 amino acids (152–184 HXB2 numbering) from V1 and V2 (including the region that was targeted by vaccinees in the RV144 trial and by other V2-directed antibodies from chronic HIV-1 infection), connected by a disulfide bond between the two cysteine residues Cys 131 and Cys 157 that close the V1 loop (Figs 1B and 2A). The folding of this peptide, with and without the carbohydrates at positions N156 and N160, was studied. We show that glycosylation has numerous effects on the stability and accessibility of this disulfide linked V1V2 peptide that could influence the antigenicity and immunogenicity of this region.

Methods

The all-atom replica exchange molecular dynamics (REMD) simulations were used to analyze conformational aspects of a peptide construct derived from the V1V2 variable loop region of HIV-1 gp120. The peptide construct contains 6 amino acids from V1 (HXB2 gp120 residues 129–134) and 33 amino acids from V2 (HXB2 gp120 residues 157–184) connected by a disulfide bond between Cys131 in V1 loop and Cys157 in V2 loop as shown in Fig 2A. The sequence of the peptide corresponds to that of the CAP45 strain, (Genbank accession number GQ999974) it was selected because a PG9 bound structure using CAP45 is available [30]. Our model included two glycans attached to the peptide at positions N156 and N160, which, as noted, have been shown to be important for recognition by several broadly neutralizing antibodies. Also, N156 is immediately adjacent to the disulfide bridge formed by cysteines 131 and 157. This specific sequence did not contain a glycan at position 130 that is found in many HIV-1 strains (Fig 1C). Several studies using virion-associated envelope or SOSIP trimer proteins have shown that the carbohydrate moieties of HIV-1 virions are almost entirely oligomannose, consisting mainly of varying proportions of Man₅₋₉GlcNAc₂. However, Man₅GlcNAc₂ is consistently present at key glycan positions in gp120, including N156 and N160 [50–53]. Given the consistent presence of Man₅GlcNAc₂, in relevant forms of HIV-1 envelope, and that this biosynthetic intermediate reflects the inherently restricted glycan processing of HIV-1 envelope, the carbohydrate moieties used in our model are high-mannose Man₅GlcNAc₂ (Fig 2B). Furthermore, molecular dynamics simulations of the BG505 SOSIP trimer with Man₅GlcNAc₂, Man₇GlcNAc₂, or Man₉GlcNAc₂ at every sequon produced concordant results in terms of overlap with volume occupied by broadly neutralizing antibodies [53]. Separate REMD simulations were carried out for the unglycosylated and glycosylated peptide construct to quantitate the effects of glycosylation. Additionally, we carried out REMD simulations on a single peptide fragment (NCSFNVTITVRDKTTK) derived from the HIV-1 subtype C consensus sequence (ConC; Los Alamos Database Consensus Alignments, 2013) with or without glycosylation at N160, and in the absence of a disulfide bond. These simulations were used to independently verify the glycosylation effects seen in the primary simulation studies on the peptide construct.

All systems were composed of peptide/glycopeptide, water, and counter-ions. A total of 84 replicas were performed for the primary simulations, with a temperature range covering 275 K

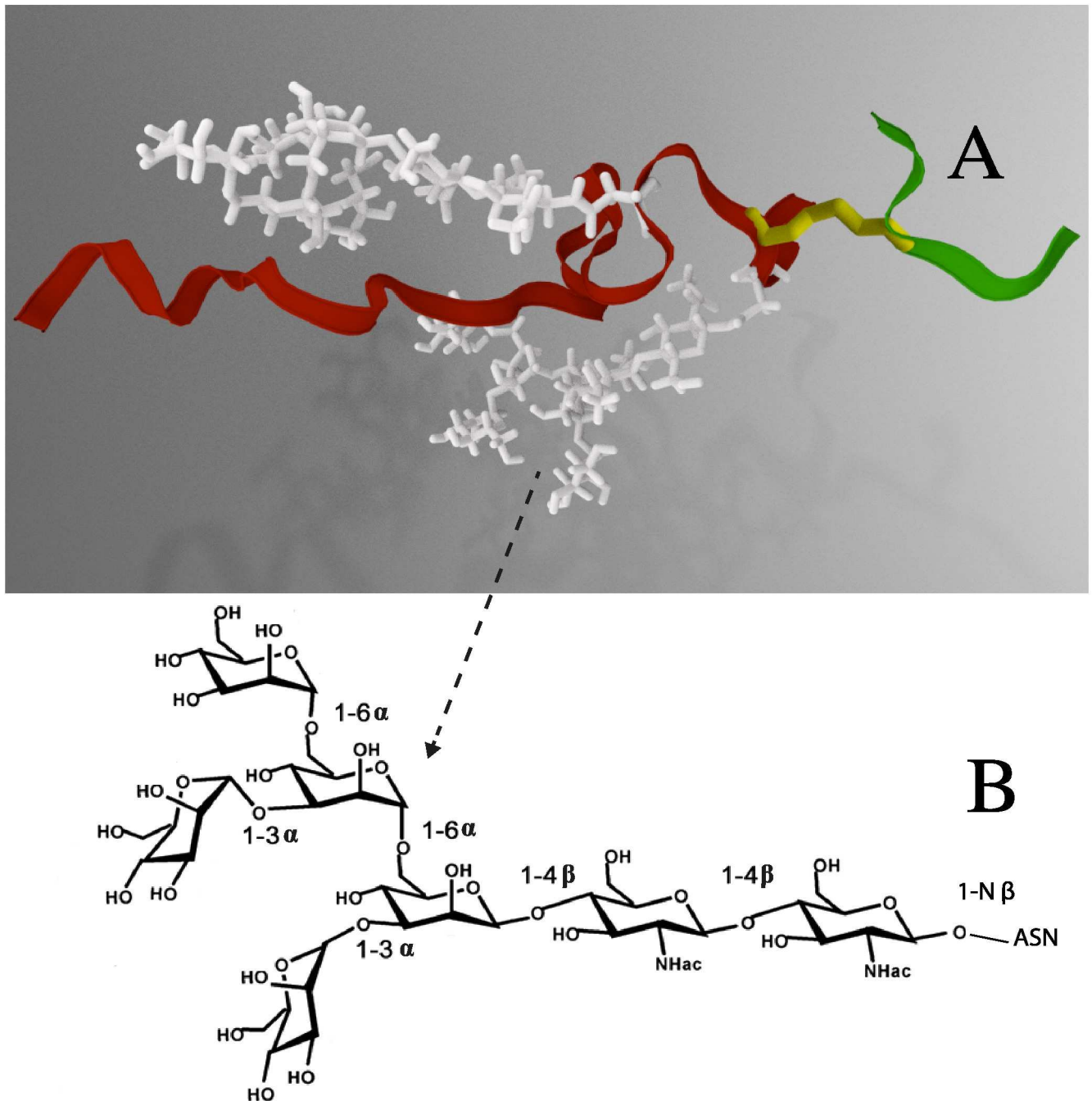


Fig 2. A three-dimensional image of the glycosylated peptide construct from the CAP45 strain. (A) A ribbon diagram shows the gp120 V1 fragment in green, and the V2 fragment in red. A disulfide bond is shown in yellow, and the carbohydrates corresponding to sites N156 and N160 are shown as sticks in white **(B)** The chemical composition of high-mannose glycan, $(\text{Man})_5(\text{GlcNAc})_2$, is shown.

doi:10.1371/journal.pcbi.1005094.g002

to 550 K. Each replica was run for 500 ns. A total of 76 replicas were included in the second set of simulations, with the temperature range covering 285 K to 558 K. Each replica was run for 100 ns. The details of the four REMD simulations are shown in [Table 1](#). The second halves of the REMD trajectories were used for analysis.

Table 1. Systems studied by MD simulations.

System	Strain	No. of Glycans	Disulfide	Waters	No. of Atoms	Method	No. of Replicas	Time, ns (replica)
V2	CAP45	0	Yes	8459	26057	REMD ^a	84	500
V2g	CAP45	2	Yes	8240	25718	REMD	84	500
V2_c	ConC	0	No	5842	17790	REMD	76	100
V2g_c	ConC	1	No	5821	17844	REMD	76	100
V2_h	CAP45	0	Yes	6944	21413	MD ^b	20	60
V2g_h	CAP45	2	Yes	7383	23051	MD	20	100

a. REMD, replica exchange molecular dynamics simulations

b. MD, conventional molecular dynamics simulations

doi:10.1371/journal.pcbi.1005094.t001

The AMBERff99SB[54] force field parameters were used for the peptide, and GLYCAM06[55] force fields were used for the carbohydrate moiety, as they have been shown to be compatible with each other for glycoprotein studies[55]. Constant temperature and constant volume (NVT) REMD simulations were conducted to sample the conformational space [56], REMD is an enhanced sampling technique based on the parallel tempering Monte Carlo method [56–59], where copies of identical systems are simulated at different temperatures. Periodically state exchange between replicas is attempted, and the acceptance rule for each move between states *i* and *j*, dictated by a Boltzmann distribution, is $P_{acc} = \min\{1, \exp[(\beta_i - \beta_j)(U(\vec{r}_i^N) - U(\vec{r}_j^N))]\}$, where $\beta = 1/k_B T$ and $U(\vec{r}_i^N)$ represents the configurational energy of the system in state *i*. Together with the exchange, the particle momenta are scaled by $(T_i/T_j)^{1/2}$, such that the kinetic energy terms in the Boltzmann factor cancel out[56]. REMD sampling can also be described in terms of umbrella sampling[60]. The temperature spacing between replicas was chosen to ensure sufficient energy distribution overlap between neighboring replicas such that exchange attempts were, on average, accepted with a 20% probability[61]. The potential energy distributions of the system were simulated at constant volume and constant temperature for 30 different temperatures to establish the distribution for the replicas[62].

In addition, thermal unfolding studies of the peptide construct were carried out to characterize the effects of glycosylation on unfolding. Conventional MD simulations at high temperature were conducted on the fragment of the V1/V2 variable loop starting from the beta-strand structure (PDB ID: 3U4E). Two sets of simulations were done, with each set having 20 replicas: one set was for peptide with two glycans at N156 and N160, and the other set was for peptide without glycans. All 20 replicas in each set started from the same conformation but ran with different random seeds at 450 K. Simulations of 60 ns were run for each one without glycan to see the peptide completely unfold, while simulations of 100 ns were run for each one with glycan, and significant amount of beta-strand remained at the end of simulations. The details of the systems are listed in [Table 1](#).

The Nose Hoover thermostat was used for the temperature coupling with a coupling time constant $\tau_T = 1.0$ ps. The protein/glycoprotein and solvent are coupled separately to thermostats with the same coupling parameters. Van der Waals interactions are treated using a 1.0 nm cutoff. The electrostatic interactions are treated by smooth particle mesh Ewald summation. All bond interactions involving hydrogen atoms are constrained using SETTLE and SHAKE to allow a 2 fs integration time step. A total of 102.4 μ s simulations have been carried out in the current study.

Configurational entropy calculations were performed following the formulation of Schlitter [63]. This approach provides an approximate value (upper bound) *S* to the true configurational

entropy S_{true} of the simulated system,

$$S_{\text{true}} < S = \frac{k_B}{2} \ln \det \left[1 + \frac{k_B T e^2}{\hbar^2} \underline{\mathbf{D}} \right]$$

where k_B is the Boltzmann's constant, T is the absolute temperature, e is Euler's number, and \hbar is Planck's constant divided by 2π . Here $\underline{\mathbf{D}}$ is the covariance matrix of mass-weighted atomic Cartesian coordinates, defined as

$$\underline{\mathbf{D}} = \langle [\underline{\mathbf{M}}^{1/2}(\mathbf{r} - \langle \mathbf{r} \rangle)] \otimes [\underline{\mathbf{M}}^{1/2}(\mathbf{r} - \langle \mathbf{r} \rangle)] \rangle$$

where \mathbf{r} is the 3N-dimensional Cartesian coordinate vector of N particles (atoms or beads) considered for the entropy calculation after least-squares fitting onto a referenced structure, $\underline{\mathbf{M}}$ is the 3N-dimensional diagonal matrix containing the masses of these particles, $\langle \dots \rangle$ denotes ensemble averaging, and the notation $\mathbf{a} \otimes \mathbf{b}$ stands for the matrix with elements μ, ν equal to $\mathbf{a}_\mu^* \mathbf{b}_\nu$. In our case, the structures after minimization were used as reference for the least square fitting of 1200 snapshots of the trajectory. Moreover, during the fitting procedure, the rotational and translational contribution was removed, considering only the internal degrees of freedom of the peptide backbone.

Finally, we carried out large-scale all-atom MD simulations of the glycosylated and the unglycosylated Env spike containing gp120 trimers [64]. Initial coordinates of the glycoprotein complex were downloaded from the protein data bank repository (PDB code 4NCO) [65]. Missing loops and residues in the structure were built using the Modeller package [66] using the full sequence of the BG505 SOSIP gp140 Env trimer in complex with broadly neutralizing antibody PGT122 [65]. The X-ray structure is partially glycosylated with mannose sugar derivatives and all glycosylation positions were completed to have the same Man₅ carbohydrate moieties. The system is represented using the same force field as used in the simulations of the V1V2 peptide fragment (described above). Env trimeric spike was placed in a cubic box of ~4000 nm³ and was solvated with 100000 water molecules.

Two independent MD simulations were carried out for the Env spike with and without glycosylation. Each simulation was run for one microsecond and carried on using the GROMACS 4.6.5 [67] molecular simulation package. All atom simulations were performed using a 2 fs time step to integrate Newton's equations of motion. The LINCS algorithm [68] was applied to constrain all bond lengths with a relative geometric tolerance of 10⁻⁴. Non-bonded interactions were handled using a twin-range cutoff scheme. Within a short-range cutoff of 0.9 nm, the interactions were evaluated every time step based on a pair list recalculated every five-time steps. The intermediate-range interactions up to a long-range cutoff radius of 1.4 nm were evaluated simultaneously with each pair list update and were assumed constant in between. A PME approach [69] was used to account for electrostatic interactions with a grid spacing set to 0.15 nm. Constant temperature (300 K) was maintained by weak coupling of the solvent and solute separately to a Berendsen heat bath [70] with a relaxation time of 1.0 ps. Similarly, an isotropic approach was used to couple the pressure of the system to 1.0 bar. Trajectories were stored every 20 ps for further analysis.

Results

Glycosylation patterns in the gp120 hyper-variable domains

HIV-1 gp120 contains between 18 to 33 N-linked glycosylation motifs, and maintains a median of about 25 despite very high levels of genetic variability [71]. Some of the glycan sites are highly conserved across all of the major HIV-1 clades and circulating recombinant forms, while a

subset show clade-specific patterns [71, 72]. However, considerable diversity in gp120 amino acid sequence and glycosylation is evident even in single individuals over time [71], and this variation can mediate antibody immune escape [34–39, 73, 74], including cases where the carbohydrate is a direct component of the epitope [30]. The range in number of glycan sites is largely a consequence of mutations, insertions and deletions that occur within the four hyper-variable domains of gp120 (V1, V2, V4 and V5) (see Fig 1A), with insertions often reflecting local imperfect direct repeats of varying lengths [75]. The V3 domain is also variable, however, it is more conserved than the four hyper-variable regions in terms of length and variation of glycosylation sites [71].

To investigate patterns of glycosylation in gp120, we first performed an updated analysis of the relationship between hyper-variable loop length and glycosylation in gp120. We included the most current set of global sequence data ($n = 4,633$) and used a bioinformatics tool that is newly available at the Los Alamos HIV database (<http://www.hiv.lanl.gov/>). Fig 3 illustrates that for the four loops that are hyper-variable in terms of insertions and deletions (V1, V2, V4, and V5), the loop length and number of glycan sites are highly variable and correlated with each other; in contrast these parameters for the V3 region are almost invariant. The V2-epitope region we have used as a basis for the peptide we model here is similar to V3 in that it is almost invariant in terms of length and number of potential glycosylation sites; the hyper-variable regions that evolve by insertion and deletion in the V1 and V2 loop flank the key epitope region represented in the peptide, in the context of the natural protein (In Fig 1A, the hyper-variable regions within the loops are indicated in blue). In contrast, there is broad net charge distribution in all of these variable regions (Fig 4), including the epitope region, which varies from -4 to +7 for V1, 2, 4 and 5, and ranges from -1 to +10 for V3. Whereas V3 has a positive charge, with a median of +4, the median charge is close to neutral for the other variable domains [76, 77].

Cysteine-proximal glycosylation sites are preferentially found at the base of the hyper-variable loops in gp120

The V1, V2, V3, and V4 loops are each delineated by an invariant cysteine-cysteine disulfide bond, and for 6 of these 8 loop-bounding cysteines, N-linked glycosylation sites tend to be immediately adjacent to one or both of the cysteines. There are two possible sequence motifs for these patterns: NCS/T or CNXS/T, where X can be any amino acid except for proline (Fig 1A and S1 Fig). Among the 4633 sequences analyzed, there are 8 conserved Cys residues at the bases of the variable loops, and on average, another 16 Cys residues per Env. Outside of the variable loops, these other Cys residues rarely have a proximal N-linked glycosylation site (Fig 5; 1357 of 73354, or 1.8%). In contrast, the majority of the conserved Cys at the bases of the variable loops are proximal to a glycan (Fig 5; 21,837 out of 37,064, 59%). The enrichment for Cys-proximal N-linked glycosylation at the bases of the hyper-variable loops is highly significant (6/8 vs 0/16, $p = 0.0002$). In particular the two N-linked glycosylation sites (N156, N197) that are proximal to the Cys residues at the base of the V2 loop (corresponding to C157, C196: Fig 5) are both very highly conserved across HIV-1 subtypes (S2 Fig). This suggests that these Cys-proximal glycans are important for Env function, and may impact the structural conformation of the variable loops in the intact protein, and V2 in particular.

V1V2 peptide construct containing both glycosylation and disulfide linkage spans a critical immunodominant region

A beta-strand of 33 amino acids that contains key neutralizing antibody epitopes is embedded in our peptide construct (Fig 1A). The portion that resides within V2 is conserved in terms of

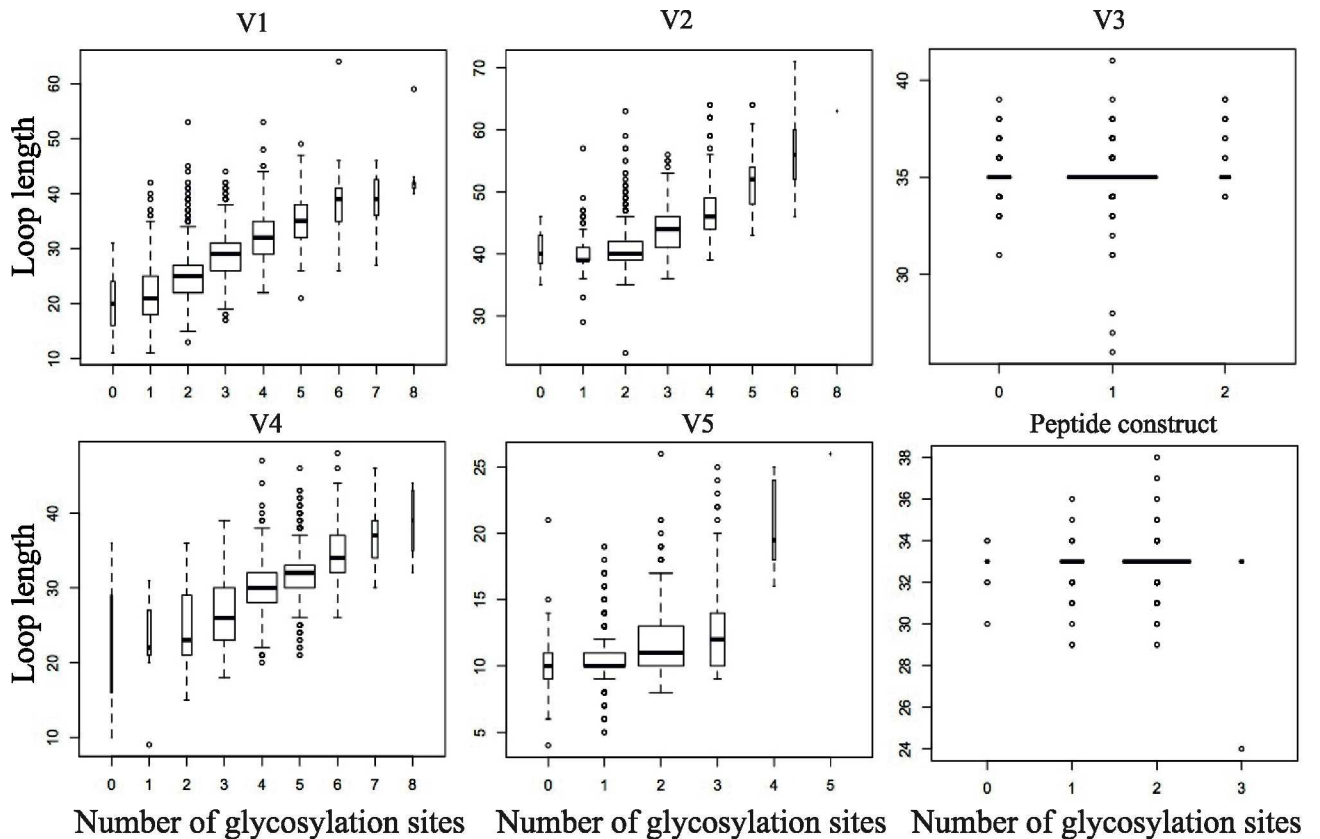


Fig 3. The relationship between loop lengths and number of glycosylation sites in the gp120 variable regions V1-V5. These plots are based on the Los Alamos Database (www.hiv.lanl.gov) alignment, which contains 4,633 curated HIV-1 Env sequences. This alignment contains intact, full length Env sequences, and includes only one sequence per sampled individual. Sequences of poor quality (frameshifts, ambiguity codes or inappropriate stop codons) were excluded from the alignment. The relative width of the box plots is proportional to the square root of the number of sequences in this set that have a given number of potential N-linked glycosylation sites, having the sequence pattern (NX[ST]), where N is an Asparagine, followed by X, any amino acid except Proline, followed by either a Serine or Threonine. Also shown is the epitope region, spanning HXB2 positions 152–184, from the V1V2 peptide construct. In the case of hyper-variable loops, V1, V2, V4, and V5, a p-value $< 2.2e-16$ was estimated for the correlation between length and number of N-linked glycosylation sites using Kendall's *tau* statistic (estimated using the R statistical package; it is non-exact due to ties and large sample sizes). The Variable Length Characteristics tool was used to evaluate these regions, and the full loop regions were included in the analysis (V1, HXB2 positions 131–157; V2 158–196; V3 296–331; V4 385–418; and V5 360–469).

doi:10.1371/journal.pcbi.1005094.g003

length and number of glycosylation sites; 94% of the sequences are 33 amino acids long in this region, and 88% carry 2 glycosylation sites (generally the highly conserved sites at N156 and N160), while 11% have only one (S2 Fig). Despite overall conservation in length and number of glycans (Fig 3 and S3 Fig), this V2 ‘epitope’ region has many highly variable positions (Fig 1C), and it can vary dramatically in net charge (Fig 4 and S3 Fig). Nevertheless, this region is of great interest from a vaccine perspective because it serves as the contact region for glycan-dependent broadly neutralizing antibodies such as PG9 and PG16 that recognize the V2 region with a preference for the quaternary structure [29, 30]. This region is also recognized by other broadly neutralizing antibodies, PGT141-PGT145 [31], CH01-CH04 [32] and the VRC26 group of mAbs [43] as well as by a number of antibodies with narrower neutralization breadth, such as C108g [78], 10/76b [79] and 2909 [80] isolated from infected or immunized animals. The two conserved N-linked glycans that are in this region form parts of these epitopes and directly contact antibodies PG9 and PG16 [29, 30]. In addition, in the RV144 vaccine trial

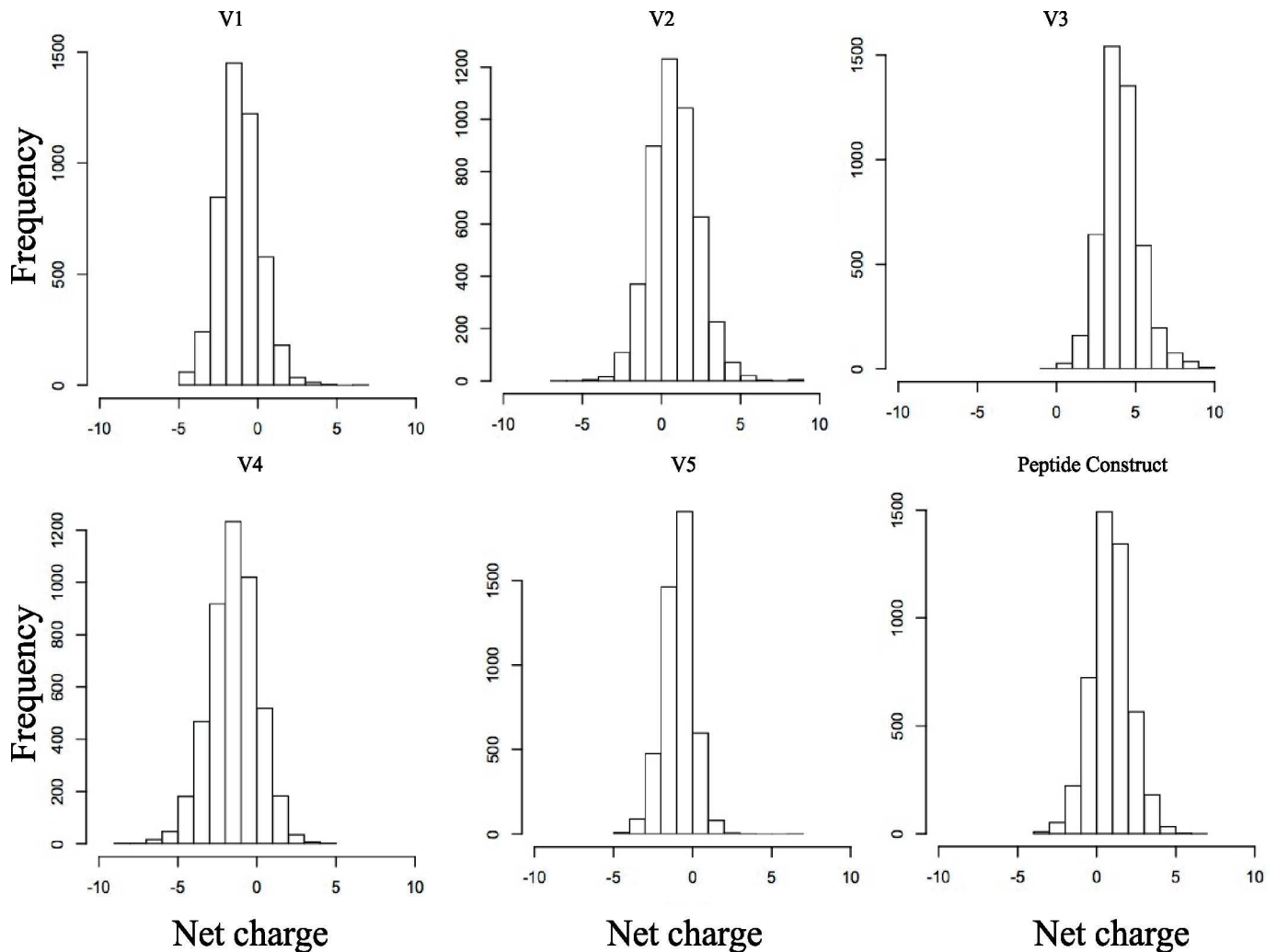


Fig 4. Net charge distribution of the variable regions. Using the same input data as in Fig 3, charge variability in the V-1V5 regions and the V2 epitope region is shown. The V3 loop and V2 epitope region, despite being conserved in terms of length and number of glycosylation sites (Fig 3), both show a great deal of variation in net charge, comparable to the level of diversity found in hyper-variable regions. Net charge is calculated as the sum of positive and negative charges, where amino acid residues E and D are assigned -1, and K, R, and H are assigned +1. The V2 epitope region, V3 and V2 tend to be positively charged; V1, V4, V5 tend to be negatively charged.

doi:10.1371/journal.pcbi.1005094.g004

immune responses to this linear epitope region were correlated with reduced risk of infection, and the protective effect was associated with antibody-dependent cellular cytotoxicity (ADCC), not neutralization [81, 82].

Glycosylation reduces the propensity of the V1V2 peptide to adopt folded secondary structures

An initial set of REMD simulations was carried out to examine how glycosylation at the conserved positions N156 and N160 affects the secondary structure propensities of the unstructured, flexible V1V2 peptide construct. This isolated peptide exists predominantly as a random coil, and the fraction of residues that form a helix, beta-strand, or turn structures as a function of temperature is plotted in Fig 6. While the fraction of residues in the glycosylated and unglycosylated peptides involved in beta-strand or turn structures decreases with increasing

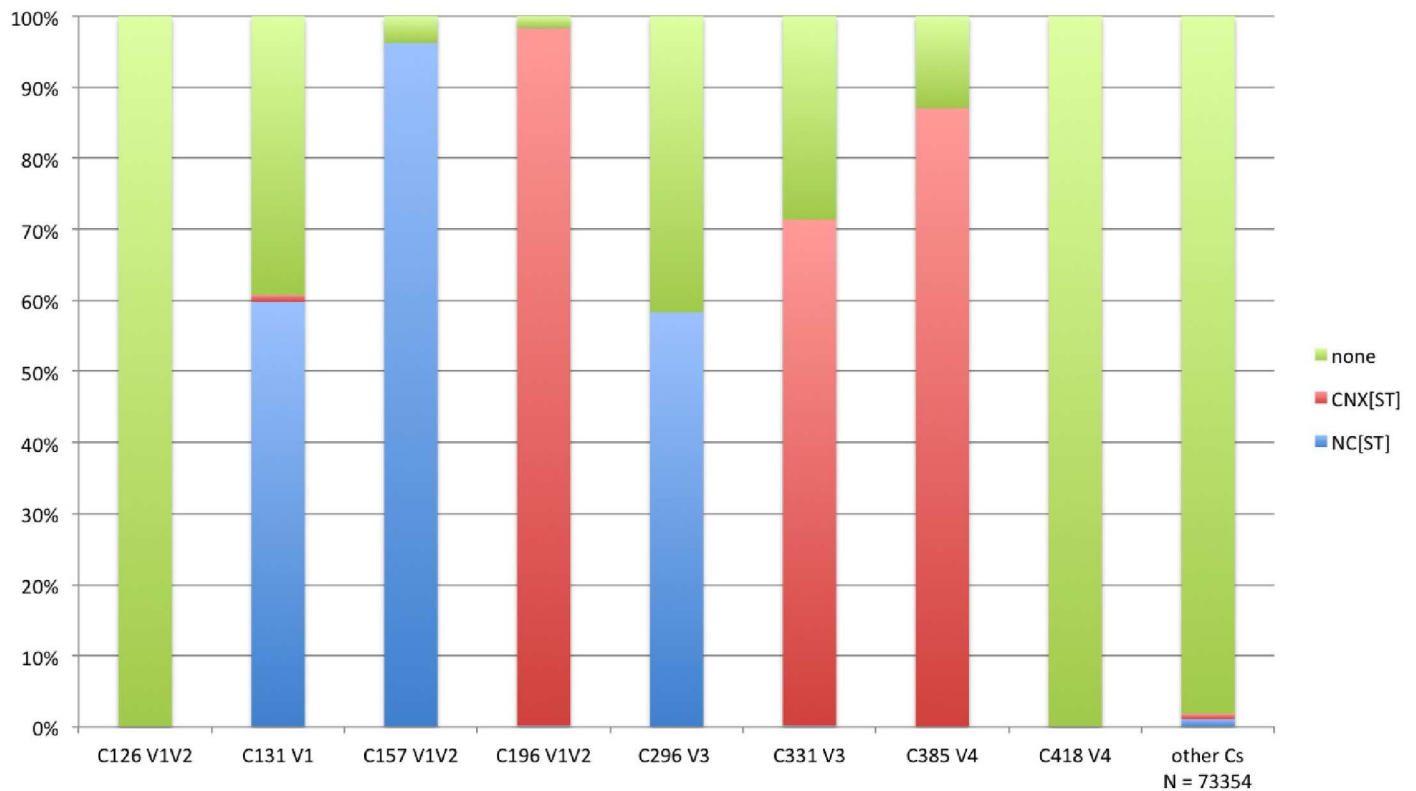


Fig 5. Six out of the eight conserved cysteines located at the base of HIV-1 variable loops V1-V4 are very frequently proximal to a N-linked glycosylation site. Of the 4633 sequences in the filtered alignment in the 2014 database, there are on average 24 Cys per gp160, including 8 that close the variable loops V1-V4 and 16 others. Among the 8 conserved Cys that form disulfide bonds at the base of V1-V4, 21,837 immediately neighbor an N-linked glycosylation site, or 59%. These are concentrated in positions Cys131, Cys157, Cys196, Cys296, Cys331, and Cys385. Among the 73,354 Cys that are not located at the base of the variable loops, proximal glycans are very rare at 1.8%. Of the 6 Cys with conserved proximal glycans in HIV, only the most conserved, Cys157 and Cys196 are also highly conserved in 14 SIVCPZ sequences.

doi:10.1371/journal.pcbi.1005094.g005

temperature, the fraction of residues that form a helix structure increases until around 440 K, followed by a slight decline. Importantly, a larger fraction of the residues are prone to form secondary structures other than random coil (helix, beta-strand, or turn structures) in the unglycosylated form compared to the glycosylated form (Fig 6). Thus, the overall probability of the V1V2 peptide to form folded secondary structures is reduced upon glycosylation.

Glycosylation modifies the free energy landscape of the V1V2 peptide construct

The free energy landscape of the V1V2 peptide in terms of end-end distance and radius of gyration is shown in Fig 7. In the absence of glycosylation, 43% of the configurations were enclosed within a single state ($R_g = 1.1$ nm, $D_{N-N} = 1.8$ nm) (Fig 7A). Upon glycosylation however, 20% of the total configurations populated two states: ($R_g = 1.2$ nm, $D_{N-N} = 1.8$ nm) and ($R_g = 1.2$ nm, $D_{N-N} = 2.5$ nm) (Fig 7B). These results suggest that the overall ensemble of the V1V2 peptide is more extended upon glycosylation. This is consistent with the reduction of secondary structural preference as seen in Fig 6. Such an effect can arise with increased entropy of a peptide backbone, disruption of intra- and inter-peptide interaction, and between peptide and glycan, each of which were subsequently investigated.

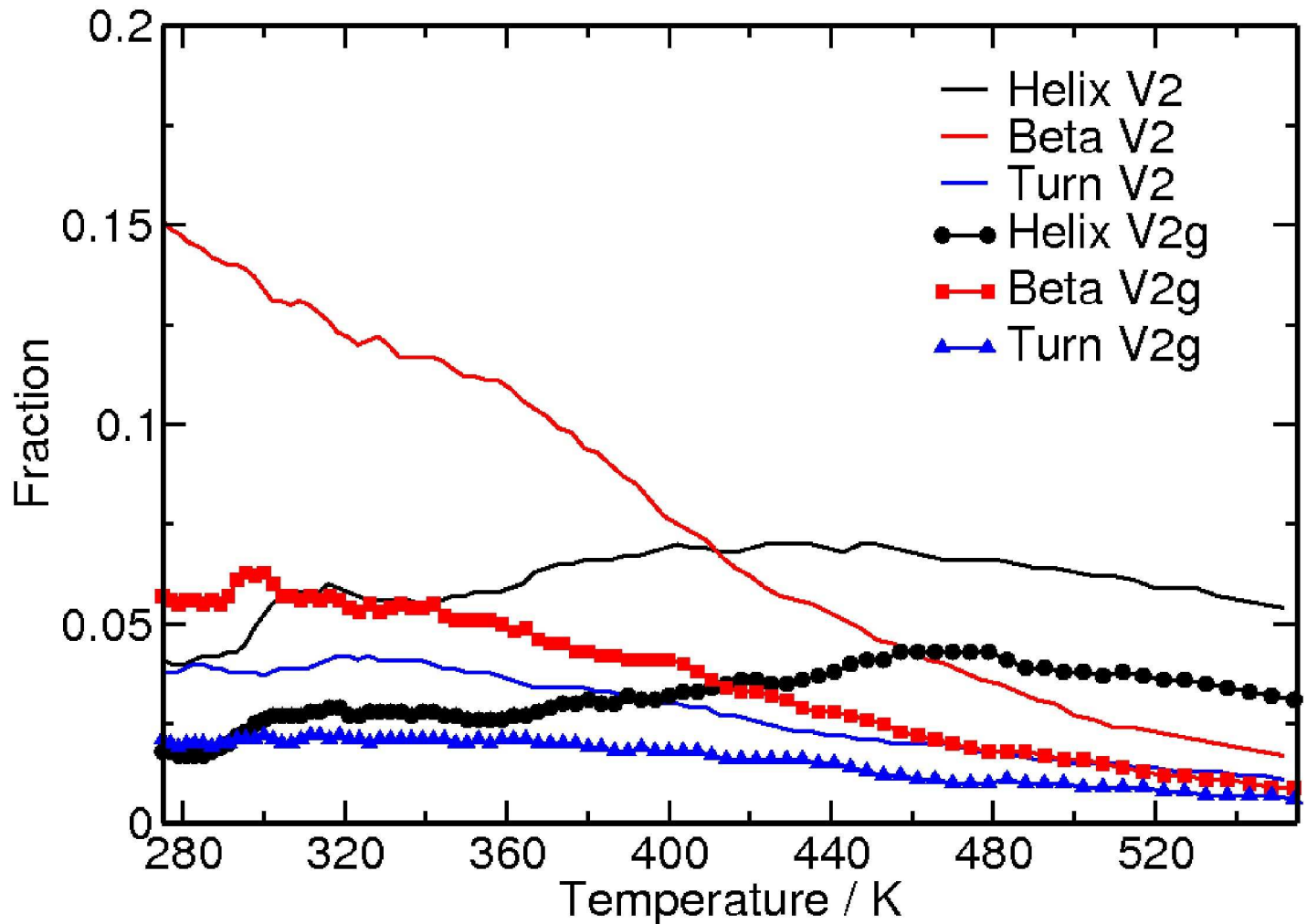


Fig 6. Fraction of residues that form helix, beta-strand or turn structure as a function of increasing temperature. Both glycosylated (symbols) and unglycosylated (no symbols) cases are shown.

doi:10.1371/journal.pcbi.1005094.g006

Glycosylation affects the backbone flexibility of the Asn156 and Asn160 residues

While N-linked glycosylation is often linked with global conformational effects, it is possible that the addition of carbohydrate could affect the N residue itself. A Ramachandran plot was therefore determined for the N156 backbone dihedral angles (Fig 8). Without glycosylation at N156, the most sampled region shows characteristics shared with polyproline II motifs (-65°, 135°). It is followed by the right-handed and left-handed alpha-helix regions (Fig 8A). Glycosylation at N156 reduces the backbone sampling of polyproline II and left-handed alpha-helix regions and enhances the sampling of the extended beta-basin (-135°, 135°) and right-handed alpha-helix (-60°, -30°) regions (Fig 8B). To further quantify the differences in sampling, Shannon entropy was calculated with and without glycosylation at N156. The equation $S = -R \sum P_i \log(P_i)$ was used to calculate Shannon entropy, with R equal to the molar gas constant and P_i equal to the probability to sample each bin. This calculation gives entropic values of $63.01 \text{ kJ K}^{-1} \text{ mol}^{-1}$ and $60.93 \text{ kJ K}^{-1} \text{ mol}^{-1}$ with and without glycosylation, respectively. The slightly higher Shannon entropy for N156 backbone sampling with glycosylation is consistent with varying bin size. Thus,

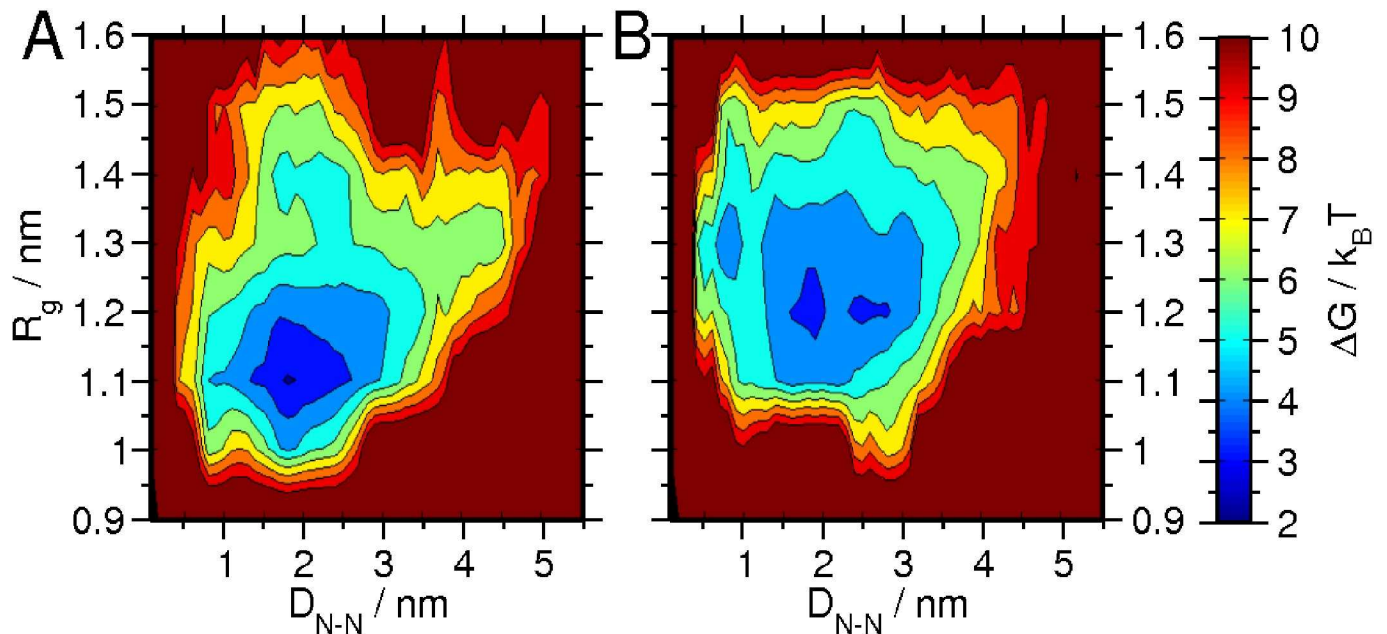


Fig 7. Influence of glycosylation on the free energy landscape of the peptide without (A) and with (B) glycosylation. Landscape is considered as a function of end-end distance and radius of gyration. The contour plots are in units of $k_B T$, and the difference between neighboring lines is $1 k_B T$.

doi:10.1371/journal.pcbi.1005094.g007

by introducing a carbohydrate moiety, increased dimensionality is added to the peptide construct, resulting in higher backbone sampling and entropy.

Glycosylation increases configurational entropy of the V1V2 peptide

Glycosylation of the V1V2 peptide at N156 and N160 could also alter the configurational entropy of the entire peptide. To investigate this, calculations were performed following the

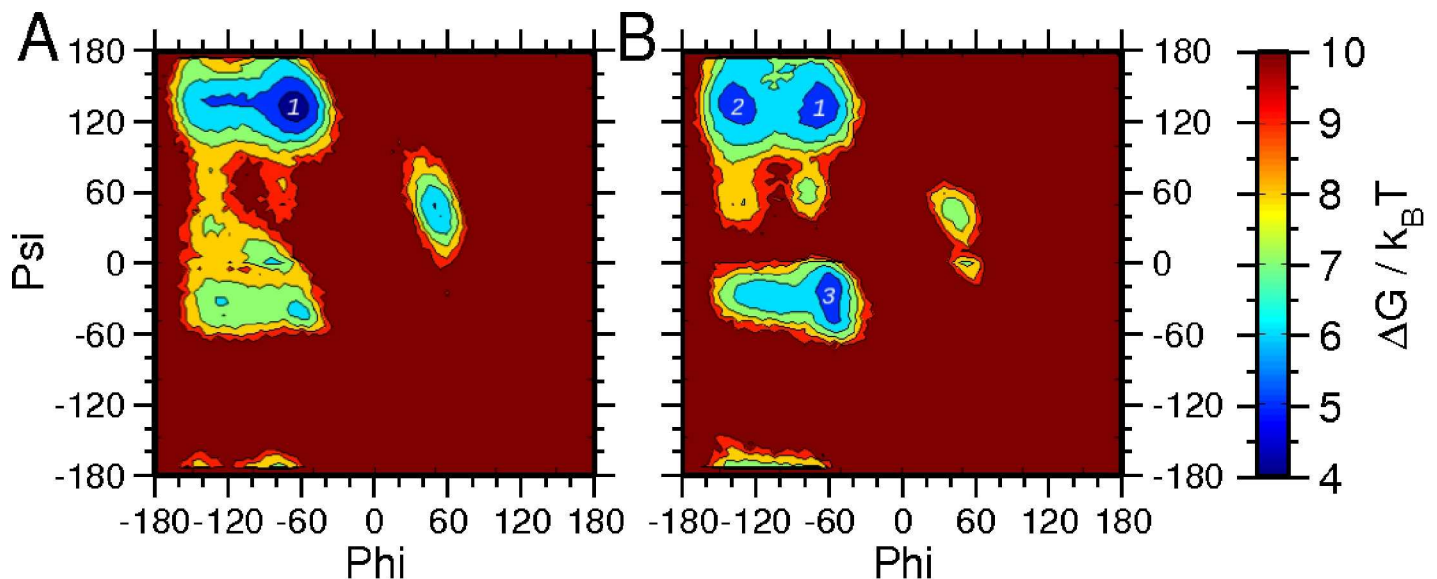


Fig 8. The backbone Ramachandran plot of the residue N156 for the two peptide systems without (A) and with (B) glycosylation is shown. The polyproline II (1), extended beta (2) and right-handed alpha-helical (3) regions are marked. The contour plots are in units of $k_B T$, and the difference between neighboring lines is $1 k_B T$.

doi:10.1371/journal.pcbi.1005094.g008

formulation of Schlitter[63]. **S4 Fig** shows the configurational entropy of the peptide backbone at 300K and 450K, averaged over a 250 ns simulation. Clearly, glycosylation increases the configurational entropy of the peptide, and is greatest at higher temperatures, consistent with the flexibility in secondary structure of the peptide. Interestingly, the inclusion of the glycans increases the entropy of the peptide backbone at higher and lower temperatures by ~ 100 and $\sim 60 \text{ J mol}^{-1} \text{ K}^{-1}$, respectively. However, such an increase in entropy of the peptide backbone is somehow compensated by the enthalpic contribution of the glycan, as discussed below.

Glycosylation disrupts intra- and inter-peptide hydrogen bonding interactions

The effect of glycosylation on intermolecular interactions can be quantified in terms of hydrogen bonding within the peptide, between peptide and solvent, and between glycan and peptide. The total number of hydrogen bonds in these different types of interactions in terms of configuration is shown in **Fig 9**. Regardless of glycosylation, as expected, peptide-solvent hydrogen bonding dominates (**Fig 9A**). Overall, there are much more hydrogen bond interactions between the peptide and water molecules compared to intra-peptide hydrogen bond interactions. However, glycosylation does disrupt hydrogen bonding between peptide and solvent, and between peptides.

Glycosylation increases the stability of peptide through enthalpic contributions

As shown above, glycosylation of the V1V2 peptide disrupts the intra- and inter-molecular hydrogen bonding of the peptide. However, this effect could potentially be compensated by the hydrogen bond interactions that arise due to introduction of glycan. Additionally, the electrostatic nature of the glycan can lead to specific interactions with charged residues of the peptide. Therefore, the different coulombic contributions to the overall energetics were considered to capture any compensating electrostatic interactions introduced by glycosylation. Accordingly, even though several hydrogen bonding interactions involving charged residues are reduced by addition of the glycans, the reduction is compensated by de novo hydrogen bonding between the glycan and other polar residues (**Fig 9B**). **Fig 10A** shows the averaged coulombic contribution as a function of time for the interactions between the peptide, glycans, and solvent. The addition of the glycan clearly affects the intra-molecular interactions of the peptide, as well as the peptide interactions with water molecules. However, the presence of the glycans adds $\sim 3000 \text{ kJ mol}^{-1}$ to the stability of the peptide (**Fig 10**, sum of the energies from panels **C** and **D**). On average, most of the contribution comes from the interaction between the two glycans and the interaction of glycans with the solvent. This observation was also recorded at a higher temperature. Overall, these results suggest that the glycan itself contributes enthalpically to the stability of the peptide in solution.

Glycan-glycan interactions are preferred over glycan-peptide interactions

The interaction between the two glycans at N156 and N160 was characterized in terms of their contact distance. The inter-glycan distance is calculated as the distance between the two C1 atoms of the first mannose in each carbohydrate moiety. The fraction of configurations as a function of inter-glycan distance is plotted in **Fig 11**. Two snapshots of the glycopeptide are also shown, one with an inter-glycan distance of 0.4 nm (left panel) and the other with a distance of 3.1 nm (right panel). A short inter-glycan distance corresponds to the glycans interacting with each other. The inter-glycan distance distribution is highest between 0.75 nm and 1.0 nm, clearly demonstrating that the smaller inter-glycan distances are preferred. It is likely that

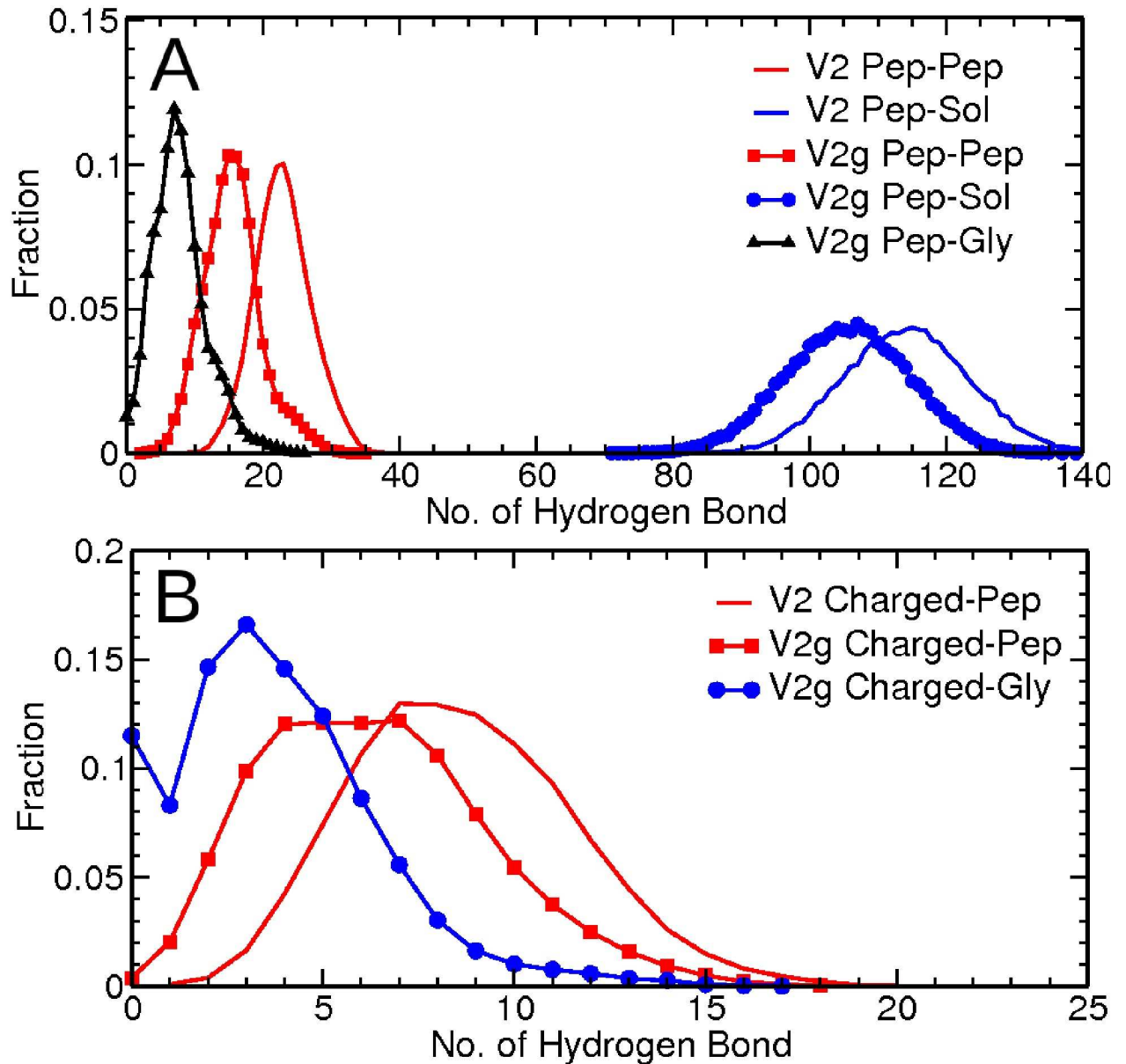


Fig 9. Fraction of configurations as a function of number of hydrogen bonds between different components of the system. (A) Total hydrogen bonding within the peptide, between peptide and solvent, and between glycan and peptide. (B) Hydrogen bonding between the glycans and charged residues. Pep stands for peptide, Sol for solvent water, Gly for glycan, and Charged for the charged residues of the peptide.

doi:10.1371/journal.pcbi.1005094.g009

antibodies that target that region can interrupt such glycan-glycan interactions. In the context of the flexible unstructured peptide construct considered in this study, spatially proximal glycans prefer to interact with each other.

Glycosylation stabilizes the pre-existing conformation of the V1V2 peptide

The V1V2 peptide construct utilized here is predominantly disordered when in solution. However, as mentioned above, it can adopt beta-strand or alpha-helical structures when bound to

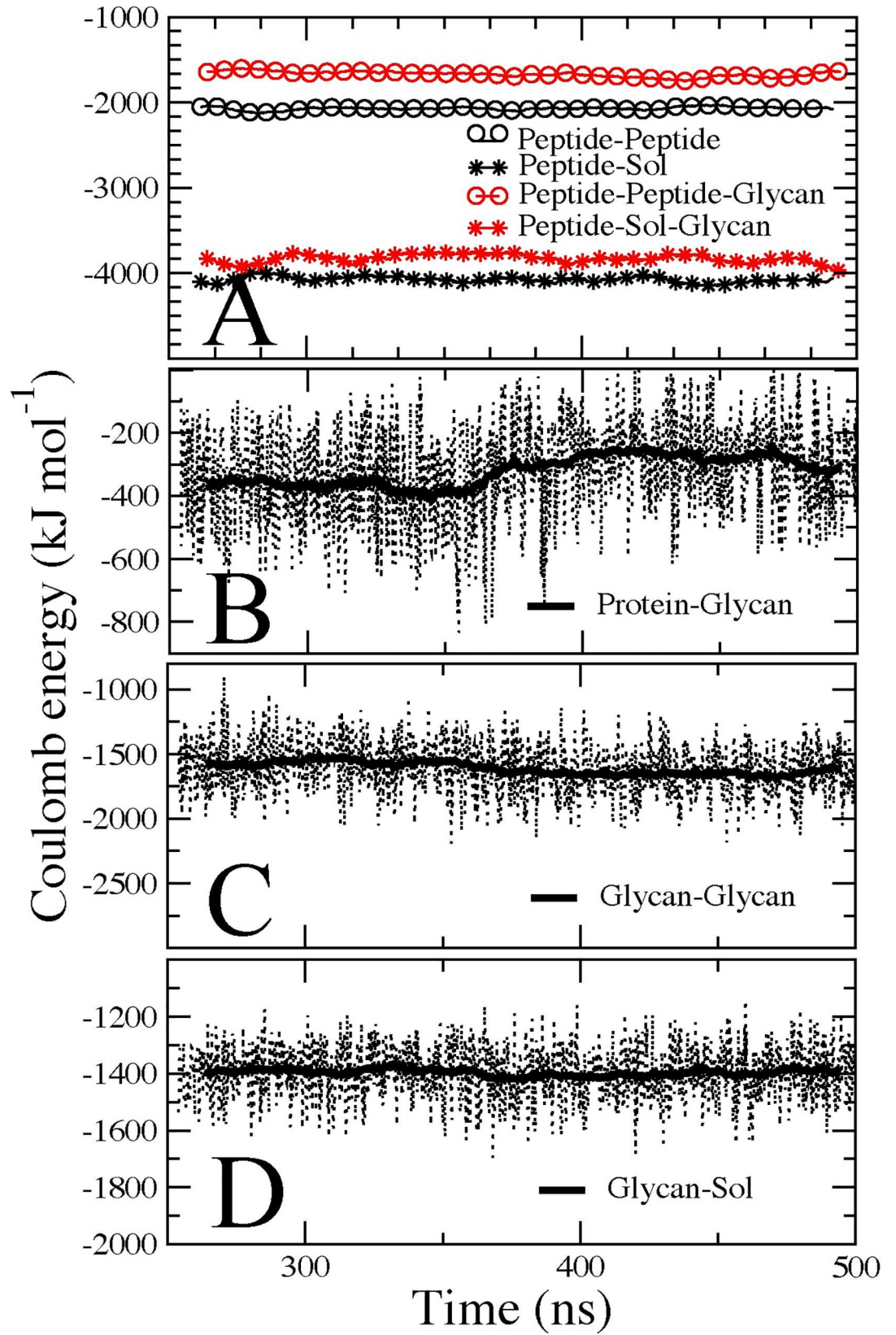


Fig 10. Electrostatic contribution from glycan to the stability of the peptide at 300K. Different contributions are depicted and averaged along the simulation time. (A) intra and inter-molecular interaction of the peptide, with or without the glycan. (B,C,D) Coulomb contributions after the addition of the glycan to the peptide.

doi:10.1371/journal.pcbi.1005094.g010

antibodies. To understand whether glycosylation can stabilize pre-formed beta-strand conformation, unfolding simulations were carried out at 450 K for the peptide construct starting from an initial beta-strand configuration similar to that in complex with the broadly neutralizing antibody PG9. Both the unglycosylated and glycosylated peptide systems were considered

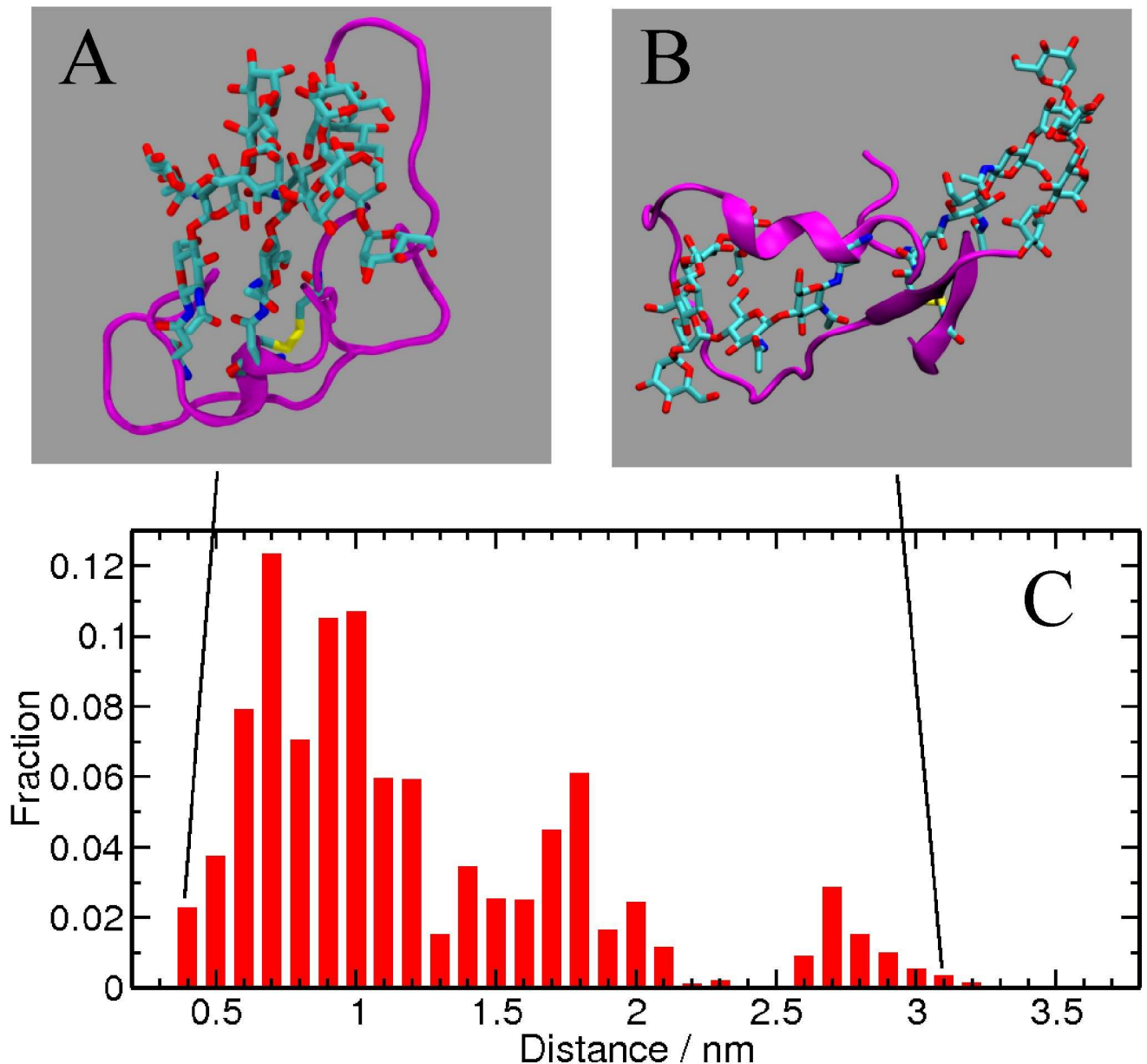


Fig 11. Fraction of configurations as a function of inter-glycan distance. Two representative configurations of shorter and longer inter-glycan distances are shown in (A) and (B). The peptide is shown as a purple ribbon, the glycans are in cyan and red stick representation, and the disulfide bond is shown in yellow. C) Cumulative fraction of the configurations as depicted in A and B. Clustering is based on glycan-glycan distance.

doi:10.1371/journal.pcbi.1005094.g011

(see [Methods](#) section). The average number of residues in beta-strand as a function of time for the two systems is plotted in [Fig 12](#). The beta-strand structures unfolded rapidly, disappearing within 60 ns of simulation for the unglycosylated peptide ([Fig 12A](#)). A fitted exponential curve provided a decay time constant of 14.56 ns. Only two residues remained in the beta-strand structure at the end of the simulation. In contrast, beta-strand structures unfolded at a much slower rate in the glycosylated system, and a significant amount of secondary structures remained at the end of the 100 ns simulation for the glycosylated peptide ([Fig 12B](#)). The decay time constant was 27.78 ns for the glycosylated peptide, and there were six residues remaining in the beta-strand at the end of the simulation. Thus, glycosylation of the V1V2 peptide retards the decay of preformed secondary structure by almost two-fold and preserves more residues with secondary structure. This potentially demonstrates the ability of glycosylation to stabilize the desired conformation of this region of V1V2 in terms of antibody recognition.

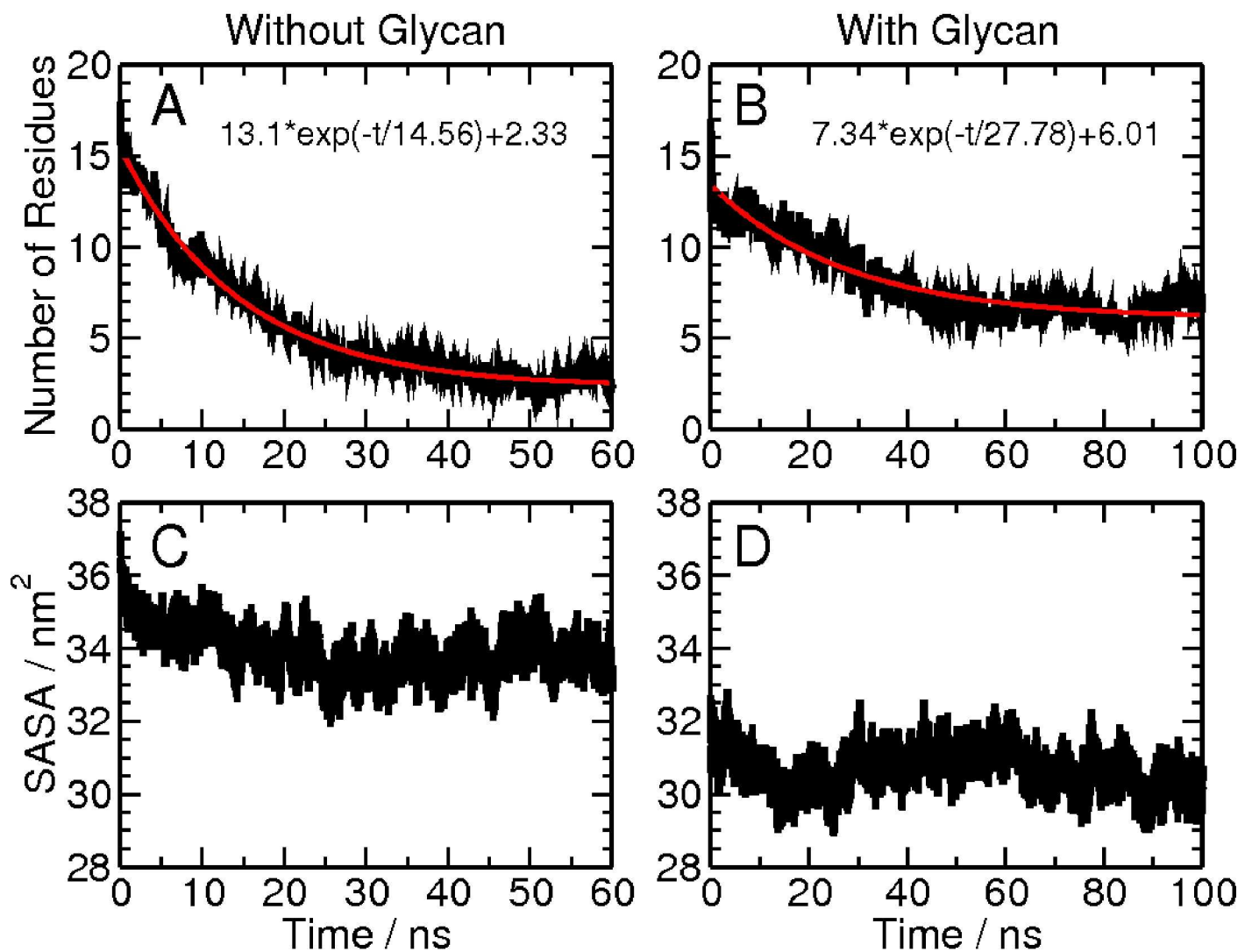


Fig 12. Thermal unfolding simulations. Panels (A) and (B) show the average number of residues in beta-strand as a function of time for the peptide alone and the glycosylated peptide; panels (C) and (D) show the average solvent accessible surface area (SASA) for the two systems as a function of time.

doi:10.1371/journal.pcbi.1005094.g012

Desolvation effects introduced by glycosylation

To explore whether some aspects of stabilization due to glycosylation could be attributed to shielding of solvent, the average surface accessible surface area (SASA) for the unglycosylated and glycosylated V1V2 peptide systems was determined, and is shown as a function of time in Fig 12C and 12D, respectively. The SASA for the glycosylated peptide is less than that of the unglycosylated peptide, likely due to glycan shielding that may also contribute to the stabilization seen above. Previous studies have shown that desolvation of helix or beta-strand peptides stabilize the conformation by strengthening the intra-peptide hydrogen bond interactions [83, 84]. In addition, interactions between the glycan and the peptide residues can also provide stability as discussed above.

Disulfide bonding promotes the formation of beta-strand structure

Next, we computationally investigated the possibility that disulfide bonds such as those found in V1V2 promote the formation of beta-strand structures in unstructured peptides by bringing two peptide regions close together. To investigate whether the disulfide bridge impacted the stabilizing effects of glycosylation on the V1V2 peptide, we considered a V2 loop fragment derived from a clade C consensus (ConC) sequence that did not contain the V1 fragment and the disulfide bond. The propensity to form secondary structures for the CAP45 V1V2 (with disulfide bond) and the ConC V2 peptide (without disulfide bond) is shown in Fig 13. In ConC V2 peptide, like CAP45 V1V2 peptide, glycosylation reduces the propensity for secondary structure formation. However, the fraction of residues that form a beta-strand structure is higher when there is a disulfide linkage, as seen in the glycosylated and unglycosylated forms of the V1V2 peptide (Fig 13A). Though, the helical content is lower for both peptides, in contrast, the fraction of residues that form a helix is reduced to the same level by glycosylation in the presence or absence of the disulfide bond (Fig 13B). Thus, our calculations tend to suggest that the propensity to form beta-strand structure is noticeably increased in the presence of the disulfide bridge.

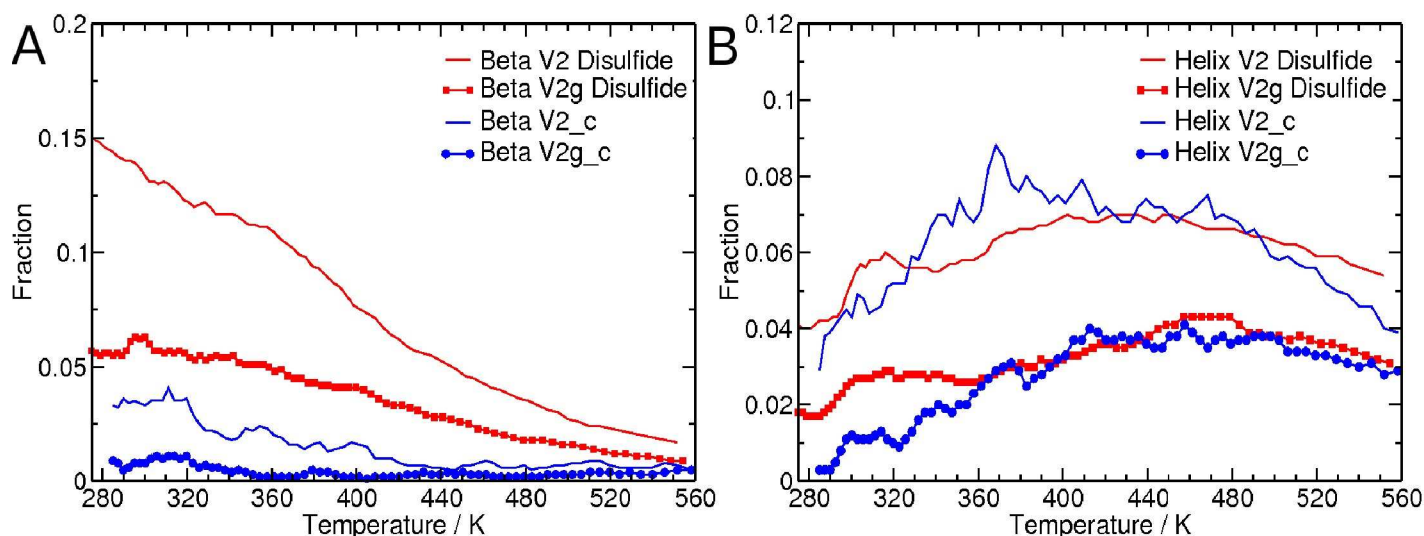


Fig 13. Secondary structure propensity of the glycosylated peptide. (A) Fraction of residues in the peptide that are in the beta-strand structure as a function of temperature for the glycosylated (V2g) and unglycosylated (V2) forms of CAP45 V1V2 peptide and the glycosylated (V2g_c) and unglycosylated (V2_cc) forms of ConC V2 peptide. (B) Fraction of residues in the peptide that are in the helix structure as a function of temperature for the glycosylated (V2g) and unglycosylated (V2) forms of CAP45 V1V2 peptide and the glycosylated (V2g_c) and unglycosylated (V2_cc) forms of ConC V2 peptide.

doi:10.1371/journal.pcbi.1005094.g013

Effect of glycosylation on the V1V2 peptide region in the context of entire Env spike

To this point, we have presented computational results from studies performed on an isolated peptide fragment from the V1V2 region of gp120. An inevitable question is whether the effects of glycosylation described for this fragment are the same in the context of the entire Env trimer spike. To address this, we performed all-atom MD simulations of the Env spike using the BG505 SOSIP gp140 Env trimer in complex with broadly neutralizing antibody PGT122 [65]. A more extensive characterization of the global effect of glycosylation and its contributions to stability of the Env trimer are in progress elsewhere (manuscript in preparation).

We find that many of the physical trends of glycosylation are preserved in the context of Env spike for the same regions of gp120 considered in the V1/V2 peptide construct. From the all-atom simulations of the Env spike, Fig 14A shows the root mean square fluctuations

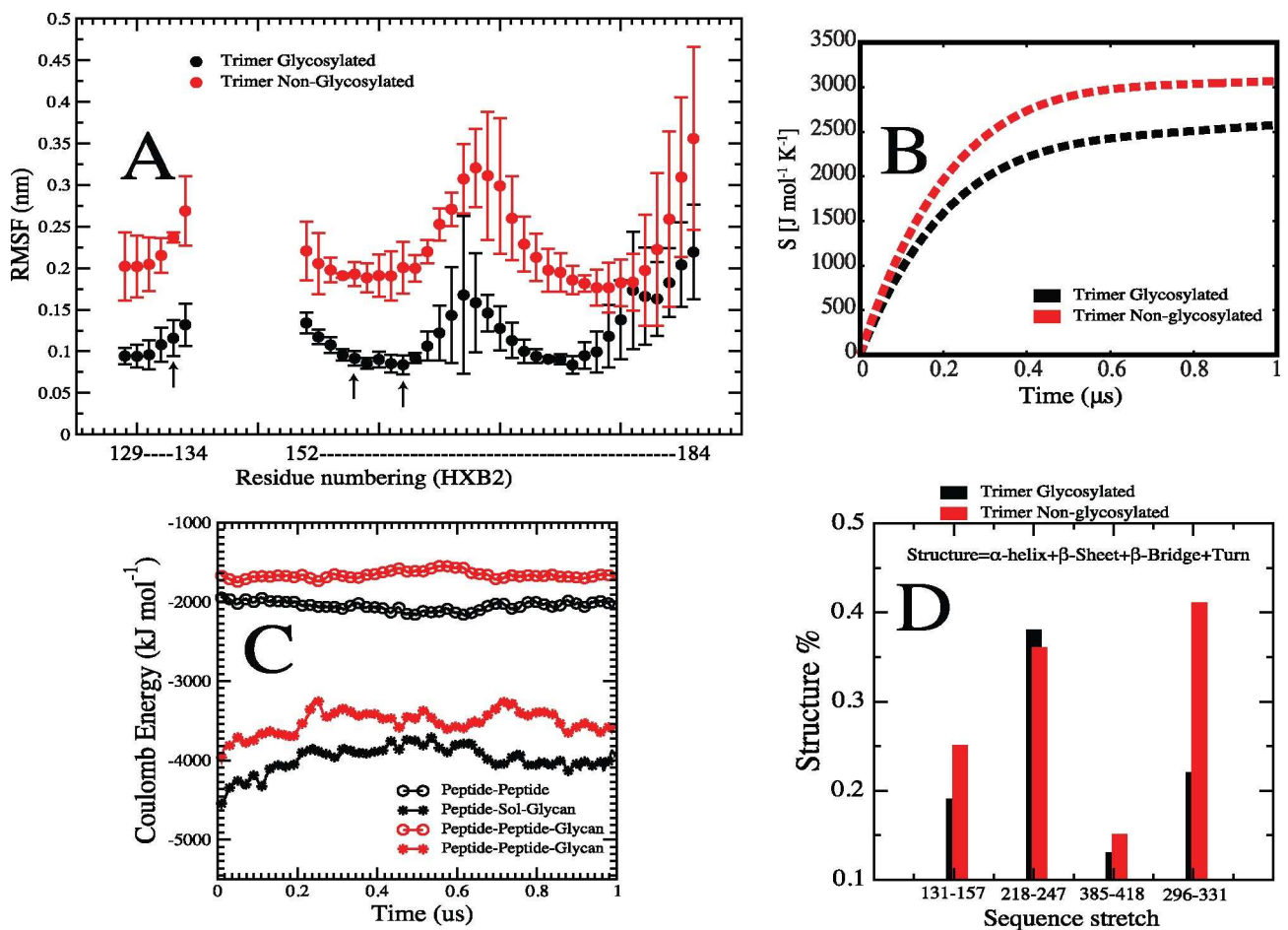


Fig 14. Effects of glycosylation on V1V2 peptide region in the context of the BG505 Env trimeric spike. (A) Root mean square fluctuation of the backbone atoms corresponding to residues 129–134 and 152–184 (HXB2 numbering) and computed for either the glycosylated (black line) and non-glycosylated (red line) protein. Error bars were estimated from calculation in each of the independent protomers. (B) Cumulative configurational entropy for the backbone atoms corresponding to the same residues as in panel A. Values were estimated by considering the total entropy from the three protomers. (C) Total interaction energy from the representative sequence as in B. The energy corresponds to the total value calculated among the three protomers and during 1 μs trajectory simulation. (D) Secondary structural percentage as computed from 1 μs MD simulations of the full Env spike. Four stretches were considered for the analysis, each featuring disulfide bonds and glycosylation sites. Computed secondary structure percentage for amino acid stretches that contain glycans adjacent to Cysteins (HXB2 numbering): 131–157 (analogous to the V1V2 peptide), 385–418 and 296–331. It further demonstrates, in the context of Env trimer, that glycosylation decreases the amount of alpha-helix, beta strands, bridge and turns in these regions.

doi:10.1371/journal.pcbi.1005094.g014

(RMSF) of the backbone atoms encompassing the V1/V2 construct sequence. The RMSF observed for the non-glycosylated sequence are at least two times greater than the glycosylated counterpart, indicating a significant reduction in flexibility imposed by glycosylation. Furthermore, the accumulated configuration entropy over time for the same V1/V2 region (**Fig 14B**) shows reduced entropy upon glycosylation. We also calculated the different coulombic contributions to the overall energetics within the V1/V2 sequence stabilization. Again, as shown in **Fig 14C** for the different interactions between the local protein region representing the peptide construct, glycans, and solvent, the results in the context of Env spike are similar to those from the isolated construct (see **Fig 10 top panel for comparison**). Therefore, whether in the context of the Env trimeric spike or as an isolated fragment, the backbone mobility of this V1V2 peptide is restricted by glycosylation. Furthermore, the presence of glycan affects the local intra-molecular interactions among protein residues well as their interactions with water molecules.

Importance of N-glycosylation sites adjacent to cysteine residues in the folding of the of the V1V2 domain

The studies described above indicate an important contribution of the highly conserved glycan at position at 156 adjacent to Cys157 in folding of the adjacent V2 peptide. The extreme conservation of this glycan in HIV-1 and among all primate lentiviruses (Cys196) also supports its role as an important element in the structural framework of the region (**S5 Fig**). These observations raise the question of whether this and other Cys-proximal glycans regulate the efficiency of processing and conformational folding in the context of the native Env protein. This was investigated by removing the 156 and 197 glycans and evaluating their effect on Env processing. To simplify this analysis, these studies were performed with SF162 Env, which conserves glycans 156 and 197 but lacks a glycan at position 160 (**Fig 15A**). For comparison, glycans in the same region that were not adjacent to the Cys residues, N136 and N188, were also mutated. The glycans at 156, 197, 136, and 188 were each eliminated by mutating the Ser or Thr residues in the corresponding motifs to Ala residues (**Fig 15A**).

Mutation of 156 and 197 glycosylation motifs resulted in a reduction in size of the corresponding gp120 proteins (**Fig 15B**). This was consistent with the loss of an N-linked glycan, and showed that both of these positions were glycosylated in the wild type (wt) Env protein. The effects of these mutations on Env folding were examined by comparing the intracellular forms of Env present after a 5 hr labeling period for the wt Env (lane 1), mutants containing the 156 or 197 mutation (lanes 2 and 3), and a 156/197 double mutant (lane 4) (**Fig 15C**). Immunoprecipitation performed with polyclonal HIVIG showed that the majority of the mutant Env protein remained in the unprocessed gp160 form. Processing to gp120 appeared to be impaired to a greater extent for the 156 glycan (lane 2) compared to the 197 glycan (lane 3), while processing was completely abrogated for the double mutation (lane 4).

Similarly impaired processing to gp120 was observed for samples immunoprecipitated with mAbs b12 and 5145A, which are specific for conformational epitopes in the CD4-binding domain (**Fig 15C**, upper panels). These antibodies recognized similar levels of the wt and mutant gPr160 precursors, indicating that these mutations did not affect the folding events required for the formation of the CD4-binding domain. However, recognition of gPr160 by mAbs to three conformational epitopes in the V2 domain was significantly reduced by mutation of the 197 glycan, and almost completely inhibited by the loss of the 156 glycan and the double mutation (**Fig 15C**, lower panels). Previous epitope mapping studies with mAb 830A indicated that this antibody recognizes a discontinuous conformational epitope that overlaps the α β -integrin binding site at positions 179–181. This epitope also involves other residues in

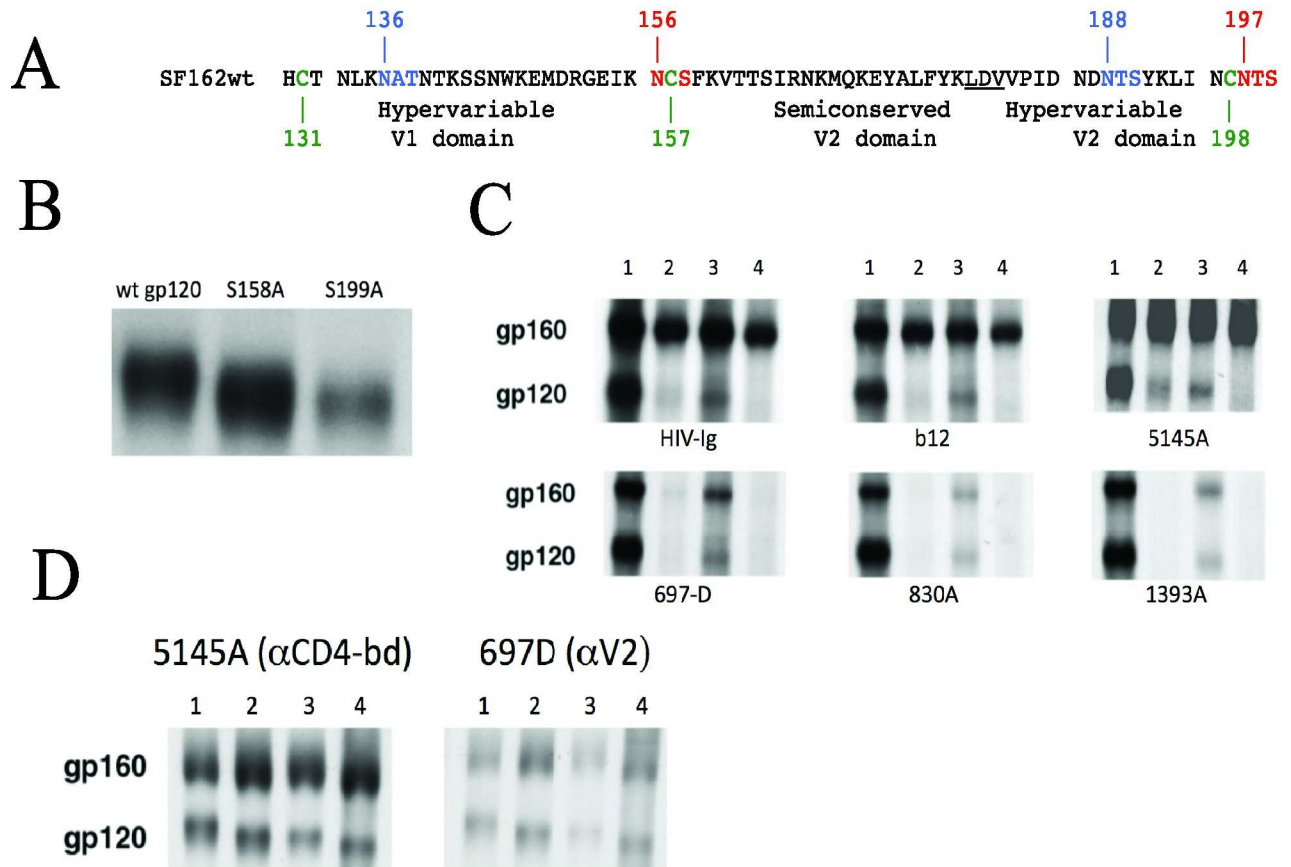


Fig 15. Effect of removing the glycosylation sites adjacent to Cys residues in V2 on processing of Env. (A) Sequence of wt SF162 Env protein. The two Cys-distal glycan motifs at positions 136 and 188 in the V1V2 hyper-variable region are shown in blue, while the two Cys-adjacent glycans at positions 156 and 197 in the semi-conserved V2 and C2 domains are indicated in red. Cys residues 131, 157, and 198 are indicated in green. (B) SDS-PAGE analysis of wt SF162 gp120 and gp120 containing mutated Cys-adjacent glycosylation sites at position 156 (S158A) and 197 (S199A). The increase in mobility is consistent with the loss of a glycan at these positions, confirming that these sites are in fact utilized. (C) Analysis of removing two Cys-adjacent glycan sites on intracellular processing of SF162 Env. Plasmids that encode wt and mutant SF162 Env proteins were transfected into 293T cells. Forty-eight hours post-transfection, the cells were radiolabeled with 35 S-cysteine for 5 hours, and cells were lysed and immunoprecipitated with polyclonal HIV+ antiserum (HIVIG), or mAbs b12 and 5145A directed against conformational epitopes in the CD4-binding domain (top panels), or mAbs 697D, 830A and 1393A directed against conformational epitopes in the V1/V2 domain (bottom panels). Lane 1—wt Env, lane 2- S158A mutant, lane 3—S199A mutant, lane 4—S158A/S199A double mutant. (D) Analysis of removing two Cys-distal glycans on intracellular SF162 Env processing. Cells transfected with wt SF162 Env (lanes 1), T138A (V1) mutant (lane 2), S190A (V2) mutant (lane 3) and T138A/S190A double mutant were labeled for 5 hrs, and then cells were lysed and labeled Env proteins immunoprecipitated with mAbs recognizing conformational epitopes in the CD4-binding domain (5145A) or in the V1/V2 domain (697D). Precursor gp160 and processed gp120 bands are indicated in (C) and (D).

doi:10.1371/journal.pcbi.1005094.g015

both the V1 and V2 domains, and a crystal structure of a V1V2 scaffolded molecule complexed with 830A demonstrated that this epitope did not include any glycans [45], suggesting that the reduced recognition of the glycosylation mutants by the V2-specific antibodies was not due to direct mutation of these epitopes. A similar analysis of the 136 and 188 glycan mutants, which are not adjacent to the Cys residues, revealed that processing of the single and double mutant proteins was similar to the wt protein (Fig 15D). These results indicate that processing of gp160 to gp120 was significantly impaired by loss of the 156 and 197 glycans, but not by loss of the 136 and 188 glycans. Furthermore, the conformation of the V1/V2 domain was altered by removal of the two Cys-adjacent glycosylation sites, such that recognition by three conformationally-dependent V2-directed mAbs was significantly reduced.

Discussion

Structural information about the effects of glycosylation on HIV-1 gp120, and in particular the hyper-variable domains, is still limited. This is due in part to the difficulties in obtaining crystal structures with glycans and hyper-variable domains intact. MD simulations can thus provide unique insight that enhances our understanding of the glycosylated gp120 structure. In this study, we used the enhanced sampling approach of replica exchange MD simulations to investigate the effects of glycosylation on a flexible unstructured region of the V1V2 domain that is of immunologic interest. We generated a peptide fragment containing portions of the V1 and V2 loops, including an antigenic region of gp120 that was recognized by antibodies generated by vaccination and during HIV-1 infection. The two regions were linked by a disulfide bond, and contained high mannose glycans at positions N156 and N160, which are targeted by a class of broadly neutralizing antibodies. These studies were undertaken to provide information about how to mimic salient structural features of V1V2 that could enhance its immunogenicity, as well as to understand the selective pressures that underlie strongly conserved features within a highly variable domain.

The addition of glycans to the V1V2 peptide caused a strong enthalpic compensation that resulted in two complementary effects on this disordered flexible fragment. First, glycosylation stabilized the pre-formed conformation of this peptide. Second, it reduced the propensity of the unstructured peptide to form secondary structures. Paradoxically, glycosylation destabilized this disordered V1V2 fragment by reducing its secondary structure propensities, while at the same time stabilizing it by preventing the peptide fragment from unfolding. It is possible that with glycosylation, the free energy of the unfolded state is lower due to the increase in entropic and enthalpic components. In addition, glycosylation also disrupts intra- and inter-peptide interactions that might be important to the folding process of the peptide. The large volume of a carbohydrate moiety could also impede the rearrangement of the V1V2 structure during its folding process. On the other hand, it is possible that the free energy of the folded structure is also lower with glycosylation due to introduction of strong interactions between glycan-glycan and glycan-solvent. Also, glycosylation reduces the solvent accessible surface area of the peptide, thereby shielding the unfolding process from the solvent.

Based on our results, we postulate that this destabilization of secondary structure is a generalized effect of the glycan attached to unstructured regions of proteins. At the same time, our results show that glycosylation can prevent a pre-formed beta-strand structure in the peptide from thermal unfolding; the unfolding process was much slower in the presence of glycosylation. One could envision that such a situation arises during oligomerization of gp120 monomers or when V1V2 is bound to an antibody.

Even though addition of glycans to the V1V2 peptide disrupted significant peptide-peptide and peptide-solvent interactions, a much larger favorable enthalpic contribution was obtained from the glycan-glycan and the glycan-solvent interactions. In fact, the large enthalpic contribution from solvation leads to better hydration of the peptide. This is reflected in the free energy landscape that favors an extended conformation of the peptide in the water solution. Additionally, a significant enthalpic contribution originates from glycan-glycan interactions in the case of glycosylation sites that are spatially proximal in the peptide, such as N156 and N160. Our studies show that if two glycans occur at spatially close sites in a flexible region of the protein, they will cluster together. There is an additional site, N130, often found next to a disulfide linkage in HIV Env ([Fig 1C](#)), although not in the CAP45 sequence that we used in this study; when it is present, it might also impact the conformation of both the peptide and the intact trimer as discussed below.

In this study, we characterized the changes in the energy landscape and thermodynamics of an isolated, disulfide bound V1V2 peptide fragment upon glycosylation. Such a characterization is critical for designing an immunogen construct involving glycosylation to ensure that the important conformational characteristics of the peptide are not significantly altered. It is also possible that the effect of glycosylation is more dramatic in a peptide construct when the scaffolding influence of the rest of protein is absent. In order to address this potential limitation, we performed extensive all-atom MD simulations of the entire Env spike and found that glycans exert similar effects on the V1V2 immunodominant region even when considered in the context of the whole Env gp120 trimer. Thus, backbone mobility is restricted when glycosylation is present in the context of the isolated peptide and the Env trimer. Furthermore, our simulations revealed that the presence of the glycan affects the local intra-molecular interactions among protein residues well as their interactions with water molecules.

All-atom MD simulations of the Env spike also demonstrated that additional glycans from the trimeric complex interact with the glycan moieties at positions 156 and 160. Preliminary results from the Env spike simulations show that the glycan at position 156 makes contact with other glycans within the same protomer, whereas the glycan at position 160 interacts with glycans from neighboring protomers at top of the spike (S6 Fig). Thus, it is likely that glycan interactions with 156 and 160 contribute to the stabilization of the V1/V2 immunodominant region and, more importantly, to the integrity of the trimer. Indeed, studies are currently underway to understand how glycan-glycan interactions contribute to the stability of the trimer.

Another study used simulations to investigate the effects of glycosylation on the mobility and conformation of the V3 domain in gp120 [85]. In that study, unglycosylated gp120 was compared with gp120 containing either a single or multiple proximal high mannose N-linked glycans. That study reported that glycans surrounding the disulfide bounded V3 domain modulate its dynamics and conformational properties. The glycans tended to constrain the movement of V3, and cause it to adopt a more narrow conformation than the non-glycosylated gp120 form. Interestingly, the glycans flanking the V3 domain are less well conserved across HIV-1 clades and circulating recombinant forms (CRFs) than those in V1V2. The N-linked glycan motif at N295, which is N-terminal to V3, is found less frequently in clades A and C compared to other group M clades and CRFs. At the C-terminal flank of V3, the N334 glycan addition site is highly conserved in CRF01_AE, but is more variable in other clades and CRFs, which tend to have higher frequencies of the N332 glycan addition site instead. Thus, the cysteine-bounded loops of gp120 are often flanked by glycans that are conserved to varying degrees, and therefore may be particularly susceptible to their influence.

Consistent with this concept is strong evidence that the placement of N-linked glycosylation sites adjacent to disulfide bonds is a highly conserved feature at the base of the V2 loop (and also modestly conserved at the base of V1, V3, and V4). Disulfide bonds are commonly found in proteins, but their effects on protein structure are still under investigation [10, 86–89]. Moreover, the effect of having a glycan juxtaposed to a disulfide bond has not been addressed from a structural perspective. In the peptide construct considered in this study, the glycan at 156 is adjacent to a disulfide bond. Past studies have shown that elimination of 156 and 160 glycans in the HIV-1 DH12 infectious molecular clone resulted in a loss of infectivity that was attributed to a defect in CD4 binding [90]. Furthermore, our studies demonstrated that removal of 156 and 160 from the HIV-1 SF162 Env impaired gp160 processing to gp120 and disrupted the conformation of V1V2. Mutation of V1V2 glycans that were not adjacent to Cys residues did not reproduce these effects. Thus, it is likely that Cys-adjacent glycans play an important role in HIV-1 Env processing, folding, and function.

We also evaluated a peptide region of V2, based on the consensus C sequence, that lacked the disulfide bond and the V1 fragment. By comparing the effects of the disulfide bond with or without glycosylation in the V1V2 and a second V2 peptide, we found that the residues adjacent to the disulfide bonded cysteine residues had a higher propensity to form beta-strand structure compared to other residues in the peptide (S7 Fig). However, addition of a glycan next to the disulfide diminished the propensity to form beta structures. In the context Env trimer, examination of other V1V2 and other disulfide regions with proximal glycans revealed that secondary structures are slightly diminished by the presence of glycosylation (Fig 14D). Taken together, these findings suggest that the preservation of the glycan next to the disulfide bond is most beneficial during folding or upon pre-forming a stable secondary conformation induced by binding to other proteins or antibodies.

Presumably, a major function of the Cys-adjacent glycans is to regulate the efficiency and specificity of disulfide bond formation. Intuitively, the large size of the glycans would limit degrees of rotation, which could regulate the orientation of the two peptide strands on either end of the disulfide. This may be particularly important in cases where there are proximal glycans to both partners of a disulfide bond, such as is observed for the Cys 131- Cys 157 disulfide bond that closes V1 and the disulfide bond at the base of the V3 loop. A somewhat under-appreciated consideration is that recombinant gp120 proteins possess considerable heterogeneity specifically in the V1V2 region [91], which could be influenced by the adjacent glycosylation. Consistent with this concept, mass spectroscopy studies have provided evidence for alternative disulfide pairing in the V1V2 region of the recombinant CON-S gp140 ΔCFI protein [91] and this may be influenced by glycosylation. There is also evidence that disulfide reorganization occurs after receptor binding and is mediated by membrane-associated protein-disulfide isomerases [92]. This could also be affected by proximal glycans. Thus, further studies are warranted to explore whether these effects are present only in lentiviral glycoproteins.

Here, we addressed the effect of glycosylation of a V1V2 peptide that generally exists in a disordered conformation in solution. From the standpoint of immunogen design, thermodynamics elucidated from the current study provide insightful strategies to stabilize the V1V2 peptide and drive it towards the formation of beta-strand structures that could be desirable for eliciting broadly neutralizing antibodies over those that recognize linear, non-neutralizing or strain-specific V1V2 epitopes. The RV144 human vaccine trial elicited cross-reactive but weakly neutralizing antibodies directed against epitopes in V2 that were inversely correlated with the rate of HIV-1 infection [93]. Structural studies indicated that the key region of V2 (residues 168–176) recognized by vaccine elicited non-neutralizing mAbs CH58 and CH59, and the V1V2-targeted broadly neutralizing mAb PG9, can exist in multiple conformations [46, 47]. PG9 appears to preferentially bind to a beta strand conformation, whereas CH58 and CH59 may recognize alternate forms [46, 47]. An additional key feature of V1V2 directed broadly neutralizing antibodies such as PG9 is their ability to bind to glycans at N156 and N160, in addition to the underlying peptide (17). Alam et al. [40] described V1V2 glycopeptide immunogens that bind with high affinity to mature V1V2 broadly neutralizing antibodies and their putative germlines, but with much lower affinity to the vaccine-elicited, strain-specific V2 antibodies [40]. Also, they elegantly showed the importance of disulfide bonds in their peptide constructs, which is consistent with our findings. Therefore, glycopeptide immunogens represent a viable strategy to elicit V1V2-directed broadly neutralizing antibodies, but the effects of glycan-proximal disulfide bonds will also need to be considered. The region that intervenes between the V2 ‘epitope’ region and the Cys involved in V2 loop closure is a hyper-variable segment. Thus, peptides and scaffolds that encompass the end of the V2 loop and include the conserved glycosylation sites at the base of the V1V2 region, such as those described in [93], may better mimic the structure found in a native Env trimer. However, this region also spans

sequences that are unique and highly distinctive between every isolate, a balance to consider in immunogen and reagent design.

Finally, one aspect that was not considered in the current study is the heterogeneity of carbohydrate forms at glycan sites, in particular at N156 and N160. A recent study using BG505 trimer protein demonstrated that N156 and N160 participate in a glycan ring at the trimer apex [52]. Both glycans were found to be predominantly of the oligomannose type, consisting of varying proportions $\text{Man}_s\text{GlcNAc}_2$. The authors proposed that glycan processing at N156 and N160 is likely to be constrained by inter-protomer contacts, resulting in minimal processing, although N160 tends to be more heterogenous than N156 [52]. Likewise, a study using a clade G SOSIP trimer demonstrated that the trimer apex is a region of glycan crowding, with less processing than glycans occurring in more dispersed regions [53]. The results of these trimer-based studies stand in contrast with that of Amin et al. [48], which evaluated cyclic V1V2 peptides that contain glycans attached at N156 (or N173 which substitutes for N156 in some isolates) and N160. Interestingly, the N-glycans at N156 and N173 have been shown to be spatially equivalent in terms of antibody recognition. The authors found that $\text{Man}_s\text{GlcNAc}_2$ glycan was required at N160 for recognition by broadly neutralizing antibodies PG9 and PG16. Furthermore, a sialylated N-glycan at the secondary site (N156 or N173) was also necessary for antibody binding to the glycopeptide. However, in the context of the BG505 trimer, PG9 binds regardless of whether the protein was produced in the presence or absence of glycan processing, supporting that these glycans are likely to be composed mainly of oligomannose forms in the native envelope trimer [94]. Taken together, these studies suggest that a better understanding of the structural features and immunogenicity of V1V2, and the conformational forms recognized by broadly neutralizing antibodies, could lead to the development of novel glycopeptide immunogens.

Conclusion

A major difficulty in the pursuit of incorporating V1V2 epitopes into HIV vaccine design is the structural heterogeneity and variable glycosylation of this immunogenic region. Limited knowledge of how glycosylation and disulfide bonds affect the conformation and dynamics of short intrinsically disordered peptides complicates the design of immunogenic peptides. Thus, the development of strategies to define and exploit optimal configurations of V1V2 epitopes is important. Here, we used extensive replica exchange and conventional MD simulations to characterize the effects of glycosylation on the free energy landscape of a disulfide bound V1V2 peptide and dissect the enthalpic and entropic components upon addition of a glycan. Our analyses demonstrated that glycosylation stabilizes the pre-existing conformation of this peptide, and reduces its propensity to form other secondary structures. However, glycosylation also stabilizes the V1V2 peptide against thermal unfolding, and exhibits specific effects in relation to the adjacent disulfide linkage. These complementary effects originate from a combination of multiple factors, including the observation that having a disulfide bond adjacent to the glycan sites further promotes the formation of beta-strand structure in this peptide. Glycosylation and disulfide linkage are therefore likely important components that contribute to the immunogenicity of this region of V1V2, and will influence whether the appropriate conformation is adopted. The observation that HIV-1 is under strong selective pressure to conserve glycans adjacent to disulfide bonds could perhaps be exploited in the design of immunogens.

Supporting Information

S1 Fig. The HXB2 reference strain gp120 sequence was used to illustrate the association between glycosylation sites and the 8 Cysteines that form disulfide bonds at the base of the

variable loops V1-V4. Of the 23 Cys residues in HXB2, 8 form the base of a variable loop, and 5 of the 8 have an adjacent N linked glycosylation site. In contrast, 0 of the 15 Cys residues that do not occur at the base of a variable loop have an adjacent N linked glycosylation site (Fisher's exact, $p = 0.0016$).
(EPS)

S2 Fig. Conservation of N-linked glycosylation sites N156 and N160 across 4633 HIV-1 group M sequences arranged by clade. Green indicates the fraction of sequences that contain both N156 and N160; red indicates that N156 is present; purple indicates that N160 is present, and blue indicates the fraction of sequences that lacks both sites. N160 tends to be absent more frequently in the B subtype, while N156 is absent more frequently in the D subtype.
(EPS)

S3 Fig. Net charge distribution of the variable regions. Using the same input data as Fig 1, here we show that there is a great deal of charge variability in all of the variable regions. The V3 loop and V2 epitope region, despite being conserved in terms of length and number of glycosylation sites (Fig 1), both show a great deal of variation in net charge, comparable the level of diversity found in hypervariable regions. Net charge is calculated as the sum of positive and negative charges, where E and D are assigned -1, and K, D, and H are assigned +1. The V2 epitope region, V3 and V2 tend to be positively charged, V1, V4, V5, negatively charged.
(EPS)

S4 Fig. Configurational entropy of the peptide backbone. The estimated entropy was obtained using 1200 frames of simulations at 300K (lines) and 450K (circles) from a total 250 ns trajectory, for both the single peptide (black) and after glycosylation (red). The small inset shows a close-up for higher resolution.
(EPS)

S5 Fig. Conservation of N-linked proximal glycosylation sites between 72 HIV-2 and non-human primate representative lentiviral sequences from the HIV database. The V2 and V3 C-terminal proximal N-linked glycosylation sites at C196 and C331 are particularly well conserved.
(EPS)

S6 Fig. Glycan-glycan interactions in the context of gp120 trimer of Env spike. Key glycan contacts in the V1V2 region are provided based on 1 us all-atom MD simulations of the fully glycosylated Env spike. The distances between the center of mass (COM) of glycans are shown. Distances are computed for inter-protomeric contacts (A) between glycans 133–156 (blue and red respectively) and a snapshot of the ensemble configuration is shown in the circled inset. Intra-protomeric contacts are shown for glycans located in position 160 (B) as well as for the pairs 185–156, 180–156 and 197–156 (C). Each protomer is denoted by a number from 1 to 3.
(EPS)

S7 Fig. Fraction of each residue in the V2 and V2g systems that form beta-strand structures. The two blue-labeled "N" residues are the two glycosylated residues and the two red-labeled "C" residues are the two residues that are disulfide bonded.
(EPS)

Acknowledgments

We would like to acknowledge the LANL institutional computing resource, which was used for carrying out all-atom molecular dynamics simulations.

Author Contributions

Conceived and designed the experiments: JT.

Performed the experiments: JT MSJ BK.

Analyzed the data: JT CAL SG MSJ AP.

Contributed reagents/materials/analysis tools: JT.

Wrote the paper: JT CAL CAD AP SG BK.

References

1. Mononen I, Karjalainen E. Structural comparison of protein sequences around potential N-glycosylation sites. *BBA Protein Struct M.* 1984; 788(3):364–7. doi: [http://dx.doi.org/10.1016/0167-4838\(84\)90050-5](http://dx.doi.org/10.1016/0167-4838(84)90050-5)
2. Beintema J. Do asparagine-linked carbohydrate chains in glycoproteins have a preference for β -bends? *Bioscience Rep.* 1986; 6(8):709–14. doi: [10.1007/BF01116537](https://doi.org/10.1007/BF01116537)
3. Myers G, Lenroot R. HIV glycosylation: what does it portend? *AIDS research and human retroviruses.* 1992; 8(8):1459–60. doi: [10.1089/aid.1992.8.1459](https://doi.org/10.1089/aid.1992.8.1459) PMID: [1466982](https://pubmed.ncbi.nlm.nih.gov/1466982/).
4. Wyatt R, Sodroski J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science.* 1998; 280(5371):1884–8. doi: [10.1126/science.280.5371.1884](https://doi.org/10.1126/science.280.5371.1884) PMID: [9632381](https://pubmed.ncbi.nlm.nih.gov/9632381/).
5. Wills C, Farmer A, Myers G. Rapid sequon evolution in human immunodeficiency virus type 1 relative to human immunodeficiency virus type 2. *AIDS research and human retroviruses.* 1996; 12(14):1383–4. doi: [10.1089/aid.1996.12.1383](https://doi.org/10.1089/aid.1996.12.1383) PMID: [8891118](https://pubmed.ncbi.nlm.nih.gov/8891118/).
6. Wormald MR, Dwek RA. Glycoproteins: glycan presentation and protein-fold stability. *Structure.* 1999; 7(7):R155–60. doi: [10.1016/s0969-2126\(99\)80095-1](https://doi.org/10.1016/s0969-2126(99)80095-1) PMID: [10425673](https://pubmed.ncbi.nlm.nih.gov/10425673/).
7. Hanson SR, Culyba EK, Hsu T-L, Wong C-H, Kelly JW, Powers ET. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proceedings of the National Academy of Sciences of the United States of America.* 2009; 106(9):3131–6. doi: [10.1073/pnas.0810318105](https://doi.org/10.1073/pnas.0810318105) PMID: [19204290](https://pubmed.ncbi.nlm.nih.gov/19204290/)
8. Imperiali B, O'Connor SE. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr Opin Chem Biol.* 1999; 3(6):643–9. doi: [10.1016/s1367-5931\(99\)00021-6](https://doi.org/10.1016/s1367-5931(99)00021-6) PMID: [10600722](https://pubmed.ncbi.nlm.nih.gov/10600722/)
9. Mitra N, Sinha S, Ramya TN, Surolia A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem Sci.* 2006; 31(3):156–63. Epub 2006/02/14. doi: [10.1016/j.tibs.2006.01.003](https://doi.org/10.1016/j.tibs.2006.01.003) PMID: [16473013](https://pubmed.ncbi.nlm.nih.gov/16473013/).
10. Bosques CJ, Imperiali B. The interplay of glycosylation and disulfide formation influences fibrillization in a prion protein fragment. *Proceedings of the National Academy of Sciences of the United States of America.* 2003; 100(13):7593–8. doi: [10.1073/pnas.1232504100](https://doi.org/10.1073/pnas.1232504100) PMID: [12805563](https://pubmed.ncbi.nlm.nih.gov/12805563/)
11. Petrescu A-J, Milac A-L, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology.* 2004; 14(2):103–14. doi: [10.1093/glycob/cwh008](https://doi.org/10.1093/glycob/cwh008) PMID: [14514716](https://pubmed.ncbi.nlm.nih.gov/14514716/)
12. O'Connor SE, Pohlmann J, Imperiali B, Saskiawan I, Yamamoto K. Probing the Effect of the Outer Saccharide Residues of N-Linked Glycans on Peptide Conformation. *J Am Chem Soc.* 2001; 123(25):6187–8. doi: [10.1021/ja010094s](https://doi.org/10.1021/ja010094s) PMID: [11414857](https://pubmed.ncbi.nlm.nih.gov/11414857/)
13. DeKoster GT, Robertson AD. Thermodynamics of Unfolding for Kazal-Type Serine Protease Inhibitors: Entropic Stabilization of Ovomuroid First Domain by Glycosylation. *Biochemistry.* 1997; 36(8):2323–31. doi: [10.1021/bi962580b](https://doi.org/10.1021/bi962580b) PMID: [9047335](https://pubmed.ncbi.nlm.nih.gov/9047335/)
14. Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proceedings of the National Academy of Sciences of the United States of America.* 2008; 105(24):8256–61. doi: [10.1073/pnas.0801340105](https://doi.org/10.1073/pnas.0801340105) PMID: [18550810](https://pubmed.ncbi.nlm.nih.gov/18550810/)
15. Banks DD. The Effect of Glycosylation on the Folding Kinetics of Erythropoietin. *J Mol Biol.* 2011; 412(3):536–50. doi: <http://dx.doi.org/10.1016/j.jmb.2011.07.061> PMID: [21839094](https://pubmed.ncbi.nlm.nih.gov/21839094/)
16. Holst B, Bruun AW, Kielland-Brandt MC, Winther JR. Competition between folding and glycosylation in the endoplasmic reticulum. *EMBO J.* 1996; 15(14):3538–46. PMID: [8670857](https://pubmed.ncbi.nlm.nih.gov/8670857/)
17. Chen MM, Bartlett AI, Nerenberg PS, Friel CT, Hackenberger CPR, Stultz CM, et al. Perturbing the folding energy landscape of the bacterial immunity protein Im7 by site-specific N-linked glycosylation.

- Proceedings of the National Academy of Sciences of the United States of America. 2010; 107 (52):22528–33. doi: [10.1073/pnas.1015356107](https://doi.org/10.1073/pnas.1015356107) PMID: [21148421](https://pubmed.ncbi.nlm.nih.gov/21148421/)
18. Elliott S, Chang D, Delorme E, Eris T, Lorenzini T. Structural Requirements for Additional N-Linked Carbohydrate on Recombinant Human Erythropoietin. *The Journal of biological chemistry*. 2004; 279 (16):16854–62. doi: [10.1074/jbc.M311095200](https://doi.org/10.1074/jbc.M311095200) PMID: [14757769](https://pubmed.ncbi.nlm.nih.gov/14757769/)
 19. Price JL, Shental-Bechor D, Dhar A, Turner MJ, Powers ET, Gruebele M, et al. Context-Dependent Effects of Asparagine Glycosylation on Pin WW Folding Kinetics and Thermodynamics. *J Am Chem Soc*. 2010; 132(43):15359–67. doi: [10.1021/ja106896t](https://doi.org/10.1021/ja106896t) PMID: [20936810](https://pubmed.ncbi.nlm.nih.gov/20936810/)
 20. Riederer MA, Hinnen A. Removal of N-glycosylation sites of the yeast acid phosphatase severely affects protein folding. *Journal of bacteriology*. 1991; 173(11):3539–46. PMID: [2045373](https://pubmed.ncbi.nlm.nih.gov/2045373/); PubMed Central PMCID: [PMC207969](https://pubmed.ncbi.nlm.nih.gov/PMC207969/).
 21. Imberty A, Pérez S. Stereochemistry of the N-glycosylation sites in glycoproteins. *Protein Eng*. 1995; 8 (7):699–709. doi: [10.1093/protein/8.7.699](https://doi.org/10.1093/protein/8.7.699) PMID: [8577698](https://pubmed.ncbi.nlm.nih.gov/8577698/)
 22. Hindsgaul OC, R. D. *Essentials of Glcobiology*: Cold Spring Harbor (NY); 1999.
 23. Ellis CR, Maiti B, Noid WG. Specific and Nonspecific Effects of Glycosylation. *J Am Chem Soc*. 2012; 134(19):8184–93. doi: [10.1021/ja301005f](https://doi.org/10.1021/ja301005f) PMID: [22524526](https://pubmed.ncbi.nlm.nih.gov/22524526/)
 24. Nishimura I, Uchida M, Inohana Y, Setoh K, Daba K, Nishimura S, et al. Oxidative Refolding of Bovine Pancreatic RNases A and B Promoted by Asn-Glycans. *J Biochem*. 1998; 123(3):516–20. doi: [10.1093/oxfordjournals.jbchem.a021966](https://doi.org/10.1093/oxfordjournals.jbchem.a021966) PMID: [9538236](https://pubmed.ncbi.nlm.nih.gov/9538236/)
 25. Jitsuhara Y, Toyoda T, Itai T, Yamaguchi H. Chaperone-Like Functions of High-Mannose Type and Complex-Type N-Glycans and Their Molecular Basis. *J Biochem*. 2002; 132(5):803–11. doi: [10.1093/oxfordjournals.jbchem.a003290](https://doi.org/10.1093/oxfordjournals.jbchem.a003290) PMID: [12417032](https://pubmed.ncbi.nlm.nih.gov/12417032/)
 26. Beckham GT, Bomble YJ, Matthews JF, Taylor CB, Resch MG, Yarbrough JM, et al. The O-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophys J*. 2010; 99(11):3773–81. Epub 2010/11/30. doi: [10.1016/j.bpj.2010.10.032](https://doi.org/10.1016/j.bpj.2010.10.032) PMID: [21112302](https://pubmed.ncbi.nlm.nih.gov/21112302/); PubMed Central PMCID: [PMC2998629](https://pubmed.ncbi.nlm.nih.gov/PMC2998629/).
 27. Cheng S, Edwards SA, Jiang Y, Gräter F. Glycosylation Enhances Peptide Hydrophobic Collapse by Impairing Solvation. *Chem Phys Chem*. 2010; 11(11):2367–74. doi: [10.1002/cphc.201000205](https://doi.org/10.1002/cphc.201000205) PMID: [20583025](https://pubmed.ncbi.nlm.nih.gov/20583025/)
 28. Lu D, Yang C, Liu Z. How Hydrophobicity and the Glycosylation Site of Glycans Affect Protein Folding and Stability: A Molecular Dynamics Simulation. *J Phys Chem B*. 2011; 116(1):390–400. doi: [10.1021/jp203926r](https://doi.org/10.1021/jp203926r) PMID: [22118044](https://pubmed.ncbi.nlm.nih.gov/22118044/)
 29. Pancera M, Shahzad-UI-Hussan S, Doria-Rose NA, McLellan JS, Bailer RT, Dai K, et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nature structural & molecular biology*. 2013; 20(7):804–13. doi: [10.1038/nsmb.2600](https://doi.org/10.1038/nsmb.2600) PMID: [23708607](https://pubmed.ncbi.nlm.nih.gov/23708607/); PubMed Central PMCID: [PMC4046252](https://pubmed.ncbi.nlm.nih.gov/PMC4046252/).
 30. McLellan JS, Pancera M, Carrico C, Gorman J, Julien J-P, Khayat R, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature*. 2011; 480(7377):336–43. <http://www.nature.com/nature/journal/v480/n7377/abs/nature10696.html>—supplementary-information. doi: [10.1038/nature10696](https://doi.org/10.1038/nature10696) PMID: [22113616](https://pubmed.ncbi.nlm.nih.gov/22113616/)
 31. Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien JP, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature*. 2011; 477(7365):466–70. doi: [10.1038/nature10373](https://doi.org/10.1038/nature10373) PMID: [21849977](https://pubmed.ncbi.nlm.nih.gov/21849977/); PubMed Central PMCID: [PMC3393110](https://pubmed.ncbi.nlm.nih.gov/PMC3393110/).
 32. Bonsignori M, Hwang KK, Chen X, Tsao CY, Morris L, Gray E, et al. Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. *Journal of virology*. 2011; 85(19):9998–10009. doi: [10.1128/JVI.05045-11](https://doi.org/10.1128/JVI.05045-11) PMID: [21795340](https://pubmed.ncbi.nlm.nih.gov/21795340/); PubMed Central PMCID: [PMC3196428](https://pubmed.ncbi.nlm.nih.gov/PMC3196428/).
 33. Wang LX. Synthetic carbohydrate antigens for HIV vaccine design. *Curr Opin Chem Biol*. 2013; 17 (6):997–1005. doi: [10.1016/j.cbpa.2013.10.001](https://doi.org/10.1016/j.cbpa.2013.10.001) PMID: [24466581](https://pubmed.ncbi.nlm.nih.gov/24466581/); PubMed Central PMCID: [PMC4100479](https://pubmed.ncbi.nlm.nih.gov/PMC4100479/).
 34. Lynch RM, Rong R, Boliar S, Sethi A, Li B, Mulenga J, et al. The B cell response is redundant and highly focused on V1V2 during early subtype C infection in a Zambian seroconverter. *Journal of virology*. 2011; 85(2):905–15. doi: [10.1128/JVI.02006-10](https://doi.org/10.1128/JVI.02006-10) PMID: [20980495](https://pubmed.ncbi.nlm.nih.gov/20980495/); PubMed Central PMCID: [PMC3020014](https://pubmed.ncbi.nlm.nih.gov/PMC3020014/).
 35. Moore PL, Gray ES, Wibmer CK, Bhiman JN, Nonyane M, Sheward DJ, et al. Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nature medicine*. 2012; 18(11):1688–92. doi: [10.1038/nm.2985](https://doi.org/10.1038/nm.2985) PMID: [23086475](https://pubmed.ncbi.nlm.nih.gov/23086475/); PubMed Central PMCID: [PMC3494733](https://pubmed.ncbi.nlm.nih.gov/PMC3494733/).
 36. Murphy MK, Yue L, Pan R, Boliar S, Sethi A, Tian J, et al. Viral escape from neutralizing antibodies in early subtype A HIV-1 infection drives an increase in autologous neutralization breadth. *PLoS*

- pathogens. 2013; 9(2):e1003173. doi: [10.1371/journal.ppat.1003173](https://doi.org/10.1371/journal.ppat.1003173) PMID: [23468623](https://pubmed.ncbi.nlm.nih.gov/23468623/); PubMed Central PMCID: [PMC3585129](https://pubmed.ncbi.nlm.nih.gov/PMC3585129/).
37. Rong R, Li B, Lynch RM, Haaland RE, Murphy MK, Mulenga J, et al. Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. *PLoS pathogens*. 2009; 5(9):e1000594. doi: [10.1371/journal.ppat.1000594](https://doi.org/10.1371/journal.ppat.1000594) PMID: [19763269](https://pubmed.ncbi.nlm.nih.gov/19763269/); PubMed Central PMCID: [PMC2741593](https://pubmed.ncbi.nlm.nih.gov/PMC2741593/).
 38. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. *Nature*. 2003; 422(6929):307–12. doi: [10.1038/nature01470](https://doi.org/10.1038/nature01470) PMID: [12646921](https://pubmed.ncbi.nlm.nih.gov/12646921/).
 39. Wibmer CK, Bhiman JN, Gray ES, Tumba N, Abdool Karim SS, Williamson C, et al. Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. *PLoS pathogens*. 2013; 9(10):e1003738. doi: [10.1371/journal.ppat.1003738](https://doi.org/10.1371/journal.ppat.1003738) PMID: [24204277](https://pubmed.ncbi.nlm.nih.gov/24204277/); PubMed Central PMCID: [PMC3814426](https://pubmed.ncbi.nlm.nih.gov/PMC3814426/).
 40. Alam SM, Dennison SM, Aussedat B, Vohra Y, Park PK, Fernandez-Tejada A, et al. Recognition of synthetic glycopeptides by HIV-1 broadly neutralizing antibodies and their unmutated ancestors. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110(45):18214–9. doi: [10.1073/pnas.1317855110](https://doi.org/10.1073/pnas.1317855110) PMID: [24145434](https://pubmed.ncbi.nlm.nih.gov/24145434/); PubMed Central PMCID: [PMC3831483](https://pubmed.ncbi.nlm.nih.gov/PMC3831483/).
 41. Aussedat B, Vohra Y, Park PK, Fernandez-Tejada A, Alam SM, Dennison SM, et al. Chemical synthesis of highly congested gp120 V1V2 N-glycopeptide antigens for potential HIV-1-directed vaccines. *J Am Chem Soc*. 2013; 135(35):13113–20. doi: [10.1021/ja405990z](https://doi.org/10.1021/ja405990z) PMID: [23915436](https://pubmed.ncbi.nlm.nih.gov/23915436/); PubMed Central PMCID: [PMC3826081](https://pubmed.ncbi.nlm.nih.gov/PMC3826081/).
 42. Morales JF, Morin TJ, Yu B, Tatsuno GP, O'Rourke SM, Theolis R Jr., et al. HIV-1 envelope proteins and V1/V2 domain scaffolds with mannose-5 to improve the magnitude and quality of protective antibody responses to HIV-1. *The Journal of biological chemistry*. 2014; 289(30):20526–42. doi: [10.1074/jbc.M114.554089](https://doi.org/10.1074/jbc.M114.554089) PMID: [24872420](https://pubmed.ncbi.nlm.nih.gov/24872420/); PubMed Central PMCID: [PMC4110267](https://pubmed.ncbi.nlm.nih.gov/PMC4110267/).
 43. Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature*. 2014; 509(7498):55–62. doi: [10.1038/nature13036](https://doi.org/10.1038/nature13036) PMID: [24590074](https://pubmed.ncbi.nlm.nih.gov/24590074/); PubMed Central PMCID: [PMC4395007](https://pubmed.ncbi.nlm.nih.gov/PMC4395007/).
 44. Gorman J, Soto C, Yang MM, Davenport TM, Guttman M, Bailer RT, et al. Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. *Nature structural & molecular biology*. 2016; 23(1):81–90. doi: [10.1038/nsmb.3144](https://doi.org/10.1038/nsmb.3144) PMID: [26689967](https://pubmed.ncbi.nlm.nih.gov/26689967/); PubMed Central PMCID: [PMC4833398](https://pubmed.ncbi.nlm.nih.gov/PMC4833398/).
 45. Pan R, Gorny MK, Zolla-Pazner S, Kong XP. The V1V2 Region of HIV-1 gp120 Forms a Five-Stranded Beta Barrel. *Journal of virology*. 2015; 89(15):8003–10. doi: [10.1128/JVI.00754-15](https://doi.org/10.1128/JVI.00754-15) PMID: [26018158](https://pubmed.ncbi.nlm.nih.gov/26018158/); PubMed Central PMCID: [PMC4505664](https://pubmed.ncbi.nlm.nih.gov/PMC4505664/).
 46. McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature*. 2011; 480(7377):336–43. doi: [10.1038/nature10696](https://doi.org/10.1038/nature10696) PMID: [22113616](https://pubmed.ncbi.nlm.nih.gov/22113616/); PubMed Central PMCID: [PMC3406929](https://pubmed.ncbi.nlm.nih.gov/PMC3406929/).
 47. Liao HX, Bonsignori M, Alam SM, McLellan JS, Tomaras GD, Moody MA, et al. Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. *Immunity*. 2013; 38(1):176–86. Epub 2013/01/15. doi: [10.1016/j.immuni.2012.11.011](https://doi.org/10.1016/j.immuni.2012.11.011) PMID: [23313589](https://pubmed.ncbi.nlm.nih.gov/23313589/); PubMed Central PMCID: [PMC3569735](https://pubmed.ncbi.nlm.nih.gov/PMC3569735/).
 48. Amin MN, McLellan JS, Huang W, Orwenyo J, Burton DR, Koff WC, et al. Synthetic glycopeptides reveal the glycan specificity of HIV-neutralizing antibodies. *Nature chemical biology*. 2013; 9(8):521–6. doi: [10.1038/nchembio.1288](https://doi.org/10.1038/nchembio.1288) PMID: [23831758](https://pubmed.ncbi.nlm.nih.gov/23831758/); PubMed Central PMCID: [PMC3730851](https://pubmed.ncbi.nlm.nih.gov/PMC3730851/).
 49. Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *The Journal of biological chemistry*. 1990; 265(18):10373–82. PMID: [2355006](https://pubmed.ncbi.nlm.nih.gov/2355006/).
 50. Doores KJ, Bonomelli C, Harvey DJ, Vasiljevic S, Dwek RA, Burton DR, et al. Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc Natl Acad Sci U S A*. 2010; 107(31):13800–5. Epub 2010/07/21. doi: [10.1073/pnas.1006498107](https://doi.org/10.1073/pnas.1006498107) PMID: [20643940](https://pubmed.ncbi.nlm.nih.gov/20643940/); PubMed Central PMCID: [PMC2922250](https://pubmed.ncbi.nlm.nih.gov/PMC2922250/).
 51. Bonomelli C, Doores KJ, Dunlop DC, Thaney V, Dwek RA, Burton DR, et al. The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. *PloS one*. 2011; 6(8):e23521. doi: [10.1371/journal.pone.0023521](https://doi.org/10.1371/journal.pone.0023521) PMID: [21858152](https://pubmed.ncbi.nlm.nih.gov/21858152/); PubMed Central PMCID: [PMC3156772](https://pubmed.ncbi.nlm.nih.gov/PMC3156772/).
 52. Behrens AJ, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, et al. Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *Cell reports*. 2016; 14(11):2695–706. doi: [10.1016/j.celrep.2016.02.058](https://doi.org/10.1016/j.celrep.2016.02.058) PMID: [26972002](https://pubmed.ncbi.nlm.nih.gov/26972002/); PubMed Central PMCID: [PMC4805854](https://pubmed.ncbi.nlm.nih.gov/PMC4805854/).

53. Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, et al. Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell*. 2016; 165(4):813–26. doi: [10.1016/j.cell.2016.04.010](https://doi.org/10.1016/j.cell.2016.04.010) PMID: [27114034](https://pubmed.ncbi.nlm.nih.gov/27114034/).
54. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins-Structure Function and Bioinformatics*. 2006; 65(3):712–25. doi: [10.1002/prot.21123](https://doi.org/10.1002/prot.21123) PMID: [WOS:000241247100017](https://pubmed.ncbi.nlm.nih.gov/WOS:000241247100017/).
55. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, et al. GLY-CAM06: A generalizable biomolecular force field. *Carbohydrates. J Comp Chem*. 2008; 29(4):622–55. doi: [10.1002/jcc.20820](https://doi.org/10.1002/jcc.20820) PMID: [17849372](https://pubmed.ncbi.nlm.nih.gov/17849372/)
56. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*. 1999; 314(1–2):141–51. doi: [10.1016/s0009-2614\(99\)01123-9](https://doi.org/10.1016/s0009-2614(99)01123-9) PMID: [WOS:000083955300022](https://pubmed.ncbi.nlm.nih.gov/WOS:000083955300022/).
57. Hansmann UHE, Okamoto Y. New Monte Carlo algorithms for protein folding. *Current Opinion in Structural Biology*. 1999; 9(2):177–83. doi: [10.1016/s0959-440x\(99\)80025-6](https://doi.org/10.1016/s0959-440x(99)80025-6) PMID: [WOS:000085219800005](https://pubmed.ncbi.nlm.nih.gov/WOS:000085219800005/).
58. Hukushima K, Nemoto K. Exchange Monte Carlo method and application to spin glass simulations. *Journal of the Physical Society of Japan*. 1996; 65(6):1604–8. doi: [10.1143/jpsj.65.1604](https://doi.org/10.1143/jpsj.65.1604) PMID: [WOS:A1996UV36200022](https://pubmed.ncbi.nlm.nih.gov/WOS:A1996UV36200022/).
59. Tian J, Garcia AE. Simulation Studies of Protein Folding/Unfolding Equilibrium under Polar and Nonpolar Confinement. *J Am Chem Soc*. 2011; 133(38):15157–64. doi: [10.1021/ja2054572](https://doi.org/10.1021/ja2054572) PMID: [21854029](https://pubmed.ncbi.nlm.nih.gov/21854029/)
60. Nymeyer H, Gnanakaran S, Garcia A. *Numerical Computer Methods, Part D; Methods in Enzymology, Vol 383*; Academic Press, Inc: San Diego, CA, 2004; pp119+.
61. Garcia AE, Herce H, Paschek D. Chapter 5 Simulations of Temperature and Pressure Unfolding of Peptides and Proteins with Replica Exchange Molecular Dynamics. In: David CS, editor. *Ann Rep Comp Chem. Volume 2*; Elsevier; 2006. p. 83–95.
62. Angel E, Garcia HH, Dietmar Paschek. Simulations of Temperature and Pressure Unfolding of Peptides and Proteins with Replica Exchange Molecular Dynamics. *Annu Rep Comp Chem*. 2006; 2:83–95.
63. Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem Phys Lett*. 1993; 215(6):617–21. [http://dx.doi.org/10.1016/0009-2614\(93\)89366-P](http://dx.doi.org/10.1016/0009-2614(93)89366-P).
64. Sanders RW, Derking R, Cupo A, Julien JP, Yasmeeen A, de Val N, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS pathogens*. 2013; 9(9):e1003618. doi: [10.1371/journal.ppat.1003618](https://doi.org/10.1371/journal.ppat.1003618) PMID: [24068931](https://pubmed.ncbi.nlm.nih.gov/24068931/); PubMed Central PMCID: [PMC3777863](https://pubmed.ncbi.nlm.nih.gov/PMC3777863/).
65. Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science*. 2013; 342(6165):1477–83. doi: [10.1126/science.1245625](https://doi.org/10.1126/science.1245625) PMID: [24179159](https://pubmed.ncbi.nlm.nih.gov/24179159/); PubMed Central PMCID: [PMC3886632](https://pubmed.ncbi.nlm.nih.gov/PMC3886632/).
66. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, et al. Comparative protein structure modeling using MODELLER. *Current protocols in protein science / editorial board, John E Coligan [et al]*. 2007;Chapter 2:Unit 2 9. doi: [10.1002/0471140864.ps0209s50](https://doi.org/10.1002/0471140864.ps0209s50) PMID: [18429317](https://pubmed.ncbi.nlm.nih.gov/18429317/).
67. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013; 29(7):845–54. doi: [10.1093/bioinformatics/btt055](https://doi.org/10.1093/bioinformatics/btt055) PMID: [23407358](https://pubmed.ncbi.nlm.nih.gov/23407358/); PubMed Central PMCID: [PMC3605599](https://pubmed.ncbi.nlm.nih.gov/PMC3605599/).
68. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*. 1997; 18(12):1463–72. doi: [10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
69. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*. 1993; 98(12):10089–92. doi: <http://dx.doi.org/10.1063/1.464397>
70. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*. 1984; 81(8):3684–90. doi: <http://dx.doi.org/10.1063/1.448118>
71. Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C, et al. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*. 2004; 14(12):1229–46. doi: [10.1093/glycob/cwh106](https://doi.org/10.1093/glycob/cwh106) PMID: [15175256](https://pubmed.ncbi.nlm.nih.gov/15175256/).
72. Travers SA. Conservation, Compensation, and Evolution of N-Linked Glycans in the HIV-1 Group M Subtypes and Circulating Recombinant Forms. *Isrn Aids*. 2012; 2012:823605. doi: [10.5402/2012/823605](https://doi.org/10.5402/2012/823605) PMID: [24052884](https://pubmed.ncbi.nlm.nih.gov/24052884/); PubMed Central PMCID: [PMC3765798](https://pubmed.ncbi.nlm.nih.gov/PMC3765798/).

73. Bar KJ, Tsao CY, Iyer SS, Decker JM, Yang Y, Bonsignori M, et al. Early low-titer neutralizing antibodies impede HIV-1 replication and select for virus escape. *PLoS pathogens*. 2012; 8(5):e1002721. doi: [10.1371/journal.ppat.1002721](https://doi.org/10.1371/journal.ppat.1002721) PMID: [22693447](https://pubmed.ncbi.nlm.nih.gov/22693447/); PubMed Central PMCID: PMC3364956.
74. Liao HX, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature*. 2013; 496(7446):469–76. doi: [10.1038/nature12053](https://doi.org/10.1038/nature12053) PMID: [23552890](https://pubmed.ncbi.nlm.nih.gov/23552890/); PubMed Central PMCID: PMC3637846.
75. Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, Gaschen B, et al. HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS pathogens*. 2009; 5(5):e1000414. doi: [10.1371/journal.ppat.1000414](https://doi.org/10.1371/journal.ppat.1000414) PMID: [19424423](https://pubmed.ncbi.nlm.nih.gov/19424423/); PubMed Central PMCID: PMC2671846.
76. Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, et al. Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell*. 1999; 96(5):667–76. doi: [10.1016/s0092-8674\(00\)80577-2](https://doi.org/10.1016/s0092-8674(00)80577-2) PMID: [10089882](https://pubmed.ncbi.nlm.nih.gov/10089882/).
77. Rizzuto CD, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA, et al. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science*. 1998; 280(5371):1949–53. doi: [10.1126/science.280.5371.1949](https://doi.org/10.1126/science.280.5371.1949) PMID: [9632396](https://pubmed.ncbi.nlm.nih.gov/9632396/).
78. Pinter A, Honnen WJ, D'Agostino P, Gorny MK, Zolla-Pazner S, Kayman SC. The C108g epitope in the V2 domain of gp120 functions as a potent neutralization target when introduced into envelope proteins derived from human immunodeficiency virus type 1 primary isolates. *Journal of virology*. 2005; 79(11):6909–17. doi: [10.1128/JVI.79.11.6909-6917.2005](https://doi.org/10.1128/JVI.79.11.6909-6917.2005) PMID: [15890930](https://pubmed.ncbi.nlm.nih.gov/15890930/); PubMed Central PMCID: PMC1112130.
79. McKeating JA, Shotton C, Cordell J, Graham S, Balfe P, Sullivan N, et al. Characterization of neutralizing monoclonal antibodies to linear and conformation-dependent epitopes within the first and second variable domains of human immunodeficiency virus type 1 gp120. *Journal of virology*. 1993; 67(8):4932–44. PMID: [7687306](https://pubmed.ncbi.nlm.nih.gov/7687306/); PubMed Central PMCID: PMC237881.
80. Krachmarov C, Lai Z, Honnen WJ, Salomon A, Gorny MK, Zolla-Pazner S, et al. Characterization of structural features and diversity of variable-region determinants of related quaternary epitopes recognized by human and rhesus macaque monoclonal antibodies possessing unusually potent neutralizing activities. *Journal of virology*. 2011; 85(20):10730–40. doi: [10.1128/JVI.00365-11](https://doi.org/10.1128/JVI.00365-11) PMID: [21835798](https://pubmed.ncbi.nlm.nih.gov/21835798/); PubMed Central PMCID: PMC3187505.
81. Pollara J, Bonsignori M, Moody MA, Liu P, Alam SM, Hwang KK, et al. HIV-1 vaccine-induced C1 and V2 Env-specific antibodies synergize for increased antiviral activities. *Journal of virology*. 2014; 88(14):7715–26. doi: [10.1128/JVI.00156-14](https://doi.org/10.1128/JVI.00156-14) PMID: [24807721](https://pubmed.ncbi.nlm.nih.gov/24807721/); PubMed Central PMCID: PMC4097802.
82. Gottardo R, Bailer RT, Korber BT, Gnanakaran S, Phillips J, Shen X, et al. Plasma IgG to linear epitopes in the V2 and V3 regions of HIV-1 gp120 correlate with a reduced risk of infection in the RV144 vaccine efficacy trial. *PLoS one*. 2013; 8(9):e75665. doi: [10.1371/journal.pone.0075665](https://doi.org/10.1371/journal.pone.0075665) PMID: [24086607](https://pubmed.ncbi.nlm.nih.gov/24086607/); PubMed Central PMCID: PMC3784573.
83. Garcia AE, Hummer G. Water penetration and escape in proteins. *Proteins-Structure Function and Genetics*. 2000; 38(3):261–72. doi: [10.1002/\(sici\)1097-0134\(20000215\)38:3<261::aid-prot3>3.0.co;2-q](https://doi.org/10.1002/(sici)1097-0134(20000215)38:3<261::aid-prot3>3.0.co;2-q) PMID: [WOS:000085160600003](https://pubmed.ncbi.nlm.nih.gov/WOS:000085160600003/).
84. Garcia AE, Sanbonmatsu KY. alpha-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99(5):2782–7. doi: [10.1073/pnas.042496899](https://doi.org/10.1073/pnas.042496899) PMID: [WOS:000174284600037](https://pubmed.ncbi.nlm.nih.gov/WOS:000174284600037/).
85. Wood NT, Fadda E, Davis R, Grant OC, Martin JC, Woods RJ, et al. The influence of N-linked glycans on the molecular dynamics of the HIV-1 gp120 V3 loop. *PLoS one*. 2013; 8(11):e80301. doi: [10.1371/journal.pone.0080301](https://doi.org/10.1371/journal.pone.0080301) PMID: [24303005](https://pubmed.ncbi.nlm.nih.gov/24303005/); PubMed Central PMCID: PMC3841175.
86. Holbourn KP, Acharya KR, Perbal B. The CCN family of proteins: structure-function relationships. *Trends Biochem Sci*. 2008; 33(10):461–73. Epub 2008/09/16. doi: [10.1016/j.tibs.2008.07.006](https://doi.org/10.1016/j.tibs.2008.07.006) PMID: [18789696](https://pubmed.ncbi.nlm.nih.gov/18789696/); PubMed Central PMCID: PMC2683937.
87. Almeida AM, Li R, Gellman SH. Parallel Beta-Sheet Secondary Structure Is Stabilized and Terminated by Interstrand Disulfide Cross-Linking. *Journal of the American Chemical Society*. 2011; 134(1):75–8. doi: [10.1021/ja208856c](https://doi.org/10.1021/ja208856c) PMID: [22148521](https://pubmed.ncbi.nlm.nih.gov/22148521/)
88. Sanders RW, Hsu ST, van Anken E, Liscajet IM, Dankers M, Bontjer I, et al. Evolution rescues folding of human immunodeficiency virus-1 envelope glycoprotein GP120 lacking a conserved disulfide bond. *Mol Biol Cell*. 2008; 19(11):4707–16. Epub 2008/08/30. doi: [10.1091/mbc.E08-07-0670](https://doi.org/10.1091/mbc.E08-07-0670) PMID: [18753405](https://pubmed.ncbi.nlm.nih.gov/18753405/); PubMed Central PMCID: PMC2575144.
89. Qin M, Zhang J, Wang W. Effects of disulfide bonds on folding behavior and mechanism of the beta-sheet protein tendamistat. *Biophys J*. 2006; 90(1):272–86. Epub 2005/10/11. doi: [10.1529/biophysj.105.063552](https://doi.org/10.1529/biophysj.105.063552) PMID: [16214873](https://pubmed.ncbi.nlm.nih.gov/16214873/); PubMed Central PMCID: PMC1367026.

90. Ogert RA, Lee MK, Ross W, Buckler-White A, Martin MA, Cho MW. N-linked glycosylation sites adjacent to and within the V1/V2 and the V3 loops of dualtropic human immunodeficiency virus type 1 isolate DH12 gp120 affect coreceptor usage and cellular tropism. *Journal of virology*. 2001; 75(13):5998–6006. doi: [10.1128/JVI.75.13.5998–6006.2001](https://doi.org/10.1128/JVI.75.13.5998-6006.2001) PMID: [11390601](https://pubmed.ncbi.nlm.nih.gov/11390601/); PubMed Central PMCID: PMC114315.
91. Go EP, Zhang Y, Menon S, Desaire H. Analysis of the disulfide bond arrangement of the HIV-1 envelope protein CON-S gp140 DeltaCFI shows variability in the V1 and V2 regions. *Journal of proteome research*. 2011; 10(2):578–91. doi: [10.1021/pr100764a](https://doi.org/10.1021/pr100764a) PMID: [21114338](https://pubmed.ncbi.nlm.nih.gov/21114338/); PubMed Central PMCID: PMC3075074.
92. Barbouche R, Miquelis R, Jones IM, Fenouillet E. Protein-disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion. *The Journal of biological chemistry*. 2003; 278(5):3131–6. doi: [10.1074/jbc.M205467200](https://doi.org/10.1074/jbc.M205467200) PMID: [12218052](https://pubmed.ncbi.nlm.nih.gov/12218052/).
93. Zolla-Pazner S, deCamp AC, Cardozo T, Karasavvas N, Gottardo R, Williams C, et al. Analysis of V2 antibody responses induced in vaccinees in the ALVAC/AIDSVAX HIV-1 vaccine efficacy trial. *PLoS one*. 2013; 8(1):e53629. Epub 2013/01/26. doi: [10.1371/journal.pone.0053629](https://doi.org/10.1371/journal.pone.0053629) PMID: [23349725](https://pubmed.ncbi.nlm.nih.gov/23349725/); PubMed Central PMCID: PMC3547933.
94. Julien JP, Lee JH, Cupo A, Murin CD, Derking R, Hoffenberg S, et al. Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110(11):4351–6. doi: [10.1073/pnas.1217537110](https://doi.org/10.1073/pnas.1217537110) PMID: [23426631](https://pubmed.ncbi.nlm.nih.gov/23426631/); PubMed Central PMCID: PMC3600498.