Delineation of antibody epitopes at the residue level is key to understanding antigen resistance mutations, designing epitope-specific probes for antibody isolation, and developing epitope-based vaccines. Ideally, epitope residues are determined in the context of the atomic-level structure of the antibody-antigen complex, though structure determination may in many cases be impractical. Here we describe an efficient computational method to predict antibody-specific HIV-1 envelope (Env) epitopes at the residue level, based on neutralization panels of diverse viral strains. The method primarily utilizes neutralization potency data over a set of diverse viral strains representing the antigen, and enhanced accuracy could be achieved by incorporating information from the unbound structure of the antigen. The method was evaluated on 19 HIV-1 Env antibodies with neutralization panels comprising 181 diverse viral strains and with available antibody-antigen complex structures. Prediction accuracy was shown to improve significantly over random selection, with an average of greater-than-8-fold enrichment of true positives at the 0.05 false-positive rate level. The method was used to prospectively predict epitope residues for two HIV-1 antibodies, 8ANC131 and 8ANC195, for which we experimentally validated the predictions. The method is inherently applicable to antigens that exhibit sequence diversity, and its accuracy was found to correlate inversely with sequence conservation of the epitope. Together the results show how knowledge inherent to a neutralization panel and unbound antigen structure can be utilized for residue-level prediction of antibody epitopes.

Broadly neutralizing antibodies (bNabs) against various antigens, such as HIV-1 envelope glycoprotein (Env) (1–4) and influenza virus hemagglutinin (HA) (5, 6), can have utility as therapeutics in the context of passive transfer (7) and as templates for the design of epitope-specific vaccines (8). The determination of the epitope targeted by an antibody of interest can assist in understanding virus resistance and escape mutations (9), offer clues for antibody affinity improvement (10, 11), and guide immunogen design for focusing the immune response toward neutralizing epitopes (12). Structure determination, by X-ray crystallography or nuclear magnetic resonance spectroscopy, can provide atomic-level resolution of epitopes and interactions in antibody-antigen complexes, but structures for many such complexes can be difficult or even infeasible to obtain (13). Cryo-electron microscopy could also be used to identify general epitope regions, but this method is typically associated with lower-resolution structures and in most cases cannot provide atomic-level information (14). A variety of other experimental methods can also be applied to epitope residue mapping, though these are typically laborious and can be limited by different factors, such as sensitivity to effects from distal residues not part of the direct antibody-antigen interactions (e.g., alanine scanning) or dependence on the presence of substantial antibody interactions in a sequentially continuous region of the antigen (e.g., pepscan) (15). The latter case is particularly limiting since most antibody epitopes are discontinuous (i.e., involving multiple sequentially noncontiguous regions) (13). In silico methods for epitope prediction are also available, but the majority focus on predicting protein residues that can be part of any epitope and are thus not antibody specific (16–19). Only a limited number of antibody-specific epitope prediction methods have been proposed thus far (20, 21). Computational docking can also be used to predict epitope residues by means of generating a structural model of the antibody-antigen complex. However, docking depends on the existence of separate antigen and antibody structures (or accurate structural models), and docking scoring functions are, in general, not optimal (22, 23) and in many cases unable to accurately predict the epitope of interest (24). Recently, a computational method was proposed for predicting the epitopes of query antibodies based on the similarity of their neutralization fingerprints to the fingerprints of antibodies with known epitopes (25). This method, however, does not provide residue-level information and is not applicable to antibodies that bind to novel epitopes. Another recent study utilized HIV-1 antibody neutralization panels to identify antigen residues functionally important for binding to specific antibodies (26). The method used in the study, however, aims at predicting a limited number of antigen residues of functional importance for a given antibody, rather than identifying the antibody epitope.

Here we present a computational method for antibody-specific prediction of epitope residues based on neutralization data from a
panel of diverse viral strains. The method is applicable to viruses that exhibit strain diversity, such as HIV-1 and influenza virus, and relies on the hypothesis that antibody neutralization potency can be expected to be affected more substantially by the choice of residue type at an epitope position but to a lesser extent by the choice of residue type at a nonepitope position. Conversely, residue positions whose variation is found to affect neutralization potency can be expected to be part of the respective antibody epitope (Fig. 1), although distal effects and residue covariation may also play a role. To explore this relationship between neutralization potency and antigen sequence variation, we applied a neutralization-based method to a set of 19 HIV-1 antibodies with published antibody-antigen complex structures and neutralization panels of diverse viral strains. The incorporation of data from diverse strains for the same virus can increase the likelihood of observing an association between changes in neutralization potency and variations in sequence. Variants of the method that further incorporated antigen structural information were also evaluated. Finally, we prospectively applied our method to the prediction of epitope residues for two HIV-1 antibodies and experimentally validated the computational predictions. Overall, our results indicate that a neutralization-based method in combination with structural information could be a useful tool for the accurate prediction of antibody epitopes of genetically diverse antigens.

### Materials and Methods

#### Data Set

A neutralization panel of 181 HIV-1 strains and antibody-antigen complex crystal structures for 19 HIV-1 broadly neutralizing antibodies downloaded from the Protein Data Bank (PDB) (27) were used to assess the performance of the computational methods (Table 1; see Table S1 in the supplemental material) (10, 25, 28–39). Antibody-mediated neutralization of 181 pseudotyped HIV-1 isolates was measured using TZM-bl cells and a luciferase reporter gene assay as described previously (40–42). Neutralization curves were fit by nonlinear regression using a 5-parameter hill slope equation as described previously (41) to determine the antibody concentrations required to inhibit infection by 50% (IC50).

Some of the neutralization data have been previously published (4, 25, 30, 33, 36, 43, 44). The gp160 amino acid sequences for the 181 HIV-1 strains were aligned using HIVAlign (http://www.hiv.lanl.gov/content/sequence/VIRALIGN/viralign.html). The antigen residues for the CD4-binding-site antibodies were defined as the residues in a ligand-free gp120 core structure (PDB ID 3TGT) (residues 44 to 124, 198 to 301, 324 to 396, and 408 to 492); for PG9 and PG16, the antigen residues were, respectively, defined as the residues in the V1/V2 loop present in the crystal structures (residues 126 to 196, except for the missing residues in the respective structures); and for MPER antibodies, the antigen residues were defined as residues in the MPER region (residues 659 to 685). Epitope residues were defined as antigen residues for which any heavy atom was within 5 Å of any heavy atom in the given antibody, as observed in the crystal structures.

#### Epitope Prediction Algorithm

Six different method variants were evaluated using the set of 19 HIV-1 antibodies. We used the true-positive (TP) rate and false-positive (FP) rate to evaluate the prediction accuracy, where the TP rate was defined as the ratio of the number of correctly predicted epitope residues to the number of real (as defined by the antibody-antigen structure [see above]) epitope residues, while the FP rate was defined as the ratio of the number of incorrectly predicted epitope residues to the number of real nonepitope residues. For methods where a set of optimal parameters (weight or radius [see below]) had to be determined, the prediction accuracy of the method was evaluated with leave-one-out cross-validation, where the optimal parameter value that maximized the average TP rate at the 0.05 FP rate level for 18 of the antibodies was used to evaluate the performance on the remaining antibody.

For each method variant, the TP rates for the predictions at FP rates of 0.05, 0.10, 0.15, 0.20, and 0.25 were evaluated. The parameters selected for

### Table 1: The 19 HIV-1 Antibodies used in the analysis

<table>
<thead>
<tr>
<th>Antibody</th>
<th>PDB ID</th>
<th>Site of recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRC01</td>
<td>3NBG</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC-PG04</td>
<td>3SE9</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>b12</td>
<td>2NY7</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>b13</td>
<td>3DX</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>F105</td>
<td>3HI1</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>NIH45-46</td>
<td>3U7Y</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>3BNC117</td>
<td>4JPV</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC-PG20</td>
<td>4LSUP</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC23</td>
<td>4H6R</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>CH103</td>
<td>4JAN</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC03</td>
<td>3E8</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC06</td>
<td>4JB9</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>VRC-CH31</td>
<td>4LSP</td>
<td>CD4-binding site</td>
</tr>
<tr>
<td>17b</td>
<td>2NXY</td>
<td>CD4i</td>
</tr>
<tr>
<td>10E8</td>
<td>4G6F</td>
<td>MPER</td>
</tr>
<tr>
<td>2F5</td>
<td>1TIJ</td>
<td>MPER</td>
</tr>
<tr>
<td>4E10</td>
<td>2FX7</td>
<td>MPER</td>
</tr>
<tr>
<td>PG9</td>
<td>3U4E</td>
<td>V1/V2</td>
</tr>
<tr>
<td>PG16</td>
<td>4DQ0</td>
<td>V1/V2</td>
</tr>
</tbody>
</table>
each method variant were the optimal parameters to maximize the average TP rate at the 0.05 FP rate level for the epitope prediction of the 19 HIV-1 antibodies. The six variants were as follows.

1. Normalized mutual information (nMI). Each antigen residue $i$ was ranked by the normalized mutual information between amino acid types and neutralization $IC_{50}$. The score for residue $i$ was then defined as

$$Score_i = nMI(X_i; Y) = \frac{MI(X_i; Y)}{H(Y)} = \frac{\sum_{x \in X, y \in Y} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right)}{-\sum_{x \in X} p(x) \log(p(x))}$$

where $X_i$ is a variable that encompasses the possible amino acid types at position $i$ (for a total of 21 possible values, i.e., the 20 natural amino acid types and a gap in the sequence alignment), while $Y$ is a binary variable defined by the respective neutralization $IC_{50}$ (for a total of 2 possible values, i.e., neutralizing [$IC_{50} < 50 \mu$g/ml] and nonneutralizing [$IC_{50} \geq 50 \mu$g/ml]). $MI(X_i; Y)$ is the standard mutual information defined between $X_i$ and $Y$. Since the mutual information is dependent on the Shannon entropy of each variable, here we normalized the mutual information by the Shannon entropy of the amino acid types at each position ($H(Y)$) (the Shannon entropy of the amino acid types at each position is a function of the selected input sequences); the mutual information was not normalized by $H(Y)$ since, for a given antibody, $H(Y)$ was constant at each position.

2. Normalized mutual information with window smoothing (nMI$_{\text{window}}$). The score for each residue was the normalized mutual information calculated as in method variant 1 averaged with the normalized mutual information of $n$ upstream and downstream (in sequence) residues:

$$Score_i = \frac{\sum_{j=i-n}^{i+n} nMI(X_j; Y)}{2n + 1}$$

Different $n$ values (1 to 5) were evaluated, and a value of 4 was selected from the leave-one-out analysis on the 19 HIV-1 antibodies.

3. Normalized mutual information with prior filtering of non-solvent-accessible residues (SA/nMI). From the antibody-free antigen structure (PDB ID 3TGT for the gp120 core structure, PDB ID 4G6F for MPER), the solvent-accessible area of each structure (PDB ID 3TGT for the gp120 core structure, PDB ID 4G6F for MPER) was determined in method variant 4:

$$Score_i = nMI(X_i; Y) + w_1PS_i + \sum_{j \in A \neq i \neq j \leq R \times \left( 1 - \frac{r_{ij}}{R} \right) \times \left( 0.8 + 0.2 \right) (nMI(X_i; Y) + w_2PS_i)$$

Different weights ($w_1 = 0.01$ to 0.20 in increments of 0.01) were evaluated, and a weight of $0.039 \pm 0.002$ (mean $\pm$ standard error) was selected from the leave-one-out analysis.

4. Prior filtering of nonaccessible residues and summation of the normalized mutual information of surrounding residues in space (SA/nMI$_{\text{proxsum}}$). The non-accessible residues were filtered out as in method variant 3. The score of each of the remaining residues was the normalized mutual information of the residue as calculated in method variant 1 plus a distance-dependent sum of the normalized mutual information of surrounding (in space) residues:

$$Score_i = nMI(X_i; Y) + \sum_{j \in A \neq i \neq j \leq R \times \left( 1 - \frac{r_{ij}}{R} \right) \times 0.8 + 0.2 \times nMI(X_j; Y)$$

where $A$ is the set containing all solvent-accessible residues of the antigen, $r_{ij}$ is the shortest distance between any side chain heavy atom of residues $i$ and $j$ (for glycine, $C_\alpha$ was used instead of a side chain atom), and $R$ is a normalization distance. A number of different values for $R$ (5.5 to 15.0 Å in increments of 0.5) were evaluated, and a value of 9 Å was selected from the leave-one-out analysis on the 19 HIV-1 antibodies. Such a simple scoring function that gradually decreases the weight of the score as the distance increases has also been used elsewhere (47, 48).

The last two method variants add an epitope propensity score (PS) to two of the previous methods (nMI$_{\text{window}}$ and SA/nMI$_{\text{proxsum}}$). An epitope log odds ratio for each residue type was proposed previously (see Table 1 in reference 16) as an estimate of the likelihood of each type of amino acid being an epitope residue. Since an ensemble of antigen sequences was used as input, the PS for each antigen position $i$ was defined as the average of epitope log odds ratios of position $i$ from all input sequences.

5. Normalized mutual information plus a weighted epitope propensity score with window smoothing (nMI$_{\text{window}}$/PS). The score composed of a window-smoothed normalized mutual information plus a weighted epitope propensity (we used $n = 4$ as determined in method variant 2):

$$Score_i = \frac{\sum_{j=i-n}^{i+n} (nMI(X_j; Y) + w_1PS_j)}{2n + 1}$$

Different weights ($w_1 = 0.01$ to 0.20 in increments of 0.01) were evaluated, and a weight of $0.039 \pm 0.002$ (mean $\pm$ standard error) was selected from the leave-one-out analysis.

6. Prior filtering of nonaccessible residues and summation of normalized mutual information and epitope propensity of residues in spatial proximity (SA/nMI$_{\text{proxsum}}$/PS). This was similar to method variant 4, but a weighted epitope propensity score (PS) was added to the normalized mutual information score (we used $R = 9.0$ as determined in method variant 4):

$$Score_i = nMI(X_i; Y) + w_1PS_i + \sum_{j \in A \neq i \neq j \leq R \times \left( 1 - \frac{r_{ij}}{R} \right) \times \left( 0.8 + 0.2 \right) (nMI(X_i; Y) + w_2PS_i)$$

Different weights ($w_1 = 0.01$ to 0.20 in increments of 0.01) were evaluated, and a weight of $0.031 \pm 0.005$ was selected from the leave-one-out analysis. All calculations were performed using a combination of Java and Perl code.

**Panel size and strain diversity analysis.** The effect of neutralization panel size on the prediction accuracy was evaluated by removing nonredundant strains from the full panel. In this analysis, strains with missing neutralization data for any antibody (see Table S1 in the supplemental material) were removed from the panel, resulting in a panel of 171 strains. The redundancy of the 171 strains was reduced by the “decrease redundancy” tool on the ExPASy server (49) based on the gp160 amino acid sequence, with maximum similarities set at 97.5%, 95%, 92.5%, 90%, 87.5%, 85%, 82.5%, and 80%, resulting in panels with sizes of 166, 165, 158, 152, 129, 96, 34, and 16 strains, respectively. The average performance of 100 randomly selected panels with the same panel sizes was also evaluated for comparison. Two clade B-specific panels were also created for 16- and 34-strain panel sizes by randomly selecting strains from the total of 39 clade B strains in the data set.

**Residue conservation analysis.** The conservation level of the epitopes was defined by the number of epitope residues that are fully conserved among all tested strains divided by the total number of epitope residues.

**Statistical analysis.** A two-tailed paired $t$ test was used to evaluate the significance of the differences in prediction accuracy (based on the TP rates at the 0.05 FP rate level) between the different method variants and random selection (TP rate equals FP rate) evaluated on the 19 HIV-1
antibodies; computed $P$ values were further adjusted using the false-discovery rate (FDR) method (50).

**Prospective epitope prediction for HIV-1 antibodies 8ANC131 and 8ANC195.** The SA/nMI$\text{proxsum}/PS$ method was applied to predict epitope residues for HIV-1 antibodies 8ANC131 and 8ANC195, using a neutralization panel of 181 HIV-1 strains (see Table S2 in the supplemental material) and the ligand-free gp120 core structure (PDB ID 3TGT) as structural input ($0.031$ was used for $w^2$ in the SA/nMI$\text{proxsum}/PS$ method). Binding of 8ANC131 to full-length gp120 (YU2) variants (wild type, R456S, K282N, I326T, and D78N) and binding of 8ANC195 to 2CC (a recombinant HXB2 gp120 core with stabilizing mutations W96C/V275C/I109C/Q428C/M95W/T257S/S375W/A433M) (38, 51) variants (wild type, deglycosylated wild type, N234S, T236K, and N276D) were determined using a previously described enzyme-linked immunosorbent assay (ELISA) method (4). Briefly, proteins were expressed by transient transfection in Freestyle 293-F cells (Invitrogen, Carlsbad, CA). gp120 proteins were purified either using DEAE-agarose resin (GE Healthcare, Little Chalfont, United Kingdom) followed by nickel chelation resin (Sigma, St. Louis, MO) or by affinity chromatography using monoclonal antibody 17b. Antibodies 8ANC131 and 8ANC195 were purified by protein A-affinity chromatography. Deglycosylated 2CC was prepared using methods described previously (38). Briefly, 2CC was expressed by transient transfection in HEK293S GnTi$^-$/H11002 cells, purified by 17b affinity chromatography, deglycosylated with endoglycosidase H, and purified with concanavalin A (ConA)-Sepharose (Sigma) and Superdex 200 (GE Healthcare) chromatography. Purified recombinant gp120 proteins were diluted in phosphate-buffered saline, and 200 ng per well was adsorbed onto Reacti-Bind 96-well plates (Pierce, Rockford, IL). Plates were blocked and incubated with a 5-fold dilution series of antibody with a starting concentration of 5 $\mu$g/ml. Horseradish peroxidase-conjugated goat anti-human IgG Fc antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) was used to detect bound antibody, and the color reaction was carried out using SureBlue 3,3',5,5'-tetramethylbenzidine (Kirkegaard & Perry Laboratories, Gaithersburg, MD). For 8ANC131, the epitope predictions were further compared to a recently determined crystal structure of the antibody-gp120 complex (52). For 8ANC195, the neutralization of the 3337.V2.C6 virus variants (wild type, N234S, T236K, and N276D) was determined as described in “Data set” above. Site-directed mutagenesis was performed by GeneImmune Biotechnology (Rockville, MD).

**Source code availability.** Source code is freely available to noncommercial users upon request.

**RESULTS**

**Epitope prediction for 19 HIV-1 antibodies.** To interrogate whether, for the purposes of epitope prediction, there is sufficient association between variation in strain sequence and changes in antibody neutralization potency on a panel of diverse viral strains, we applied a neutralization-based approach to a set of 19 HIV-1 Env antibodies for which antibody-antigen complex crystal structures were available (Table 1; see Table S1 in the supplemental material). Specifically, putative epitope residues were selected based on the normalized mutual information (nMI) scores between neutralization potency for a given antibody and the sequence variation at a given antigen residue for the given set of strains (see Materials and Methods). A higher nMI score for a given residue position was taken to indicate a possible epitope residue. The prediction accuracy using the nMI approach was significantly better than random ($P = 0.0005$) (see Table S3 in the supplemental material), with an average true-positive (TP) rate of 0.199 at the 0.05 false-positive (FP) rate level (Fig. 2; see Fig. S1 in the supplemental material).
Window smoothing of the nMI scores (nMI$_{\text{window}}$; see Materials and Methods) further improved the prediction (average TP rate of 0.264 at the 0.05 FP level).

Next, we evaluated the effects of augmenting the epitope prediction of the neutralization-based method with antigen structural information. Filtering out non-solvent-accessible residues based on the antigen structure before scoring the remaining residues by nMI (SA/nMI; see Materials and Methods) gave minimal improvement to the prediction accuracy (Fig. 2; see Fig. S1 in the supplemental material). However, averaging in the nMI scores of structure-proximal residues in conjunction with filtering out non-solvent-accessible residues (SA/nMI$_{\text{proxsum}}$; see Materials and Methods) significantly improved the performance compared to those of the nMI and nMI$_{\text{window}}$ methods ($P < 0.0005$ and $P < 0.005$, respectively) (see Table S3 in the supplemental material). Specifically, the average TP rate at the 0.05 FP rate level was 0.385, an almost 8-fold improvement over random and a 1.75-fold improvement over using only nMI as a predictor. For 17 out of the 19 HIV-1 antibodies, SA/nMI$_{\text{proxsum}}$ gave predictions equal to or better than those by nMI at the 0.05 FP rate level (see Table S4 in the supplemental material).

The effect of adding an epitope propensity score (see Materials and Methods) to two of the method variants, nMI$_{\text{window}}$ (nMI$_{\text{window}}$/PS) and SA/nMI$_{\text{proxsum}}$ (SA/nMI$_{\text{proxsum}}$/PS), was also evaluated (Fig. 2; see Fig. S1 in the supplemental material). Adding an epitope propensity score gave modest improvements to the two methods (TP rate increases of 0.264 to 0.283 and 0.385 to 0.403, respectively).

<table>
<thead>
<tr>
<th>Antibody</th>
<th>No. of:</th>
<th>TP rate at an FP rate of:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Antigen residues</td>
<td>Epitope residues</td>
</tr>
<tr>
<td>VRC01</td>
<td>342</td>
<td>31</td>
</tr>
<tr>
<td>VRC-PG04</td>
<td>342</td>
<td>27</td>
</tr>
<tr>
<td>b12</td>
<td>342</td>
<td>31</td>
</tr>
<tr>
<td>b13</td>
<td>342</td>
<td>21</td>
</tr>
<tr>
<td>F105</td>
<td>342</td>
<td>31</td>
</tr>
<tr>
<td>NIH45-46</td>
<td>342</td>
<td>42</td>
</tr>
<tr>
<td>3BNC117</td>
<td>342</td>
<td>32</td>
</tr>
<tr>
<td>VRC-PG20</td>
<td>342</td>
<td>42</td>
</tr>
<tr>
<td>VRC23</td>
<td>342</td>
<td>32</td>
</tr>
<tr>
<td>CH103</td>
<td>342</td>
<td>23</td>
</tr>
<tr>
<td>VRC03</td>
<td>342</td>
<td>40</td>
</tr>
<tr>
<td>VRC06</td>
<td>342</td>
<td>39</td>
</tr>
<tr>
<td>VRC-CH31</td>
<td>342</td>
<td>24</td>
</tr>
<tr>
<td>17b</td>
<td>342</td>
<td>16</td>
</tr>
<tr>
<td>10E8</td>
<td>27</td>
<td>12</td>
</tr>
<tr>
<td>2F5</td>
<td>27</td>
<td>10</td>
</tr>
<tr>
<td>4E10</td>
<td>27</td>
<td>11</td>
</tr>
<tr>
<td>PG9</td>
<td>52</td>
<td>14</td>
</tr>
<tr>
<td>PG16</td>
<td>41</td>
<td>11</td>
</tr>
</tbody>
</table>

FIG 3 Effect of the number of viral strains and strain diversity on epitope prediction. The average epitope prediction accuracy for 19 HIV-1 antibodies by the SA/nMI$_{\text{proxsum}}$/PS method for panels of different sizes and different degrees of diversity is shown as the average TP rate at the 0.05 FP rate level. The red line represents predictions using different smaller panels constructed by removing sequence-redundant strains with ExPASy. The blue line represents the average prediction accuracy of 100 randomly generated panels for the same panel sizes (error bars indicate standard errors). The orange line represents the prediction accuracy using strains selected only from clade B (16- and 34-strain panels only). The dendrograms display the distribution of the viral strains used in the nonredundant 16-strain panel selected from ExPASy (red) and the clade B 16-strain panel (orange) among all 171 strains (gray).
spectively, at the 0.05 FP rate level. The prediction accuracy of SA/nMI\textsubscript{proxsum}/PS was significantly better than that of nMI\textsubscript{window}/PS ($P < 0.01$), and this was the best method among the six variants in terms of average prediction accuracy, although the improvement over the prediction accuracy of SA/nMI\textsubscript{proxsum} was not significant.

Table 2 displays the prediction accuracy of the SA/nMI\textsubscript{proxsum}/PS method on the 19 HIV-1 antibodies. For 14 of the antibodies, the SA/nMI\textsubscript{proxsum}/PS method resulted in an enrichment of TP rate over random of greater than 5-fold at the 0.05 FP rate level; for 2 of the antibodies, the enrichment over random was between 1- and 5-fold; and for the remaining 3 antibodies, the SA/nMI\textsubscript{proxsum}/PS method performed worse than random. The 14 antibodies with greater-than-5-fold enrichment by the SA/nMI\textsubscript{proxsum}/PS method included all 3 of the MPER and all 2 of the V1/V2 antibodies as well as 9 of the 13 CD4-binding-site antibodies, but not the sole CD4i antibody.

**Prediction accuracy versus neutralization panel size and strain diversity.** Since larger panels of viral strains can be expected to contain more information with respect to sequence diversity, we investigated how the size of the neutralization panels may affect the prediction accuracy of our method. To that end, we generated subsets of the 181-strain panel at different cutoffs for sequence identity among included strains (see Materials and Methods). The effect of reducing the size of the neutralization panel by removing sequence-redundant strains is shown in Fig. 3 (red line). The average TP rate for the SA/nMI\textsubscript{proxsum}/PS method at the 0.05 FP rate level decreased 31.6% for the panel of 16 strains (corresponding to a sequence identity cutoff of 80.0%) and 20% for the panel of 34 strains (82.5% cutoff), while the prediction accuracy was similar (within 10%) to that for the full-size panel for the panels of 96 strains or larger (≥85% cutoff).

The effect of reducing the size of the neutralization panel by random selection was also examined (Fig. 3). When the prediction accuracies from 100 randomly selected panels were averaged for each panel size, the average TP rate for the SA/nMI\textsubscript{proxsum}/PS method at the 0.05 FP rate level decreased 49.33% for random panels of 16 strains, 38.13% for random panels of 34 strains, and 12.40% for random panels of 96 strains, while the prediction accuracy was similar (within 10%) to that for the full-size panel for random panels of 129 strains or larger. Smaller panels performed significantly worse than the full-size panel when comparing the average prediction accuracies of randomly selected panels (see Table S5 in the supplemental material).

The effect of strain diversity on prediction accuracy of the neutralization-based algorithm was seen by comparing the nonredundant panels with clade B-specific panels, which are composed of sequences with lower sequence variation. The prediction by the SA/nMI\textsubscript{proxsum}/PS method using the panel of 16 clade B strains was significantly worse than the prediction using the panel of 16 nonredundant strains ($P < 0.05$), with a 56.16% decrease in terms of the average TP rate at the 0.05 FP rate level (Fig. 3). This result indicates that prediction accuracy can be improved when using diverse strains compared to using strains with lower sequence variation.

**Prediction accuracy versus epitope sequence conservation.** The prediction accuracy of the SA/nMI\textsubscript{proxsum}/PS method described here was highly dependent on the variability of each residue among the different strains. Of note, the prediction accuracy was inversely correlated with the conservation level of the epitope residues (Fig. 4; see Table S6 in the supplemental material). The residues in the epitopes for antibodies b12 and b13, for example, had higher levels of sequence conservation and were less accurately predicted than other CD4-binding-site antibodies (see Fig. S2 in the supplemental material).

**Prospective prediction of epitope residues for HIV-1 antibody 8ANC131.** The SA/nMI\textsubscript{proxsum}/PS method was applied to prospectively predict the epitope for gp120 antibody 8ANC131, a CD4-binding-site antibody (2). The top 10 predicted residues (see Table S7 in the supplemental material) were located in three distinct regions of the gp120 structure (Fig. 5A and B): region 1 (residues 456, 466, 280, 96, 282, and 461), region 2 (residues 78, 79, and 80), and region 3 (residue 326). We selected residues from each region for mutation analysis (residues 456 and 282 from region 1, residue 78 from region 2, and residue 326 from region 3); for each position, the YU2 gp120 wild-type residue was mutated to R456S and K282N resulted in a substantial decrease in binding of 8ANC131 to YU2 gp120 (Fig. 5C; see Fig. S3 in the supplemental material).
material). These results and the fact that region 1 had the largest number of residues in the top 10 predictions suggested that the predicted residues from region 1 are likely part of the 8ANC131 epitope. As a further validation, the computational predictions were further compared to the actual epitope residues defined by a recently determined crystal structure of the antibody-gp120 complex (52). The SA/nMIproxsum/PS method demonstrated multiple fold increases of true positive rates at different FP rate levels compared to random selection (Fig. 5D). In particular, TP rates were improved by 6.98-fold at the 0.05 FP rate level compared to random selection. Among the top 10 predicted residues, five of the six region 1 residues (456, 280, 96, 282, and 461) but none of the region 2 and region 3 residues are part of the actual antibody epitope.

Prospective prediction of epitope residues for HIV-1 antibody 8ANC195. The SA/nMIproxsum/PS method was applied to prospectively predict the epitope for gp120 antibody 8ANC195, which was isolated using 2CC, a gp120-based probe stabilized in the CD4-bound conformation (38, 51). This antibody does not possess the typical characteristics of a CD4-binding-site antibody, and its precise epitope is currently unknown (2). The 10 top-scoring residues were focused on a patch of the gp120 surface encompassing a loop preceding the α1 helix, loops A and D, and the β23/V5/β24 strand/loop/strand region (Fig. 6; see Table S8 in the supplemental material). Three of the top 10 residues (234, 236, and 276) affected the presence of an N-linked glycosylation sequon (for glycans at positions 234 and 276), and since enzymatic deglycosylation of 2CC knocked out its binding to 8ANC195 (Fig. 6C), we selected these three residues for mutation analysis: for each position, the 2CC residue was mutated to an amino acid type preferentially found in 8ANC195-resistant strains (53). The selected mutations (N234S, T236K, and N276D) each removed a single N-linked glycosylation site (Fig. 6C; see Fig. S4 in the supplemental material). Each of the three mutations knocked out 8ANC195 binding (Fig. 6C), in agreement with the computational predictions, suggesting that 8ANC195 is a glycan-reactive antibody targeting a novel epitope on gp120. Additionally, we tested the effects of N234S, T236K, and N276D on neutralization by 8ANC195 by introducing each of these three mutations into 3337.V2.C6, the strain most sensitive to 8ANC195 in the 181-strain neutralization panel. All three mutant viruses were resistant (IC50 of >50 μg/ml) to neutralization by 8ANC195 (Fig. 6D; see Fig. S5 in the supplemental material), further confirming that the N-linked glycans at positions 234 and 276 are important for viral recognition by this antibody.

DISCUSSION

Most of the currently available computational methods for epitope prediction are based on analyzing sequence or biochemical properties of the general protein surfaces and are not antibody specific. Based on the hypothesis that sequence variation of epitope residues likely affects antibody neutralization, the neutralization-based method described here provides an efficient way to identify epitope residues in cases where antibody neutralization panels of diverse viral strains are available. Neutralization information has been used elsewhere to classify broadly neutralizing antibodies and to understand the antigenic signature associated with these antibodies (25, 54). We have demonstrated that this method can predict the epitopes of HIV-1 antibodies with high accuracy. Experimental validation of the prospective predictions for antibody 8ANC195 revealed dependence of antigen recogni-
tion by this antibody on the presence of N-linked glycans at residues 234 and 276, indicating that these residues might indeed be part of the antibody epitope. To our knowledge, there is no other known HIV-1 broadly neutralizing antibody reported to target this epitope, making 8ANC195 an interesting target for further structural characterization. Similar findings with respect to glycan reactivity of antibody 8ANC195 were also reported recently (26).

Although the availability of neutralization panels over diverse viral strains is a prerequisite for neutralization-based methods, such panels are often performed as one of the first steps in the evaluation of an antibody of interest, in order to define its neutralization breadth and potency. Furthermore, although some variability in the measurements of neutralization potency can be observed upon multiple repeats of the experiment for the same antibody against the same virus, such variability is expected to have minimal effects on the prediction accuracy, since neutralization potency was treated as a binary value. Indeed, this was the case when our method was applied to predict the epitopes of antibodies VRC01 and VRC-PG04 based on two neutralization panels with an overlapping set of strains (see Fig. S6 and Table S9 in the supplemental material). We further note that, in addition to neutralization panels, the computational framework described here can also be used with other experimental measurements, such as binding affinity panels.

A number of factors that could affect the performance of the methods were explored further. Although the neutralization panel used in this study contained 181 viral strains, a panel of around 100 non-sequence-redundant viral strains could be sufficient to achieve similar prediction accuracy, and panels a third that size did not perform significantly worse. Hence, while smaller panels may not fully capture the true diversity of the viral protein, they nevertheless seem to possess sufficient information for the prediction of epitope residues. Since the variability of the residue types at each position is dependent on the input sequences, the mutual information scores were normalized with the Shannon entropy of the residue type at each position (see Materials and Methods) instead of using mutual information directly as the scoring function. The prediction accuracy using this normalized mutual information was slightly better than that of its unnormalized counterpart in the low-FP-rate region (see Fig. S7 in the supplemental material).

Although alternative cutoffs may also be used, we note that, for the

**FIG 6** Epitope prediction for HIV-1 antibody 8ANC195 using the SA/nMIproxsum/PS method. (A) The top 10 predicted residues (yellow spheres) mapped on a gp120 core crystal structure (red) (PDB ID 3TGT), fitted in an unliganded HIV-1 Env trimer cryo-EM map (EMD accession code 5019, pink). (B) Outline of the top 10 predicted epitope residues (yellow) for 8ANC195 shown on a transparent surface and cartoon representation of gp120 (red). Two N-linked glycosylation sites, at residues 234 and 276, were among the selected residues; the N-acetylglycosamine moieties linked directly to the two residues are shown in green sticks to highlight their location. (C) ELISA binding of 8ANC195 to the stabilized gp120 core 2CC, 2CC N234S, 2CC T236K, 2CC N276D, and deglycosylated 2CC (2CC degly). 8ANC195 showed good binding to the wild type but did not bind deglycosylated 2CC. The mutations N234S, T236K, and N276D, each of which resulted in the removal of an N-linked glycan, knocked out binding of 2CC to 8ANC195. (D) 8ANC195 neutralization of wild-type, N234S, T236K, and N276D 3337.V2.C6. Wild-type 3337.V2.C6 was potently neutralized by 8ANC195, while all of the N234S, T236K, and N276D mutant viruses were resistant to neutralization by that antibody.
current set of antibodies, the prediction accuracy with an IC_{50} cutoff at 50 μg/ml was higher than that using other IC_{50} cutoffs with lower values (see Fig. S8 in the supplemental material). A potential explanation for this observation may be that a cutoff at the detection limit of the assay (50 μg/ml) may give a more robust classification of neutralized versus nonneutralized strains for each antibody, compared to using a lower IC_{50} as a cutoff. Nevertheless, we note that the cutoff must be lowered for pan-neutralizing antibodies (for which the neutralization IC_{50} is less than 50 μg/ml for all viral strains), since otherwise the mutual information scores between residue types at each position and neutralization potency will always be zero. We also note that the computation of mutual information scores dictates the use of discrete classes, since classes for sequence variation were formed based on amino acid types. A binary variable was selected for the neutralization potency, although discretization into a larger number of classes could also be performed; however, using a ternary variable reduced the prediction accuracy for the current set of antibodies (see Fig. S9 in the supplemental material). Finally, while an inverse trend between the number of epitope residues and the prediction accuracy was observed, the correlation was not statistically significant (see Fig. S10 in the supplemental material). It is interesting to speculate that a possible explanation for such a trend may be that, in general, there exists only a small number of epitope residues whose variation has a more substantial effect on antibody neutralization and that can thus be more easily picked up by the neutralization-based epitope prediction method, and this may have a more pronounced effect within the context of a smaller versus a larger epitope. Nevertheless, Fig. S10 in the supplemental material shows that, with few exceptions, the TP rates for most antibodies fall within the range of approximately 0.4 to 0.6, indicating that epitope size is not a major determinant for prediction accuracy.

Incorporating information from unbound antigen structures, which are available for many antigens (such as for different regions of HIV-1 Env and for influenza virus HA), was found to enhance the neutralization-based epitope prediction. In addition to removing surface-inaccessible residues, we conjectured that residues with lower nMI scores could nevertheless be a part of the epitope if they are surface proximal to patches that contained residues with higher nMI scores, since antibody epitopes are generally formed by one or more contiguous surface patches. Indeed, scoring each residue not by its nMI score alone but by adding a weighted sum of the nMI scores for structure-proximal residues improved the predictions in most of the test cases. On the other hand, if a residue with a high nMI score is in an isolated patch surrounded only by residues with low nMI scores, it may actually not be part of the epitope and its score may be due to some distal effects responsible for antibody resistance or to an artifact of, e.g., sequence covariance with epitope residues. For the method variants without antigen structure incorporation, averaging in the nMI scores of sequence-neighboring residues achieved similar but lesser effects. Adding the epitope propensity score developed by Haste Andersen et al. (16), which reflects the tendency of a particular amino acid type to be an epitope residue, into the scoring function also slightly improved the prediction accuracy. The score is statistically derived and scores amino acid types based on the frequency with which they are observed in epitope regions compared to non-epitope regions (e.g., asparagine, arginine, proline, and lysine have higher propensity scores since they are over-represented in epitope regions, while cysteine, alanine, leucine, valine, and phenylalanine have lower scores since they are under-represented in epitope regions). Hence, the inclusion of the pre-computed propensity scores does not substantially increase computational complexity. Since the SA/nMIproxsum/PS method had the highest average prediction accuracy at the 0.05 FP rate level for the 19 HIV-1 antibodies in our data set, we selected it as the default method variant. However, it should be noted that, statistically, this method does not perform significantly better than SA/nMIproxsum (see Table S3 in the supplemental material), indicating that the prediction power of SA/nMIproxsum/PS still derives mainly from summing the mutual information scores of spatially proximal residues. Further improvements in prediction accuracy may be achieved by incorporating additional sequence and structural information from mining known antibody-antigen interfaces.

The neutralization-based methods described here are limited to observing the direct effects on neutralization of sequence variation within an antibody epitope. For example, the methods failed to predict epitope residues for the CD4i antibody 17b, possibly due to the dependence of neutralization by 17b on the requirement for conformational change of the viral protein (55), a secondary effect that is not taken into account by the neutralization-based methods and that can misguide the epitope predictions. Another major caveat of using mutual information between residue types and neutralization potency to predict epitope residues is the pronounced inverse dependence of prediction accuracy on the conservation level of epitope residues. For example, residue D368 of HIV-1 gp120, which is important for the binding of various VRC01 class antibodies, could not be predicted as an epitope residue by calculating the mutual information alone, since that residue was conserved among all viral strains in our data set. Thus, the neutralization-based epitope prediction methods described here are more applicable for antigens (and epitopes) showing larger diversity.

We further note that approximately half of the residues incorrectly predicted (FP rate = 0.05) by the SA/nMIproxsum/PS method to be part of the epitope for the set of 19 antibodies were found within 5 Å within an epitope residue, and ~65% were within 10 Å (see Fig. S11 in the supplemental material). This observation could indicate that some of these false-positive residues may nevertheless be important for interactions with the given antibody, either through direct interactions explained by strain-to-strain epitope variation, or in a supporting role for other neighboring epitope residues. Indeed, although we evaluated the prediction accuracy of the algorithm relative to antibody-antigen complex structures, the nature of the neutralization-based methods should also allow the identification of residues that are not necessarily part of, or even in proximity to, the epitope defined by the antibody-antigen complex structures but that may still, through distal effects, be of functional importance for antibody recognition.

Despite their limitations, the neutralization-based methods generated promising results, with significant enrichment and multiple-fold improvement over random epitope selection for the majority of the tested antibodies. It would also be of utility if these methods could be associated with a measure of the reliability for the epitope predictions for a given antibody. To that end, the neutralization-based methods can be augmented with the epitope-specific antibody clustering method described previously (25), which is applicable to both conserved and diverse antigens but is dependent on the existence of a reference set of antibodies with known epitopes and cannot make predictions at the residue level.
level since no antigen sequence or structural information is incorporated. Where applicable, the antibody clustering method can be applied to predict the general epitope region for a query antibody, possibly also providing an indication of the level of conservation of the epitope. In turn, this can serve as a measure of the reliability of the neutralization-based residue-level epitope predictions. Alternatively, in the general case of a novel epitope where prior knowledge of the sequence conservation level of the epitope residues is not available, it may be beneficial to examine other factors that may affect the reliability of the predictions. Hence, we determined that there is a significant relationship between residue score distribution and prediction accuracy for a given epitope: the larger the ratio of the median of the top residue scores to the median of all residue scores, the higher the prediction accuracy (see Fig. S12 in the supplemental material). Further work will be necessary to evaluate variants of the residue score distribution and other factors as measures of reliability for the epitope predictions for a given antibody.

The epitope prediction methods described here show that it is possible to obtain detailed structural information on HIV-1 Env antibody recognition through analysis of functional readouts such as antibody-antigen neutralization data. It is interesting to speculate that the inverse information flow could also be exploited: for example, given a target epitope region for which currently no known neutralizing antibodies exist, it may be possible to devise a (hypothetical) antibody neutralization fingerprint that matches that epitope. Such epitope-specific neutralization fingerprints (25) can then be used to analyze donor sera, as a way to identify new broad and potent antibodies targeting the epitopes of interest. Similarly, one can start by first devising hypothetical antibody neutralization fingerprints, followed by prediction of the epitopes that may match such fingerprints. It is likely that not all generated fingerprints could represent real antibodies, so a search over the neutralization fingerprint space may be necessary. This process could give insight into the susceptibility of a specific virus to broad neutralization and can help identify novel broadly neutralizing epitopes as new targets for immunogen design.

The identification of neutralizing antibody epitopes plays an important role in understanding how these antibodies bind to their antigen and is a key step to designing epitope-specific immunogens. Since structure determination of antibody-antigen complexes is not always feasible, the wealth of information hidden in antibody neutralization panels can be used to efficiently extract antibody epitope information at the residue level and can be augmented with other experimental methods (such as cryo-electron microscopy and alanine or peptide scanning) for confirmation or more complete delineation of the epitopes for antibodies of interest. The identification of a novel epitope for a highly potent and broad neutralizing antibody can also be used as a starting point for designing immunogens to elicit similar antibodies. The computational framework described here therefore provides an efficient way to identify epitope residues on antigens characterized by sequence diversity and could further improve the understanding of strain-specific antibody resistance and facilitate the development of epitope-based vaccines.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the Vaccine Research Center, NIAID, the International AIDS Vaccine Initiative’s Neutralizing Antibody Consortium, the Office of AIDS Research, NIH, and the following grants to M.C.N.: Bill and Melinda Gates Foundation grant 38619s and NIH grants AI 100663-01 and AI 100148-01. M.C.N. is an HHMI investigator.

We thank J. Stucker for assistance with graphics and the Structural Biology Section, Structural Bioinformatics Core, Humoral Immunology Section, and Humoral Immunology Core at the NIH Vaccine Research Center for helpful discussions or comments on the manuscript. We thank J. Baalwa, D. Ellenberger, D. Gabuzda, F. Gao, B. Hahn, K. Hong, J. Kim, F. McCutchan, D. Montefiori, L. Morris, J. Overbaugh, E. Sanders-Buell, G. Shaw, R. Swannstrom, M. Thomson, S. Tovanabutra, C. Williamson, and L. Zhang for contributing the HIV-1 envelope plasmids used in our neutralization panels.

REFERENCES


Residue-Level Antibody Epitope Prediction


