Mapping mutational effects along the evolutionary landscape of HIV envelope

Hugh K. Haddox[1,2,†], **Adam S. Dingens**[1,2,†], **Sarah K. Hilton**[1,3], **Julie Overbaugh**[4], **Jesse D. Bloom**[1,3,*]

*For correspondence:
jbloom@fredhutch.org (JDB)

[†]These authors contributed equally to this work

[1]Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA; [2]Molecular and Cellular Biology PhD program, University of Washington, Seattle, WA; [3]Department of Genome Sciences, University of Washington, Seattle, WA; [4]Human Biology Division and Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, WA

Abstract   The immediate evolutionary space accessible to HIV is largely determined by how single amino-acid mutations affect fitness. These mutational effects can shift as the virus evolves. However, the prevalence of such shifts in mutational effects remains unclear. Here we quantify the effects on viral growth of all amino-acid mutations to two HIV envelope (Env) proteins that differ at >100 residues. Most mutations similarly affect both Envs, but the amino-acid preferences of a minority of sites have clearly shifted. These shifted sites usually prefer a specific amino acid in one Env, but tolerate many amino acids in the other. Surprisingly, shifts are only slightly enriched at sites that have substituted between the Envs—and many occur at residues that do not even contact substitutions. Therefore, long-range epistasis can unpredictably shift Env's mutational tolerance during HIV evolution, although the amino-acid preferences of most sites are conserved between moderately diverged viral strains.

Introduction

HIV's envelope (Env) protein evolves very rapidly. The major group of HIV-1 that is responsible for the current pandemic originated from a virus that entered the human population ~100 years ago (*Sharp and Hahn, 2011*; *Worobey et al., 2008*; *Faria et al., 2014*). The descendants of this virus have evolved so rapidly that their Envs now have as little as 65% protein identity (*Lynch et al., 2009*). For comparison, protein orthologs shared between humans and mice have only diverged to a median identity of 78% over 90 million years (*Waterston et al., 2002*; *Hedges et al., 2006*).

Env's rapid evolution has dire consequences for anti-HIV immunity, since it erodes the efficacy of most neutralizing antibodies (*Albert et al., 1990*; *Wei et al., 2003*; *Richman et al., 2003*; *Burton et al., 2005*). Because of this public-health importance, numerous studies have experimentally characterized aspects of the "evolutionary landscape" that Env traverses. The immediate evolutionary space accessible to any given Env is largely defined by the effects on viral fitness of all single amino-acid mutations to Env. Most mutational studies have measured how just a small number of these mutations affect viral growth in cell culture, although it has recently become possible to use deep mutational scanning to measure the effects of many (*Al-Mawsawi et al., 2014*; *Duenas-Decamp et al., 2016*) or even all (*Haddox et al., 2016*) single amino-acid mutations mutations to an Env variant.

But interpreting these studies in the context of Env evolution requires addressing a fundamental question: How informative are mutational studies of a single protein variant about constraints on long-term evolution? During protein evolution, substitutions at one site can change the effect of mutations at other sites (*Natarajan et al., 2013*; *Gong et al., 2013*; *Harms and Thornton, 2014*; *Podgornaia and Laub, 2015*; *Starr and Thornton, 2016*; *Klink and Bazykin, 2017*). We will follow the nomenclature of *Pollock et al.* (*2012*) to refer to these changes in mutational effects as *shifts* in a site's amino-acid preferences. Such shifts can accumulate as substitutions become entrenched via epistatic interactions with subsequent changes (*Starr et al., 2017*; *Pollock et al., 2012*; *Shah et al., 2015*; *Bazykin, 2015*)—although the magnitude of these shifts is usually limited (*Doud et al., 2015*; *Chan et al., 2017*; *Ashenberg et al., 2013*; *Risso et al., 2014*).

Given that the Envs of circulating HIV strains represent a vast collection of homologs that often differ at >100 residues, shifts in amino-acid preferences could make the outcome of any study highly dependent on the Env used. Indeed, epistasis among a few combinations of Env mutations has been experimentally demonstrated (*da Silva et al., 2010*), and epistatic fitness landscapes have been computationally inferred for a variety of HIV proteins (*Kouyos et al., 2012*; *Ferguson et al., 2013*; *Mann et al., 2014*; *Barton et al., 2015*) including Env (*Louie et al., 2018*). However, the only protein-wide experimental studies of how amino-acid preferences shift during evolution have examined proteins that are structurally far simpler than Env, which forms a large heavily glycosylated heterotrimeric complex that transitions through multiple conformational states (*Munro et al., 2014*; *Ozorowski et al., 2017*).

Here we use an improved version of a previously described deep mutational scanning strategy (*Haddox et al., 2016*) to measure the effects on viral growth of all single amino-acid mutations to two transmitted-founder virus Envs that differ by >100 mutations. We compare these complete maps of mutational effects to identify sites that have shifted in their amino-acid preferences between the Envs. Most sites show no detectable shifts, but 30 sites have clearly shifted preferences. These shifted sites usually prefer a specific amino acid in one Env but have shifted to tolerate many amino acids in the other Env. The shifted sites cluster in structure, but are often distant from any amino-acid substitutions that distinguish the two Envs, demonstrating the action of long-range epistasis. By aggregating our measurements for both Envs, we identify sites that evolve faster or slower in nature than expected given the functional constraints measured in the lab, probably due to pressure for immune evasion. Overall, our work provides complete across-strain maps of mutational effects that inform analyses of Env's evolution and function.

Results

Two Envs from clade A transmitted-founder viruses

The viruses most relevant to HIV's long-term evolution are those which are transmitted from human-to-human. However, the only prior work that has measured how all Env amino-acid mutations affect HIV growth is a study by some of us (*Haddox et al., 2016*) that used a late-stage lab-passaged CXCR4-tropic virus (LAI; *Peden et al., 1991*). The properties of Env can vary substantially between such late-stage viruses and the transmitted-founder viruses relevant to HIV's long-term evolution (*Sagar et al., 2006*; *Wilen et al., 2011*; *Parrish et al., 2013*; *Ronen et al., 2015*).

We therefore selected Envs from two transmitted-founder viruses, BG505.W6M.C2.T332N and BF520.W14M.C2 (hereafter referred to as BG505 and BF520), that were isolated from HIV-infected infants shortly after mother-to-child transmission (*Nduati et al., 2000*; *Wu et al., 2006*; *Goo et al., 2014*). The BG505 Env has been extensively studied from a structural standpoint (*Julien et al., 2013*; *Lyumkis et al., 2013*; *Pancera et al., 2014*; *Huang et al., 2014*; *Sanders et al., 2015*; *Stewart-Jones et al., 2016*; *Gristick et al., 2016*), and variants of this Env are being tested as a vaccine immunogens (*Sanders et al., 2013*, *2015*; *de Taeye et al., 2015*). We used the T332N variant of BG505 Env because it has a common glycosylation site that is targeted by many anti-HIV antibodies (*Sanders et al., 2013*). The BF520 Env was isolated from an infant who developed an early broad anti-HIV

Figure 1. Phylogenetic tree showing the relationship of BG505 and BF520 to other clade A Envs. The tree shows the 69 Envs in the alignment in *Figure 1*-source data 1, which is a subsample of clade A sequences from the group M alignment in the Los Alamos HIV sequence database (http://www.hiv.lanl.gov). Sites not mutagenized in our experiments (the signal peptide and cytoplasmic tail) or that are poorly alignable were masked as indicated in *Figure 1*-source data 2, leaving 616 alignable sites. The pairwise identity of BG505 and BF520 to other sequences at alignable sites is in *Figure 1–Figure Supplement 1*. The tree topology was inferred using `RAxML` (*Stamatakis, 2014*) under the GTRCAT model of nucleotide substitution, and branch lengths were optimized under the M0 Goldman-Yang model (*Yang et al., 2000*) using `phydms` (*Hilton et al., 2017*).

Figure 1–Figure supplement 1. Pairwise identity of all Env sequences to BG505 and BF520.

Figure 1–source data 1. The alignment of clade A *env* coding sequences is in `cladeA_alignment.fasta`.

Figure 1–source data 2. The 240 Env sites masked in all phylogenetic analyses because they were not mutagenized in our experiments or are poorly alignable are listed in `alignment_mask.csv`.

---

antibody response (*Goo et al., 2014*; *Simonich et al., 2016*). We have previously created comprehensive codon-mutant libraries of the BF520 Env and used them to map HIV antibody escape (*Dingens et al., 2017*), but these BF520 libraries have not been characterized with respect to how mutations affect viral growth.

Both BG505 and BF520 are from clade A of the major (M) group of HIV-1. *Figure 1* shows the phylogenetic relationship among these two Envs and other clade A sequences. BG505 and BF520 are identical at 721 of the 836 pairwise-alignable protein sites (86.2% identity). However, in our experiments we mutagenized only the ectodomain and transmembrane domain of Env, and excluded the signal peptide and cytoplasmic tail. The reason is that we measure how Env mutations affect viral growth, which is influenced both by the functionality of Env protein molecules and their expression level. Mutations in the signal peptide and cytoplasmic tail commonly affect Env expression level (*Chakrabarti et al., 1989*; *Yuste et al., 2004*; *Li et al., 1994*), so we excluded these regions with the goal of reducing the degree to which we simply identified mutations that affected Env expression. In the ectodomain and transmembrane domains of Env, BG505 and BF520 are identical at 549 of the 616 sites (89.1% identity) that are alignable across clade A Envs (*Figure 1*-source data 1, *Figure 1*-source data 2). The divergence between BG505 and BF520 therefore offers ample opportunity to investigate mutational shifts during Env evolution.

Deep mutational scanning of each Env

We have previously described a deep mutational scanning strategy for measuring how all amino-acid mutations to Env affect HIV growth in cell culture, and applied this strategy to the late-stage lab-adapted LAI strain (*Haddox et al., 2016*). Here we made several modifications to this earlier strategy to apply it to transmitted-founder Envs and to reduce the experimental noise. This last consideration is especially important when comparing Envs, since it is only possible to reliably detect differences that exceed the magnitude of the experimental noise. Our modified deep mutational scanning strategy is in *Figure 2*A. This approach had the following substantive changes: instead of SupT1 cells, we used SupT1.CCR5 cells (SupT1 cells that express CCR5 in addition to CXCR4 (*Boyd et al., 2015*)) to support growth of viruses with transmitted-founder, CCR5-tropic Envs; we used more virions for the first passage ($\geq 3 \times 10^6$ versus 5×10^5 infectious units per library) to avoid bottlenecking library diversity; and rather than performing a full second passage we just did a short high-MOI infection to enable recovery of *env* genes from infectious virions without bottlenecking (*Figure 2*A). We performed this deep mutational scanning in full biological triplicate for both BG505 and BF520 (*Figure 2*B). Our libraries encompassed all codon mutations to all sites in Env except for the signal peptide and cytoplasmic tail.

The deep mutational scanning effectively selected for functional Envs as evidenced by strong purifying selection against stop codons. *Figure 3*A shows the average frequency of mutations across Env in the plasmid mutant libraries, the mutant viruses, and wildtype controls as determined from the deep sequencing. The mutant viruses show clear selection against stop codons and

Figure 2. Deep mutational scanning workflow. **(A)** We made libraries of proviral HIV plasmids with random codon-level mutations in the *env* gene. The number of mutations per gene approximately followed a Poisson distribution with a mean between 1 and 1.5 (***Figure 2–Figure Supplement 1***). We transfected the plasmids into 293T cells to generate mutant viruses, which lack a genotype-phenotype link since cells are multiply transfected. To establish a genotype-phenotype link and select for Env variants that support HIV growth, we passaged the libraries in SupT1.CCR5 cells for four days at a low multiplicity of infection (MOI) of 0.01. To isolate the *env* genes from only viruses that encoded a functional Env protein, we infected the passaged libraries into SupT1.CCR5 cells at high MOI and harvested reverse-transcribed non-integrated viral DNA after 12 hours. We then deep sequenced the *env* genes from these final samples as well as the initial plasmid library, using molecular barcoding to reduce sequencing errors. We also deep sequenced identically handled wildtype controls to estimate error rates. Using these sequencing data, we estimated the preference for each of the 20 amino acids at each site in Env. These data are represented in logo plots, with the height of each letter proportional to that site's preference for that amino acid. **(B)** We conducted this experiment in full biological triplicate for both BG505 and BF520, beginning each replicate with independent creation of the plasmid mutant library. These replicates therefore account for *all* sources of noise and error in the experiments.

Figure 2–Figure supplement 1. Sanger sequencing of selected clones from the mutant plasmid libraries.

Figure 3. The deep mutational scanning selects for functional Envs and yields measurements that are well correlated among replicates. **(A)** The average per-codon mutation frequency when sequencing plasmids encoding wildtype Env (*DNA*), plasmid mutant libraries (*mutDNA*), mutant viruses after the final infection (*mutvirus*), and virus generated from wildtype plasmids (*virus*). Mutations are categorized as nonsynonymous, synonymous, or stop codon. The *DNA* samples show that sequencing errors are rare, and the *virus* samples show that viral-replication errors are well below the frequency of mutations in the *mutDNA* samples. Comparing the *mutvirus* to *mutDNA* shows clear purifying selection against stop codons and some nonsynonymous mutations, particularly after subtracting the background error rates given by the *virus* and *DNA* samples (***Figure 3**-source data* 2). More extensive plots from the analysis of the deep sequencing data are in Supplementary files 1 and 2. **(B)** Correlations between replicates in the measured preferences of each site in Env for all 20 amino acids. Blue indicates replicate measurements on BF520, red indicates replicate measurements on BG505, and gray indicates across-Env measurements of BF520 versus BG505. *R* is the Pearson correlation coefficient. The numerical values for the preferences are in ***Figure 3**-source data* 1. ***Figure 3–Figure Supplement 1*** shows the correlations using contour rather than scatter plots.

Figure 3–Figure supplement 1. Correlations plotted on a contour rather than a scatter plot.

Figure 3–source data 1. Preferences for each replicate and averages are in `all_prefs_unscaled.zip`.

Figure 3–source data 2. Average frequencies of nonsynonymous, synonymous, and stop-codon mutations as plotted in ***Figure 3*** are in `avgmutfreqs.csv`. There is only one *DNA* sample for BF520 which is listed three times with each BF520 replicate. We calculate the error-corrected *pre*-selection mutation frequency as the *mutDNA* minus *DNA*, and the error-corrected *post*-selection mutation frequency as the *mutvirus* minus the *virus*. We use these error-corrected frequencies to calculate the *percent* of mutations remaining after selection.

---

many nonsynonymous mutations (***Figure 3***A). This selection is more apparent if we correct for the background error rates estimated from the wildtype controls (***Figure 3**-source data* 2). The error-corrected frequencies of stop codons drop to 3%–16% of their original values (***Figure 3**-source data* 2), with the residual stop codons probably due to some non-functional virions surviving due to complementation by other co-infecting virions. The error-corrected frequencies of non-synonymous mutations also drop substantially (43%–49% of their original values), whereas the frequencies of synonymous mutations drop only slightly (85%–95% of their original values). These trends are consistent with the fact that nonsynonymous mutations are often deleterious, whereas synonymous mutations often (although certainly not always, see ***Zanini and Neher, 2013***) have only mild effects on viral growth. ***Figure 3***A only summarizes one aspect of the deep mutational scanning data, but Supplementary files 1 and 2 contain detailed plots showing all aspects of the data (read depth, per-site mutation rate, etc) as generated by the `dms_tools2` software (***Bloom, 2015***, https://jbloomlab.github.io/dms_tools2/).

We used the deep mutational scanning data to estimate the preference of each site in Env for each amino acid via the analysis method described in ***Bloom*** (***2015***). As graphically illustrated

in *Figure 2*A, the preferences for each site are normalized to sum to one. Our libraries were mutagenized at 670 sites in BG505 and 662 sites in BF520, so $670 \times 20 = 13,400$ and $662 \times 20 = 13,240$ preferences were estimated for each Env, respectively. The correlations between the preferences from different experimental replicates are in *Figure 3*B, and the preferences themselves are in *Figure 3*-source data 1. These replicate-to-replicate correlations are substantially higher than those for the deep mutational scanning of LAI Env by *Haddox et al.* (*2016*), which had replicate-to-replicate Pearson correlations of only $R = 0.45$ to 0.50.

While the replicates are well correlated across all replicates for both BG505 and BF520, the replicates for BG505 are more correlated with each other than with replicates for BF520, and vice versa (*Figure 3*B, compare red and blue versus gray plots). This fact hints that there are some shifts in amino-acid preferences between the two Envs—something that is investigated with more statistical rigor later in this paper. Note also that there is a trend for highly preferred amino acids to be more strongly preferred in BG505 than BF520 (most high-preference points in the gray plots in *Figure 3*B fall above the diagonal); however, this trend does not necessarily reflect differences between the Envs. Rather, there were modest differences in the stringency of selection between our deep mutational scans of BG505 and BF520 (*Figure 3*-source data 2 shows that purifying selection better purged stop codons in BG505). In the next section, we correct for these experimental differences by calibrating each dataset to match the stringency of selection in nature.

Amino-acid preferences of the Envs and their relationship to HIV evolution

The most immediate question is how authentically the experimental measurements describe the actual selection on Env function in nature. Direct comparisons between experimentally measured amino-acid preferences and amino-acid frequencies in natural sequences are confounded by the fact that the natural sequences are evolutionarily related. This problem can be overcome by making the comparison in a phylogenetic context to account for the evolutionary relationships among sequences.

Specifically, we used our deep mutational scanning data to construct experimentally informed codon models (ExpCM's) for Env's evolution. An ExpCM is a phylogenetic substitution model that incorporates the functional constraints measured in a deep mutational scanning experiment (*Hilton et al., 2017*). If the experiment captures much of the actual evolutionary constraint on a gene, then an ExpCM will describe the gene's natural evolution better than a standard phylogenetic codon substitution model. The reason is that standard codon substitution models (such as the commonly used Goldman-Yang style models; *Yang et al., 2000*) only model functional constraint via a single parameter that represents the rate of fixation of nonsynonymous protein-altering mutations relative to synonymous ones; this parameter is called dN/dS or ω. In contrast, an ExpCM accounts for the preference of each site for each of the 20 amino acids under the functional selection in the deep mutational scan, and then additionally adds an ω parameter that represents the relative rate of nonsynonymous to synonymous substitutions after accounting for these functional constraints (*Bloom, 2017*; *Hilton et al., 2017*). Importantly, since we expect some sites in Env to be under diversifying selection from immunity, we extended the ExpCM's described in *Hilton et al.* (*2017*) to draw ω from a gamma distribution as is commonly done for codon-substitution models (*Yang et al., 2000*).

Table 1 shows that ExpCM's informed by the deep mutational scanning of either BG505 or BF520 describe the natural evolution of Env vastly better than a standard codon substitution model. In addition to the improved fit of the ExpCM's, we can also interpret the ω parameter. Recall that for standard codon substitution models, ω is simply the rate of fixation of nonsynonymous mutations relative to synonymous ones. For such models, the gene-wide average ω is almost always <1, since purifying selection purges many functionally deleterious amino-acid mutations even for adaptively evolving proteins (*Murrell et al., 2015*). Indeed, *Table 1* shows that Env's gene-wide average ω is <1 for a standard model. But for ExpCM's, ω is the relative rate of nonsynonymous to synonymous substitutions *after* accounting for functional constraints measured in the deep

Model	ΔAIC	LogLikelihood	nParams	stringency	$\overline{\omega}$	ω_α	ω_β	nsites $\omega_r > 1$	nsites $\omega_r < 1$
ExpCM BF520	0.0	-35218.8	7	2.8	1.4	1.0	0.7	66	35
ExpCM BG505	269.0	-35353.3	7	2.1	1.3	0.9	0.7	65	53
Goldman-Yang M5	3455.1	-36941.4	12	nan	0.8	0.6	0.7	14	211

Table 1. Evolutionary models informed by the deep mutational scanning describe HIV's evolution in nature much better than a standard substitution model. Shown are the results of maximum likelihood fitting of substitution models to the clade A phylogeny in *Figure 1*. Experimentally informed codon models (ExpCM, *Hilton et al., 2017*) utilizing the across-replicate average of the deep mutational scanning describe Env's natural evolution far better than a standard codon substitution model (the M5 model of *Yang et al., 2000*) as judged by comparing the Akaike information criteria (ΔAIC, *Posada and Buckley, 2004*). Both ExpCM's have a stringency parameter >1. All models draw ω from a gamma distribution, and the table shows the mean ($\overline{\omega}$) and shape parameters (ω_α and ω_β) of this distribution. The last two columns show the number of sites evolving faster ($\omega_r > 1$) or slower ($\omega_r < 1$) than expected at a false discovery rate of 0.05, as determined using the approach in *Bloom* (*2017*) (see also the last section of the Results). Analyses were performed using `phydms` (*Hilton et al., 2017*, http://jbloomlab.github.io/phydms/). *Table 1*-source-data 1 shows the results for additional substitution models.

Table 1–source data 1. Results for phylogenetic models where ω is not drawn from a gamma-distribution or where the preferences are averaged across sites to eliminate the site specificity are in `modelcomparison.md`.

Figure 4. Amino-acid preferences for the BG505 Env. At each site, the height of the letter is proportional for that site's preference for that amino acid. The top color bar indicates the region of Env (gp120 variable loop, gp120 not variable loop, or gp41). The lower color bar indicates the evidence that the site evolves faster ($\omega_r > 1$) or slower ($\omega_r < 1$) than expected given the experiments (see the last section of the Results and *Bloom, 2017*). We report the *P*-value for $\omega_r \neq 1$ rather than the value of ω_r itself since point estimates of ω_r are unreliable for individual sites due to low numbers of observations, making the *P*-value a better indicator of the strength of the statistical evidence for faster or slower than expected evolution (*Kosakovsky Pond and Frost, 2005*; *Murrell et al., 2012*). The letters above the logos indicate the wildtype amino acid in BG505. Sites are numbered using the HXB2 scheme (*Korber et al., 1998*). This logo plot shows the site-specific amino-acid preferences for BG505 after averaging the replicates and re-scaling by the stringency parameter in *Table 1*. The figure was generated using `dms_tools2` (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/), which in turn utilizes `weblogo` (*Crooks et al., 2004*). The numerical values of the preferences are in *Figure 4*-source data 1, the mapping from sequential to HXB2 numbering is in *Figure 4*-source data 2, and the ω_r values are in *Figure 4*-source data 3.

Figure 4–source data 1. The numerical values of the amino-acid preferences plotted in this figure are in `rescaled_BG505_prefs.csv`.

Figure 4–source data 2. The sequence of BG505 Env and mapping from sequential (*original* column) to HXB2 numbering (*new* column) is in `BG505_to_HXB2.csv`.

Figure 4–source data 3. The ω_r values and associated *P*-values for BG505 in HXB2 numbering are in `BG505_omegabysite.tsv`.

Figure 5. Amino-acid preferences for the BF520 Env. This figure is the same as *Figure 4* except that it shows the data for BF520 instead of BG505. The numerical values of the preferences are in *Figure 5*-source data 1, the mapping from sequential to HXB2 numbering is in *Figure 5*-source data 2, and the ω_r values are in *Figure 5*-source data 3.

Figure 5–source data 1. The numerical values of the amino-acid preferences plotted in this figure are in `rescaled_BF520_prefs.csv`.

Figure 5–source data 2. The sequence of BF520 Env and mapping from sequential (*original* column) to HXB2 numbering (*new* column) is in `BF520_to_HXB2.csv`.

Figure 5–source data 3. The ω_r values and associated *P*-values for BF520 in HXB2 numbering are in `BF520_omegabysite.tsv`.

mutational scanning (*Bloom, 2017*). For the ExpCM's, the gene-wide average ω is >1 (*Table 1*), indicating that external selection (e.g., from immunity) drives Env to fix amino-acid mutations faster than expected under a null model that only accounts for functional constraints on the protein.

ExpCM's also have a stringency parameter that relates selection in the experiments to that in nature. Essentially, this parameter indicates how strongly natural selection prefers the amino acids that are preferred in the deep mutational scanning (*Hilton et al., 2017*). A stringency parameter >1 indicates that natural selection prefers the same amino acids as the experiments, but with greater stringency. Both ExpCM's have stringency parameters >1 (*Table 1*)—a finding that makes sense, since the stop-codon analysis in the previous section suggests that the experimental selections are more lax than natural selection on HIV.

For the entire rest of the paper, we use the experimentally measured preferences re-scaled by the stringency parameters in *Table 1*. The reason we do this is to distinguish genuine differences

Figure 6. Env sites with shifted amino-acid preferences between BG505 and BF520. Note that the preferences have been re-scaled using the stringency parameters in *Table 1* to enable direct comparison across Envs. **(A)** Calculation of the corrected distance between the amino-acid preferences of BG505 and BF520 at four example sites. We have triplicate measurements for each Env. We calculate the distance between each pair of replicate measurements, and group these into comparisons *between* the two Envs and *within* replicates for the same Env. We compute the root-mean-square distance (RMSD) for both sets of comparisons, which we denote as $RMSD_{between}$ and $RMSD_{within}$. The latter quantity is a measure of experimental noise. The noise-corrected distance between Envs at a site, $RMSD_{corrected}$, is simply the distance between the two Envs minus this noise. **(B)** The bottom distribution (orange) shows the corrected distances between BG505 and BF520 at all alignable sites (see *Figure 6*-source data 1 for numerical values). The next distribution (blue) is a null generated by computing the corrected distances on all randomizations of the replicates among Envs. The top two distributions (green) compare Env to the non-homologous influenza hemagglutinin (HA) protein (*Doud and Bloom, 2016*) simply putting sites into correspondence based on sequence number. We compute the *P*-value that a site has shifted between BG505 and BF520 as the fraction of the null distribution that exceeds that shift, and identify significant shifts at a false discovery rate (FDR) of 0.1 using the method of *Benjamini and Hochberg* (*1995*). Using this approach, 30 of the 659 sites have significant shifts (corrected distance ≥0.22). **(C)** All sites that have significantly shifted their amino-acid preferences at an FDR of 0.01. For each site, the logo stacks show the across-replicate average preferences for BG505 and BF520. The wildtype amino acid for that Env is indicated using the small black letters above each logo plot; note how the wildtype amino acid is frequently but not always the most preferred one. The sites are sorted by the magnitude of the shift.

Figure 6–source data 1. The corrected distances between BG505 and BF520 at each site are in `BG505_to_BF520_prefs_dist.csv`.

---

between the two Envs from mere variation in the strength of selection between the two sets of experiments. Re-scaling both sets of preferences to optimally describe Env evolution in nature is a principled way to standardize the measurements; see *Hilton et al.* (*2017*) and the Methods section entitled "Re-scaling the preferences" for a more detailed explanation.

A qualitative way to assess if the deep mutational scanning authentically describes selection on Env function is to visually compare the measurements with existing knowledge. *Figure 4* and *Figure 5* show the re-scaled across-replicate average of the amino-acid preferences for each Env. At sites of known functional importance, these preferences are usually consistent with prior knowledge. For instance, residues T257, D368, E370, W427, and D457 are important for Env binding to CD4 (*Olshevsky et al., 1990*), and all these amino acids are highly preferred in our deep mutational scanning (*Figure 4* and *Figure 5*). Likewise, Env has 10 disulfide bonds (linking sites 54-74, 119-205, 126-196, 131-157, 218-247, 228-239, 296-331, 378-445, 385-418, and 598-604), most of which are important for function (*van Anken et al., 2008*)—and the cysteines at these sites are highly preferred in our deep mutational scanning. The deep mutational scanning is also consistent with prior knowledge about sites that are tolerant of mutations. For instance, Env has five variable loops that mostly evolve under weak constraint in nature (*Starcich et al., 1986*; *Zolla-Pazner and Cardozo, 2010*)—and most sites in these loops are mutationally tolerant in our deep mutational scanning (see sites indicated by gray overlay bars in *Figure 4* and *Figure 5*, such as 132 to 195). It is beyond the scope of this paper to catalog associations between our measurements and all other prior mutational studies of Env, but the concordance of our findings with the above mutational studies, and the fact that our data improve phylogenetic models of Env's natural evolution, suggest that our experiments do a reasonable job of authentically measuring functional selection on Env.

Shifts in amino-acid preferences between BG505 and BF520

The most fundamental question that we seek to address is how similar the amino-acid preferences are between the two Envs. We have already noted that *Figure 3*B shows that the preferences are more correlated for replicate measurements on the same Env than for replicate measurements on different Envs. However, simply comparing correlation coefficients does not identify specific sites where mutational effects have shifted, nor does it quantify the magnitude of any shifts.

We therefore used a more rigorous approach to identify sites where the amino-acid preferences differ between BG505 and BF520 by an amount that exceeds the noise in our experiments. We first re-scaled the preferences from each experimental replicate by the stringency parameter for that Env from *Table 1* to calibrate all measurements to the stringency of natural selection. We then identified the 659 sites in the mutagenized regions of Env that are pairwise alignable between BG505 and BF520 (*Figure 6*-source data 1). For each site, we calculated the shift in amino-acid

preferences between Envs using an approach similar to that of *Doud et al.* (*2015*) as illustrated in *Figure 6*A. This approach calculates the magnitude of the shift after correcting for experimental noise by comparing the differences in preferences between replicates for BG505 and BF520 to the differences between replicates for the same Env. *Figure 6*A shows this calculation for a site that has not shifted (site 598, which strongly prefers cysteine in both Envs), the most shifted site (512, which shifts from being mutationally tolerant in BG505 to strongly preferring alanine in BF520), and two other sites with more intermediate behaviors.

The overall distribution of shifts between BG505 and BF520 is shown in *Figure 6*B. Most sites have relatively small shifts (close to zero), although there is a long tail of sites with large shifts. This tail reaches its upper value with site 512, which has a shift of 0.52 out of a maximum possible of 1.0. How should we interpret this distribution—have mutational effects shifted a lot, or not very much? We can establish an upper-bound for how much sites might shift by comparing Env to a *non*-homologous protein. *Figure 6*B shows the distribution of shifts when comparing Env to influenza's hemagglutinin protein, which has previously had its amino-acid preferences measured by deep mutational scanning (*Doud and Bloom, 2016*) . Most sites have large shifts between Env and hemagglutinin, with the typical shift being ~0.4 and some approaching the maximum value of 1.0. We can also establish a lower-bound by creating a null distribution for the expected shifts if all differences are simply due to experimental noise. This null distribution is created by randomizing the experimental replicates among Envs. *Figure 6*B shows that the null distribution is more peaked at zero than the real distribution, and does not have the same prominent tail of sites with large shifts. The answer to the question of how much mutational effects have shifted is therefore nuanced: they have substantially shifted at some sites, but remain vastly more similar between the two Envs than between two unrelated proteins.

We can use the null distribution to identify sites where the shifts between BG505 and BF520 are significantly larger than the noise in our experiments (*Figure 6*B). There are 30 such sites at a false discovery rate of 0.1. *Figure 6*C shows the amino-acid preferences of these significantly shifted sites for each Env. For the majority of shifted sites, one Env prefers a specific amino acid whereas the other Env tolerates many amino acids; for instance, see sites 512, 516, 599, 165, 605 and 505 in *Figure 6*C. Such broadening and narrowing of a site's mutational tolerance is frequently linked to changes in protein stability, with a more stable protein typically being more mutationally tolerant (*Wang et al., 2002*; *Bloom et al., 2006*; *Gong et al., 2013*; *Kumar et al., 2017*). Work with engineered Env protein in the form of "SOSIP" trimer (*Binley et al., 2000*; *Sanders et al., 2002*) has shown that BG505 SOSIP is more thermostable than BF520 SOSIP (*Verkerke et al., 2016*). Consistent with this fact, sites with altered mutational tolerance are often (although not always, see sites 165 and 520 in *Figure 6*C) more mutationally tolerant in BG505. Differences in Env's expression level might also contribute to a general broadening or narrowing of tolerance to subsequent mutations. The reason is that our experiments select for viral growth (which is affected by both Env function and expression), so it is possible that some of the shifts are due to epistatic mutational effects on expression rather than function.

However, not all of the significantly shifted sites show a simple pattern of broadening or narrowing of mutational tolerance. For instance, site 288 does not alter its mutational tolerance but rather flips its rather narrow amino-acid preference from phenylalanine in BG505 to leucine in BF520 (*Figure 6*C). Thus, there is variation in both the extent and types of shifts observed.

Structural and evolutionary properties of shifted sites

What distinguishes the sites that have undergone significant shifts? First, we analyzed the distribution of shifted sites in context of Env's three-dimensional structure. Env's structure is highly conformationally dynamic and undergoes large changes upon receptor binding and membrane fusion. In an effort to account for these dynamics, we examined multiple conformational states of Env: the closed pre-fusion state (*Stewart-Jones et al., 2016*), the open CD4-bound state (*Ozorowski et al., 2017*), and the post-fusion six-helix bundle (*Weissenhorn et al., 1997*). *Figure 7*A shows the

Figure 7. Characteristics of significantly shifted sites. **(A)** One monomer of the closed pre-fusion Env trimer (PDB 5FYL; *Stewart-Jones et al., 2016*) is colored from white to red according to the magnitude of the mutational shift at each site (red indicates large shift). Sites that are significantly shifted according to *Figure 6*B are in spheres, and all other sites are in cartoon representation. **(B)** There is no significant difference in the relative solvent accessibility of sites that have and have not undergone significant shifts. This observation holds for both the closed trimer conformation in (A) and the CD4-bound trimer conformation (PDB 5VN3; *Ozorowski et al., 2017*). The absolute solvent accessibility of each site was calculated using DSSP (*Kabsch and Sander, 1983*) and normalized to a relative solvent accessibility using the absolute accessibilities from *Tien et al.* (*2013*). **(C)** Sites of significant shifts are clustered in the structures of both the closed and open Env trimers. The left two plots show the distance of each significantly shifted and not-shifted site to the closest other shifted site in the indicated structure. The right-most plot shows the minimum distance across both conformations. The trend for shifts to cluster becomes stronger when considering the minimum distance, suggesting multiple conformations contribute to this trend. **(D)** Large mutational shifts are *not* strongly enriched at sites that have substituted between BG505 and BF520, or at sites that contact sites that have substituted. The plots show the magnitudes of the shifts among structurally resolved sites that have substituted between BG505 and BF520, the non-substituted sites that physically contact a substitution in the indicated structure(s) (any non-hydrogen atom within 3.5Å), and all other sites. *Figure 7*-source data 1 shows that there is a borderline-significant tendency of significantly shifted sites to have substituted. All plots only show sites that are structurally resolved in the indicated structure(s). Structural distances and solvent accessibilities were calculated using all monomers in the trimer. P values were calculated using the Mann-Whitney U test. *Figure 7–Figure Supplement 1* and *Figure 7–Figure Supplement 2* zoom in on some relevant clusters of sites.

Figure 7–Figure supplement 1. Cluster of shifted sites in the post-fusion six-helix bundle of Env.
Figure 7–Figure supplement 2. Clusters of shifted sites in highly dynamic regions of Env.
Figure 7–source data 1. The sites of significant shifts in *Figure 6*B are somewhat more likely to have substituted between BG505 and BF520. This association is borderline statistically significant, with $P = 0.055$ using a Fisher's exact test on the contingency table in `shifts_vs_subs_table.csv`.

---

locations of the shifted sites on the crystal structure of Env in the closed pre-fusion state. There is no visually obvious tendency for shifted sites to preferentially be on Env's surface or in its core, and statistical analysis of both the closed and open states of Env (*Figure 7*B) finds no association between a site's relative solvent accessibility and whether its amino-acid preferences have shifted. We did not attempt to analyze the association between solvent accessibility and shift for the post-fusion six-helix bundle because crystal structures of this conformation only contain ~80 Env residues (*Weissenhorn et al., 1997*; *Chan et al., 1997*; *Tan et al., 1997*).

However, *Figure 7*A does suggest that the sites of significant shifts tend to cluster in Env's structure. A statistical analysis confirms that there is clustering of shifted sites for the closed and open conformations, with the effect being strongest when we define contacts based on the closest intra-residue distance across these two conformations (*Figure 7*C). Therefore, the factors that drive shifts in Env's mutational tolerance often affect physically interacting clusters of residues in a coordinated fashion. We also investigated clustering of shifted sites in the post-fusion six-helix bundle. Because structures of this conformation only resolve the coordinates of ~80 residues, we did not perform a statistical analysis. However, a qualitative analysis revealed that three of the four shifted sites that are resolved in the post-fusion conformation cluster at one end of the helical bundle (*Figure 7–Figure Supplement 1*).

An obvious hypothesis is that strongly shifted sites have substituted between BG505 and BF520, or physically contact such substitutions. According to this hypothesis, substitutions would alter the local physicochemical environment of the substituted site and its neighbors, thereby shifting the amino-acid preferences of sites in the physical cluster. But surprisingly, for both the closed and open conformations, the typical magnitude of shifts is not significantly larger at sites that have substituted, or at sites that contact sites that have experienced substitutions (*Figure 7*D). For the six-helix bundle, there are five structurally resolved substituted sites, one of which is adjacent to the cluster of shifted sites (*Figure 7–Figure Supplement 1*). The number of resolved shifted and substituted sites in this structure is too small for a meaningful statistical analysis of the type in *Figure 7*D. However, the cluster of shifted and substituted sites in the six-helix bundle is also present

Figure 8. Entrenchment of substitutions during Env evolution. There are 12,521 possible amino-acid mutations at the 659 mutagenized sites alignable between BG505 and BF520. The blue densities show the effects of all these mutations to each Env. The orange densities show the effects of just the 92 mutations that convert BG505 to BF520 or vice versa. In the absence of entrenchment, mutating a site in BG505 to its identity in BF520 should have the opposite effect of mutating the site in BF520 to its identity in BG505. In this case, we would expect the BF520→BG505 distribution to be the mirror image of the BG505→BF520 distribution—and both distributions should be centered around zero if the two Envs are equivalently functional. Instead, mutating a site in either Env to its identity in the other Env tends to be deleterious, indicating that substitutions are often entrenched in the Env in which they have fixed. The effect of a mutation is quantified as the log of the ratio of the site's preference for the mutant amino acid to the preference for the wildtype amino acid.

in the closed and open states (*Figure 7–Figure Supplement 1*), and so is included in the statistical analyses in *Figure 7*D.

There is a borderline trend for the significantly shifted sites to be more likely to have substituted between BG505 and BF520 (*Figure 7*-source data 1), but most shifted sites have not substituted (only 8 of the 30 shifted sites differ in amino-acid identity between the two Envs). The lack of strong enrichment in shifts at substituted sites contrasts with previous protein-wide experimental (*Doud et al., 2015*) and simulation-based (*Pollock et al., 2012*; *Shah et al., 2015*) studies of shifting amino-acid preferences, which found that shifts were dramatically more pronounced at sites of substitutions. The difference may arise because these earlier studies examined proteins that are fairly conformationally static (absolutely so in the case of the simulations). That fact that Env is extremely complex and conformationally dynamic (*Munro et al., 2014*; *Ozorowski et al., 2017*) may increase the opportunities for long-range epistasis to enable substitutions at one site to shift the amino-acid preferences of distant sites.

Indeed, many of the shifted sites cluster within regions of Env that are highly conformationally dynamic. *Figure 7–Figure Supplement 2* shows the structural context of these clusters in finer detail. One cluster is at the trimer apex where two of Env's variable loops pack against one another and against an adjacent protomer. These interactions are likely involved in regulating the transition between conformational states, and upon CD4 binding, these loops become highly disordered (*Guttman et al., 2014*; *Ozorowski et al., 2017*). Mutations at two of the shifted sites in this cluster (165 and 307) have been shown to cause Env to assume aberrant conformations, suggesting that these sites can strongly modulate Env's dynamics (*Lee et al., 2017*). Strikingly, this cluster of shifted sites may reflect previously observed differences in the conformational dynamics of this regions between these two Envs; the V2 region of BF520 SOSIP trimer is more accessible to deuterium exchange than the BG505 SOSIP trimer (*Verkerke et al., 2016*). The other cluster of shifted sites is near a network of hydrophobic amino acids that has been proposed to help transmit the large-scale conformational change that takes place upon CD4 binding (*Ozorowski et al., 2017*). One of the shifted sites (site 69) overlaps with this network, and mutations at another (site 64) have been shown to strongly modulate the relative stability of the open and closed conformations (*de Taeye et al., 2015*). In total, these two clusters consist of nearly half of the shifted sites (13 out of 30). One hypothesis why so many shifted sites cluster in these regions is that their dynamic nature allows long-range epistatic interactions to be readily propagated between substituted sites and distant shifted sites. It is difficult to discern exactly how these interactions might occur, but there is certainly a trend for sites that are conformationally dynamic to also be sites that show shifts in their amino-acid preferences during evolution.

Entrenchment of substitutions modestly contributes to mutational shifts

One idea that has recently gained support in the protein-evolution field is that substitutions become "entrenched" by subsequent evolution (*Pollock et al., 2012*; *Shah et al., 2015*; *Starr et al., 2017*). Entrenchment is the tendency of a mutational reversion to become increasingly unfavorable as a sequence evolves. Given two homologs, if there is no entrenchment then the effect of mutating a

Figure 9. Sites in Env that evolve faster or slower in nature than expected given the functional constraints measured in the lab. We calculated the statistical evidence that each site evolves faster ($\omega_r > 1$) or slower ($\omega_r < 1$) than expected given the experimentally measured amino-acid preferences using the method of *Bloom* (*2017*). **(A)** One monomer of the Env trimer (PDB 5FYL; *Stewart-Jones et al., 2016*) is colored from blue to white to red based on the strength of evidence that sites evolve slower than expected (blue), as expected (white) or faster than expected (red) given the BG505 experiments. Sites for which we lack ω_r estimates are colored black. Sites where the rate of evolution is significantly different than expected at a false discovery rate of 0.05 are shown in spheres. **(B)** Like (A) but using the data from the BF520 experiments. For both Envs, sites that evolve significantly slower or faster than expected are often on Env's surface (*Figure 9–Figure Supplement 1*). **(C)** The results are similar regardless of whether the BG505 or BF520 experiments are used. Many of the sites of slower-than-expected evolution are asparagines in N-linked glycosylation motifs (*Figure 9–Figure Supplement 2*). All sites that evolve slower than expected for both experimental datasets are in *Figure 9–Figure Supplement 3*. **(D)** A large cluster of sites that evolve slower than expected is likely involved in Env's transition between open and closed conformational states. Gray boxes indicate sites that *Ozorowski et al.* (*2017*, PDB 5VN3) proposed form a hydrophobic network that regulates the conformational change; blue boxes and sticks indicate sites that evolve slower than expected. All analyses used the phylogenetic tree in *Figure 1*. The ω_r and Q-values are in *Figure 9*-source data 1.

Figure 9–Figure supplement 1. Relative solvent accessibilities of sites evolving faster or slower than expected.
Figure 9–Figure supplement 2. Amino-acid preferences and alignment frequencies for glycosylation motifs.
Figure 9–Figure supplement 3. Amino-acid preferences and alignment frequencies of sites that evolve slower than expected.
Figure 9–source data 1. The ω_r and Q-values are in `merged_omegabysite.csv`.

site in the first homolog to its identity in the second will simply be the opposite of mutating the site in the second homolog to its identity in the first. But if there is entrenchment, then both mutations will be unfavorable, since the site is entrenched at its preferred identity in each homolog.

Figure 8 shows the distribution of effects for mutating all sites that differ between BG505 and BF520 to the identity in the other Env. As expected under entrenchment, the average effect of these mutations is deleterious—although there are a substantial number of sites where the mutational flips are not deleterious. We can get some sense of the magnitude of the entrenchment by comparing the effects of the BG505↔BF520 mutations to the distribution of effects of all possible amino-acid mutations (*Figure 8*). This comparison shows that even unfavorable inter-Env mutational flips are generally more favorable than random amino-acid mutations. Therefore, entrenchment occurs for some but not all substitutions that distinguish BG505 and BF520, and the magnitude of entrenchment is less than the effect of a typical random mutation. Entrenchment of substitutions therefore contributes to some of the mutational shifts. But given that many of these shifts occur at sites that do not even differ between the Envs (*Figure 7*D), entrenchment of substitutions is clearly not the only cause of the shifting amino-acid preferences.

Comparing selection in the lab to natural selection
Our experiments measure the effects of mutations on viral growth in a T-cell line in the lab. But HIV actually evolves in humans, where additional selection pressures on Env are undoubtedly present. For instance, antibody pressure might increase the rate of evolution at some sites (*Albert et al., 1990*; *Wei et al., 2003*; *Richman et al., 2003*), whereas pressure to mask certain epitopes (*Kwong et al., 2002*) might add constraint at other sites. Comparing selection in our experiments to natural selection can identify sites that are under such additional pressures during HIV's actual evolution in humans.

We determined whether each site in Env evolves faster or slower in nature than expected given three models: that evolution is purely neutral (all nonsynonymous and synonymous mutations have equivalent effects), that sites are under the protein-level constraint measured in our experiments with BG505, or that sites are under the constraint measured with BF520. The first model used a standard dN/dS test (the "FEL" method of *Kosakovsky Pond and Frost, 2005*), whereas the other two models are conceptually similar but account for the experimentally measured amino-acid

preferences as described by *Bloom* (*2017*). All three models test if individual sites evolve faster or slower than expected, but they "expect" different things: the dN/dS model expects nonsynonymous and synonymous mutations fix at the same rate, while the ExpCM expects the rate at a site to depend on the experimentally measured functional constraints. In all cases, the evidence that a site r evolves differently than expected is statistically summarized by the P-value that ω_r is > or < 1. The standard dN/dS model finds hundreds of sites that evolve slower than expected under neutral evolution (*Table 1*, $\omega_r < 1$), and only a handful of sites that evolve faster than expected under neutral evolution (*Table 1*, $\omega_r > 1$). This finding is unsurprising, since it is well known that Env is under functional constraint. In contrast, ExpCM's that test the rates of evolution relative to the experimentally measured constraints find far fewer sites that evolve slower than expected, but many more sites that evolve faster (*Table 1*).

The sites that evolve slower or faster than expected from the experiments are shown in *Figure 9*A,B, and overlaid on the logoplots in *Figure 4* and *Figure 5* as the ω_r values. The identified sites are similar regardless of whether we use the experiments with BG505 or BF520 (*Figure 9*C). The reason the results are similar for both experimental datasets is that (as discussed above) the amino-acid preferences of *most* sites are similar in both Envs, suggesting that either dataset provides a reasonable approximation of the site-specific functional constraints across the clade A Envs in *Figure 1*.

What causes some sites to evolve faster or slower in nature than expected from the experiments? The answer in both cases is likely to be immune selection. Most of the sites of faster-than-expected evolution are on the surface of Env (*Figure 9*A,B and *Figure 9–Figure Supplement 1*). Env's escape from autologous neutralizing antibodies often involves amino-acid substitutions in surface-exposed regions (*Moore et al., 2009*), including at many of the sites that evolve faster than expected. Since our deep mutational scanning did not impose antibody pressure, sites where substitutions are antibody-driven will evolve faster in nature than expected from the experiments.

Interestingly, immune selection also offers a plausible explanation for the sites that evolve *slower* than expected. In addition to escaping immunity via substitutions at antibody-binding footprints, Env is notorious for employing a range of more general strategies to reduce its susceptibility to antibodies. These strategies include shielding immunogenic regions with glycans (*Wei et al., 2003*; *Stewart-Jones et al., 2016*; *Gristick et al., 2016*) or hiding them by adopting a closed protein conformation (*Kwong et al., 2002*; *Guttman et al., 2015*; *Ozorowski et al., 2017*). Sites that contribute to such general immune-evasion strategies will be under a constraint in nature that is not present in our experiments—and indeed, such sites evolve more slowly than expected from our experiments. For instance, we find very little selection to maintain most glycans in our cell-culture experiments. Of the 21 N-linked glycosylation sites shared between BG505 and BF520, only four are under strong selection to maintain the glycan in our experiments—despite the fact that most are conserved in nature (*Figure 9*C and *Figure 9–Figure Supplement 2*). This finding concords with prior literature suggesting that these glycans are selected primarily for their role in immune evasion (*Pugach et al., 2004*; *Wang et al., 2013*; *Rathore et al., 2017*). Similarly, a network of sites that help regulate Env's transition between open and closed conformations that have different antibody susceptibilities (*Figure 9*D) also evolve slower in nature than expected from our experiments. Therefore, we can distinguish evolutionary patterns that are shaped by simple selection for Env function from those that are due to the additional complex pressures imposed during human infections.

Discussion

We have experimentally measured the preference for each amino acid at each site in the ectodomain and transmembrane domain of two Envs under selection for viral growth in cell culture. These amino-acid preference maps are generally consistent with prior knowledge about sites that are important for protein properties such as receptor binding or disulfide-mediated stability. However, the main value of these maps comes not from comparing them with prior knowledge, but from the fact that such prior knowledge encompasses just a small fraction of the vast mutational space

available to Env. Because Env evolves so rapidly, every study of this protein must be placed in an evolutionary context, and our comprehensive amino-acid preference maps potentially enable this in ways that prior piecemeal studies of mutations cannot.

But these maps come with a potentially serious caveat: each one is measured for just a single Env variant. The major question that our study aimed to answer is whether the maps are still useful for evolutionary questions, or whether Env's amino-acid preferences shift so rapidly that each map only applies to the specific HIV strain for which it was measured. This question is reminiscent of one that was grappled with in the early days of protein crystallography, when it first became possible to build maps of a protein's structure. Because it was not (and is still not) possible to crystallize every variant of a protein, it was necessary to determine whether protein structures could be usefully generalized among homologs. Fortunately for the utility of structural biology, it soon became apparent that closely homologous proteins have similar structures (*Chothia and Lesk, 1986*; *Sander and Schneider, 1991*). This rough generalizability of protein structures holds even for a protein as conformationally complex as Env—for although there are many examples of mutations that alter aspects of Env's conformation and dynamics (*Kwong et al., 2000*; *White et al., 2010*; *Almond et al., 2010*; *Davenport et al., 2013*), SOSIP trimer structures from diverse Env strains remain highly similar in most respects (*Julien et al., 2015*; *Pugach et al., 2015*; *Stewart-Jones et al., 2016*; *Verkerke et al., 2016*; *Gristick et al., 2016*).

Our results show that amino-acid preference maps of Env also have a useful level of conservation for many purposes. From a qualitative perspective, the amino-acid preferences look mostly similar between BG505 and BF520, and so provide a valuable reference for estimating which mutations are likely to be tolerated at each site in diverse HIV strains. Indeed, we anticipate that the complete maps of mutational effects in *Figure 4* and *Figure 5* will be useful for future sequence-structure-function studies. From an analytical perspective, a powerful use of our maps is to identify sites that evolve differently in nature than is required by the simple selection for viral growth imposed in our experiments—and the identified sites are largely the same regardless of whether the analysis uses an amino-acid preference map from BG505 or BF520.

Of course, from the perspective of protein evolution, the most interesting sites are the exceptions to the general conservation of amino-acid preferences. Consistent with studies of other proteins (*Natarajan et al., 2013*; *Harms and Thornton, 2014*; *Doud et al., 2015*; *Starr et al., 2017*), we find a subset of sites that change markedly in which mutations they tolerate. Some shifted sites simply accommodate more amino acids in the more stable BG505 Env—a type of shift that has been well-documented for other proteins (*Wang et al., 2002*; *Bloom et al., 2006*; *Gong et al., 2013*; *Kumar et al., 2017*). But interestingly, there is no strong trend for shifts to be enhanced at sites that differ between BG505 and BF520. Recent studies of protein evolution have focused on the idea that substitutions become "entrenched" as sites shift to accommodate new amino acids (*Pollock et al., 2012*; *Shah et al., 2015*; *Bazykin, 2015*; *Starr et al., 2017*). Indeed, a prior protein-wide comparison of amino-acid preferences across homologs of influenza nucleoprotein found a significant enrichment of shifts at sites of substitutions (*Doud et al., 2015*). But although there is some entrenchment of differences between BG505 and BF520, this is not the major factor behind the shifts in amino-acid preferences: the majority of sites that have shifted between BG505 and BF520 actually have the same wildtype amino acid in both Envs even though the preferences have shifted. This rather surprising result might be due to Env's exceptional conformational complexity—mutations can cause long-range alterations in Env's conformation (*Kwong et al., 2000*; *White et al., 2010*; *Almond et al., 2010*; *Davenport et al., 2013*), so it seems plausible that they might also shift mutational tolerance at distant sites. Regardless of the exact mechanism, our large-scale datasets of mutational effects in multiple viral strains should be useful for efforts to computationally parameterize "fitness landscapes" of Env (*Kouyos et al., 2012*; *Ferguson et al., 2013*; *Mann et al., 2014*; *Barton et al., 2015*; *Louie et al., 2018*).

Our experiments provide highly quantitative data on the mutational tolerance of Env under selection for viral growth in cell culture. These data are amenable to rigorous functional and

evolutionary analyses. Here we have shown how these data can be compared between Envs to identify sites where mutational tolerance shifts with viral genotype, or between experiments and nature to identify sites under different pressure in the lab and in humans. Future experiments that modulate selection pressures in other relevant ways should provide further insight into the forces that drive and constrain HIV's evolution.

Methods and Materials

Creation of codon-mutant libraries

Our codon mutant libraries mutagenized all sites in *env* to all 64 codons, except that the signal peptide and cytoplasmic tail were not mutagenized. The rationale for excluding these regions is that they are not part of Env's ectodomain, and are prone to mutations that strongly modulate Env's expression level (*Chakrabarti et al., 1989*; *Yuste et al., 2004*; *Li et al., 1994*).

The codon-mutant libraries were generated using the approach originally described in *Bloom* (*2014a*), with the modification of *Dingens et al.* (*2017*) to ensure more uniform primer melting temperatures. The computer script used to design the mutagenesis primers (along with some detailed implementation notes) is at https://github.com/jbloomlab/CodonTilingPrimers. For BF520, the three libraries are the same ones described by *Dingens et al.* (*2017*). For BG505, we created three libraries for this study. The wildtype BG505 sequence used for these libraries is in Supplemental file 3. The BG505 mutagenesis primers are in Supplemental file 4.

The end primers for the BG505 mutagenesis were: 5'-tgaaggcaaaactactggtccgtctcgagcagaagac agtggcaatgaga-3' and 5'-gctacaaatgcatataacagcgtctcattctttccctaacctcaggcca-3'. As with BF520, we cloned the BG505 *env* libraries into the *env* locus of the full-length proviral genome of HIV strain Q23 (another subtype-A transmitted/founder virus; *Poss and Overbaugh, 1999*) using the high-efficiency cloning vector described in *Dingens et al.* (*2017*). For this cloning, we digested the cloning vector with BsmBI, and then used PCR to elongate the amplicons to include 30 nucleotides at each end that were identical in sequence to the ends of the BsmBI-digested vector. The primers for this PCR were: 5'-agataggttaattgagagaataagagaaagagcagaagacagtggcaatgagagtgatgg-3' and 5'-ctcctggtgctgct ggaggggcacgtctcattctttccctaacctcaggccatcc-3'. Next, we used NEBuilder HiFi DNA Assembly (NEB, E2621S) to clone the *env* amplicons into the BsmBI-digested plasmids. We purified the assembled products using Agencourt AMPure XP beads (Beckman Coulter, A63880) using a bead-to-sample ratio of 1.5, and then transformed the purified products into Stellar electrocompetent cells (Takara, 636765). The transformations yielded between 1.5-3.6 million unique clones for each of the three replicate libraries, as estimated by plating 1:2,000 dilutions of the transformations. We scraped the plated colonies and maxiprepped the plasmid DNA; unlike in *Dingens et al.* (*2017*), we did not include a 4-hour outgrowth step after the scraping step. For the wildtype controls, we maxiprepped three independent cultures of wildtype BG505 *env* cloned into the same Q23 proviral plasmid. See *Figure 2–Figure Supplement 1* and *Figure 3*A for information on the average mutation rate in these libraries as estimated by Sanger sequencing and deep sequencing, respectively.

Generation and passaging of viruses

For BG505, we generated mutant virus libraries from the proviral plasmid libraries by transfecting 293T cells in three 6-well plates (so 18 wells total per library) with a per-well mixture of 2 μg plasmid DNA, 6 μl FuGENE 6 Transfection Reagent (Promega, E269A), and 100 μl DMEM. The 293T cells were seeded at 5×10^5 cells/well in D10 media (DMEM supplemented with 10% FBS, 1% 200 mM L-glutamine, and 1% of a solution of 10,000 units/mL penicillin and 10,000 μg/mL streptomycin) the day before transfection, such that they were approximately 50% confluent at the time of transfection. In parallel, we generated wildtype viruses by transfecting one 6-well plate of 293T cells with each wildtype plasmid replicate. At 2 days post-transfection, we harvested the transfection supernatant, passed it through a 0.2 μm filter to remove cells, treated the supernatant with DNAse to digest residual plasmid DNA as in *Haddox et al.* (*2016*), and froze aliquots at -80°C. We thawed and titered aliquots using the TZM-bl assay in the presence of 10 μg/mL DEAE-dextran as described in *Dingens et al.* (*2017*).

We conducted the low MOI viral passage illustrated in *Figure 2*A in SupT1.CCR5 cells (obtained from Dr. James Hoxie; *Boyd et al., 2015*). During this passage, cells were maintained in R10 media, which has the same composition as the D10 described above, except RPMI-1640 (GE Healthcare Life Sciences, SH30255.01) is used in the place of DMEM. In addition, the media contained 10 μg/mL

DEAE-dextran to enhance viral infection. We infected cells with 4 million (for replicate 1) or 5 million (for replicates 2 and 3) TZM-bl infectious units of mutant virus at an MOI of 0.01, with cells at a starting concentration of 1 million cells/mL in vented tissue-culture flasks (Fisher Scientific, 14-826-80). At day 1 post-infection, we pelleted cells, aspirated the supernatant, and resuspended cell pellets in the same volume of fresh media still including the DEAE-dextran. At 2 days post-infection, we doubled the volume of each culture with fresh media still including DEAE-dextran. At 4 days post-infection, we pelleted cells, passed the viral supernatant through a 0.2 μm filter, concentrated the virus ~30 fold using ultracentrifugation as described in *Dingens et al.* (*2017*), and froze aliquots at -80°C. In parallel, for each replicate, we also passaged 2×10^5 (for replicate 1) or 5×10^5 (for replicates 2 and 3) TZM-bl infectious units of wildtype virus using the same procedure. To obtain final titers for our concentrated virus, we thawed one of the aliquots stored at -80°C and titered using the TZM-bl assay in the presence of 10 μg/mL DEAE-dextran.

For the final short-duration infection illustrated in *Figure 2*A, for each replicate we infected 10^6 TZM-bl infectious units into 10^6 SupT1.CCR5 cells in the presence of 100 μg/mL DEAE-dextran (note that this is a 10-fold higher concentration of DEAE-dextran than for the other steps, meaning that the effective MOI of infection is higher if DEAE-dextran has the expected effect of enhancing viral infection). Three hours post-infection, we pelleted the cells and resuspended them in fresh media without any DEAE-dextran. At 12 hours post-infection, we pelleted cells, washed them once with PBS, and then used a miniprep kit to harvest reverse-transcribed unintegrated viral DNA (*Haddox et al., 2016*).

The generation, passaging and deep sequencing of BF520 was done in a highly similar fashion, except that we only had a single replicate of the wildtype control. Note that the final passaged BF520 mutant libraries analyzed here actually correspond to the "no-antibody" controls described in *Dingens et al.* (*2017*), but that study did not analyze the initial plasmid mutant libraries relative to these passaged viruses, and so was not able to provide measurements of the amino-acid preferences.

Illumina deep sequencing

We deep sequenced all of the samples shown in *Figure 3*A: the plasmid mutant libraries and wildtype plasmid controls, and the cDNA from the final mutant viruses and wildtype virus controls. In order to increase the sequence accuracy, we used a barcoded-subamplicon sequencing strategy. This general strategy was originally applied in the context of deep mutational scanning by *Wu et al.* (*2014*), and the specific protocol used in our work is described in *Doud and Bloom* (*2016*) (see also https://jbloomlab.github.io/dms_tools2/bcsubamp.html).

The primers used for BG505 are in Supplementary file 5. The primers used for BF520 are in *Dingens et al.* (*2017*). The data generated by the Illumina deep sequencing are on the Sequence Read Archive under the accession numbers provided at the beginning of the Jupyter notebook in Supplementary files 1 and 2.

Analysis of deep-sequencing data

We analyzed the deep-sequencing data using the `dms_tools2` software package (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/, version 2.2.4). The algorithm that goes from the deep-sequencing counts to the amino-acid preferences is that described in *Bloom* (*2015*) (see also https://jbloomlab.github.io/dms_tools2/prefs.html). A Jupyter notebook that performs the entire analysis including generation of most of the figures in this paper is in Supplementary file 1. An HTML rendering of this notebook is in Supplementary file 2. A repository containing all of this code is also available at https://github.com/jbloomlab/EnvMutationalShiftsPaper (*Haddox et al., 2018*, copy archived at https://github.com/elifesciences-publications/EnvMutationalShiftsPaper).

The Jupyter notebooks in Supplementary files 1 and 2 also contain numerous plots that summarize relevant aspects of the deep sequencing such as read depth, per-codon mutation frequency,

mutation types, etc. Supplementary file 1 also contains text files and CSV files with the numerical values shown in these plots.

Citations are also owed to `weblogo` (*Crooks et al., 2004*, http://weblogo.threeplusone.com/) and `ggseqlogo` (*Wagih, 2017*, https://omarwagih.github.io/ggseqlogo/), which were used in the generation of the logoplots.

Alignments and phylogenetic analyses of Env sequences

A basic description of the process used to generate the clade A sequence alignment in *Figure 1*-source data 1, the alignment mask in *Figure 1*-source data 2, and the phylogenetic tree in *Figure 1* are provided in the legend to that figure. An algorithmic description of how the alignment and tree were generated are in Supplementary files 1 and 2.

For fitting of the phylogenetic substitution models, we used `phydms` (*Hilton et al., 2017*, http://jbloomlab.github.io/phydms/, version 2.2.1) to optimize the substitution model parameters and branch lengths on the fixed tree topology in *Figure 1*. The Goldman-Yang (or YNGKP) model used in *Table 1* is the M5 variant described by *Yang et al.* (*2000*), with the equilibrium codon frequencies determined empirically using the CF3x4 method (*Pond et al., 2010*). For the ExpCM shown in *Table 1*, we extended the models with empirical nucleotide frequencies described in *Hilton et al.* (*2017*) to also allow ω to be drawn from discrete gamma-distributed categories exactly as for the M5 model. These ExpCM with gamma-distributed ω were implemented in `phydms` using the equations provided by *Yang* (*1994*) (see also http://jbloomlab.github.io/phydms/implementation.html#models-with-a-gamma-distributed-model-parameter). The preferences were re-scaled by the stringency parameters in *Table 1* as described in *Hilton et al.* (*2017*). For both the M5 model and the ExpCM with a gamma-distributed ω, we used four categories for the discretized gamma distribution.

Table 1-source-data 1 shows the results for a wider set of models than those used in *Table 1*. These include the M0 model of *Yang et al.* (*2000*), ExpCM without a gamma-distributed ω, and ExpCM in which the amino-acid preferences are averaged across sites as a control to ensure that the improved performance of these models is due to their site-specificity. Note how for these Env alignments, using a gamma-distributed ω is very important in order for the ExpCMs to outperform the M5 model—we suspect this is because there are many sites of strong diversifying selection.

For detection of sites with faster or slower than expected evolution, we used the approach in *Bloom* (*2017*), which is exactly modeled on the FEL approach of *Kosakovsky Pond and Frost* (*2005*) but extended to ExpCM. This approach estimates a P-value that ω_r is not equal to one for each site r using a likelihood-ratio test. The actual point estimates of ω_r are unreliable for individual sites due to the limited number of observations, so we report the P-value that ω_r is not equal to one, which is a better indication of the strength of the statistical evidence for faster or slower than expected evolution (*Kosakovsky Pond and Frost, 2005*; *Murrell et al., 2012*). For the Q-values and false discovery rate testing, we considered the tests for $\omega_r > 1$ and $\omega_r < 1$ separately.

Supplementary files 1 and 2 contains the code that runs `phydms` to reproduce all of these analyses.

Re-scaling the preferences

The amino-acid preferences that are directly extracted from the deep sequencing data essentially give the enrichment / depletion of each mutation, normalized to sum to one at each site (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/prefs.html). However, the extent that any mutation is enriched or depleted is a combination of two factors: the inherent effect of that mutation, and the "stringency" of the experimental selection. For instance, if the selection is weak, then deleterious mutations will only be slightly depleted; conversely, if selection is strong, then deleterious mutations will be greatly depleted. The fact that the preferences depend on the stringency of the experimental selection is important if we want to compare results between Envs. The reason is that our goal is to identify differences in the inherent effects of mutations between Envs, not simply find differences

due to variation in experimental stringency. Of course, we have done our best to perform the experiments for BG505 and BF520 equivalently, but because these are different viruses with different growth rates, it is impossible to exactly match the experimental stringencies. This can be seen in *Figure 3*-source data 2, which shows that stop codons were more depleted for BG505 than BF520, indicating that selection in our experiments was more stringent for BG505.

How should we best re-scale the preferences? Raising them to a power is a sensible approach. To see why, imagine a mutation that is depleted 3-fold after 2 rounds of viral growth. If our experiment instead allowed $2^2 = 4$ rounds of viral growth, then the mutation would be depleted $3^2 = 9$-fold. More generally, if a mutation is enriched in frequency by ϕ-fold after n rounds of viral growth, then it will be enriched in frequency by ϕ^β-fold after $\beta \times n$ rounds of viral growth. Since the amino-acid preferences are conceptually equivalent to the re-normalized enrichments of mutations (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/prefs.html), it therefore makes sense that the re-scaled preference $\pi_{r,a}$ for amino-acid a at site r should be related to the directly measured preference $\hat{\pi}_{r,a}$ by $\pi_{r,a} \propto \left(\hat{\pi}_{r,a}\right)^\beta$. And indeed, this is exactly the re-scaling scheme described in *Hilton et al.* (*2017*) that we use to re-scale our preferences for BG505 and BF520.

The last point is how to choose the re-scaling parameter β for each Env. It turns out that the features that we have described above for our experiments are also a feature of natural evolution: the expected frequency of a substitution during evolution depends not only on the inherent fitness effect of that mutation, but also on the effective population size, which is conceptually somewhat similar to the stringency of selection. It turns out that in a mutation-selection phylogenetic model of evolution, if the amino-acid preferences are taken to represent the "fitness effects" of mutations, then the exponential scaling parameter β is proportional to the effective population size (*Halpern and Bruno, 1998*; *McCandlish and Stoltzfus, 2014*; *Bloom, 2014b*). Therefore, fitting the β parameter using a phylogenetic approach enables standardization of the preferences for the two Envs, and re-scales the preferences so that they best match with the actual stringency of selection observed in nature (*Hilton et al., 2017*).

Note that in practice this re-scaling scheme is roughly equivalent to a more heuristic approach that has been used by *Gray et al.* (*2017*) and others. In this heuristic approach, the log-transformed enrichment ratios from different experiments are adjusted so that the distributions have equal spreads. Since multiplying log-transformed enrichment ratios is equivalent to exponentiating amino-acid preferences, these two re-scaling procedures apply the same mathematical transformation.

Identifying sites of shifted amino-acid preference

When identifying shifts in amino-acid preferences between the two Envs, we needed a way to quantify differences between the Envs while accounting for the fact that our measurements are noisy. The approach we use is based closely on that of *Doud et al.* (*2015*), and is illustrated graphically in *Figure 6*A. The RMSD$_{\text{corrected}}$ value is our measure of the magnitude of the shift. *Figure 6*A, its legend, and the associated text completely explains these calculations with the following exception: they do not detail how the "distance" between any two preference measurements was calculated. The distance between preferences at each site was simply defined as half of the sum of absolute value of the difference between preferences for each amino acid. Specifically, for a given site r, let $\pi^i_{r,a}$ be the re-scaled preference for amino-acid a in homolog i (e.g., BG505) and let $\pi^j_{r,a}$ be the re-scaled preference for that same amino acid in homolog j (e.g., BF520). Then the distance between the homologs at this site is simply $D^{i,j}_r = \frac{1}{2} \sum_a |\pi^i_{r,a} - \pi^j_{r,a}|$. The factor of $\frac{1}{2}$ is used so that the maximum distance will always fall between zero and one.

Analysis of entrenchment

For the analysis in *Figure 8*, the results are presented in terms of the mutational effects rather than the amino-acid preferences. If $\pi_{r,a}$ is the preference of site r for amino-acid a and $\pi_{r,a'}$ is the preference for amino-acid a' (both re-scaled by the stringency parameters in *Table 1*), then the estimated effect of the mutation from a to a' is simply $\log\left(\frac{\pi_{r,a'}}{\pi_{r,a}}\right)$.

Data and code availability

All code and input data required to reproduce all analyses in this paper are in Supplementary file 1 (see also Supplementary file 2). A repository containing all of this code is also available at https://github.com/jbloomlab/EnvMutationalShiftsPaper (*Haddox et al., 2018*, copy archived at https://github.com/elifesciences-publications/EnvMutationalShiftsPaper). The deep sequencing data are on the Sequence Read Archive with the accession numbers listed in Supplementary files 1 and 2.

References

Al-Mawsawi LQ, Wu NC, Olson CA, Shi VC, Qi H, Zheng X, Wu TT, Sun R. High-throughput profiling of point mutations across the HIV-1 genome. Retrovirology. 2014; 11(1):124.

Albert J, Abrahamsson B, Nagy K, Aurelius E, Gaines H, Nyström G, Fenyö EM. Rapid development of isolate-specific neutralizing antibodies after primary HIV-1 infection and consequent emergence of virus variants which resist neutralization by autologous sera. AIDS. 1990; 4(2):107–112.

Almond D, Kimura T, Kong X, Swetnam J, Zolla-Pazner S, Cardozo T. Structural conservation predominates over sequence variability in the crown of HIV type 1's V3 loop. AIDS research and human retroviruses. 2010; 26(6):717–723.

van Anken E, Sanders RW, Liscaljet IM, Land A, Bontjer I, Tillemans S, Nabatov AA, Paxton WA, Berkhout B, Braakman I. Only five of 10 strictly conserved disulfide bonds are essential for folding and eight for function of the HIV-1 envelope glycoprotein. Molecular Biology of the Cell. 2008; 19(10):4298–4309.

Ashenberg O, Gong LI, Bloom JD. Mutational effects on stability are largely conserved during protein evolution. Proceedings of the National Academy of Sciences. 2013; 110(52):21071–21076.

Barton JP, Kardar M, Chakraborty AK. Scaling laws describe memories of host–pathogen riposte in the HIV population. Proceedings of the National Academy of Sciences. 2015; 112(7):1965–1970.

Bazykin GA. Changing preferences: deformation of single position amino acid fitness landscapes and evolution of proteins. Biology Letters. 2015; 11(10):20150315.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B. 1995; p. 289–300.

Binley JM, Sanders RW, Clas B, Schuelke N, Master A, Guo Y, Kajumo F, Anselma DJ, Maddon PJ, Olson WC, et al. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. Journal of Virology. 2000; 74(2):627–643.

Bloom JD. An experimentally determined evolutionary model dramatically improves phylogenetic fit. Mol Biol Evol. 2014; 31(8):1956–1978.

Bloom JD. An experimentally informed evolutionary model improves phylogenetic fit to divergent lactamase homologs. Molecular Biology and Evolution. 2014; 31(10):2753–2769.

Bloom JD. Software for the analysis and visualization of deep mutational scanning data. BMC Bioinformatics. 2015; 16:168.

Bloom JD. Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. Biology Direct. 2017; 12(1):1.

Bloom JD, Labthavikul ST, Otey CR, Arnold FH. Protein stability promotes evolvability. Proceedings of the National Academy of Sciences. 2006; 103(15):5869–5874.

Boyd DF, Peterson D, Haggarty BS, Jordan AP, Hogan MJ, Goo L, Hoxie JA, Overbaugh J. Mutations in HIV-1 envelope that enhance entry with the macaque CD4 receptor alter antibody recognition by disrupting quaternary interactions within the trimer. Journal of Virology. 2015; 89(2):894–907.

Burton DR, Stanfield RL, Wilson IA. Antibody vs. HIV in a clash of evolutionary titans. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(42):14943–14948.

Chakrabarti L, Emerman M, Tiollais P, Sonigo P. The cytoplasmic domain of simian immunodeficiency virus transmembrane protein modulates infectivity. Journal of Virology. 1989; 63(10):4395–4403.

Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. Cell. 1997; 89(2):263–273.

Chan YH, Venev SV, Zeldovich KB, Matthews CR. Correlation of fitness landscapes from three orthologous TIM barrels originates from sequence and structure constraints. Nature Communications. 2017; 8:14614.

Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. The EMBO Journal. 1986; 5(4):823.

Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Research. 2004; 14(6):1188–1190.

Davenport TM, Guttman M, Guo W, Cleveland B, Kahn M, Hu SL, Lee KK. Isolate-specific differences in the conformational dynamics and antigenicity of HIV-1 gp120. Journal of Virology. 2013; 87(19):10855–10873.

Dingens AS, Haddox HK, Overbaugh J, Bloom JD. Comprehensive mapping of HIV-1 escape from a broadly neutralizing antibody. Cell Host & Microbe. 2017; 21:777–787.

Doud MB, Ashenberg O, Bloom JD. Site-specific amino acid preferences are mostly conserved in two closely related protein homologs. Molecular Biology and Evolution. 2015; 32(11):2944–2960.

Doud MB, Bloom JD. Accurate measurement of the effects of all amino-acid mutations to influenza hemagglutinin. Viruses. 2016; 8:155.

Duenas-Decamp M, Jiang L, Bolon D, Clapham PR. Saturation Mutagenesis of the HIV-1 Envelope CD4 Binding Loop Reveals Residues Controlling Distinct Trimer Conformations. PLoS Pathogens. 2016; 12(11):e1005988.

Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pépin J, et al. The early spread and epidemic ignition of HIV-1 in human populations. Science. 2014; 346(6205):56–61.

Ferguson AL, Mann JK, Omarjee S, Ndung'u T, Walker BD, Chakraborty AK. Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. Immunity. 2013; 38(3):606–617.

Gong LI, Suchard MA, Bloom JD. Stability-mediated epistasis constrains the evolution of an influenza protein. eLife. 2013; 2:e00631.

Goo L, Chohan V, Nduati R, Overbaugh J. Early development of broadly neutralizing antibodies in HIV-1-infected infants. Nature Medicine. 2014; 20(6):655–658.

Gray VE, Hause RJ, Fowler DM. Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. Genetics. 2017; 207(1):53.

Gristick HB, von Boehmer L, West Jr AP, Schamber M, Gazumyan A, Golijanin J, Seaman MS, Fätkenheuer G, Klein F, Nussenzweig MC, et al. Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. Nature Structural & Molecular Biology. 2016; 23(10):906–915.

Guttman M, Cupo A, Julien J, Sanders R, Wilson I, Moore J, Lee K. Antibody potency relates to the ability to recognize the closed, pre-fusion form of HIV Env. Nature Communications. 2015; 6:6144–6144.

Guttman M, Garcia NK, Cupo A, Matsui T, Julien JP, Sanders RW, Wilson IA, Moore JP, Lee KK. CD4-induced activation in a soluble HIV-1 Env trimer. Structure. 2014; 22(7):974–984.

Haddox HK, Dingens AS, Bloom JD. Experimental estimation of the effects of all amino-acid mutations to HIV's envelope protein on viral replication in cell culture. PLoS Pathogens. 2016; 12(12):e1006114.

Haddox HK, Dingens AS, Hilton SK, Overbaugh J, Bloom JD. Computer code for "Mapping mutational effects along the evolutionary landscape of HIV envelope". Github. 2018; commit ec358fd. https://github.com/jbloomlab/EnvMutationalShiftsPaper.

Halpern AL, Bruno WJ. Evolutionary distances for protein-coding sequences: modeling site-specific residue frequencies. Molecular Biology and Evolution. 1998; 15(7):910–917.

Harms MJ, Thornton JW. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature. 2014; 512(7513):203.

Hedges SB, Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics. 2006; 22(23):2971–2972.

Hilton SK, Doud MB, Bloom JD. phydms: Software for phylogenetic analyses informed by deep mutational scanning. PeerJ. 2017; 5:e3657.

Huang J, Kang BH, Pancera M, Lee JH, Tong T, Feng Y, Imamichi H, Georgiev IS, Chuang GY, Druz A, et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature. 2014; 515(7525):138–142.

Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, Klasse PJ, Burton DR, Sanders RW, Moore JP, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science. 2013; 342(6165):1477–1483.

Julien JP, Lee JH, Ozorowski G, Hua Y, de la Peña AT, de Taeye SW, Nieusma T, Cupo A, Yasmeen A, Golabek M, et al. Design and structure of two HIV-1 clade C SOSIP. 664 trimers that increase the arsenal of native-like Env immunogens. Proceedings of the National Academy of Sciences. 2015; 112(38):11947–11952.

Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983; 22(12):2577–2637.

Klink G, Bazykin G. Parallel evolution of metazoan mitochondrial proteins. Genome Biology and Evolution. 2017; 9:1341–1350.

Korber B, Foley BT, Kuiken C, Pillai SK, Sodroski JG, et al. Numbering positions in HIV relative to HXB2CG. Human Retroviruses and AIDS. 1998; 3:102–111.

Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular Biology and Evolution. 2005; 22(5):1208–1222.

Kouyos RD, Leventhal GE, Hinkley T, Haddad M, Whitcomb JM, Petropoulos CJ, Bonhoeffer S. Exploring the complexity of the HIV-1 fitness landscape. PLoS Genetics. 2012; 8(3):e1002551.

Kumar A, Natarajan C, Moriyama H, Witt CC, Weber RE, Fago A, Storz JF. Stability-mediated epistasis restricts accessible mutational pathways in the functional evolution of avian hemoglobin. Molecular Biology and Evolution. 2017; 34(5):1240–1251.

Kwong PD, Doyle ML, Casper DJ, Cicala C, Leavitt SA, Majeed S, Steenbeke TD, Venturi M, Chaiken I, Fung M, et al. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. Nature. 2002; 420(6916):678–682.

Kwong PD, Wyatt R, Majeed S, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. Structure. 2000; 8(12):1329–1339.

Lee JH, Andrabi R, Su CY, Yasmeen A, Julien JP, Kong L, Wu NC, McBride R, Sok D, Pauthner M, et al. A broadly neutralizing antibody targets the dynamic HIV envelope trimer apex via a long, rigidified, and anionic β-hairpin structure. Immunity. 2017; 46(4):690–702.

Li Y, Luo L, Thomas DY, Kang OY. Control of expression, glycosylation, and secretion of HIV-1 gp120 by homologous and heterologous signal sequences. Virology. 1994; 204(1):266–278.

Louie RH, Kaczorowski KJ, Barton JP, Chakraborty AK, McKay MR. Fitness landscape of the human immunodeficiency virus envelope protein that is targeted by antibodies. Proceedings of the National Academy of Sciences. 2018; p. 201717765.

Lynch RM, Shen T, Gnanakaran S, Derdeyn CA. Appreciating HIV type 1 diversity: subtype differences in Env. AIDS Research and Human Retroviruses. 2009; 25(3):237–248.

Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, Burton DR, Sanders RW, Moore JP, Carragher B, et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science. 2013; 342(6165):1484–1490.

Mann JK, Barton JP, Ferguson AL, Omarjee S, Walker BD, Chakraborty A, Ndung'u T. The fitness landscape of HIV-1 gag: advanced modeling approaches and validation of model predictions by in vitro testing. PLoS Computational Biology. 2014; 10(8):e1003776.

McCandlish DM, Stoltzfus A. Modeling evolution using the probability of fixation: history and implications. The Quarterly Review of Biology. 2014; 89(3):225–252.

Moore PL, Gray ES, Morris L. Specificity of the autologous neutralizing antibody response. Current Opinion in HIV and AIDS. 2009; 4(5):358.

Munro JB, Gorman J, Ma X, Zhou Z, Arthos J, Burton DR, Koff WC, Courter JR, Smith AB, Kwong PD, et al. Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. Science. 2014; 346(6210):759–763.

Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP, Smith DM, et al. Gene-wide identification of episodic selection. Molecular Biology and Evolution. 2015; 32(5):1365–1371.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Pond SLK. Detecting individual sites subject to episodic diversifying selection. PLoS Genetics. 2012; 8(7):e1002764.

Natarajan C, Inoguchi N, Weber RE, Fago A, Moriyama H, Storz JF. Epistasis among adaptive mutations in deer mouse hemoglobin. Science. 2013; 340(6138):1324–1327.

Nduati R, John G, Mbori-Ngacha D, Richardson B, Overbaugh J, Mwatha A, Ndinya-Achola J, Bwayo J, Onyango FE, Hughes J, et al. Effect of breastfeeding and formula feeding on transmission of HIV-1: a randomized clinical trial. JAMA. 2000; 283(9):1167–1174.

Olshevsky U, Helseth E, Furman C, Li J, Haseltine W, Sodroski J. Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. Journal of Virology. 1990; 64(12):5701–5707.

Ozorowski G, Pallesen J, de Val N, Lyumkis D, Cottrell CA, Torres JL, Copps J, Stanfield RL, Cupo A, Pugach P, et al. Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. Nature. 2017; 547(7663):360–363.

Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, Huang J, Acharya P, Chuang GY, Ofek G, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature. 2014; 514(7523):455–461.

Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, Zajic L, Iyer SS, Decker JM, Kumar A, et al. Phenotypic properties of transmitted founder HIV-1. Proceedings of the National Academy of Sciences. 2013; 110(17):6626–6633.

Peden K, Emerman M, Montagnier L. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1 LAI, HIV-1 MAL, and HIV-1 ELI. Virology. 1991; 185(2):661–672.

Podgornaia AI, Laub MT. Pervasive degeneracy and epistasis in a protein-protein interface. Science. 2015; 347(6222):673–677.

Pollock DD, Thiltgen G, Goldstein RA. Amino acid coevolution induces an evolutionary Stokes shift. Proceedings of the National Academy of Sciences. 2012; 109(21):E1352–E1359.

Pond SK, Delport W, Muse SV, Scheffler K. Correcting the bias of empirical frequency parameter estimators in codon models. PLoS One. 2010; 5(7):e11230.

Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology. 2004; 53(5):793–808.

Poss M, Overbaugh J. Variants from the diverse virus population identified at seroconversion of a clade A human immunodeficiency virus type 1-infected woman have distinct biological properties. Journal of Virology. 1999; 73(7):5255–5264.

Pugach P, Kuhmann SE, Taylor J, Marozsan AJ, Snyder A, Ketas T, Wolinsky SM, Korber BT, Moore JP. The prolonged culture of human immunodeficiency virus type 1 in primary lymphocytes increases its sensitivity to neutralization by soluble CD4. Virology. 2004; 321(1):8–22.

Pugach P, Ozorowski G, Cupo A, Ringe R, Yasmeen A, de Val N, Derking R, Kim HJ, Korzun J, Golabek M, et al. A native-like SOSIP. 664 trimer based on an HIV-1 subtype B env gene. Journal of Virology. 2015; 89(6):3380–3395.

Rathore U, Saha P, Kesavardhana S, Kumar AA, Datta R, Devanarayanan S, Das R, Mascola JR, Varadarajan R. Glycosylation of the core of the HIV-1 envelope subunit protein gp120 is not required for native trimer formation or viral infectivity. Journal of Biological Chemistry. 2017; 292(24):10197–10219.

Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proceedings of the National Academy of Sciences. 2003; 100(7):4144–4149.

Risso VA, Manssour-Triedo F, Delgado-Delgado A, Arco R, Barroso-delJesus A, Ingles-Prieto A, Godoy-Ruiz R, Gavira JA, Gaucher EA, Ibarra-Molero B, et al. Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. Molecular Biology and Evolution. 2014; 32(2):440–455.

Ronen K, Sharma A, Overbaugh J. HIV transmission biology: translation for HIV prevention. AIDS. 2015; 29:2219–2227.

Sagar M, Wu X, Lee S, Overbaugh J. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. Journal of Virology. 2006; 80(19):9586–9598.

Sander C, Schneider R. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins: Structure, Function, and Bioinformatics. 1991; 9(1):56–68.

Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, Kim HJ, Blattner C, de la Peña AT, Korzun J, 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.

Sanders RW, Van Gils MJ, Derking R, Sok D, Ketas TJ, Burger JA, Ozorowski G, Cupo A, Simonich C, Goo L, et al. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science. 2015; 349(6244):aac4223.

Sanders RW, Vesanen M, Schuelke N, Master A, Schiffner L, Kalyanaraman R, Paluch M, Berkhout B, Maddon PJ, Olson WC, et al. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. Journal of Virology. 2002; 76(17):8875–8889.

Shah P, McCandlish DM, Plotkin JB. Contingency and entrenchment in protein evolution under purifying selection. Proceedings of the National Academy of Sciences. 2015; 112(25):E3226–E3235.

Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harbor Perspectives in Medicine. 2011; 1(1):a006841.

da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE. Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. Genetics. 2010; 185(1):293–303.

Simonich CA, Williams KL, Verkerke HP, Williams JA, Nduati R, Lee KK, Overbaugh J. HIV-1 neutralizing antibodies with limited hypermutation from an infant. Cell. 2016; 166(1):77–87.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30:1312–1313.

Starcich BR, Hahn BH, Shaw GM, McNeely PD, Modrow S, Wolf H, Parks ES, Parks WP, Josephs SF, Gallo RC, et al. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell. 1986; 45(5):637–648.

Starr TN, Flynn JM, Mishra P, Bolon DN, Thornton JW. Pervasive contingency and entrenchment in a billion years of Hsp90 evolution. bioRxiv. 2017; p. 189803.

Starr TN, Thornton JW. Epistasis in protein evolution. Protein Science. 2016; 25(7):1204–1218.

Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, Thomas PV, Wagh K, Zhou T, Behrens AJ, et al. Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. Cell. 2016; 165(4):813–826.

de Taeye SW, Ozorowski G, de la Peña AT, Guttman M, Julien JP, van den Kerkhof TL, Burger JA, Pritchard LK, Pugach P, Yasmeen A, et al. Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. Cell. 2015; 163(7):1702–1715.

Tan K, Liu Jh, Wang Jh, Shen S, Lu M. Atomic structure of a thermostable subdomain of HIV-1 gp41. Proceedings of the National Academy of Sciences. 1997; 94(23):12303–12308.

Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. Maximum allowed solvent accessibilites of residues in proteins. PloS One. 2013; 8(11):e80635.

Verkerke HP, Williams JA, Guttman M, Simonich CA, Liang Y, Filipavicius M, Hu SL, Overbaugh J, Lee KK. Epitope-independent purification of native-like envelope trimers from diverse HIV-1 isolates. Journal of Virology. 2016; 90(20):9471–9482.

Wagih O. ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics. 2017; 33(22):3645–3647.

Wang W, Nie J, Prochnow C, Truong C, Jia Z, Wang S, Chen XS, Wang Y. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology. 2013; 10(1):1.

Wang X, Minasov G, Shoichet BK. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. Journal of molecular biology. 2002; 320(1):85–95.

Waterston R, Lindblad-Toh K, Birney E, Rogers J, Abril J, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002; 420(6915):520–562.

Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, et al. Antibody neutralization and escape by HIV-1. Nature. 2003; 422(6929):307–312.

Weissenhorn W, Dessen A, Harrison S, Skehel J, Wiley D. Atomic structure of the ectodomain from HIV-1 gp41. Nature. 1997; 387(6631):426.

White TA, Bartesaghi A, Borgnia MJ, Meyerson JR, de la Cruz MJV, Bess JW, Nandwani R, Hoxie JA, Lifson JD, Milne JL, et al. Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure. PLoS Pathogens. 2010; 6(12):e1001249.

Wilen CB, Parrish NF, Pfaff JM, Decker JM, Henning EA, Haim H, Petersen JE, Wojcechowskyj JA, Sodroski J, Haynes BF, et al. Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins. Journal of Virology. 2011; 85(17):8514–8527.

Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe JJ, Kabongo JMM, Kalengayi RM, Van Marck E, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature. 2008; 455(7213):661–664.

Wu NC, Young AP, Al-Mawsawi LQ, Olson CA, Feng J, Qi H, Chen SH, Lu IH, Lin CY, Chin RG, et al. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. Scientific Reports. 2014; 4:4942.

Wu X, Parast AB, Richardson BA, Nduati R, John-Stewart G, Mbori-Ngacha D, Rainwater SM, Overbaugh J. Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant. Journal of Virology. 2006; 80(2):835–844.

Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. Journal of Molecular Evolution. 1994; 39(3):306–314.

Yang Z, Nielsen R, Goldman N, Pedersen AMK. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 2000; 155(1):431–449.

Yuste E, Reeves JD, Doms RW, Desrosiers RC. Modulation of Env content in virions of simian immunodeficiency virus: correlation with cell surface expression and virion infectivity. Journal of Virology. 2004; 78(13):6775–6785.

Zanini F, Neher RA. Quantifying selection against synonymous mutations in HIV-1 env evolution. Journal of Virology. 2013; 87(21):11843–11850.

Zolla-Pazner S, Cardozo T. Structure–function relationships of HIV-1 envelope sequence-variable regions refocus vaccine design. Nature Reviews Immunology. 2010; 10(7):527–535.

Supplementary file 1. The code to perform all steps in the analysis is in `analysis_code.zip`. Specifically, this file contains a Jupyter notebook that performs the analysis, all required input data, and all reasonably sized output files. The Jupyter notebook downloads the deep sequencing data, processes it with the `dms_tools2` software (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/), and also performs a variety of downstream analyses that generate most of the figures for this paper.

Supplementary file 2. An HTML rendering of the Jupyter notebook that performs the computational analysis. The actual notebook is in Supplementary file 1, but if you just want to look at the analysis rather than run it, then you may prefer this file instead. In particular, the notebook contains plots detailing the deep sequencing data analysis as generated using the `dms_tools2` software (*Bloom, 2015*, https://jbloomlab.github.io/dms_tools2/).

Supplementary file 3. The sequence of the wildtype BG505 *env* used in our study is in FASTA format in the file `BG505_env.fasta`.

Supplementary file 4. The sequences of the primers used for the BG505 codon mutagenesis are in the file `BG505_codon_mutagenesis_primers.txt`.

Supplementary file 5. The primers used for the BG505 barcoded-subamplicon sequencing are in the file `BG505_bcsubamp_primers.txt`.

Figure 1–Figure supplement 1. The histograms show the pairwise amino-acid identity of each Env to all other sequences in the clade A alignment in **Figure 1**-source data 1 after masking the sites delineated in **Figure 1**-source data 2. There are 616 non-masked sites. The pairwise protein identity between BG505 and BF520 is 86.2% (721 of 836 sites identical) when considering *all* sites, and 89.1% (549 of 616 sites identical) when considering just the non-masked sites.

Figure 2–Figure supplement 1. We Sanger sequenced 44 clones of BG505 Env sampled roughly evenly from each of the three replicate mutant plasmid libraries. **(A)** There was an average of 1.5 mutant codons per clone, with the number of mutations per clone roughly following a Poisson distribution. **(B)** The mutant codons had a mix of single-, double-, and triple-nucleotide changes, with an elevated number of single-nucleotide changes than expected. **(C)** Nucleotide frequencies were fairly uniform in the mutant codons. **(D)** Mutations were distributed roughly evenly along the mutagenized region of *env* (30-699 in the sequential numbering scheme used in this plot). **(E)** For clones with multiple mutations, we computed the pairwise distance in primary sequence between each codon mutation and plotted the cumulative distribution of these distances (red line). We also simulated the expected distribution of pairwise distances if mutations occurred entirely independently (blue line). The observed distribution is close to the expected distribution. Comparable data for the BF520 libraries is provided in *Dingens et al.* (*2017*).

Figure 3–Figure supplement 1. This figure shows the same data as in **Figure 3**B except that it shows KDE contour plots rather than points. As is obvious from this representation, the vast majority of the points fall near the origin (very low preference in both replicates), and most of the correlation signal is therefore due to the relatively modest number of amino acids that have high preference at any given site. This is expected, since most amino acid mutations at most sites will be strongly deleterious, and so have low preference.

Figure 7–Figure supplement 1. The left-most panel shows the six-helix bundle of Env (PDB 1ENV; *Weissenhorn et al., 1997*), focusing on the central three helices, each of which is derived from a different Env monomer and which pack together to form the bundle's core. The other three helices that make up this bundle are located in another part of the structure. There are ~80 residues per monomer that are resolved in this structure, including four significantly shifted sites and five substituted sites. Of these, three of the shifted sites (582, 583, and 587) and one substituted site (588) form a cluster at one end of the bundle. We show the side chains for each of these sites colored according to whether the site has shifted (red) or substituted (yellow). We wondered whether this cluster was unique to the six-helix bundle, but found that it is also present in roughly the same structural arrangement in both the closed pre-fusion conformation (center panel) and the open CD4-bound conformation (right panel; note that some of the surrounding residues in both the closed and open structures have been hidden for the sake of clarity). As shown in **Figure 6**C, each of the shifted sites are more tolerant of mutations in BG505 than in BF520. The two shifted sites that point inward towards the bundle's core (583 and 587) may be important for packing. Based on these structures, it is difficult to discern if or how the substituted site may impact the preferences at the shifted sites nearby.

Figure 7–Figure supplement 2. A fine-grained view of two clusters of shifted sites, both of which are in conformationally dynamic regions of Env. In each structure, we only show side chains for shifted sites or other sites of interest, which include substituted sites and sites that are proposed to take part in a hydrophobic network that helps control Env's dynamics (*Ozorowski et al., 2017*). These side chains are colored according to the key in the upper right. **(A)** A cluster of four shifted sites at the apex of the closed pre-fusion conformation of Env (PDB 5FYL; *Stewart-Jones et al., 2016*), a region that is not resolved in the open CD4-bound conformation. These sites are in Env's first/second variable loop (164 and 165) and its third variable loop (307 and 309), which pack against one another and against an adjacent Env protomer (the inter-protomer boundary is indicated by a dashed line). Each of the shifted sites has substituted. And these sites are in the immediate vicinity of other non-shifted, but substituted sites. Thus, it seems likely that these shifts at least partially arise from short-range epistatic interactions within this cluster. However, longer-range epistatic interactions with more distant sites also seem plausible given the highly dynamic nature of this region. **(B)** A cluster of shifted sites that are adjacent to a network of hydrophobic residues (blue/purple) that help mediate the conformational change between the closed pre-fusion state to the open CD4-bound state (*Ozorowski et al., 2017*). One of the shifted sites (purple) is in this network. These sites are shown in context of both structural states, the latter of which is from (PDB 5VN3; *Ozorowski et al., 2017*). Nearly one third of the shifted residues cluster within this structural region (9/30). In contrast to panel (A), only a few of the shifted sites are adjacent to substitutions or have substituted themselves. This trend is found in both the closed and open conformations, suggesting that the shifts may be primarily due to long-range epistatic interactions.

Figure 9–Figure supplement 1. Sites are grouped by whether they have $\omega_r > 1$ (diversifying selection) or $\omega_r < 1$ (purifying selection) at $Q < 0.05$ in *both* Envs, or whether they fall into neither of these categories. Relative solvent accessibilities were calculated as in **Figure 7**. As can be seen from these box plots with overlaid points for each site, sites of both diversifying and purifying selection tend to have higher relative solvent accessibility than other sites.

Figure 9–Figure supplement 2. This plot shows all sites that are N-linked glycosylation motifs in both BG505 and BF520. Each motif is named by the residue number of the asparagine, and the preferences (averaged across replicates) for each Env are shown as well as the frequencies of amino acids across the clade A alignment at each of the three positions in the motif (N-X-S/T). As can be seen from this plot, the experiments measure relatively broad mutational tolerance at many sites where the natural Env sequences have a strongly conserved motif. We suspect this is because many glycans serve as a shield against immunity in nature—a function that is not required in cell culture.

Figure 9–Figure supplement 3. This plot shows all sites that are evolving more slowly than expected in natural sequences given the preferences measured in both Envs. Specifically, it shows all sites with $\omega_r < 1$ at $Q < 0.05$ for the ExpCMs for *both* BG505 and BF520. For each site, the plots show the preferences averaged across replicates and re-scaled for each Env, as well as the frequencies of amino acids in the clade A Env alignment. The Q-value indicated is the maximum of that for the two Envs.