Inference of epistatic effects leading to entrenchment and drug resistance in HIV-1 protease

William F. Flynn,[1,2] Allan Haldane,[2,3] Bruce E. Torbett,[4] and Ronald M. Levy[2,3,*]

[1] *Department of Physics and Astronomy, Rutgers University*

[2] *Center for Biophysics and Computational Biology, Temple University*

[3] *Department of Chemistry, Temple University*

[4] *Department of Molecular and Experimental Medicine, The Scripps Research Institute*

(Dated: November 2, 2016)

Abstract

Understanding the complex mutation patterns that give rise to drug resistant viral strains provides a foundation for developing more effective treatment strategies for HIV/AIDS. Multiple sequence alignments of drug-experienced HIV-1 protease sequences contain networks of many pair correlations which can be used to build a (Potts) Hamiltonian model of these mutation patterns. Using this Hamiltonian model we translate HIV protease sequence covariation data into quantitative predictions for the probability of observing specific mutation patterns which are in agreement with the observed sequence statistics. We find that the statistical energies of the Potts model are correlated with the fitness of individual proteins containing therapy-associated mutations as estimated by *in vitro* measurements of protein stability and viral infectivity. We show that the penalty for acquiring primary resistance mutations depends on the epistatic interactions with the sequence background. Primary mutations which lead to drug resistance can become highly advantageous (or entrenched) by the complex mutation patterns which arise in response to drug therapy despite being destabilizing in the wildtype background. Anticipating epistatic effects is important for the design of future protease inhibitor therapies.

* Corresponding author: ronlevy@temple.edu

I. INTRODUCTION

The ability of HIV to rapidly mutate leads to antiretroviral therapy (ART) failure among infected patients. Enzymes coded by the *pol* gene play critical roles in viral maturation and have been key targets of several families of drugs used in combination therapies. The protease enzyme is responsible for the cleavage of the Gag and Gag-Pol polyproteins into functional constituent proteins and it has been estimated that resistance develops in as many as 50% of patients undergoing monotherapy (Richman et al. 2004) and as many as 30% of patients undergoing modern combination antiretroviral therapy (c-ART) (Gupta et al. 2008).

The combined selective pressures of the human immune response and antiretroviral therapies greatly affect the evolution of targeted portions of the HIV-1 genome and give rise to patterns of correlated amino acid substitutions. As an enzyme responsible for the maturation of the virion, the mutational landscape of HIV protease is further constrained due to function, structure, thermodynamics, and kinetics (Bloom et al. 2010, Haq et al. 2012, Lockless et al. 1999, Zeldovich and Shakhnovich 2008, Zeldovich et al. 2007). As a consequence of these constraints, complex mutational patterns often arise in patients who have failed c-ART therapies containing protease inhibitors (PI), with mutations located both at critical residue positions in or near the protease active site and others distal from the active site (Chang and Torbett 2011, Flynn et al. 2015, Fun et al. 2012, Haq et al. 2012). In particular, the selective pressure of PI therapy gives rise to patterns of strongly correlated mutations generally not observed in the absence of c-ART, and more therapy-associated mutations accumulate under PI therapy than under all other types of ART therapies (Shafer 2006, Shafer and Schapiro 2008, Wu et al. 2003). In fact, the majority of drug-experienced subtype B protease sequences in the Stanford HIV Drug Resistance Database (HIVDB) have more than 4 PI-therapy-associated mutations (see Figure S1). Within the Stanford HIVDB are patterns of multiple resistance mutations, and in order to overcome the development of resistance, understanding these patterns is critical.

A mutation's impact on protein stability or fitness depends on the genetic background in which it is acquired. Geneticists call this phenomenon "epistasis". It is well understood that major drug resistance mutations in HIV protease destabilize the protease in some way, reducing protein stability or enzymatic activity, which can greatly alter the replicative and transmissive ability, or *fitness*, of that viral strain (Bloom et al. 2010, Boucher et al. 2016, Grenfell et al. 2004, Wang et al. 2002). To compensate for this fitness loss, protease accumulates accessory mutations which have been shown to restore stability or activity (Chang and Torbett 2011, Fun et al. 2012, Martinez-Picado et al. 1999). But it is unclear how the acquisition and impact of primary and accessory mutations are modulated in the presence of the many different genetic backgrounds observed, especially those present in the complex

resistant genotypes that arise under inhibitor therapy.

Coevolutionary information derived from large collections of related protein sequences can be used to build models of protein structure and fitness (Burger and van Nimwegen 2010, Göbel et al. 1994, Hinkley et al. 2011, Liu et al. 2009, Lockless et al. 1999, Socolich et al. 2005). Given a multiple sequence alignment (MSA) of related protein sequences, a probabilistic model of the network of interacting protein residues can be inferred from the pair correlations encoded in the MSA. Recently, probabilistic models, called Potts models, have been used to assign scores to individual protein sequences which correlate with experimental measures of fitness (Ferguson et al. 2013, Figliuzzi et al. 2015, Haq et al. 2012, Hopf et al. 2015, Mann et al. 2014). These advances build upon previous and ongoing work in which Potts models have been used to extract information from sequence data regarding tertiary and quaternary structure of protein families (Barton et al. 2016a, Haldane et al. 2016, Jacquin et al. 2016, Marks et al. 2012, Morcos et al. 2011, 2014, Sulkowska et al. 2012, Sutto et al. 2015, Weigt et al. 2009) and sequence-specific quantitative predictions of viral protein stability and fitness (Barton et al. 2016b, Butler et al. 2016, Haq et al. 2012, Shekhar et al. 2013).

In this study, we show how such models can be constructed to capture the epistatic interactions involved in the evolution of drug resistance in HIV-1 protease. The acquisition of resistance mutations which accumulate under the selective pressure of inhibitor therapy leave many residual correlations observable in MSAs of drug-experienced sequences (Hoffman et al. 2003, Rhee et al. 2007, Wu et al. 2003), and we use the pair correlations that can be extracted from MSAs to construct a Potts model of the mutational landscape of drug experienced HIV-1 protease. We first provide several tests which demonstrate that our inferred model faithfully reproduces several key features of our original MSA including higher order correlations. We then compare the Potts model statistical energies with experimental measurements of fitness, including structural stability and relative infectivity of individual HIV protease variants which contain resistance mutations. Finally, the Potts scores are used to describe the epistatic mutational landscape of three primary resistance mutations. We observe strong epistatic effects. The primary mutations are destabilizing in the context of the wildtype background, but become stabilizing on average as other resistance mutations accumulate in the background, similar to the concept of entrenchment in systems biology (Gong et al. 2013, Pollock et al. 2012, Shah et al. 2015). Furthermore, we find that entrenchment is modulated by the collective effect of the entire sequence, including mutations at polymorphic residues, and the variance of the statistical energy cost of introducing a primary mutation increases as resistance mutations accumulate; this heterogeneity is another manifestation of epistasis (Barton et al. 2016b, McCandlish et al. 2015). These findings provide a framework for exploring mutational resistance mechanisms using probabilistic models.

II. RESULTS

A. Model inference and dataset

Given a multiple sequence alignment (MSA), we can infer a statistical model $P(\vec{\sigma})$ for the probability of finding a protein sequence with sequence identity $\vec{\sigma}$ which takes the form $P(\vec{\sigma}) \propto \exp(-E(\vec{\sigma}))$ from the statistical properties of the MSA. The maximum entropy model which reproduces the first and second order marginal distributions of the MSA, $P_i(\sigma_i)$ and $P_{ij}(\sigma_i, \sigma_j)$ of residue positions i and position pairs i, j, is given by the Potts Hamiltonian $E(\vec{\sigma}) = \sum_i h_i(\sigma_i) + \sum_{i<j} J_{ij}(\sigma_i, \sigma_j)$, where the fields $h_i(\sigma_i)$ and couplings $J_{ij}(\sigma_i, \sigma_j)$ represent the preference for residue σ_i at position i and residue pair $\sigma_i\sigma_j$ at positions i, j, respectively. The Potts model is fit to the bivariate marginals of the MSA such that it recovers the correlated pair information $C_{ij}(\sigma_i, \sigma_j) = P_{ij}(\sigma_i, \sigma_j) - P_i(\sigma_i)P_j(\sigma_j)$.

The Potts model captures epistatic effects; in contrast an independent model of a multiple sequence alignment can be constructed by summing the logarithm of the univariate marginals $E_{ind}(\vec{\sigma}) = \sum_i \log P_i(\sigma_i)$. As described in the following section, the ability to reproduce higher order marginals of the MSA (beyond second order) is a true predictive test of the Potts model, one which the independent model fails.

As described in the introduction, protease sequence evolution under protease inhibitor selective pressure produces correlations between amino acid substitutions that are larger in magnitude than those that occur in the absence of drug pressure (seen in Figure S2) (Gupta and Adami 2016, Rhee et al. 2007, Wu et al. 2003). Although correlations between some drug-associated sites can be identified through analysis of drug-naive sequences, or structural and/or evolutionary constraints (Butler et al. 2016, Hoffman et al. 2003), the most complete model of the epistatic landscape of drug-resistance mutations is constructed using the correlations found in a varied set of drug-experienced sequences. As we demonstrate in later sections, correlations among the primary, accessory, and polymorphic mutations which arise under c-ART therapy all contribute to protease fitness. Starting with an MSA constructed from 5610 HIV-1 subtype B drug-experienced protease sequences obtained from the Stanford HIVDB, we have inferred a Potts model using a Markov Chain Monte Carlo (MCMC) method implemented on GPUs (see Materials and Methods and the supplemental information of Haldane et al. (Haldane et al. 2016) for more details).

B. Recovery of the observed sequence statistics – marginal probabilities

We can gauge the accuracy of the model by examining how well the model reproduces various statistics of the MSA. The most direct test is the reproduction of higher order

correlations observed in the multiple sequence alignment beyond pair correlations. Shown in Figure 1A is the recovery of the marginal probabilities of the most common subsequences observed in the dataset across varying subsequence lengths. The recovery of the bivariate marginals (pair frequencies) is not predictive but it demonstrates the quality of fit of the Potts model. The results shown in Figure 1 demonstrate that the Potts model is able to predict the frequencies of higher order marginals with accuracy. The Pearson correlation coefficient for the observed probabilities compared with the Potts model prediction remains above $R^2 \geq 0.95$ for subsequence lengths as large as 14. In contrast the independent model correlation coefficient is significantly worse ($R^2 \to 0.22$).

Figure 2 shows the probability distribution of sequences that differ from the consensus by k mutations as predicted by the Potts and independent models compared to the observed distribution derived from the MSA. The Potts model predicts a distribution of mutations per sequence which is very close to the observed distribution whereas the independent model incorrectly predicts a multinomial distribution centered about 8 mutations from consensus. The very good agreement between the higher order sequence statistics of the Potts model and the observed statistics from the MSA provides additional evidence that the Potts model is not overfit.

The Potts model also captures the observed statistics for larger subsequences, but as subsequence lengths increase, observed marginal probabilities in our MSA approach the sampling limit of the alignment ($1/N \approx 2 \times 10^{-4}$), meaning comparisons with the observed data at this level become dominated by noise. Tests with synthetic data (not shown) confirm that for longer subsequences the discrepancy between observed higher order marginals and the Potts model are consistent with effects caused by the finite sample size (5610 sequences) of the MSA (Haldane et al. 2016). In the following section, we compare Potts model statistical energies with experimentally determined measurements of protease fitness.

C. Protease mutations, protein stability, and replicative capacity

Two experimental tests used to quantify the effects of protease mutations on viral fitness are thermal stability of the folded protein and replicative capacity (Chang and Torbett 2011, Louis et al. 2011, Muzammil et al. 2003). Chang and Torbett demonstrate that stability is compromised by the acquisition of primary mutations and this loss of stability can be rescued by known compensatory mutations, sometimes in excess of the reference stability. Muzammil et al. and Louis et al. have shown that patterns of up to 10 or more resistance mutations do not necessarily suffer from reduced stability relative to the wildtype, and that non-active site mutations can lead to resistance in certain sequence contexts. In Figure 3A the change in statistical Potts energies, $\Delta E = E - E_{ref}$ is plotted versus the

change in thermal stability, where E and E_{ref} are the statistical energies of the mutated and reference sequences corresponding to each pair of stability measurements. We observe a strong correlation between Potts ΔE and the change in stability as reflected by the change in melting temperature ($R = -0.85$, $p = 0.0003$). In contrast, the change in stability computed using the independent model shows no correlation (Figure S3A).

We have extracted results for viral replicative capacity in which 29 single Protease mutants were studied by Henderson et al. (Henderson et al. 2012) and an additional small set of more complex sequence variants (van Maarseveen et al. 2006) that were tested relative to the wildtype sequence. As with the stability measurements, we find the relative Potts energy correlates well with infectivity ($r = -0.64$, $p < 10^{-5}$), shown in Figure 3B. The same comparison using the independent model computed fitness again shows no predictive power (Figure S3B). Complementary to the RC assay presented in their study, Henderson et al. presented a SpIn assay and an additional assay measuring drug concentrations which inhibit protease function (EC50). Potts fitness predictions against these data are shown in Figure S4. While this additional comparison does not show statistically significant correlation, probably because the observed measurements span a much smaller range of values, they do exhibit the same negative trends observed in Figure 3. All data shown in Figures 3, S3, S4 can be found in Supplementary Data 1.

The results presented here are reinforced by other recent studies of protein evolutionary landscapes (Ferguson et al. 2013, Figliuzzi et al. 2015, Hopf et al. 2015, Mann et al. 2014) where varying measures of experimental fitness are compared to statistical energies derived from correlated Potts models constructed from multiple sequence alignments. The range of statistical energies and the correlation with fitness are qualitatively similar to those presented by Ferguson et al. and Mann et al. where statistical energies of engineered HIV-1 Gag variants generated using a similar inference technique are compared with replicative fitness assays. The same can be said for correlations between Potts scores and relative folding free energies of Beta Lactamase TEM-1 presented by Figliuzzi et al.. This collection of studies demonstrate that Potts model statistical energies correlate with the fitness of protein sequences in different contexts, including protein families evolving under weak selection pressure (Figliuzzi et al. 2015, Hopf et al. 2015), viral proteins evolving under immune pressure (Ferguson et al. 2013, Mann et al. 2014), and as presented here, viral proteins evolving under drug pressure.

D. Inference of epistasis among therapy-associated mutations

The sequences present in the Stanford HIVDB have been deposited at many stages of HIV infection and treatment, showcasing a variety of resistance patterns spanning from wildtype to patterns of more than 15 mutations at PI-associated positions. In this section,

we describe how Potts statistical energies can be used to infer epistatic effects on the major HIV protease resistance mutations.

Although all current PIs are competitive active site inhibitors, major resistance mutations can be found both inside and outside of the protease active site; the substrate envelope hypothesis suggests this arises because PIs have a larger interaction surface with protease compared to that of its natural substrates (King et al. 2004, Özen et al. 2011, Prabu-Jeyabalan et al. 2002). V82 and I84 are positions inside the substrate cleft and major resistance mutations V82A and I84V have been shown to directly affect binding of inhibitors (Chellappan et al. 2007, King et al. 2002, Lefebvre and Schiffer 2008). L90 is a residue located outside of the substrate cleft and flap sites. Mutations at position 90, specifically L90M, have been shown to allow shifting of the aspartic acids of the active site catalytic triad (D25) on both chains, subsequently allowing for larger conformational changes at the dimer interface and active site cleft that reduce inhibitor binding (Kovalevsky et al. 2006, Mahalingam et al. 2004, Ode et al. 2006).

Given a sequence containing one of the 3 mutants V82A, I84V, and L90M, we can determine the context-dependence of that mutation in its background by calculating the change in statistical energy associated with reversion of that mutation back to wildtype. This corresponds to computing $\Delta E = E_{obs} - E_{rev}$ where E_{obs} is the Potts energy of an observed sequence with one of these primary mutations and E_{rev} is the Potts energy of that sequence with the primary mutation reverted to its consensus amino acid type. Due to the pairwise nature of the Potts Hamiltonian, this computation reveals a measure of epistasis for a sequence $\vec{\sigma}$ containing mutant $X \to Y$ at position k

$$\Delta E(\vec{\sigma}_{k,Y}) = h_k(Y) - h_k(X) + \sum_{i \neq k} (J_{ik}(\sigma_i, Y) - J_{ik}(\sigma_i, X)) \tag{1}$$

where the pair terms J_{ik} are the couplings between the mutation site and all other positions in the background. When this measure is positive, the background imparts a fitness penalty for the reversion of the primary resistance mutation to the wildtype and when negative, the sequence regains fitness with reversion to wildtype. Using this measure, we computed ΔE for every sequence in our HIVDB MSA containing V82A, I84V, L90M and have arranged the energies versus sequence hamming distance from the consensus including only PI-associated sites, shown in Figure 4. As more mutations accumulate in the background, the preference for each primary resistance mutation to revert to wildtype is lost and the primary mutation becomes preferred over the wildtype on average when enough background mutations have accumulated. These crossover points are 6, 9, and 7 mutations for V82A, I84V, and L90M, respectively. When a sufficient number of mutations have accumulated, the primary resistance mutation becomes *entrenched*, meaning a reversion to wildtype at that position

is destabilizing in most sequences; the primary mutation becomes more entrenched as more background mutations are acquired. The effect is largest for L90M; for sequences containing > 7 PI-associated mutations, on average the L90M primary mutation is $\approx$100 times more likely than the wildtype leucine at position 90. In contrast, this primary mutation is $\approx$80 times less likely than the wildtype residue in the subtype B consensus sequence background. In other words, there is an $\approx 8,000$ fold difference in the probability of observing the mutation L90M depending on the background sequence. The trend shared for V82A, I84V, and L90M is representative of the larger class of primary mutations; mutations V32I, M46L, I47V, G48V, I50V, I54V, L76V, and others become less destabilizing as the number of background mutations increases (see Figure S5).

Why are primary resistance mutations much more likely in some backgrounds and not others? Are these effects caused by a small set of epistatic interactions with the primary resistance mutation or the collective effect of many small epistatic interactions?

To answer these questions, we compared the sequence backgrounds which most entrench primary mutations from those sequences which most prefer wildtype instead of the primary mutation. Using as an example a fixed hamming distance of 10 from the subtype B consensus sequence, we examined the differences between the sequences among the top 10% and bottom 10% of ΔE values in the $h = 10$ column in each of the subplots of Figure 4. The $h = 10$ column was chosen as it is the column with the most data for the primary mutations V82A, I84V, and L90M. These two groups of sequences, top 10% and bottom 10%, are referred to as "most entrenched" (ME) and "least entrenched" (LE) sequences, respectively.

One might expect that the accumulation of accessory mutations in a sequence will lead to the entrenchment of a primary mutation and, under this assumption, the most entrenched sequences should contain more accessory mutations than the least entrenched sequences. We observe more accessory mutations in the most entrenched sequences on average, but the difference is not significant and a large number of accessory mutations accumulate in the least entrenching sequences for V82A, I84V, and L90M as shown in Figure 5. In other words, simply counting accessory mutations in a sequence is unlikely to predict whether that sequence will entrench a primary mutation.

Previous research has identified significant correlations between various primary and accessory mutations and the primary resistance mutations under study here (Flynn et al. 2015, Rhee et al. 2007, Wu et al. 2003). We find that the presence of these accessory mutations alone cannot account for the separation of the most entrenched sequences from the least entrenched sequences. The most striking example is the double mutant G73S-L90M. G73S is present in 75% of the ME sequences and never present in the LE sequences; however, reversion of G73S in the sequences with the double mutation only results in a shift of ΔE equivalent to 15% of the difference between the mean ΔEs in the ME and LE sequences.

This suggests that while G73S certainly helps to entrench L90M, it is not required for the entrenchment of L90M and is not solely responsible for the entrenchment of L90M when present. Similar effects are observed for mutation I54V in the entrenchment of V82A and M46I and L90M in the entrenchment of I84V.

To uncover the clearest patterns of mutations that differentiate the LE sequences from the ME sequences, we performed principal component analysis (PCA) on the combined set of ME and LE sequences at PI-associated sites. The projections of the ME and LE sequences onto the first 3 principal components are shown in Figures 6 and S6. The first three principal components explain approximately 40% of the total variance when performed on the data corresponding to V82A, I84V, L90M (39.5%, 42.5%, 37.4% respectively). In the case of L90M, the first principal component clearly separates the most entrenched sequences from the least entrenched sequences while the second principal component separates variation within both groups. For V82A and I84V, a linear combination of the first two principal components separates the ME from the LE sequences, most likely due to variation between and within the most and least entrenching sequences being similarly large (which can be seen in the plots of hamming distance in Figure S6).

Examination of the first principal component (PC) eigenvector shows that the residues of at least 11 PI-associated sites contribute to the differentiation of the most entrenched (ME) sequences from the least entrenched (LE) sequences for primary mutation L90M, with residues K20F/I/V, M46I, G73S, V82V, and I84V contributing most strongly. Sequences from the two classes for which the first PC explains the most variation, measured as the hamming distance captured by the first PC, can be found in Table S1. Contributions from 11 sites is consistent with the average pairwise hamming distance of 11 between the most and least entrenched sequences, as seen in Figure 6 inset. Similarly, sets of 14 and 16 residues among the first two principal eigenvectors are responsible for the separation of ME and LE sequences for V82A and I84V, respectively (see Figure S6). These observations reinforce the point that while previously identified primary-accessory mutation pairs are important for acquisition and fixation of primary mutations, a model which captures epistatic effects collectively, like the Potts model, is needed to identify sequence backgrounds most likely to accommodate primary mutations.

Non-PI-associated polymorphisms also appear to modulate the entrenchment of primary resistance mutations, though the effect is secondary to that of PI-associated mutations. There exist sets of sequences, each with the same pattern of PI-associated mutations, that differ in entrenchment scores by as much as $\Delta\Delta E \approx 3$, which corresponds to observable probabilities differing by more than an order of magnitude. This appears to be the result of strong positive and negative couplings that arise between non-PI-associated polymorphisms and certain PI-associated mutations. For example, we find that non-PI-associated muta-

tions V11I, K43R/N, I66V, C67F/L/Q/E, I72V/L, T74A, P79A, and C95F all appear to regulate the entrenchment of L90M. Some of these residues lie in the hydrophobic core of the protease dimer, and subtle conformational changes in the hydrophobic core by these residues may play an important role in inhibitor binding (Mittal et al. 2012). A demonstration of this modulation is shown in Figure S7, where a common background sequence of 10 PI-associated mutations is shared by several observed sequences in the original MSA with varying number of additional polymorphisms. Two of these sequences are shown in Figure S7B, and contain one and six additional mutations respectively. Despite the complicated network of interactions, the presence of the additional five polymorphic mutations in the second sequence increases the entrenchment of L90M, with $\Delta\Delta E = 2.39$ when reverting L90M to L, which corresponds to ~ 10 fold increase in frequency.

These results present testable predictions, and we have included three pairs of sequences that we predict will be most and least entrenching for the primary mutations discussed here, which can be found in Table I. Using either replicative capacity or melting temperature as a proxy for fitness, it should be possible to verify experimentally whether the Potts model correctly predicts the relative frequencies upon reverting the primary mutation to wildtype for the selected sequences pairs listed in Table I.

III. DISCUSSION

The evolution of viruses under drug selective pressure induces mutations which are correlated due to constraints on structural stability and function that contribute to fitness. The correlations induce epistatic effects, a primary or accessory resistance mutation can be either stabilizing or destabilizing depending on the genetic background. Recently epistasis has become a focus for analysis in structural biology and genomics as researchers have begun to successfully link the coevolutionary information in collections of protein sequences with the structural and functional fitness of those proteins (Barton et al. 2016b, Butler et al. 2016, Ferguson et al. 2013, Figliuzzi et al. 2015, Hinkley et al. 2011, Hopf et al. 2015, Mann et al. 2014). In the current study, we have used the correlated mutations encoded in a multiple sequence alignment of drug-experienced HIV-1 protease sequences to parametrize a Potts model of sequence statistical energies that can be used as an estimator of stability and relative replicative capacity of individual protease sequences containing drug resistance mutations.

The most entrenching sequences are those at local fitness maxima, and accumulating mutations, as seen here as increasing hamming distance from the subtype B consensus sequence, unlocks pathways to these local fitness maxima (Gupta and Adami 2016). Up to 100–1000 times more probable than sequences that favor reversion to the consensus genotype,

these highly resistant sequences observed in our MSA present a significant risk for the transmission of drug resistance to new hosts as they incur large fitness penalties for reversion. Indeed, we find that the entrenchment effect is strongest for L90M, which has been shown to revert very slowly in drug naive patients with transmitted drug resistance (Yang et al. 2015).

This work builds upon a large literature, ranging from experimental work (Chang and Torbett 2011, Henderson et al. 2012) and statistical analyses of covarying pairs of mutations (Rhee et al. 2007, Wu et al. 2003) to more advanced statistical models of patterns of mutations at many positions (such as Potts models) (Butler et al. 2016, Haq et al. 2009, 2012), to strengthen our understanding of the emergent properties of drug resistance in HIV-1 protease. We demonstrate that, while very important, the information conveyed by pairs of primary and accessory mutations only tells a small part of the story; the context of the full sequence background is really necessary to understand how primary resistance mutations become stabilized. The results presented here advance recent work in the field of using Potts models to study HIV evolution (Barton et al. 2016b, Butler et al. 2016) by providing systematic prospective predictions quantifying the influence of specific multi-residue patterns on the tolerance of drug resistance mutations.

Recent publications have reported that mutations near or distal to Gag cleavage sites play a role in promoting cleavage by drug-resistant and enzymatically deficient proteases, by selecting for mutations that increase substrate contacts with the protease active site, altering the flexibility of the cleavage site vicinity, or by as of yet unknown mechanisms (Breuer et al. 2011, Flynn et al. 2015, Fun et al. 2012, Kolli et al. 2009, Parry et al. 2011, Prabu-Jeyabalan et al. 2002). This suggests that viral coevolution of Gag with selective protease mutations may further stabilize multiple resistance mutations; thus, the analysis of protease mutation patterns can be extended to include amino acid substitutions within Gag and the Gag-Pol polyprotein. Furthermore, this type of analysis is not limited to protease and may be used to study the development of resistance in other HIV drug targets, such as reverse transcriptase and integrase, as well as other biological systems that develop resistance to antibiotic or antiviral therapies.

The Potts model is a powerful tool for interrogating protein fitness landscapes as it captures the correlated effects of many mutations collectively. The analysis presented here provides a framework to examine the structural and functional fitness of individual viral proteins under drug selection pressure. Elucidating how patterns of viral mutations accumulate and understanding their epistatic effects has the potential to impact design strategies for the next generation of c-ART inhibitors and therapies.

IV. MATERIALS AND METHODS

Sequence Data

Sequence information (as well as patient and reference information) was collected from the Stanford University HIV Drug Resistance Database (`http://hivdb.stanford.edu`) (Shafer 2006) using the Genotype-Rx Protease Downloadable Dataset (`http://hivdb.stanford.edu/pages/geno-rx-datasets.html`) that was last updated on 29/04/2013 (there now exists a more recent sequence alignment updated May 2015). There are 65,628 protease isolates from 59,982 persons in this dataset. From this dataset, 5,824 drug-experienced, subtype B, non-mixture, non/recombinant, and unambiguous sequences were extracted. Sequences with more than 1 gap and MSA columns with more than 1% gaps (positions 1–5 and 99) were removed, resulting in $N = 5,610$ sequences of length $L = 93$.

For the comparison made in Figure S2, drug-naive subtype B non/mixture, non-/recombinant, and unambiguous sequences were extracted from the same downloadable dataset. As with drug/experienced sequences, gap-containing sequences and columns were removed, resulting in 13,350 sequences of length 89.

Mutations considered PI-associated were extracted from (Johnson et al. 2013) and `https://hivdb.stanford.edu/dr-summary/resistance-notes/PI/`:
L10I/F/V/C/R, V11I, G16E, K20R/M/I/T/V, L24I, D30N, V32I, L33I/F/V,
E34Q, M36I/L/V, K43T, M46I/L, I47V/A, G48V, I50L/V, F53L/Y, I54V/L/A/M/T/S,
Q58E, D60E, I62V, L63P, I64L/M/V, H69K/R, A71V/I/T/L, G73S/A/C/T, T74P, L76V,
V77I, V82A/F/T/S/L/I, N83D, I84V, I85V, N88D/S, L89I/M/V, L90M, I93L/M.

Marginal Reweighting

Weights (w^k) reciprocal to the number of sequences contributed by each patient were computed and assigned to each sequence. With these weights, estimates of the bivariate marginal probabilities were computed from the MSA of N sequences:

$$P_{ij}(\sigma_i, \sigma_j) = \frac{1}{N} \sum_{k=1}^{N} w^k \delta\big(\sigma_i^k, \sigma_i\big) \delta\big(\sigma_j^k, \sigma_i\big) \tag{2}$$

where σ_i^k is the residue identity at position i of the kth sequence $\vec{\sigma}^k$, $0 < w^k \leq 1$ is the weight of sequence k, and delta $\delta(\alpha, \beta)$ equals one if $\alpha = \beta$ and is otherwise zero.

Otherwise, all sequences are assumed independent; no reweighting was done to account for shared ancestry among these sequences. Phylogenetic trees of drug-naive and drug-treated HIV-infected patients have been show to exhibit star-like phylogenies (Gupta and

Adami 2016, Keele et al. 2008), and thus phylogenetic corrections are not needed. Further, phylogenetic corrections based on pairwise sequence similarity cut-offs of 40% of sequence length or more as are common in studies utilizing direct coupling analysis (DCA) (Morcos et al. 2011, 2014, Weigt et al. 2009) of protein families would drastically reduce the number of effective sequences in our MSA and would lead to mischaracterization of the true underlying mutational landscape. Potts models of other HIV protein sequences under immune pressure have been parameterized with no phylogenetic corrections (Barton et al. 2016b, Ferguson et al. 2013, Mann et al. 2014).

Alphabet Reduction

It has been shown that "reduced alphabets" consisting of 8 or 10 groupings of amino acids capture most of the information contained in the full 20 letter alphabet (Murphy et al. 2000). We expand on this notion by computing an alphabet reduction that has the least effect on the statistical properties of our MSA. In the context of model building, a reduced alphabet decreases the number of degrees of freedom to be modeled. This leads to a more efficient model inference (Barton et al. 2016a, Haldane et al. 2016).

Given the empirical bivariate marginal distribution for each pair of positions in the MSA using 21 amino acid characters (20 + 1 gap), the procedure begins by selecting a random position i. All possible alphabet reductions from 21 to 20 amino acid characters at position i are enumerated for every pair of positions ij, where $j \neq i$, by summing the bivariate marginals corresponding to each of the 210 possible combinations of amino acid characters at position i. The reduction which minimizes the root square mean difference (RMSD) in mutual information (MI) content:

$$\sqrt{\frac{1}{N}\sum_{ij}\left(\mathrm{MI}_{ij}^{Q=21} - \mathrm{MI}_{ij}^{Q=Q'}\right)^2} \tag{3}$$

between all pairs of positions ij with the original alphabet size $Q = 21$ and reduced alphabet size $Q = 20$ is selected. The alphabet at each position i is reduced in this manner until all positions have position-specific alphabets of size $Q = 20$. This process is then repeated for each position by selecting the merger of characters which minimizes the RMSD in MI between all pairs of positions ij with the original alphabet size $Q = 21$ and reduced alphabet size $Q = Q'$, and is stopped once $Q = 2$.

Due to residue conservation at many loci in the HIV protease genome, the average number of characters per position is 2, and several previous studies of HIV have used a binary alphabet to extract meaningful information from sequences (Ferguson et al. 2013, Flynn et al. 2015, Shekhar et al. 2013, Wu et al. 2003). However, using a binary alphabet marginalizes

potentially informative distinctions between amino acids at certain positions, especially PI-associated sites, that acquire multiple mutations from the wildtype. We found that an alphabet of 4 letters substantially reduces the sequence space to be explored during the model inference while providing the necessary discrimination between different types of mutant residues at each position. Additionally, the information lost in this reduction is minimal; Pearson's R^2 between the mutual information (MI) of the bivariate marginal distributions in 21 letters and in 4 letters is ≈0.995 (Figures S8, S9).

The original MSA was then re-encoded using the reduced per-position alphabet, and the bivariate marginals (Eq. 2) were recalculated using the reduced alphabet. Small pseudocounts are added to the bivariate marginals, as described (Haldane et al. 2016). Briefly, instead of adding a small flat pseudocount such as $1/N$, we add pseudocounts which correspond to a small per-position chance μ of mutating to a random residue such that the pseudocounted marginals P^{pc} are given by

$$P_{ij}^{pc}(\sigma_i, \sigma_j) = (1-\mu)^2 P_{ij}(\sigma_i, \sigma_j) + \frac{(1-\mu)\mu}{Q}(P_i(\sigma_i) + P_j(\sigma_j)) + \frac{\mu^2}{Q^2} \tag{4}$$

where we take $\mu \approx 1/N$.

Maximum Entropy Model

Following (Mora and Bialek 2011), we seek to approximate the unknown empirical probability distribution $P(\vec{\sigma})$ which describes HIV-1 protease sequences $\{\vec{\sigma}\}$ of length L where each residue is encoded in an alphabet of Q states by a model probability distribution $P^m(\vec{\sigma})$. The model distribution we choose is the maximum entropy distribution, e.g. the distribution which maximizes

$$S = -\sum_{k=1}^{Q^L} P^m\left(\vec{\sigma}^k\right) \log P^m\left(\vec{\sigma}^k\right) \tag{5}$$

and has been derived by (Barton et al. 2016a, Ferguson et al. 2013, Mézard and Mora 2009, Morcos et al. 2011, Weigt et al. 2009) and others satisfying the following constraints:

$$\sum_{k}^{Q^L} P^m\left(\vec{\sigma}^k\right) = 1 \tag{6}$$

$$\sum_{k}^{Q^L} P^m\left(\vec{\sigma}^k\right)\delta\left(\sigma_i^k, \sigma_i\right) = P_i(\sigma_i) \tag{7}$$

$$\sum_{k}^{Q^L} P^m\left(\vec{\sigma}^k\right)\delta\left(\sigma_i^k, \sigma_i\right)\delta\left(\sigma_j^k, \sigma_j\right) = P_{ij}(\sigma_i, \sigma_j) \tag{8}$$

i.e. such that the empirical univariate and bivariate marginal probability distributions are preserved. Through a derivation using Lagrange multipliers not presented here (but can be found in (Ferguson et al. 2013, Mora and Bialek 2011)), the maximum entropy model takes the form of a Boltzmann distribution

$$P^m(\vec{\sigma}) = \frac{1}{Z}\exp\left(-\beta E(\vec{\sigma})\right) \tag{9}$$

$$E(\vec{\sigma}) = \sum_{i}^{L} h_i(\sigma_i) + \sum_{i<j}^{L(L-1)/2} J_{ij}(\sigma_i, \sigma_j) \tag{10}$$

where the quantity $E(\vec{\sigma})$ is the Potts Hamiltonian, which determines the statistical energy of a sequence $\vec{\sigma}$, $1/Z$ is a normalization constant, and the inverse temperature $\beta = 1/k_B T$ is such that $k_b T = 1$. This form of the Potts Hamiltonian consists of Lq field parameters h_i and $\binom{L}{2}Q^2$ coupling parameters J_{ij} which describe the system's preference for each amino acid character at site i and each amino acid character pair at sites i, j, respectively. In the way we present the Boltzmann distribution $P^m \propto \exp(-E)$, negative fields and couplings signify favored amino acids preferences.

Not all the model parameters are independent. Due to the relationship between bivariate marginals P_{ij}, P_{ik}, P_{jk} and the fact that the univariate marginals can be derived entirely from the bivariate marginals, only $L(Q-1) + \binom{L}{2}(Q-1)^2$ of these $LQ + \binom{L}{2}Q^2$ parameters are independent. Several schemes have been developed and used by others to fully constrain the Hamiltonian (see (Morcos et al. 2011, Weigt et al. 2009), for example). Further, the fully-constrained Potts Hamiltonian is "gauge invariant" such that the probablity $P^m\left(\vec{\sigma}^k\right)$ is unchanged by (a) a global bias added to the fields, $h_i(\sigma_i) \to h_i(\sigma_i) + b$, (b) a per-site bias added to the fields $h_i(\sigma_i) \to h_i(\sigma_i) + b_i$, (c) rearrangement of field and coupling contributions such that $J_{ij}(\sigma_i, \sigma_j) \to J_{ij}(\sigma_i, \sigma_j) + b_{ij}(\sigma_j)$ and $h_i(\sigma_i) \to h_i(\sigma_i) - \sum_{j \neq i} b_{ij}(\sigma_j)$, or (d) a combination thereof. Due to this gauge invariance, model parameters are over-specified and thus not unique until a fully-constrained gauge is specified, but the properties P^m and ΔE, among others, are gauge invariant and unique among fully-constrained gauges.

Model Inference

Finding a suitable set of Potts parameters $\{h, J\}$ fully determines the total probability distribution $P^m(\vec{\sigma})$ and is achieved by obtaining the set of fields and couplings which yield bivariate marginal estimates $P^m(\sigma_i, \sigma_j)$ that best reproduce the empirical bivariate marginals $P^{obs}(\sigma_i, \sigma_j)$. Previous studies have developed a number of techniques to do this (Balakrishnan et al. 2011, Barton et al. 2016a, Cocco and Monasson 2011, Ekeberg et al. 2013, Ferguson et al. 2013, Haq et al. 2012, Jones et al. 2012, Mézard and Mora 2009, Morcos

et al. 2011, Weigt et al. 2009). Following (Ferguson et al. 2013), we estimate the bivariate marginals given a set of fields and couplings by generating sequences through Markov Chain Monte Carlo (MCMC) where the Metropolis criterion for a generated sequence is proportional to the exponentiated Potts Hamiltonian. The optimal set of parameters $\{h, J\}$ are found through multidimensional Newton search, where bivariate marginal estimates are compared to the empirical distribution to determine descent steps. Unlike several inference methods referenced above, this method avoids making explicit approximations to the model probability distribution, though approximations are made in the computation of the Newton steps, and this method is limited by sampling error of the input emperical marginal distributions and by the need for the simulation to equilibrate. Also, the method is computationally intensive. A brief description of the method follows; see the supplemental information of Haldane et al. (Haldane et al. 2016) for a full description of the method.

Determining the schema for choosing the Newton step is crucial. In (Ferguson et al. 2013), a quasi-newton parameter update approach was developed, in which updates to J_{ij} and h_i are determined by inverting the system's Jacobian, to minimize the difference between model-estimated and empirical marginals. To simplify and speed up this computation, we take advantage of the gauage invariance of the Potts Hamiltonian to infer a model in which $h_i = 0 \ \forall \ i$, and we compute the expected change in the model marginals $\Delta \ P_{ij}$ (dropping the m superscript) due to a change in J_{ij} to first order by

$$\Delta P_{ij}(\sigma_i, \sigma_j) = \sum_{kl,\sigma_k\sigma_l} \frac{\partial P_{ij}(\sigma_i, \sigma_j)}{\partial J_{kl}(\sigma_k, \sigma_l)} \Delta J_{kl}(\sigma_k, \sigma_l) + \sum_{k,\sigma_k} \frac{\partial P_{ij}(\sigma_i, \sigma_j)}{\partial h_k(\sigma_k)} \Delta h_k(\sigma_k) \tag{11}$$

with a similar relation for $\Delta P_i(\sigma_i)$. The challenge is to compute the Jacobian $\frac{\partial P_{ij}(\sigma_i,\sigma_j)}{\partial J_{kl}(\sigma_k,\sigma_l)}$ and invert the linear system in Equation 11, and solve for the changes ΔJ_{ij} and Δh_i given ΔP_{ij} which we choose as

$$\Delta P_{ij} = \gamma\big(P_{ij}^{emp} - P_{ij}\big) \tag{12}$$

given a damping parameter γ chosen small enough for the linear (and other) approximations to hold.

The computational cost of fitting $\binom{93}{2} \times (4-1)^2 + 93 \times (4-1) = 38,781$ model parameters on 2 NVIDIA K80 or 4 NVIDIA TitanX GPUs is approximately 4 hours. For a more thorough description of the inference methodology, see the supplementary information of Haldane et al. (Haldane et al. 2016).

Experimental Comparison

Experimentally derived values for either melting temperature (T_m) or viral infectivity via replicative capacity (RC) were mined from the results presented in (Chang and Torbett

2011, Henderson et al. 2012, Louis et al. 2011, Muzammil et al. 2003, van Maarseveen et al. 2006). A csv file of the resulting mined data can be found in Supplementary Data 1.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health (P50 GM103368 to W.F.F., B.E.T., R.M.L; R01 GM30580 to A.H., R.M.L.; S10 OD020095 to W.F.F., A.H., R.M.L.). We thank the supportive collaborative environment provided by the HIV Interaction and Viral Evolution (HIVE) Center at the Scripps Research Institute (`http://hive.scripps.edu`).

Balakrishnan, S., Kamisetty, H., Carbonell, J. G., Lee, S.-I., and Langmead, C. J. Learning generative models for protein fold families. *Proteins*, 79(4):1061–78, Apr 2011. doi: `10.1002/prot.22934`.

Barton, J. P., De Leonardis, E., Coucke, A., and Cocco, S. ACE: adaptive cluster expansion for maximum entropy graphical model inference. *Bioinformatics*, page btw328, Jun 2016a. doi: `10.1093/bioinformatics/btw328`.

Barton, J. P., Goonetilleke, N., Butler, T. C., Walker, B. D., McMichael, A. J., and Chakraborty, A. K. Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. *Nat Commun*, 7:11660, May 2016b. doi: `10.1038/ncomms11660`.

Bloom, J. D., Gong, L. I., and Baltimore, D. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science*, 328(5983):1272–5, Jun 2010. doi: `10.1126/science.1187816`.

Boucher, J. I., Bolon, D. N. A., and Tawfik, D. S. Quantifying and understanding the fitness effects of protein mutations: Laboratory versus nature. *Protein Sci*, 25(7):1219–1226, Jul 2016. doi: `10.1002/pro.2928`.

Breuer, S., Sepulveda, H., Chen, Y., Trotter, J., and Torbett, B. E. A cleavage enzyme-cytometric bead array provides biochemical profiling of resistance mutations in HIV-1 Gag and protease. *Biochemistry*, 50(20):4371–4381, 2011. doi: `10.1021/bi200031m`.

Burger, L. and van Nimwegen, E. Disentangling direct from indirect co-evolution of residues in protein alignments. *PLoS Comput Biol*, 6(1):e1000633, Jan 2010. doi: `10.1371/journal.pcbi.1000633`.

Butler, T. C., Barton, J. P., Kardar, M., and Chakraborty, A. K. Identification of drug resistance mutations in hiv from constraints on natural evolution. *Phys. Rev. E*, 93:022412, Feb 2016.

doi: 10.1103/PhysRevE.93.022412.

Chang, M. W. and Torbett, B. E. Accessory mutations maintain stability in drug-resistant HIV-1 protease. *J Mol Biol*, 410(4):756–60, Jul 2011. doi: 10.1016/j.jmb.2011.03.038.

Chellappan, S., Kairys, V., Fernandes, M. X., Schiffer, C., and Gilson, M. K. Evaluation of the substrate envelope hypothesis for inhibitors of HIV-1 protease. *Proteins: Struct, Funct, Bioinf*, 68(2):561–567, May 2007. doi: 10.1002/prot.21431.

Cocco, S. and Monasson, R. Adaptive cluster expansion for inferring boltzmann machines with noisy data. *Phys Rev Lett*, 106(9):090601, Mar 2011. doi: 10.1103/PhysRevLett.106.090601.

Ekeberg, M., Lövkvist, C., Lan, Y., Weigt, M., and Aurell, E. Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Phys. Rev. E*, 87(1):012707, Jan 2013. doi: 10.1103/PhysRevE.87.012707.

Ferguson, A. L., Mann, J. K., Omarjee, S., Ndung'u, T., Walker, B. D., and Chakraborty, A. K. Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. *Immunity*, 38(3):606–17, Mar 2013. doi: 10.1016/j.immuni.2012.11.022.

Figliuzzi, M., Jacquier, H., Schug, A., Tenaillon, O., and Weigt, M. Coevolutionary landscape inference and the context-dependence of mutations in beta-lactamase TEM-1. *Mol Biol Evo*, 33(1):msv211, Oct 2015. doi: 10.1093/molbev/msv211.

Flynn, W. F., Chang, M. W., Tan, Z., Oliveira, G., Yuan, J., Okulicz, J. F., Torbett, B. E., and Levy, R. M. Deep sequencing of protease inhibitor resistant HIV patient isolates reveals patterns of correlated mutations in Gag and Protease. *PLoS Comput Biol*, 11(4):e1004249, 2015. doi: 10.1371/journal.pcbi.1004249.

Fun, A., Wensing, A. M. J., Verheyen, J., and Nijhuis, M. Human Immunodeficiency Virus Gag and protease: partners in resistance. *Retrovirology*, 9(1):63, Jan 2012. doi: 10.1186/1742-4690-9-63.

Göbel, U., Sander, C., Schneider, R., and Valencia, A. Correlated mutations and residue contacts in proteins. *Proteins*, 18(4):309–317, Apr 1994. doi: 10.1002/prot.34018042.

Gong, L. I., Suchard, M. A., Bloom, J. D., and Pascual, M. Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife*, 2:e00631, 2013. doi: 10.7554/eLife.00631.

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., and Holmes, E. C. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, 303 (5656):327–32, Jan 2004. doi: 10.1126/science.1090727.

Gupta, A. and Adami, C. Strong Selection Significantly Increases Epistatic Interactions in the Long-Term Evolution of a Protein. *PLoS Genet*, 12(3):e1005960, Mar 2016. doi: 10.1371/

journal.pgen.1005960.

Gupta, R., Hill, A., Sawyer, A. W., and Pillay, D. Emergence of drug resistance in HIV type 1 infected patients after receipt of first-line highly active antiretroviral therapy: a systematic review of clinical trials. *Clin Infect Dis*, 47(5):712–22, Sep 2008. doi: 10.1086/590943.

Haldane, A., Flynn, W. F., He, P., Vijayan, R., and Levy, R. M. Structural propensities of kinase family proteins from a potts model of residue co-variation. *Protein Sci*, May 2016. doi: 10.1002/pro.2954.

Haq, O., Levy, R. M., Morozov, A. V., and Andrec, M. Pairwise and higher-order correlations among drug-resistance mutations in HIV-1 subtype B protease. *BMC Bioinformatics*, 14(10): 1–14, 2009. doi: 10.1186/1471-2105-10-S8-S10.

Haq, O., Andrec, M., Morozov, A. V., and Levy, R. M. Correlated electrostatic mutations provide a reservoir of stability in HIV protease. *PLoS Comput Biol*, 8(9):e1002675, Jan 2012. doi: 10.1371/journal.pcbi.1002675.

Henderson, G. J., Lee, S.-K., Irlbeck, D. M., Harris, J., Kline, M., Pollom, E., Parkin, N., and Swanstrom, R. Interplay between single resistance-associated mutations in the HIV-1 protease and viral infectivity, protease activity, and inhibitor sensitivity. *Antimicrob Agents Ch*, 56(2): 623–33, Feb 2012. doi: 10.1128/AAC.05549-11.

Hinkley, T., Martins, J., Chappey, C., Haddad, M., Stawiski, E., Whitcomb, J. M., Petropoulos, C. J., and Bonhoeffer, S. A systems analysis of mutational effects in HIV-1 protease and reverse transcriptase. *Nat Gen*, 43(5):487–9, May 2011. doi: 10.1038/ng.795.

Hoffman, N. G., Schiffer, C. A., and Swanstrom, R. Covariation of amino acid positions in HIV-1 protease. *Virology*, 314(2):536–548, Sep 2003. doi: 10.1016/S0042-6822(03)00484-7.

Hopf, T. A., Ingraham, J. B., Poelwijk, F. J., Springer, M., Sander, C., and Marks, D. S. Quantification of the effect of mutations using a global probability model of natural sequence variation. *arXiv*, Oct 2015.

Jacquin, H., Gilson, A., Shakhnovich, E., Cocco, S., and Monasson, R. Benchmarking Inverse Statistical Approaches for Protein Structure and Design with Exactly Solvable Models. *PLoS Comput Biol*, 12(5):e1004889, May 2016. doi: 10.1371/journal.pcbi.1004889.

Johnson, V. a., Calvez, V., Gunthard, H. F., Paredes, R., Pillay, D., Shafer, R. W., Wensing, A. M., and Richman, D. D. Update of the drug resistance mutations in HIV-1: March 2013. *Topics in Antiviral Medicine*, 21(1):6–14, 2013.

Jones, D. T., Buchan, D. W. A., Cozzetto, D., and Pontil, M. PSICOV: precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments. *Bioinformatics*, 28(2):184–90, Jan 2012. doi: 10.1093/bioinformatics/btr638.

602 Keele, B. F., Giorgi, E. E., Salazar-Gonzalez, J. F., Decker, J. M., Pham, K. T., Salazar, M. G.,
603 Sun, C., Grayson, T., Wang, S., Li, H., et al. Identification and characterization of transmitted
604 and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci*, 105(21):
605 7552–7, May 2008. doi: 10.1073/pnas.0802203105.
606 King, N. M., Melnick, L., Prabu-Jeyabalan, M., Nalivaika, E. A., Yang, S.-S., Gao, Y., Nie, X.,
607 Zepp, C., Heefner, D. L., and Schiffer, C. A. Lack of synergy for inhibitors targeting a multi-
608 drug-resistant HIV-1 protease. *Protein Sci*, 11(2):418–429, Apr 2002. doi: 10.1110/ps.25502.
609 King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., and Schiffer, C. A. Combating Susceptibility
610 to Drug Resistance: Lessons from HIV-1 Protease. *Chem Biol*, 11(10):1333–1338, 2004. doi:
611 10.1016/j.chembiol.2004.08.010.
612 Kolli, M., Stawiski, E., Chappey, C., and Schiffer, C. A. Human immunodeficiency virus type 1
613 protease-correlated cleavage site mutations enhance inhibitor resistance. *Journal of virology*,
614 83(21):11027–42, Nov 2009. doi: 10.1128/JVI.00628-09.
615 Kovalevsky, A. Y., Tie, Y., Liu, F., Boross, P. I., Wang, Y. F., Leshchenko, S., Ghosh, A. K.,
616 Harrison, R. W., and Weber, I. T. Effectiveness of nonpeptide clinical inhibitor TMC-114
617 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *Journal of*
618 *Medicinal Chemistry*, 49(4):1379–1387, Feb 2006. doi: 10.1021/jm050943c.
619 Lefebvre, E. and Schiffer, C. A. Resilience to resistance of HIV-1 protease inhibitors: profile of
620 darunavir. *AIDS reviews*, 10(3):131–42, 2008.
621 Liu, Z., Chen, J., and Thirumalai, D. On the accuracy of inferring energetic coupling between
622 distant sites in protein families from evolutionary imprints: illustrations using lattice model.
623 *Proteins*, 77(4):823–31, Dec 2009. doi: 10.1002/prot.22498.
624 Lockless, S. W., Ranganathan, R., Kukic, P., Mirabello, C., Tradigo, G., Walsh, I., Veltri, P.,
625 Pollastri, G., Socolich, M., Lockless, S. W., et al. Evolutionarily conserved pathways of
626 energetic connectivity in protein families. *BMC Bioinformatics*, 15(5438):295–299, 1999. doi:
627 10.1186/1471-2105-15-6.
628 Louis, J. M., Aniana, A., Weber, I. T., and Sayer, J. M. Inhibition of autoprocessing of natural
629 variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors.
630 *Proc Natl Acad Sci USA*, 108(22):9072–7, May 2011. doi: 10.1073/pnas.1102278108.
631 Mahalingam, B., Wang, Y.-F., Boross, P. I., Tozser, J., Louis, J. M., Harrison, R. W., and Weber,
632 I. T. Crystal structures of HIV protease V82A and L90M mutants reveal changes in the
633 indinavir-binding site. *European journal of biochemistry / FEBS*, 271(8):1516–24, Apr 2004.
634 doi: 10.1111/j.1432-1033.2004.04060.x.
635 Mann, J. K., Barton, J. P., Ferguson, A. L., Omarjee, S., Walker, B. D., Chakraborty, A., and
636 Ndung'u, T. The fitness landscape of HIV-1 Gag: Advanced modeling approaches and valida-

tion of model predictions by in vitro testing. *PLoS Comput Biol*, 10(8):e1003776, Aug 2014. doi: 10.1371/journal.pcbi.1003776.

Marks, D. S., Hopf, T. a., and Sander, C. Protein structure prediction from sequence variation. *Nat Biotechnol*, 30(11):1072–80, Nov 2012. doi: 10.1038/nbt.2419.

Martinez-Picado, J., Savara, A. V., Sutton, L., and D'Aquila, R. T. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. *J Virol*, 73(5):3744–52, May 1999.

McCandlish, D. M., Otwinowski, J., and Plotkin, J. B. Detecting epistasis from an ensemble of adapting populations. *Evolution*, 69(9):2359–70, Sep 2015. doi: 10.1111/evo.12735.

Mézard, M. and Mora, T. Constraint satisfaction problems and neural networks: A statistical physics perspective. *J Physiol*, 103(1-2):107–13, 2009. doi: 10.1016/j.jphysparis.2009.05.013.

Mittal, S., Cai, Y., Nalam, M. N. L., Bolon, D. N. A., and Schiffer, C. A. Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease. *J Am Chem Soc*, 134(9):4163–4168, Mar 2012. doi: 10.1021/ja2095766.

Mora, T. and Bialek, W. Are biological systems poised at criticality? *J Stat Phys*, 144(2):268–302, Jun 2011. doi: 10.1007/s10955-011-0229-4.

Morcos, F., Pagnani, A., Lunt, B., Bertolino, A., Marks, D. S., Sander, C., Zecchina, R., Onuchic, J. N., Hwa, T., and Weigt, M. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc Natl Acad Sci USA*, 108(49):E1293–301, Dec 2011. doi: 10.1073/pnas.1111471108.

Morcos, F., Schafer, N. P., Cheng, R. R., Onuchic, J. N., and Wolynes, P. G. Coevolutionary information, protein folding landscapes, and the thermodynamics of natural selection. *Proc Natl Acad Sci USA*, 111(34):12408–13, Aug 2014. doi: 10.1073/pnas.1413575111.

Murphy, L. R., Wallqvist, A., and Levy, R. M. Simplified amino acid alphabets for protein fold recognition and implications for folding. *Protein Eng*, 13(3):149–152, 2000. doi: 10.1093/protein/13.3.149.

Muzammil, S., Ross, P., and Freire, E. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. *Biochemistry*, 42(3):631–638, Jan 2003. doi: 10.1021/bi027019u.

Ode, H., Neya, S., Hata, M., Sugiura, W., and Hoshino, T. Computational simulations of HIV-1 proteases: Multi-drug resistance due to nonactive site mutation L90M. *J Am Chem Soc*, 128 (24):7887–7895, 2006. doi: 10.1021/ja060682b.

Özen, A., Halilolu, T., and Schiffer, C. A. Dynamics of preferential substrate recognition in HIV-1 protease: Redefining the substrate envelope. *Journal of Molecular Biology*, 410(4):726–744,

2011. doi: 10.1016/j.jmb.2011.03.053.

Parry, C. M., Kolli, M., Myers, R. E., Cane, P. A., Schiffer, C. A., and Pillay, D. Three residues in HIV-1 matrix contribute to protease inhibitor susceptibility and replication capacity. *Antimicrob Agents Ch*, 55(3):1106–13, Mar 2011. doi: 10.1128/AAC.01228-10.

Pollock, D. D., Thiltgen, G., and Goldstein, R. A. Amino acid coevolution induces an evolutionary Stokes shift. *Proc Natl Acad Sci USA*, 109(21):E1352–9, May 2012. doi: 10.1073/pnas.1120084109.

Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure*, 10(3):369–81, Mar 2002. doi: 10.1016/S0969-2126(02)00720-7.

Rhee, S.-Y., Liu, T. F., Holmes, S. P., and Shafer, R. W. HIV-1 subtype B protease and reverse transcriptase amino acid covariation. *PLoS Comput Biol*, 3(5):e87, May 2007. doi: 10.1371/journal.pcbi.0030087.

Richman, D. D., Morton, S. C., Wrin, T., Hellmann, N., Berry, S., Shapiro, M. F., and Bozzette, S. A. The prevalence of antiretroviral drug resistance in the United States. *AIDS*, 18(10):1393–401, Jul 2004.

Shafer, R. W. Rationale and uses of a public HIV drug-resistance database. *J Infect Dis*, 194 Suppl 1:S51–S58, Sep 2006. doi: 10.1086/505356.

Shafer, R. W. and Schapiro, J. M. HIV-1 Drug Resistance Mutations: an Updated Framework for the Second Decade of HAART. *AIDS Review*, 10(2):67–84, 2008.

Shah, P., McCandlish, D. M., and Plotkin, J. B. Contingency and entrenchment in protein evolution under purifying selection. *Proc Natl Acad Sci USA*, 112(25):E3226–35, Jun 2015. doi: 10.1073/pnas.1412933112.

Shekhar, K., Ruberman, C., Ferguson, A., Barton, J., Kardar, M., and Chakraborty, A. Spin models inferred from patient-derived viral sequence data faithfully describe HIV fitness landscapes. *Phys Rev E*, 88(6):062705, Dec 2013. doi: 10.1103/PhysRevE.88.062705.

Socolich, M., Lockless, S. W., Russ, W. P., Lee, H., Gardner, K. H., and Ranganathan, R. Evolutionary information for specifying a protein fold. *Nature*, 437(7058):512–8, Sep 2005. doi: 10.1038/nature03991.

Sulkowska, J. I., Morcos, F., Weigt, M., Hwa, T., and Onuchic, J. N. Genomics-aided structure prediction. *PNAS*, 109(26):10340–10345, 2012. doi: 10.1073/pnas.1207864109.

Sutto, L., Marsili, S., Valencia, A., and Gervasio, F. L. From residue coevolution to protein conformational ensembles and functional dynamics. *Proc Natl Acad Sci USA*, 112(44):13567–72, Nov 2015. doi: 10.1073/pnas.1508584112.

van Maarseveen, N. M., de Jong, D., Boucher, C. A. B., and Nijhuis, M. An increase in viral replicative capacity drives the evolution of protease inhibitor-resistant human immunodeficiency virus type 1 in the absence of drugs. *J Acq Imm Def*, 42(2):162–8, Jun 2006. doi: 10.1097/01.qai.0000219787.65915.56.

Wang, X., Minasov, G., and Shoichet, B. K. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *Journal of molecular biology*, 320(1):85–95, Jun 2002. doi: 10.1016/S0022-2836(02)00400-X.

Weigt, M., White, R. A., Szurmant, H., Hoch, J. A., and Hwa, T. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc Natl Acad Sci USA*, 106(1): 67–72, Jan 2009. doi: 10.1073/pnas.0805923106.

Wu, T. D., Schiffer, C. A., Gonzales, M. J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A. R., Fessel, W. J., and Shafer, R. W. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J Virol*, 77(8):4836–4847, 2003. doi: 10.1128/JVI.77.8.4836.

Yang, W.-L., Kouyos, R. D., Böni, J., Yerly, S., Klimkait, T., Aubert, V., Scherrer, A. U., Shilaih, M., Hinkley, T., Petropoulos, C., Bonhoeffer, S., et al. Persistence of transmitted HIV-1 drug resistance mutations associated with fitness costs and viral genetic backgrounds. *PLoS Pathog*, 11(3):e1004722, Mar 2015. doi: 10.1371/journal.ppat.1004722.

Zeldovich, K. B. and Shakhnovich, E. I. Understanding protein evolution: From protein physics to darwinian selection. *Annu Rev Phys Chem*, 59(1):105–127, May 2008. doi: 10.1146/annurev.physchem.58.032806.104449.

Zeldovich, K. B., Chen, P., and Shakhnovich, E. I. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci USA*, 104(41):16152–7, Oct 2007. doi: 10.1073/pnas.0705366104.

FIGURES AND FIGURE LEGENDS

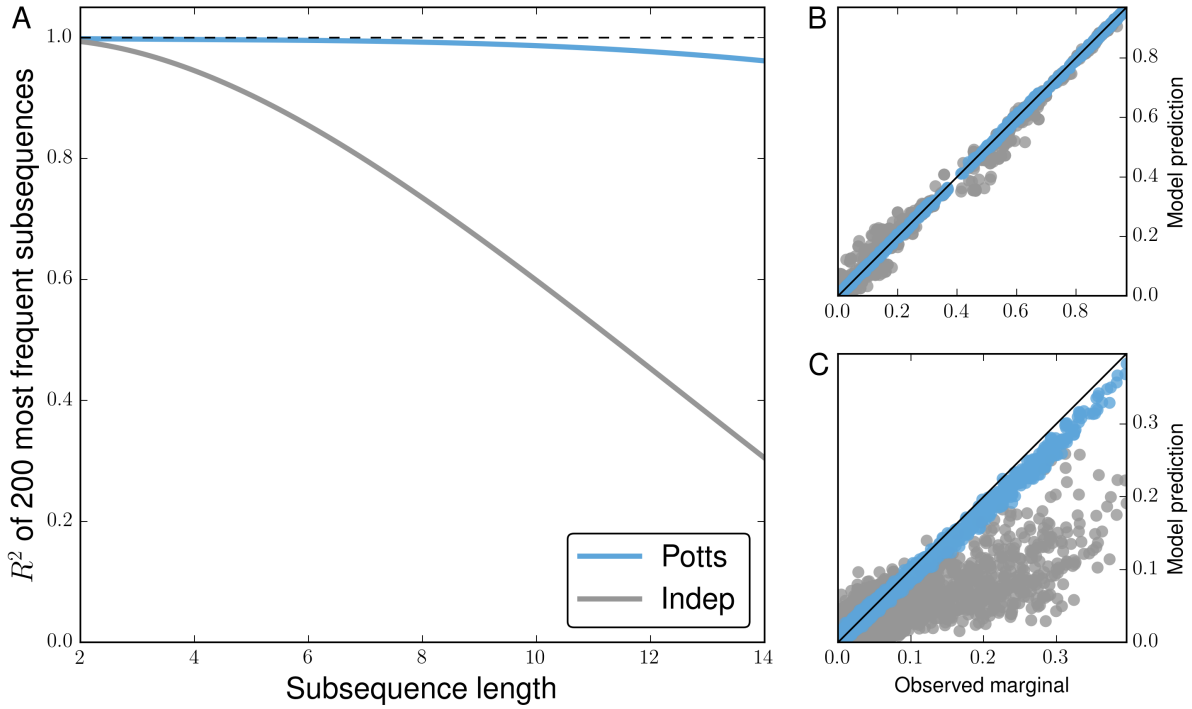


FIG. 1: **Potts model is predictive of higher order sequence statistics.** For each subsequence length varying from 2 to 14, subsequence frequencies determined by counting occurences in the MSA are computed for all observed subsequences at 500 randomly chosen combinations among 36 PI-associated positions. (A) Pearson R^2 of the 200 most probable observed subsequence frequencies (marginals) with corresponding predictions by Potts (blue) and independent (gray) models for varying subsequence lengths. (B) 2^{nd} and (C) 14^{th} order observed marginals predicted by both models. Shown in (B,C) are observed frequencies at the 500 randomly chosen combinations of 2 and 14 positions among 36 PI-associated sites, with approximately 2500 and 5600 subsequence frequencies greater than 0.01 visible, respectively.

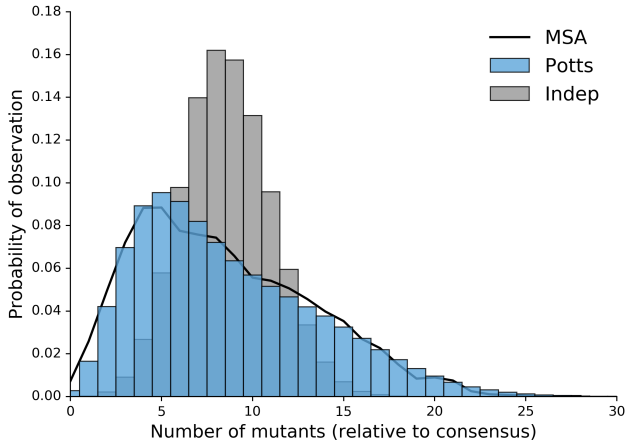


FIG. 2: **Potts model captures properties of full length sequence ensemble.** Probabilities of observing sequences with any k mutations relative to the consensus sequence as observed in original MSA (black) and predicted by the Potts (blue) and independent (gray) models.

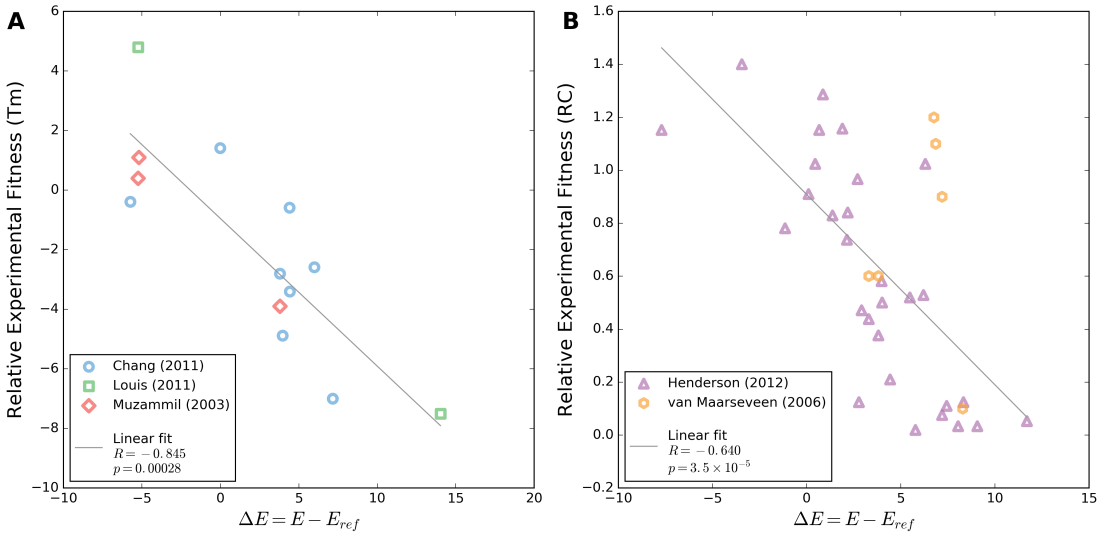


FIG. 3: **Change in Potts energy correlates with change in experimental fitness.** (A) Changes in melting temperature (T_m) for individual sequences relative to a reference sequence extracted from literature (Chang and Torbett 2011, Louis et al. 2011, Muzammil et al. 2003). These sequences differ from the wildtype by 1–2 mutations (Chang and Torbett 2011) up to 10–14 mutations (Louis et al. 2011, Muzammil et al. 2003). (B) Change in relative infectivty as measured by replicative capacity assay for individual sequences containing only single point mutations (Henderson et al. 2012) and 1–5 mutations (van Maarseveen et al. 2006). In both panels a linear regression fit with Pearson's R and associated two-tailed p-value are provided in the legend.

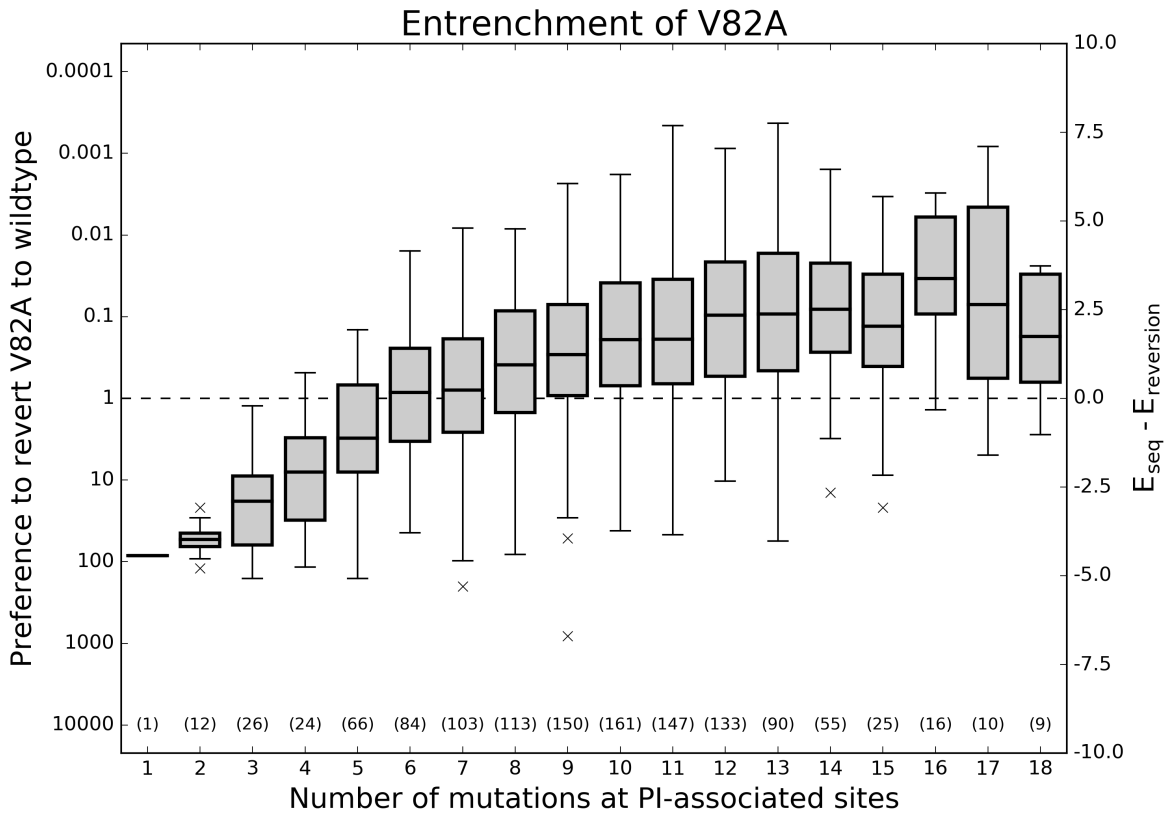
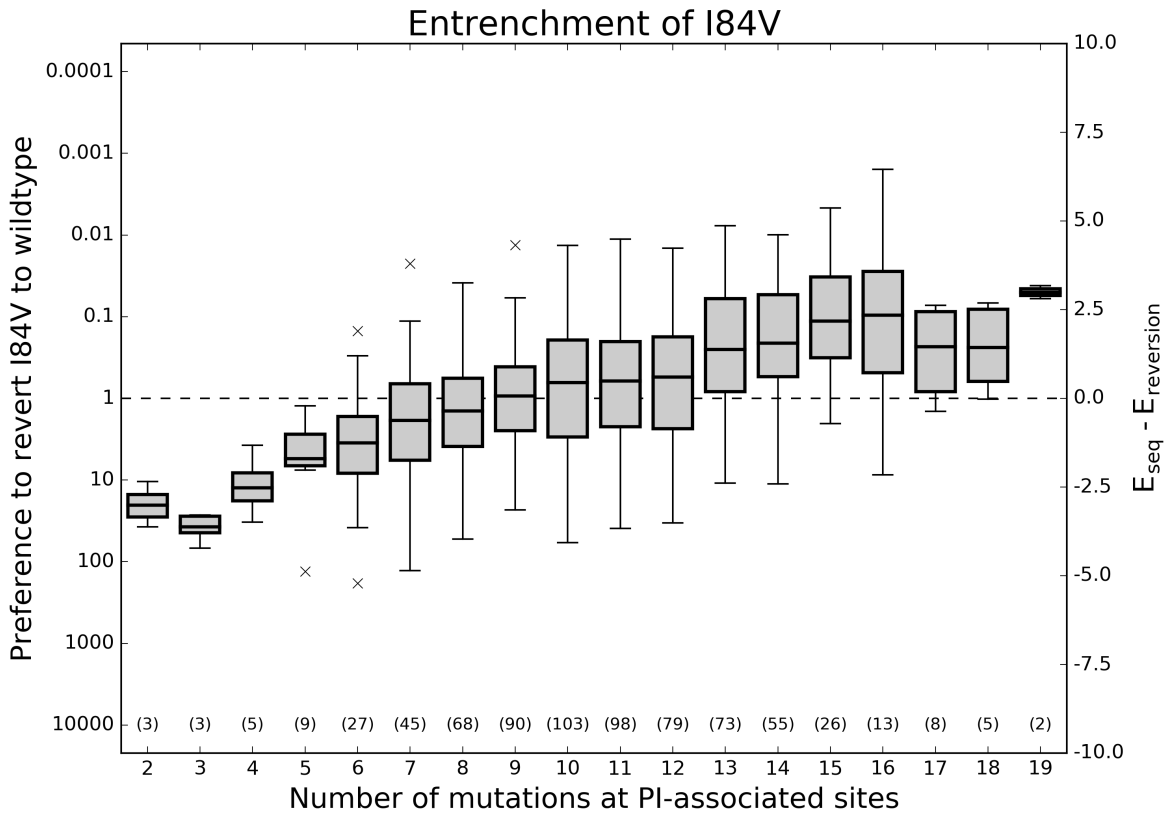
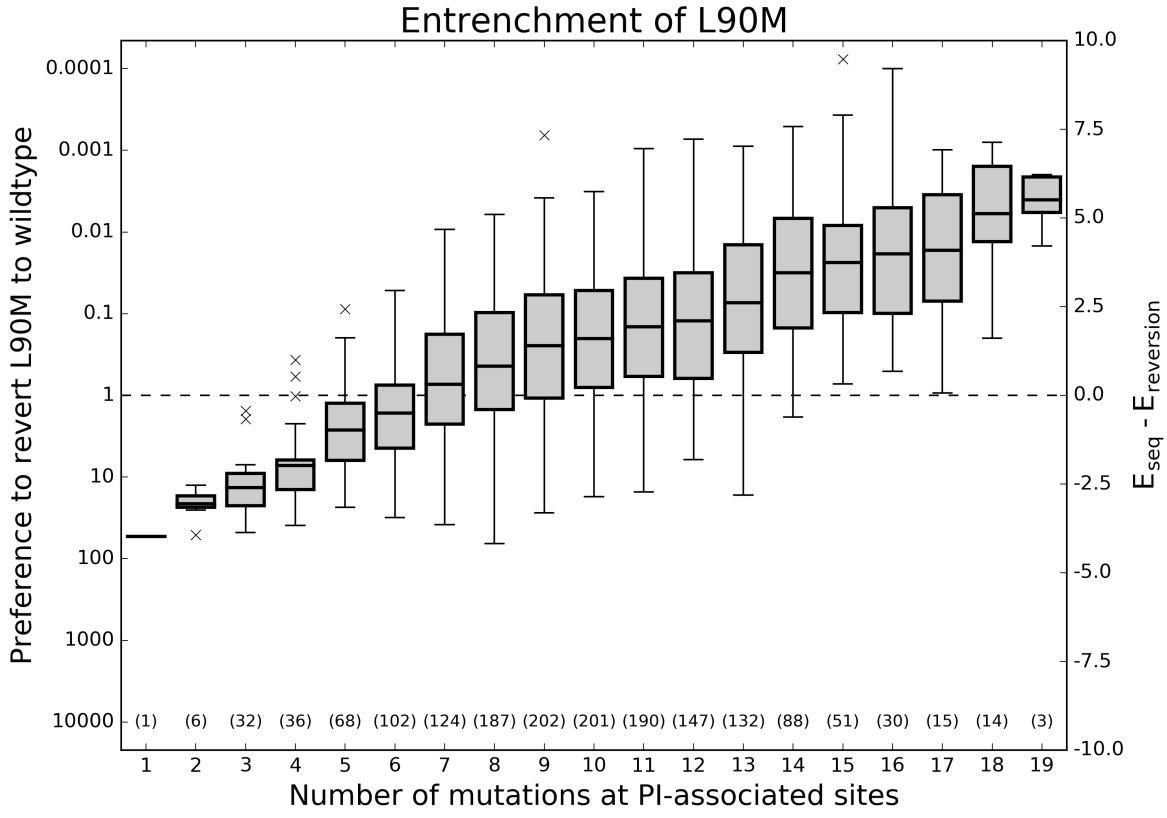


FIG. 4: **Effect of epistasis on the fitness penalty incurred by primary resistance mutations.** For each of the 3 primary HIV protease mutations described in (Chang and Torbett 2011), two Potts statistical energies are computed for all observed sequences containing that mutation: E_{seq}, the energy of the sequence with that mutation and $E_{reversion}$, the energy with that primary mutation reverted to wildtype. This Potts energy difference, $\Delta E = E_{seq} - E_{reversion}$ is shown versus hamming distance from the wildtype including only PI-assocated positions. Ordinate scales are given in both relative probability of reversion $\exp(-\Delta E)$ (left) and ΔE (right). Values below (above) the dashed line on the ordinate correspond to fitness gain (penalty) upon reversion to wildtype. Although primary resistance mutations initially destabilize the protease, as mutations accumulate, the primary resistance mutations become entrenched, meaning their reversion becomes destabilizing to the protein.





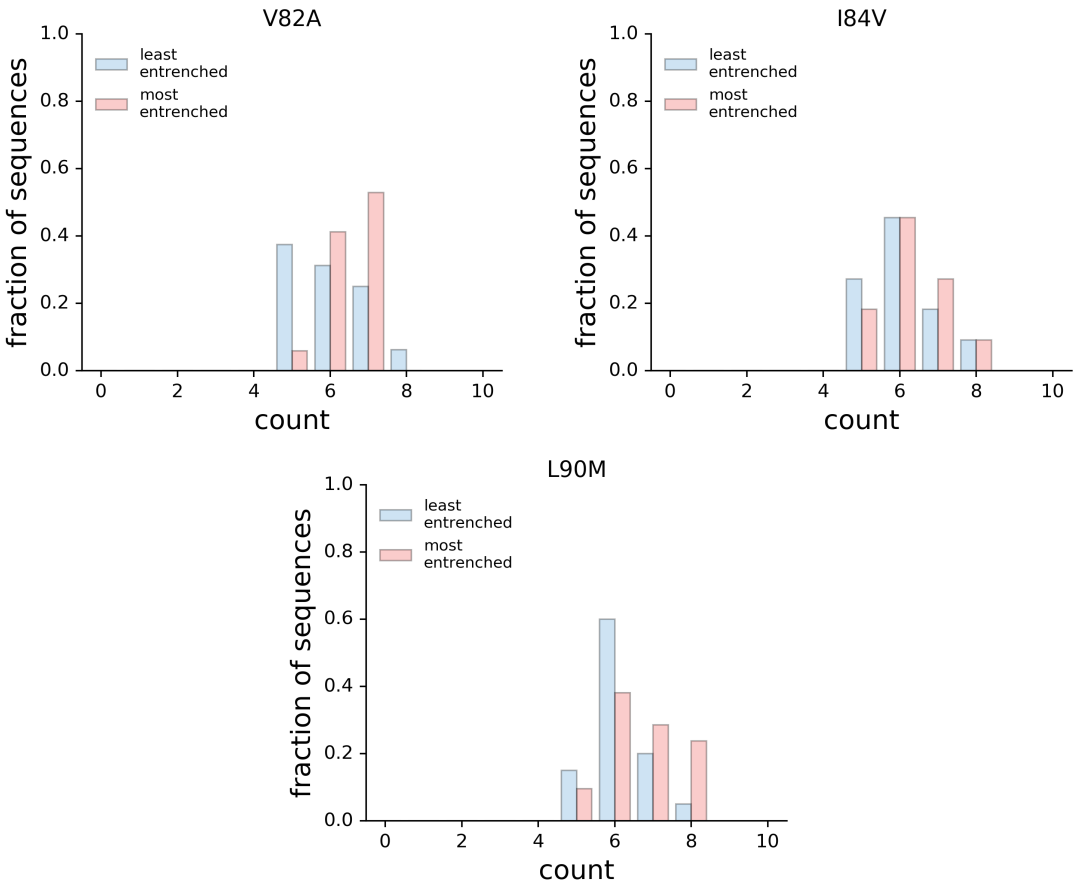


FIG. 5: **Distributions of accessory mutations in most and least entrenching sequences.** The number of accessory mutations among the 10% most and least entrenching sequences for the primary mutations V82A, I84V, and L90M with a fixed hamming distance of 10 from consensus. In all three cases, the distributions are not significantly different (Mann Whitney $U_{V82} = 92.5$, $U_{V84} = 53.0$, $U_{L90} = 145.5$, all with $p > 0.05$).

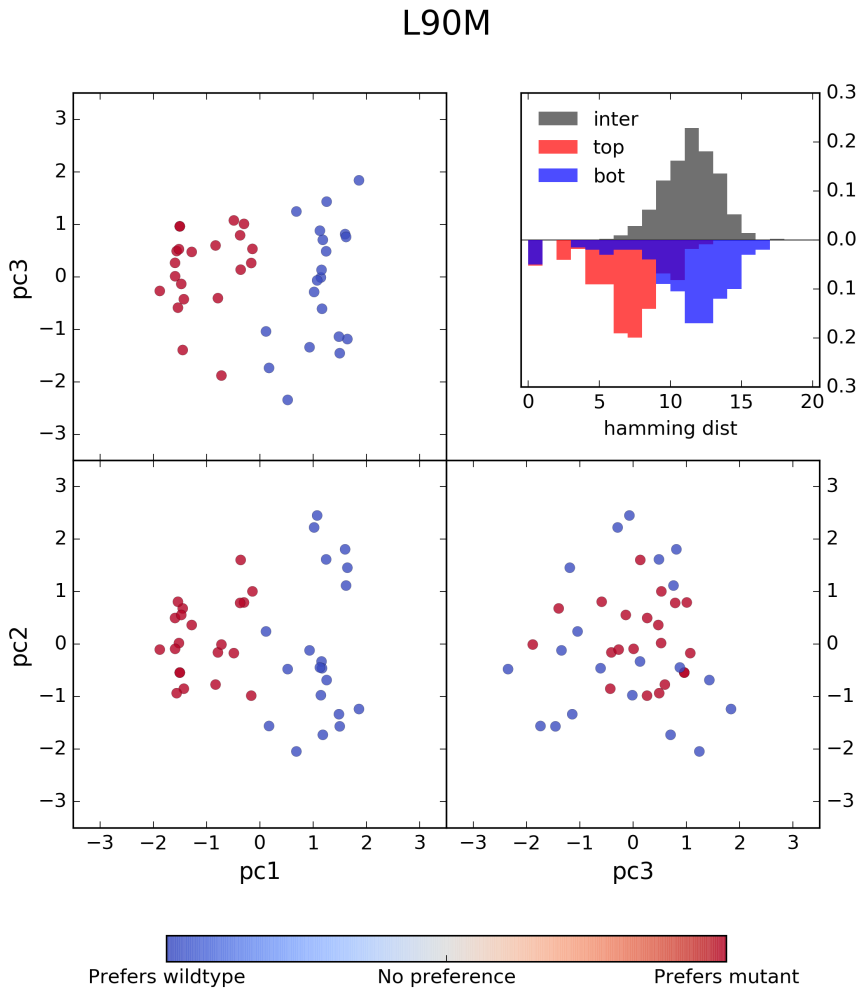


FIG. 6: **PCA analysis of most and least entrenching sequence backgrounds for primary resistance mutation L90M.** Sequences from the 10^{th} and 90^{th} percentiles in ΔE of the sequences containing L90M and with a hamming distance of 10 from the consensus were labeled as "least entrenching" and "most entrenching", respectively, and pooled. These sequences of length $L = 93$ encoded with a $Q = 4$ alphabet were transformed to bit vectors of length LQ and Principal component analysis (PCA) was performed on this set of transformed sequences. The projection of these sequences onto their first 3 principal compenents are shown above with the least entrenching sequences colored blue and most entrenching sequences colored red. The first principal component clearly separates the most from the least entrenching sequence backgrounds for L90M while the other two components explain variation within the two groups of sequences. Shown in the inset are the distributions of hamming distances between (gray) and within the most entrenching (red) and least entrenching (blue) sequences.

731 **TABLES**

TABLE I: Combinations of a most and least entrenching sequence corresponding to the entrenchment of the primary mutations V82A, I84V, and L90M.

		V82A		I84V		L90M	
Position[a,b]	Consensus	ME[c]	LE[d]	ME	LE	ME	LE
10	L	I	I	F	L	I	L
13	I	I	V	I	V	I	I
20	K	R	K	K	R	I	K
24	L	I	L	L	L	L	L
30	D	D	D	D	N	D	D
33	L	L	F	F	L	L	L
35	E	D	E	E	D	D	E
36	M	I	M	M	I	M	I
37	N	N	N	D	S	S	D
41	R	K	R	R	R	R	K
46	M	L	M	I	M	I	M
48	G	G	G	G	G	G	V
54	I	V	V	V	I	I	V
57	R	R	R	R	K	R	R
58	Q	Q	Q	Q	E	Q	Q
62	I	I	I	I	V	V	V
63	L	P	P	P	P	P	P
67	C	C	F	C	C	C	C
69	H	H	H	H	H	H	Y
71	A	V	V	V	T	I	V
72	I	I	M	I	I	I	I
73	G	G	S	G	G	S	G
74	T	T	T	P	T	T	T
77	V	V	I	V	V	V	V
82	V	A	A	V	V	V	A
84	I	I	V	V	V	V	I
88	N	N	N	N	D	N	N
90	L	L	M	M	M	M	M
93	I	L	I	L	I	L	L
	$\Delta\Delta E$	6.93		5.80		5.52	
	Relative probability ME/LE	1022		330		250	

[a] The residue at positions not listed is the subtype B consensus residue

[b] PI-associated positions are shown in bold

[c] most entrenching

[d] least entrenching