Tuning the course of evolution on the biophysical fitness landscape of an RNA virus

Assaf Rotem[1,#], Adrian W.R. Serohijos[2,10,#], Connie B. Chang[1,9], Joshua T. Wolfe[3], Audrey E. Fischer[3], Thomas S. Mehoke[3], Huidan Zhang[1,4], Ye Tao[1], W. Lloyd Ung[1], Jeong-Mo Choi[2], Abimbola O. Kolawole[5], Stephan A. Koehler[1], Susan Wu[3], Peter M. Thielen[3], Naiwen Cui[1], Plamen A. Demirev[3], Nicholas S. Giacobbi[6], Timothy R. Julian[7,12], Kellogg Schwab[7], Jeffrey S. Lin[3], Thomas J. Smith[8], James M. Pipas[6], Christiane E. Wobus[5], Andrew B. Feldman[11], David A. Weitz[1,*], and Eugene I. Shakhnovich[2,*]

[1] School of Engineering and Applied Sciences and Department of Physics, Harvard University, 9 Oxford Street, Cambridge, MA 02138, USA
[2] Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA
[3] Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA
[4] Department of Cell Biology, Key Laboratory of Cell Biology, Ministry of Public Health, and Key Laboratory of Medical Cell Biology, Ministry of Education, China Medical University, Shenyang 110001, China
[5] Department of Microbiology and Immunology, University of Michigan Medical School, 1150 West Medical Center Drive, Ann Arbor, MI 48109, USA
[6] Department of Biological Sciences, University of Pittsburgh, 4249 Fifth Avenue, Pittsburgh, PA 15260, USA.
[7] Environmental Health Sciences and the Hopkins Water Institute, Johns Hopkins Bloomberg School of Public Health , Baltimore, Maryland 21231, USA
[8] Department of Biochemistry and Molecular Biology, University of Texas Medical Branch at Galveston, 301 University Boulevard, Galveston, TX 77555, USA
[9] Chemical and Biological Engineering Department, Montana State University, Bozeman, Montana, USA
[10] Département de Biochimie et Centre Robert-Cedergren en Bioinformatique et Génomique, Université de Montréal, Quebec, Canada
[11] Department of Emergency Medicine, Johns Hopkins Medicine, 5801 Smith Avenue, Suite 3220, Davis Building, Baltimore, MD, 21209, USA
[12] Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), 8600 Dübendorf, Switzerland.

[#]These authors contributed equally.
*Corresponding authors:
D.A.W. (weitz@seas.harvard.edu)
E.I.S. (shakhnovich@chemistry.harvard.edu)

ABSTRACT

Predicting viral evolution remains a major challenge with profound implications for public health. Viral evolutionary pathways are determined by the fitness landscape, which maps viral genotype to fitness. However, a quantitative description of the landscape and the evolutionary forces on it remain elusive. Here, we apply a biophysical fitness model based on capsid folding stability and antibody binding affinity to predict the evolutionary pathway of norovirus escaping a neutralizing antibody. The model is validated by experimental evolution in bulk culture and in a drop-based microfluidics device, the "Evolution Chip", which propagates millions of independent viral sub-populations. We demonstrate that along the axis of binding affinity, selection for escape variants and drift due to random mutations have the same direction. However, along folding stability, selection and drift are opposing forces whose balance is tuned by viral population size. Our results demonstrate that predictable epistatic tradeoffs shape viral evolution.

INTRODUCTION

Responding to viral pandemics or to the emergence of new microbial pathogens is a major challenge to public health [1]. A critical component to this response is the prediction of the course of microbial evolution. One approach to this prediction uses available genomic samples and branching patterns of genealogical trees to statistically project future dominant strains [2]. However, this approach can only infer the likelihood that an existing viral strain will dominate, and cannot predict emergence of novel strains that might become dominant due to *de novo* beneficial mutations. It is these novel strains that are often the most virulent, posing the greatest hazard to public health. Predicting the evolution of novel strains is the key to addressing this challenge and requires knowledge of the relationship between mutations in the viral genome and the fitness for individual organisms [3]. This relationship is the fitness landscape, which is a complex, multidimensional function; however, this can rarely be quantitatively determined. Nevertheless, it is essential for predicting selection of the most probable mutants. Moreover, the fate of mutations is also a function of population structure [4]. In particular, population size changes the balance between the impact of random mutations on fitness and that of selection [5]

and is thought to affect both the rate and direction of evolution [6]. Recent studies use microbial fitness landscapes [7-9] or population structure [10,11] to predict the course of evolution, but to date, none links these elements together. Without this link, further progress in predicting the course of viral evolution is significantly hindered.

In this paper, we quantitatively determine a fitness landscape for an RNA virus subjected to the environmental pressure of a neutralizing antibody, and use it to account for the evolution of the virus under conditions that constrain population size. The experimentally measured fitness landscape is correctly described by two biophysical parameters: the thermodynamics of folding of the capsid protein and its binding to the antibody. We probe the evolution of a model norovirus both in bulk, where population size is large, and in a microfluidic evolution chip which uses small drops to perform millions of experiments [12-16] that probe evolution in very small population sizes. We show that the dynamics of viral adaptation is strongly dependent on population size. These results can be quantitatively described by a theoretical framework that combines protein biophysics and population genetics, providing the critical link between fitness landscape and population structure that enables prediction of evolution.

We focus in this work on Murine Norovirus (MNV), a model for human RNA viruses which are the major cause of epidemics in the world [17-19]. MNV is a non-enveloped RNA virus that consists of 180 copies of the capsid protein assembled around a 7.5kb long positive-strand RNA genome. It mutates at ~1 base per genome per replication cycle and produces ~10^4 progenies in a single cell infection, of which ~100 are infectious viral particles, or plaque forming units (pfu)[12].

RESULTS

Lab evolution of Norovirus in large and small population sizes

To study viral evolution, we propagate a viral isolate (MNV-1, denoted *wt*) in the presence of a neutralizing antibody (mAb6.2, [20]) that binds to the protruding domain (P-domain) of the capsid, and prevents virus entry into the host cell [21,22]. This set-up allows us to study the way the virus evolves to adapt to a new environment. To investigate the dynamics of this escape from the antibody, we sequence a 376 bp fragment of the genome encoding the outermost part of the P-domain (residues 281 – 412 of VP1) and follow the frequency of 37,244 unique haplotypes

(**Figure 1-figure supplement 1** and **Table S1**) observed over several passages, allowing us to follow the evolution over several generations. First, we propagate *wt* in standard bulk culture conditions, using ~10^6 virions per passage under Ab pressure (**Figure 1A**). The population is initially dominated by the *wt* (~90% of the population) with the rest of the viral quasi-species consisting of single and double mutants (**Table S1**). After 2 passages the total number of surviving viruses has decreased significantly due to the neutralizing effect of the Ab (**Figure 1-figure supplement 2**); however, three single mutants E296K (*A*), D385G (*B*), T301I (*C*) as well as their double mutants (*AB*, *AC*, *BC*) occur at higher frequencies than the other haplotypes. By the fourth passage the triple mutant *ABC*, which first arises on passage 2, occurs at a frequency even higher than the other mutant haplotypes (86%); moreover, the total number of viruses increases to levels comparable to those observed after the first passage (**Figure 1-figure supplement 2**). This suggests that *ABC* is an escape variant.

A central tenet of evolutionary theory is that the way organisms explore their fitness landscape depends on the size of their population, which controls the balance between random mutational drift (i.e., direction of *randomly arising* mutations) in the population and fitness driven selection (i.e., direction of beneficial mutations) [4,5]. This balance determines the most likely evolutionary pathways on a given fitness landscape. Indeed, the population size may be particularly important for noroviruses where a single viral particle is sufficient to infect the host animal [23]; thus it is possible that viruses propagate in very small populations as they adapt to a new environment prior to the emergence of an epidemic.

The consequences of a smaller population size are captured by the S. Wright's shifting balance theory [6], which hypothesizes that evolution can proceed more efficiently in three phases: i) a population is divided into subpopulations to weaken selection and increase drift, ii) each subpopulation evolves independently, whereupon iii) the subpopulations are mixed back into a large population and then all individuals compete. We can directly probe this hypothesis experimentally by drastically reducing population size compared to typical laboratory bulk cultures, which propagate ~10^6 to 10^8 viruses (**Figure 1-figure supplement 2**). To evolve viruses in small population sizes we use a novel microfluidics system, the "Evolution Chip", which propagates ~10^6 subpopulations of 1-10 infectious particles (pfu) in distinct and non-mixing compartments (**Figure 1B**, **Figure 1-figure supplement 3**, **Movie S1** and **Movie S2**).

The microfluidics system allows us to drastically reduce the population size without reducing the total number of viruses sampled, thereby maintaining the statistics comparable to that of a bulk experiment.

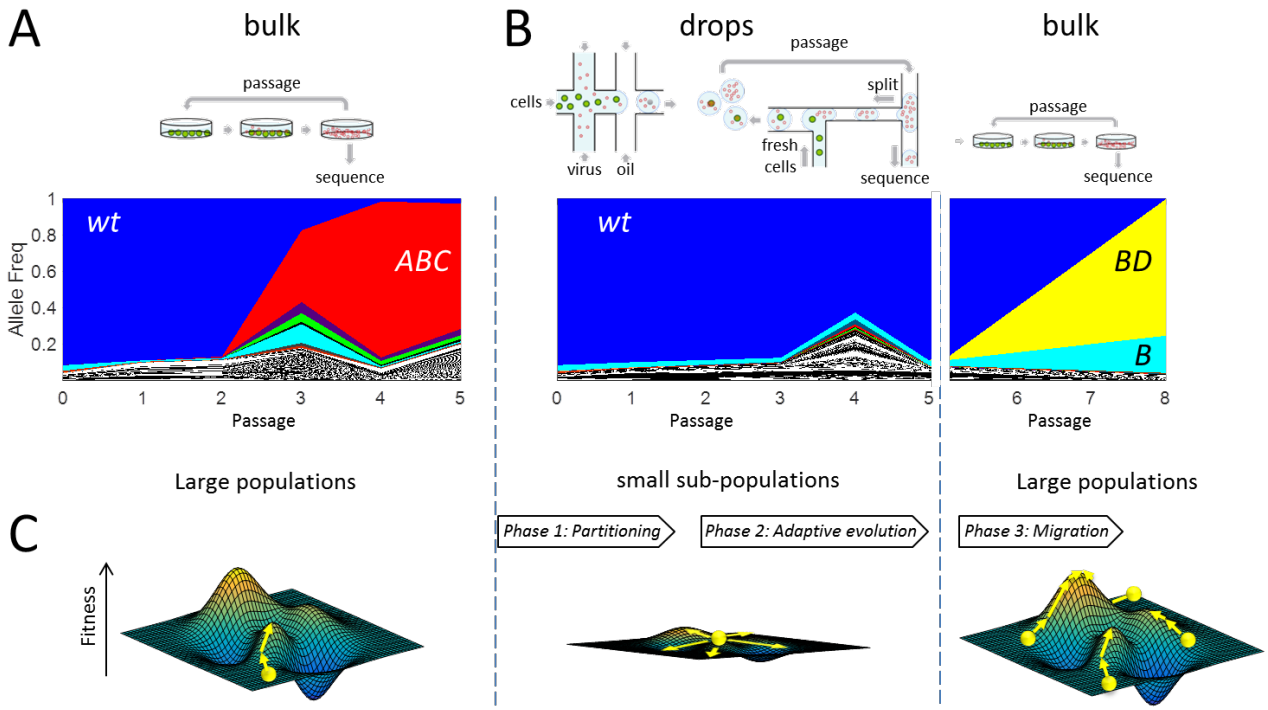

Figure 1. Wright's shifting balance theory (SBT) - Viral evolution in large and small population sizes.

A) Top: 10^6 viruses evolving against a neutralizing antibody in bulk by serial propagation. **Bottom:** The allele frequencies of 1,364 distinct P-domain haplotype sequences are plotted per passage (see also Figure S2**). B) Viral evolution in small populations. Top:** Experimental scheme of the three phases of Wright's SBT - *Phase 1*: 10^6 pico-liter drops are loaded with on average 1 virus and 2 host cells per drop. *Phase 2:* viruses evolve in drops for five passages. *Phase 3:* the emulsion is broken and all lineages evolve together in bulk for 3 additional passages (frequencies for passages 6 and 7 are interpolated from passages 5 and 8, see also Figure S3 and Movie S1, S2). **Bottom:** The allele frequencies of 620 distinct P-domain haplotypes from B (*Phase 2, center panel, phase 3 right panel*) are plotted per passage. (Figure S2). **C) Exploration of the fitness landscape for the three phases of SBT** - *phase 1 (left)*: a previously adapted population (lower peak on left fitness landscape) is partitioned into smaller subpopulations. *phase 2 (middle)*: small sub-populations adapt in isolated conditions where selection pressure is reduced, allowing an extended exploration of the fitness landscape. *phase 3 (right)*: subpopulations migrate, mix and compete, evolving new and more fit variants that were explored in isolation (highest peak). Haplotype legend: A: E296K, B: D385G, C: T301I, D: A382V.

The following supplements are available for Figure 1:

Figure 1-figure supplement 1. Overview of the experimental set-up

A) RAW 264.7 cells exposed to MNV-1 virus are infected and viral progenies are released into the supernatant. To passage viruses, a small volume fraction of the supernatant is transferred to a new sample of fresh cells. The viral genomes in the remaining supernatant are prepared for sequencing. **B)** The 7.5kb long MNV-1 genome includes 4 main reading frames with the capsid protein gene (VP1) residing in the second frame. For our purposes, we sequence a 376bp long fragment in VP1 that encodes part of the protruding domain (P domain) of the capsid protein. **C)** *Left:* Cryo-EM structure of MNV capsid, adapted from [21]. *Right:* Model of MNV made by fitting the P-domain of MNV-1 [24] and the shell domain from the Norwalk virus structure[25] into the cryo-EM density of MNV-1 [22].

Figure 1-figure supplement 2. Allele frequency and viral RNA concentration from experimental viral evolution in bulk and drops

A) The allele frequencies of 1,364 distinct P-domain haplotypes of the evolution in bulk (Figure 1A) are plotted per passage, with the y-axis plotted in log-scale to show the polyclonal structure of the viral quasi-species. **B)** The allele frequencies of 620 distinct P-domain haplotypes of the evolution in drops (Figure 1B) are plotted per passage, with the y-axis plotted in log-scale to show the polyclonal structure of the viral quasi-species. We note that in (A) and (B), similar to primary Figure 1, dominant haplotypes are individually color coded but do not appear due to log transformation of Allele frequency. **C, D)** The concentration of viral RNA in the experimental samples is plotted per passage for both the evolution in bulk (C) and that in drops (D).

Figure 1-figure supplement 3. "Evolution Chip": a system for evolving virus in very small populations

Left: Virus and host cells co-encapsulation. *Middle:* Drop incubation. *Right:* Drop passaging. All scale bars are 100 μm.

Table S1. Complete list of MNV-1 variants from deep sequencing

Movie S1. Encapsulation of virus and cells in drops

Movie S2. "Evolution Chip" for serial passaging in drops

In stark contrast with the bulk experiments, amplification and hence growth of potential escape variants that sweep the population is precluded when each variant is confined in a single drop with just two host cells; as a result, the *wt* remains the dominant fraction of the observed viruses through all passages. Potential escape viruses are present, but are in complete isolation from each other, at population sizes of just a single infection event per generation. This microfluidic experiment (**Figure 1B**) implements the first two phases of the shifting balance theory [6]: i) partitioning the population in drops to weaken selection and increase genetic drift and ii) evolving the sub-populations without competition between drops. However, the absence of competition precludes detection of potential escape variants. To overcome this, we break the emulsion after five passages in the evolution chip, mix the contents of all the drops and propagate the sample in the presence of Ab under bulk conditions for three additional passages [26]. This enables new escape variants that evolved in isolation to take over the integrated population and facilitates their identification, isolation and characterization. This also implements the third phase of the shifting balance theory: iii) mixing the subpopulations back into a large population. Remarkably, after mixing, a double mutant, D385G-A382V (*BD*), sweeps the population; (**Figure 1B**) this is in sharp contrast with the standard bulk cultures where the triple mutant *ABC* sweeps the population.

Next, we determine if the escapees in drops are more fit than the escapees in bulk. To address this question, we engineered the mutations into the infectious clone and recovered mutant viruses. Next, we performed head-to-head competition of *wt*, *BD*, and *ABC* variants. We show in **Figure 2** the frequency of each of the three clones at the end of the competition. Indeed, *BD* and *ABC* are true escape variants since they outcompete *wt* under neutralizing antibody. However, without antibody, *wt* is more fit than both *BD* and *ABC*, which explains the observation that neither *BD* nor *ABC* spontaneously arise in serial passaging without Ab. More importantly, in the competition between *BD* and *ABC*, we find that the escapee from droplets is more fit compared to the escapee from bulk (Figure 2). Despite being more fit, *BD* is not detected in any of the bulk serial passaging under neutralizing Ab, while *ABC* was observed thrice. We hypothesize that the initial acquisition of the mutation D (enroute to *BD*) is limited in large populations, because epistatic interactions dictate that the virus traverse a low fitness regime, before climbing up a local fitness peak. Indeed, such unlikely pathways on the landscape can be accessed if selection is weakened by decreasing population size as Wright originally envisioned.

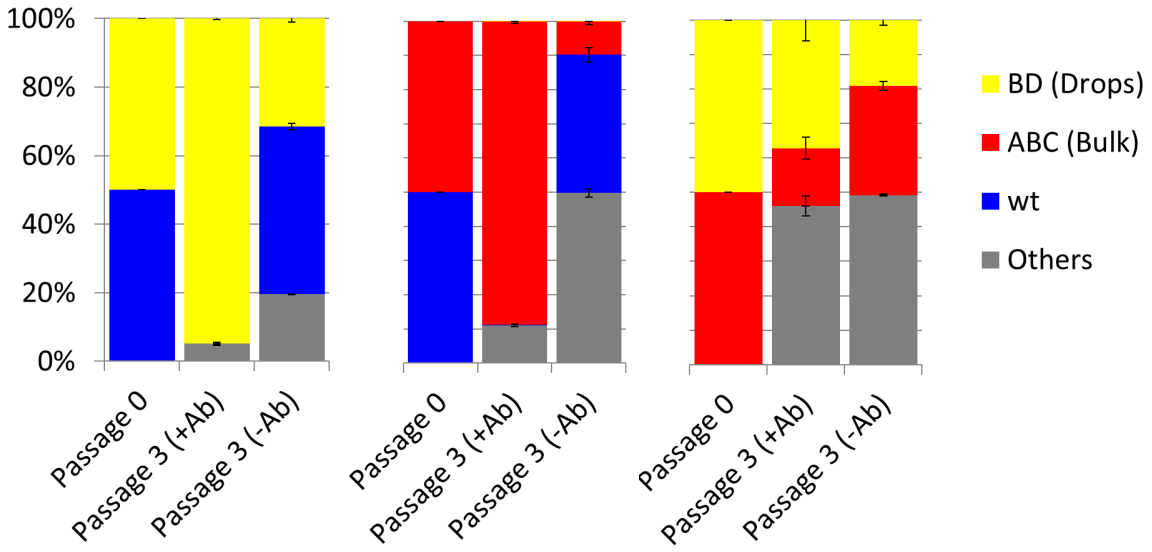

Figure 2. Allele frequency of the clone after 3 passages in competition assays. To perform pairwise competition of the clones, we mixed equal titers of the clone, propagate them for 3 passages, and then perform deep sequencing. Error bars are S.E. of 3 biological replicates for each measurement. See also Table S2.

The following supplement is available for Figure 2:

Table S2. Analysis for head to head competitions

Fitness landscape of norovirus escaping an antibody is projected onto the biophysical properties of its capsid domain

These results provide strong qualitative support to Wright's shifting balance theory showing that the evolutionary dynamics and outcome dramatically depend on the population structure. To understand these dynamics quantitatively we must first develop a tractable model for the fitness landscape of the virus. In general, fitness is expected to be a complex function of multiple traits. Instead we focus on the dependence of viral fitness in the presence of a neutralizing antibody on two biophysical properties of the P-domain: The folding energy, which is a measure of stability, and the binding affinity to the antibody, which is a measure of neutralization. While the importance of binding affinity to antibody is apparent, the universal importance of protein folding stability for bacterial and viral fitness was also shown [8,27]. This choice of variables is further supported by the fact that all the mutations of the dominant escape variants we observe in

our experiments are located within the binding site between the P-domain and the Ab, as shown by mapping the mutations on the 3D structure of the *wt* P-domain in **Figure 3A**.

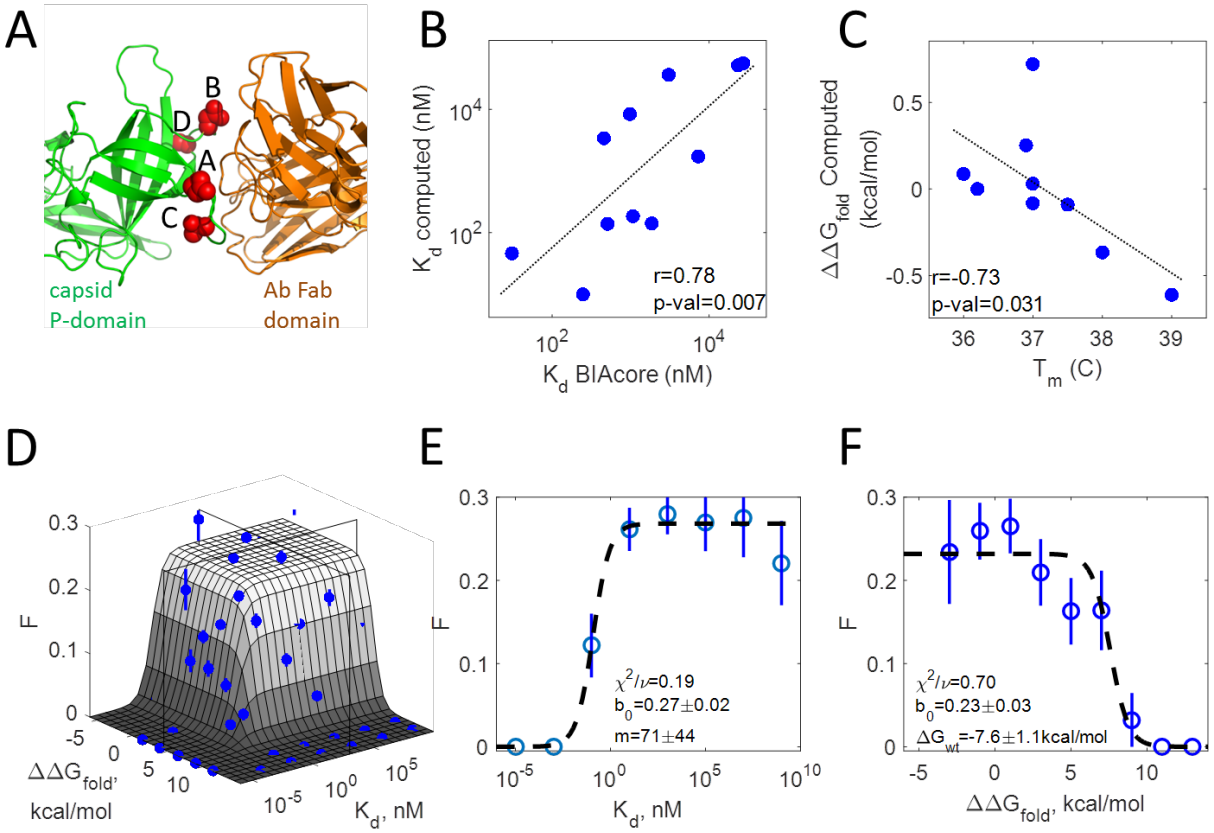

Figure 3. Fitness landscape of norovirus escaping a neutralizing antibody

A) The P-domain Antibody complex structure. The SNPs of all dominant P-domain variants (red circles) are located on the docking site of the P-domain-antibody complex (PDB ID: 3LQE). **B)** A high correlation exists between Ab dissociation constant K_d that was experimentally measured using surface plasmon resonance (SPR) and the one computed from force field calculations **C)** The anti-correlation between the experimentally measured P-Domain melting temperature (T_m) and the folding stability computed from force field calculations. Two outlier variants were excluded from the analysis. **D)** A 3D plot of the probability of infection F averaged over 2,076 distinct haplotypes binned according to their dissociation constant K_d and folding stability ΔG_{fold} (blue points) overlaid with the theoretical fit according to Eq. 1 (gray surface). Cross sections (black frames) demark the regions used for the projections in B and C. **E)** The probability of infection for all haplotypes with ΔG_{fold}<4.5Kcal/mol (cross section parallel to K_d axis in A) is projected on the K_d-F plane, binned according to their K_d (blue points) and overlaid with the theoretical fit to Eq. 1 (dashed line). **F)** The probability of infection for all haplotypes with K_d>10^3nM (cross section parallel to ΔG_{fold} axis in A is projected on the ΔG_{fold}-F plane, binned according to their ΔG_{fold} (blue points) and overlaid with the theoretical fit to Eq. 1 (dashed line). F

is determined from deep sequencing lysates of *in vitro* experiments in the presence of neutralizing antibody. K_d and ΔG_{fold} are estimated from mapping the haplotype mutations to the 3D structure of the capsid P-domain in complex with the neutralizing antibody.

The following figure supplements are available for Figure 3:

Figure 3-figure supplement 1. Distribution of viral fitness estimated from deep sequencing

A) Distribution of 3,770 growth rates calculated from the first 3 passages of all evolution experiments. **B)** The number of haplotypes that were binned in each 2D bin of **Figure 3D** is color coded with a logarithmic gray scale representation.

Figure 3-figure supplement 2. Fitness landscape of MNV-1 in the absence of a neutralizing antibody

A) A 3D plot of the probability of infection (z-axis) averaged over ~2,000 distinct haplotypes binned according to their dissociation constant K_d and folding stability $\Delta\Delta G_{fold}$ (blue error bars) overlaid with the theoretical fit according to **Equation 1** (gray surface). Cross sections (black boxes) border the regions used for the projections in B and C. **B)** The probability of infection for all haplotypes with $\Delta\Delta G_{fold} < 8kcal/mol$ (cross section parallel to K_d axis in A) projected on the K_d-z plane, binned according to their K_d (blue error bars) and overlaid with the theoretical fit according to **Equation 1** (dashed line). **(C)** The probability of infection for all haplotypes is projected on the $\Delta\Delta G_{fold}$-z plane, binned according to their $\Delta\Delta G_{fold}$ (blue error bars) and overlaid with the theoretical fit according to **Equation 1** (dashed line). Probability of infection is determined from deep sequencing lysates of *in vitro* experiments in the absence of neutralizing antibody K_d and $\Delta\Delta G_{fold}$ are estimated from mapping the haplotypes' mutations to the 3D structure of the capsid P-domain in complex with the neutralizing antibody.

To calculate the folding energy of the P-domain and its binding affinity to the antibody for each haplotype sequence, we use force field calculations based on the structural mapping in **Figure 3A**[26,28] to determine the change in folding energy $\Delta\Delta G_{fold}$ between the mutant and the *wt*; from this we determine ΔG_{fold} of the mutant by adding the folding energy of the *wt*, ΔG_{wt}. We also determine the change in binding energy, $\Delta\Delta G_{bind}$, between the mutated P-domain-Ab complex and the *wt*; from this we determine the dissociation constant $K_d = K_0 \exp\left(\beta\Delta\Delta G_{bind}\right)$ where K_0 is the dissociation constant of the *wt*. We test the accuracy of our calculations by comparing the calculated biophysical properties of the escape haplotypes to experimentally

measured properties. To accomplish this, we express and purify the P-domain of each haplotype[29] and measure its binding to the Ab to extract K_d; we also measure the melting temperature of the P-domain, T_m, which correlates inversely to ΔG_{fold} (**Table S3** and ref. [30]). The measured values of the biophysical properties of the dominant escape haplotypes correlate strongly with the calculated values of the same haplotypes, as shown in **Figure 3B-C**. Importantly, we reverse engineer the escape viruses with their haplotype sequences on the background of the *wt* for the rest of the virus and confirm that the observed mutations in the P-domain are directly responsible for their increase in fitness both *in vitro* (**Figure 4A** and **Figure 4-figure supplement 1A**) and *in vivo* in mice (**Figure 4B** and **Figure 4-figure supplement 1B**); thus, our biophysical variables are relevant for viral fitness inside the real animal host.

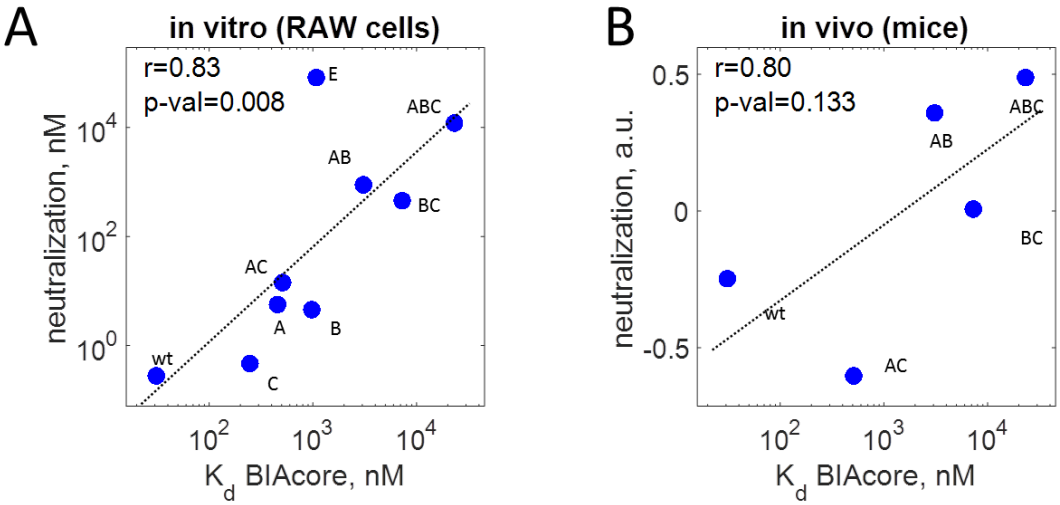

Figure 4. MNV-1 neutralization versus binding affinity of the P-domain to neutralizing antibody

A) *in vitro* neutralization of dominant haplotypes correlates to their K_d and the average ratio between them is ~120, in good agreement with the modeled value of m≈70. E: L386F **B)** *in vivo* neutralization of dominant haplotypes in mice correlates to their K_d.

The following figure supplement is available for Figure 4:

Figure 4-figure supplement 1. MNV-1 neutralization versus folding stability of the P-domain

(A) MNV-1 neutralization in RAW cells correlates with T_m. Using reverse engineering, we introduce the dominant SNPs (**Figure 3A**) on the background of *wt* MNV-1 and measure their effective neutralization (see *Methods*). **(B)** MNV-1 neutralization in mice weakly correlates with T_m (see *Methods*).

The biophysical fitness landscape describes the dependence of viral fitness in the presence of a neutralizing antibody on ΔG_{fold} and $1/(mK_d)$, where the parameter m accounts for the multiple binding sites of the capsid. To formulate the viral fitness we assume that the *wt* P-domain occurs in three specific states: folded and unbound, folded and bound, and unfolded (which is always unbound). The virus infects only when the P-domain is folded and unbound, hence, we can express the viral infectivity F at a given concentration of antibody $[Ab]$ as [31]:

$$F = b_0 \frac{e^{-\beta \Delta G_{fold}}}{1 + e^{-\beta \Delta G_{fold}} + \frac{[Ab]}{mK_d} e^{-\beta \Delta G_{fold}}}, \qquad \textbf{(Equation 1)}$$

where the numerator is the Boltzmann probability of being folded and unbound and the denominator is the partition function that sums over the probability of all three states, and $\beta=1/k_BT$ where k_B is the Boltzmann constant and T is the temperature. The function F has two regimes as shown by the surface in **Figure 3D**. For low binding affinities and stable P-domain structures, viruses are expected to infect host cells at some fixed probability, $0<b_0<1$, determined by the average effect of all remaining viral properties on the infection process, and $F=b_0$. By contrast, when the binding to the Ab is strong or when the P-domain is unstable, the virus cannot infect its host and $F=0$.

To compare the model to experiment, we use sequencing data to determine the growth rate of each virus from the change in genome haplotype frequencies between successive generations [9]. The growth rates distribute into two distinct groups with 87% of haplotypes exhibiting little or no growth and the rest exhibiting considerably larger growth. We take the first group to be non-infective, and take the second group to be infective (**Figure 3-figure supplement 1A**). For each haplotype sequence, we map the mutations to the 3D structure of the *wt* P-domain [28] and use Eris force field calculations to determine the change in folding energy $\Delta\Delta G_{fold}$ between the mutant and the *wt*; from this we determine stability of the mutant $\Delta G_{fold}=\Delta G_{fold,wt}+\Delta\Delta G_{fold}$, where ΔG_{wt} is the folding energy of the *wt*. We also determine the change in binding energy, $\Delta\Delta G_{bind}$, between the mutated P-domain-Ab complex and the *wt*; from this we determine the dissociation constant $K_d=K_0e^{\beta\Delta\Delta G_{bind}}$ where K_0 is the dissociation constant of the *wt*. We bin the haplotypes using $\Delta\Delta G_{fold}$ and K_d and calculate F for each bin from the fraction of infective haplotypes. This

binning exploits the large number of unique haplotypes to reduce the effects of errors in the calculations and of contributions from other biophysical properties. We fit the model by varying the three unknown parameters, b_0, ΔG_{wt} and the multiplier m.

We obtain excellent agreement between the model and the data, as shown by the dashed line in **Figure 3E-F**. The infectivity of haplotypes is zero at low K_d or high ΔG_{fold}, while at high K_d and low ΔG_{fold}, the landscape plateaus at F~0.25 independent of either of the biophysical coordinates. The value of $1/m \approx 1.4\%$ obtained from the fit reflects the fact that only about 3 of the 180 P-domains on the capsid have to be blocked by the Ab to prevent infection (**Figure 3E** and ref. [12]). The value of $F \sim 0.25$ at the plateau is significantly less than the expected value of b_0=1; this points to the role of factors not included in the model, such as the interaction of the capsid with the host-cell receptor, in successful infection. In the absence of the neutralizing antibody, we do not expect the fitness landscape to depend on K_d, which is indeed the case (**Figure 3-figure supplement 2**). Altogether, these results demonstrate that binding and folding energies are very good predictors of viral extinction; however they are less successful in predicting infectivity. This suggests that antibody escape and folding stability of the capsid protein are necessary but not sufficient for viral infection.

Most likely pathways on the fitness landscape predicted by protein biophysics and population genetics

To determine how the virus evolves on the fitness landscape, we use population genetics theory and calculate the ratio dN/dS, where dN is the rate of non-synonymous evolutionary rate and dS is the synonymous evolutionary rate [32] (For greater details, see section on Balance between selection and mutational drift on the fitness landscape in the Materials and Methods). As a rule, mutants affect the population more in regions where non-synonymous mutations are beneficial (high dN/dS), driving it down the gradient of dN/dS towards regions where non-synonymous mutations are deleterious (low dN/dS) and thus do not affect the population. The dN/dS ratio is a function of population size and of the fitness effect of a mutation or the selection coefficient $s=(F_{mutant}-F_{widltype})/F_{widltype}$ [32,33]. Using the fitness landscape from **Equation 1**, we calculate the dN/dS ratio on the viral fitness landscape and how it depends on population size (see Methods

and [34,35] for details). In large population sizes, dN/dS exhibits a strong gradient towards high K_d and high folding stability (**Figure 5A**). Consequently, a *wt* population that is initially unstable and neutralized by Ab is expected to evolve resistance by increasing both K_d and stability. However, in small population sizes the gradient of dN/dS is directed only towards high K_d and the same initially unstable population will evolve resistance to Ab without increasing folding stability (**Figure 5B**).

The effect of population size on the course of evolution can be explained by analyzing the balance between selection and mutational drift, the two forces driving evolution. The direction of drift and selection along the trait of folding stability and binding affinity to the antibody can be inferred from protein engineering and systematic studies on effects of random mutations on proteins. Along the axis of folding stability ΔG_{fold}, beneficial mutations increase folding stability, but random de novo mutations in proteins tend to decrease stability[36-39]. Thus, selection and drift act in opposite directions (**Figure 5A-B**, arrows), leading to mutation-selection balance. Along the axis of binding affinity K_d, beneficial mutations for the virus lead to escape from Ab (towards high K_d), and random mutations on protein interfaces perturb binding (also towards high K_d). Thus, selection and drift act along the same direction (**Figure 5A-B**, arrows).

To validate the difference in expected pathways to antibody resistance based on population size, we simulate the trajectory for 4 sequential fixations of a single mutation starting from the position of the *wt* virus population on the viral fitness landscape (see [31] for details). For large N, the simulations show that the norovirus population increases both its folding stability and dissociation constant, K_d, as it escapes the neutralizing antibody, following the trajectory shown by the solid points in **Figure 5A-B** (see also **Figure 5-figure supplement 1** for simulations in the polyclonal regime). After 4 sequential single mutations the model shows a rise of 5 orders of magnitude in K_d and an increase of 3kcal/mol in ΔG_{fold} towards stabilization corresponding to a change of 9°C in T_m. On the other hand, for small N the simulations show that the norovirus population increases its K_d to escape the antibody but exhibits only a small change in folding stability of less than 1kcal/mol, corresponding to an increase in T_m of about 2°C, as shown by the trajectory in **Figure 5B** (see also **Figure 5-figure supplement 1**).

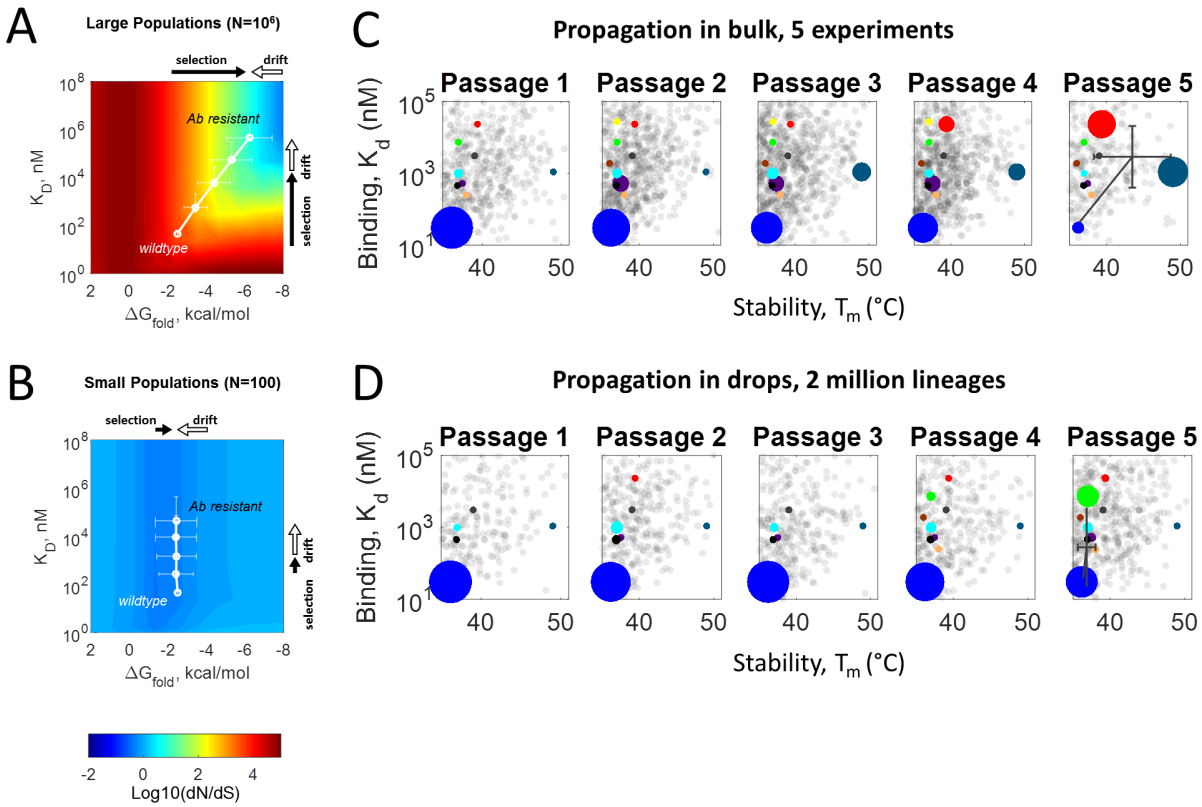

Figure 5. Most likely pathways on the fitness landscape predicted by protein biophysics and population genetics. A,B) Average stringency of selection for several population sizes (see Text and). For large population sizes, the increase in K_d is strongly coupled to the increase in T_m. However, for small population sizes, the selection for K_d is decoupled from the selection for folding stability. The white lines are the predicted trajectories from forward evolutionary simulations of an MNV population escaping an Ab, but with a P-domain which is unstable. Each trajectory is the average of 1000 independent simulations. The direction of selection (black arrows) is towards greater folding stability and weaker affinity to the antibody. Selection is strong when the P-domain is unstable and/or is tightly bound to the Ab. Selection pressure is approximately zero when the fitness landscape is flat (neutral). Along the direction of folding stability, most random mutations are destabilizing which lead to a mutational drift (white arrows) towards protein destabilization. Along binding affinity axis, most random mutations perturb the protein-protein interaction that leads to a mutational drift towards weaker binding. **C,D)** Density plots of all haplotypes grouped according to passages. Color circles denote dominant haplotypes whose biophysical properties were measured, while the remaining gray circles denote haplotypes whose biophysical properties were calculated. The size of the circle represents the allele frequency of each haplotype. Gray cross at passage 5 denotes mean and s.d. of all haplotypes.

The following figure supplement is available for Figure 5:

Figure 5-figure supplement 1. Simulations of evolutionary dynamics on the fitness landscape with a polyclonal viral population

The average folding stability and K_d is shown for 5 simulated passages of populations of 10 and 10^6 viruses. The simulation was repeated once for the large population and 1,000 times for the small population and the average of all repeats are shown (SSE error bars are not visible at<0.1%). For this simulation we used the initial conditions: ΔG_{fold}=-2.5kcal/mol, K_d=46 nM, [*Ab*]=8 nM and μ_r=0.1.

To compare the theory to experiment we plot the position of viral haplotypes evolving in 5 independent bulk passaging experiments as a function of K_d and T_m, denoting their frequencies by the size of the circles for each passage. To that end we include the calculated biophysical values K_d and T_m (gray haplotypes), using the correlation curve in **Figure 3B** to relate between ΔG_{fold} and T_m, as well as the biophysically measured values for select haplotypes (colored symbols). There is a clear trajectory as the intermediate variants evolve, having increasingly weaker affinities and higher T_m, with the escape variants at passage 5 ultimately having the weakest affinity, with an overall average of K_d ~3,000 nM and the highest P-domain stability with an average of T_m~43.5°C, as shown in **Figure 5C**. We also plot the position of viral haplotypes evolving in ~2 million independent drop passaging experiments. We observe a clear trajectory of increasingly weaker affinities while maintaining original T_m, with the escape variants at passage 5 ultimately having an overall average of K_d ~300 nM and T_m~36.8°C, as shown in **Figure 5D**. The direction of these trajectories parallels that of the simulated trajectories, supporting our choice of these explicit biophysical properties as valid and useful coordinates for the fitness landscape. This also provides us a theoretical framework with which to interpret new experimental data and to test fundamental concepts of evolutionary theory.

DISCUSSION

It is a central concept in evolutionary biology theory that the population size determines the balance between two evolutionary forces, selection and mutational drift [5,34]. The results presented here provide direct experimental evidence in support of this central concept. Moreover, the role of Wright's shifting balance theory in real evolution has been contentious and controversial because of lack of direct experimental demonstration[40]. Using the segregation in

the Evolution Chip, we show that isolated viral populations starting at a fitness valley in the presence of a neutralizing Ab are able to explore unlikely pathways on the adaptive fitness landscape and, following migration, shift the whole population to new fitness peaks (**Figure 1C**); this is a direct demonstration of the shifting balance theory.

The ability to change the course of evolution using the Evolution Chip and to predict its direction on the biophysical fitness landscape with the parameters of protein stability and binding may be helpful in addressing pandemics. These biophysical parameters are more generally applicable to other viruses; for example, both binding [41] and folding stability [8] are relevant traits for the evolution of an influenza virus in the presence of a neutralizing antibody. Prediction of future diversification of circulating viral mutants is key to developing and potentially fielding an effective vaccine prior to, or in the early stages of a pandemic. The methodology presented here may assist in proactive exploration of viral diversification to inform selection of novel viruses prior to their natural emergence, for developing viral therapeutics.

MATERIALS AND METHODS

Microfluidic devices:

We fabricate polydimethylsiloxane (PDMS) devices using photolithography and coat them with fluorophilic Aquapel (Rider, MA, USA) to prevent wetting of drops on the channel walls. Electrodes are fabricated on chip using low melting temperature solder [42]. The designs used to fabricate the devices are available in ACAD format (**Supplementary File S1**). We use OEM syringe pumps (KD Scientific, MA, USA) to drive the fluidics and a fast camera (HiSpec1, Fastec Imaging,USA) to image encapsulation and drop fusion.

Reagents

For the inert carrier oil we use HFE-7500 (3M, USA) with 1% w/w of a block co-polymer surfactant of perfluorinated polyethers (PFPE) and polyethyleneglycol (PEG)[43]. A compatible surfactant is available commercially (008-FluoroSurfactant, Ran Biotechnologies, USA). To separate the emulsion, we use a commercially available demulsifier (1H,1H,2H,2H-perfluoro-1-octanol, CAS # 647-42-7). Other chemicals were purchased from Sigma Aldrich (St. Louis, MO).

Clones and Antibodies

The plaque-purified MNV-1 clone (GV/MNV-1/2002/USA) MNV-1.CW3 [44] (referred herein as *wt*) was used at passage 6 for in vitro passaging experiments. Recombinant MNV-1 viruses containing P-domain point mutants E296K (*A*), D385G (*B*), T301I (*C*), A382V (*D*), L386F (*E*), E296K-T301I (*AC*), E296K-D385G (*AB*), T301I-D385G (*BC*), A382V-D385G (*BD*) and E296K-T301I-D385G (*ABC*) were generated as previously described [20]. The isotype control IgG directed against Coxsackievirus B4 (CV) (clone 204-4) was purchased from ATCC (HB 185). Anti-MNV-1 monoclonal antibody (mAb) A6.2 (IgG2a) was grown in Bioreactor CELLine CL 1000 flasks (Sigma-Aldrich) following the manufacturer's recommendations. Antibodies were purified over a HiTrap protein A column (GE Healthcare) according to the manufacturer's instruction, dialyzed against PBS, and stored at -20 °C.

MNV-1 P-domain mutant expression and purification

All recombinant proteins was expressed and purified as previously reported [21] with some modifications. Briefly, the P domain of MNV-1 (residues 225 to 541) was cloned into a pUC57 expression vector with NH2-terminal 6-histidine tag. The protein was expressed overnight at 20°C in Escherichia coli. The cells were subsequently lysed, and the protein was obtained from the supernatant, two steps purification by Ni column and by gel filtration on a Superdex 75 column (GE Healthcare). The proteins were dialyzed overnight at 4°C against a phosphate buffer (pH=7.8 and 20 mM NaCl).

Folding stability measurement by thermofluorescence

Thermal denaturation was carried out using melt-curve module of BioRad CFX96, and Sypro Orange dye as a probe for unfolding as described earlier [45]. The dye was added to the final concentration of 5× in a 25 µl reaction volume containing 4 µM of protein in 10 mM sodium phosphate buffer pH 7.8. To control for concentration dependence, we also performed the experiment using 2 µM concentration of protein. The T_m is estimated as the extremum of the derivative of the fluorescence signal. We report the average T_m of eight technical repeats, four for 4 µM of protein and another four for 2 µM of protein.

Binding kinetics by surface plasmon resonance

Realtime biomolecular analysis was performed by surface plasmon resonance (SPR) using a BIAcore 3000 instrument equipped with nitrilotriacetic acid (NTA) sensor chip. Purified

monoclonal antibody A6.2 was immobilized on the surface while the MNV-1 P-domain variants were the analytes. To perform single kinetic measurement, we (1) Quickinjected 20 µl of buffer (10 mM sodium phosphate buffer pH 7.8 with 0.005% surfactant to prevent minimizing non-specific interaction); (2) Kinjected 50 uL of protein with 100 s dissociation; (3) Quickinjected 10 µl of regeneration buffer (pH 2.0); and (4) Quickinjected 20 µl of phosphate buffer. For each MNV variant, we performed kinetic measurements with concentrations in the range of 0.5 to 100 µM. Data analysis was conducted with BIAevaluation package. Curve fittings were done with the 1:1 Langmuir binding model.

Cell Culture

RAW 264.7 (murine macrophage) cells were purchased from ATCC (Manassas, VA) and maintained as described previously [21,46]. Adherent cell culture medium (RAW medium) contains Dulbecco's Modified Eagle's Medium, 4 mM L-glutamine, 100 µg/mL penicillin, 100 µg/mL streptomycin, 10 mM HEPES, 10% heat-inactivated fetal bovine serum. The RAW 264.7 were adapted to suspension culture in spinner flasks for these experiments, for compatibility with drop-based microfluidics. Suspension cell culture medium (Suspension Medium) contains adherent RAW medium supplemented with sodium bicarbonate (7.5%). BSRT7 (BHK cells expressing T7 polymerase) cells were cultured in DMEM supplemented with 10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin[47]. For all viral infection experiments, suspension RAW 264.7 cells were suspended in Enhanced Suspension Medium (ESM) comprised of Suspension Medium supplemented with 15% optiprep.

Viral evolution in bulk

RAW 264.7 suspension cells are centrifuged for 5 min at 3,000 rpm and re-suspended in fresh medium at a concentration of $6x10^6$ cells/mL. In antibody neutralization experiments, virus is first incubated for 30 min at 37°C with mAb A6.2 in 200 µL PBS prior to dilution into ESM at $4x10^6$ pfu/mL, such that the final Ab concentration is 8.57 nM. One mL of cell suspension and 1 mL of virus suspension are mixed in a single well of a 12-well dish containing a sterile stir bar and incubated on stir plate in a 37°C incubator, 5% CO2, for 24 hrs. To passage viral progeny to the next generation (the progeny of the first inoculum is considered passage 0 (P0)), 250 uL of the supernatant from cell lysates of the previous generation is supplemented to 2 mL of a fresh suspension of $2x10^6$ RAW cells/mL in ESM in a 12-well dish containing a sterile stir bar and incubated on the stir plate in a 37°C incubator, 5% CO2, for 24 hrs. In antibody neutralization experiments, 8.57 nM mAb A6.2 is supplemented to the fresh suspension of RAW cells prior to

passaging. Cell lysates were harvested by 2 rounds of freeze/thaw and centrifugation for 5 min at 5,000 x g.

Viral evolution in drops

To evolve viruses in small population sizes we use a novel microfluidics system, the "Evolution Chip", which propagates ~10^6 subpopulations of 1-10 infectious particles (pfu) in distinct and non-mixing compartments. We encapsulate ~2 host cells and ~1 pfu in 100µm diameter aqueous drops in inert oil at a rate of millions per hour. The resultant emulsion is incubated under physiological conditions for one viral life cycle (~24 hours). Successful replication of the virus leads to the death of the host cell and the release of viral progeny within the drop. Thus, to enable continued passage, each drop is split and ~10% of its volume, containing viral progeny from the previous generation, is merged with a new drop containing a fresh host cell for the next generation.

Suspension-adapted RAW 264.7 cells were centrifuged for 5 min at 3,000 rpm, re-suspended in ESM at $8x10^6$ cells/mL. The suspension of cells is co-flowed at a 1:1 ratio with MNV-1 virus diluted into ESM at $4x10^6$ pfu/mL. The two aqueous phases - cell suspension and buffer - meet immediately before passing through the microfluidic drop making junction so that they only mix inside the 100um drops containing them (**Supplementary Movie 1**). For the continuous phase dispersing the drops, we use HFE-7500 Oil with 1% surfactant. Typical flow rates are 8 mL/hr for the oil and 2 mL/hr for cells and virus. Drops were collected in 1.5 mL tubes and incubated at 37°C, in 5% CO_2. Following 18-24 hr incubation at 37°C, drops are re-injected into the "evolution-chip" microfluidic device where ~10% of their volume is split off and fused with freshly formed drops containing freshly prepared suspension of RAW cells in ESM. The synchronization in the device ensures that in >95% of cases, the split from one drop fuses with exactly one newly formed drop, enabling the viral lineages to propagate in isolation. Typical flow rates are 4 mL/hr for the oil, 1 mL/hr for the re-injected drops, and 1 mL/hr for the fresh cells. The newly formed drops are collected in 1.5 mL tubes and incubated at 37°C, in 5% CO_2, while the content of the old, split drops are extracted for analysis (The progeny of the first encapsulation is considered passage 0 (P0)). In antibody neutralization experiments, 8.57 nM mAb A6.2 is supplemented to the fresh suspension of RAW cells prior to re-injection of drops.

Head to head competitions between viral strains

Next, we determine if the escapees in drops are more fit than the escapees in bulk. To address this question, we performed head-to-head competition of *wt*, *BD*, and *ABC* strains. We mixed equal viral titers of each pair of strains, passaged them for 3 rounds and then performed deep

sequencing. The competition was performed under two conditions, with and without neutralizing antibody, and over three replicates. The sequencing results of the competitions are given in **Table S2**.

RAW 264.7 suspension cells are centrifuged for 5 min at 3,000 rpm and 2 mL of a fresh suspensions of $1x10^6$ RAW cells/mL in RAW medium is dispensed in 6-well plates. In antibody neutralization experiments, virus is first incubated for 30 min at 37°C with mAb A6.2 in 200 µL PBS prior to adding it into the wells, such that the final Ab concentration is 0.08 nM. Cells and virus are incubated in a 37°C incubator, 5% CO_2, for 48 hrs. To passage viral progeny to the next generation (the progeny of the first inoculum is considered passage 0 (P0)), 200 uL of the supernatant from cell lysates of the previous generation is supplemented to 2 mL of a fresh suspension of $1x10^6$ RAW cells/mL in RAW medium in a 6-well dish and incubated in a 37°C incubator, 5% CO_2, for 24 hrs. In antibody neutralization experiments, 0.08 nM mAb A6.2 is supplemented to the fresh suspension of RAW cells prior to passaging. Cell lysates were harvested by 2 rounds of freeze/thaw and centrifugation for 5 min at 5000 x g.

In-vitro neutralization measurements

To measure neutralization in-vitro we followed Fischer et. al [12] to obtain a neutralization curve which we then fit by the equation $y = 1\Big/\left(1+\frac{k_d}{X}\right)$ where y is the viral titer, X is the *Ab* concentration and k_d is the fitting parameter.

In-vivo neutralization measurements

Mouse studies were performed in accordance with local and federal guidelines as outlined in the "Guide for the Care and Use of Laboratory Animals" of the National Institutes of Health. Protocols were approved by the University of Michigan Committee on Use and Care of Animals (UCUCA Number: 09710). Viral strains were neutralized in $STAT^{-/-}$ mice injected with 500ug mAb A6.2 and compared to their infection in mice injected with an isotype as described in [48]. The decrease in viral titers was first standardized for each tissue across viral strains, before the average over all tissues within each strain was taken as its final neutralization score. The study was performed in biological triplicates. No randomization or blinding was used.

Measurement of infectious virus titer

Viral titer was determined by either plaque assay, as described previously[49], or by TCID50 assay, as described previously [50]. TCID 50% infectivity was translated to plaque forming unit (pfu) using the conversion of 0.7 TCID50 units to 1 pfu.

Quantification of viral genome

Viral RNA isolated from experimental samples by MagMAX Viral RNA Isolation Kit (Life Technologies) was evaluated by qPCR using MNV-1-specific primer/probe sequences (Forward: GTGCGCAACACAGAGAAACG, Reverse: CGGGCTGAGCTTCCTGC, and probe: FAM/CTAGTGTCTCCTTTGGAGCACCTA/TAMARA) [21]. Takara One Step reagent (Mountain View, CA) was used to measure the genome copy of viral samples. qRT-PCR was performed on ABI real time PCR machine (OneStep Plus) using the following thermal cycling parameters: 5 min at 42° C and 10 sec at 95° C, 40 cycles of 5 sec at 95° C and 34 sec at 60° C. Titered MNV-1 viral stock and PicoGreen (quBit)-quantitated pT7 MNV 3'RZ plasmid [50] were used as standards for pfu/mL and genome copies/mL analysis, respectively.

Amplicon preparation and sequencing

Viral RNA was isolated using the MagMax viral RNA kit (Life Technologies), then subjected to DNAse digest using DNAfree Turbo DNAse (Life Technologies). SuperScript III reverse transcriptase (Life Technologies) was then used to convert vRNA to first strand DNA with specific primers tiling the entirety of MNV-1 at ~1kb increments using the following primers:

AGCCGATCACAGGCTCCTTGGC

CCATGTTGGATAAGAGGGCTGGC

ACGCACTTCCTCAACTCAGCCG

GGCCATGCTGATCCTGGCCA

CCACCAGGATGCCATCCGAGA

GTCGACATCAGCGCGTGGTATGA

CAACAGGGTGGGCACCACGTC

CAACAACAGGGCTCTCAGCATAAACCAG

Library preparation for Illumina sequencing on the miSeq platform was carried out in accordance with Illumina tech note 15044223 Rev. A using KAPA HotStart HiFi DNA polymerase 2x master mix (KAPA Biotechnologies), with primers for MNV-1 capsid amplification corresponding to MNV-1 VP1 nucleotides 848-1275 (Forward: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**GTTCATGGGTGTCCTGCTTT**, Reverse: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GGGGAGAAAGGGACCAATT**;

gene specific portion in bold). After purification with Ampure XP (Beckman Genomics), samples were quantified using Qubit high sensitivity reagents (Life Technologies), then pooled in equimolar ratios assuming specific amplification during adapter addition. Final pools of up to 96 libraries were quantified using high sensitivity Bioanalyzer reagents (Agilent) prior to sequencing 2X250 bp long paired end reads of the amplicons with miSeq (v2 chemistry, miSeq control software v2.3.0.3). Samples were supplemented with 10% phiX control library (Illumina). All reagents were used according to manufacturer's recommendations, and obtained from Life Technologies unless otherwise noted.

Amplicon sequencing analysis

Paired end reads were obtained from the MiSeq after going through the MiSeq Reporter Generate FASTQ Workflow, which demultiplexes the raw data by MID pair (removing adapter sequences in the process), and creates a separate pair of output files for each paired-end sample. These paired-end reads are aligned to the MNV-1 reference sequence (RefSeq: NC_008311) using bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/) and only reads that include the amplicon range (nucleotides 5918 to 6293 of MNV-1) without insertions and deletions are analyzed. SNPs in each of the remaining reads are saved for further analysis together with their Illumina reported quality. SNPs in overlapping regions of the paired end reads are called using the read with the higher quality. If the call in both forward and reverse reads match, the quality of SNP is taken as the sum of qualities, otherwise, the difference between the higher and lower quality is taken. Sequences with more than 20 substitutions were discarded and the remaining reads were then clustered into unique candidate haplotypes and tested for statistical significance using a hypothesis test. The null hypothesis presumed that a haplotype candidate is generated by sequencing errors in the reading of one of its "potential ancestors". Potential ancestors are haplotypes that are both "nearest neighbors" - have the smallest hamming distance to the candidate - and are more frequent - have more reads than the candidate. We used a simple generative model using a base-calling substitution rate equal to the lowest read quality recorded for any of the relevant substitutions in the candidate reads and then calculated the probability that the observed candidate was derived from this distribution (see next section on Determination of Significance).

Determination of significance from sequencing

Initially, the most abundant haplotype is assigned P-value=0. Haplotypes are tested in descending order of their frequency. For each haplotype, a group of "potential ancestor haplotypes" was defined as those haplotypes that are both "nearest neighbors" - have the smallest hamming distance to it - and are more frequent - have more reads than this haplotype. The single event probability that a copy of the candidate haplotype was generated while sequencing a copy of each potential ancestor is calculated using the quality of reads of the nucleotides that differ

between the two haplotypes: the score for each substitution address was calculated as the lowest quality over all reads sequenced in that address for the candidate haplotype. The final score was calculated from the product of all single substitution scores. The probability P_{ij} that a candidate haplotype i was generated by sequencing errors in the reading of a potential ancestor j was determined using the complementary cumulative density (CCDF) of a binomial function, with the number of trials equal to the frequency of the potential ancestor, the number of successes equal to the frequency of the candidate haplotype and the success rate taken as the single event probability. This probability is calibrated according to the probability of ancestor P_j so that $P_{ij}^{'} = 1-\left(1-P_{ij}\right)\left(1-P_j\right)$. Finally, the P-value of haplotype i is calculated as the maximum probability over all potential ancestors: $P_i = \max\left\{P_{ij}^{'}\right\}$.

Calculation of probability of infection P_{infect} from growth rates

To obtain the fitness landscape in **Figure 3**, we define the probability of infection of each haplotype as $P_{inf} = \Theta(\rho - 10^{-3})$, where ρ is the growth rate of the haplotype calculated as described below. Applying this threshold is motivated by the fact that the distribution of growth rates is bimodal – most haplotypes are either infective to some extent (ρ>=0.1) or not (ρ=0) with less than 1% of growth rates fall within the intermediate region (see **Figure 3-figure supplement 1A**). Consequently, the setting of the threshold is insensitive in the region (0,0.1).

Calculation of growth rates from haplotype frequencies

We calculate growth rates by comparing between haplotype frequencies of consecutive generations. To avoid sequencing errors, we only consider frequencies with P-value$\leq 10^{-4}$. Additionally, since our growth rates are relative to that of the wildtype *wt*, we only use data from the first 2 passages of evolution, where the *wt* is still the most abundant haplotype.

For each pair of consecutive generations that were sequenced, we calculate the growth rate ρ_i of all haplotypes that were detected in the samples using a simple linear model:

$$n_i^{'} = n_i\, p_i \qquad \text{(Equation 2)}$$

where n_i and n_i' are the copy numbers of haplotype i in the first and second generation respectively. We ignore mutation rates since mutation rates of MNV (~10^{-4}) are smaller than the sequencing errors (~10^{-3}). Assuming unbiased sampling of original copies during sequencing, we substitute:

$$n_i = f_i \cdot \Sigma n \qquad \text{(Equation 3)}$$

$$n_i' = f_i' \cdot \Sigma n' = f_i' \Sigma(n_i \cdot \rho_i) = f_i' \Sigma(f_i \cdot \rho_i \Sigma n)$$

where f_i and f_i' are the frequency of haplotype i measured by sequencing the first and the second sample respectively. When we use these substitutions in the original equation (1) we get:

$$f_i' \Sigma(f_i \cdot \rho_i \Sigma n) = f_i \cdot \rho_i \Sigma n \qquad \text{(Equation 4)}$$

$$f_i' \Sigma(f_i \cdot \rho_i) = f_i \cdot \rho_i$$

the resulting set of linear equations, one per haplotype, is a homogenous system:

$$\boldsymbol{AP} = 0 \qquad \text{(Equation 5)}$$

where the parameter matrix is $A_{ij} = f_i' f_j - \delta_{ij}$ and the growth rates vector is $P_j = \rho_j$.

The non-trivial solution solves for all growth rates, assuming a growth rate of 1 for the wild type.

Since the same haplotype may exist in more than one pair of consecutive generations, multiple growth rates can be assigned to the same haplotype, in which case the average growth rate is calculated for this haplotype. The distribution of all growth rates is plotted in **Figure 3-figure supplement 1A**.

<u>Viral fitness function based on thermodynamics of protein folding and binding</u>

We assume that the fitness F is proportional to fraction of MNV-1 P-domain that are folded and free from the antibody. If we further assume that the MNV-1 P-domain exists in three states— folded unbound, folded bound and unfolded (always unbound)— then we can express fitness F as

$$F \sim P_{\inf} = \frac{e^{-\beta \Delta G_{fold}}}{1 + \left(\frac{[Ab]}{mC_0}\right) e^{-\beta \Delta G_{bind}} e^{-\beta \Delta G_{fold}} + e^{-\beta \Delta G_{fold}}} = \frac{1}{1 + \left(\frac{[Ab]}{mK_d}\right) + e^{\beta \Delta G_{fold}}} \qquad \text{(Equation 6)}$$

where ΔG_{fold} and ΔG_{bind} are the free energies of folding and binding, $[Ab]$ is the antibody concentration, C_0 is the standard reference concentration and $\beta = 1/k_B T$. The parameter

m accounts for the added entropy due to the fact that any of the 180 P-domains on the viral capsid have a similar contribution to the neutralization of the capsid.

We note that this formulation takes into account the known experimental fact that binding induces folding. Specifically, the total fraction of folded proteins is the sum of proteins that are folded and bound to the *Ab* and proteins that are folded and free, and then normalized by total number of proteins in solution:

$$P_{folded} = \frac{\left(\frac{[Ab]}{mK_d}\right)e^{-\Delta G_f} + e^{-\Delta G_f}}{1+\left(\frac{[Ab]}{mK_d}\right)e^{-\Delta G_f} + e^{-\Delta G_f}} \qquad \text{(Equation 7)}$$

Indeed, $P_f \to 1$ when $[Ab] \to \infty$ or $K_d \to 0$. That is, when there is a sufficient amount of antibody or when the binding to the Ab is very tight, all P-domains will eventually bind to the antibody and be folded.

<u>Fitness effects of mutations (selection coefficient) on the viral fitness landscape</u>

Conceptually, evolution may be thought of as a hill-climbing process, whereby the rate of ascent is proportional to the effect of mutations on fitness. The fitness effect of a mutation is quantified by the selection coefficient $s=(F_{mut}-F_{wt})/F_{wt}$. The subscript "*mut*" and "*wt*" refers to mutant and wildtype, respectively. At the protein level, a non-synonymous mutation will change the folding and binding free energies as $\Delta G_{fold,mut}=\Delta G_{fold,wt}+\Delta\Delta G_{fold}$ and $\Delta G_{bind,mut}=\Delta G_{bind,wt}+\Delta\Delta G_{bind}$. For the fitness function (**Equation 1** and also **Equation 6**), we can derive a simple expression for the selection coefficient.

$$F_{mut} = \frac{1}{1+\frac{[Ab]}{mC_0}e^{-\beta\left(\Delta G_{bind,wt}+\Delta\Delta G_{bind}\right)} + e^{\beta\left(\Delta G_{fold,wt}+\Delta\Delta G_{fold}\right)}} \qquad \text{(Equation 8)}$$

Consider the case when $\Delta\Delta G_{fold}<<1$ and $\Delta\Delta G_{bind}<<1$. Then, using the expansion $e^x \approx 1+x$, when $x<<1$:

$$F_{mut} \approx \frac{1}{1+\frac{[Ab]}{mC_0}e^{-\beta(\Delta G_{bind,wt})}\beta(1-\Delta\Delta G_{bind})+e^{\beta(\Delta G_{fold,wt})}\beta(1+\Delta\Delta G_{fold})}$$

$$\approx F_{wt}+F_{wt}\left(\beta\Delta\Delta G_{bind}\frac{[Ab]}{mC_0}e^{-(\Delta G_{bind,wt})}-\beta\Delta\Delta G_{fold}e^{(\Delta G_{fold})}\right) \quad \text{(Equation 9)}$$

Thus, the selection coefficient is

$$s=\frac{F_{mut}-F_{wt}}{F_{wt}}$$

$$\approx \beta\Delta\Delta G_{bind}\frac{[Ab]}{mC_0}e^{-(\Delta G_{bind,wt})}-\beta\Delta\Delta G_{fold}e^{(\Delta G_{fold})}$$

$$=\beta\Delta\Delta G_{bind}\frac{[Ab]}{mK_{d,wt}}-\beta\Delta\Delta G_{fold}e^{(\Delta G_{fold,wt})} \quad \text{(Equation 10)}$$

From **Equation 10**, the magnitude of the selection coefficient of a given mutation depends on the wildtype folding stability ($\Delta G_{fold,wt}$) and binding affinity ($K_{d,wt}$); hence, there is strong epistasis on the fitness landscape. Moreover, because a single mutation can affect both folding and binding—mutations are pleiotropic in their effect to proteins—both terms in **Equation 10** contribute to the selection coefficient. This epistasis-pleiotropy link at the level of protein biophysics is the origin of the coupling between the selection for folding stability and binding.

The difference in sign for the two terms in **Equation 10** also denotes the opposite direction of selection for the two biophysical traits. Along folding stability, selection points towards more stable proteins (more negative values of ΔG_{fold}); along binding affinity, selection points towards weaker binding (larger values of K_d) (**Figure 5A**). By contrast, randomly occurring mutations tend to destabilize proteins and protein-protein complexes [38,39], rendering both $\Delta\Delta G_{fold}$ and $\Delta\Delta G_{bind}$ more positive. Thus, along folding stability, mutational drift points towards less stable proteins, opposite to the selective force; however, along binding affinity, drift also points towards weaker binding (**Figure 5A**), parallel to the selective force. We present a formal treatment of the balance between selection and drift in the next section.

<u>Balance between selection and mutational drift on the fitness landscape: Theoretical analysis</u>

We now formally combine the contribution of drift and selection and quantitatively determine the course MNV-1 evolution on the fitness landscape. We first limit ourselves to the simple monoclonal regime, where each generation is dominated by just one strain with a probability equal to its probability of fixation. To that end, we compare the probability of fixation of a non-

synonymous mutation that changes folding stability and binding with the probability of fixation of a synonymous mutation, which we assume is neutral (s=0). We denote this ratio as ω. The underlying motivation for this ratio is that when ω=1, mutations are neutral at the level of fitness, and mutational drift dominates. Specifically, [51]

$$\omega = \frac{dN}{dS} = N_e P_{fix}(s, N_e) = N_e \left[\frac{1 - \exp(-2s)}{1 - \exp(-2N_e s)} \right] \qquad \text{(Equation 11)}$$

where $P_{fix}(s, N_e)$ is the probability of fixation[4], N_e is the population size and s is the selection coefficient (**Equation 10**). s is a function of the wildtype folding stability ΔG_{wt} and binding affinity K_0 and the changes induced by the mutation to folding stability and binding $\Delta\Delta G_{bind}$ and $\Delta\Delta G_{fold}$. Because the effect of random mutations on folding stability and binding, denoted by $p(\Delta\Delta G_{fold})$ and $p(\Delta\Delta G_{bind})$, are well-characterized [38,39], we can derive the rate of protein evolution averaged over all possible effect of random mutations on folding and binding, that is,

$$\overline{\omega}(\Delta G_{fold,wt}; K_{d,wt}; N_e) = \iint \omega p(\Delta\Delta G_{fold}) p(\Delta G_{bind}) d(\Delta\Delta G_{fold}) d(\Delta\Delta G_{bind}) \qquad \text{(Equation 12)}$$

Eq. S10 defines the average stringency of Darwinian selection as a function of background, and hence the strength of epistatic interaction on the fitness landscape defined by protein folding and binding.

We show in **Figure 5A-B** (colormap) the value of the integral over the range of $K_{d,wt}$ and $\Delta G_{fold,wt}$ values. Indeed, positive selection is strongest when the *wt* P-domain is unstable (low ΔG) or tightly bound to the antibody. At high population sizes, purifying selection and positive selection dominates in the regime of low folding stability and tight binding affinity (red colors in **Figure 5A**). The prediction is that if the population starts in the unstable regime, it would migrate away from the regime where selection is strong and towards the regime where it is more neutral.

<u>Simulations of evolutionary dynamics on the fitness landscape: Monoclonal regime</u>

We complement our theoretical analysis above by performing explicit simulation of population dynamics. This evolutionary dynamics allows for the occurrence of random mutation and then selection. The algorithm is as follows:

Initialization

We start with a viral population of size N. Since the population is monoclonal, the entire population is uniquely determined by single values of folding stability and binding affinity. That

is, the population is a single point defined by the ordered pair (ΔG_{fold}, K_d). For the simulation shown in **Figure 5A-B**, the population starts at (-2.5 kcal/mol, 46 nM).

Propagation

1. A non-synonymous mutation occurs in the MNV-1 p-domain at a rate of $\mu_r = 0.1$ per replication [52]. This mutation changes the folding and binding free energies as $\Delta G_{fold,mut} = \Delta G_{fold,wt} + \Delta\Delta G_{fold}$ and $\Delta G_{bind,mut} = \Delta G_{bind,wt} + \Delta\Delta G_{bind}$. The value $\Delta\Delta G_{fold}$ is drawn from a Gaussian distribution with mean 1 kcal/mol and s.d. 1.7 kcal/mol [37,38]. The value $\Delta\Delta G_{bind}$ is drawn from a Gaussian distribution with mean 1 kcal/mol and s.d. 2 kcal/mol [37,38]. Consequently, $K_{d,mut} = K_{d,wt} e^{(\beta\Delta\Delta G_{bind})}$.
2. Calculate the fitness of the mutant F_{mut} using **Equation 1**.
3. Calculate the selection coefficient $s = (F_{\text{mut}} - F_{wt})/F_{wt}$.
4. Determine the probability of fixation [4]

$$P_{fix}(s, N_e) = \frac{1 - \exp(-2s)}{1 - \exp(-2N_e s)} \qquad \text{(Equation 13)}$$

5. Draw a random number r between 0 and 1. If $r <= P_{fix}$, the mutation is fixed and the folding stabilities and binding affinities are updated; otherwise, keep the wildtype values.
6. Repeat steps 1-5 for a specified number of fixations (In **Figure 5A-B**, we ran simulations until fixation of 4 amino acid substitutions).

<u>Simulations of evolutionary dynamics on the fitness landscape: Polyclonal regime</u>

The analyses in the previous section are performed in the monoclonal regime, where a closed formula of the stringency of selection is tractable. However, a true viral quasi-species is necessarily polyclonal because of the high mutation rate, thus, we also perform forward polyclonal regime simulation, for which the closed form formula for the stringency of selection is no longer accurate. Results are shown in **Figure 5-figure supplement 1**.

Initialization

We start with a viral population of size N. Each of the N virion in the population is characterized by its folding stability and binding affinity. Thus, the entire population constitutes a distribution of ΔG_{fold} and K_d values and migrates as a "cloud" on the fitness landscape. For the simulation shown in Fig. S4, we start with a population of N randomly chosen variants from the sequenced MNV-1 viral stock.

Propagation

1. Each of the N viruses has a chance to infect and replicate with a probability given by the fitness function **Equation 1**.

2. Each replicating virus mutates at a rate of $\mu_r = 0.1$ per p-domain per replication [52]. If a mutation occurs, the folding stability and binding affinity are updated similar to Step #1 in the monoclonal regime. The number of progenies is calibrated to N to maintain a constant population size.

3. If no viruses are chosen for replication, the sample is extinct.

4. Repeat from steps 1-3 for desired number of generations. To compare the number of generations to the number of fixations in the monoclonal simulations, we assume that each fixation requires 4 generations.

ACKNOWLEDGEMENTS

The authors would like to acknowledge C. Wilke and J. Plotkin for very helpful comments. This work was supported by Defense Advanced Research Projects Agency (DARPA) contract # HR0011-11-C-0093, by the National Science Foundation (DMR-1310266 to DW and MCB-1243837 to ES), NIH GM068670 (ES), by the Harvard Materials Research Science and Engineering Center (DMR-1420570) and by the National Natural Science Foundation of China (HZ, 81372496).

COMPETING INTERESTS

The authors declare none.

REFERENCES

1 Taubenberger, J. K., Morens, D. M. & Fauci, A. S. The next influenza pandemic: can it be predicted? *Jama* **297**, 2025-2027, doi:10.1001/jama.297.18.2025 (2007).

2 Luksza, M. & Lassig, M. A predictive fitness model for influenza. *Nature* **507**, 57-61, doi:10.1038/nature13087 (2014).

3 de Visser, J. A. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nature reviews. Genetics* **15**, 480-490, doi:10.1038/nrg3744 (2014).

4 Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624-626 (1968).

5 Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* **302**, 1401-1404, doi:10.1126/science.1089370302/5649/1401 [pii] (2003).

6 Wright, S. Evolution in Mendelian Populations. *Genetics* **16**, 97-159 (1931).

7 Sanjuan, R., Moya, A. & Elena, S. F. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8396-8401, doi:10.1073/pnas.0400146101[pii] (2004).

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

9 Acevedo, A., Brodsky, L. & Andino, R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**, 686-690, doi:10.1038/nature12861 (2014).

10 Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571-574, doi:10.1038/nature12344 (2013).

11 Nahum, J. R. *et al.* A tortoise-hare pattern seen in adapting structured and unstructured populations suggests a rugged fitness landscape in bacteria. *Proc Natl Acad Sci U S A* **112**, 7530-7535, doi:10.1073/pnas.1410631112 (2015).

12 Fischer, A. E. *et al.* A high-throughput drop microfluidic system for virus culture and analysis. *Journal of virological methods* **213**, 111-117, doi:10.1016/j.jviromet.2014.12.003 (2015).

13 Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146-2155, doi:10.1039/c2lc21147e (2012).

14 Tao, Y. *et al.* Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. *Lab on a chip* **15**, 3934-3940, doi:10.1039/c5lc00556f (2015).

15 Tao, Y. *et al.* Artifact-free Quantification and Sequencing of Rare Recombinant Viruses Using Drop-based microfluidics. *Chembiochem : a European journal of chemical biology*, doi:10.1002/cbic.201500384 (2015).

16 Zhang, H. *et al.* Isolation and Analysis of Rare Norovirus Recombinants from Coinfected Mice Using Drop-Based Microfluidics. *Journal of virology* **89**, 7722-7734, doi:10.1128/JVI.01137-15 (2015).

17 Wobus, C. E., Thackray, L. B. & Virgin, H. W. t. Murine norovirus: a model system to study norovirus biology and pathogenesis. *J Virol* **80**, 5104-5112, doi:10.1128/JVI.02346-05 (2006).

18 Ettayebi, K. *et al.* Replication of human noroviruses in stem cell-derived human enteroids. *Science* **353**, 1387-1393, doi:10.1126/science.aaf5211 (2016).

19 Jones, M. K. *et al.* Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* **346**, 755-759, doi:10.1126/science.1257147 (2014).

20 Kolawole, A. O. *et al.* Flexibility in Surface-Exposed Loops in a Virus Capsid Mediates Escape from Antibody Neutralization. *Journal of virology* **88**, 4543-4557, doi:Doi 10.1128/Jvi.03685-13 (2014).

21 Taube, S. *et al.* High-resolution x-ray structure and functional analysis of the murine norovirus 1 capsid protein protruding domain. *J Virol* **84**, 5695-5705, doi:10.1128/JVI.00316-10 (2010).

22 Katpally, U., Wobus, C. E., Dryden, K., Virgin, H. W. t. & Smith, T. J. Structure of antibody-neutralized murine norovirus and unexpected differences from viruslike particles. *J Virol* **82**, 2079-2088, doi:10.1128/JVI.02200-07 (2008).

23 Teunis, P. F. *et al.* Norwalk virus: how infectious is it? *Journal of medical virology* **80**, 1468-1476, doi:10.1002/jmv.21237 (2008).

24 Prasad, B. V. *et al.* X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**, 287-290 (1999).

25 Katpally, U. *et al.* High-resolution cryo-electron microscopy structures of murine norovirus 1 and rabbit hemorrhagic disease virus reveal marked flexibility in the receptor binding domains. *J Virol* **84**, 5836-5841, doi:10.1128/JVI.00314-10 (2010).

26 Koonin, E. V. & Wolf, Y. I. Constraints and plasticity in genome and molecular-phenome evolution. *Nat Rev Genet* **11**, 487-498, doi:nrg2810 (2010).

27 Rodrigues, J. V. *et al.* Biophysical principles predict fitness landscapes of drug resistance. *Proc Natl Acad Sci U S A* **113**, E1470-1478, doi:10.1073/pnas.1601441113 (2016).

28 Yin, S., Ding, F. & Dokholyan, N. V. Eris: an automated estimator of protein stability. *Nature methods* **4**, 466-467, doi:nmeth0607-466 (2007).

29 Wolf, Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V. & Lipman, D. J. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7273-7280, doi:0901808106 (2009).

30 Privalov, P. L. Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241 (1979).

31 Cheron, N., Serohijos, A. W., Choi, J. M. & Shakhnovich, E. I. Evolutionary dynamics of viral escape under antibodies stress: A biophysical model. *Protein Sci* **25**, 1332-1340, doi:10.1002/pro.2915 (2016).

32 Yang, Z. & Nielsen, R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* **19**, 908-917 (2002).

33 Swanson, W. J., Nielsen, R. & Yang, Q. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* **20**, 18-20 (2003).

34 Serohijos, A. W. R., Rimas, Z. & Shakhnovich, E. I. Protein Biophysics Explains Why Highly Abundant Proteins Evolve Slowly. *Cell Reports* **2**, 249-256, doi:10.1016/j.celrep.2012.06.022 (2012).

35 Dasmeh, P., Serohijos, A. W., Kepp, K. P. & Shakhnovich, E. I. The influence of selection for protein stability on dN/dS estimations. *Genome biology and evolution* **6**, 2956-2967, doi:10.1093/gbe/evu223 (2014).

36 Kumar, M. D. *et al.* ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic acids research* **34**, D204-206, doi:34/suppl_1/D204 [pii]10.1093/nar/gkj103 (2006).

37 Zeldovich, K. B., Chen, P. & Shakhnovich, E. I. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 16152-16157, doi:0705366104 [pii]

10.1073/pnas.0705366104 (2007).

38 Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L. & Tawfik, D. S. The stability effects of protein mutations appear to be universally distributed. *J Mol Biol* **369**, 1318-1332, doi:S0022-2836(07)00434-2 [pii]10.1016/j.jmb.2007.03.069 (2007).

39 Moal, I. H. & Fernandez-Recio, J. SKEMPI: a Structural Kinetic and Energetic database of Mutant Protein Interactions and its use in empirical models. *Bioinformatics* **28**, 2600-2607, doi:10.1093/bioinformatics/bts489 (2012).

40 Crow, J. F. Mid-century controversies in population genetics. *Annual review of genetics* **42**, 1-16, doi:10.1146/annurev.genet.42.110807.091612 (2008).

41 Fonville, J. M. *et al.* Antibody landscapes after influenza virus infection or vaccination. *Science* **346**, 996-1000, doi:10.1126/science.1256427 (2014).

42 Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nature protocols* **8**, 870-891, doi:10.1038/nprot.2013.046 (2013).

43 Holtze, C. *et al.* Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab on a Chip* **8**, 1632-1639, doi:Doi 10.1039/B806706f (2008).

44 Thackray, L. B. *et al.* Murine noroviruses comprising a single genogroup exhibit biological diversity despite limited sequence divergence. *J Virol* **81**, 10460-10473, doi:10.1128/JVI.00783-07 (2007).

45 Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols* **2**, 2212-2221, doi:10.1038/nprot.2007.321 (2007).

46 Wobus, C. E. *et al.* Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS biology* **2**, e432, doi:10.1371/journal.pbio.0020432 (2004).

47 Buchholz, U. J., Finke, S. & Conzelmann, K. K. Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. *J Virol* **73**, 251-259 (1999).

48 Kolawole, A. O. *et al.* Flexibility in surface-exposed loops in a virus capsid mediates escape from antibody neutralization. *J Virol* **88**, 4543-4557, doi:10.1128/JVI.03685-13 (2014).

49 Gonzalez-Hernandez, M. B., Bragazzi Cunha, J. & Wobus, C. E. Plaque assay for murine norovirus. *Journal of visualized experiments : JoVE*, e4297, doi:10.3791/4297 (2012).

50 Arias, A., Urena, L., Thorne, L., Yunus, M. A. & Goodfellow, I. Reverse genetics mediated recovery of infectious murine norovirus. *Journal of visualized experiments : JoVE*, doi:10.3791/4145 (2012).

51 Nielsen, R. & Yang, Z. Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Molecular biology and evolution* **20**, 1231-1239, doi:10.1093/molbev/msg147msg147 [pii] (2003).

52 Bull, R. A., Eden, J. S., Rawlinson, W. D. & White, P. A. Rapid evolution of pandemic noroviruses of the GII.4 lineage. *PLoS pathogens* **6**, e1000831, doi:10.1371/journal.ppat.1000831 (2010).

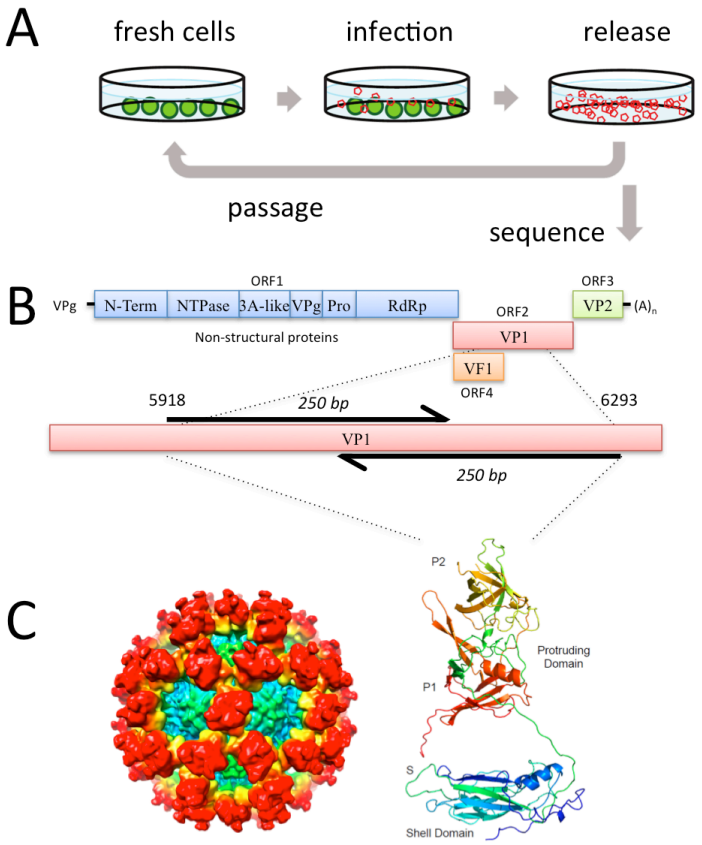

Figure 1-figure supplement 1.

1

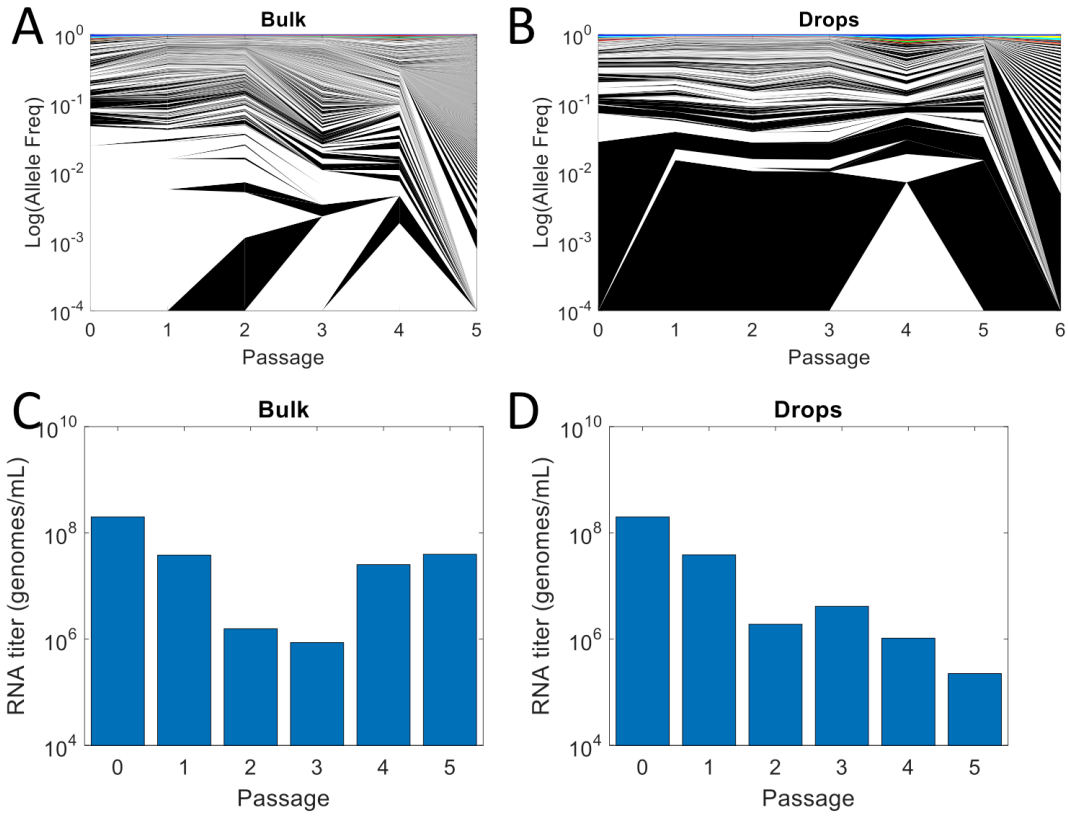

Figure 1-figure supplement 2.

1

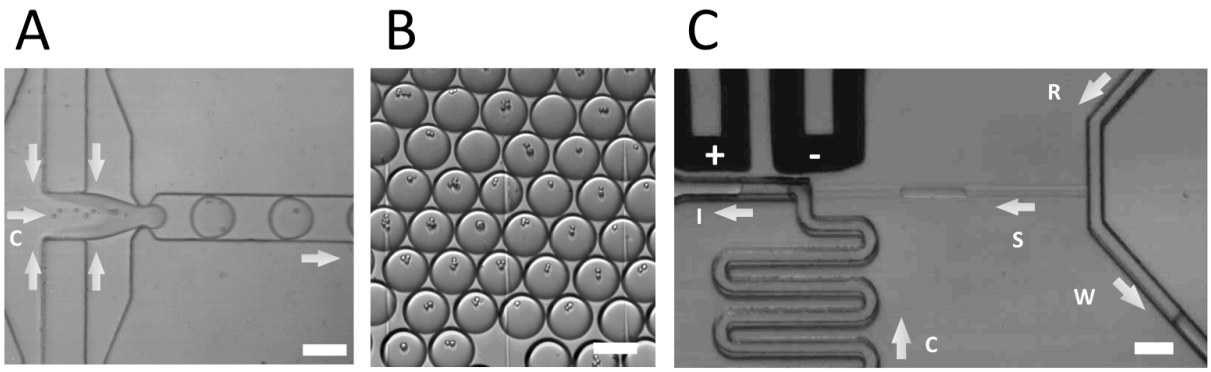

Figure 1-figure supplement 3.

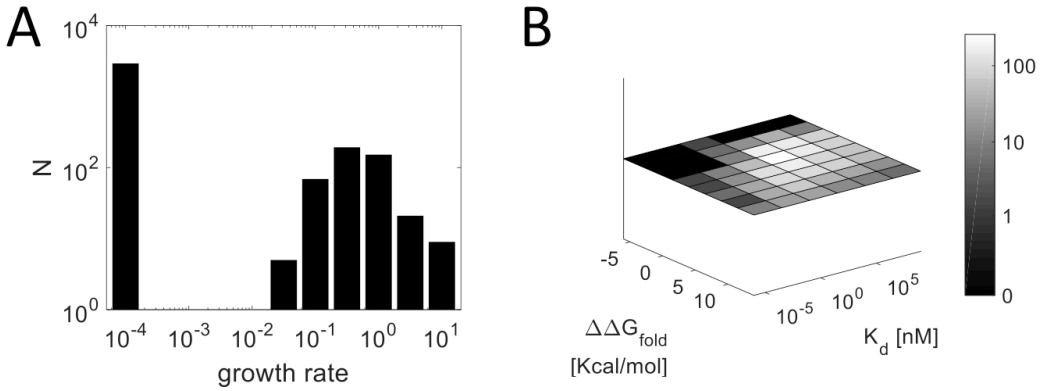

Figure 3-figure supplement 1.

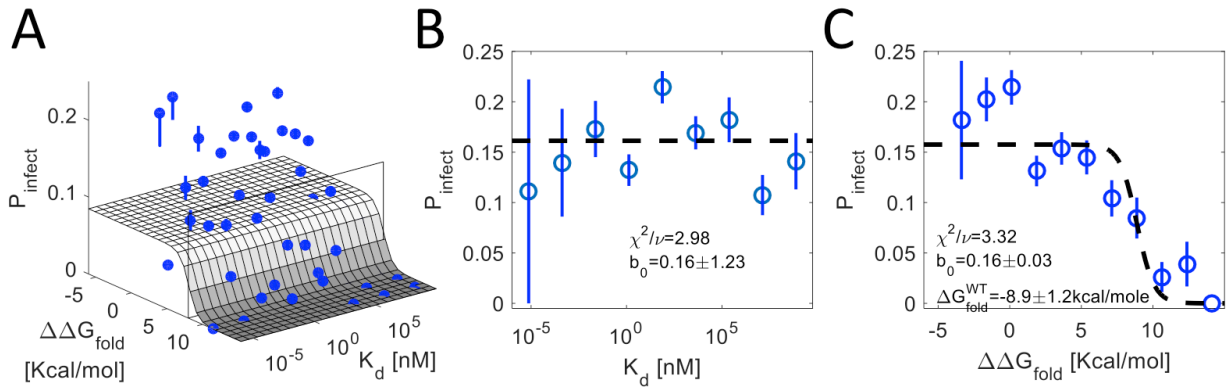

Figure 3-figure supplement 2.

1

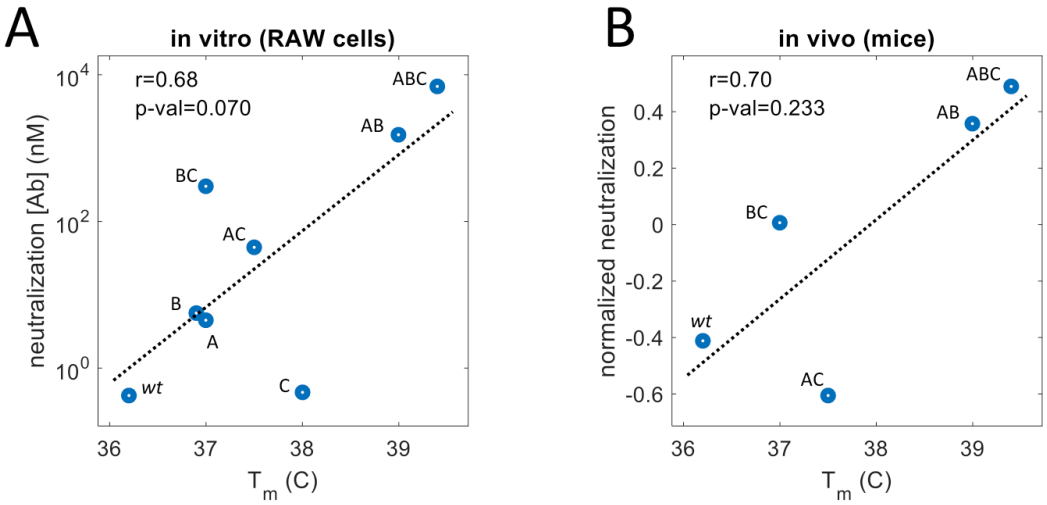

Figure 4-figure supplement 1.

1

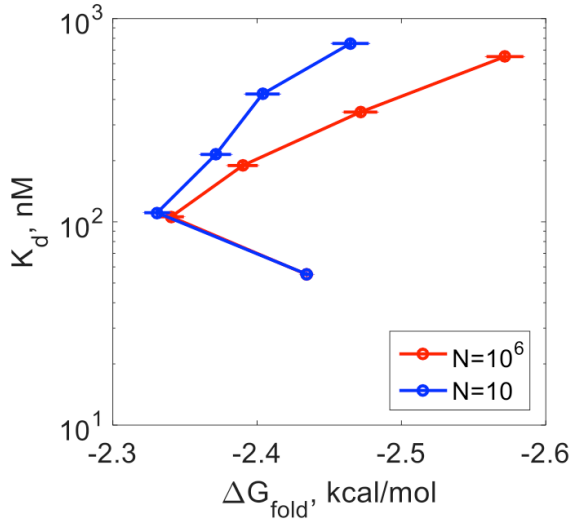

Figure 5-figure supplement 1.