Supplementary Information

Modeling how many envelope-glycoprotein trimers per virion participate in human immunodeficiency virus infectivity and its neutralization by antibody

Per Johan Klasse

Address correspondence to:

P.J.Klasse

Department of Microbiology and Immunology

Cornell University

Weill Medical College

1300 York Avenue, Box 62

New York, NY 10021

Telephone: 212 746 4827

Fax: 212 746 8340

Email: pek2003@med.cornell.edu

Redundancy of models

Supplementary Figure 1A outlines a thought experiment: imagine a homogenous population of virions with eight trimers each. At 50% functional Env (p=0.5), the average virion would have one of each possible kind of homo-trimer and three of each possible kind of hetero-trimer. We now postulate three distinct combinations of thresholds at the level of the trimer and the virion: S=1, L=7; S=2, L=4 and S=3, L=1, *i.e.* three liminal models. But all three widely different models imply exactly the same outcome: the virion is one functional trimer short of infectivity. Supplementary Figure 1B shows that pairs of functions for different combinations of thresholds at the two levels, when n remains constant, are empirically indistinguishable. Supplementary Figure 1C shows that for different combinations of n and thresholds at the trimer and virion levels, curves may also approximately superimpose. This means that even more precise data for many more x values than we currently have would not in practice be able to pick out the correct theoretical function.

Bridging incremental and liminal models

Because of the redundancy among models, data alone cannot evaluate them. But there are other criteria for judging them: some assumptions do not make biological sense. The hypotheses S=2 and S=3 are biophysically peculiar: they mean, respectively, that the first or second non-functional protomer in the trimer has zero effect on trimer function, whereas the second or third completely abrogates it, from all to nothing. Thus the S=[2,3] hypotheses do not even allow any loss in the probability of docking onto receptors when

one or two protomers lack that function. There are more plausible rivals of S=1 than S=[2,3].

One such possibility is shown in Supplementary Figure 2, where protomers contribute independently to virion infectivity. Initially, the model still postulates absolute thresholds at the virion level. Five different thresholds are illustrated. With n=9 (*i.e.*, 27 protomers), absolute thresholds are seen to yield curves far too steep to fit empirical data. Therefore, a sixth possibility is also explored: each inactivation of a protomer reduces the infectivity with the same amount but only down to the level where the virion has completely lost its infectivity and further incorporation of inactive protomers adds nothing. This model represents a bridge between absolute thresholds and incremental effects of each inactive protomer.

Supplementary Figure 3 presents other such hybrids between liminal and incremental models. There the threshold S=1 at the trimer level is kept constant. Instead, mixed liminal-incremental effects of non-functional trimers are explored at the level of the virion. Some curves indicate that, as the number of non-functional trimers rises, the first few non-functional ones do not affect infectivity. But above a certain number of non-functional trimers each additional one decreases the infectivity in a proportionally incremental manner (L=[5,…,9] and L=[3,…,9]). Other curves illustrate the converse possibility, that every non-functional trimer in the virion has a proportionally incremental inactivating effect to a limit: the further accumulation of inactive trimers has no effect, because the virion has already lost all its infectivity (L=[1,…,5] and L=[1,..,7]). Lastly, an intermediate case is shown: a broad middle zone of thresholds with equal, decremental effect of each inert trimer, omitting inactivating effects of only the first and last two non-

functional trimers (L=[3,…,7]). Such averaging over several but not all thresholds was found to fit some data sets excellently as shown in Table 3 in the main article and Supplementary Table 2.

As argued in the main article, viral heterogeneities can soften thresholds. For example heterogeneous distributions of trimers over the virion sphere may co-exist with real incremental effects, which are mathematically indistinguishable from those that merely blur apparent thresholds. In Supplementary Figure 4, an additional form of heterogeneity is illustrated. The total number of functional trimers, n, and the threshold of inactivation, L, are varied such that the minimal number of trimers required for infectivity, T, is kept constant. Obviously, the value of n greatly affects the curve. Such effects are restricted to the liminal component in models. Notably, the slope of the middle of the curve is not affected as much as in Supplementary Figure 3. A skew variation of n might arise when trimers start getting inactivated in a population of newly budded virions. As also shown in Supplementary Figure 4, this would result in a downward displacement of most of the curve.

Incremental protomeric contributions to trimer function

As part of the search for a more realistic alternative to S=1 than either S=2 or S=3 provides, Supplementary Table 1 further explores different possibilities of gradual protomeric contributions to trimer function.

The special case of totally incremental contributions by each protomer gives the linear curve that goes through (x=0, y=0) and (x=1, y=1). Its function is I=p. But one can also envisage some deviations in the slope of cumulative protomeric contributions. Then

the function to fit is *I=kp-m (Eq. 11)*. As shown in the top part of Supplementary Table 1, the coefficient *k* varies little for the data sets, between 0.9 and 1.1. The parameter *m*, determining the threshold at the lower end of the curve, varies more, from 0.024 to 0.21. This threshold indicates that virions with some active protomers are totally devoid of infectivity. It is striking that for some data sets this simple model yields excellent fits (even though the normalized endpoints (x=0, y=0) and (x=1, y=1) are included in the fitting). For example, the simple linear models fit the data better for three out of five mutants analyzed by Yang et al. (Yang et al., 2006a) than does the model proposed in that study.

The middle rows of Supplementary Table 1 present a model that makes two assumptions. First, infectivity is proportional to trimer function, *i.e.* there is no threshold at the virion level. Second, the loss of one, two or three functional protomers is allowed to have gradual effects on trimer function. Thus nonlinear regression gives the thresholds *S*=1, *S*=2 and *S*=3 different weights relative to each other. The TCLA neutralization and cleavage-defective data fit best to a simple incremental *S*=1 model. For PI neutralization and PI YU2 cleavage-defective data there are small deviations giving some weight to the *S*=2 and *S*=3 thresholds, respectively. For the PI JR-FL cleavage-defective data, however, substantial weight is given to *S*=2.

The data on receptor-binding and fusion-segment mutants all give substantial deviations from *S*=1. But only PI $CD4^-$ and PI $CCR5^-$ do so more than PI JR-FL cleavage-defective mutants.

The bottom part of Table 1 gives the R^2 for the two different models that have been proposed for these kinds of data (Yang et al., 2005a; Yang et al., 2005b; Yang et al.,

2006a), S=1 or S=2. Although these models were not given in the trimer-incremental form, because of the combined conditions n=1 and T=1 they are mathematically identical to incremental models. Thus altogether four models are listed, one of which has been advocated previously for each kind of mutant Env. But only for two out the ten data sets is that previously proposed model the best-fitting one of these four models. Still, as Figure 3 (main article) shows, to ignore thresholds at the level of the virion for n>1, as these four models do, is to over-simplify.

Compensatory effects at the two levels of liminal-incremental modeling

Supplementary Table 2 shows that well-fitting liminal-incremental models with both high (L=8) and intermediate (L=5) thresholds can fit the data in (Yang et al., 2006a). However, when L is high, h is low and vice versa. This is another illustration of the redundancy of models: a correct model cannot unambiguously be identified. However, if one assumes that the number of active trimers required should be the similar to what best fits the other kinds of mutations, *i.e.* around L=5 rather than L=8, this would strengthen the support for the hypothesis that the receptor-binding and fusion-segment mutants do have less dominant effects on trimer function than the cleavage-defective mutants.

A case of exact model redundancy arises from a corresponding compensatory effect at the two levels of modeling: The n=1, T=1, L=1 model in its S=1 form, is obviously mathematically identical to another liminal model, *viz.* n=3, L=1, T=3 and h=1 ($\Leftrightarrow$ q=p), as well as to the pure incremental model with S=1. But these three differ sharply in their virological interpretations and test implications.

Legends for Supplementary Figures

Supplementary Figure 1

A. The cartoon illustrates the potential redundancy of models that arises from the compensatory effects of opposing changes in the premises set for the two levels of the model: *i.e.,* at the level of the trimer and the level of the virion. Active subunits of the trimers are shown in red, inert ones in blue. A homogenous population of virions with the binomially determined distribution of 8 trimers is represented, with the proportion of functional Env subunits set at p=0.5. Three combinations of threshold conditions at the trimer level, S, and at the virion level, L, are presented. The condition S=1, L=7 means that only two active trimers are required for infectivity, but the depicted virion falls one trimer short. For S=2, L=4, five active trimers are required for infectivity but this virion again falls one trimer short. For S=3, L=1, all eight trimers are required for infectivity but again the depicted virion falls one trimer below the requisite level. In summary, in none of the three cases would the virion be infectious: in each one the virion falls below the requirement for infectivity by the same margin of one trimer, despite the drastic differences among the three models.

B. The pairs of curves represent theoretically distinct models that would be difficult to distinguish empirically because of their proximity. This situation can arise with combinations of different intra-trimer compositions and thresholds for virions that have the same number of trimers: [n=5, L=1, S=3] (red squares) and [n=5, L=5, S=1] (blue

squares); [n=9, L=6, S=1] (red triangles) and [n=9, L=2, S=2] (blue triangles); [n=9, L=8, S=2] (red circles) and [n=9, L=4, S=3] (blue circles).

C. Approximate redundancy can occur for distinct values of n: [n=4, L=2, S=2] (red squares) and [n=9, L=1, S=3] (blue squares); [n=4, L=3, S=2] (red circles) and [n=9, L=9, S=1] (blue circles) (C).

Supplementary Figure 2

With nine trimers per virion, n=9, i.e., 27 protomers, and absolute thresholds at the virion level but none at the trimer level (*i.e.*, each subunit functions independently of its neighbors), steep sigmoid curves result. The situation is principally similar to when n is high in Figure 2B (main article). Different values of protomeric thresholds at the level of the virion are illustrated: 1 (red circles), 8 (blue circles), 14 (green circles), 20 (red squares) and 27 (blue squares). In addition, the average for only the 14 lowest threshold values, [1,…,14], (magenta triangles) is shown. In this model, every inactivation of a protomer would decrease the infectivity of the virion in equally small steps unto a limit (specifically, when 14 protomers are non-functional). When this occurs, the virion would be completely inert. Hence any further inactivation of protomers has no additional effect.

Supplementary Figure 3

The figure shows hybrids of models with liminal and incremental effects at the virion level, but with a constant protomeric threshold at the trimer level (set to S=1). The total number of trimers per virion is also constant, n=9. In the first case, the first five inert trimers each cause an equal degree of infectivity impairment, but increasing the number of inert trimers beyond five has no further effect (L=[1,…,5]), green circles. In the second case, the first seven inert trimers have equal adverse effects on infectivity, but adding more inert trimers is without any consequence (L=[1,…,7]), yellow circles. In the third case, the virion is sensitive to inactivation in a broad, intermediate zone of inert trimer accumulation (L=[3,…,7]), blue circles. In the fourth case, the addition of the first two inert trimers has no effect, but from then on, each additional non-functional trimer decreases infectivity to an equal extent (L=[3,…,9]), black circles. Finally, the first four inert trimers added do not impair infectivity, but from five onwards, each additional non-functional trimer causes an equal decrease in infectivity (L=[5,…,9]), red circles.

Supplementary Figure 4

If n varies within the virion population but the number of trimers necessary for infection by each virion is constant (*i.e.*, L varies correspondingly so that n-L and thus T are kept constant), the effect is again to soften the apparent average threshold. The curves for the two extreme cases and the median one of pure populations of virions are shown: n=9, L=7, red circles, n=6, L=4, blue squares and n=3, L=1, green triangles. Interspersed among these curves are others illustrating the modulating effects of different virion heterogeneities. Functions for all the seven different values of n are given different

weights by the parameter *b,* in order to represent their proportions in the population of virions: An equal mix of these seven forms of virions (b=1) is shown as orange triangles. A symmetrical distribution around the median virion, n-L=2, n=[3,…,9], $I=(I_{n=3;L=1}+bI_{n=4;L=2}+b^3I_{n=5;L=3}+b^9I_{n=6;L=4}+b^3I_{n=7;L=5}+bI_{n=8;L=6}+I_{n=9;L=7}/(2+2b+2b^3+b^9)$, b=1.3 is shown (pink circles). Other values for b in this function are not represented since the curves for b=1 (orange triangles), b=1.3 (pink circles) and homogenous n=6, L=4 (blue squares, which high b values approach) are only marginally separated, and then only in the upper part of the curve. The curve for a skew heterogeneity, which could arise in a newly budded population of virions that is losing functional trimers, is also shown (black squares): n-L=2, n=[3,…,9], $I=(I_{n=3;L=1}+bI_{n=4;L=2}+b^2I_{n=5;L=3}+b^4I_{n=6;L=4}+b^6I_{n=7;L=5}+b^8I_{n=8;L=6}+b^{10}I_{n=9;L=7}/(1+b+b^2+b^4+b^6+b^8+b^{10})$, b=1.3.

Supplementary Table 1
Comparison of protomer and trimer incremental models

		Neutralization [I]		Cleavage defective[I]			Receptor-fusion defective[I]				
		TCLA	**PI**	**TCLA**	**PI YU2**	**PI JRFL**	**TCLA CD4$^-$**	**TCLA CXCR4$^-$**	**PI CD4$^-$**	**PI CCR5$^-$**	**PI fusion$^-$**
Linear[II] (Protomer incre-mental)	***k***	0.90	0.94	0.94	1.0	1.1	1.0	1.0	1.1	1.1	1.0
	m	0.13	0.12	0.21	0.18	0.12	0.091	0.092	0.024	0.064	0.070
	R^2	0.80	0.85	0.81	0.89	**0.94**	**0.97**	**0.97**	0.95	0.98	**0.98**
Trimer incre-mental[III]	***S=1***	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0[III]	1.0	1.0
	u, S=2	0	0	0	0.28	2.0	1.6	1.1	1.0	16	1.1
	v, S=3	0	0.015	0	0	0	0.032	0.1	0.21	0	0.28
	R^2	0.99	**0.9983**	0.99	**1.0**	**0.97**	**1.0**	**1.0**	**0.98**	**0.999**	**0.99**
Previous models[IV]	***S=1***	*0.99*	0.998	*0.99*	0.97	0.84	0.78	0.80	0.30	0.56	0.73
	S=2	0.34	0.50	0.51	0.67	**0.93**	0.92	0.90	0.97	0.997	0.93
Number superior[V]		0	1	0	1	3	2	2	1	1	2

[I] The neutralization data are from (Yang et al., 2005a): the mean *I* for the TCLA HXBc2 neutralized by 8 NAbs was calculated; the data for the PIs YU2, ADA and KB9 were pooled (mean of *I* for 2-3 NAbs). The data on cleavage defective TCLA and PI YU2 were derived from (Yang et al., 2005b); those PI JRFL are given in (Herrera et al., 2006). The mean *I* values for receptor-binding- and fusion-defective mutants are to be found in (Yang et al., 2006): the TCLA CD4$^-$ is HxB2 with the Env mutation D368R; TCLA CXCR4$^-$ is the HxB2 Env V3 mutant R308L; PI CD4$^-$ is YU2 D368R; PI CCR5$^-$ is the double V3 mutant R315G/L317S; PI fusion$^-$ is the TM mutant L520E. Results have been rounded off to two significant digits except where it would obscure a small difference in R^2 between the models.

[II] The equation of the linear model is $I=kp-m$ *(Eq.11)*.

[III] This model assumes proportionally incremental effects of trimers on the infectivity of the virion $I=(p^3+u(3p^2-2p^3)+v(3p-3p^2+p^3))/(1+u+v)$ *(Eq.12)*. At the level of the trimer it allows for differential relative effects of the loss of the functionality of one, two or three hetero-dimeric subunits. With the PI-CD4$^-$ data, the first round of nonlinear regression gave such large values of *u* and *v* for the PI CD4$^-$ data that the constant coefficient 1 for the *S*=1 threshold became negligible. Therefore the coefficient for the *S*=2 threshold was instead set to 1 and the relative weights given in the table were obtained.

[IV] The two incremental models with *S*=1 and *S*=2 are mathematically identical to the two special-case threshold models that have been advocated: One of those, postulating *n*=1, and estimating *L*=1, *T*=1 and *S*=1, has been proposed for neutralization data (Yang et al., 2005a) and the mixed cleavage-defective mutants (Yang et al., 2005b). In contrast, for data obtained with a different set of mutants affecting receptor interactions and the fusion-segment integrity of Env, the parameter values *T*=1 and *S*=2 were proposed (Yang et al., 2006). Unlike in the main article Table 3, the form of these models tested here allowed no flexibility: $I=q$. Therefore the R^2 values for these previous models are somewhat lower than for the corresponding flexible versions in main article Table 3. The R^2 values of these rigid previous models are given in underlined italics when they are higher than the other options in the table. The R^2 values of superior alternative models are given in bold.

(V) The last row of the table gives the number of times out of three possible that alternative models had higher R^2 values than those of the n=1 T=1 S=1 and n=1 T=1 S=2 models, respectively.

Supplementary Table 2
Illustration of compensatory effects at the trimer and virion levels of mixed liminal-incremental models

		Receptor-fusion defective mutants[I]				
		TCLA CD4-	**TCLA CXCR4-**	**PI CD4-**	**PI CCR5-**	**PI fusion-**
High-virion-threshold model	R^2	**1.0** (II,III)	**1.0** (II,III)	**1.0** (II,III)	**1.0** (II,III)	**1.0** (II,III)
Threshold	L	**8**	**8**	**8**	**8**	**8**
Weight	b	**2.3**	**2.0**	**8.8**	**4.1**	**2.0**
Protomer spectrum	h	**0.21**	**0.24**	**0.28**	**0.22**	**0.31**
Intermediate-virion-threshold model	R^2	**1.0** (II,IV)	**1.0** (II,IV)	**0.97** (II,IV)	**1.0** (II,IV)	**0.99** (II,IV)
Threshold	L	**5**	**5**	**5**	**5**	**5**
Weight	b	**1.4**	**1.3**	**2.6**	**2.0**	**1.2**
Protomer spectrum	h	**0.70**	**0.65**	**1.0**	**0.94**	**0.73**

(I) The modeled data are from (Yang et al., 2006).

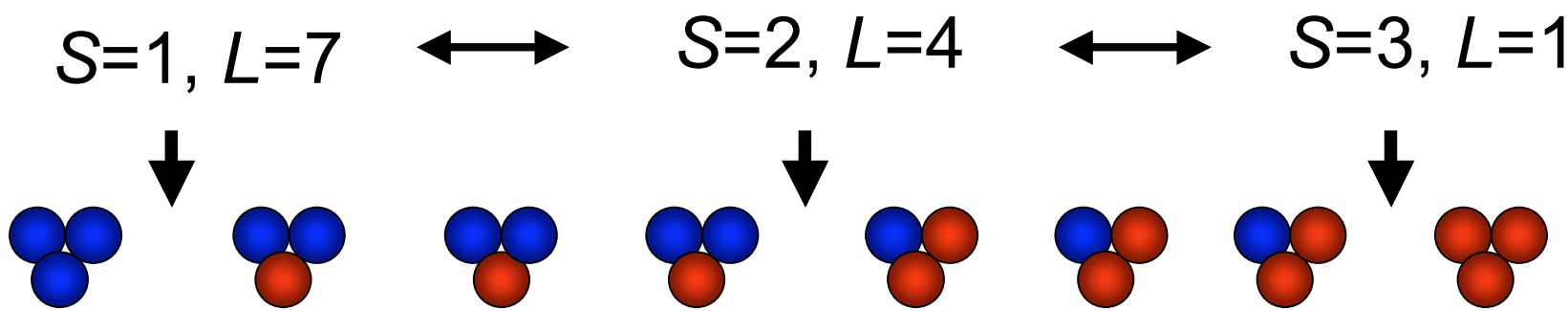
(II) The trimer function is described by the spectrum $q=hp+(1-h)p^3$ *(Eq.5)* in all equations. The infectivity equation has the general form $I=(c_1I_{L=1}+c_2I_{L=2}+c_3I_{L=3}+c_4I_{L=4}+c_5I_{L=5}+c_6I_{L=6}+c_7I_{L=7}+c_8I_{L=8}+c_9I_{L=9})/(c_1+c_2+c_3+c_4+c_5+c_6+c_7+c_8+c_9)$ *(Eq.4)*, where $I_{L=}\Sigma^{9}_{r=(9-L+1)}\,{}^{9}C_r q^r(1-q)^{9-r}$ *(Eq.1 with n=9)*. The coefficients c_L give different weights to the nine possible thresholds.

(III) $c_1=b^0$, $c_2=b^0$, $c_3=b^0$ $c_4=b^0$, $c_5=b^0$, $c_6=b^0$, $c_7=b^1$, $c_8=b^2$, $c_9=b^1$

(IV) $c_1=b^0$, $c_2=b^0$, $c_3=b^1$ $c_4=b^2$, $c_5=b^3$, $c_6=b^2$, $c_7=b^1$, $c_8=b^0$, $c_9=b^0$

Supplementary Figure 1A

Redundancy of models: empirical equivalence of different thresholds

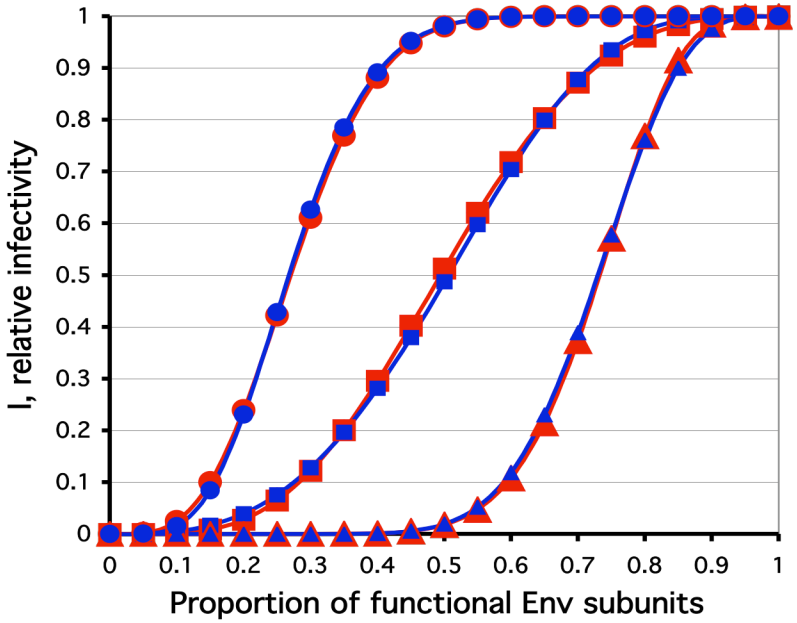


=active
=inert

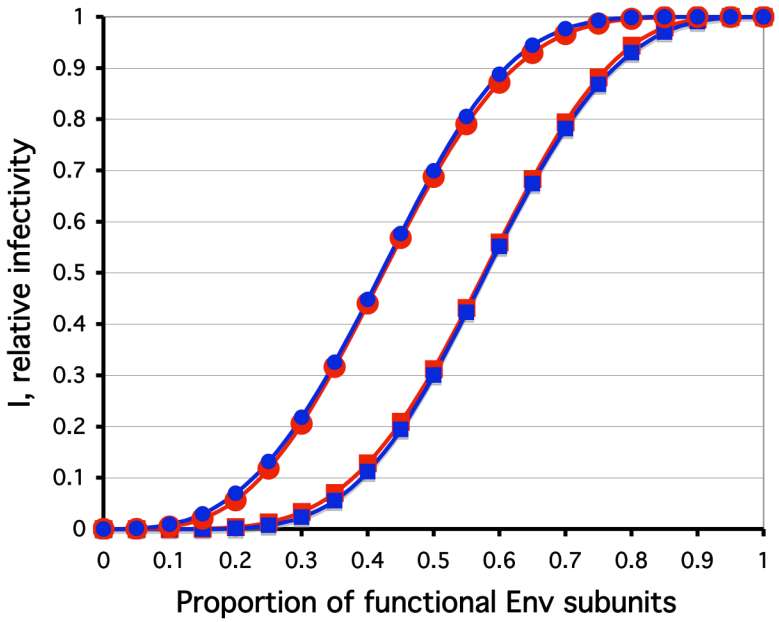
p=0.5, n=8

Supplementary Figure 1 (cont.)

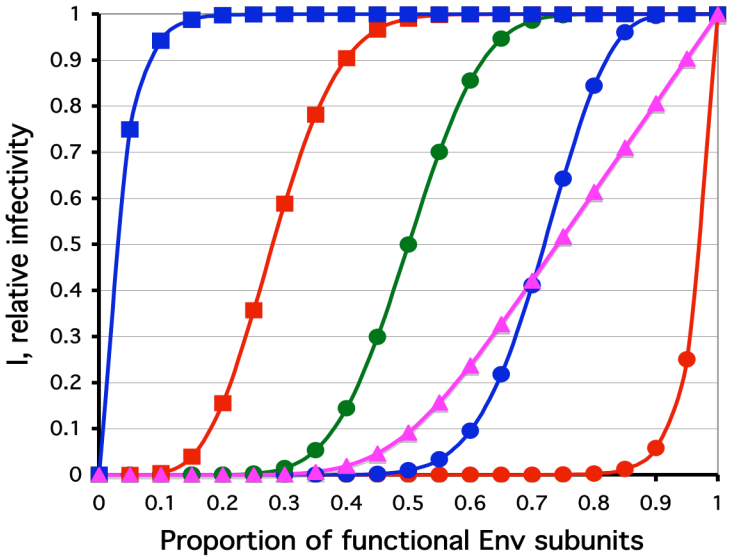
B



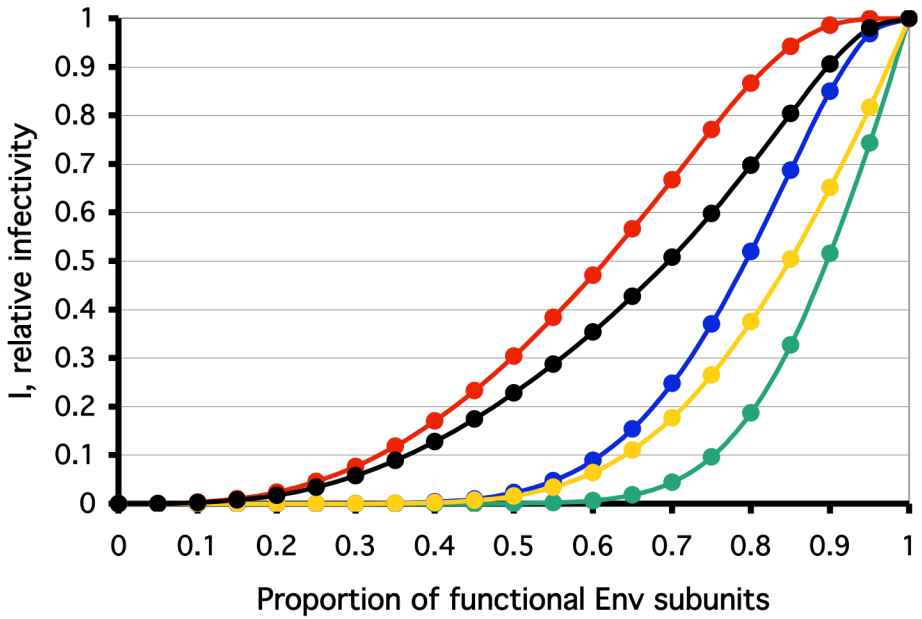
C



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4