# A coarse-grained model of affinity maturation indicates the importance of B-cell receptor avidity in epitope subdominance

**Victor Ovchinnikov** [1,*] **and Martin Karplus** [1,2,*]

[1] *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, 02138, USA*

[2] *Laboratoire de Chimie Biophysique, ISIS, Université de Strasbourg, 67000 Strasbourg, France*

Correspondence*:
Victor Ovchinnikov, 12 Oxford St, Department of Chemistry and Chemical Biology, Cambridge, MA, 02138, USA
ovchinn@fas.harvard.edu or ovchinnv@georgetown.edu

Martin Karplus, 12 Oxford St, Department of Chemistry and Chemical Biology, Cambridge, MA, 02138, USA, marci@tammy.harvard.edu

## ABSTRACT

The elicitation of broadly neutralizing antibodies (bnAbs) is a major goal in the design of vaccines against rapidly-mutating viruses. In the case of influenza, many bnAbs that target conserved epitopes on the stem of the hemagglutinin protein (HA) have been discovered. However, these antibodies are rare, are not boosted well upon reinfection, and often have low neutralization potency, compared to strain-specific antibodies directed to the HA head. Different hypotheses have been proposed to explain this phenomenon. We use a coarse-grained computational model of the germinal center reaction to investigate how B-cell receptor binding valency affects the growth and affinity maturation of competing B-cells. We find that receptors that are unable to bind antigen bivalently, and also those that do not bind antigen cooperatively, have significantly slower rates of growth, memory B-cell production, and, under certain conditions, rates of affinity maturation. The corresponding B-cells are predicted to be outcompeted by B-cells that bind bivalently and cooperatively. We use the model to explore strategies for a universal influenza vaccine, *e.g.*, how to boost the concentrations of the slower growing cross-reactive antibodies directed to the stem. The results suggest that, upon natural reinfections subsequent to vaccination, the protectiveness of such vaccines would erode, possibly requiring regular boosts. Collectively, our results strongly support the importance of bivalent antibody binding in immunodominance, and suggest guidelines for developing a universal influenza vaccine.

Keywords: germinal center, simulation, influenza, hemagglutinin, vaccination

## 1 INTRODUCTION

Viral respiratory infections remain a high source of morbidity and mortality around the world. Universal vaccines against highly-mutable viruses such as influenza and HIV are therefore highly desirable. However,
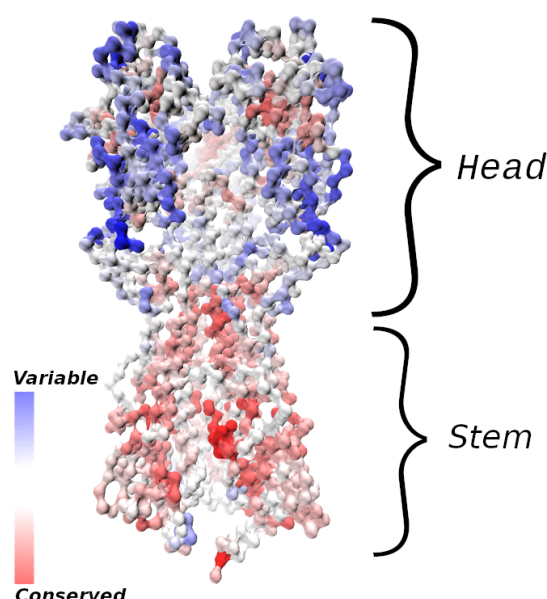
**Figure 1.** Influenza hemagglutinin (HA) spike protein colored by residue conservation. Sequences of avian, swine and human influenza type A spike proteins were downloaded from the NIH influenza research database(3); conservation was computed in MATLAB(4) after clustering the sequences to 97% identity and multiple alignment. The HA structure was taken from PDB entry 3LZG(5), and the image was generated using Visual Molecular Dynamics.(6) Only the $C_\alpha$ atoms are shown.

23  their development is complicated by the high antigenic drift of the viral genomes, which stems from low
24  replicative fidelity of viral NA polymerases. In the case of influenza, there also exists the possibility of
25  antigenic 'shifts', whereby genome segments from different viral strains are reassorted in a host co-infected
26  with multiple viruses. Such reassortments can result in pandemic strains against which preexisting immunity
27  in the affected population is low. For example, the H1N1 1918 influenza pandemic was likely caused by a
28  reassortment between avian and human influenza strains, the 2009 H1N1 pandemic was probably caused
29  by a triply reassorted human, avian, and swine strains,(1) and the SARS-CoV2 pandemic is believed to
30  have originated from a bat virus adapted to infect humans and other hosts.(2)

31    Despite the high mutation potential of many viruses, epitopes that are relatively well conserved between
32  different strains, and antibodies that bind to them, have been discovered. In the case of the influenza
33  hemagglutinin (HA) protein (see Fig. 1), the majority of conserved epitopes are located on the stem or base
34  of the HA(7, 8); less frequently, they are found in the HA head, *e.g.*, at the interface between the heads
35  of HA monomers(9, 10) in a homotrimer, or in the vicinity of the sialic acid receptor(11, 12). There are
36  different plausible reasons for the increased sequence conservation of these regions, such as (i) maintenance
37  of function (*e.g.* the sialic acid receptor is required for host cell entry, and a properly folded and probably
38  somewhat rigid HA helices are needed for fusion with the host cell) and (ii) a lack of evolutionary pressure
39  to drive escape mutants(11); the stem epitopes and the occluded epitope(9) are less accessible to antibodies,
40  and thus experience less selective pressure than the exposed HA head. Because antibodies (Abs) that bind
41  conserved epitopes could confer protection against many different viral strains, the elicitation of such
42  cross-reactive, or broadly-neutralizing Abs (bnAbs) has been a goal of vaccine development.(7) BnAbs
43  have been isolated from animal and human subjects in response to natural infection or vaccination,(8)
44  and some have been engineered(13) to have very high breadth, providing simultaneous protection against
45  phylogenetically distant group I and II HAs.(14, 15, 13)

46    Unfortunately, such bnAbs tend to be rare and are not well boosted in infections(15, 16, 17, 18). Further,
47  most of them have lower neutralization potency, compared to strain-specific Abs. However, bnAbs have
48  different mechanisms of protection, *e.g.* labeling infected cells for destruction by Ab-directed cytotoxicity
49  (ADCC), or inhibiting membrane fusion after phagocytosis by the host cell,(7) rather than blockage of
50  viral binding to cells, so that direct comparisons between different Abs are not always informative. Many
51  bnAbs have been shown to provide heterosubtypic protection in passive immunization studies,(8) indicating
52  that their elicitation by a vaccine could provide broad immunity. In the case of influenza, titers of stem-
53  directed Abs have been shown to increase with age, correlating inversely with incidence of symptomatic
54  infection.(19)

55    The antigenic epitopes targeted by rare bnAbs are labeled immunosubdominant, because they usually
56  are not the primary targets of immune responses. Since durable vaccine responses against highly-mutable
57  pathogens will need to overcome immunodominance, *i.e.* to focus on conserved subdominant epitopes,
58  various possible causes for subdominance have been considered. They include (i) epitope autoreactivity,(12)
59  (ii) low frequency of precursor antibodies able to bind the epitopes in question,(20) (iii) epitope shielding
60  by glycans,(10) (iv) preexisting immune memory,(20, 18) poor steric accessibility of epitopes(21), (v)
61  inability to recruit sufficient T-cell help in germinal centers (GCs) due to a lack of compatible MHCII
62  (major histocompatibility complex of type 2) epitopes,(17) and (iv) entropic dominance of distracting
63  epitopes.(22, 21, 23) Although multiple effects are likely to contribute, it is of interest to identify the most
64  important ones. In the context of influenza HA, especially in regard to the stem and the occluded interfacial
65  epitope of Watanabe et al. (9), low steric accessibility appears to be a particularly important factor.

66    Andrews et al. (20) found that some stem-directed Abs bound virus particles with an affinity that was
67  an order of magnitude lower than that for recombinant HA, which was interpreted in terms of reduced
68  accessibility of stem epitopes of whole virions. Amitai et al. (21) used molecular simulations to map the
69  accessibility of HA epitopes to antibody binding on model virus-like particles (VLPs), and predicted that
70  the rate of bivalent (avid) antibody binding to stem epitopes is much lower than that for head epitopes. The
71  importance of bivalent binding in antibody maturation was demonstrated elegantly by Kanekiyo et al. (24),
72  who co-displayed HA receptor binding domains (RBDs) from different influenza strains on nanoparticles
73  (NPs). The heterotypic NPs (RBDs from multiple strains on an NP) elicited Abs with higher breadth than a
74  mixture of homotypic NPs (RBD from a single strain on an NP). Because the vaccines differed only in the
75  geometric arrangement of the antigens, and not in composition or proportion, the results imply that mosaic
76  display confers an advantage to cross-reactive Abs *via* bivalent binding(24). Similar results were obtained
77  recently for the SARS-CoV2 RBD.(25)

78    Motivated by these studies, we employed a coarse-grained computer model of affinity maturation (AM)
79  in GCs to investigate how avidity differences in the binding of B-cell receptors (BCRs) to their cognate
80  antigens (AGs) could influence the patterns of immunodominance observed experimentally. Despite the
81  coarse-graining, the model allows comparisons between multiple competing epitope/paratope pairs, in
82  terms of B-cell and Ab production and affinity for antigen. We first show that the model is in qualitative
83  agreement with experimental observations of the basic properties of GC reactions, and of the subdominance
84  of influenza stem epitopes, whose predominant mode of binding to BCR/Ab is assumed to be monovalent.
85  We then use the model to simulate multiple exposures to antigen and interpret the results to propose
86  strategies for overcoming immunodominance. We suggest that bnAbs that target subdominant epitopes are
87  most likely to be elicited by increasing effective epitope concentration *via* design of custom immunogens
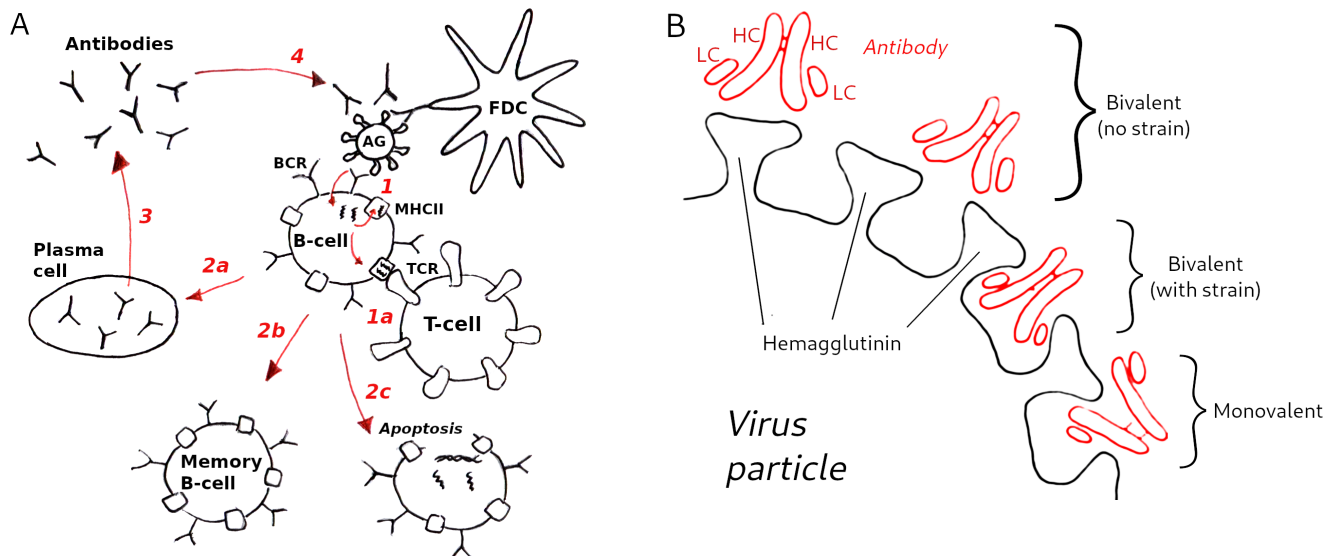
**Figure 2.** Schematic of the GC model used in this study. (A) Model overview: (1) B-cells are activated upon binding to antigen presented on follicular dendritic cells (FDCs, not explicitly modeled); (1a) in the optional T-cell help model (see text) B-cells are activated when the major histocompatibility complex receptor (MHC2) binds to the T-cell receptor (TCR); B-cell activation rescues B-cells from apoptosis, allowing them to mutate and proliferate; (2) depending on the activation signal, B-cells can differentiate to plasma cells (2a), to memory B-cells (2b), or undergo apoptosis (2c); (3) Plasma cells secrete antibodies (Abs), which also compete with B-cell receptors for antigen (4), which is the essential aspect of the Ab feedback model(26) (see text). (B) Hypothetical modes of antibody binding to influenza spikes; bivalent binding without strain (top) corresponds to cooperative binding by antibody arms; bivalent binding with strain (middle) corresponds to noncooperative binding; monovalent binding (bottom) is assumed to be the dominant mode of binding of anti-HA stem antibodies (see Methods for details).

88  or cocktail composition. The present model also predicts the resulting immune memory to be short-
89  lived, which suggests that regular boosts may be required for vaccines composed of immunosubdominant
90  epitopes.

91     Because the model and its computer implementation involve many technical aspects, we include the
92  detailed Methods section at the end of the paper, after the Results and Discussion, which are presented
93  next.

## 2 RESULTS

94  A visual schematic of the model is given in in Fig. 2, and a brief overview is presented below. Complete
95  details of the model equations and parametrization are given in Methods, and the parameter values used
96  are listed in Table 1. Our approach is related to the differential equation models of Kepler and Perelson
97  (27) and listed Oprea and Perelson (28). These models achieve a balance between biological detail and
98  computational complexity: (i) they are sufficiently accurate to model the population sizes of B-cell receptors
99  (BCRs) with different affinities, represented by discretized affinity classes; (ii) they contain a relatively
100 small number of parameters, and (iii) they are simple enough to permit a large number of simulations to
101 study the effects of different initial conditions or model parameters.(27, 28) More sophisticated approaches
102 exist, which model individual interactions between immune complexes (ICs), Ags, BCRs and T-cells,
103 aiming to capture GC reactions in higher detail.(29, 30) However, their use would be computationally
104 prohibitive for this study. For example, modeling the case in which the initial BCR affinity to an antigen
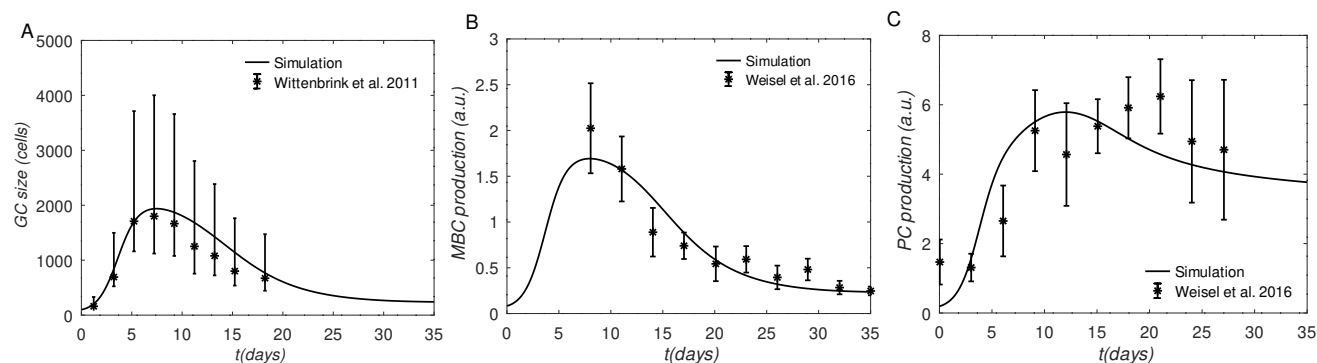
**Figure 3.** Comparison of simulation and experiments. A: Total B-cells, B: Memory B-cell production rate, C: Plasma cell production rate. Experimental data for panel A was generated from the GC cross-sectional areas plotted in Fig. S1B of Ref. 36, and converted to B cell counts as done in Ref. 37; The lower and upper error bars in panel A correspond to 30% and 70% quantiles, respectively ; experimental data for panels B & C was taken from Fig. 4 of Ref. 37, who obtained raw data from Weisel et al. (38); the error bars in B & C correspond to approximately one SD, estimated from the range of the data, which we assumed to be Gaussian-distributed.

105 was drawn randomly from a distribution involved 8400 germinal center simulations, each 35 days in
106 duration (see Results).

107    As in other models of affinity maturation,(31, 32) the interaction between BCRs and AGs is represented
108 by a Langmuir isotherm that uses two equilibrium association constants, which represent binding of the
109 first and second antibody arm. Unlike the more detailed models,(33, 21) the present model does not have an
110 explicit structural component, and the effect of bivalency is reflected in the assignment of the association
111 constants. Further, we simplify the biology by not distinguishing between centroblasts and centrocytes,
112 nor between light and dark zones (LZ *vs.* DZ). Instead we consider the overall proliferation and death
113 rates. This choice is motivated by the findings that the differences between centroblasts and centrocytes,
114 and, more generally between the LZ and DZ, are smaller than previously thought;(34) *i.e.*, the physical
115 boundary between LZ and DZ is rather diffuse, selection and proliferation can occur in both zones, albeit
116 with different rates, and T-helper cells appear to be present in both zones, albeit in different proportions.
117 Because we are interested in the dependence of the immune response on the number of competing BCRs
118 with different binding valency and affinity, the model is parametrized to predict the following quantities: (i)
119 the number of GC B-cells, (ii) the number of memory B-cells (MBCs), some of which can be recruited into
120 secondary GCs, (iii) the number of plasma cells (PCs) that secrete Abs, (iv) the number of Abs, which
121 compete with BCRs for antigen and implicitly regulate GC size,(26) and, optionally, (v) the T-helper cell
122 population (which is discussed in SI Appendix). The species (i)–(iv) are distributed into affinity classes,
123 while the T-cell model uses two equilibrium affinities, those between TCRs and MHCIIs loaded or unloaded
124 with peptides. The T-cell model is related to the T-cell expansion model of Mayer et al. (35). We consider
125 the model to be optional for the simulations performed here, because qualitatively the same results were
126 obtained by assuming that the amount of T-cell help is equal to BCR activation by binding to antigen.

127    We first show that the model captures some experimentally determined properties of GCs. In Fig. 3
128 we compare the model results obtained with a single BCR/Ag pair to the average GC size reported by
129 Wittenbrink et al. (36), and the MBC and PC production rates determined by Weisel et al. (38). The
130 experimental data were also used by Pélissier et al. (37) to parametrize their stochastic GC model, which
131 was used to explore the mechanism of clonal bursts. We note that the Wittenbrink et al. (36) observed very
132 high CG size variability, as reflected in the experimental error bars (Fig. 3); thus, the agreement between

133  our model and the experimental average should be considered as a qualitative validation, as the model does
134  not capture GC size heterogeneity.

135    One apparent disagreement is that rate of PC production in our model is accelerated by several days,
136  compared to the data of Weisel et al. (38) (see Fig. 3C). Although the experimental MBC and PC production
137  in Fig. 3B & C, respectively, correspond to the same time, matching the PC production rate in the simulation
138  required shifting the experimental measurements by several days (compare Fig. 3B with Fig. 3C). However,
139  the discrepancy is not expected to affect our results significantly, because we are interested in comparing the
140  relative MBC output from different lineages at the end of GC reactions, which are simulated consistently,
141  *i.e.* using the same model.

142    Next, we investigate the effects of BCR binding avidity under different scenarios. First, we compare
143  the growth rates of three noncompeting B-cell lineages, which differ only in the value of the second-arm
144  equilibrium binding constant $K_{eq}^2$ (written as $K^{i2}$ for lineage $i$ in Eq. (4) of Methods). This idealized
145  scenario corresponds to three GCs evolving independently, which is the noninteracting, $o$=0, case (see
146  Methods, Sec. 4.0.5), each completely dominated by a single B-cell lineage.

147    We compare the behavior of lineages with three regimes of bivalency, corresponding to $K_{eq}^2$=0, $K_{eq}^2$=$K_{eq}^1$,
148  and $K_{eq}^2$=10$K_{eq}^1$, which we denote, respectively, as the monovalent, noncooperative, and cooperative
149  binding cases. The justification for the chosen values is discussed in the Methods Sec. 4.0.2.

150    The results in Fig. 4A show that the monovalent BCR has a significantly slower rate of growth compared
151  with the bivalent BCRs, due to the reduced activation by Ag assumed by the model. In the monovalent case,
152  the maximum B-cell count is less than a third of that in the bivalent cooperative case, and less than half of
153  that in the noncooperative case; this lower maximum is also reached later; *i.e.*, after 20 days, *vs.* 5 to 7 days.
154  Further, the MBC count at the end of the GC is reduced by half in the monovalent case, compared with
155  the cooperative case (Fig. 4B), and the final average affinity of the B-cell population is several-fold lower
156  (Fig. 4C). The delayed peak in the monovalent response is qualitatively consistent with the observations of
157  Tan et al. (17), who noted that the anti-influenza stem Ab response was delayed by a week relative to the
158  overall AB response, and that the B-cell response was several-fold lower in magnitude.

159    Although the noninteracting GC case is illustrative, it is idealized, because in affinity maturation there
160  will generally be many different B-cell lineages competing for epitopes on the Ags, or for T-cell help
161  signals. Thus, a more realistic model of a single GC needs to account for competition between different
162  lineages. As described in Sec. 4.0.5 we model this competition using an *occlusion* parameter $o$, with $o = 0$
163  for the fully noninteracting case, presented above, and $o = 1$ for the fully interacting (competing) case.

164    We describe the results of the fully interacting case ($o$=1) next; however, we note that a more realistic
165  description of the overall GC reaction would probably involve intermediate values of $o$. For example, if,
166  in some cases, different B-cells are able to bind to different epitopes on the same AG simultaneously,
167  the occlusion parameter would need to be less than 1, which would correspond to decreased competition
168  between the B-cells. Moreover, when modeling multiple GCs, it may be necessary to include the possibility
169  that the Abs secreted by the plasma cells can diffuse across many GCs, and compete with the 'local' BCRs
170  for AG.(26) This scenario would correspond to indirect competition between different B cell lineages in
171  different GCs, which might be modeled with some optimized intermediate (though generally unknown)
172  occlusion value. A possible starting point for estimating such a value could involve competitive binding
173  experiments using antibodies specific for the different epitopes.

174    The simulation of the fully interacting case shows that the population disadvantage of the monovalent
175  lineage is further increased (Fig. 4)D–F. The peak B-cell concentration of the monovalent lineage is more
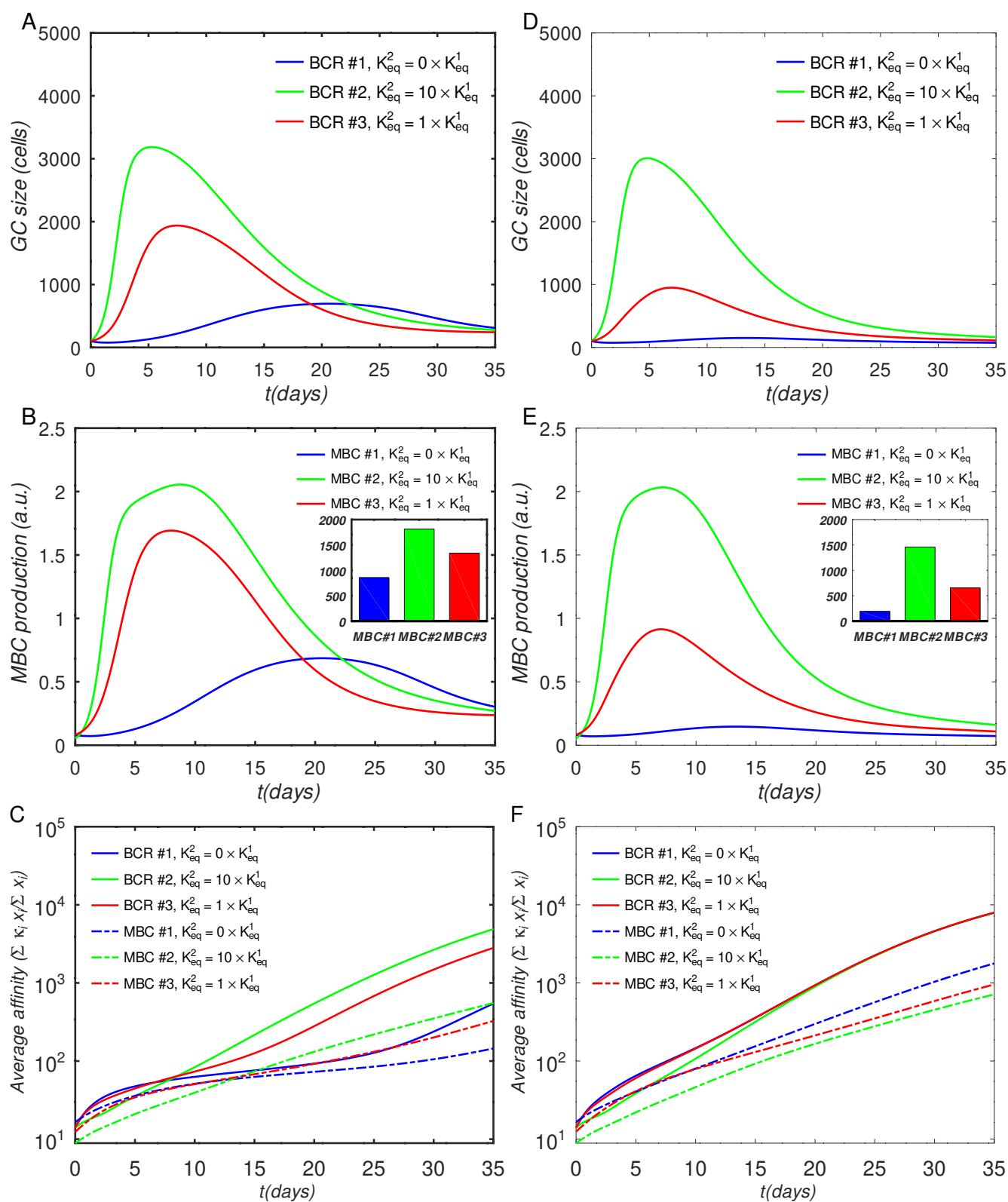
**Figure 4.** Effect of antibody valency and epitope occlusion on GC properties. Left column (A–C): noninteracting B-cell case ($o = 0$); Right column (D–F): fully interacting B-cell case ($o = 1$); A,D: Total B cells; B,E: Memory cell production rate; insets: total MBC population at end of simulation; C,F: Average affinity of B-cells and MBCs. For the definition of occlusion $o$, see Sec. 4.0.5.

176 than 10-fold lower than that of the bivalent cooperative lineage (panel D). These results are to be expected,
177 because the more rapidly proliferating lineage occludes the Ag, effectively reducing the amount of epitope
178 available to the the monovalent lineage. The average monovalent MBC production is lower by about a
179 factor of eight relative to the bivalent cooperative case (panel E).

180  It is noteworthy that the average affinities of the three BCRs are indistinguishable after about 20 days
181 after initiation of the GC reaction in the fully competing case (panel F). Further, the BCR affinities in this
182 case at the end of the simulation are higher than those in the $o$=0 case (Fig. 4C *vs*. Fig. 4F). This result is
183 understandable in terms of increased competition for survival inside the interacting ($o$=1) GC. The fact
184 that binding by one BCR occludes access to other epitopes implies that the effective epitope availability
185 is decreased for all BCRs. A decrease in the available binding sites increases the selection pressure on
186 the BCRs, leading to the survival of the higher-affinity lineages. We will return to this point when we
187 investigate the effects of varying epitope concentration explicitly. For completeness, simulation results
188 with intermediate values of occlusion are shown in Fig. S1.

189  As discussed in the introduction, the reason for targeting immunosubdominant epitopes such as the
190 influenza HA stem or the interfacial epitope(9) in vaccinations is their association with the elicitation of
191 bnAbs, which are likely to provide protective immunity against future strains. In the context of the present
192 simulations, such pre-existing protective immunity can be modeled by increasing the initial affinity of the
193 monovalent antibodies, while keeping the other affinities unchanged. This gives the monovalent antibodies
194 a survival advantage. A physiological rationale of setting a higher initial affinity of monovalent Abs could
195 be that a significant proportion of monovalent (*e.g.*, anti-HA stem) MBCs are recruited into secondary GCs,
196 where the higher initial affinity allows the MBC-derived blasts to compete more effectively with naïve
197 bivalent B-cells. To model this scenario, we shifted the initial affinity distribution of the monovalent Abs
198 by about two orders of magnitude towards higher values (see the distributions in Fig. 5D), and repeated the
199 simulations for the fully interacting $o$=1 case (the $o$=0 case is shown in Fig. S2).

200  Figure 5A shows that the population of the monovalent B-cells increased several-fold as compared
201 with Fig. 4D (no advantage), such that, at their peak, these B-cells were almost as numerous as the
202 bivalent noncooperative ones. However, even the large affinity advantage was insufficient to overcome
203 the dominance of the cooperatively-binding Abs in terms of the total MBC response, which was still
204 significantly lower in the monovalent case (Fig. 5B). Nachbagauer et al. (39) found that anti-HA stem
205 immunity can be elicited or boosted upon immunization with chimeric HA constructs with HA heads
206 to which the host is naïve, fused to HA stems against which there is preexisting immunity. In other
207 studies,(40, 16) it was reported that boosting with HAs from pandemic, rather than with seasonally-drifted
208 strains, boosted anti-stem immunity more effectively. The authors' interpretation of the results was that the
209 vaccinations boosted preferentially anti-stem responses derived from MBCs, which were able to outcompete
210 the naïve response to the HA head. Further, Ellebedy et al. (16) also found that immunosubdominance of
211 the stem reemerged after repeat immunization with the same pandemic strain.

212  To test whether the above findings could be explained with the present model, we systematically repeated
213 the preceding simulations for different numbers of distinct BCR/epitope pairs (2 to 15), different occlusion
214 values $o$=[0,0.5,0.9,1], and three different values of affinity advantage provided to BCR#1 (given below);
215 BCR#1 bound monovalently ($K_{eq}^{12}$=0) or bivalently ($K_{eq}^{12}$=10$K_{eq}^{11}$). The remaining BCR#$i$ ($2 \leq i \leq 15$)
216 were modeled as bivalent with $K_{eq}^{i2}$=10$K_{eq}^{11}$. These simulations are discussed below, and their results are
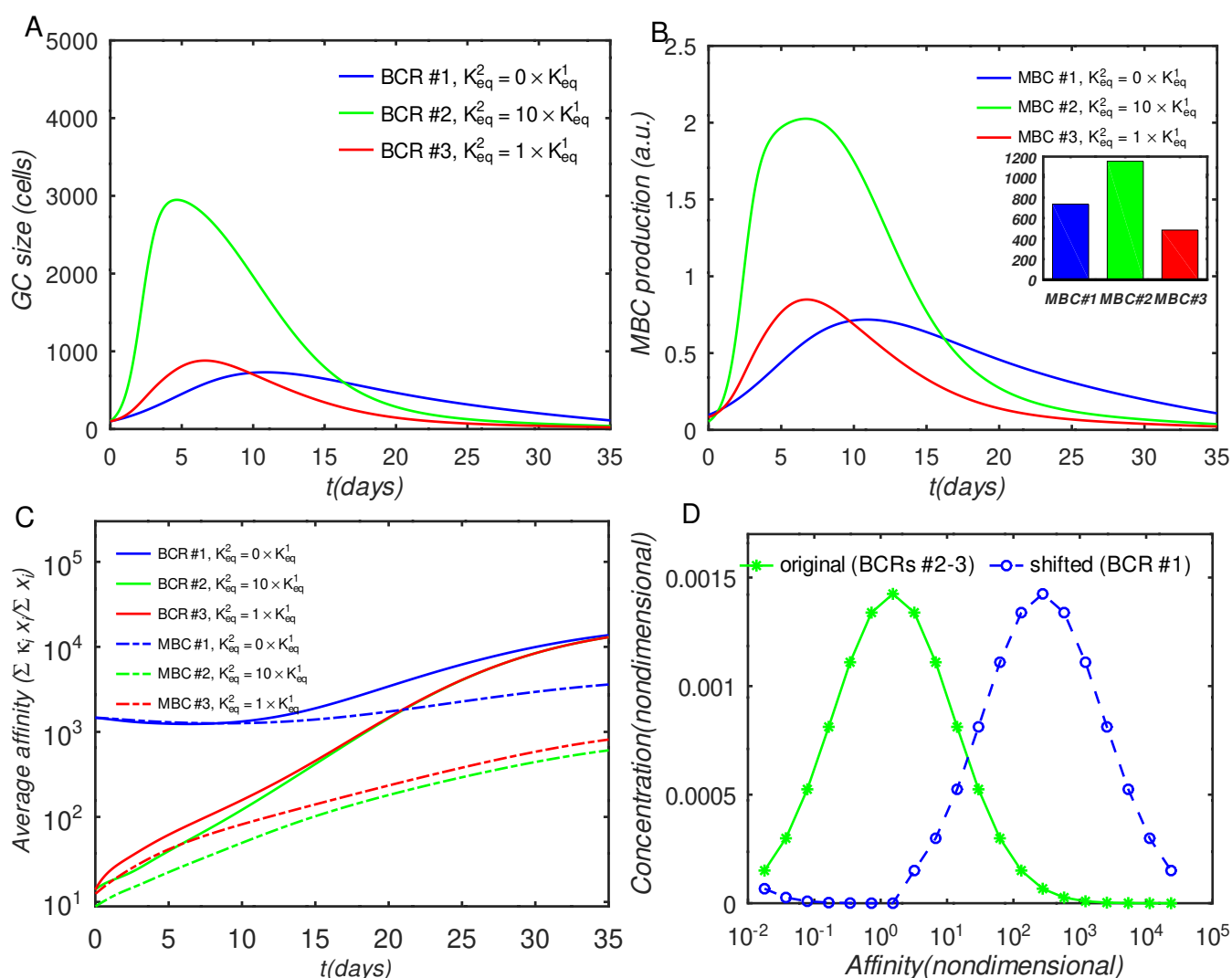217 shown in Fig. 6.

**Figure 5.** Effect of initial affinity advantage on the growth of monovalent B-cells in the fully interacting B-cell case ($o = 1$). Panels A–C show the same quantities as Fig. 4A–C; The affinity distribution corresponding to BCR#1 was shifted toward higher values relative to BCR#2 and BCR#3 (panel D).

218    The goal of the simulations is to approximate conditions in which one monovalent or bivalent anti-stem
219 BCR (#1) lineage is evolving concurrently with 1-14 bivalent anti-head BCRs. In the first vaccination,
220 all BCRs start from the same affinity distribution peaked at $K_{eq}^1 \simeq 1.53$ (Fig. 5D). To model the effect of
221 stem conservation after the first vaccination, BCR#1 is given a (multiplicative) affinity advantage over the
222 remaining BCRs of $\Delta K_{eq}$=10, 100, or 1000. After the first (prime) simulation, each boost is initialized with
223 a combination of 25% MBCs taken at the end of the previous simulation, and 75% naïve B-cells having
224 the same initial distribution as that used for the prime. We assume that the previously-generated anti-head
225 MBCs are poorly matched to the boosting Ag and shift their affinity distribution toward lower values
226 by a factor of 1000, essentially eliminating any advantage of previous maturation. For the presumptive
227 stem-directed BCR#1, we assume that the previously-generated MBCs are better matched to the boosting
228 Ag, and shift their affinity downward only by a factor of 100, 10, or 1 (unchanged), to explore the effect of
229 the mismatch; (thus, the affinity advantage $\Delta K_{eq}$ of BCR#1 corresponds to 1000 divided by 100, 10, or
230 1). We found that the proportion MBC#1 reaches a plateau by about five immunizations (see Fig. S3). In
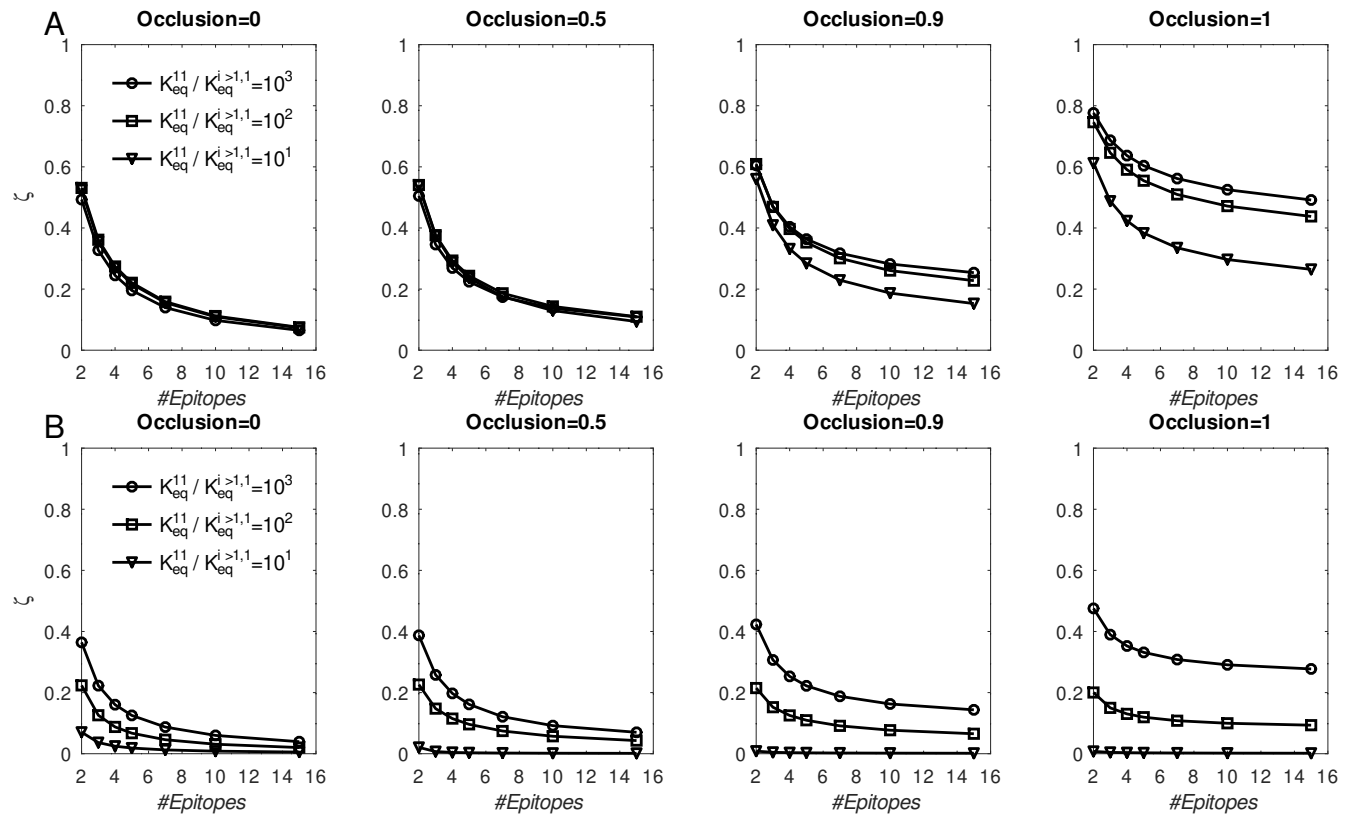231 Fig. 6 we show the fraction of MBCs#1 after the sixth simulation.

**Figure 6.** Fraction of MBC#1 ($\zeta$, defined in the text) at the end of six GC simulations for different initial affinity advantage values *vs.* total number of BCR/Epitope pairs. A: BCR#1 is cooperatively bivalent ($K_{eq}^{12}=10K_{eq}^{11}$); B: BCR#1 is monovalent ($K_{eq}^{12}=0$).

First, we discuss the results of the cooperative bivalent anti-stem case (Fig. 6A). Here, BCR#1 does not have an avidity disadvantage (since it is cooperatively bivalent with $K_{eq}^{12}=10K_{eq}^{11}$), relative to the remaining BCRs, and has an affinity advantage, as described above. In the noncompeting case ($o$=0) the MBC fraction $\zeta=MBC_1/\sum_{i=1}^{N_B} MBC_i$ is not very sensitive to the affinity advantage, because the Abs are maturing independently and the concentration of each epitope is the same. As the competition between the BCR lineages is increased, the affinity advantage becomes more important. For example, in the high occlusion cases $o \geq 0.9$, with 9 competing low-affinity BCRs, a 1 to 3 order of magnitude affinity advantage results in ~20% to ~50% of the final MBC population being derived from BCR#1 ($\zeta \in [0.2, 0.5]$ in the two right panels of Fig. 6A).

In contrast, the monovalent anti-stem response (Fig. 6B) produces markedly lower, though still significant MBC#1 proportions. For the lowest initial affinity advantage ($\times 10$), the proportion of MBC#1 is vanishingly small for all interacting cases. For the higher advantage values ($\times 100$ and $\times 1000$), the proportion of MBC#1 with 9 competing low-affinity BCRs at $o \geq 0.9$ is in the range $\sim 10\% - \sim 30\%$ ($\zeta \in [0.1, 0.3]$ in the two right panels of Fig. 6B).

We interpret these results to suggest that a previous response to a conserved epitope could be boosted to dominate the subsequent response, even in the presence of a significant number of poorly-conserved 'distracting' epitopes. This is consistent with the chimeric vaccination results of Nachbagauer et al. (39). However, the extent of boosting is critically dependent on the affinity advantage of the preexisting immunity. In the case of monovalent antibodies, the affinity advantage needs to be high to overcome the proliferation

                                                            

251  disadvantage caused by monovalency, and the entropic disadvantage caused by distracting epitopes that
252  generally outnumber conserved ones. Ellebedy et al. (16) noted that vaccination with a pandemic strain
253  against which anti-HA head immunity is low, preferentially boosted anti-HA stem antibodies. However,
254  upon reimmunization with the same antigen, the anti-HA head Abs were boosted preferentially. These
255  results can be rationalized in the present model by the differential affinity advantages of the anti-HA stem
256  Abs. Specifically, for the first immunization, the anti-stem immunity has a sufficient affinity advantage
257  to overcome the growth-related and entropic disadvantages. For the second immunization, the affinity
258  advantage is eroded because the anti-HA head immunity has undergone affinity maturation caused by the
259  first immunization, which leads to the restoration of stem epitope subdominance.

260      Because the affinity advantage of a conserved epitope cannot be predicted accurately if antigenic drift is
261  present, a more realistic approach would be to treat $\Delta K_{eq}$ as a random variable. To investigate this scenario,
262  we performed a round of simulations in which $\Delta K_{eq}$ was sampled from the lognormal distribution; the
263  results are described in SI Appendix S1. The differences between the two cases were similar to those in the
264  simulations with fixed $\Delta K_{eq}$. For example, in the bivalent case, some MBC#1 cells were always present,
265  while in some simulations of the monovalent case the MBC#1 proportion is near zero (see Fig. S4).

266      The above simulations suggest that an affinity advantage alone will not always be sufficient to overcome
267  the disadvantage of slower proliferation and entropic distraction. This result, together with other factors,
268  such as low germline precursor frequency or T-cell help insufficiency,(17) could explain the lower
269  prevalence of anti-HA stem immunity.

270      In natural infections, the epitope concentrations are predetermined by the Ag itself (*e.g.*, the solvent-
271  accessible surface area of the HA head on an influenza virion is about twice that of the stem; see Fig. S5 in
272  SI Appendix S2). However, the vaccination setting allows use of designed antigens with some epitopes
273  masked by glycosylation(10, 41, 42) or even completely removed using protein engineering(43, 44).
274  Additionally, one can administer concurrently multiple Ags, in which some Ag epitopes are sufficiently
275  conserved between the antigens that they can be considered to represent a single epitope with a higher
276  effective concentration. (24, 45, 25, 23). In the next set of simulations, we investigate the interplay of
277  epitope concentration and BCR binding valency. Figure 7 shows the results of a GC simulation involving
278  three competing BCR/Ag pairs, for comparison with the previous 3-BCR simulations (Figs. 4 and 5);
279  the noninteracting case can be found in Fig. S6. The concentration of Ag#1, which corresponds to the
280  monovalent BCR#1, was twice that of the other two Ags, while all other parameters were the same as in
281  the original 3-BCR simulation. The peak monovalent B-cell population increases several-fold relative to
282  the uniform concentration case (Fig. 4D); it is similar in magnitude to that in the affinity-advantaged case
283  of Fig. 5, but occurs later, at ~23 *vs.* ~11 days. This behavior was expected, because, although BCR#1
284  has an impaired growth rate due to monovalency, it also has more antigen available, allowing it to grow
285  for longer times. Even though the other (bivalent) BCRs are occluding, the higher concentration of Ag#1
286  overcomes the occlusion disadvantage, albeit with a slow growth rate. The resulting MBC#1 population is
287  much closer to that of the bivalent cooperative case. However, although the greater abundance of Ag#1
288  amplifies the total BCR#1 response, it also reduces the competition for this Ag, which results in a lower
289  overall affinity of the resulting B-cells (Fig. 7C).

290      For completeness, in SI Appendix S3 we repeated the above simulation while systematically varying
291  the number of BCR/epitope pairs (2-15), occlusion parameter value (0,0.5,0.9,1), with three values of
292  affinity advantage provided to the BCR#1, and with the BCR#1s bound monovalently or bivalently. The
293  results are summarized in Fig. S7. In all cases, increasing Ag concentration leads to greater MBC output,
294  with the increase being larger if the corresponding BCR also has a significant affinity advantage. The
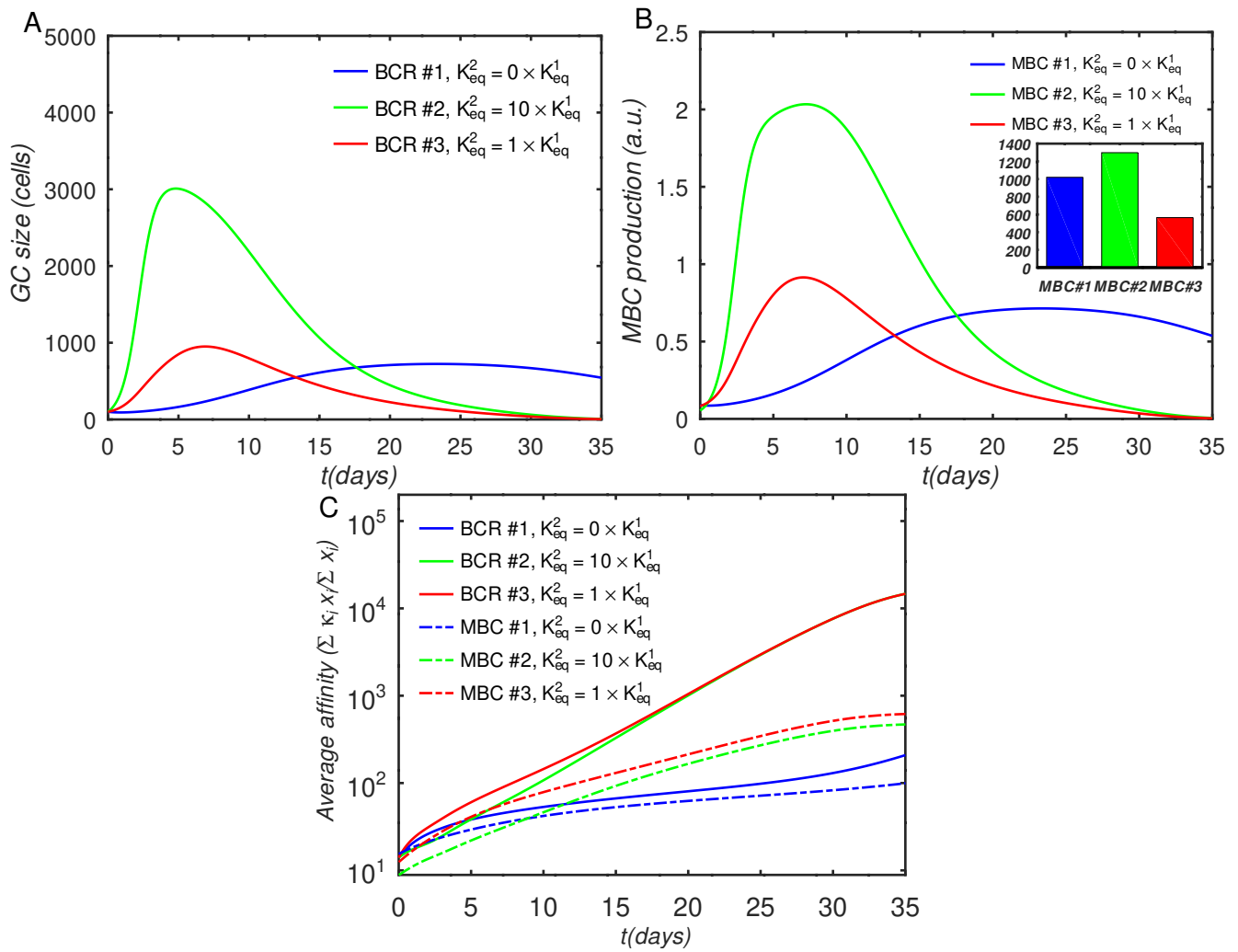
**Figure 7.** Effect of BCR valency and epitope concentration on GC evolution, with $o = 1$ (fully competitive case). Panels A–C show the same quantities as Fig. 4A–C; $\alpha_1^T = 2$, $\alpha_2^T = 1$, $\alpha_3^T = 1$; $\alpha^T$ is the nondimensional total Ag concentration (see Sec. 4.0.6).

295  increased MBC output is associated with decreased affinity, however, and the affinity decrease is larger
296  for monovalent than bivalent BCR#1s. The results therefore suggest that epitope subdominance can be
297  overcome by increasing epitope concentration in vaccinations with cocktails of designed antigens, as
298  proposed by others.(24, 25, 23) However, this is achieved at the expense of a reduction in the affinity of the
299  resulting MBCs. For vaccine design, the precise Ag concentrations may need to be optimized to achieve a
300  compromise between MBC population size and affinity for antigen.

301      Finally, to investigate whether the outcome of multiple vaccinations can be optimized by manipulating the
302  epitope concentrations corresponding to monovalent BCRs, we simulated six consecutive immunizations
303  under the same initial conditions as described before, except that the concentration of Ag#1 was increased
304  in some, but not all, of the simulations. Specifically, we considered three immunization scenarios, in which
305  the total nondimensional concentration of Ag#1, $\alpha_1^T$=[Ag#1/Ag#i>1], in the six consecutive immunizations
306  was (1, 1, 1, 1, 1, 1), (2, 1, 1, 2, 1, 1), and (2, 2, 1.5, 1.1, 1, 1). These three scenarios were chosen to
307  determine whether increased Ag#1 occurring early in a vaccination regimen would translate into superior
308  responses in later exposures. Figures 8 and S8 show the MBC output and affinity, respectively, at the end
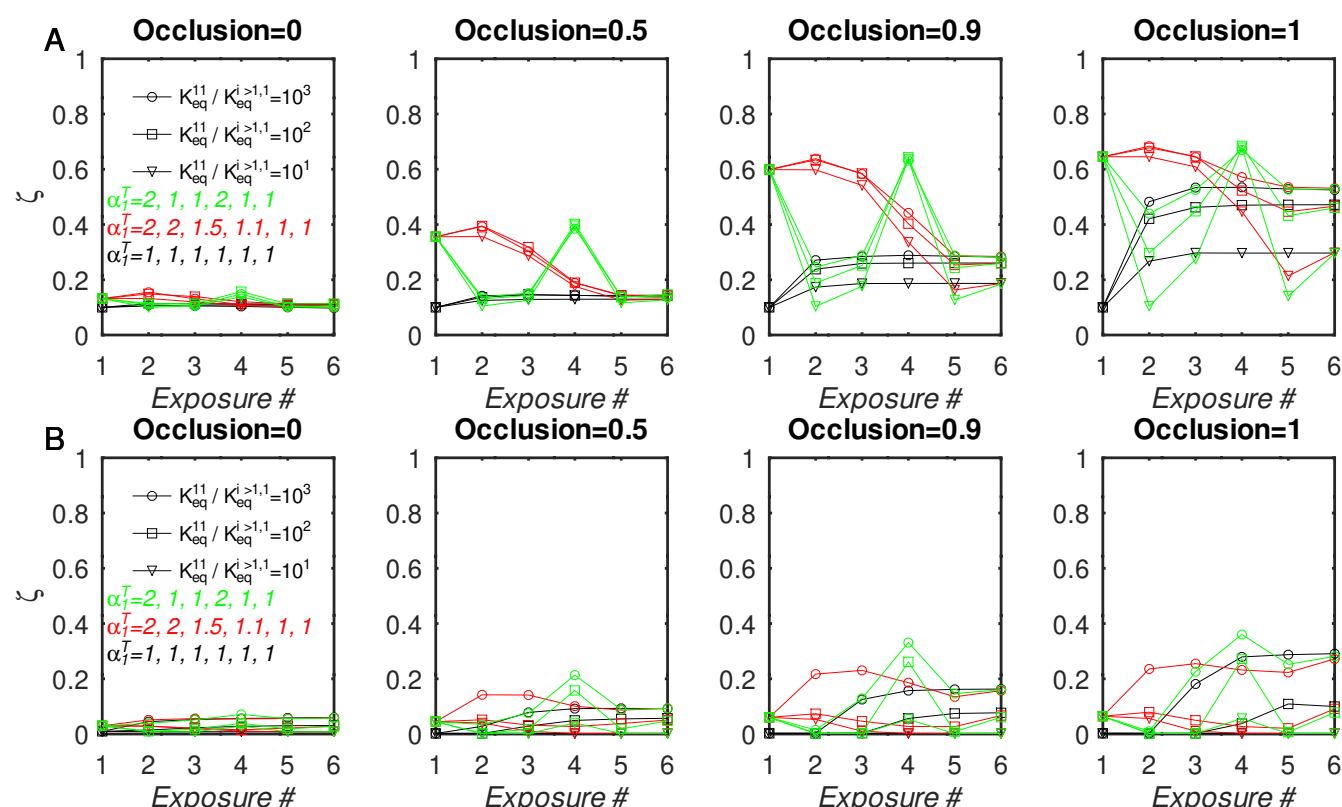
**Figure 8.** Fraction of MBC#1 *vs.* number of sequential GC simulations for different initial affinity advantage values and different Ag#1 concentrations, with 10 total BCR/Epitope pairs. A & B: MBC#1 fraction ($\zeta$) at the end of simulations. A: BCR#1 is cooperatively bivalent ($K_{eq}^{12}=10K_{eq}^{11}$); B: BCR#1 is monovalent ($K_{eq}^{12}=0$). The four sets of panels A and B show the effect of increasing occlusion from $o$=0 (no competition) to $o$=1 (full competition).

309   of each immunization for ten BCR/epitope pairs (other cases are omitted for clarity, but are qualitatively
310   similar). Consistent with the previous results, MBC#1 output after a particular immunization increases
311   if Ag#1 used in that immunization is increased (and vice versa), and the MBC#1 affinity decreases if
312   Ag#1 used in that immunization is increased (and vice versa). However, the differences between the three
313   protocols essentially disappear after the final exposure, hence the present model suggests that there may not
314   be a significant long term immunological effect of simply manipulating antigen concentration in a vaccine.

## 3 DISCUSSION

315   Rapidly mutating and proliferating viruses such as influenza, HIV, and, more recently, SARS-CoV-2,
316   accumulate escape mutations that can render existing host immunity obsolete. However, because such
317   pathogens must maintain infectivity to survive, some mutations are highly improbable, as they would
318   significantly reduce or even eliminate viral fitness. In addition, selection pressure from the immune system
319   is variable along the antigenic sequence (*e.g.* solvent-exposed regions are more susceptible to antibodies
320   than buried ones). The resulting differences in mutation propensities make it possible to partition the viral
321   topology into variable and conserved epitopes. Unfortunately, variable epitopes tend to be immunodominant,
322   *i.e.* they are the main targets of adaptive immunity. Immunodominance, in itself, may be the result of
323   viral adaptation; for example, the large highly-variable head of influenza hemagglutinins is an entropic
324   distraction to the immune system. A major focus of current vaccine research is to elicit a potent and durable

immunity to conserved, immunosubdominant, epitopes. Such vaccines would lower HIV infection rates, or eliminate the need for a yearly influenza vaccine.

Here, we employed a coarse-grained model of affinity maturation (AM) parametrized using experimental data on germinal centers (GCs)(36, 38, 37) to determine whether differences in B-cell receptor (BCR) binding valency could explain the subdominance of certain epitopes. The main assumption of the model is that B-cell activation increases with the amount of equilibrium-bound BCR to antigen (Ag). More specifically, when both arms (Fabs) of BCR bind the antigen displayed on follicular dendritic cells (FDCs), the probability of internalizing the Ag increases, even if the affinity of each receptor for Ag is weak. This assumption appears to be in accord with the experiments; Arevalo et al. (18) interpreted their vaccination boosting data by suggesting that many weak BCR/Ag interactions are sufficient to activate B-cells. Compelling indirect evidence comes from co-display of different Ags on nanoparticles, which was  demonstrated to preferentially elicit broadly-neutralizing Abs. These findings(24, 25) imply that avid bivalent binding confers a proliferation advantage, which is consistent with the present model. However, as B-cell activation by Ags is a complicated process, involving cross-linking of the BCRs, it is not clear to what extent avid binding would increase crosslinking. Future experiments and simulations may be needed to shed more light on the activation process.

The present simulations indicate that monovalent B-cells always grow more slowly than bivalent ones, and are therefore easily dominated by B cells that are able to bind bivalently and cooperatively. When given an initial affinity advantage over bivalent B-cells, as might be expected to occur upon recruitment of monovalent memory cells (MBCs) into secondary GCs, the affinity advantage was often insufficient to overcome the slower growth. The monovalent B cells outcompeted bivalent ones only if the affinity advantage was more than an order of magnitude (see Fig. 6B). These results are in agreement with influenza vaccination experiments, which show that a boost with a pandemic strain for which the host has little immunity against epitopes in the HA head, produces high anti-HA stem titers; a subsequent boost with the same vaccine elicits anti-HA head Abs.(16) We have interpreted these experimental data by assuming that anti-HA stem Abs bind monovalently with a high affinity advantage in the first vaccination, but not the second. In the second vaccination, the bivalent anti-HA head Abs are able to overcome the advantage *via* maturation induced by the first shot.

Rather than relying solely on an affinity advantage, a more robust method to boost monovalent B-cells is to increase the concentration of their cognate epitope(s). The simulations indicate that this approach results in the highest number of monovalent MBCs (see Figs. 7 and 8 in Sec. 2). The finding is not surprising, since the presence of foreign antigen is what initiates and sustains GC reactions in the first place. However, because selection among B-cells is driven by competition for Ag, an increase in available Ag will allow lower-affinity BCRs to survive. Thus, although the corresponding monovalent Abs become more numerous with increased epitope concentration, they evolve lower average affinity. When we simulated a subsequent GC reaction initialized with MBCs from such a memory pool, but *without* the concentration advantage (designated by $\alpha_1^T = 1$ in the results), as would occur in a natural infection, the monovalent B-cell population rapidly decreased, such that after two such consecutive GC reactions, there was no difference when compared to vaccinations in which the concentration of the Ag cognate to the monovalent Ab was never increased (Fig. 8).

These results suggest that increasing Ag concentration might only provide a temporary advantage to the cognate BCRs, as Ag levels are easy to manipulate in vaccination, but not in infection. Nevertheless, the strategy could prove useful to expand the number of initial low-affinity B-cell lineages targeting rare epitopes against which high-quality B-cell precursors are rare, such as group I and II influenza stem

369 epitopes. If subsequent exposures *via* natural infection restore immunosubdominance,(16) regular vaccine
370 boosting with higher concentrations of subdominant epitopes could be required.

371     We note that multi-antigen vaccination cocktails have been designed,(24, 25, 23) in which the epitopes
372 that are conserved between the antigens are at effectively higher concentration than variable epitopes.
373 Simulations performed here suggest that such cocktails have promise to elicit Abs to conserved epitopes
374 *via* a concentration advantage. Future experiments are needed to address whether the resulting immunity
375 would persist after multiple rounds of natural infections.

376     The model used here uses simple assumptions to show that immunosubdominance can be explained to a
377 large degree by differences in the antibody binding valency. However, other factors, such as low natural
378 germline precursor frequency, repertoire filtering due to self-reactivity, or reduced T-cell help, are likely to
379 contribute. These factors could be incorporated into the model in an approximate way in future studies.
380 More sophisticated approaches that explicitly model BCR evolution in sequence space and/or compute
381 BCR/Ag binding affinity using atomic models(46, 47), are likely to be required to design actual vaccine
382 antigens and their cocktails.

## 4 METHODS

### Model of Germinal Center Affinity Maturation

383

384     The model of affinity maturation (AM) used here is based on the work of Kepler and Perelson (27),
385 who used systems of coupled differential equations to calculate the concentrations of B-cells of different
386 discrete affinities for an antigen (Ag). We have generalized the model to simulate the maturation of multiple
387 B-cell lineages, each binding to its cognate antigen mono- or bivalently, and producing memory B-cells
388 and plasma cells, which secrete antibodies of the same affinity for the cognate Ag. The model does not
389 have any geometric or topological component to represent binding, and, where there is no ambiguity, we
390 sometimes use the terms antigen and epitope interchangeably. We will also sometimes use the abbreviation
391 BCR (B-cell receptor) to refer to B-cells, with the implicit scaling assumption that each B-cell has $\sim 10^5$
392 BCRs on its surface.(48, 49)

393     We assume that the simulated germinal center(s) (GCs) have been seeded by $N_B$ B-cell lineages $B^i$,
394 and that each lineage can bind only to its cognate antigen $a_i$. Though the model is based on the system of
395 differential equations of Kepler and Perelson (27), it has several important additions, notably memory cell,
396 plasma cell, and antibody production. For clarity, we first present a minimal version of the model, which is
397 close to the original B-cell model(27) and describe the modifications in subsequent subsections.

### 4.0.1 Basic model

398

399     For each B-cell lineage, we explicitly model its binding affinity distribution. Specifically, we assume that
400 (1) the equilibrium binding affinity $K_{eq}^i$ of any B-cell $B^i$ derived from the lineage $i$ for its cognate antigen
401 $a_i$ is in the range $K_{eq}^{min} \leq K_{eq}^i \leq K_{eq}^{max}$ and (2) that the affinities can be represented by a discrete set of
402 values, as follows. We assign to each $B^i$ a binding energy index $j \geq 1$ such that

$$\log K_{eq}^{min} \leq (j-1)\Delta E + \log K_{eq}^{min} \leq \log K_{eq}^i < j\Delta E + \log K_{eq}^{min} \leq \log K_{eq}^{max} + \Delta E. \qquad (1)$$

403 Equation (1) corresponds to a uniform discretization of binding affinities in logarithmic space with energy
404 grid spacing $\Delta E$, or exponential discretization in affinity space.

405      As done in Ref. 27, we will refer to the energy bins $j$ as *affinity classes*. Their number, $N_\epsilon$, is related to
406   $\Delta E$ by $\log K_{eq}^{max} - \log K_{eq}^{min} = (N_\epsilon - 1)\Delta E$. We take $N_\epsilon$=20 (see Table 1), which implicitly determines
407   $\Delta E$ once $K_{eq}^{min}$ and $K_{eq}^{max}$ are chosen. In the simulations, we only allow values of $K_{eq}$ that correspond to
408   the bin edges, *i.e.*, $K_j^i = K_{eq}^{min} e^{(j-1)\Delta E}$, $1 \leq j \leq N_\epsilon$, where, for brevity, we replaced the subscript $eq$ by
409   the affinity class index $j$.

410      To simulate affinity maturation, we compute the time evolution of B-cell populations in each affinity
411   class $j$ using

$$\frac{d[B_j^i]}{dt} = \left\{ -k_d(1 - h_j^i) - k_p \right\} [B_j^i] + 2k_p \sum_{k=1}^{N_\epsilon} m_{kj}[B_k^i], \tag{2}$$

$$\text{for } 1 \leq i \leq N_B,$$

412   where $k_p$ and $k_d$ are proliferation and death constants, respectively, $m_{kj}$ is the probability for a BCR $B^i$ in
413   affinity class $k$ to transition to affinity class $j$ *via* somatic mutations that take place during AM, $h_j^i$ is B-cell
414   activation function (discussed below), and $N_B$ is the number of different B-cell lineages.
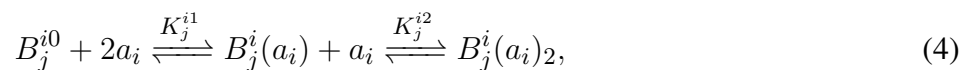
415      The class transition probabilities $m_{jk}$ are assumed to be independent of the lineage $i$, and are defined as

$$m_{jk} = \frac{[\mu(1 - p_L)]^{|k-j|}}{|k-j|!} \frac{\exp(-\mu)}{1 + \Lambda^{2(k-j)}}, \quad j \neq k$$

$$m_{jj} = 1 - \sum_{k \neq j}^{N_\epsilon} m_{jk}, \tag{3}$$

416   where $\Lambda$ determines the ratio of advantageous to non-lethal deleterious mutations, $p_L$ is the probability
417   of lethal mutations, and $\mu$ is the probability of an expressed (*i.e.* nonsilent) mutation per generation.(27)
418   Kepler and Perelson (27) used an oscillating function for $\mu(t)$ to mimic the effect of interconversion of
419   centroblasts and centrocytes on the mutation rate, which was optimized to maximize a 'total' affinity $A(t)$
420   of mature B-cells of a single lineage $i$=1 ($A(t) = \sum_j^{N_\epsilon} [B_j^1] K_j^1$). We use a constant average mutation rate
421   $\mu$=0.1, which is appropriate for comparing growth rates of different BCR lineages within the framework of
422   this coarse-grained model; otherwise one would need to specify the phases of oscillation for each lineage,
423   which are unknown, and might furthermore be stochastic. The constant value $\mu$=0.1 was also used by
424   Oprea and Perelson (28).

425      The activation function $h_j^i$ is derived from the proportion of the B cells, $B_j^i$, that receive a survival signal
426   *via* binding to antigen and/or helper T-cells (Tfh). In the simpler model, which involves only activation by
427   antigen (Ag), $h_j^i$ is computed from the equilibrium fraction of receptors $B_j^i$ bound to the cognate Ag $a_i$.
428   Because each BCR has two binding arms (Fabs), we assume the binding reactions

$$B_j^{i0} + 2a_i \underset{\longleftarrow}{\overset{K_j^{i1}}{\rightleftharpoons}} B_j^i(a_i) + a_i \underset{\longleftarrow}{\overset{K_j^{i2}}{\rightleftharpoons}} B_j^i(a_i)_2, \tag{4}$$

429   corresponding to sequential binding of the first and second Ags to free B-cells ($B_j^{i0}$). From the conservation
430   of total B cells $B_j^i$,

$$[B_j^{i0}] + [B_j^i(a_i)] + [B_j^i(a_i)_2] = [B_j^i], \tag{5}$$

431  we have

$$[B_j^i(a_i)] = \frac{K_j^{i1}[B_j^i][a_i]}{1 + K_j^{i1}[a_i](1 + K_j^{i2}[a_i])} \tag{6}$$

432  and

$$[B_j^i(a_i)_2] = \frac{K_j^{i1}K_j^{i2}[a_i]^2}{1 + K_j^{i1}[a_i](1 + K_j^{i2}[a_i])}, \tag{7}$$

433  which allows us to compute the fraction of bound BCR arms (two per BCR), provided that the concentration
434  of free Ag, $[a_i]$, is known, *i.e.*

$$\theta_j^i = \frac{[B_j^i(a_i)] + 2[B_j^i(a_i)_2]}{2[B_j^i]}. \tag{8}$$

435  In the above, $K_j^{i1}$ and $K_j^{i2}$ are equilibrium binding constants, corresponding to binding by a first BCR arm
436  and the second, respectively. $K_j^{i1}$ is equal to the affinity of the Fab, *i.e.*, $K_j^{i1}=K_j^i$, and $K_j^{i2}$ is taken to be
437  proportional to it, $K_j^{i2}=C_iK_j^i$, where the constant $C_i$ is affinity-independent, and reflects geometry-related
438  factors that influence the binding of the second arm, such as excluded volume (entropy), and deformation
439  strain (energy) required to position the second arm for binding. Monovalent binding corresponds to $C_i = 0$.
440  Other values of $C_i$ used in the simulations were 1 and 10, which are discussed in the next subsection.

441  To evaluate Eqs. (6) and (7), $[a_i]$ is needed. Writing conservation of total antigen $[a_i^T]$, which is prescribed
442  at the beginning of the GC reaction, and possibly evolves during the reaction, we have

$$[a_i] + 2\sum_{j=1}^{N_\epsilon}[B_j^i]\theta_j^i = [a_i^T]. \tag{9}$$

443  We solve Eq. (9) for $[a_i]$ iteratively using the Newton-Raphson method.(50)

444  With $\theta_j^i$ determined, we can compute the activation $h_j^i$. Kepler and Perelson (27) modeled a single lineage
445  and used $h_j^1=2\theta_j^1$ with $K_j^{12}=0$, *i.e.*, they treated all BCRs as monovalent. In the present model, we modified
446  the functional form of $h$ to match the observations of GC size evolution of Wittenbrink et al. (36), while
447  keeping most of the parameters from the KP93 model(27). Specifically, we defined the activation function
448  as

$$h_j^i = \rho(\theta_j^i)^\epsilon \tag{10}$$

449  and performed least squares optimization to improve agreement with the average GC data(36), to obtain
450  $\rho$=0.94 and $\epsilon$=0.7 (see SI Appendix S4 & Fig. S9). Because $\theta_j^i < 1$, the prefactor reduces the activation
451  upper bound to $\rho$; the exponent $\epsilon$¡1 increases activation for smaller binding fraction values, reducing the
452  competitive advantage of higher affinity BCRs.

453  Two comments on Eq. (2) are necessary. First, the proliferation is split into two terms(27) to expose the
454  fact that mutations in a cell of lineage $B_j^i$ during division will lead to the loss of the parent cell and a gain of
455  two daughter cells. Second, the proliferation is not *activated* (*i.e.* not proportional to $h_j^i$). Whether activation
456  by Ag and T-cells mainly rescues B-cells from apoptosis (death rate proportional to [1-$h$]),(51, 31) or
457  actually increases the rates of proliferation (growth rate proportional to $h$) has been a matter of some debate,
458  with more recent evidence in favor of activated proliferation.(52) In this study, however, parametrizing
459  the model with activated proliferation would not change the main conclusions; the main difference in the
460  activated proliferation model parameters was that the rate constants $k_p^{max}$ and $k_d$ had to be increased and

461 reduced, respectively, to fit experimental data (see SI Appendix S5). The fact that in the *nonactivated*
462 proliferation model the B-cell death rate constant is higher than the proliferation rate (Table 1) reflects
463 the importance of rescue from apoptosis to B-cell survival for this model. Comparison of the activated
464 proliferation model results to the experiments(36, 38) is given in Figs. S10 and S11. A study of the two
465 types of proliferation models was performed by Amitai et al. (53), whose main finding was that activated
466 proliferation reduces clonal diversity.

## 4.0.2 Avidity of simulated BCRs

468 To examine the effect of BCR avidity on the evolution of B cells within the GC, we compare the behavior
469 of lineages with three regimes of bivalency, corresponding to $K_{eq}^2 = 0$, $K_{eq}^2 = K_{eq}^1$, and $K_{eq}^2 = 10K_{eq}^1$, which we
470 denote, respectively, as the monovalent, noncooperative, and cooperative binding cases. We can rationalize
471 the chosen values as follows. The binding constant can be expressed in terms of the free energy difference
472 ($\Delta F$) between reactants and products, which is approximately decomposable into rotational, translational
473 (equivalently, concentration) and configurational components(54); Because the two binding arms (Fabs)
474 are identical, we assume that they populate identical conformational ensembles, and therefore their $\Delta F$
475 of binding can only differ in the rotational and translational entropies, and, possibly, in the energetic
476 strain needed to move the second Fab into its binding position. The translational entropy penalty of
477 binding of the second Fab (Fab2) will generally be much smaller than that of the first (Fab1), because
478 the volume accessible to unbound Fab2 is restricted by the binding of Fab1 to its epitope, whereas the
479 volume accessible to an unbound BCR is of the order of the GC volume. More specifically, we can estimate
480 the volume available to Fab2 when Fab1 is bound to be the volume occupied by an antibody, which is of
481 the order $(10nm)^3 = 10^{-24}m^3$. The effective AG concentration is inversely proportional to this volume
482 (*i.e.* we assume that one antigenic site is available to Fab2, restrained by Fab1). In contrast, when Fab1
483 is unbound, we take the AG concentration to be inversely proportional to the volume of the GC light
484 zone, approximated as 50% of the GC volume ($= 0.5 \times [80\mu m]^3 = 2.56 \times 10^{-13}m^3$) and proportional
485 to the number of individual Ags presented on FDCs. Because antigen is generally abundant in GCs,(55)
486 we take the number of Ags available to bind BCRs as 1000 times the typical B-cell count in a maturing
487 GC, which is around 2000 from Fig. 3. The ratio of antigen concentrations for bound *vs.* unbound Fab1
488 is then $10^{24} \times 2.56 \times 10^{-13}/(1000 \times 2000) \simeq 10^5$. The difference in the rotational entropy penalty due
489 to binding is expected to be much smaller, because antibodies appear to be sufficiently flexible to permit
490 considerable independent rotation of the individual Fabs.(56) In particular, we expect the difference to be
491 less than an order of magnitude, and neglect this contribution. If energetic strain (*i.e.* $\Delta E$) is needed to
492 accommodate binding of Fab2, it will reduce the binding affinity by the factor $\exp(\Delta E/[k_B T])$. In the
493 absence of experimental data, we assume a strain energy in the range 1-5 kcal/mol, which corresponds to
494 the reduction of $K_{eq}^2$ by a factor in the range $\exp(1/[k_B T]) - \exp(5/[k_B T]) \simeq 5$–4160, where $k_B T \simeq 0.6$
495 at $T = 300K$. The above crude estimates suggest that, even with a substantial antibody strain of 5kcal/mol, a
496 bivalency binding advantage of $\times 25$ would be present. For simplicity, and to include a margin of safety in
497 our results, we assume a slightly lower binding advantage factor of 10, *i.e.* $K_{eq}^2 = 10K_{eq}^1$. We also include
498 the monovalent case, $K_{eq}^2 = 0$, which can also be interpreted as requiring infinite strain energy for Fab2
499 binding, and an intermediate case $K_{eq}^2 = K_{eq}^1$, which we label noncooperative.

500 A possibility that is beyond the scope of this work is to compute strain energy from molecular dynamics
501 simulations of bivalent Ab/Ag binding. However, such simulations are expected to be difficult because of
502 the large sizes of the antibody and antigen molecules involved.

### 4.0.3 Memory cell production

The basic AM model, Eq. (2), follows the populations of B-cell lineages ($i$) with different affinities ($j$). However, we are also interested in the memory B-cell (MBC) populations produced by different lineages, since these MBCs will be activated in the host upon repeat infections. As in our earlier modeling work,(57) we assume that some of the B-cells exit the GC reaction as MBCs or plasma cells (PCs). MBCs are discussed here and PCs, in the next subsection.

Using $M_j^i$ to denote MBCs of lineage $i$ and affinity $j$, the corresponding evolution equation is

$$\frac{d[M_j^i]}{dt} = C_M h_j^i (1 - h_j^i)[B_j^i] - k_d^M [M_j^i], \tag{11}$$

where the first term corresponds to differentiation from GC B-cells, and the second, to apoptosis. The fraction $h_j^i(1 - h_j^i)$ preferentially selects B-cells of intermediate affinity, reflecting the observation that higher-affinity B-cells are more likely to recycle into the dark zone, rather than exit as MBCs.(58) The value of $C_M$ is 0.3 (discussed further later) and the death rate constant $k_d^M$ is 0.02/day(59).

### 4.0.4 Plasma cell and antibody production

Under the assumption of constant Ag concentration, Eqs. (2) converge to a steady-state solution, in which the GC is composed of highest-affinity B-cells, surviving indefinitely.(27) However, it is known that GCs shrink to 5% of their maximum size after about a month.(60) While the assumption of constant antigen used in Ref. 27 is likely to be unrealistic, Ag consumption in the GC is not the main cause of GC shrinkage. It is known that immune complexes presented by follicular dentritic (and other) cells(34) persist for very long times, which is probably necessary for immune memory maintenance.(61)

To account for Ag consumption, we follow Rundell et al. (59) and model it by exponential decay,

$$\frac{d[a_i^T]}{dt} = -k_d^a [a_i^T], \tag{12}$$

with $k_d^a$=0.011/day, which corresponds to a half life of about 63 days.

To model GC shrinkage consistently with the experimental observations of GC size,(36) we adopt the antibody feedback model of Zhang et al. (26). The essential concept is that some of the maturing B-cells differentiate into PCs, which secrete Abs. These Abs can diffuse throughout the GCs and compete with BCRs for antigen. Once the Abs are sufficiently numerous and of high affinity, the GC shrinks. Although the process of GC shrinkage is probably considerably more complicated, involving regulatory T cells and various signaling molecules, we employ the Ab feedback model here because it requires few additional variables and parameters (see below), and is able to capture GC size evolution over time, as shown here and in Ref. 26.

Since antibodies are secreted by PCs, we began with the PC evolution equation

$$\frac{d[P_j^i]}{dt} = C_P h_j^i (1 - h_j^i)[B_j^i] - k_d^P [P_j^i]. \tag{13}$$

However, our attempts at fitting this model to the PC production data of Weisel et al. (38) did not yield good agreement. The main reason is that the MBC and PC production rates have different evolution profiles in the experiments (see Fig. 3B *vs.* Fig. 3C), with the MBC rate decreasing rapidly, while the

535  PC rate remains essentially constant within the experimental uncertainty. In contrast, our models for the
536  two quantities (Eqs. (11) and (13)) are the same, except for the numerical values of the parameters. To
537  improve agreement in the PC rate, we added an additional, semiempirical, source term to Eq. (13), $k_M^P[M_j^i]$,
538  to mimic low-level differentiation of MBCs activated by immune complexes carrying antigen into late
539  B-blasts, which differentiate into long-lived PCs.(62, 59) The resulting PC evolution equation is

$$\frac{d[P_j^i]}{dt} = C_P h_j^i (1 - h_j^i)[B_j^i] + k_M^P[M_j^i] - k_d^P[P_j^i], \tag{14}$$

540  which maintains some PC production even if the B-cell (but not MBC) population vanishes.

541  In Eq. (14) the value of the death rate constant $k_d^P$ is 0.0336/day(59), the differentiation constant $C_P$ is 0.7,
542  and the production constant $k_M^P$ is 0.18/day. $C_P$ and $k_M^P$ were first set by trial and error and subsequently
543  refined by least squares fitting to reproduce the average GC dynamics. While we could not obtain a
544  biologically-motivated value for $k_M^P$, we note that it is about an order of magnitude lower than the B-cell
545  proliferation constant (see Table 1), consistent with its role as a lesser source of PCs. However, its value
546  still appears to be unphysically high, especially in comparison to the MBC death rate of 0.02/d (Eq. (11)),
547  possibly reflecting deficiencies or oversimplifications in the differentiation components of the model. For
548  example, the probabilities of a B-cell exiting the GC to differentiate into an MBC *vs.* a PC are kept constant
549  ($C_M$=0.3 *vs.* $C_P$=0.7). Recent data suggests that PCs tend to be produced more frequently in later stages of
550  the GC reaction,(38) and that a PC is more likely to result than an MBC if the B-cell has higher affinity for
551  antigen, and/or receives more T-cell help.(58) However, because quantitative data describing the relative
552  MBC/PC output is scarce and imprecise, we do not implement an affinity dependence in the MBC/PC
553  differentiation choice, and instead use the same preference for B-cells of intermediate affinity in both cases
554  *via* the factor $h_j^i(1 - h_j^i)$ in Eqs. (11) and (14). More sophisticated affinity-based cell fate decisions are
555  probably needed in these equations. They can be modeled using other functions of $h_j^i$, or by introducing
556  other biological species or signaling molecules, as more precise data become available.

557  Antibodies are secreted by PCs and and their removal is modeled with exponential decay

$$\frac{d[A_j^i]}{dt} = k_p^A[P_j^i] - k_d^A[A_j^i]. \tag{15}$$

558  The Ab death rate constant $k_d^A$ is obtained from the half-life of 10d,(26) and an approximate secretion $k_p^A$
559  rate constant is obtained as follows. We assume that every PC secretes $1.7 \times 10^8$ Ab molecules/day(63).
560  However, the Abs are allowed to diffuse freely in and out of GCs, and, assuming that the diffusion
561  is fast enough to establish equilibrium, we scale this rate by the ratio of internal to external volumes
562  corresponding to a single GC. The internal volume is taken as the volume of a sphere of radius $80\mu m$,(36)
563  and the external volume is taken to be $0.04 mL$(26), which gives $k_p^A$=8500/day. Starting from this value, and
564  the PC differentiation parameter $C_P$=0.5, we used least squares fitting to improve the agreement between
565  the model and the average GC sizes of Wittenbrink et al. (36) (an example of parameter fitting is shown in
566  SI Appendix S4). The optimized values were $k_p^A$=35000/day and $C_P$=0.7, corresponding to higher values
567  of Ab production needed to achieve faster GC shrinkage.

568    The effect of Ab competition is incorporated by modifying the Ag conservation Eq. (9) to include binding
569  to Abs,

$$[a_i] + 2\sum_{j=1}^{N_\epsilon}([A_j^i/\sigma] + [B_j^i])\theta_j^i =$$

$$[a_i] + \sum_{j=1}^{N_\epsilon}[C_j^i]\theta_j^i = [a_i^T],$$
(16)

570  where we assumed that Abs bind antigen in the same manner as do BCRs, and defined a total receptor
571  concentration $[C_j^i] \equiv 2([A_j^i/\sigma] + [B_j^i])$; the scaling factor $\sigma = 10^5$ appears because each B-cell ($B_j^i$) is
572  assumed to have $10^5$ BCRs,(48) which have the same binding valency as Abs.

573    At this stage, the model, as written in Eq. (2), does not have explicit limitations on the maximum GC size.
574  The shape of the B-cell population curve is governed entirely by proliferation, death and competition with
575  Abs. To obtain a close match to the peak in the experimental B-cell count,(36) we follow others (28, 59)
576  and introduce a maximum lineage size $B_{max} = 5000$ cells. The proliferation rate is modified as follows,

$$k_p = k_p^{max} \times \left(1 - \frac{\sum_j[B_j]}{B_{max}}\right),$$
(17)

577  where $k_p^{max}$ is the maximum proliferation rate. A similar idea was used by Amitai et al. (21), who increased
578  the cell death rate, as a critical B-cell population was approached.

### 4.0.5  Clonal competition via epitope occlusion

580    Thus far, we have described a model which has competition only within each clonal lineage, *i.e.*, higher-
581  affinity cells outcompete lower-affinity cells, and are themselves eventually outcompeted by growing
582  numbers of high-affinity Abs derived from them. However, GCs are seeded by multiple lineages, and it is
583  therefore important to consider the effects of interclonal competition. In the context of influenza, it would
584  be of interest to model how anti-HA-head Abs could directly compete with anti-HA-stem Abs.

585    Toward this end, we generalize the model by introducing a distinction between epitopes and antigens.
586  Specifically, we recognize that a single antigen can present different epitopes. For example, an entire viral
587  spike may be considered an antigen with many different epitopes, each targeted by a different activated
588  B cell. We postulate that the binding of a BCR or Ab of type $k$ to its cognate epitope $a_k$ can reduce the
589  accessibility of epitope $a_i$, so that the effective concentration of $a_i$ available to bind is reduced by a fraction
590  of the bound concentration of $a_k$. We label this reduction $\Delta^k[a_i]$, which is

$$\Delta^k[a_i] = -O_{ik}\sum_{j=1}^{N_\epsilon}[C_j^k]\theta_j^k, \quad \text{with } 0 \le O_{ik} \le 1, \ i \ne j,$$
(18)

591  where we introduced the *occlusion* tensor $O_{ik}$, which models the effect of epitope $k$ occupancy on epitope
592  $i$. In particular, $O_{ik}$=0 corresponds to the absence of interaction, and $O_{ik}$=1 implies that binding of $a_k$
593  completely prevents binding to $a_i$ (full occlusion). Setting $O_{ii} \equiv 1$, we write the modified Ag conservation
594  equations as

$$[a_i] + \sum_{j=1}^{N_\epsilon}\sum_{k=1}^{N_B} O_{ik}[C_j^k]\theta_j^k = [a_i^T],$$
(19)

595   in which the concentrations $[a_i]$ are now coupled *via* the occlusion tensor (unlike in Eq. (16), in which they
596   are independent). The components of $O$ can in principle be set independently, provided that care is taken to
597   avoid component values so large that negative concentrations could result. For simplicity, we begin with a
598   constant occlusion independent of the antigen identity, *i.e.*, $O_{ik} = o$, for $0 \le o \le 1$, and $i \ne k$, and reduce
599   its value to prevent negative AG concentrations. Specifically, if $[a_k^T] > [a_i^T]$ we set

$$O_{ik} = o \times \min \left\{ 1, \frac{[a_i^T]}{[a_k^T]} \right\}. \tag{20}$$

600   In the simplified case of two BCR/epitope pairs, Eq. (20) can be justified as follows. We combine the
601   conservation equations,

$$[a_1] + \sum_{j=1}^{N_\epsilon}[C_j^1]\theta_j^1 + O_{12}[C_j^2]\theta_j^2 = [a_1^T],$$

$$[a_2] + \sum_{j=1}^{N_\epsilon}[C_j^2]\theta_j^2 + O_{21}[C_j^1]\theta_j^1 = [a_2^T] \tag{21}$$

602   to obtain

$$[a_1] + (1 - O_{12}O_{21}) \sum_{j=1}^{N_\epsilon}[C_j^1]\theta_j^1 = [a_1^T] - O_{12}([a_2^T] - [a_2]). \tag{22}$$

603   Because $1 - O_{12}O_{21} \ge 0$ and the bound fraction $\theta^1$ is zero only if $[a_1] = 0$ (we assume that the binding
604   constants are not both zero), to ensure $[a_1] \ge 0$, it is sufficient to require

$$[a_1^T] - O_{12}([a_2^T] - [a_2]) \ge 0, \tag{23}$$

605   or

$$O_{12} \le \frac{[a_1^T]}{[a_2^T] - [a_2]}. \tag{24}$$

606   Eq. (20) satisfies this condition, since we also assume that $[a_2]$ is nonnegative.

### 4.0.6   Integration of model equations

608       In this section we describe the numerical procedures used to compute the time-dependent concentrations
609   of the Ags, cells, and Abs in the GC. First, following Kepler and Perelson (27), we make all concentrations
610   nondimensional using the total concentration of one of the antigens at the beginning of simulation.
611   For single-epitope (validation) simulations, we use the sole epitope. For multi-epitope simulations, we
612   aribitrarily elected to use the second epitope, to be able to vary the concentrations of the first epitope
613   for studying the effects of epitope concentration in maturation. Thus, the nondimensional variables are
614   $\alpha_i \equiv [a_i]/[a_2^T]$, $x_j^i \equiv [B_j^i]/[a_2^T]$, $y_j^i \equiv [A_j^i]/[a_2^T]$, $p_j^i \equiv [P_j^i]/[a_2^T]$, $w_j^i \equiv [M_j^i]/[a_2^T]$, and $\kappa_j^{ik} = K_j^{ik}[a_2^T]$.

In the new variables, the nondimensional evolution equations are

$$\frac{dx_j^i}{dt} = \left\{ -k_d(1 - h_j^i) - k_p \right\} x_j^i + 2k_p \sum_{k=1}^{N_\epsilon} m_{kj} x_k^i, \tag{25}$$

$$\frac{dw_j^i}{dt} = C_M h_j^i (1 - h_j^i) x_j^i - k_d^M w_j^i, \tag{26}$$

$$\frac{dp_j^i}{dt} = C_P h_j^i (1 - h_j^i) x_j^i + k_M^P w_j^i - k_d^P x_j^i, \tag{27}$$

$$\frac{dy_j^i}{dt} = k_p^A p_j^i - k_d^A y_j^i, \tag{28}$$

$$\frac{d\alpha_i^T}{dt} = -k_d^a \alpha_i^T, \quad \text{for} \quad 1 \le i \le N_B, \ 1 \le j \le N_\epsilon. \tag{29}$$

and the nondimensional AG conservation equations are

$$\alpha_i + 2 \sum_{j=1}^{N_\epsilon} \sum_{k=1}^{N_B} O_{ik}(x_j^k + \sigma^{-1} y_j^k)\theta_j^k = \alpha_i^T, \tag{30}$$

$$\theta_j^i = \frac{1}{2} \cdot \frac{\kappa_j^{i1} x_j^i (1 + 2\kappa_j^{i2} x_j^i)}{1 + \kappa_j^{i1} x_j^i (1 + \kappa_j^{i2} x_j^i)}, \tag{31}$$

$$O_{ik} = o \times \min \left\{ 1, \frac{\alpha_i^T}{\alpha_k^T} \right\}, \tag{32}$$

615   and the expressions for $m_{ij}$ and $h$ are the unchanged.

616   We follow Kepler and Perelson (27) and take the dimensionless affinities $\kappa^{min} = 7.5^{-2}$, $\kappa^{max} = 7.5^5$,
617   but use $N_\epsilon$=20 affinity bins, compared to their 8, for a finer discretization. This corresponds to $\Delta E \simeq 0.74$
618   in Eq. (1). Eq. (25) – Eq. (29) were integrated in Octave(64) using the explicit Euler method,(65) with the
619   time step $dt = 0.01 \times$ day. At each iteration, Eq. (30) was solved for $\alpha_i$ using the iterative Newton-Raphson
620   (NR) method(50). The initial guess at the current iteration was taken as the corresponding value in the
621   previous iteration, which ensured that convergence required only a few NR iterations. At the beginning
622   of the simulation, the initial guesses were $\alpha_i = \alpha_i^T$. Simulation duration was 35 days, requiring less
623   than a minute of computer time using a single computing node with an Intel Xeon E5 2.3 GHz Haswell
624   CPU. However, the simulation cost was approximately linearly dependent on the number of BCR/Ag pairs
625   simulated. The simulation parameters and their values are listed in Table 1.

### 626   4.0.7   Initial conditions

627   Initial values correspond to time $t$=0. For single-epitope simulations $\alpha_1^T$=1. For multi-epitope simulations,
628   $\alpha_2^T$=1, by definition of the normalization and $\alpha_i^T$=1 for $i > 2$. The concentration of the first epitope
629   $\alpha_1^T$ was varied between 1 and 2, depending on simulation to investigate the effects of AG concentration
630   (see Results). The initial population of each B-cell lineage was 100 cells, because we assumed that a
631   mutation-free expansion of each seeding B-cell has already taken place prior to simulation. The initial
632   distribution of naive B-cell binding constants was log-normal; specifically, we used a Gaussian in the
633   energy space, centered on the 7th class ($\kappa \simeq 1.53$) with standard deviation $\sigma$=0.6 × Λ. This choice was

**Table 1.** Model and Simulation Parameters. Time is measured in days.
[†] Used only for computing cell counts *a posteriori*, *i.e.* does not impact simulations;
[‡] The affinity bounds apply only to the first binding constant $\kappa_j^{i1}$; the second binding constant was defined as different multiples of the first to investigate avidity effects (see Sec. 2);
[*] The ratio of advantageous to deleterious but nonlethal mutations is given by $1/(1 + \Lambda^2)$; It is computed by logarithmic scaling of the value $\Lambda^0$=30 of Kepler and Perelson (27) as $\log \Lambda = \Delta E \log \Lambda^0 / \log 7.5$, which accounts for the difference that Kepler and Perelson (27) used an 8-point disretization, or 8 affinity classes (corresponding to $\Delta E = \log 7.5$), while we use 20 classes.

| Parameter | Description | Value | Source |
|---|---|---|---|
| $k_p^{max}$ | maximum B-cell proliferation rate | 2.2 | adjusted from 4(Ref. 27) to fit data(36) |
| $B_{max}$ | Maximum allowed B-cell count | 5000 | fit to data(36) |
| $k_d$ | B-cell death rate | 4.125 | adjusted from 4(27) to fit data(36) |
| $\mu$ | BCR mutation rate | 0.1 | Oprea and Perelson (28) |
| $\sigma$ | BCRs per B-cell | $10^5$ | Casten and Pierce (48) |
| $p_L$ | Fraction of lethal mutations | 0.5 | Kepler and Perelson (27) |
| $C_M$ | MBC differentiation constant | 0.3 | fit to data(36, 38) |
| $C_P$ | PC differentiation constant | 0.7 | fit to data(36, 38) |
| $k_d^M$ | MBC death rate | 0.02 | Rundell et al. (59) |
| $k_d^a$ | Decay of AG presented to BCRs | 0.111 | Rundell et al. (59) |
| $k_d^P$ | PC death rate | 0.0336 | Rundell et al. (59) |
| $k_M^P$ | PC production rate from MBCs | 0.17 | fit to data(38) |
| $k_p^A$ | AB secretion from PCs | 35000 | based on Refs 63, 26 (see text) |
| $k_d^A$ | AB death rate | 0.069 | Zhang et al. (26) |
| $k_d^a$ | AG removal rate | 0.011 | Rundell et al. (59) |
| $\rho$ | activation prefactor in $h = \rho\theta^\epsilon$ | 0.94 | fit to data(36, 38) |
| $\epsilon$ | activation exponent in $h = \rho\theta^\epsilon$ | 0.7 | fit to data(36, 38) |
| $\xi$ | Concentration corresponding to one cell[†] | 1e-4 | Kepler and Perelson (27) |
| $\kappa^{min}$ | Smallest (nondimensional) affinity[‡] | $7.5^{-2}$ | Kepler and Perelson (27) |
| $\kappa^{max}$ | Largest (nondimensional) affinity[‡] | $7.5^5$ | Kepler and Perelson (27) |
| $\Lambda$ | Mutation distribution constant[*] | $\simeq 3.5$ | based on Ref. (27) |
| $N_\epsilon$ | Number of affinity classes | 20 | up from 8(Ref. 27) for higher resolution |

634 made to approximate by a smooth function the discrete Dirac mass of $\kappa$=7.5, used by Kepler and Perelson
635 (27). The distribution is shown in Fig. 6D of the Results. The above distribution appears to us to be more
636 physical than the sharply peaked distribution of Kepler and Perelson (27). However, we note that we
637 could also obtain a satisfactory fit to the experimental data using the latter distribution, albeit with small
638 parameter adjustments that affect the total GC size and time to GC peak.

639     For simulations whose purpose was to investigate the effect of an initial affinity advantage on the rate of
640 B-cell proliferation (see Sec. 2), the initial distribution was shifted by scaling the abscissa by a prescribed
641 advantage ($\Delta K_{eq}$), and linearly reinterpolated onto the initial grid in energy space (see Results and
642 Fig. 6D). All initial distributions (of each lineage) were normalized to 100 B-cells.

## CONFLICT OF INTEREST STATEMENT

643 The authors declare that the research was conducted in the absence of any commercial or financial
644 relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

## FUNDING

## ACKNOWLEDGMENTS

## SUPPLEMENTAL DATA

Supplementary Material accompanies this article, including text appendices S1–S7 and figures S1–S13. The computer code to reproduce the simulation data and figures in this paper can be found online at `https://github.com/ovchinnv/affinity-maturation-paper2021.git`

## REFERENCES

1 .Taubenberger J, Kash J, Morens D. the 1918 Influenza Pandemic: 100 Years of Questions Answered and Unanswered. *Sci. Trans. Med.* **11** (2019) 0. doi:10.1126/scitranslmed.aau5485.

2 .Andersen K, Rambaut A, Lipkin W, Holmes E, Garry R. the Proximal Origin of SARS-CoV-2. *Nat. Med.* **26** (2020) 450–452. doi:10.1038/s41591-020-0820-9.

3 .Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, et al. The Influenza Virus Resource at the National Center for Biotechnology Information. *J. Virol.* **82** (2008) 596–601. doi:10.1128/JVI.02005-07.

4 .MATLAB. *Version 7.10.0 (R2010a)* (Natick, Massachusetts: The MathWorks Inc.) (2010).

5 .Xu R, Ekiert D, Krause J, Hai R, Crowe J, Wilson I. Structural Basis of Preexisting Immunity to the 2009 H1N1 Pandemic Influenza Virus. *Science* **328** (2010) 357–60. doi:10.1126/science.1186430.

6 .Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **14** (1996) 33–38.

7 .Nachbagauer R, Choi A, Hirsh A, Margine I, Iida S, Barrera A, et al. Defining the Antibody Cross-Reactome Directed Against the Influenza Virus Surface Glycoproteins. *Nat. Immunol.* **18** (2017) 464–473. doi:10.1038/ni.3684.

8 .Corti D, Cameroni E, Guarino B, Kallewaard N, Zhu Q, Lanzavecchia A. Tackling Influenza with Broadly Neutralizing Antibodies. *Curr. Opin. Virol.* **24** (2017) 60–69. doi:10.1016/j.coviro.2017.03.002.

9 .Watanabe A, McCarthy K, Kuraoka M, Schmidt A, Adachi Y, Onodera T, et al. Antibodies to a Conserved Influenza Head Interface Epitope Protect by an IgG Subtype-Dependent Mechanism. *Cell* **177** (2019) 1124–1135.e16. doi:10.1016/j.cell.2019.03.048.

10 .Bajic G, Maron M, Adachi Y, Onodera T, McCarthy K, McGee C, et al. Influenza Antigen Engineering Focuses Immune Responses to a Subdominant but Broadly Protective Viral Epitope. *Cell host & microbe* **25** (2019) 827–835.e6. doi:10.1016/j.chom.2019.04.003.

11 .Raymond D, Bajic G, Ferdman J, Suphaphiphat P, Settembre E, Moody M, et al. Conserved Epitope on Influenza-Virus Hemagglutinin Head Defined by a Vaccine-Induced Antibody. *Proc. Natl. Acad. Sci. USA* **115** (2018) 168–173. doi:10.1073/pnas.1715471115.

12 .Bajic G, van der Poel C, Kuraoka M, Schmidt A, Carroll M, Kelsoe G, et al. Autoreactivity Profiles of Influenza Hemagglutinin Broadly Neutralizing antibodies. *Sci. Rep.* **9** (2019) 3492. doi:10.1038/s41598-019-40175-8.

13 .Kallewaard N, Corti D, Collins P, Neu U, McAuliffe J, Benjamin E, et al. Structure and Function Analysis of an Antibody Recognizing All Influenza a Subtypes. *Cell* **166** (2016) 596–608. doi:10.1016/j.cell.2016.05.073.

14 .Corti D, Voss J, Gamblin S, Codoni G, Macagno A, Jarrossay D, et al. a Neutralizing Antibody Selected from Plasma Cells That Binds to Group 1 and Group 2 Influenza a Hemagglutinins. *Science (New York, N.Y.)* **333** (2011) 850–6. doi:10.1126/science.1205669.

15 .Li G, Chiu C, Wrammert J, McCausland M, Andrews S, Zheng N, et al. Pandemic H1N1 Influenza Vaccine Induces a Recall Response in Humans That Favors Broadly Cross-Reactive Memory B Cells. *Proc. Natl. Acad. Sci. USA* **109** (2012) 9047–52. doi:10.1073/pnas.1118979109.

16 .Ellebedy A, Krammer F, Li G, Miller M, Chiu C, Wrammert J, et al. Induction of Broadly Cross-Reactive Antibody Responses to the Influenza Ha Stem Region Following H5N1 Vaccination in Humans. *Proc. Natl. Acad. Sci. USA* **111** (2014) 13133–8. doi:10.1073/pnas.1414070111.

17 .Tan H, Jegaskanda S, Juno J, Esterbauer R, Wong J, Kelly H, et al. Subdominance and Poor Intrinsic Immunogenicity Limit Humoral Immunity Targeting Influenza Ha Stem. *The J. Clin. Invest.* **129** (2019) 850–862. doi:10.1172/JCI123366.

18 .Arevalo C, Le Sage V, Bolton M, Eilola T, Jones J, Kormuth K, et al. Original Antigenic Sin Priming of Influenza Virus Hemagglutinin Stalk Antibodies. *Proc. Natl. Acad. Sci. USA* **117** (2020) 17221–17227. doi:10.1073/pnas.1920321117.

19 .Nachbagauer R, Choi A, Izikson R, Cox M, Palese P, Krammer F. Age Dependence and Isotype Specificity of Influenza Virus Hemagglutinin Stalk-Reactive Antibodies in Humans. *mBio* **7** (2016) e01996–15. doi:10.1128/mBio.01996-15.

20 .Andrews S, Huang Y, Kaur K, Popova L, Ho I, Pauli N, et al. Immune History Profoundly Affects Broadly Protective B Cell Responses to Influenza. *Sci. Trans. Med.* **7** (2015) 316ra192. doi:10.1126/scitranslmed.aad0522.

21 .Amitai A, Sangesland M, Barnes R, Rohrer D, Lonberg N, Lingwood D, et al. Defining and Manipulating B Cell Immunodominance Hierarchies to Elicit Broadly Neutralizing Antibody Responses Against Influenza Virus. *Cell Syst.* **11** (2020) 573–588.e9. doi:10.1016/j.cels.2020.09.005.

22 .Wang S. Optimal Sequential Immunization Can Focus Antibody Responses Against Diversity Loss and Distraction. *PLoS Comput. Biol.* **13** (2017) 1–27. doi:10.1371/journal.pcbi.1005336.

23 .Ives S, Bürckert JP, Pettus C, Peñate K, Hovde R, Bayless N, et al. a General Solution to Broad-Spectrum Vaccine Design for Rapidly Mutating Viruses. *submitted* (2020). doi:10.21203/rs.3.rs-100459/v1.

24 .Kanekiyo M, Joyce M, Gillespie R, Gallagher J, Andrews S, Yassine H, et al. Mosaic Nanoparticle Display of Diverse Influenza Virus Hemagglutinins Elicits Broad B Cell Responses. *Nat. Immunol.* **20** (2019) 362–372. doi:10.1038/s41590-018-0305-x.

25 .Cohen A, Gnanapragasam P, Lee Y, Hoffman P, Ou S, Kakutani L, et al. Mosaic Nanoparticles Elicit Cross-Reactive Immune Responses to Zoonotic Coronaviruses in Mice. *Science (New York, N.Y.)* **371** (2021) 735–741. doi:10.1126/science.abf6840.

26 .Zhang Y, Meyer-Hermann M, George L, Figge M, Khan M, Goodall M, et al. Germinal Center B Cells Govern Their Own Fate Via Antibody Feedback. *J. Exp. Med.* **210** (2013) 457–64. doi:10.1084/jem.20120150.

27 .Kepler T, Perelson A. Somatic Hypermutation in B Cells: An Optimal Control Treatment. *J. Theor. Biol.* **164** (1993) 37–64. doi:10.1006/jtbi.1993.1139.

28 .Oprea M, Perelson AS. Somatic Mutation Leads to Efficient Affinity Maturation When Centrocytes Recycle Back to Centroblasts. *J. Immunol.* **158** (1997) 5155–5162.

29 .Meyer-Hermann M. a Mathematical Model for the Germinal Center Morphology and Affinity Maturation. *J. Theor. Biol.* **216** (2002) 273–300. doi:10.1006/jtbi.2002.2550.

30 .Meyer-Hermann M, Mohr E, Pelletier N, Zhang Y, Victora G, Toellner K. a Theory of Germinal Center B Cell Selection, division, and exit. *Cell Rep.* **2** (2012) 162–74. doi:10.1016/j.celrep.2012.05.010.

31 .Zhang J, Shakhnovich E. Optimality of Mutation and Selection in Germinal Centers. *PLoS Comput. Biol.* **6** (2010) e1000800. doi:10.1371/journal.pcbi.1000800.

32 .Wang S, Mata-Fink J, Kriegsman B, Hanson M, Irvine D, Eisen H, et al. Manipulating the Selection Forces During Affinity Maturation to Generate Cross-Reactive HIV Antibodies. *Cell* **160** (2015) 785–797. doi:10.1016/j.cell.2015.01.027.

33 .Robert P, Marschall A, Meyer-Hermann M. Induction of Broadly Neutralizing Antibodies in Germinal Centre Simulations. *Curr. Opin. Biotech.* **51** (2018) 137–145. doi:10.1016/j.copbio.2018.01.006.

34 .Allen C, Okada T, Cyster J. Germinal-Center Organization and Cellular Dynamics. *Immunity* **27** (2007) 190–202. doi:10.1016/j.immuni.2007.07.009.

35 .Mayer A, Zhang Y, Perelson A, Wingreen N. Regulation of T Cell Expansion by Antigen Presentation Dynamics. *Proc. Natl. Acad. Sci. USA* **116** (2019) 5914–5919. doi:10.1073/pnas.1812800116.

36 .Wittenbrink N, Klein A, Weiser A, Schuchhardt J, Or-Guil M. Is There a Typical Germinal Center? a Large-Scale Immunohistological Study on the Cellular Composition of Germinal Centers During the Hapten-Carrier-Driven Primary Immune Response in Mice. *J. Immunol. (Baltimore, Md. : 1950)* **187** (2011) 6185–96. doi:10.4049/jimmunol.1101440.

37 .Pélissier A, Akrout Y, Jahn K, Kuipers J, Klein U, Beerenwinkel N, et al. Computational Model Reveals a Stochastic Mechanism Behind Germinal Center Clonal Bursts. *Cells* **9** (2020) 0. doi:10.3390/cells9061448.

38 .Weisel F, Zuccarino-Catania G, Chikina M, Shlomchik M. a Temporal Switch in the Germinal Center Determines Differential Output of Memory B and Plasma Cells. *Immunity* **44** (2016) 116–130. doi:10.1016/j.immuni.2015.12.004.

39 .Nachbagauer R, Liu W, Choi A, Wohlbold T, Atlas T, Rajendran M, et al. a Universal Influenza Virus Vaccine Candidate Confers Protection Against Pandemic H1N1 Infection in Preclinical Ferret Studies. *NPJ vaccines* **2** (2017) 26. doi:10.1038/s41541-017-0026-4.

40 .Krammer F, Pica N, Hai R, Tan G, Palese P. Hemagglutinin Stalk-Reactive Antibodies Are Boosted Following Sequential Infection with Seasonal and Pandemic H1N1 Influenza Virus in Mice. *J. Virol.* **86** (2012) 10302–7. doi:10.1128/JVI.01336-12.

763 **41** .Eggink D, Goff P, Palese P. Guiding the Immune Response Against Influenza Virus Hemagglutinin
764    Toward the Conserved Stalk Domain by Hyperglycosylation of the Globular Head Domain. *J. Virol.* **88**
765    (2014) 699–704. doi:10.1128/JVI.02608-13.

766 **42** .Zhou T, Doria-Rose N, Cheng C, Stewart-Jones G, Chuang G, Chambers M, et al. Quantification
767    of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation. *Cell Rep.* **19** (2017) 719–732.
768    doi:10.1016/j.celrep.2017.04.013.

769 **43** .Yassine H, Boyington J, McTamney P, Wei C, Kanekiyo M, Kong W, et al. Hemagglutinin-Stem
770    Nanoparticles Generate Heterosubtypic Influenza Protection. *Nat. Med.* **21** (2015) 1065–70. doi:10.
771    1038/nm.3927.

772 **44** .Corbett K, Moin S, Yassine H, Cagigi A, Kanekiyo M, Boyoglu-Barnum S, et al. Design of
773    Nanoparticulate Group 2 Influenza Virus Hemagglutinin Stem Antigens That Activate Unmutated
774    Ancestor B Cell Receptors of Broadly Neutralizing Antibody Lineages. *mBio* **10** (2019) 0.
775    doi:10.1128/mBio.02810-18.

776 **45** .Boyoglu-Barnum S, Hutchinson G, Boyington J, Moin S, Gillespie R, Tsybovsky Y, et al. Glycan
777    Repositioning of Influenza Hemagglutinin Stem Facilitates the Elicitation of Protective Cross-Group
778    Antibody Responses. *Nature communications* **11** (2020) 791. doi:10.1038/s41467-020-14579-4.

779 **46** .Conti S, Kaczorowski K, Song G, Porter K, Andrabi R, Burton D, et al. Design of Immunogens to
780    Elicit Broadly Neutralizing Antibodies Against HIV Targeting the CD4 Binding Site. *Proc. Natl. Acad.*
781    *Sci. USA* **118** (2021) 0. doi:10.1073/pnas.2018338118.

782 **47** .Sprenger K, Conti S, Ovchinnikov V, Chakraborty A, Karplus M. Multiscale affinity maturation
783    simulations to elicit broadly neutralizing antibodies against hiv. *submitted, PLoS Comput. Biol.* (2021).

784 **48** .Casten L, Pierce S. Receptor-Mediated B Cell Antigen Processing. Increased Antigenicity of a
785    Globular Protein Covalently Coupled to Antibodies Specific for B Cell Surface Structures. *J. Immunol.*
786    *(Baltimore, Md. : 1950)* **140** (1988) 404–10.

787 **49** .Alberts B, Johnson A, Lewis J, Raf M, Roberts K, Walter P. *Molecular Biology of the Cell* (New York:
788    Garland Science), fourth edn. (2002).

789 **50** .Press W, Flannery B, Teukolsky S, Vetterling W. *Numerical Recipes in FORTRAN 77: The Art of*
790    *Scientific Computing* (Cambridge University Press), second edn. (1992).

791 **51** .Anderson S, Khalil A, Uduman M, Hershberg U, Louzoun Y, Haberman A, et al. Taking Advantage:
792    High-Affinity B Cells in the Germinal Center Have Lower Death Rates, but similar rates of division,
793    compared to low-affinity cells. *J. Immunol. (Baltimore, Md. : 1950)* **183** (2009) 7314–25. doi:10.4049/
794    jimmunol.0902452.

795 **52** .Gitlin A, Mayer C, Oliveira T, Shulman Z, Jones M, Koren A, et al. HUMORAL IMMUNITY. T Cell
796    Help Controls the Speed of the Cell Cycle in Germinal Center B Cells. *Science (New York, N.Y.)* **349**
797    (2015) 643–6. doi:10.1126/science.aac4919.

798 **53** .Amitai A, Mesin L, Victora G, Kardar M, Chakraborty A. a Population Dynamics Model for Clonal
799    Diversity in a Germinal Center. *Front. Microbiol.* **8** (2017) 1693. doi:10.3389/fmicb.2017.01693.

800 **54** .Hill TL. *an Introduction to Statistical Thermodynamics* (New York: Dover) (1986).

801 **55** .El Shikh M, El Sayed R, Sukumar S, Szakal A, Tew J. Activation of B Cells by Antigens on Follicular
802    Dendritic Cells. *Trends. Immunol.* **31** (2010) 205–11. doi:10.1016/j.it.2010.03.002.

803 **56** .Saphire E, Parren P, Pantophlet R, Zwick M, Morris G, Rudd P, et al. Crystal Structure of a Neutralizing
804    Human IGG Against HIV-1: A Template for Vaccine Design. *Science (New York, N.Y.)* **293** (2001)
805    1155–9. doi:10.1126/science.1061692.

806 **57** .Ovchinnikov V, Louveau J, Barton J, Karplus M, Chakraborty A. Role of Framework Mutations
807 and Antibody Flexibility in the Evolution of Broadly Neutralizing Antibodies. *eLife* **7** (2018) 0.
808 doi:10.7554/eLife.33038.

809 **58** .Ise W, Kurosaki T. Plasma Cell Differentiation During the Germinal Center Reaction. *Immunol. Rev.*
810 **288** (2019) 64–74. doi:10.1111/imr.12751.

811 **59** .Rundell A, DeCarlo R, HogenEsch H, Doerschuk P. the Humoral Immune Response to Haemophilus
812 Influenzae Type B: A Mathematical Model Based on T-Zone and Germinal Center B-Cell Dynamics. *J.*
813 *Theor. Biol.* **194** (1998) 341–81. doi:10.1006/jtbi.1998.0751.

814 **60** .Kelsoe G. *in Situ* studies of the germinal center reaction. *Adv. Immunol.* **60** (1995) 267–288.

815 **61** .Murphy K. *Janeway's Immunobiology* (New York: Garland Science), eighth edn. (2012).

816 **62** .Liu Y, Johnson G, Gordon J, MacLennan I. Germinal Centres in T-Cell-Dependent Antibody Responses.
817 *Immunol. Today* **13** (1992) 17–21. doi:10.1016/0167-5699(92)90199-H.

818 **63** .Leanderson T, Källberg E, Gray D. Expansion, selection and mutation of antigen-specific b cells in
819 germinal centers. *Immunol. Rev.* **126** (1992) 47–61. doi:10.1111/j.1600-065x.1992.tb00630.x.

820 **64** .Eaton JW, Bateman D, Hauberg S, Wehbring R. *GNU Octave Version 4.0.0 Manual: A High-Level*
821 *Interactive Language for Numerical Computations* (2015).

822 **65** .Moin P. *Fundamentals of Engineering Numerical Analysis* (Cambridge University Press) (2001).

## FIGURE CAPTIONS

**Figure . 1.** Influenza hemagglutinin (HA) spike protein colored by residue conservation. Sequences of avian, swine and human influenza type A spike proteins were downloaded from the NIH influenza research database(3); conservation was computed in MATLAB(4) after clustering the sequences to 97% identity and multiple alignment. The HA structure was taken from PDB entry 3LZG(5), and the image was generated using Visual Molecular Dynamics.(6) Only the $C_\alpha$ atoms are shown.

**Figure . 2.** Schematic of the GC model used in this study. (A) Model overview: (1) B-cells are activated upon binding to antigen presented on follicular dendritic cells (FDCs, not explicitly modeled); (1a) in the optional T-cell help model (see text) B-cells are activated when the major histocompatibility complex receptor (MHC2) binds to the T-cell receptor (TCR); B-cell activation rescues B-cells from apoptosis, allowing them to mutate and proliferate; (2) depending on the activation signal, B-cells can differentiate to plasma cells (2a), to memory B-cells (2b), or undergo apoptosis (2c); (3) Plasma cells secrete antibodies (Abs), which also compete with B-cell receptors for antigen (4), which is the essential aspect of the Ab feedback model(26) (see text). (B) Hypothetical modes of antibody binding to influenza spikes; bivalent binding without strain (top) corresponds to cooperative binding by antibody arms; bivalent binding with strain (middle) corresponds to noncooperative binding; monovalent binding (bottom) is assumed to be the dominant mode of binding of anti-HA stem antibodies (see Methods for details).

**Figure . 3.** Comparison of simulation and experiments. A: Total B-cells, B: Memory B-cell production rate, C: Plasma cell production rate. Experimental data for panel A was generated from the GC cross-sectional areas plotted in Fig. S1B of Ref. 36, and converted to B cell counts as done in Ref. 37; The lower and upper error bars in panel A corresponds to 30% and 70% quantiles, respectively ; experimental data for panels B & C was taken from Fig. 4 of Ref. 37, who obtained raw data from Weisel et al. (38); the error bars in B & C correspond to approximately one SD.

**Figure . 4.** Effect of antibody valency and epitope occlusion on GC properties. Left column (A–C): noninteracting B-cell case ($o = 0$); Right column (D–F): fully interacting B-cell case ($o = 1$); A,D: Total B cells; B,E: Memory cell production rate; insets: total MBC population at end of simulation; C,F: Average affinity of B-cells and MBCs. For the definition of occlusion $o$, see Sec. 4.0.5.

**Figure . 5.** Effect of initial affinity advantage on the growth of monovalent B-cells in the fully interacting B-cell case ($o = 1$). Panels A–C show the same quantities as Fig. 4A–C; The affinity distribution corresponding to BCR#1 was shifted toward higher values relative to BCR#2 and BCR#3 (panel D).

**Figure . 6.** Fraction of MBC#1 ($\zeta$, defined in the text) at the end of six GC simulations for different initial affinity advantage values *vs.* total number of BCR/Epitope pairs. A: BCR#1 is cooperatively bivalent ($K_{eq}^{12}=10K_{eq}^{11}$); B: BCR#1 is monovalent ($K_{eq}^{12}=0$).

**Figure . 7.** Effect of BCR valency and epitope concentration on GC evolution, with $o = 1$ (fully competitive case). Panels A–C show the same quantities as Fig. 4A–C; $\alpha_1^T = 2$, $\alpha_2^T = 1$, $\alpha_3^T = 1$; $\alpha^T$ is the nondimensional AG concentration (see Sec. 4.0.6).

**Figure . 8.** Fraction of MBC#1 *vs.* number of sequential GC simulations for different initial affinity advantage values and different AG#1 concentrations, with 10 total BCR/Epitope pairs. A & B: MBC#1 fraction ($\zeta$) at the end of simulations. A: BCR#1 is cooperatively bivalent ($K_{eq}^{12}=10K_{eq}^{11}$); B: BCR#1 is monovalent ($K_{eq}^{12}=0$). The four sets of panels A and B show the effect of increasing occlusion from $o=0$ (no competition) to $o=1$ (full competition).

## CONTRIBUTION TO THE FIELD

823  Viral pathogens such as influenza, coronavirus, and HIV remain a high source of global mortality and
824  morbidity, because their rapid rate of mutation presents serious challenges to vaccine design. Many anti-
825  viral vaccines are currently in development, but to be effective against future strains, they must raise
826  host immunity to structurally conserved regions of the virus. Unfortunately, viral adaptation conceals
827  such conserved epitopes from the immune system and guides host immunity towards highly variable,
828  distracting epitopes. Success in developing vaccines against rapidly mutating viruses depends critically on
829  understanding and learning how to overcome this distraction. We have developed and validated a computer
830  model of adaptive immunity, in which we simulate the development of high-affinity antibodies to antigenic
831  epitopes with different properties. We find that the mode of antibody binding to antigens can explain
832  experimental results that show that the immune system is easily directed from certain conserved epitopes
833  toward variable ones. We performed additional simulations to explore the influence of antigen concentration
834  on the competition between different antibodies, and whether raised immunity to conserved (but concealed)
835  epitopes could be long lasting. Our results are expected to inform design of universal vaccines.