Published in final edited form as:

Phys Biol. 2013 April ; 10(2): 025001. doi:10.1088/1478-3975/10/2/025001.

A one-shot germinal center model under protein structural stability constraints

Sana Raoof^{1,3}, Muyoung Heo², and Eugene I Shakhnovich^{1,4}

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

² Department of Physics, Pusan National University, Busan 609-735, Korea

³ Harvard College.

Abstract

The germinal center reaction is the process by which low-affinity B cells evolve into potent, immunoglobulin-secreting plasma and memory B cells. Since the recycling hypothesis was created, experimental studies have both tracked movement of a small population of B cells from the light zone into the dark zone, supporting the recycling model, and parallel to the light zone–dark zone interface, indicating a one-way trajectory. We present a novel, sequence-based *ab initio* model of protein stability and protein interactions. Our model contains a dark zone region of clonal expansion and somatic hypermutation and a light zone site of antigenic selection. We show not only that a one-shot model is sufficient to achieve biologically-realistic rates of affinity growth, population dynamics, and silent:non-silent mutation ratios in the complementary determining region and framework region of antibodies, but also that a stochastic recycling program with or without realistic constraints on the structural stabilities of GC antibodies cannot produce biologically-observed affinity growth, population dynamics or silent:non-silent mutation profiles. The effect of recycling erases affinity gains made by potent antibodies cycling back from the light zone and causes B cells to pool in the dark zone under high replication rates.

Introduction

Germinal centers (GC) are transient microenvironments in secondary lymphoid organs where B cells producing antibodies (Abs) specific to an incoming antigen (Ag) undergo clonal expansion, somatic hypermutation and antigenic selection [1]. This ‘germinal center reaction’ results in up to 10-fold affinity maturation in the course of the humoral immune response [2]. By 10 days post-immunization, the germinal center is observed to contain histologically-distinct compartments, including a ‘dark zone’ (DZ) and a ‘light zone’ (LZ) [1]. Exponential growth of B cells in the DZ, centroblasts, is accompanied by somatic hypermutation (SHM) of B cell receptor genes. Centroblasts downregulate the expression of surface immunoglobulins (‘B cell receptors,’ or Igs) [1] and, therefore, are not subject to selection at this stage [3]. Hypermutation of centroblasts in the DZ creates a diverse repertoire of mutated sequences of the variable (Ag-specific) regions of Igs. Antigenic

⁴ Author to whom any correspondence should be addressed. Eugene@belok.harvard.edu.

selection takes place in the LZ, where follicular dendritic cells (FDCs) presenting specific Ags reside. These FDCs express the toll like receptor 4 (TLR4) [4]. While B cells expressing surface Igs with higher affinity have higher chance of being positively selected and eventually exit the GC as plasma or memory cells, those expressing lower affinity surface Igs have higher chance of being removed through an apoptotic pathway [3, 5–8]. Recent studies implicate the TLR4 pathway in GC growth and neogenesis, interaction between Abs and Ags in the LZ, and volume of plasma cells produced per GC [4]. The exact spatiotemporal trajectory of cells in the GC reaction still remains unknown [8, 9].

The driving force for organization of the germinal center was explored by Cyster *et al* in 2004 [10]. CXCR4^{+/+} or CXCR4^{-/-} fetal liver cells were transferred into lethally irradiated mice. After reconstitution and immunization CXCR4 was found to be necessary for the normal expression of chemokine CXCL13 and location of CD23⁺CD35⁺FDCs in the LZ, and location of centroblasts (dividing B cells) in the DZ at 5 h postimmunization. Overall, the authors found a DZ in fewer than 7% of CXCR4^{-/-} GCs compared to over 75% of CXCR4^{+/+} GCs, indicating that CXCR4 is necessary for GC compartmentalization [10]. The structural development of the GC was further explored in three two-photon microscopy studies. Together, they suggest that Brownian motion with a directional persistence time of ~1 min accounts for the observed migration of B cells in the GC. [11] Only about 5–10% of cells are reported to move between zones within an hour-long period; hence, theoretical GC models based on separate DZ–LZ compartments rely on chemotaxis to explain the directional movement of B cells. Figge *et al* provide a statistical GC kinetics model to support that persistent random walks are sufficient to account for DZ–LZ migration frequencies. The authors then offer a functional GC model, using three-dimensional GC structure and individual Ab–Ag interaction duration times. This more complex model suggests that low to moderate chemotactic signal strengths (of CXCL12 and CXCL13) are needed to form a DZ–LZ structure. To prevent unrealistic B cell accumulation in the GC, the authors conjecture that B cells are desensitized to chemotactic signals over time [11]. As a follow up, Boer *et al* showed that, even after accounting for artifacts in cell-tracking data, low chemotactic signal strength is sufficient to explain movement from DZ to LZ [12]. Straight-track movement into the LZ applies to nearly half of B cells in the DZ, as observed from experimental imaging data [12]. In their study, B cell speeds and migration angles with respect to the DZ–LZ boundary are found to heavily influence the arrival time of B cells in the LZ. The authors do not find evidence for cyclic reentry of centrocytes into the DZ, although they point out that perhaps only a small population of cells moves back into the DZ, or their dataset on LZ cells is insufficient to observe recycling trends [12].

The 1993 recycling hypothesis by Kepler and Perelson conjectures that B cells cycle between centrocyte and centroblast states several times before differentiating or undergoing apoptosis. Centroblasts are subject to regular bouts of diversifying mutation [13, 14]. The recycling hypothesis remains a highly contested model since other theoretical models have been developed which produce affinity growth while relying on ‘one-shot’ dynamics. In 2006, Moreira *et al* showed that an experimentally-observed volume of antibody-forming cells could be generated by a theoretical one-shot germinal center model by relying on a realistic two-level selection scheme (antigenic and T-cell selection) [14]. The results of this

model correlate significantly with experimental data on the humoral immune response to hapten (4-hydroxy-3-nitrophenyl)acetyl (NP) [14], which is a well-established experimental model system for studying the GC reaction [15].

Real-time imaging experiments reveal many of the mechanistic aspects of GC affinity maturation [3, 8, 9, 16]. Victoria *et al* exploited the difference in cell surface markers present on DZ and LZ B cells to sort the populations by flow cytometry. LZ cells were in S phase and enriched in activation markers, cell-surface molecules, and apoptotic regulators, whereas DZ cells were in S, G2 and M phases and contained mitotic regulatory genes, indicating that competition and selection occur in the LZ, whereas proliferation occurs in the DZ. Furthermore, the authors photoactivated green fluorescent protein (GFP) using two photon microscopy in specific regions of the GC to track movement of DZ and LZ B cells over time. They report a flux of cells from the DZ to LZ and a smaller flux in the reverse direction, indicating that only certain centrocytes are selected to ‘recycle’ between zones. Conversely, Hauser *et al* also used two photon microscopy to show that most B cell movement is parallel to the DZ–LZ boundary, and that B cell proliferation does occur in the LZ [9], while some B cell selection occurs in DZ [16]. These results suggest selection and proliferation may occur in both compartments, making recycling unnecessary.

A basic, sequence-based, multi-scale model of immunity was presented in [17]. The model addresses the basic question of dynamics of affinity maturation in GC in response to viral infection. Here, we extend the model much further to consider affinity maturation with and without recycling. We present an explicitly sequence-based ab initio model of protein stability and protein interactions, with distinct dark zone and LZ regions within the GC and apoptotic pathways responding to structural stabilities of Abs. We show not only that a one-shot model is sufficient to achieve biologically-realistic rates of affinity growth, population dynamics, R/S ratios in the CDR (complementarity determining region) and framework region (FR), and monoclonality, but also that optimal mutation rates, sufficient affinity growth and realistic RS ratios cannot be achieved by stochastic recycling with or without realistic constraints on the structural stabilities of GC antibodies.

Results

We model three regions: a DZ, LZ and a GC exterior. Every protein (Ab and Ag) in our model corresponds to an explicit gene sequence in a coarse-grained lattice protein model, making it possible to exactly determine the strength of interaction between proteins (Abs and Ags) and the structural stability of single proteins (Abs). For simplicity, antibodies, B cells and immunoglobulins all refer to the same entity in our model, although we refer to them specifically where appropriate. B cells in our model can be recruited into the germinal center, replicate, mutate, die, and exit the germinal center. We assume the following. First, stronger interaction energies between Abs and Ags are a selective advantage for the Abs’ chance of survival in the LZ. Second, B cells without B cell receptors in DZ cannot be subject to antigenic selection. Finally, we assume that loss of stable Ab structure post-mutation will result in B cell death due to detrimental structure-function consequences.

A diverse set of randomly chosen Abs is generated, all with different random affinities to a single antigenic sequence used in the simulation. Antibodies with higher affinities to the antigen are more likely to be recruited into the GC, and a lower limit of initial P_{int} for GC entry is set at 0.3. Once B cells enter the DZ, they undergo monoclonal expansion at a high replication rate and then somatic hypermutation. These processes create a large and diverse pool of DZ B cells, which enter the LZ at an experimental rate r_{LZ} = once per 6 h [18]. LZ cells are subject to antigenic selection; high affinity LZ cells exit the GC as memory/plasma cells or, in models with recycling, reenter the DZ. There is a lower replication rate (once per day) in the LZ and no mutation.

We simulated germinal center reactions with and without recycling and constraints on the structural stability of Abs in the reaction, and birth rates that are equal to the biologically observed rate. Our models are summarized in table 1.

Recycling erases affinity maturation

Centroblasts, or large, dividing GC B cells, are mainly present in the DZ, whereas small B cells, or centrocytes, predominate in the LZ. Since our model does not treat spatial distribution of cells explicitly, the movement of B cells from the DZ to LZ in our simulation is implicit in centroblast differentiation. In the one-shot model presented here (A), we find that optimal affinity maturation occurs in the range of realistic B cell replication rates (~6 h in the DZ) and centroblast differentiation rates (~6 h) [18]. In the recycling model (B), no affinity maturation is observed at any rates. The affinity gains caused by antigenic selection in the LZ allow a small number of cells in our recycling model to reenter the DZ. The affinity maturation achieved by this potent population is lost to detrimental mutations in SHM and the high DZ replication rate. Since our simulation calculates protein structural stability after every round of mutation, the harmful effects of recycling are illuminated: affinity maturation is erased in every round of DZ reentry. This iterative loss of affinity is because there is no antigenic selection in the DZ, so hypermutation and population expansion of all DZ B cells ‘washes out’ the population of high-affinity centrocytes that recycled back from the LZ. Similarly, we see no affinity maturation in recycling model C, where the structural stability constraints are released. In this model, P_{int} grows to 0.3 as B cells are initially recruited into the DZ; hypermutation lowers P_{int} drastically until a small population of cells enters the LZ and reestablishes P_{int} to 0.3. The P_{int} does not grow beyond this threshold for entry since most cells reside in the DZ: as they not express mature antibodies in DZ their affinity is not high enough to differentiate into centrocytes, and those few that do occasionally differentiate will recycle back into the DZ anyway. Other theoretical models fail to consider protein integrity through sequence analysis [13, 14]. Hence, the optimal region of B cell affinity maturation in our model disappears after adding recycling with or without a protein structural stability constraint (models B, C) (figure 1). To control for the effect of our particular recycling schedule, we also simulated a ‘slowed down’ recycling model, in which B cells spent, on average, twice as much time in the LZ as cells in model B. This steady loss of centrocytes to apoptotic pathways and to the DZ, where the population is rapidly expanding and ‘diluting’ the population of the LZ, stunts P_{int} at a pre-LZ level (0.3). Since B cells that recycle to the DZ are not guaranteed to re-enter the LZ, recycling represents an effective death of some high affinity B cells.

Population dynamics

Only our one-shot model achieves population dynamics that are clearly consistent with experimental observations (figure 2) [1]. The DZ population reaches a peak of ~3000 B cells by 5 days post-immunization, when centroblasts differentiate into centrocytes in the LZ. At this time, there is pronounced LZ population growth, allowing for competitive selection to achieve the affinity growth seen in model A. In the recycling model, the DZ persists for 15 days post-immunization since LZ cells are moving back to the DZ and replicating at high rates. Moreover, in model C, cells in the LZ are removed by apoptosis due to the constraint on their structural stability following cycling of deleterious hypermutation, which limits LZ growth. The additional loss of B cells to the DZ in recycling models B and C further explains why the LZ does not grow, so affinity maturation is not seen. When structural stability requirements are lifted, the LZ continues to lose cells recycling back to the DZ, but a new source of population loss is introduced. Unstable Ab proteins whose binding surfaces were mutated in SHM survive the transition to the LZ in this model, but their low affinity interactions with antigenic proteins cause apoptosis in the LZ. These very unstable Abs are removed quickly due to low P_{int} , preventing the LZ from growing. This ‘pooling’ of B cells in the DZ causes affinity stagnation, as DZ B cells do not express B cell receptors and aren't subject to competitive selection.

Positive and negative selection in the CDR and FR

The measure of selection pressure on proteins is the R/S ratio between replacement and silent mutations (the ratio of mutations which code for amino acid change to those mutations which do not translate to amino acid change). The R/S ratio for neutral mutations is 3. Experimental data indicates an R/S ratio > 3 in the CDR region and $R/S < 3$ in the FR region once Abs enter the LZ, indicating positive selection in the CDR to evolve potent Igs and negative selection in the FR to preserve their structural integrities [19]. In our one-shot model (A), we find strongly fluctuating R/S ratios for the first 5 days post-immunization, since B cells are still in the DZ subject to SHM without selection, as seen in figure 2. Once cells move into the LZ, there is initial negative selection reflecting death of low-affinity Abs through apoptosis. With additional time in the LZ to replicate slowly while undergoing affinity maturation, however, there is clear positive selection in the CDR and negative selection in the FR, (figure 3) as expected. After about 9 days, R/S ratios fall under 3 in both regions, as this is the period for plasma and memory B cells to leave the GC altogether (figure 3). We can trace the average number of mutations over time in a single GC to support this explanation (figure 4). Even after the DZ disappears and SHM terminates, the average number of mutations in any Ab sequence changes due to the death of low-affinity antibodies in the LZ. In the DZ, cells accumulate an average of six mutations during SHM and before moving into the LZ. Once the DZ disappears around day seven, the diversified, high affinity B cells survive antigenic selection, whereas the immature B cells in the DZ die, increasing the average mutation count to nine per B cell. In our recycling model (B), affinity progress cannot be made because the structural integrity of proteins is compromised in every iteration. Hence, R/S ratios remain strongly negative. If protein structural constraints are released, (model C) CDR R/S ratios decrease once cells begin SHM in the DZ (before day 3 post-immunization) and increase during the first bout of antigenic selection in the LZ. Since

the majority of the GC cells are ‘pooled’ in the DZ in this model, their structural integrity is compromised harshly by hypermutation, which may cause death due to P_{nat} constraints (model B) or apoptosis due to low affinity in the LZ (models B and C). Hence, ‘positive’ selection is seen in both the CDR and FR once B cells enter the LZ, although empirical R/S ratios in the FR are <3 [19]. Judgment of selection pressures by R/S ratios in this situation is not straightforward due to the volatile structural conformation of the rapidly mutating cells pooling in the DZ.

Discussion

The explicit, sequence-based genotype-phenotype relationship in our model allows us to reexamine the feasibility of the assumptions about the affinity landscape used in previous models [13, 20–22]. The earlier phenomenological models promoting the recycling hypothesis assumed a simple string model of affinity whereby affinity is achieved by exact matching of complementary pairs of corresponding amino acids on the antigen and antibody, rendering such affinity models in a class of exact ‘string matching.’ In contrast, our model considers interactions between Ab and Ag surfaces using a realistic interaction scheme for all 20 types of amino acids based on empirical amino acid contact energies [23].

A coarse-grained presentation of proteins as lattice polymers has been employed, with success, in many studies of fundamental aspects of protein folding thermodynamics and kinetics [24–26]. Lattice models, despite their simplicity, reproduce key fundamental properties of proteins—their vast configuration space and unique sequence–structure relationship. Famous ‘Levinthal paradox’ of protein folding has been resolved in a lattice model study [27], which showed that sequences selected to be stable in their native configurations can reach their native states on realistic times scales much shorter than required for exhaustive search over all configurations. The advantage of many lattice models is in their exact nature whereby all configurations can be enumerated and statistical mechanics can be evaluated exactly and in simple kinetics. Recently lattice models were used to study more complex systems such as interacting proteins in cytoplasm [28], evolution of protein–protein interactions and abundances [29] and dynamics of affinity maturation in GC [17]. Lattice models provide a coarse-grained representation of protein structure, where amino acids are reduced to interacting centers occupying lattice sites. A phenomenological potential knowledge-based potential (KBP) derived from statistics of contacts between amino acids in proteins [23] is often used to represent energetics of intra- and inter-molecular interactions in lattice proteins. KBPs are widely used in various approaches to protein design and protein structure prediction [30]: they provide a physically realistic presentation of interactions in protein systems [31]. In particular recent study demonstrated that the use of Miyazawa–Jernigan KBPs to model evolution of protein–protein interactions provided realistic amino acid compositions on interfaces between interacting proteins [29]. As any simplified modeling lattice presentation of proteins has its limitations. For example, they cannot provide a detailed geometry of Ab–Ag interface and as such are not suitable to predict several detailed aspects of evolutionary trajectories of virus–antibody ‘arms race’ such as dominant escape routes for viruses. Nevertheless modeling three-dimensional structure of antibody and antigen and the exact nature of sequence–structure and sequence–interaction affinity relationships in lattice

models provides a unique opportunity to study fundamental aspects of immune dynamics within a microscopic ab initio multiscale approach.

The present version of the model does not consider the spatial dimension of the problem. While we make a clear distinction between compartments in GC, it concerns the features of corresponding cell types rather than their actual physical separation. In the future we plan to extend the model to include spatial separation between DZ and LZ and cell migration between them. This will add more realism to the analysis and will make it possible to analyze spatiotemporal events in greater detail.

Our model emphasizes that total free energy (potential of mean force) of antibody-antigen interaction determines binding affinity rather than one-by-one amino acid contacts between an Ab and Ag. A low *total* interaction energy between an Ab and Ag can be achieved by multiple combinations of amino acid sequences in CDRs rather than via a unique string-matching combination as assumed in earlier phenomenological models [21, 32]. This key physical aspect of protein-protein interactions gives rise to significant plasticity of Ig sequence space. An additional layer of realism in our model is that it considers the constraint on Ig stability explicitly, providing a realistic description of interplay of selection for affinity with the constraint to maintain stable FRs of Abs. We find that the enormous size of the Ig sequence space provides enough 'room' for affinity maturation via SHM despite these constraints. The obvious concern is that SHM may prevent B cells from stable proliferation because of the genetic load of deleterious mutations. The genetic instability caused by a high mutation rate of GC B cells in phenomenological models, along with GC compartmentalization, motivated Kepler and Perelson to propose the phasic mutation schedule as an optimal strategy for B cell proliferation during affinity maturation [13, 20, 22]. The phasic mutation schedule can drive B cells to be diversified during the 'on' phase of SHM and to stably proliferate during the 'off' phase. In our model, the stability of Ig sequences is monitored explicitly and the effect of mutations on stability and affinity is evaluated exactly; we find that the genetic load of deleterious mutations is not as severe as postulated in *a priori* models because of the effect of compensating beneficial mutations which affect affinity or stability of Igs. Our simulations show that a 'one-shot' SHM program with a phasic mutation schedule results in potent and stable Ig molecules.

In order to check the effect of protein structural constraints on affinity maturation of Abs in a one-shot model, we incorporated the constraint that B cells are removed when the structural stabilities of their Igs drop below a certain value. Realistic mutation rates, centroblast differentiation rates, and P_{nat} constraints yield optimal affinity maturation in the one-shot model presented here, while affinity maturation is not achieved under realistic parameters for the recycling model. Previous theoretical studies of the GC reaction did not consider the biologically important implications of mutations on antibody active site conformation or antibody stability, both of which can lead to apoptosis. In model C, the effect of protein stability is disregarded, but the low affinity Abs that result from SHM without structural constraints leads to a high death rate by antigenic selection in the LZ. We show that a one-shot model is necessary and sufficient for affinity maturation using experimentally observed parameters and results in realistic affinity maturation ($P_{\text{int}} \sim 0.8$).

Using our microscopic model, we showed how an effective adaptive immune response, which is very similar to a real adaptive immune response within an organism, develops. Most importantly, we demonstrated that our model of adaptive dynamic immune response captures many complexities of real immune systems such as optimal rates of SHM and apoptosis. The striking qualitative and quantitative similarity of many general features of the immune response found in our model to experimentally observed aspects of humoral immune response attest to the robustness of the physical principles of adaptive immunity.

Experimental approaches to test this model's predictions may look at *in vivo* germinal center reactions at different time-points post-immunization. The dissociation constant (K_d) of Abs must be measured to test whether affinity growth from GC recruitment to final LZ exit is comparable to that predicted by our model. B cell DNA should be sequenced to find R/S ratios in the CDR and FR. The total number of mutations acquired should be compared to our prediction of about nine mutations after SHM.

Insights into the spatiotemporal schedule of the GC reaction may guide medical investigations into GC-related disorders, such as Hodgkin's and non-Hodgkin's lymphoma. Hodgkin's lymphoma patients exhibit swollen lymph nodes and lack lymph node structure, and the cause of their disease is unknown [33]. Kuppers *et al* report that Ab gene rearrangements in 16 of 17 Hodgkin's patients were found to be monoclonal [34]. Other studies support the clonality of tumor cells in Hodgkin's patients [35, 36]. Further examination of the gene coding of the variable region of Abs reveals significant diversification (and monoclonality), which is uniquely attributed to SHM in the GCR [19], and suggests that Hodgkin's cells have been subject to antigenic selection. In four cases, mutations generating a 'stop codon' were found; such detrimental mutations normally cause apoptosis in GC B cells [34] (reflected by the clearing of B cells in our model if P_{nat} falls below 0.6 or if the native structure is changed post-mutation). The ability of 'crippled' Abs to survive selection pressures and leave the GC evidences oncological transformation, a phenomenon the authors conjecture may be prevalent in Hodgkin's cells [34]. Our one-shot model may be modified so that structural stability constraints are removed and changes in conformation of Ab proteins are allowed. This will allow crippled Abs to enter into the LZ and compete for antigenic interaction. Although experimental studies elucidating the LZ selection mechanism in Hodgkin's lymphoma patients do not exist, it is only necessary for one malignant, crippled B cell to escape the germinal center microenvironment to generate lymphoma, since it is monoclonal. We may assign a rate at which cells undergo oncological transformation in the DZ and, if a malignant Abs enters the LZ, it may be subject to a lower apoptotic rate or be free from antigenic selection. In such a model, a crippled centrocyte may instead exit the GC as a memory/plasma cell with some probability.

The incidence of both non-Hodgkin's lymphoma (NHL) and Hodgkin's lymphoma in patients is higher than that predicted by chance; in such patients, NHL is predominantly seen in B cell follicles and has identical clonal origins to Hodgkin's cells [37]. Moreover, an average of six of the seven mutations in the Hodgkin's cells and 21 mutations in NHL cells were common, indicating that a post-SHM GC B cell is the common 'seeder cell' for both diseases [37]. Disparate selection of cells during competition for antigenic interactions in the LZ is likely involved in the divergence of phenotypes, such as the huge, multinuclear cells

lacking Ab expression in Hodgkin's lymphoma only [37]. NHL cells resemble GC B cells morphologically and create a normal follicular structure [38]. The force driving this GCR-like trajectory of NHL may be that malignant B cells adhere to GCs in the same way as healthy B cells [38]. Once inside the GC environment, NHL cells are commonly mutated to overexpress bcl2, blocking apoptosis and driving lymphadenopathy, and are associated with the lack of a mantle zone—GC boundary, and a centrocytic phenotype [39]. Our model may be modified in the following ways: P_{nat} requirements for GC entry will be relaxed to model a fluid mantle zone—GC border, once recruited into the GC, a B cell will proliferate for an extended time (we may adjust we may τ , the relaxation time in the exponential term of the DZ birth rate), will escape apoptosis post-SHM despite a low P_{nat} or structural change with some small probability, and will enter the LZ and leave with some probability to re-enter another GC to represent adhesion to hyperproliferative germinal centers.

Methods

Model properties

The detailed structure of this model has been presented earlier [17]. Here, we give a brief review of its design. Proteins are modeled as 27-unit compact polymers so that their thermodynamic properties, stabilities and interactions can be calculated exactly from their primary sequences. We consider a $3 \times 3 \times 3$ lattice model proteins and reduce protein structural space from the full set of 103 346 maximally compact (cubic) 27-mer structures [40] to randomly and uniformly chosen 10 000 structures in order to save computer time. Each protein folds into the native structure, which has the minimum energy out of these 10 000 structures. Only amino acids that are nearest neighbors in structure on the lattice, but are not covalently linked can interact. For each protein conformation we define the energy as sum over all interactions:

$$E_i(g) = \sum_{k=1}^{20} \sum_{l=1}^{20} \varepsilon(k, l) n(k, l), \quad (1)$$

where k and l correspond to the amino acid types involved in an interaction (there are total 20m different types of amino acids, like in real proteins), $\varepsilon(k, l)$ is the energy parameter describing the energy of interaction between amino acids of type k and l in Miyazawa–Jernigan 20*20 interaction potential matrix, and $n(k, l)$ is the total number of contacts between amino acids of types k and l in the current structure.

Protein stability is calculated in terms of the Boltzmann probability (P_{nat}) of the native (i.e. lowest energy) structure of the protein at a given temperature,

$$P_{\text{nat}}(g) = \frac{\exp[-E_0(g)/T]}{\sum_{i=1}^{10\,000} \exp[-E_i(g)/T]}, \quad (2)$$

where g is a protein sequence in a 20 amino acid type alphabet, $E_0(g)$ and $E_i(g)$ are energies of a protein in the native (i.e. lowest energy) and i th structure respectively, and T is an environmental temperature scaled by the Boltzmann constant.

In order to calculate the interaction strength between two proteins, we consider rigid-body docking between two $3 \times 3 \times 3$ lattice proteins (figure 5). Each cubic lattice protein has six binding surfaces and four-fold rotational symmetry of a binding surface (only the strongest binding positions when two faces overlap completely, making nine interaction bonds, are taken into account.) Therefore, a pair of lattice proteins has $6 \times 6 \times 4 = 144$ binding conformations. A diagram of two lattice proteins interacting is provided in figure 5.

Protein interaction strength is defined by the Boltzmann probability (P_{int}) of specific binding between two proteins, g_1 and g_2 , as follows:

$$P_{\text{int}}(g_1, g_2) = \frac{\exp[-f \cdot E_0(g_1, g_2)/T]}{\sum_{j=1}^{144} \exp[-f \cdot E_j(g_1, g_2)/T]}, \quad (3)$$

where $E_0(g_1, g_2)$ is the binding energy of two proteins in the binding mode corresponding to the lowest energy out of all 144 possible binding modes, and f is a pre-factor that takes into account possible different strengths of intra-protein and inter-protein interactions. $E_0(g_1, g_2)$ is defined as:

$$E_0(g_1, g_2) = \sum_{k=1}^{20} \sum_{l=1}^{20} \varepsilon(k, l) m(k, l), \quad (4)$$

where $m(k, l)$ represents the total number of contacts between protein interfaces of proteins g_1 and g_2 , and k and l are the indices for an amino acid in the protein interface of proteins g_1 and g_2 , respectively. In these calculations, we use the Miyazawa–Jernigan pairwise contact potential for both protein structural and interaction energies, but scale protein–protein interactions by a constant factor [23]. We chose $T = 0.85$ in Miyazawa–Jernigan potential dimensionless energy units.

Each Ab protein is modeled individually by keeping track of its basepair and amino acid sequences. The model discards Abs of stability $P_{\text{nat}} < 0.6$ from the simulation. Also, if a mutation changes the lowest energy structure of a protein or its binding site, the protein is discarded.

Conversion of simulation time steps to real-time

We set the division rate of B cells to a constant, $b = 0.01$. Because the standard division cycle of eukaryotic cells takes 24 h [41], 69.6 time steps in our simulation correspond to 1 day. Once B cells are recruited into the GC DZ, they assume exponentially decaying birth rates in the dark zone (described below in equation 7) which are initially very fast; the B cell activation factor, α , in equation (7) is set to 5.0 to obtain 3–4-fold faster division at moderate antibody–antigen interactions strengths, $P_{\text{int}}(\text{Ab}, \text{Ag})$, ranging from 0.4 to 0.6.

Simulation setup

At the beginning of the simulation, we randomly chose one sequence for the antigen, which is fixed throughout the whole simulation. A diverse pool of 941 different antibodies is randomly generated by subjecting a single Ab sequence to point mutations. Their initial

affinities to a single antigenic sequence are random. Ab proteins with $P_{\text{nat}} < 0.4$ are discarded from the initial pool, and the average initial $P_{\text{nat}} = 0.51$. The entry rate of germline B cells into GC is given in equation (5).

$$r^{DZ} = \begin{cases} r_0^{DZ} \left(1 - \exp \left(-\frac{P_{\text{int}}(Ab, Ag) - P_{\text{int}}^{\text{entry}}}{\sigma^{\text{entry}}} \right) \right) & \text{for } P_{\text{int}}(Ab, Ag) \geq P_{\text{int}}^{\text{entry}} \\ 0 & \text{for } P_{\text{int}}(Ab, Ag) \leq P_{\text{int}}^{\text{entry}} \end{cases} \quad (5)$$

where r_0^{DZ} is a base entry rate of B cells into DZ of GC, $P_{\text{int}}(Ab, Ag)$ is the thermal probability that an antibody and an antigen interact with each other in a predetermined binding mode as a measure of the B cell's affinity for the antigen, and $P_{\text{int}}^{\text{entry}}$ is a threshold affinity required for germline B cells to be recruited into DZ (set at 0.3). σ^{entry} controls the steepness of the logistic function of the entry rate and reflects the time window for DZ population growth reported experimentally [1]. Equation (5) mimics the initial dynamics of GC formation: the antigen presented by the follicular dendritic cells (FDCs) recruits potent B cells that initially have moderate affinity to the antigen. It reflects the condition that cells should have some initial threshold value of affinity in order to enter the GC.

We considered three zonal compartments: outside of the GC, the DZ and the LZ. Initially, B-cells are outside of the germinal center and are recruited to the germinal center depending on the affinity of their antibodies to the antigen at an average rate of $1/20 \text{ d}^{-1}$.

Upon entering the germinal center, antibodies undergo monoclonal expansion in the dark zone. The replication rate of the B cells in this phase is four times greater than the normal B cell replication rate, which is taken to be 1/day: typical of eukaryotic division rate. After three days of the monoclonal expansion phase, B cells go through somatic hypermutation (SHM), when the mutation rates are around 0.5 to 0.7 mutations per replication. We modeled this behavior by setting B cell switching rate into the SHM phase to an average of once per three days. SHM lasts for an average of 6 h, reflecting the experimental centroblast differentiation rate. We assume that processes in the DZ are affinity-independent because centroblasts do not express cell-surface receptors to Ag proteins. After finishing SHM, centroblasts in the dark zone differentiate into centrocytes and migrate to the LZ to be selected to form memory or plasma B cells or undergo apoptosis. We modeled these phenomena by setting the B cell death rate in LZ (apoptosis), d_0 in equation (6) below, to six times per day and the B cell differentiation rate (plasma cell formation or GC exit rate) to once per 20 days. To represent recycling, B cells in the LZ have a probability of returning to the DZ at the rate that is 5/8 of the rate of centroblast differentiation, approximating experimental observations [3]. We assume centroblast differentiation is not affinity dependent (because centroblasts do not express B cell receptors of stable structure), so our rate of recycling is independent of P_{int} . Upon re-entering the DZ, B cells immediately rejoin the pool of cells undergoing SHM at the high replication rate characteristic of the DZ (~4 times/day). There is no limitation on the number of times a single B cell may recycle between zones since the processes of centroblast differentiation and centrocyte recycling are both stochastic. This recycling scheme is comparable to that of Oprea and Perelson, in which B cells in the LZ are faced with a constant probability of DZ reentry, with the maximal affinity being achieved by a recycling probability between 0.6–0.7 [21]. Low

affinity B cells in the LZ are removed by an apoptotic pathway. The death rate in the LZ is given in equation (6):

$$d^{Ab} = d_0^{Ab} (1 - P_{int}(Ab, Ag)), \quad (6)$$

where d_0^{Ab} is the base death rate of B cells, set to 6/day, and $P_{int}(Ab, Ag)$ is the affinity of the antibody in a B cell to the antigen. An affinity-dependent apoptosis rate has been established previously [15].

We also incorporated another selection mechanism for B cells in the LZ by assuming that the B cells with high affinities for the antigen replicate fastest [42]. Whereas LZ centrocytes may be chosen to replicate based on the affinity of their surface receptors, a dynamic birth rate is chosen for DZ centroblasts. Without affinity-dependence in the DZ, we use an exponential function to capture the exponential population growth during monoclonal expansion [1]. Thus, we can describe the replication rate of B cells within the GC as follows:

$$b^{Ab} = \begin{cases} b_0^{DZ} e^{-t/\tau} \left[1 - \frac{N_{DZ}}{N_{max}^{DZ}} \right] \\ b_0^{LZ} [1 + \alpha P_{int}(Ab, Ag)] \end{cases} \quad (7)$$

where b_0^{DZ} and b_0^{LZ} are base B cell replication rates in the DZ and LZ and both base rates are set so that the replication rate of B cells in the DZ is greater than in the LZ. N^{DZ} is the DZ population, N_{max}^{DZ} (set at 3000) is the maximum DZ population [43], and $P_{int}(Ab, Ag)$ is the Boltzmann probability of a specific interaction of antibody protein, Ab, with an antigen, Ag, respectively, α is the B-cell activation factor (set at 5.0), and T is the environmental temperature scaled by the Boltzmann constant. There is no standard experimental estimate of the maximum DZ size at 5 days post-immunization, so a theoretical estimate was used [43]. These Boltzmann probabilities serve as a measure of the strength of interactions between different entities in the model. In the context of the immune system, an increase in the antibody replication rate with increasing antigen binding strength is a conjectured phenomenon of B cell activation by an antigen [32]. We always set $b_0^{DZ} = 4b_0^{LZ}$, representing the experimental 6 h DZ replication rate and the 24 h eukaryotic cell replication, which we hypothesize applies to the LZ [41, 43, 44].

References

1. Allen CDC, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity*. 2007; 27:190–202. [PubMed: 17723214]
2. Ulrich HD, et al. The interplay between binding energy and catalysis in the evolution of a catalytic antibody. *Nature*. 1997; 389:271–5. [PubMed: 9305839]
3. Victora GD, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell*. 2010; 143:592–605. [PubMed: 21074050]
4. Garin A, et al. Toll-like receptor 4 signaling by follicular dendritic cells is pivotal for germinal center onset and affinity maturation. *Immunity*. 2010; 33:84–95. [PubMed: 20643339]
5. Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intracloal generation of antibody mutants in germinal centres. *Nature*. 1991; 354:389–92. Engl. transl. [PubMed: 1956400]

6. Berek C, Berger A, Apel M. Maturation of the immune response in germinal centers. *Cell*. 1991; 67:1121–9. [PubMed: 1760840]
7. Batista FD, Neuberger MS. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity*. 1998; 8:751–9. Engl. transl. [PubMed: 9655489]
8. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity*. 2007; 27:190–202. Engl. transl. [PubMed: 17723214]
9. Allen CD, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. *Science*. 2007; 315:528–31. [PubMed: 17185562]
10. Allen CDC, et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat. Immunol.* 2004; 5:943–52. [PubMed: 15300245]
11. Figge MT, et al. Deriving a germinal center lymphocyte migration model from two-photon data. *J. Exp. Med.* 2008; 205:3019–29. [PubMed: 19047437]
12. Beltman JB, Allen CD, Cyster JG, de Boer RJ. B cells within germinal centers migrate preferentially from dark to light zone. *Proc. Natl Acad. Sci. USA*. 2011; 108:8755–60. Engl. transl. [PubMed: 21555569]
13. Kepler TB, Perelson AS. Cyclic re-entry of germinal center B cells and the efficiency of affinity maturation. *Immunol. Today*. 1993; 14:412–5. [PubMed: 8397781]
14. Moreira JS, Faro J. Re-evaluating the recycling hypothesis in the germinal centre. *Immunol. Cell Biol.* 2006; 84:404–10. Engl. transl. [PubMed: 16834575]
15. Jacob J, Kelsoe G. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl: II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. *J. Exp. Med.* 1992; 176:679–87. Engl. transl. [PubMed: 1512536]
16. Hauser AE, et al. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity*. 2007; 26:655–67. Engl. transl. [PubMed: 17509908]
17. Heo M, Zeldovich K, Shakhnovich E. Diversity against adversity: how adaptive immune system evolves potent antibodies. *J. Stat. Phys.* 2011; 144:241–67.
18. Meyer-Hermann M. Does recycling in germinal centres exist. *Immunol. Cell Biol.* 2002; 80:30–5. [PubMed: 11869360]
19. Klien U, et al. Somatic hypermutation in normal and transformed human B cells. *Immunol. Rev.* 1998; 162:261–80. [PubMed: 9602370]
20. Kepler TB, Perelson AS. Somatic hypermutation in B cells: an optimal control treatment. *J. Theor. Biol.* 1993; 164:37–64. [PubMed: 8264243]
21. Oprea M, Perelson AS. Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J. Immunol.* 1997; 158:5155–62. [PubMed: 9164931]
22. Kepler TB, Perelson AS. Modeling and optimization of populations subject to time-dependent mutation. *Proc. Natl Acad. Sci. USA*. 1995; 92:8219–23. Engl. transl. [PubMed: 7667271]
23. Miyazawa S, Jernigan RL. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* 1996; 256:623–44. [PubMed: 8604144]
24. Sali A, Shakhnovich E, Karplus M. How does a protein fold. *Nature*. 1994; 369:248–51. Engl. transl. [PubMed: 7710478]
25. Abkevich VI, Gutin AM, Shakhnovich EI. Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry*. 1994; 33:10026–36. [PubMed: 8060971]
26. Camacho CJ, Thirumalai D. Kinetics and thermodynamics of folding in model proteins. *Proc. Natl Acad. Sci. USA*. 1993; 90:6369–72. Engl. transl. [PubMed: 8327519]
27. Shakhnovich EI. Proteins with selected sequences fold into unique native conformation. *Phys. Rev. Lett.* 1994; 72:3907–10. Engl. transl. [PubMed: 10056327]
28. Deeds, EJ.; Ashenberg, O.; Gerardin, J.; Shakhnovich, EI. *Proc. Natl Acad. Sci. USA*. Vol. 104. 14952–7: 2007. Robust protein protein interactions in crowded cellular environments. Engl. transl.
29. Heo M, Maslov S, Shakhnovich E. Topology of protein interaction network shapes protein abundances and strengths of their functional and nonspecific interactions. *Proc. Natl Acad. Sci. USA*. 2011; 108:4258–63. Engl. transl. [PubMed: 21368118]

30. Zhang Y, Arakaki AK, Skolnick J. TASSER: an automated method for the prediction of protein tertiary structures in CASP6. *Proteins*. 2005; 61:91–8. Engl. transl. [PubMed: 16187349]
31. Finkelstein AV, Gutin AM, Badretdinov A. Boltzmann-like statistics of protein architectures. Origins and consequences. *Subcell Biochem*. 1995; 24:1–26. Engl. transl. [PubMed: 7900172]
32. Oprea M, van Nimwegen E, Perelson AS. Dynamics of one-pass germinal center models: implications for affinity maturation. *Bull. Math. Biol.* 2000; 62:121–53. [PubMed: 10824424]
33. Victora GD, Nussenzweig MC. Germinal centers. *Annu. Rev. Immunol.* 2012; 30:429–57. Engl. transl. [PubMed: 22224772]
34. Küppers R, Hansmann ML, Rajewsky K. Clonality and germinal centre B-cell derivation of Hodgkin/Reed–Sternberg cells in Hodgkin's disease. *Ann. Oncol.* 1998; 9:S17–20. [PubMed: 9926232]
35. Marafioti T, et al. Origin of nodular lymphocyte-predominant Hodgkin's disease from a clonal expansion of highly mutated germinal-center B cells. *New Engl. J. Med.* 1997; 337:453–8. [PubMed: 9250847]
36. Ohno T, et al. Clonality in nodular lymphocyte-predominant Hodgkin's disease. *New Engl. J. Med.* 1997; 337:459–66. [PubMed: 9250848]
37. Marafioti T, et al. Classical Hodgkin's disease and follicular lymphoma originating from the same germinal center B cell. *J. Clin. Oncol.* 1999; 17:3804–9. [PubMed: 10577852]
38. Freedman A, et al. Follicular non-Hodgkin's lymphoma cell adhesion to normal germinal centers and neoplastic follicles involves very late antigen-4 and vascular cell adhesion molecule-1. *Blood*. 1992; 79:206–12. [PubMed: 1370204]
39. Sangle N. Lymphoma-B cell neoplasms. PathologyOutlines.com. 2011 PathologyOutlines.com
40. Shakhnovich EI, Gutin A. Enumeration of all compact conformations of copolymers with random sequence links. *J. Chem. Phys.* 1990; 93:5967–71.
41. Alberts, B. *Molecular Biology of the Cell*. 3rd edn. Garland; New York: 1994. p. 864-66.
42. MacLennan ICM. Germinal centers. *Annu. Rev. Immunol.* 1994; 12:117–39. [PubMed: 8011279]
43. Meyer-Hermann M. A mathematical model for the germinal center morphology and affinity maturation. *J. Theor. Biol.* 2002; 216:273–300. [PubMed: 12183119]
44. DeFranco, AL. *Immunity: The Immune Response in Infectious and Inflammatory Disease*. Oxford University Press; Oxford: 2007. p. 350

Figure 1.

The phase diagrams of models. The color code corresponds to P_{int} . (a) One-shot model: optimal affinity maturation occurs in the range of realistic centrocyte differentiation rates (~6 h) and B cell mutation rates (0.6–0.8). (b) Recycling with P_{nat} constraints: the affinity gains caused by antigenic selection in the LZ allow a small number of cells to reenter the DZ. This small population is then subject to destabilizing SHM, leading to apoptosis if $P_{int} < 0.6$. No affinity maturation is observed. (c) Recycling without P_{nat} constraints models: hypermutation lowers P_{int} drastically, until a small population of cells enters the LZ and P_{int} grows to about 0.3. The P_{int} does not grow beyond this GC entry threshold since most cells are pooled in the DZ, and those that enter the LZ may recycle back into the DZ.

Figure 2.

Population dynamics in DZ and LZ and the affinity maturation. Population dynamics of B cells are plotted with respect to time and mutation rate, $m = 0.7$ and centroblast differentiation, $r_{LZ} = 0.03$. The red and green lines are DZ population and LZ population, respectively; the dotted blue line is P_{int} . (a) One-shot: the DZ population reaches a peak of ~3000 B cells by five days post-immunization, when centroblasts differentiate into centrocytes in the LZ. At this time, there is pronounced LZ population growth to ~10 000 cells, which matches experimental reports [18]. (b) Recycling with P_{nat} constraints: the dark zone persists for 15 days post-immunization since LZ cells are moving back to the DZ and replicating there at high rates. (c) Recycling without P_{nat} constraints models: unstable Ab proteins whose binding surfaces were mutated in SHM survive the transition to the LZ (since there is no stability requirement), but their low affinity interaction with antigenic proteins triggers the apoptotic pathway in the LZ.

Figure 3.

R/S ratios for all models. All data are averaged with 100 ensembles obtained by different simulations. The RS ratio of CDR (red) and FR (green) of antibodies are plotted with respect to days post-immunization at $m = 0.7$ and $r_{LZ} = 0.03$. In one-shot model A, positive selection is seen only in the CDR region at the time of centroblast differentiation. In recycling model B with $P_{nat} > 0.6$, there is no positive selection in either region since repeated SHM ‘erases’ affinity gains made in the LZ. In recycling model C, CDR R/S ratios decrease once cells begin SHM in the DZ and increase during the first bout of antigenic selection in the LZ. Interestingly, ‘positive’ selection is seen in both the CDR and FR once B cells enter the LZ, although empirical R/S ratios in the FR are < 3 [19]. Judgment of selection pressures by R/S ratios in this situation is difficult due to the volatile structural conformation of the rapidly mutating cells pooling in the DZ.

Figure 4.

Mutation count versus time for the one-shot model. The red line is the number of mutations accumulated in Abs in the GCR (left axis). The dotted green line is P_{int} . Centroblasts accumulate mutations in SHM and move into the LZ. When the DZ disappears around day seven post-immunization, the average mutation count per B cell increases sharply, as shown here. The diversified high-affinity B cells survive in the LZ while low-affinity B cells without beneficial mutations die off.

Figure 5. Interaction of two lattice proteins. Maximally compact $3 \times 3 \times 3$ cubic lattice proteins interact. Each cube has six binding surfaces and four-fold rotational symmetry of each binding mode. Each pair of interacting faces is subject to nine amino acid–amino acid binding energies. The lowest energy binding mode represents the correct Ab–Ag binding conformation [17].

Table 1

We consider three different models depending on their population dynamics between dark zone and light zone (recycling or one-shot) and the protein structural stability (P_{nat}) constraints. The biologically observed replication rate in LZ is 1 /day [41].

Model	Recycling	P_{nat} constraints
A	No	0.6
B	Yes	0.6
C	Yes	0.0