The challenges of modelling antibody repertoire dynamics in HIV infection

Shishi Luo\textsuperscript{1,2} and Alan S. Perelson\textsuperscript{3}

\textsuperscript{1}Department of Electrical Engineering and Computer Science, and \textsuperscript{2}Department of Statistics, UC Berkeley, Berkeley, CA 94110, USA
\textsuperscript{3}Los Alamos National Laboratory, Theoretical Biology and Biophysics Group, Los Alamos, NM 87545, USA

Antibody affinity maturation by somatic hypermutation of B-cell immunoglobulin variable region genes has been studied for decades in various model systems using well-defined antigens. While much is known about the molecular details of the process, our understanding of the selective forces that generate affinity maturation are less well developed, particularly in the case of a co-evolving pathogen such as HIV. Despite this gap in understanding, high-throughput antibody sequence data are increasingly being collected to investigate the evolutionary trajectories of antibody lineages in HIV-infected individuals. Here, we review what is known in controlled experimental systems about the mechanisms underlying antibody selection and compare this to the observed temporal patterns of antibody evolution in HIV infection. We describe how our current understanding of antibody selection mechanisms leaves questions about antibody dynamics in HIV infection unanswered. Without a mechanistic understanding of antibody selection in the context of a co-evolving viral population, modelling and analysis of antibody sequences in HIV-infected individuals will be limited in their interpretation and predictive ability.

1. Introduction

The evolution of the antibody repertoire in response to HIV infection, where the viral population co-evolves with the antibody response, is not well understood. The use of high-throughput antibody sequence data in characterizing antibody repertoire dynamics is also in its infancy, with methods still being standardized and basic questions investigated [1]. Notwithstanding, antibody sequence data generated from HIV patients are being used to study the adaptive immune response to HIV infection. In particular, spurred by the discovery of potent antibodies that neutralize a broad panel of HIV strains, large antibody repertoire datasets are being collected from HIV-infected individuals (e.g. [2]). One goal of this work is to identify antibody affinity maturation pathways that might be mimicked by vaccine protocols [3].

With the collection of data in these areas racing ahead, it is important to understand the complexities of antibody dynamics where the viral population evolves throughout the course of infection, as in the case of HIV (shown schematically in figure 1). To illustrate these complexities, we first review the mechanisms that are known to underlie antibody selection in the presence of a single antigen [5,6]. We then review empirically observed temporal dynamics of the antibody response in HIV-infected patients. We argue that the mechanisms determined from single-antigen studies are insufficient to guide the analysis and modelling of antibody evolution during HIV infection. In particular, understanding antibody–virus co-evolutionary dynamics requires knowledge of the B-cell response when an antigen changes over time and of B-cell competition in the presence of multiple co-circulating antigens. Neither of these processes can be deduced from single-antigen studies. In the absence of experimental evidence, any analysis or modelling of the antibody response to HIV infection will require making assumptions about these processes. The analysis and interpretation of antibody sequence data will therefore be limited by the validity of these assumptions.
2. Antibody selection in germinal centres

Selection of B cells, and their expressed antibodies, chiefly occurs in sites called germinal centres (reviewed in [5]). Germinal centres are dynamic structures that form approximately one week after infection when antigen-activated B cells migrate to B-cell follicles in lymph nodes and the spleen. Once inside germinal centres, B cells undergo proliferation and mutation in a region of the germinal centre, known as the dark zone, and selection in another region, called the light zone. Replication has been measured to occur every 6–14 h [5] and somatic mutations accumulate with each replication. In mice infected with influenza, this mutation occurs at a rate of $10^{-3}$ per base per generation [7].

The mechanisms of B-cell selection in germinal centres are becoming increasingly understood in murine model systems in which the immune system is exposed to a single known antigen. Recent work by Gitlin et al. [6] has experimentally demonstrated that the length of time a B cell spends replicating and mutating in the dark zone of the germinal centre is correlated with the affinity between its expressed antibody and the invading antigen. Affinity-based selection of B cells occurs in the light zone of the germinal centre, where B cells collect antigen displayed on follicular dendritic cells. B cells expressing receptors with higher affinity to the antigens will collect more antigen from follicular dendritic cells than B cells with low-affinity receptors. The antigen collected by B cells are internalized and then presented as peptides to follicular helper T cells (T<sub>FH</sub> cells). B cells that present more peptides to T<sub>FH</sub> cells survive longer and spend more time in the dark zone and undergo more replication and mutation.

These experimental studies of germinal centre reactions have allowed explicit computational modelling of antibody selection in the case where there is a single type of antigen [8,9]. However, equally detailed modelling of antibody selection in the case of a co-evolving viral population is currently difficult. Indeed, recent models, such as those by Meyer-Hermann et al. [8] and Wang et al. [9] do not include viral evolution, although the latter model uses multiple non-evolving viral strains to simulate vaccination protocols. The difficulty of constructing models of co-evolving virus and antibody populations lies in the lack of experimental understanding about germinal centre reactions during chronic infection. Critical quantities such as the frequency with which B cells cycle through germinal centres over the course of infection, how long it takes mutant viral strains in other parts of the body to appear in germinal centres, and what fraction of B cells leaving germinal centres are plasma cells versus memory cells, are unknown. In addition to these unknown quantities in chronic infection, HIV poses a further complication in that T<sub>FH</sub> cells are CD4<sup>+</sup> cells and hence targets of HIV infection. Because CD8<sup>+</sup> T cells are generally excluded from B-cell follicles, infected CD4<sup>+</sup> T<sub>FH</sub> are substantially shielded from CD8-mediated suppression [10]. The effects of HIV infection on T<sub>FH</sub> cell function in germinal centre reactions are currently being explored, but impairments in T<sub>FH</sub> function have been suggested to contribute to the relative inefficiency of broadly neutralizing antibody generation in HIV-infected individuals [11,12].

In the absence of this information, one study that models both antibody and viral evolutionary dynamics avoids explicitly modelling germinal centre reactions altogether [13]. Instead, virus–antibody co-evolution is modelled phenomenologically, with antibodies and viral epitopes modelled as character strings that replicate and mutate. Antibody affinity is evaluated by a string-matching procedure. Both antibodies with relatively high affinity and viral strains escaping most antibodies have high relative fitness and tend to persist and evolve further.

It would be preferable if the analysis of HIV antibody sequence data was based on more mechanistically explicit models of antibody selection with a co-evolving viral population. To this end, we describe observed temporal patterns of antibody evolution in HIV infection with which such models would need to be consistent. These temporal patterns further highlight properties of antibody selection in HIV infection that current experimental results cannot explain.

3. Initial stages of HIV infection

Early in HIV infection the viral population expands and then settles to an approximate steady-state size, called the set-point viral load (figure 1). The virus in the acute stage of an infection caused by sexual transmission is largely homogeneous and the mutations that occur in the viral population in these early stages appear to be random [14,15]. Indeed, there is little selection pressure exerted on the virus, as the earliest HIV-specific antibodies are not detected until around two to three weeks after infection [16], or 13 days after detectable plasma viraemia [17]. These antibodies predominantly target epitopes on the HIV envelope gp41, which is not exposed on functional HIV envelope and are therefore non-neutralizing [16,17].

It is puzzling that in the same time that it takes the adaptive immune system to mount a neutralizing antibody response to a typical infection [18], it only mounts a non-neutralizing response against HIV. One possible reason is that the contact region between gp41 and gp120 on HIV-1 envelope is immunogenic and often exposed on disassembled gp41–gp120 complexes that are abundant throughout HIV infection [16]. Viral particles that display this immunogenic region may therefore also be abundant.
on follicular dendritic cells, leading to selection for B cells that recognize this non-functional region.

Another possibility is that gp41-specific antibodies arise from memory B-cell populations primed by gut bacteria. Trama et al. [19] showed that gp41 antibodies cross-react with gut commensal bacteria and found antibodies with similar characteristics to gp41-specific antibodies in HIV-negative individuals. They suggested that since memory B cells respond more rapidly than naive B cells, the non-neutralizing antibody response to HIV could be due to the clonal expansion of B cells previously primed by gut bacteria, bacteria that coincidentally share features with gp41.

Thus, one challenge in studying antibody repertoire dynamics in HIV infection is that the part of the antibody repertoire that is initially elicited, or that responds most rapidly in early infection, is non-neutralizing. Because determination of antibody structure from sequence data is still a difficult and low-throughput process, there is no straightforward way to determine from sequence data alone what epitopes these non-neutralizing antibodies are targeting. While these antibodies are being produced, what is happening to antibodies that target neutralizing epitopes? Whether they are present but at undetectable levels, or not present at all at this early stage is an open question.

4. Antibody neutralization of the infecting strain and rapid viral escape

Neutralizing antibodies specific to the transmitted HIV strain, called autologous antibodies, eventually arise, but are typically not detected until months after infection, with some patients taking more than a year [2,16,20,21]. The epitopes targeted by these neutralizing antibodies tend to be in variable regions, e.g. the V3 loop, which, owing to being prominently exposed on the HIV envelope, may be more immunogenic than other neutralizing targets [16,22].

Autologous neutralizing ability is demonstrated in vitro by assaying virus sampled from a patient early in infection against plasma collected later in infection (e.g. [21,23]). However, when the same assay is carried out for the viral population contemporaneous with this same neutralizing plasma, the contemporaneous virus population is not inhibited. In other words, by the time a neutralizing antibody response against a virus strain is detectable in the blood, the viral population has already mutated to escape it. Figure 2 shows the lag in response time of neutralizing plasma against the viral population from four patients [21]. The lag, which varies over time and among patients, ranges from about six to 16 months.

Because viral escape occurs, the initial neutralizing antibodies do not clear the infection. In fact, whether these antibodies play any role in controlling the viral population is unclear [24]. This is slightly counterintuitive because for immune escape variants to replace the infecting strain, the viral population would need to have been under selective pressure. Yet there are no obvious clinical manifestations of this selection, such as a notable decrease in the viral population. This could be due to plasma and viral samples being taken too infrequently to detect such changes, or that the effects of the antibody response manifest themselves in tissues rather than peripheral blood.

In any case, the above suggests that while it takes substantial time for the antibody population to mount a neutralizing response, the viral population can rapidly escape neutralizing antibodies. This gives the impression (figure 2) that the antibody response is consistently targeting past viral strains, not current ones. Therefore, another challenge to studying antibody repertoire dynamics in HIV infection is a lack of knowledge of how a change in the viral strain is reflected in the antibody selection process. A recent study demonstrated how follicular dendritic cells retain intact antigen for long periods of time [25]. It is possible, then, that viral particles displayed on follicular dendritic cells during HIV infection are older relative to the viral strains abundant in the rest of the body, forcing affinity maturation towards older strains. However, until experiments on germinal centre reactions with temporally changing antigens are conducted, it is unclear whether to model the delay in a neutralizing response as the result of restrictive properties of follicular dendritic cells or as the result of inherently slower rates of mutation and selection during antibody affinity maturation.

5. Co-evolution in later stages of infection

Time-series antibody and viral sequence data for later stages of infection are most detailed in patients who develop broadly neutralizing antibodies. The pattern common to the cases studied to date is a repeating cycle of immune escape by the virus and increased antibody neutralization breadth, with the antibody response lagging behind the viral population.

Moore et al. [26] studied two HIV-1 patients who developed broad antibodies, where breadth in both patients was based on binding to a glycan on the envelope protein. The glycan target was not on the transmitted strain, and glycan shielding is a known HIV immune escape mechanism [20]. Thus, the appearance of the glycan in the viral population at around six months after infection was hypothesized to be an escape mutation in response to an initial autologous antibody response. At around 12–15 months post infection, antibodies that target the glycan were detected in the individual’s sera. This suggests that selection for antibodies against the escape variant occurred, albeit with a lag of more than six months. Finally, at 2 years post infection, the glycan disappeared from the viral population, suggesting that the virus again escaped the antibody response.

A similar chain of events was observed by Wibmer et al. [27], where three waves of increasingly broad neutralizing antibody responses occurred in a single individual. The first wave of antibodies emerged at around 16 months and targeted the V2 loop, another region known to be variable and immunogenic [22]. Escape from this first wave occurred via deletion of a glycan on which the first-wave antibody depended. This deletion simultaneously exposed an epitope on the CD4 binding site, leading to a second wave of broadly neutralizing antibodies at around 28 months that relied on CD4 binding site neutralization. Immune escape and a subsequent third wave of broadly neutralizing antibodies followed, though the location of the escape mutations and the targets of the antibodies were not identified in the study.

Perhaps the most detailed study to date, with resolution at the level of antibody clonal lineages, comes from [28] in the case of patient CH505 [2]. Gao and colleagues identified a
specific clonal lineage of B cells that selected for escape variants of the founder HIV strain. These variants contained mutations that enhanced reactivity towards a different B-cell lineage, one that ultimately led to a broadly neutralizing antibody. Their results imply that rather than a purely serial process of neutralization and escape, escape from one antibody lineage led to selection of a concurrent antibody lineage that developed broad neutralizing breadth.

The sequential selection scenarios in the first two studies again raise the question of how affinity maturation occurs in the context of temporally changing antigens. The results of [28] raise a further complication: B cells specific to different viral strains compete with each other in HIV infection. This question has been investigated in HIV with a mouse model by Forsell et al. [29]. They compared antibody responses in mice vaccinated with a wild-type HIV strain with mice vaccinated with a modified strain. The modified strain had a mutation that caused the V3 loop, a known immunodominant region, to be masked, i.e. glycosylated. Forsell and colleagues were interested in knowing whether the B-cell response to other epitopes would be increased as a result of masking the more immunogenic V3 loop. They found that while the antibody response to the V3 loop was indeed suppressed, there was no evidence of a concomitant increase in the antibody response to other epitopes. The work of Forsell et al. [29] thus suggests a specific mechanism of B-cell competition: one where the clonal expansion of a B-cell lineage is determined by a target epitope’s absolute immunogenicity, rather than its immunogenicity relative to other epitopes. This result also raises questions regarding antibody competition early in infection: if the targets of the initial non-neutralizing antibodies are highly immunogenic, how are neutralizing antibodies to less immunogenic epitopes competing with them? Also, there is the issue of competition between anti-Env B cells and B cells specific for other HIV proteins that remains to be quantitatively addressed.

6. Conclusion

High-resolution antibody sequence data are increasingly being collected in studies of HIV and antibody co-evolution. The current goal of inferring dynamics specific to each individual from these data is an important first step in understanding the different paths that evolution might take in generating antibody repertoires towards HIV. However, to identify general characteristics of antibody–virus co-evolution during HIV infection, we will need to construct models that incorporate selective mechanisms that drive antibody repertoire dynamics. We have outlined here the gap between what is known experimentally about antibody selection in single-strain systems and what has been observed empirically for HIV infections. In particular, antibody selection in the presence of an evolving viral population and in the presence of multiple distinct viral strains is not well understood. Until the mechanisms of selection in these situations are determined experimentally, interpretation and prediction based on high-throughput antibody sequence data in HIV infection will be limited by the modelling assumptions made about these mechanisms.

Authors’ contributions. S.L. and A.S.P. wrote the article.

Competing interests. We declare we have no competing interests.

Funding. This work was performed under the auspices of the US Department of Energy under contract DE-AC52-06NA25396 and supported by the Center of Nonlinear Studies at Los Alamos National Laboratory, NIH grant nos R01-AI028433, R01-OD011095 and the Center for HIV AIDS Vaccine Immunology-Immunogen Design (CHAVI-ID) grant UMI-AI100645.