Antibody-guided vaccine design: identification of protective epitopes
Antonio Lanzavecchia\textsuperscript{1,2}, Alexander Frühwirth\textsuperscript{1}, Laurent Perez\textsuperscript{1} and Davide Corti\textsuperscript{3}

In the last decade, progress in the analysis of the human immune response and in the isolation of human monoclonal antibodies have provided an innovative approach to the identification of protective antigens which are the basis for the design of vaccines capable of eliciting effective B-cell immunity. In this review we illustrate, with relevant examples, the power of this approach that can rapidly lead to the identification of protective antigens in complex pathogens, such as human cytomegalovirus and \textit{Plasmodium falciparum}, and of conserved sites in highly variable antigens, such as influenza hemagglutinin and HIV-1 Env. We will also discuss how the genealogical analysis of antigen-stimulated B cell clones provides the basis to delineate the best suitable prime-boost vaccination strategy for the induction of broadly neutralizing antibodies.

Addresses
\textsuperscript{1} Institute for Research in Biomedicine, Via Vincenzo Vela 6, 6500 Bellinzona, Switzerland
\textsuperscript{2} Institute of Microbiology, ETH Zürich, Vladimir-Prelog-Weg 1, 8093 Zürich, Switzerland
\textsuperscript{3} Humabs BioMed, Via Mirasole 1, 6500 Bellinzona, Switzerland

Corresponding author: Lanzavecchia, Antonio (lanzavecchia@irb.usi.ch)

Introduction
When our immune system is challenged by an infectious agent, a polyclonal antibody response is generated against multiple protein and non-protein antigens. The extent of the response reflects the immunogenicity of the individual components, which is determined by multiple factors such as their abundance, their complexity and their capacity to bind to cellular receptor and trigger innate immunity, not to mention the influence of pre-existing immunity. In general, only a fraction of the antibodies produced exerts protective activity by binding to molecules that are required for invasion or virulence or by eliciting effector mechanisms. For instance, the largest fraction of antibodies produced in response to a virus infection is directed against internal or surface proteins in a denatured or post-fusion conformation and is therefore devoid of neutralizing activity [1,2,3]. Thus, the immunogenicity of the most abundant proteins offers to the pathogen the opportunity to limit the most effective response by a mechanism of antigenic competition.

In this review we discuss how recent advances in the characterization of the neutralizing antibody response to human pathogens, combined with advanced protein engineering methods, have provided a new way to solve the problem of antigenic competition and have led to the production of vaccine candidates that elicit antibody responses of high magnitude and specific activity.

Analytic vaccinology: the cases of HCMV and malaria
Human cytomegalovirus (HCMV) is a herpes virus that establishes a lifelong infection in healthy individuals, but causes serious pathology in the fetus and in immunosuppressed patients and has been associated with immune senescence and atherosclerosis [4]. HCMV uses multiple glycoprotein complexes for binding and fusion to host cells and has a broad cell tropism, being able to infect fibroblasts as well as epithelial, endothelial and myeloid cells [5]. The fusion protein gB has been considered the most obvious vaccine candidate but clinical trials with recombinant gB (in the post-fusion conformation) has shown limited efficacy [6]. To identify the most potent HCMV vaccine, an analytic vaccinology approach was used to isolate, from memory B cells of naturally infected donors, a large panel of monoclonal antibodies selected for their capacity to neutralize HCMV infection of multiple cell types [7] (Figure 1). This approach led to the identification of a new class of antibodies that were 1000 fold more potent that antibodies to gB in neutralizing HCMV infection of epithelial, endothelial and myeloid cells. These antibodies were mapped to nine distinct sites on the gH/gL/UL128-131A complex, a pentameric complex that was previously found to be required for infection of those cell types [8]. The identification of the pentamer as the target of the most potent antibodies was subsequently confirmed by several studies [9,10,11]. As an example, a soluble pentamer produced by stably a transfected CHO cell line elicited in mice antibody titers that persisted to high levels over time.
and that were 300-1000 fold higher than those found in individuals that recovered from primary HCMV infection [10**]. Importantly, the antibodies elicited by the pentamer vaccine prevented cell-to-cell spread and viral dissemination from endothelial cells to leukocytes and neutralized infection of both epithelial cells and fibroblasts due to the production of antibodies to the gH glycoprotein, which is required for fibroblasts infection.

The target-agnostic approach can be particularly useful to identify targets in complex pathogens such as bacteria and parasites. An interesting example regards the identification of variant surface antigens (VSAs) which are present on the surface of *Plasmodium falciparum* (*Pf*)-infected erythrocytes and mediate adhesion to endothelia, leading to pathology. The VSAs are encoded by more than 200 genes that are polymorphic and clonally
expressed, thus providing the pathogen with a powerful mechanism of escape from the antibody response. Human monoclonal antibodies isolated from multiparous women have been used to identify VAR2CSA as the primary target of antibodies that protect from placental malaria [12]. In a more recent study, Tan et al. isolated several antibodies that broadly react with erythrocytes infected with different Pf isolates and identified the target antigens as distinct RIFINs [13**]. Surprisingly, these antibodies acquired their broad reactivity through a novel mechanism of insertion of a large DNA fragment between the V and DJ segments. The insert originates from chromosome 19 and encodes the extracellular domain of the collagen-binding inhibitory receptor LAIR-1, which is both necessary and sufficient for binding to RIFINs. Importantly, the LAIR-1 domain carries somatic mutations that abolish binding to collagen and increase binding to infected erythrocytes. These findings illustrate, with a biologically relevant example, a novel mechanism of antibody diversification and demonstrate the existence of conserved epitopes as candidates for the development of a malaria vaccine [14].

**Structure-based vaccinology: the case of HRSV**

Human respiratory syncytial virus (HRSV) is the most prominent viral agent of pediatric respiratory infections. Currently, the only treatment available is Palivizumab [15], a humanized monoclonal antibody that binds to an epitope that is conserved in the pre-fusion and post-fusion conformation of the HRSV F protein [16]. On the basis of the properties of Palivizumab, the initial efforts towards an HRSV vaccine were focused on the use of the post-fusion F-protein that was engineered to increase its solubility, resulting into highly stable immunogen forms, that were tested either as soluble antigens or displayed onto virus-like particles (VLPs) [17,18].

Two recent studies used the structural information on the Palivizumab linear epitope, a 24 amino acid helix–loop–helix structure, to develop epitope-focused vaccines. Correia et al. used computational protein design to generate stable protein scaffolds that accurately mimic the Palivizumab epitope and, when chemically linked to VLPs, induce HRSV-neutralizing antibodies in macaques [19*]. In a subsequent study, Milich and coworkers used an empirical approach by directly inserting the Palivizumab epitope in VLPs composed of woodchuck hepatitis virus core (WHcAg) protein, which were then selected with the antibody and found to be immunogenic in rodents [20*]. While these studies provide for the first time a proof-of-principle for epitope-focused vaccinology (Figure 1), it is not evident what could be the advantage of limiting the response to a single epitope, given the presence of several conserved epitopes in the HRSV F protein.

The analysis of the human antibody response to HRSV [21,22] and the isolation of human neutralizing monoclonal antibodies [23,24] showed that the large majority of HRSV neutralizing antibodies elicited by natural infection are specific for the pre-fusion F protein conformation and, unlike Palivizumab, do not cross-react with the post-fusion conformation. On the basis of these observations and on the crystal structure of the pre-fusion F protein in complex with a neutralizing antibody [25], Kwong and coworkers used structure-based design to produce a stabilized HRSV F protein through the introduction of cysteine residues and the filling of hydrophobic cavities [26**,27]. This protein (dubbed DS-Cav1) maintained binding to potent neutralizing antibodies and was able to elicit in mice and macaques levels of HRSV-neutralizing antibodies that exceeded the protective threshold. These results provide a clear example of structure-driven vaccinology to generate stable immunogens capable of inducing a polyclonal response to multiple neutralizing epitopes of viral fusion proteins, overcoming the problem of their intrinsic instability (Figure 1).

**Epitope-focused vaccinology: the cases of influenza A**

Influenza A and HIV-1 are the prototypes of viruses that evade the antibody response by continuously mutating surface glycoproteins. However, certain sites are relatively conserved since they are required for infectivity and fusion and can be therefore exploited to design vaccines capable of inducing broad protection. In the last decade the isolation of broadly neutralizing antibodies has provided a powerful platform to identify such epitopes, thus speeding up vaccine design efforts [28,29].

The sialic acid binding pocket of influenza hemagglutinin (HA) is a conserved site that can be targeted by antibodies, such as CH65, that carry at the tip of HCDR3 a distinct motif that mimics sialic acid [30**]. Importantly, these antibodies arise from diverse germline origins and affinity maturation pathways, suggesting that such an antibody response could be rapidly generated by appropriate immunogens. However, the breadth of these antibodies is somewhat limited due to the presence of highly variable residues that flank the receptor binding site.

The HA stem is a more conserved region which has been shown to be recognized by different families of broadly neutralizing antibodies that recognize multiple subtypes in group 1 [31–34], group 2 [35] and, in some cases, both group 1 and group 2 influenza A viruses [36*,37,38]. Recently, Pappas et al. showed that most individuals make antibodies that broadly recognize group 1 subtypes. This public antibody response relies exclusively on the H chain, which is encoded by VH1-69 alleles with phenylalanine at position 54 and a short HCDR3 with tyrosine at position 98. Interestingly, through the reconstruction of the developmental pathways of several such clones, it was
found that high affinity binding was achieved in most cases by a single mutation [39**]. These findings explain the high frequency of antibodies specific for the stem of group 1 HAs. In contrast, antibodies that recognize both group 1 and group 2 HAs are much less frequent since they use different VH/VL genes. Furthermore they are generated through a complex developmental pathway, being first selected for their germ-line reactivity against group 1 and subsequently acquiring group 2 reactivity through somatic mutations [36]. The latter finding suggests that the rare pan-influenza A neutralizing antibodies might be elicited using an appropriate heterologous prime-boost strategy.

The identification of the HA stem as a conserved site of influenza HA has raised the possibility of developing a universal influenza vaccine [40]. Two approaches have been developed towards this ambitious goal. The first involves a prime-boost strategy using chimeric HA molecules carrying the same stem combined with heterologous heads [41,42] and is based on the classical principle of the original antigenic sin that represents in this case a natural mechanism of immunofocusing. The second approach of epitope-focused vaccinology involves the production of headless HAs [43] (Figure 1). In a recent report, Yassine et al. used iterative cycles of structure-based design to produce a stabilized stem immunogen, HA-SS, genetically fused to ferritin nanoparticles, which is recognized by stem-specific monoclonal antibodies and elicits, in mice and ferrets, broadly cross-reactive antibodies [44**]. In general, the assemblies of multiple copies of subunit antigens in well-ordered arrays, such as in the case of ferritin or VLPs, offer the advantage of mimicking the repetitiveness of most natural pathogens surface proteins potentially providing improved antigen stability and immunogenicity [45,46]. In an another study, Impagliazzo et al. used a rational design combined with a library approach to generate a HA stem antigen called mini-HA that protects mice and non-human primates from lethality and symptoms [47**]. It should be noted that in both studies the neutralizing antibody levels induce by the vaccine were modest, suggesting that in vivo protection may be achieved through a combination of neutralization and Fc-dependent effector function, which is known to be elicited by anti-stem antibodies [36*,48,49]. The headless HA vaccine has been shown to work for group 1 influenza viruses and should be extended to group 2 viruses, and possibly to influenza B, in order to fully realize the dream of a universal influenza vaccine.

**Factors that limit the antibody response: the challenge of HIV-1**

An effective antibody response is limited by the frequency of antigen-specific B cells in the naïve repertoire, by the number of mutations required to reach high affinity or breadth, and by the availability of T cell help. Thus, an ideal vaccine should contain a sufficiently high number of T and B cell epitopes and should be formulated to effectively prime the germinal center reaction [50]. In addition, the finding that preexisting immunity can shape the antibody response to new structurally related antigens [51–53] suggests modalities of prime-boost immunization to optimize the antibody response [54].

Recent technological advances are increasingly used to investigate the developmental pathways leading to potent and broadly neutralizing antibodies. These include the in depth analysis of the antibody response through the isolation of multiple cells within the same lineage, possibly implemented by next generation sequencing and deconvolution of the serum antibody repertoire through LC–MS/MS analysis [39*,55,56]. The systematic reconstruction of the genealogy trees of antibody lineages has shown that neutralizing antibodies to influenza or HRV mature rapidly [39**] or may not even need affinity maturation, while somatic mutations are critical to achieve breadth [23,36**]. The situation is strikingly different in the case of HIV-1, where viruses and B cells co-evolve over a period of years, leading to a slow and infrequent development of broadly neutralizing antibodies [57]. In the last decade many new broadly neutralizing antibodies of high potency have been isolated and new sites of vulnerability on the Env protein have been defined. Importantly, in the last few years cryo-EM and crystallography were used to finally solve the structure of a stabilized soluble Env trimer molecule (i.e. BG505 SOSIP), alone or in combination with several broadly neutralizing antibodies [58–60]. These constructs are now considered a relevant mimic of the native functional trimer being recognized preferentially by neutralizing antibodies. The HIV-1 broadly neutralizing antibodies are found only in a fraction of infected individual and have a slow and complex developmental pathway or derive from very rare precursors characterized by extremely long HCDR3. These observations have led to the proposal of a prime-boost strategy with different antigens that target the naïve precursors and different stages of the antibody developmental pathway defined as ‘B cell-lineage vaccine design’ [61*].

These hurdles represent a significant challenge to the development of an effective HIV-1 vaccine. However, these efforts contributed to the development of novel vaccine design strategies that represent a fertile ground for significant advances in the generation of new and better vaccines against other pathogens.

**Acknowledgements**

We thank Siro Bianchi for the help in the preparation of the figure.

The work in Lanzavecchia’s laboratory is supported by the European Research Council (grant no. 250348 IMMUNEexplore and 679955 BROADimmune), the Swiss National Science Foundation (grant no. 160279), the Swiss Vaccine Research Institute.
References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
●● of outstanding interest


With [12] provides evidence for the immunogenicity of a HCMV pentameric vaccine.


Broadly reactive antibodies to malaria antigens generated by a new mechanism of DNA transposition.


Together with [20] provides an example of epitope-focused vaccination.


Together with [19] provides an example of epitope-focused vaccination.


A remarkable example of structure-based design of a stabilized prefusion RSV F protein vaccine.


Describes a common motif used by antibodies that bind to the receptor binding site of influenza HA.


37. First example of a pan-influenza A neutralizing antibody.


41. Describes the requirements and developmental pathway of broadly influenza neutralizing antibodies.


With [47**] shows an example of epitope-focused vaccine based on influenza HA stem.


With [44**] shows an example of epitope-focused vaccine based on influenza HA stem.