

Immune epitope mapping in the post-genomic era: lessons for vaccine development

Alessandro Sette and Bjoern Peters

Technological advances have allowed for the creation of ever more complete maps of targets of immune responses in infectious pathogens. The evidence accumulating from such recent studies points to a broader range of targets recognized than previously expected, in terms of both numbers and characteristics of the targeted antigens. Also, multiple studies report a substantial variation in the targets recognized in different human individuals. These findings are not in conflict with the concept of immunodominance, because there are still only a few targets recognized compared with the multitude of potential targets available in a complex pathogen. However, they raise the question if vaccines, which try to emulate protective natural immune responses, should elicit an equally broad range of responses to efficiently convey protection.

Addresses

La Jolla Institute for Allergy and Immunology, Division of Vaccine Discovery, 9420 Athena Circle, La Jolla, CA 92037, USA

Corresponding author: Sette, Alessandro (alex@liai.org)

Current Opinion in Immunology 2007, **19**:106–110

This review comes from a themed issue on
Antigen processing and recognition
Edited by José Villadangos and Peter Van Endert

Available online 17th November 2006

0952-7915/\$ – see front matter

© 2006 Elsevier Ltd. All rights reserved.

DOI [10.1016/j.coi.2006.11.002](https://doi.org/10.1016/j.coi.2006.11.002)

Introduction

Infectious pathogens contain a large number of possible targets for immune responses, only few of which are recognized following an immunization or natural infection. The reasons for the immunodominance of the targets recognized can be either intrinsic to the targets themselves, such as differences in their affinity for relevant receptors of the immune system, or extrinsic, such as competitive effects that suppress the response against one target in favor of another. The mechanisms that contribute to immunodominance have been studied extensively, and the first excellent handbook on the subject has been published recently [1].

Instead of duplicating existing overviews, herein we will review recent studies aiming to characterize the full spectrum of immune response targets in infectious pathogens. Even within this narrow scope, the field of

immunodominance and related topics is broad. Because of this, and owing to space constraints, we admittedly focus only on a subset of all potentially relevant articles, of subjective and arbitrary interest to us. For this reason, we apologize in advance for the many important and outstanding references that are not discussed herein.

Earlier studies that characterized targets of immune responses typically focused on relatively simple pathogens, or on a few antigens for which previous knowledge about their antigenicity existed. Although the latter is a useful strategy to identify some targets of immune responses, it is obviously not suited to an exhaustive discovery of all targets. Technical advances and knowledge of the whole genomic composition of pathogens have recently made it possible to study an even broader range of targets with a manageable experimental effort. Herein, we will review such studies that were performed in the past two years, and will reveal a larger than expected breadth of immune responses in terms of both numbers and characteristics of the antigens recognized.

The range of targets recognized in a protective natural immune response provides guidance on what a response induced through vaccination should look like. Reverse vaccinology approaches provide similar guidance by systematically testing antigens for their ability to provide protection. Likewise, we review some recent reverse vaccinology studies, and will discuss how these results and those from mapping natural immune responses might be combined in the process of vaccine development.

Mapping T-cell responses in human donors

The top part of [Table 1](#) lists several studies from the past two years that aimed to identify the majority of targets of CD8⁺ and/or CD4⁺ T-cell responses detected following viral infections in human donors. In all of these studies, T cells were taken from an immunized donor and were presented with fragments of the pathogen that could thereby be tested for recognition. The actual targets recognized by T cells are typically peptide epitopes presented by MHC molecules. Because the number of unique peptides contained in a complex pathogen easily exceeds tens of thousands, it is impractical to test all of them individually for recognition. This is a particular problem in the case of individual human donors, for which the sample amounts are limited. In two studies using smaller viruses — hepatitis C virus (HCV) [2] and dengue virus (DV) [3] — testing all candidates was nevertheless achieved by using a two-step approach in which pools of

Table 1

Summary of selected recent studies mapping targets of immune responses in infective pathogens.

Cell type	Organism	Immunization	Pathogen	Tested targets	CD4 or CD8	Proteins recognized/ proteins tested	Epitopes recognized/ epitopes tested	Refs
T cell	Human	Spontaneously resolved infection	HCV	301 peptides overlapping the HCV polyprotein	CD4	6/6	89/301	[2]
		Acute secondary infection	DV	260 peptides overlapping 6 proteins	CD4 CD8	6/6	47/260	[3]
		Seropositive donors	HCMV	213 peptide pools representing all ORFs	CD4 CD8	125/213 107/213	– –	[4**]
		Seropositive donors	HSV-2	48 peptide pools representing 48/78 ORFs	CD8	48/48	–	[5]
		Vaccinated (Dryvax)	VACV	Peptides predicted to bind MHC (in [6], initial screen with VACV genomic library)	CD8	43/258	63	[6,7*,8]
T cell	Mice	Infected C57BL/6 mice	MCMV	170 ORFs in plasmids	CD8	27/170	24	[9*]
		Infected C57BL/6 mice	VACV	258 ORFs in plasmids	CD8	5/258	5	[10]
		Infected C57BL/6 mice	VACV	2256 peptides predicted to bind MHC	CD8	43/258	49/2256	[11**]
Antibody	Human	Vaccinated, pooled serum	VACV	194 ORFs	–	14/194	–	[14]
		Infected	<i>C. trachomatis</i>	156 ORFs	–	27/156	–	[16]
	Rabbit	Vaccinated, live vaccine (EV76)	<i>Y. pestis</i>	149 ORFs	–	13/149	–	[15]

peptides were first tested, followed by tests of individual peptides from the pools that were identified as positive for T-cell recognition. For two larger viruses — human cytomegalovirus (HCMV) [4**] and herpes simplex virus (HSV) [5] — peptides were tested and were pooled by their open reading frame (ORF) of origin, without an additional deconvolution step. This effectively maps targets of T-cell responses at the level of individual ORFs.

Three recent studies have mapped CD8⁺ T-cell responses in vaccinia virus (VACV) using samples from human donors immunized with Dryvax, a small pox vaccine preparation of live-VACV. The studies differ in the effector cells used and in their approaches to target selection, but all map responses to individual peptide epitopes. One study established cytotoxic T lymphocyte (CTL) lines through limited dilution cloning [6], which were tested for recognition of peptides selected based on the presence of relevant human leukocyte antigen-binding motifs. A second study tested similarly selected peptides using peripheral-blood mononuclear cells taken directly *ex vivo* [7*]. A third study used target cells that expressed polypeptide fragments from a library of vaccinia genomic DNA, which were tested for recognition by both CTL clones and bulk CTL [8]. Comparing the two

types of effector cells used, the authors of this last study found that bulk cells provide higher throughput in epitope discovery without loss of precision.

The range of targets recognized in these T-cell response studies in terms of epitopes and their source proteins is also summarized in Table 1. For the small viruses (HCV and DV), responses against all proteins examined were detected in at least one individual. Although this result might have been expected, a surprisingly broad response was also detected in the case of larger viruses. For HCMV, 107 of the 213 ORFs tested were found to be recognized by CD8⁺ cells and 125 by CD4⁺ cells. In the HSV study, all 48 ORFs tested elicited responses in at least one donor. Even though the ORFs in that study were pre-selected to include all proteins that are immediate-early expressed or part of virion particles, only 30 HSV ORFs were not tested; this means that more than half of all possible ORFs evoked a response. Similarly, when the results from the three VACV studies that were performed in humans are compiled, it is apparent that epitopes from at least 43 ORFs were recognized. Importantly, for HCMV and VACV, new ORFs were discovered as being recognized with nearly every new donor tested. Thus, it is likely that the currently known set of recognized proteins for these viruses is not yet complete.

T-cell responses in individual human donors

Which and how many targets were recognized by individual donors varied dramatically in all studies, with at least a ten-fold difference between the minimum and maximum number of targets recognized in any individual. The highest number of targets recognized was found in the HSV study [5], in which an individual donor responded to 46 of the 48 tested ORFs. By contrast, the reported median numbers of targets recognized per donor was between 8 and 12 for all studies, which is a surprisingly narrow range given the diversity of the pathogens studied (HCV, DV, HCMV, HSV-2 and VACV) and experimental approaches utilized.

Only a minority of targets was recognized across multiple donors. As each donor carries an individual and possibly unique set of MHC alleles, a divergence of recognized targets was expected. In addition, the history of each donor in encountering pathogens is likely to influence the recognition pattern in these studies. Finally, with the exception of the Dryvax immunization studies [6,7,8], genetic variations in the viral strains infecting each donor could further individualize the responses in each donor. It is currently difficult to assess the relative importance of these factors. However, in the HCMV study [4**] a pair of identical twins were enrolled, and displayed a remarkably similar response profile. This indicates that variations in the genetic background — including but not limited to the MHC allele composition — might contribute a dominant influence on the level and pattern of immune responses observed in any given individual.

T-cell responses in mice

One advantage of studying immune responses to experimental infections in inbred strains of mice is the large supply of T cells, which overcomes the limiting amount of sample available from individual human donors. Table 1 lists three recent studies that have attempted to map the entire T-cell response against a viral pathogen and have deconvoluted the targets of these responses to the peptide level.

Munks *et al.* [9*] tested target cells transfected with plasmids encoding predicted ORFs of the murine cytomegalovirus (MCMV) for recognition by T cells from immunized mice. Of the 170 predicted MCMV ORFs, responses were detected against 27. For 18 of these, responses could be further mapped to 24 individual peptide epitopes, which accounted for most (~85%) of the MCMV-specific response. In a study using a similar ORF expression library to screen for VACV epitopes, only 5 epitopes were identified out of the 258 studied [10]. However, these only accounted for about 40% of the VACV-specific response. In a consequent study [11**], 2256 VACV-derived peptides, which were predicted to bind murine MHC class I H-2^b molecules with high affinity, were tested for recognition by T cells from

VACV-infected mice. This resulted in the discovery of an additional 44 epitopes that, in combination with the original 5, accounted for 95% of the VACV-specific response. These studies demonstrate that it is now possible to account for the majority of T-cell response targets in inbred mice for complex viral pathogens.

Mapping targets of B-cell responses

When mapping the targets of B-cell responses, one major hurdle is the fact that many antibodies bind to conformational epitopes (i.e. amino acids in a protein sequence that are brought together in the three-dimensional protein fold but are disparate in its linear sequence). To map such epitopes, it is important to test proteins in their native conformation. One recently developed approach that allows for high-throughput screening of targets while maintaining a reasonable degree of native conformation is the use of microarrays spotted with viral proteins (reviewed in [12]). Three such studies are listed at the bottom of Table 1.

Davies *et al.* [13*] studied antibody responses to VACV in a set of 194 proteins. 14 proteins were recognized by antibodies from vaccinated humans, and an additional 13 antigens were recognized by non-immunized individuals. The latter are thought to be recognized by cross-reactive antibodies specific for other environmental antigens. The study further compared the results obtained using human sera with sera from VACV-infected mice and macaques. 11 of the antigens specifically recognized by immunized humans are also recognized in mice and macaques, which is a remarkable level of agreement between species. Antibodies against one of these proteins were subsequently shown to provide protection against lethal challenges with VACV in mice [14].

Li *et al.* [15] identified antibody targets recognized following vaccination of rabbits with the *Yersinia pestis* live vaccine EV76. They found that 50 out of 149 proteins are recognized, 37 of which also show cross-reactivity with pre-immunized sera. Several of the newly identified targets elicited equal or stronger responses than the previously identified V antigen. Sharma *et al.* [16] used an array of 156 ORFs, which were selected as a representative sampling of the *Chlamydia trachomatis* genome. Screening sera of 15 women that were urogenitally infected with *C. trachomatis* gave a broad range of responses targeting between 2 and 18 proteins. 7 ORFs were recognized by antibodies from more than half of the donors, which makes them more frequently recognized than the major outer membrane protein (MOMP) antigen, which was previously considered to be the immunodominant antigen.

Characteristics of antigens recognized by natural immune responses

It is commonly assumed that certain characteristics of an antigen, such as its function or expression kinetics,

influence its recognition by immune responses. When examining the ORFs recognized in the studies compiled in Table 1, it becomes apparent that such a correlation cannot be absolute: ORFs from all functional categories, times of expression and cellular localizations are found to evoke responses, in some cases for both antibody and T-cell responses. Because of the large number of targets recognized, this is not surprising. Thus, it might be more meaningful to correlate the frequency with which responses are detected with characteristics of the ORFs recognized.

For T-cell responses, one important correlate with ORF recognition is the size of the ORF [4^{••},5,7[•],11^{••}]. It is indicated that the higher the number of peptides encoded by an ORF, the more likely it is that one of them will be recognized. Accordingly, it is important to control for the number of peptides encoded in a class of ORFs when analyzing if recognition of this class is more or less frequent than that of other ORFs. After taking this into account, several studies found more frequent responses to proteins that are expressed early in the infection cycle [4^{••},7[•],9[•],11^{••}]. Also, two studies [3,4^{••}] found a positive correlation between inter-strain conservancy of amino acid sequences and their frequency of recognition. It makes intuitive sense that the targets that are preferred for immune responses are those that are expressed early in an infection and those that are conserved between strains.

For antibody responses, the majority of recognized antigens characterized in [13[•]] were structural proteins, more than half of which were envelope proteins. However, antigens that are not part of the envelope also elicited strong responses. As the authors state, the contribution of these antibodies to protection remains to be determined. For both *Y. pestis* and *C. trachomatis*, several of the ORFs eliciting strong antibody responses have not yet been assigned a function. For this reason, these are targets of high interest for further studies by immunologists and bacteriologists alike. This highlights how mapping of immune responses can also contribute to a better basic understanding of biology of the particular pathogen under examination.

Reverse vaccinology

Mapping immune responses following infection characterizes targets of natural immune responses. From a vaccine development perspective, it should be emphasized that the main goal is to induce a protective immune response, not necessary one that mimics natural immune responses. Testing this directly is the goal of the reverse vaccinology approach, which uses genomic information regarding pathogens to systematically screen their components for protective efficacy. This is typically done by immunizing mice with individual proteins and testing their ability to elicit protective antibodies, as reviewed in [17,18[•]].

In [19], Giuliani *et al.* describe the use of a combination vaccine consisting of five meningococcus antigens formulated in aluminum hydroxide, each previously found to elicit protective responses using a reverse vaccinology approach [20,21]. The antibodies elicited in mice by the combined vaccine reacted with 66 out of 85 clinical isolates of meningococcus serogroup B. Using the same vaccine, antigens adjuvanted by aluminum hydroxide combined with CpG oligonucleotides, oil-in-water emulsion MF59 or complete Freund's adjuvant led to increased reactivities of 92%, 94% and 98%, respectively (which serves as a reminder of the important role of adjuvants in effective vaccine design).

The response induced by the combination of five antigens is broader than that elicited by each individual antigen. Importantly, it is also broader than that of previously developed vaccines based on biochemically extracted mixtures of proteins from outer membrane vesicles of meningococcus serogroup B, which elicited responses against 17–18 of the 85 clinical isolates. One explanation for the relatively low cross-strain protection provided by bacterium-extracted outer membrane vesicles is the presence of proteins in the extracted vesicles that are unlikely to be found on the bacterial surface and therefore might 'distract' the immune system [22]. Another explanation is that such extracts focus the response on immunodominant antigens such as PorA, which is highly variable between strains and elicits a strong but strain-specific response.

Conclusions

In summary, the immune response mapping studies assembled in this review show that the targets of natural immune responses have a broader distribution than previously appreciated. At the same time, a smaller subset of targets is recognized more frequently than others. This suggests that a binary classification into recognized and non-recognized targets is less appropriate than a classification into more- and less-frequently recognized targets, the former corresponding to immunodominant antigens.

Vaccine design based on reverse vaccinology approaches has shown that combining multiple antigens in a vaccine provides broader protection against different strains of a pathogen than individual antigens. This finding is easily explained by variations in the antigens between different strains of the same pathogens. This also provides an example of why vaccines associated with extreme immunodominance and focusing responses on a single target might not be desirable. However, it is unclear if vaccine-induced responses need to have the same breadth of targets as natural responses. To answer this question, future studies might need to systematically evaluate the role of targets in natural immune responses by testing their ability to contribute to protection following immunization or by

correlating response patterns in individuals with their disease outcomes.

As high-throughput immune response mapping techniques generate an ever larger amount of data, it becomes increasingly cumbersome to assemble and compare their results. This raises the need to centrally store and make accessible this data, which is the goal of the immune epitope database [23] and other related online resources. We believe that, in the future, such bioinformatics resources will play an important role in the generation of scientific hypothesis and in the selection of biological and immunological targets for experimental testing.

Acknowledgements

This study was supported by National Institutes of Health Contract HHSN266200400024C and National Institutes of Health Grant RO1-AI-56268.

References and recommended reading

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

- of special interest
- of outstanding interest

1. Frelinger JA (Ed): *Immunodominance: the choice of the immune system*. Wiley; 2005.
This first handbook on the subject provides a detailed overview of the current understanding of mechanisms contributing to immunodominance.
2. Schulze zur Wiesch J, Lauer GM, Day CL, Kim AY, Ouchi K, Duncan JE, Wurcel AG, Timm J, Jones AM, Mothe B *et al.*: **Broad repertoire of the CD4⁺ Th cell response in spontaneously controlled hepatitis C virus infection includes dominant and highly promiscuous epitopes**. *J Immunol* 2005, **175**:3603-3613.
3. Simmons CP, Dong T, Chau NV, Dung NT, Chau TN, Thao le TT, Dung NT, Hien TT, Rowland-Jones S, Farrar J: **Early T-cell responses to dengue virus epitopes in Vietnamese adults with secondary dengue virus infections**. *J Virol* 2005, **79**:5665-5675.
4. Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Rucht F, Sleath PR, Grabstein KH, Hosken NA, Kern F *et al.*: **Broadly targeted human cytomegalovirus-specific CD4⁺ and CD8⁺ T cells dominate the memory compartments of exposed subjects**. *J Exp Med* 2005, **202**:673-685.
This study presents the first map of the totality of CD4⁺ and CD8⁺ T-cell responses to a complex pathogen such as HCMV on the level of the ORF recognized.
5. Hosken N, McGowan P, Meier A, Koelle DM, Sleath P, Wagener F, Elliott M, Grabstein K, Posavad C, Corey L: **Diversity of the CD8⁺ T-cell response to herpes simplex virus type 2 proteins among persons with genital herpes**. *J Virol* 2006, **80**:5509-5515.
6. Terajima M, Cruz J, Leporati AM, Demkowicz WE Jr, Kennedy JS, Ennis FA: **Identification of vaccinia CD8⁺ T-cell epitopes conserved among vaccinia and variola viruses restricted by common MHC class I molecules, HLA-A2 or HLA-B7**. *Hum Immunol* 2006, **67**:512-520.
7. Oseroff C, Kos F, Bui HH, Peters B, Pasquetto V, Glenn J, Palmore T, Sidney J, Tschärke DC, Bennink JR *et al.*: **HLA class I-restricted responses to vaccinia recognize a broad array of proteins mainly involved in virulence and viral gene regulation**. *Proc Natl Acad Sci USA* 2005, **102**:13980-13985.
This study maps T-cell responses to a complex pathogen on a peptide level, testing more than 6000 peptides in human donors.
8. Jing L, Chong TM, McClurkan CL, Huang J, Story BT, Koelle DM: **Diversity in the acute CD8 T cell response to vaccinia virus in humans**. *J Immunol* 2005, **175**:7550-7559.
9. Munks MW, Gold MC, Zajac AL, Doom CM, Morello CS, Spector DH, Hill AB: **Genome-wide analysis reveals a highly diverse CD8 T cell response to murine cytomegalovirus**. *J Immunol* 2006, **176**:3760-3766.
The authors map the majority of T-cell responses against MCMV in mice to the level of peptides recognized, finding a similar breadth as in [11••]
10. Tschärke DC, Karupiah G, Zhou J, Palmore T, Irvine KR, Haeryfar SM, Williams S, Sidney J, Sette A, Bennink JR *et al.*: **Identification of poxvirus CD8⁺ T cell determinants to enable rational design and characterization of smallpox vaccines**. *J Exp Med* 2005, **201**:95-104.
11. Moutaftsi M, Peters B, Pasquetto V, Tschärke DC, Sidney J, Bui HH, Grey H, Sette A: **A consensus epitope prediction approach identifies the breadth of murine T(CD8⁺)-cell responses to vaccinia virus**. *Nat Biotechnol* 2006, **24**:817-819.
This study demonstrates that the majority of T-cell responses are targeted against peptides that can be identified by bioinformatic predictions of MHC binding.
12. Robinson WH: **Antigen arrays for antibody profiling**. *Curr Opin Chem Biol* 2006, **10**:67-72.
13. Davies DH, Liang X, Hernandez JE, Randall A, Hirst S, Mu Y, Romero KM, Nguyen TT, Kalantari-Dehaghi M, Crotty S *et al.*: **Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery**. *Proc Natl Acad Sci USA* 2005, **102**:547-552.
This study applies protein chip technology to map antibody responses on a whole-proteome basis, showing a remarkably similar profile of responses between humans, primates and mice.
14. Davies DH, McCausland MM, Valdez C, Huynh D, Hernandez JE, Mu Y, Hirst S, Villarreal L, Felgner PL, Crotty S: **Vaccinia virus H3L envelope protein is a major target of neutralizing antibodies in humans and elicits protection against lethal challenge in mice**. *J Virol* 2005, **79**:11724-11733.
15. Li B, Jiang L, Song Q, Yang J, Chen Z, Guo Z, Zhou D, Du Z, Song Y, Wang J *et al.*: **Protein microarray for profiling antibody responses to *Yersinia pestis* live vaccine**. *Infect Immun* 2005, **73**:3734-3739.
16. Sharma J, Zhong Y, Dong F, Piper JM, Wang G, Zhong G: **Profiling of human antibody responses to *Chlamydia trachomatis* urogenital tract infection using microplates arrayed with 156 chlamydial fusion proteins**. *Infect Immun* 2006, **74**:1490-1499.
17. Scarselli M, Giuliani MM, Adu-Bobie J, Pizza M, Rappuoli R: **The impact of genomics on vaccine design**. *Trends Biotechnol* 2005, **23**:84-91.
18. Mora M, Donati C, Medini D, Covacci A, Rappuoli R: **Microbial genomes and vaccine design: refinements to the classical reverse vaccinology approach**. *Curr Opin Microbiol* 2006, **9**:532-536.
This review describes advances in the reverse vaccinology approach driven by the availability of multiple related microbial genomes, making it possible to rationally design vaccines for cross-strain reactivity.
19. Giuliani MM, Adu-Bobie J, Comanducci M, Arico B, Savino S, Santini L, Brunelli B, Bambini S, Biolchi A, Capecchi B *et al.*: **A universal vaccine for serogroup B meningococcus**. *Proc Natl Acad Sci USA* 2006, **103**:10834-10839.
20. Pizza M, Scarlato V, Masignani V, Giuliani MM, Arico B, Comanducci M, Jennings GT, Baldi L, Bartolini E, Capecchi B *et al.*: **Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing**. *Science* 2000, **287**:1816-1820.
21. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, Eisen JA, Ketchum KA, Hood DW, Peden JF, Dodson RJ *et al.*: **Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58**. *Science* 2000, **287**:1809-1815.
22. Ferrarini G, Garaguso I, Adu-Bobie J, Doro F, Taddei AR, Biolchi A, Brunelli B, Giuliani MM, Pizza M, Norais N *et al.*: **Outer membrane vesicles from group B *Neisseria meningitidis* Δgna33 mutant: proteomic and immunological comparison with detergent-derived outer membrane vesicles**. *Proteomics* 2006, **6**:1856-1866.
23. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Flieri W, Kronenberg M, Kubo R, Lund O *et al.*: **The immune epitope database and analysis resource: from vision to blueprint**. *PLoS Biol* 2005, **3**:e91.