nature biotechnology

A consensus epitope prediction approach identifies the breadth of murine T_{CD8+}-cell responses to vaccinia virus

Magdalini Moutaftsi[1], Bjoern Peters[1], Valerie Pasquetto[1], David C Tscharke[2], John Sidney[1], Huynh-Hoa Bui[1], Howard Grey[1] & Alessandro Sette[1]

© 2006 Nature Publishing Group http://www.nature.com/naturebiotechnology

The value of predictive algorithms for identifying CD8+ T (T_{CD8+})-cell epitopes has not been adequately tested experimentally. Here we demonstrate that conventional bioinformatic methods predict the vast majority of T_{CD8+}-cell epitopes derived from vaccinia virus WR strain (VACV-WR) in the H-2^b mouse model. This approach reveals the breadth of T-cell responses to vaccinia, a widely studied murine viral infection model, and may provide a tool for developing comprehensive antigenic maps of any complex pathogen.

T_{CD8+} cells recognize peptides associated with major histocompatibility complex (MHC) class I molecules on the surfaces of cells harboring intracellular pathogens such as viruses, and destroy the infected cells. Therefore, comprehensive analysis of the antigen specificity of host-T-cell responses to pathogen infection is crucial to understanding host-pathogen interactions and developing diagnostic tools and vaccines. Here we determined the extent to which epitopes predicted by bioinformatics account for the immune response to vaccinia virus (VACV) after infection of C57BL/6 mice, a well-characterized model in which about one-third of all T_{CD8+} cells are activated at the peak of the response to infection.

Using an expression library approach, we previously identified five VACV-derived, H-2^b–restricted epitopes that account for 40% of the total anti-VACV T_{CD8+}-cell response[1]. In the present study, we examined whether additional epitopes accounting for the remaining 60% could be identified by a consensus epitope prediction approach. Four scoring-matrix prediction methods were applied to rank all possible 8-, 9- and 10-mers from the VACV-WR sequence according to their predicted abilities to bind H-2 Kb (8- and 9-mers) and Db (9- and 10-mers) MHC class I molecules (**Fig. 1**; see **Supplementary Methods** and **Supplementary Data 1** online). The median rank in the four predictions was used to select the top 1% of peptides within each size-based class, comprising 2,256 peptides (see **Supplementary Data 2** online). These were split into pools that were each tested for antigenicity by interferon (IFN)-γ-ELISPOT assays (see **Supplementary Methods** online). Representative data from screening the K^b 8-mers, where 15 positive pools were identified, are shown in **Figure 1a**. Screening of K^b 9-mer, D^b 9-mer and D^b 10-mer pools identified 66 positive pools (data not shown). Peptides within positive pools were then tested individually. In the case of the K^b 8-mer–positive pools,

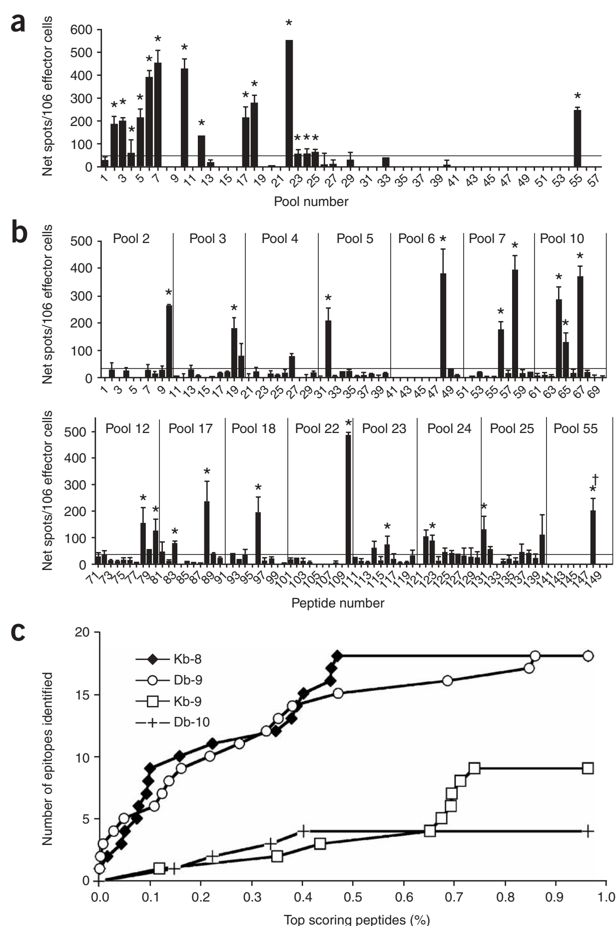

Figure 1 Identification of H-2^b–restricted epitopes derived from VACV-WR–infected C57BL/6 mice. (**a**) A representative screen of H-2 K^b 8-mer peptide pools. The asterisks indicate pools that were selected for deconvolution. (**b**) A representative deconvolution of the positive H-2 K^b 8-mer peptide pools. The asterisks indicate the identified VACV-WR–derived epitopes. (**c**) Evaluation of bioinformatics prediction accuracy. Each curve represents the prediction for one MHC allele and peptide size. The *y*-axis depicts the number of epitopes that were identified when testing the fraction of peptides contained in VACV-WR indicated on the *x*-axis.

[1]La Jolla Institute for Allergy and Immunology, 3030 Bunker Hill Street, Suite 326, San Diego, California 92109, USA. [2]Division of Immunology and Infectious Diseases, Queensland Institute of Medical Research, Herston, Australia. Correspondence should be addressed to A.S. (alex@liai.org).

Received 6 February; accepted 19 April; published online 11 June 2006; doi:10.1038/nbt1215

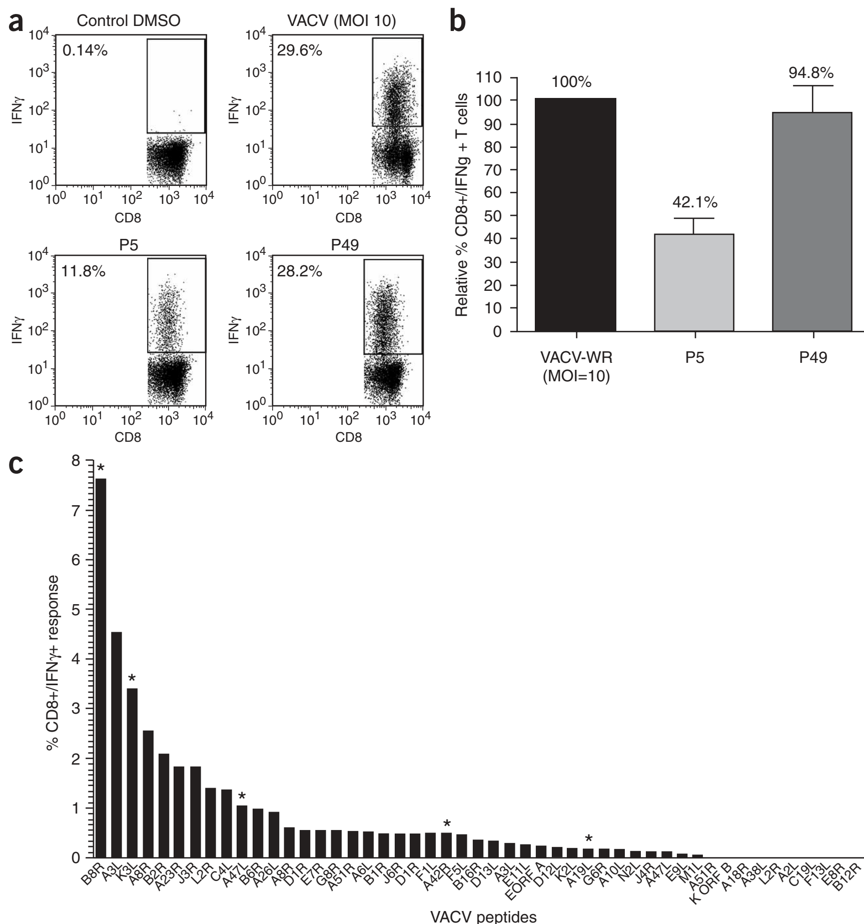

Figure 2 The identified epitopes account for the majority of the T_{CD8+}-cell response to VACV-WR. (**a**) Representative dot plots from IFN-γ–producing, VACV-WR–specific T_{CD8+}-cells after stimulation with DC2.4 cells either pulsed with peptide-pools or infected with VACV-WR (MOI = 10). (**b**) Relative contributions of a pool of the 5 previously identified epitopes (P5) and a pool of all 49 identified epitopes (P49) to the total VACV-WR-specific T_{CD8+}-cell response (represented by stimulation with infected DC2.4 cells). (**c**) Individual peptides were plotted on a linear scale according to their net contribution to the total IFN-γ response to VACV-WR (from highest to lowest response).

© 2006 Nature Publishing Group http://www.nature.com/naturebiotechnology

this identified 19 positive peptides (**Fig. 1b**). Comparable results were obtained for the K^b 9-mer, D^b 9-mer and D^b 10-mer pools (data not shown). This analysis identified 49 H-2^b–restricted T-cell epitopes (18 K^b 8-mers, 9 K^b 9-mers, 18 D^b 9-mers and 4 D^b 10-mers), including all 5 previously described epitopes[1]. Assays in individual mice did not reveal significant differences in the breadth of responses (data not shown).

These 49 VACV peptides were tested for their capacities to bind purified K^b and D^b molecules *in vitro*. All 49 epitopes bound to the predicted MHC molecule with IC_{50} values of 2,600 nM or better. Of these, 33 (67% of the total) bound with very high affinity (≤20 nM). A summary of the characteristics of all 49 epitopes is provided (see **Supplementary Table 1** online).

Various algorithms are available for predicting potential T-cell epitopes (reviewed in ref. 3), but thus far, no experimental studies appear to have fully evaluated the power of these predictions. In general, prediction of peptide binding to MHC molecules alone cannot reliably predict T-cell epitopes, as other factors, such as expression of viral proteins, antigen processing, immunodominance and the presence of a T-cell repertoire, are also important. However, because peptide binding to MHC class I molecules is always necessary for antigen presentation, efficient prediction of this association should markedly reduce the number of peptides to be tested. Screening all possible 8-mer, 9-mer and 10-mer peptides from VACV-WR would require evaluating 175,716 peptides, whereas testing overlapping 15-mers that span the entire VAVC-WR sequence would require 12,212 peptides.

To examine the effectiveness of our predictive approach, we plotted each identified epitope according to its allele/size combination as a function of the fraction of all possible peptides (see **Supplementary Table 1** online and **Fig. 1c**). All of the epitopes identified were included within the set of 503 peptides with the highest scores (0.9% of the total), suggesting that all predictable epitopes were identified. The 300 highest-ranking predicted epitopes (0.5% of the total) would have been sufficient to identify all 18 K^b 8-mers, 15 out of 18 D^b 9-mers and all 4 D^b 10-mers. Together, these correspond to 93% of all VACV-WR epitopes. In contrast, only 3 of the 9 K^b 9-mers ranked within the top 300 predicted epitopes, possibly as a smaller data set was available for generating the K^b 9-mers algorithm. Comparison of the performance of the consensus prediction with the individual predictions on which the consensus method is based reveals that the consensus approach is superior to any single predictive strategy tested (see **Supplementary Fig. 2** online).

To determine what fraction of the total VACV-WR–specific T_{CD8+}-cell response is accounted for by the identified epitopes, we stimulated splenocytes from VACV-WR–infected mice with DC2.4 target cells either pulsed with pools of peptides or infected with VACV-WR. Their IFN-γ responses were evaluated by intracellular cytokine staining (ICCS) (see **Supplementary Methods** online). Of all T_{CD8+}-cells, 29.6% were IFN-γ^+ when tested against VACV-WR–infected cells (**Fig. 2a**). As already described, the pool containing the five previously identified epitopes (P5) stimulated 11.8% of T_{CD8+} cells (about 40% of the total response) (**Fig. 2b**). More importantly, the pool corresponding to all 49 epitopes (P49) stimulated 28.2% of T_{CD8+} cells, thus accounting for the vast majority (94.8%) of the total VACV-WR T_{CD8+}-cell response (**Fig. 2b**). The percentage of T_{CD8+}-cells involved in the response to each of the epitope is shown in **Fig. 2c**, with the five previously identified epitopes marked by asterisks. Notably, although no overwhelmingly dominant epitope was found, a clear hierarchy of dominance emerges. The previously identified dominant $B8R_{20-27}$ epitope retains its eminence and accounts for about 25% of the total response.

With respect to the nature of the VACV-WR proteins recognized, our data are in accordance with our previous published data describing epitopes identified in humans and transgenic mice[4,5] (see **Supplementary Table 2** online). Most VACV antigens recognized by T_{CD8+} cells were >100 residues long and belonged to the group of

early-expressed proteins. However, late-expressed proteins were also recognized (see **Supplementary Table 3a**,**b** online). The percentage of recognized antigens within each functional category (virulence factors, structural protein, viral regulation) did not differ significantly from their frequency within the total proteome (see **Supplementary Table 3c** online).

In summary, our data indicate that the vast majority of VACV-specific, T_{CD8+}-cell epitopes can be predicted by bioinformatics. Unconventional epitopes, which evade prediction, do not make a significant contribution to the total immune response and therefore should not prevent the development of comprehensive antigenic maps for complex pathogens. However, it should be noted that only one well-characterized inbred mouse strain with only two well-characterized MHC restriction elements was used in this study. Therefore, these data may better represent the potential of immunoinformatics, rather than the current state of epitope predictive methods in general.

Our study has general implications in terms of the breadth of anti-viral T_{CD8+}-cell responses. Contrary to mouse responses to smaller viruses[6,7], where one or very few epitopes dominate the anti-viral response, we observed a diverse response in which immunodominance was a subtle phenomenon that did not preclude substantial heterogeneity in T_{CD8+}-cell responses. It is reasonable to suggest that even greater diversity exists in human populations[4,8–10]. The breadth of response against VACV, coupled with the virus's high-fidelity DNA polymerase[11], suggests that poxviruses are unlikely to evade immune responses through epitope mutation. This stands in sharp contrast to small RNA viruses[12–14]. Regarding concerns about bioterrorism, provided the same results hold true in humans, it may be virtually impossible to engineer a mutated poxvirus void of immunogenic epitopes.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS
This study was funded by the National Institutes of Health (contract no. HHSN266200400024C, and RO1 grant no. RO1-AI-56268. D.C.T. is the recipient of an National Health and Medical Research Council (Australia) Howard Florey Centenary Fellowship no. 224273.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Tscharke, D.C. *et al. J. Exp. Med.* **201**, 95–104 (2005).
2. McCraith, S. *et al. Proc. Natl. Acad. Sci. USA* **97**, 4879–4884 (2000).
3. De Groot, A.S. & Berzofsky, J.A. *Methods* **34**, 425–428 (2004).
4. Oseroff, C. *et al. Proc. Natl. Acad. Sci. USA* **102**, 13980–13985 (2005).
5. Pasquetto, V. *et al. J. Immunol.* **175**, 5504–5515 (2005).
6. Wallace, M.E. *et al. J. Virol.* **73**, 7619–7626 (1999).
7. Kast, W.M. *et al. Proc. Natl. Acad. Sci. USA* **88**, 2283–2287 (1991).
8. Smith, C.L. *et al. J. Immunol.* **175**, 8431–8437 (2005).
9. Terajima, M. *et al. J. Exp. Med.* **197**, 927–932 (2003).
10. Jing, L. *et al. J. Immunol.* **175**, 7550–7559 (2005).
11. Bray, M. & Miller, M. *Clin. Infect. Dis.* **38**, 882–889 (2004).
12. Allen, T.M. *et al. J. Virol.* **79**, 12952–12960 (2005).
13. Bowen, D.G. *et al. J. Exp. Med.* **201**, 1709–1714 (2005).
14. Klenerman, P. & Zinkernagel, R.M. *Nature* **394**, 482–485 (1998).

© 2006 Nature Publishing Group http://www.nature.com/naturebiotechnology