

CD4⁺ T Cell Epitope Discovery and Rational Vaccine Design

Daniela Santoro Rosa · Susan Pereira Ribeiro ·
Edecio Cunha-Neto

Received: 16 June 2009 / Accepted: 8 August 2009 / Published online: 14 February 2010
© L. Hirszfeld Institute of Immunology and Experimental Therapy, Wrocław, Poland 2010

Abstract T cell epitope-driven vaccine design employs bioinformatic algorithms to identify potential targets of vaccines against infectious diseases or cancer. Potential epitopes can be identified with major histocompatibility complex (MHC)-binding algorithms, and the ability to bind to MHC class I or class II indicates a predominantly CD4⁺ or CD8⁺ T cell response. Furthermore, an epitope-based vaccine can circumvent evolutionary events favoring immune escape present in native proteins from pathogens. It can also focus on only the most relevant epitopes (i.e. conserved and promiscuous) recognized by the majority of the target population. Mounting evidence points to the critical role of CD4⁺ T cells in natural antigen encounter and active immunization. In this paper the need for CD4⁺ T cell help in vaccine development, the selection of CD4⁺

T cell epitopes for an epitope-based vaccine, and how the approach can be used to induce a protective effect are reviewed.

Keywords Vaccine design · CD4⁺ T cells · T cell help · Epitope prediction · Epitope-based vaccines

Introduction

High-throughput genome sequencing of pathogens and tumor antigens and advances in bioinformatics have disclosed a massive number of potential protein targets of immune responses. The use of this new knowledge for the development of novel vaccines is termed “reverse vaccinology” (Rappuoli 2000). The use of major histocompatibility complex (MHC)-binding algorithms has allowed the identification of novel T cell epitopes that could be used in vaccine design against infectious diseases and cancer, the so-called epitope-driven vaccine design. Epitope-based vaccines have the ability to focus the immune response on highly antigenic epitopes, free from the original protein scaffold, which can be selected to be conserved and amply recognized by the target population while increasing safety. Given the fundamental role played by CD4⁺ T cells in determining the functional status of both innate and adaptive immune responses, the inclusion of appropriate CD4⁺ T cell epitopes may be essential for vaccine efficacy. Here we describe the background underlying the need for CD4⁺ T cell help in vaccine development and how to select CD4⁺ T cell epitopes for an epitope-based vaccine. We also provide examples of how the approach can be used to successfully overcome barriers observed in whole protein-based vaccines.

D. S. Rosa · S. P. Ribeiro · E. Cunha-Neto
Division of Clinical Immunology and Allergy,
Department of Internal Medicine,
University of São Paulo School of Medicine,
São Paulo, SP, Brazil

E. Cunha-Neto
Heart Institute (InCor),
University of São Paulo School of Medicine,
São Paulo, SP, Brazil

D. S. Rosa · S. P. Ribeiro · E. Cunha-Neto
Institute for Investigation in Immunology-INCT,
São Paulo, SP, Brazil

E. Cunha-Neto (✉)
Faculdade de Medicina da Universidade de São Paulo,
LIM60, Avenida Dr. Arnaldo, 455, sala 3209, Cerqueira César,
São Paulo, SP 01246-903, Brazil
e-mail: edecunha@gmail.com

Antigen Presentation and T Cell Recognition

T cells recognize antigen via the binding of T cell receptors (TCRs) to self or foreign proteins (Braga-Neto and Marques 2006). This phenomenon is dependent on the previous binding of the antigenic peptides to cell-surface glycoproteins, the MHC proteins. The human MHC locus (HLA—human leukocyte antigens) encodes three HLA class I molecules (HLA-A, -B, -C) and three HLA class II molecules (HLA-DR, -DQ, -DP). The structures of the MHC class I and class II molecules complexed with different peptides were solved by X-ray crystallographic studies. CD8⁺ T cell epitopes bound by MHC class I range from 8 to 11 residues, while CD4⁺ T cell epitopes bound to HLA class II range from 11 to more than 20 residues in length. Bound peptides are buried in the antigen-binding groove formed between the helices of the MHC molecule, leaving only a few of their side chains available for direct TCR contact (Bjorkman et al. 1987; Stern et al. 1994). Only nine residues from a given peptide bind to MHC class II within the antigen-binding groove, while flanking residues may also interact with the TCR (Carson et al. 1997). Specific binding of certain peptides to a MHC molecule comes from the interaction of peptide side chains with the irregular surface of the floor and sides of the groove, the “pockets” and ridges formed by the protrusion of MHC residues. The major pockets in the floor of the groove of HLA-DR molecules are occupied by the side chains of residues 1, 4, 6, and 9 of the bound peptide (Stern et al. 1994).

HLA molecules are highly polymorphic. To date, more than 2,215 HLA class I and 986 HLA class II allelic sequences have been identified (Robinson et al. 2009). This polymorphism is concentrated in the region encoding the peptide-binding groove, yielding very diverse amino-acid sequences in this region among different HLA alleles. Thus each pocket in the peptide-binding groove of each HLA molecule is shaped by clusters of polymorphic residues. It follows that each allelic HLA molecule only binds peptides with amino-acid sequences that are capable of interacting with its antigen-binding groove. This implies that T cells from different individuals may be able to recognize distinct epitopes from a given protein.

Escape from Presentation and Recognition by “Molecular Evolution”

For millions of years, pathogens have evolved molecular mechanisms to escape effective presentation and recognition by the immune system. Tumor cells and pathogens evolve accumulating mutations in protein antigens, either flanking or inside an epitope, to escape immune pressure by one of several mechanisms: (1) abrogation of epitope

binding to host MHC or TCR (De Groot et al. 2008), (2) loss of sites for sequence-specific processing proteases (Moudgil et al. 1996), and (3) antagonism or partial agonism of TCR signaling (Franco et al. 1995). Thus the immune response, or lack thereof, against pathogen protein antigens may reflect the evolutionary success of the parasite. Immunization with substituted synthetic peptides or DNA/proteins can present neopeptides or alter the hierarchy of dominant/cryptic T cell epitopes, bypassing escape mechanisms (Cunha-Neto 1999).

Immunodominance

The hallmarks of the adaptive immune response are specificity and memory, for which T cells are indispensable. Protein antigens typically contain multiple epitopes capable of binding MHC class II molecules, but T cell responses are limited to only a small number of these determinants in each individual. The ability of the immune system to regulate and focus T cell responses to a select number of epitopes is termed immunodominance (Berzofsky 1988). As the immune response is always mounted against immunodominant epitopes of an antigen, exposure to these regions will result in efficient priming of the immune system, which can generate protection on subsequent challenge. In other words, immunodominant epitopes, or collections of them, can be potential vaccine candidates. Such epitope-based vaccines exploit small but useful antigenic regions of a protein and ignore portions that are poorly immunogenic or can cause a harmful response (Gowthaman and Agrewala 2008).

A number of mutually nonexclusive hypotheses have been put forth to explain this very restricted specificity of T cells, including: (1) endosomal antigen processing may restrict the array of peptides available to recruit CD4⁺ T cells, (2) MHC molecules may be able to bind only a limited subset of antigenic peptides that are released during antigen processing, the so-called determinant selection and (3) the TCR repertoire may be limited and only able to detect some MHC:peptide combinations (holes in the repertoire) (Sant et al. 2005). Processing reactions within an antigen-presenting cell may also influence immunodominance since gross changes in antigen structure may modulate the efficiency of an epitope’s presentation (Li et al. 2009).

While dominant peptides efficiently elicit recall responses in animals primed with whole antigens, non-dominant or cryptic peptides are only immunogenic when directly administered in vivo. Epitope immunodominance is an MHC-restricted phenomenon (i.e. individuals with distinct MHCs will select different immunodominant

epitopes). Immunodominance of certain pathogen epitopes can also induce a detrimental effect by restricting the breadth of recognized T cell epitopes. In a recent phase I clinical trial, subjects immunized with a recombinant adenovirus 5 vector encoding the whole human immunodeficiency virus (HIV-1) proteins Gag, Pol, and Nef recognized on average a single epitope per HIV protein, a number clearly insufficient to provide protection to the highly variable HIV-1 (Watkins et al. 2008).

The Importance of CD4⁺ T Cell Help

CD4⁺ T cells play a central role in a functional adaptive immune response. They promote the optimal expansion of cytotoxic CD8⁺ T cells, maintain CD8⁺ T cell memory, and communicate with innate immune cells. Furthermore, they promote B cell differentiation into plasma cells to produce neutralizing antibodies and assist memory B cells for a swift recall response to re-infection (Yang and Yu 2009). The help provided by CD4⁺ T cells is essential for the generation of a robust primary and memory CD8⁺ T cell response and protective immunity against various viral and bacterial infections (Bevan 2004; Novy et al. 2007). Besides, CD4⁺ T cells can themselves act as antiviral effector cells either by killing infected cells directly or by secretion of antiviral cytokines, such as IFN- γ and TNF- α , and can become memory helper T cells. The importance of CD4⁺ T cells in immunity to infection is further underscored in that mice with absent or defective CD4⁺ T cell help have an impaired ability to clear viral and protozoan pathogens, such as lymphocytic choriomeningitis virus (Khanolkar et al. 2004), herpes simplex virus-1 (Rajasagi et al. 2009), and *Plasmodium* (Xu et al. 2002).

The mechanisms underlying CD4⁺ T cell help to CD8⁺ T cells are not completely understood, but recent evidence suggests that CD4⁺ T cells facilitate the activation and development of CD8⁺ T cell responses either directly through the provision of cytokines or by the major pathway, i.e. dendritic cell (DC) licensing (Lanzavecchia and Sallusto 2001; Smith et al. 2004; Wodarz and Jansen 2001). In DC licensing, the CD40 ligand-CD40 interaction results in IL-12 and IL-15 production and up-regulation of costimulatory molecules, which leads to subsequent activation and maturation of DCs, making them competent to stimulate an antigen-specific CD8⁺ T cell response (Bennett et al. 1998; Ridge et al. 1998; Schoenberger et al. 1998; Smith et al. 2004; Zhang et al. 2009). In addition to signaling via surface molecules, CD4⁺ T cell-derived IL-2 is an important component of help for CD8⁺ T cell immunity to pathogens (Livingstone et al. 2009).

Viruses with a tropism for helper T cells, such as HIV, can potentially impair the CD4⁺ cell response, resulting in

compromised cytotoxic T lymphocyte (CTL) activity and persistent infection. The importance of HIV-specific CD4⁺ T cells and its association with a protective antiviral immune response has been demonstrated (Gandhi and Walker 2002). Preservation of memory CD4⁺ T cells correlated with primate survival after challenge with simian immunodeficiency virus (SIV) (Letvin et al. 2006), further showing the importance of memory CD4⁺ T cells in protection.

Tumor-associated antigen-specific CD4⁺ T cell helper activity is exerted for the induction and maintenance of CTLs, in addition to other immune cells. Moreover, CD4⁺ T cells can also possess cytotoxic ability and directly kill tumor cells (Guo et al. 2005) and thus may have a role in cancer immunotherapy (Ohkuri et al. 2009). Taken together, evidence indicates that effective CD4⁺ T cell help or direct effector action is essential for anti-infection or anticancer immunity.

Prediction of CD4⁺ T Cell Epitopes

Given the fundamental role of CD4⁺ T cell function on anti-infection/anticancer immunity, the identification and characterization of CD4⁺ T cell epitopes is a crucial step in vaccine design. It is also important for studying the immunobiology of autoimmunity, allergy, and transplantation along with immune diagnostics (Moise and De Groot 2006; Valentino and Frelinger 2009). A conventional approach to identifying T cell epitopes is to synthesize many overlapping peptides (usually 15-mers) spanning the full length of the target antigen and test for immunogenicity using T cell assays. However, this approach is time consuming, expensive (especially for whole-proteome analyses), and may not disclose all the longer CD4⁺ T cell epitopes. On the other hand, bioinformatic/in silico tools can predict which peptides are more likely to contain T cell epitopes, greatly reducing the number of candidate sequences (Gowthaman and Agrewala 2008).

T cell epitope prediction dates back to the 1980s, when the first algorithm was developed based on the identification of amphipathic helical regions on protein antigens (Berzofsky et al. 1987). Since then, new methods based initially on MHC peptide-binding motifs and, more recently, on MHC-binding properties based on scores calculated from actual peptide-binding assays have been developed. MHC class II binding prediction methods are more complex than those for MHC-I binding (Yang and Yu 2009).

A range of bioinformatic algorithms have been developed to predict MHC class II epitopes (Table 1). Since the most selective requirement for a peptide to be immunogenic is its ability to bind to the MHC molecule, most prediction methods focus on this stage of the pathway. The algorithms

Table 1 Examples of in silico tools for class II epitope prediction

Program	Prediction algorithm	URL	References
TEPITOPE	Matrix	http://www.vaccinome.com	Sturniolo et al. (1999)
ProPred	Matrix	http://www.imtech.res.in/raghava/propred/	Singh and Raghava (2003)
NetMHCII pan	ANN	http://www.cbs.dtu.dk/services/NetMHCIIpan	Nielsen et al. (2008)
MULTIPRED	ANN	http://antigen.i2r.a-star.edu.sg/multipred1	Zhang et al. (2007)
EpiMatrix	Matrix	http://www.epivax.com	De Groot et al. (1997)
RANKPEP	Matrix	http://bio.dfci.harvard.edu/Tools/rankpep.html	Reche et al. (2002)
PRED ^{BALB/c}	Matrix	http://antigen.i2r.a-star.edu.sg/predBalbc	Zhang et al. (2005)
SYFPEITHI	Matrix	http://www.syfpeithi.de	Rammensee et al. (1999)
MHC Pred	Matrix	http://www.jenner.ac.uk/MHCPred	Guan et al. (2003)

ANN artificial neural networks, URL uniform resource locator

that are used vary in complexity and accuracy. Several of them rely on the fact that most pockets in the MHC class II binding groove are shaped by clusters of polymorphic residues and thus have distinct chemical and size characteristics in different MHC class II alleles. MHC class II binding prediction methods are categorized into two main groups: quantitative and qualitative. Qualitative matrices determine the binding status (whether a peptide is a “binder” or “non-binder”) based on the predictive score (based on position-specific binding profiles) (e.g. SYFPEITHI, RANKPEP, MULTIPRED), while quantitative approaches predict the strength of binding as well (e.g. TEPITOPE, PROPREP, NetMHCII pan, SVRMHC, ARB, and SMM-align) (Rajapakse et al. 2007; Wan et al. 2007). Several studies have compared the performances of MHC class II binding prediction methods (Gowthaman and Agrewala 2008; Lin et al. 2008; Wang et al. 2008). Briefly, the conclusions of these studies were similar, indicating that despite the difference in datasets used, the contemporary methods have shown little improvement over the older TEPITOPE algorithm (Lin et al. 2008). The TEPITOPE HLA-DR binding prediction algorithm (Sturniolo et al. 1999) uses the concept that each HLA-DR pocket can be characterized by “pocket profiles,” a quantitative representation of the interaction of all natural amino-acid residues with a given pocket (Hammer et al. 1994). A small database of pocket profiles was sufficient to generate a large number of virtual HLA-DR matrices representing a significant proportion of HLA-DR peptide-binding specificity. Such matrices were incorporated in the TEPITOPE software. For each HLA-DR specificity, TEPITOPE generated a binding score corresponding to the algebraic sum of the strength of interaction between each residue and pocket, which correlated to binding affinity. Peptide scores along a scanned protein sequence were normalized for each HLA-DR as the proportion of the best binder peptides. Since the software displays a significant number of HLA-DR specificities, it is also capable of predicting promiscuous HLA class II ligands (Panigada et al.

2002). The TEPITOPE prediction model has been successfully applied to the identification of T cell epitopes in the context of several human diseases (Bian and Hammer 2004).

The applications of computational HLA class II epitope prediction include vaccine candidate discovery, the study of pathogenesis (infectious diseases, autoimmunity, and cancer), allergy treatment, drug development, engineering of therapeutic proteins, and diagnostics. A brief account of some studies identifying epitopes with the aid of prediction algorithms performed over the last decade is presented in Table 2.

Our group and others have successfully used TEPITOPE for mapping promiscuous CD4⁺ T cell epitopes in different pathogen antigens, including HIV-1 whole proteome (Fonseca et al. 2006), *Schistosoma mansoni* Sm14 and paramyosin (Fonseca et al. 2005a; Fonseca et al. 2005b), *Plasmodium vivax* MSP-1 (Rosa et al. 2006), *Paracoccidioides brasiliensis* gp43 (Iwai et al. 2007), several *Mycobacterium tuberculosis* proteins, cytomegalovirus pp65 and glyP86, and SIV whole proteome (unpublished observations). Peptide-HLA DR binding assays confirmed the capacity of the predicted epitopes to bind to several HLA class II molecules in that most peptides bound to at least 50% of the HLA-DR molecules tested. Furthermore, a significant correlation was observed between the TEPITOPE-predicted promiscuity (i.e. the number of HLA-DR molecules predicted to bind to a certain peptide) and the promiscuity observed in binding assays to multiple HLA-DR molecules (Fig. 1). Individual peptides were recognized by 25–50% of patients; combined T cell recognition of such peptides was detected in a high proportion of patients (75–90%).

Predicted CD4⁺ Epitopes and Vaccine Design

Immunization strategies that focus solely on CD8⁺ T cell immunity might prove to be insufficient because even though they might stimulate vigorous early responses, they

Table 2 Applications of computational class II epitope prediction

Applications of computational class II epitope prediction	Antigen mapped*	References
Infectious diseases pathogenesis and vaccine design	Whole HIV-1 proteome	Fonseca et al. (2006)
	HIV-1 (Gag, Pol, Nef, Rev, Tat, Vif, Vpr, and Vpu)	Wilson et al. (2001)
	Whole HIV proteome	Cohen et al. (2006)
	Merozoite surface protein 1 (<i>Plasmodium vivax</i>)	Rosa et al. (2006)
	Duffy binding protein (<i>Plasmodium vivax</i>)	Saravia et al. (2008)
	Pre-erythrocytic-stage antigens (CSP, SSP2, EXP-1, LSA-1) (<i>Plasmodium falciparum</i>)	Doolan et al. (2000)
	Translated poxvirus genome sequences	Mitra-Kaushik et al. (2007); Calvo-Calle et al. (2007)
	MPT63 (<i>Mycobacterium tuberculosis</i>)	Mustafa (2009)
	Chlamydial genome	Barker et al. (2008)
	Hemagglutinin (influenza)	Panigada et al. (2002)
	Glycoprotein 43 (gp43) (<i>Paracoccidioides brasiliensis</i>)	Iwai et al. (2007)
	Fatty acid-binding protein (Sm14) (<i>Schistosoma mansoni</i>)	Fonseca et al. (2005a); Garcia et al. (2008)
	Paramyosin (<i>Schistosoma mansoni</i>)	Fonseca et al. (2005b)
Oncology pathogenesis and vaccine design	EBNA1, LMP1, LMP2 (EBV latency II malignancies)	Depil et al. (2007)
	MAGEA-4	Ohkuri et al. (2009)
	MAGE-6	Tatsumi et al. (2003)
	MAGE-3	Kobayashi et al. (2001)
	CEA (carcinoembryonic antigen)	Kobayashi et al. (2002)
	Kallikrein-4 (prostate-specific antigen)	Hural et al. (2002)
	PAP (prostatic acid phosphatase)	McNeel et al. (2001)
Allergy novel immunotherapy strategies	<i>Lol p5</i> (rye grass pollen)	de Lalla et al. (1999)
Autoimmunity	TSHR (TSH receptor, Grave's disease)	Inaba et al. (2006)
	hTg (human thyroglobulin- Thyroiditis)	Flynn et al. (2004)
	Tyrosine phosphatase IA-2 (Type I diabetes)	Honeyman et al. (1997)
	TYR, TRP1, TRP2, and Pmel17 (Vogt-Koyanagi-Hatada's uveitis)	Damico et al. (2005)
Engineering of new therapeutics	Human FPX	Koren et al. (2007)
Diagnostics	GAD65 (Type I diabetes)	Mallone and Nepom (2004)

* Protein, used in prediction algorithms to identify potential MHC-II binding peptides

will be unable to confer long-term protective immunity (Khanolkar et al. 2007). Although considerable information has been gathered on CD8⁺ T cell epitopes, comparatively few CD4⁺ T cell epitopes have been identified so far, no matter whether the origin is pathogens or tumor-associated proteins. The rational selection of protein sequences that function as promiscuous CD4⁺ epitopes in vaccine formulations is crucial for successful application of this vaccination strategy. Given their major role in determining the functional status and memory of effector responses, appropriate CD4⁺ T cell epitopes should be an essential part of any candidate vaccine. In this scenario it is thus necessary to identify the still missing CD4⁺ T cell epitopes recognized by the majority of individuals.

The advent of whole-genome sequencing and advances in bioinformatics marked the beginning of a new era that

approaches vaccine development starting from genomic information, a process named “reverse vaccinology” (Rappuoli 2000). In just a few years, reverse vaccinology applied to *Neisseria meningitidis* has resulted in the identification of more vaccine candidates than those discovered during the previous four decades of research by conventional methods (Pizza et al. 2000; Serruto and Rappuoli 2006). A related approach, termed “epitope-driven vaccine design” (De Groot et al. 2001), originally named “reverse immunogenetics” by Davenport and Hill, employs T cell-epitope mapping tools for finding new protein candidates for vaccines and diagnostic tests (Davenport and Hill 1996) (Fig. 2). Epitope-driven vaccine design allows the discovery of previously unknown and undescribed antigens and epitopes as vaccine candidates. Furthermore, predicted CD4⁺ T cell epitopes should be validated using in vitro

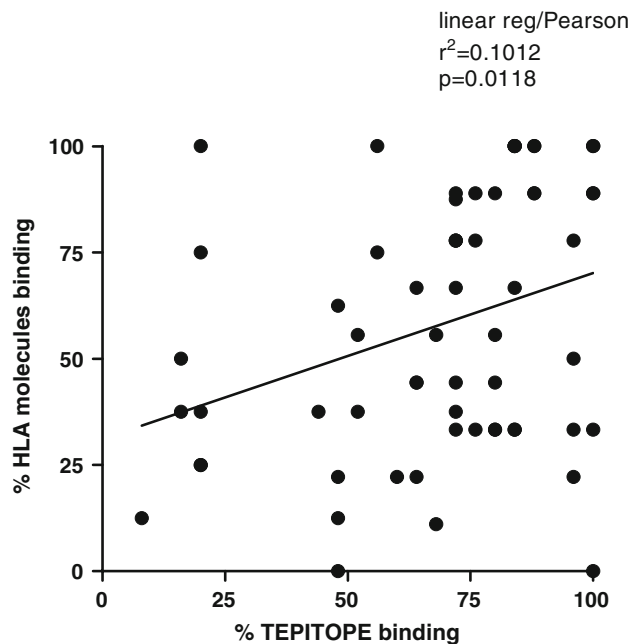


Fig. 1 Correlation between the prediction of promiscuous epitopes by TEPITOPE and the promiscuity assessed by binding assays to nine prevalent HLA-DR molecules. TEPITOPE prediction was performed using peptides selected at a 3% threshold (25 HLA-DR molecules scanned). Data from peptides derived from *Schistosoma mansoni* (paramyosin and Sm14); *Paracoccidioides braziliensis* (gp43); *Mycobacterium tuberculosis* (groEL2, PBP-1, 19kAg, 19Kmp, mtp40, and mce-1) and cytomegalovirus (pp65 and glyP86). HLA-DR binding data (DRB1*0101, *0301, *0401, *0405, *0701, *1101, *1302, *1501, DRB5*0101) were generated by Alessandro Sette and John Sidney (La Jolla Institute of Allergy and Immunology, USA)

evaluation of HLA binding, ex vivo assays with T cells from sensitized hosts (ELISpot assays, MHC-tetramers, flow cytometry for cytokine production and proliferation), and in vivo validation using HLA class II transgenic mice.

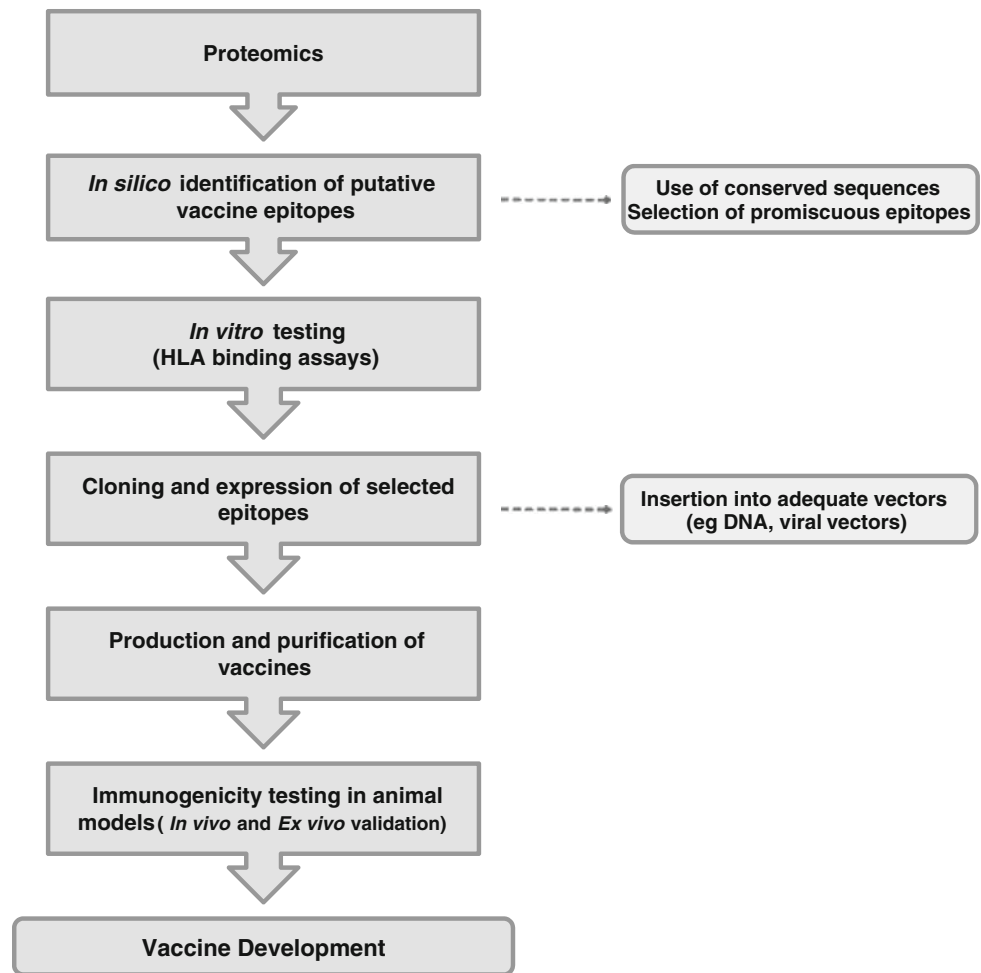
The rational design of promiscuous epitopes has been applied to create an artificial T_H-cell epitope by inserting anchor residues for diverse HLA-DR molecules, yielding a powerful pan-DR epitope (PADRE) that is presented by multiple HLA DR molecules (Alexander et al. 1994). The use of synthetic epitopes such as PADRE to optimize CD8⁺ function (Kim et al. 2008) and antibody production (Rosa et al. 2004) holds promise for a new generation of highly efficacious vaccines.

The epitope-based vaccine approach allows focusing immune responses on relevant epitopes, the rational engineering of epitopes for increased potency and breadth, as well as increasing safety (Sette and Fikes 2003). The major disadvantage of the epitope-based approach is that algorithms may fail to predict all the relevant epitopes (Iwai et al. 2003). Moreover, only a limited set of HLA class I specificities have algorithms designed to them. Furthermore, whole-protein immunization resembles infection-associated immunity and may offer epitopes that can be

recognized by patients bearing such “uncharted” HLA class I specificities. This may be minimized, however, by the use of a significant number of longer, promiscuous CD4⁺ epitopes containing CD8⁺ T cell epitopes.

The use of the available matrix-based MHC-binding algorithms allows the potential selection of high-affinity binding peptides, the ones with a higher chance of eliciting T cell responses. From the vaccine immunologist’s point of view, however, the identification of MHC allele-specific T cell epitopes may not be enough, since one is searching for epitopes for vaccine epitopes that can effectively cover the human population, with its large diversity of HLA alleles (Wilson et al. 2001). This implies the identification of multiple “promiscuous” epitopes that can bind to multiple distinct HLA alleles (Sturniolo et al. 1999) whose combined frequency in the population approaches 100%. The alignment of peptides binding to several distinct HLA-DR molecules with TEPITOPE or other quantitative matrix MHC class II prediction algorithm can identify such potentially “promiscuous” epitopes. Since many pathogens exhibit high mutation rates, it is important to select conserved protein sequences for scanning with MHC-binding algorithms (Khan et al. 2006).

Our group recently identified a group of such epitopes from the whole proteome of the HIV-1 B-subtype consensus sequence using the TEPITOPE algorithm. This led to the identification of 18 sequences predicted to bind significantly to at least two-thirds of the HLA-DR molecules covered by the TEPITOPE algorithm, representing ca. 90% of the Caucasian population (Fonseca et al. 2006). Binding assays of the 18 selected peptides with the 9 most prevalent HLA-DR molecules in the general population confirmed the ability of the selected peptides to bind to multiple HLA-DR molecules. Peripheral blood mononuclear cells (PBMCs) from over 90% of a group of HIV-1-infected patients recognized at least one of the promiscuous peptides. All 18 peptides were recognized, and the PBMCs from most of the patients recognized multiple peptides, demonstrating that the epitopes are presented in vivo during infection. Similar responses were obtained in CD8⁺ T cell-depleted PBMCs. Together, this may suggest that this epitope combination could have good potential for use as an immunogen against HIV. Polyepitopic recombinant vaccines, containing multiple epitopes in tandem, may alter the hierarchy of dominant epitopes, circumventing the mechanisms of escape from processing and presentation built into sequences of native viral proteins (Fu et al. 1997). In addition, a polyepitopic vaccine in which each epitope is promiscuous increases the chances that each individual in a genetically heterogeneous population acquires immunity to multiple epitopes, a necessary measure in a vaccine against the highly polymorphic HIV virus. We thus constructed a polyepitopic DNA vaccine encoding the 18 HIV-1 CD4⁺ T

Fig. 2 Epitope-driven vaccine design flowchart

cell epitopes. Preliminary data from immunogenicity studies in mice indicate the vaccine elicits a powerful CD4⁺ and CD8⁺ T cell response, detectable against all the contained epitopes (unpublished observations). Significantly, this showed that epitope prediction by TEPITOPE can also predict epitopes recognized in nonhuman species. Rosa et al. (2006) used the TEPITOPE algorithm to identify a promiscuous CD4⁺ T cell epitope in *P. vivax* merozoite surface protein-1. Mice immunized with a recombinant *P. vivax* MSP1₁₉ protein containing the above promiscuous epitope developed significantly higher anti-MSP19 IgG antibody titers than those immunized with intact MSP1₁₉.

In addition to immunogenicity studies, several recent studies suggest that CD4⁺ T cell epitope-based vaccines may indeed have an increased protective effect. Mice immunized with TEPITOPE-selected peptides derived from *S. mansoni* Sm14 protein displayed partial protection against infectious challenge (Garcia et al. 2008). Barker et al. (2008) scanned the *Chlamydia* proteome with the SYFPEITHI and PRO-PRED algorithms to identify proteins with promiscuous epitopes. Adoptive transfer of antigen-primed CD4⁺ T cells

was found to confer significant reduction in shedding of chlamydial organisms and duration of infection. Moreover, serum from immunized mice was found to neutralize *Chlamydia* infection of a cell monolayer in vitro, demonstrating the importance of CD4 help in generating appropriate antibody responses. Despite the importance of CD4⁺ T cells, it must be kept in mind that antibodies, CD8⁺ T cells, and the delayed-type hypersensitivity reaction are the major end-effectors of acquired immunity. Thus CD4⁺ T cell-based vaccines must necessarily engage these effectors to achieve any degree of protection.

Conclusion

Increasing evidence demonstrates the critical role of CD4⁺ T cells in natural antigen encounter and immunization. We have reviewed the essential role of CD4⁺ T cell responses in the acquired immune response and the unavoidable need to include a CD4 epitope component to provide cognate help in vaccines that aim also to induce cytotoxic CD8⁺ T cells and neutralizing antibodies. In light of the availability of epitope

prediction tools, and the recent demonstration of the important immunogenicity of epitope-based vaccines, we anticipate that the near future will show an increasing number of candidate vaccines containing cassettes of promiscuous, conserved CD4⁺ T cell epitopes, with consequent improvement of immunogenicity and protection.

Acknowledgments This work was supported by the Brazilian National Research Council (CNPq), the São Paulo State Research Funding Agency (FAPESP), International Centre of Genetic Engineering and Biotechnology (ICGEB), the Ministry of Health (Brazil), and the National Institutes of Health/NIAID (grant number R03 AI 66961-03).

References

- Alexander J, Sidney J, Southwood S et al (1994) Development of high potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides immunity 1:751–761
- Barker CJ, Beagley KW, Hafner LM et al (2008) In silico identification and in vivo analysis of a novel T-cell antigen from *Chlamydia*, NrdB. *Vaccine* 26:1285–1296
- Bennett SR, Carbone FR, Karamalis F et al (1998) Help for cytotoxic-T-cell responses is mediated by CD40 signalling. *Nature* 393:478–480
- Berzofsky JA (1988) Immunodominance in T lymphocyte recognition. *Immunol Lett* 18:83–92
- Berzofsky JA, Cease KB, Cornette JL et al (1987) Protein antigenic structures recognized by T cells: potential applications to vaccine design. *Immunol Rev* 98:9–52
- Bevan MJ (2004) Helping the CD8(+) T-cell response. *Nat Rev Immunol* 4:595–602
- Bian H, Hammer J (2004) Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE. *Methods* 34:468–475
- Bjorkman PJ, Saper MA, Samraoui B et al (1987) The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329:512–518
- Braga-Neto UM, Marques ET Jr (2006) From functional genomics to functional immunomics: new challenges, old problems, big rewards. *PLoS Comput Biol* 2:e81
- Calvo-Calle JM, Strug I, Nastke MD et al (2007) Human CD4+ T cell epitopes from vaccinia virus induced by vaccination or infection. *PLoS Pathog* 3:1511–1529
- Carson RT, Vignali KM, Woodland DL et al (1997) T cell receptor recognition of MHC class II-bound peptide flanking residues enhances immunogenicity and results in altered TCR V region usage. *Immunity* 7:387–399
- Cohen WM, Pouvelle-Moratille S, Wang XF et al (2006) Scanning the HIV genome for CD4+ T cell epitopes restricted to HLA-DP4, the most prevalent HLA class II molecule. *J Immunol* 176:5401–5408
- Cunha-Neto E (1999) MHC-restricted antigen presentation and recognition: constraints on gene, recombinant and peptide vaccines in humans. *Braz J Med Biol Res* 32:199–205
- Damico FM, Cunha-Neto E, Goldberg AC, Iwai LK, Marin ML, Hammer J, Kalil J, Yamamoto JH (2005) T cell recognition and cytokine profile induced by melanocyte epitopes in HLA-DRB1*0405-positive and negative Vogt-Koyanagi-Harada uveitis patients. *Invest Ophthalmol Vis Sci* 46:2465–2471
- Davenport MP, Hill AV (1996) Reverse immunogenetics: from HLA-disease associations to vaccine candidates. *Mol Med Today* 2:38–45
- De Groot AS, Jesdale BM, Szu E et al (1997) An interactive web site providing major histocompatibility ligand predictions: application to HIV research. *AIDS Res Hum Retrovir* 13:529–531
- De Groot AS, Bosma A, Chinai N et al (2001) From genome to vaccine: in silico predictions, ex vivo verification. *Vaccine* 19:4385–4395
- De Groot AS, McMurry J, Moise L (2008) Prediction of immunogenicity: in silico paradigms, ex vivo and in vivo correlates. *Curr Opin Pharmacol* 8:620–626
- de Lalla C, Sturniolo T, Abbruzzese L et al (1999) Cutting edge: identification of novel T cell epitopes in Lol p5a by computational prediction. *J Immunol* 163:1725–1729
- Depil S, Moralès O, Castelli FA et al (2007) Determination of a HLA II promiscuous peptide cocktail as potential vaccine against EBV latency II malignancies. *J Immunother* 30:215–226
- Doolan DL, Southwood S, Chesnut R et al (2000) HLA-DR-promiscuous T cell epitopes from plasmodium falciparum pre-erythrocytic-stage antigens restricted by multiple HLA class II alleles. *J Immunol* 165:1123–1137
- Flynn JC, McCormick DJ, Brusica V et al (2004) Pathogenic human thyroglobulin peptides in HLA-DR3 transgenic mouse model of autoimmune thyroiditis. *Cell Immunol* 229:79–85
- Fonseca CT, Cunha-Neto E, Goldberg AC et al (2005a) Human T cell epitope mapping of the *Schistosoma mansoni* 14 kDa fatty acid-binding protein using cells from patients living in areas endemic for schistosomiasis. *Microbes Infect* 7:204–212
- Fonseca CT, Cunha-Neto E, Goldberg AC et al (2005b) Identification of paramyosin T cell epitopes associated with human resistance to *Schistosoma mansoni* infection. *Clin Exp Immunol* 142:539–547
- Fonseca SG, Coutinho-Silva A, Fonseca LA et al (2006) Identification of novel consensus CD4 T-cell epitopes from clade B HIV-1 whole genome that are frequently recognized by HIV-1 infected patients. *AIDS* 20:2263–2273
- Franco A, Ferrari C, Sette A et al (1995) Viral mutations, TCR antagonism and escape from the immune response. *Curr Opin Immunol* 7:524–531
- Fu TM, Friedman A, Ulmer JB et al (1997) Protective cellular immunity: cytotoxic T-lymphocyte responses against dominant and recessive epitopes of influenza virus nucleoprotein induced by DNA immunization. *J Virol* 71:2715–2721
- Gandhi RT, Walker BD (2002) Immunologic control of HIV-1. *Annu Rev Med* 53:149–172
- Garcia TC, Fonseca CT, Pacifico LG et al (2008) Peptides containing T cell epitopes, derived from Sm14, but not from paramyosin, induce a Th1 type of immune response, reduction in liver pathology and partial protection against *Schistosoma mansoni* infection in mice. *Acta Trop* 106:162–167
- Gowthaman U, Agrewala JN (2008) In silico tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion. *J Proteome Res* 7:154–163
- Guan P, Doytchinova IA, Zygouri C et al (2003) MHCpred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl Bioinformatics* 2:63–66
- Guo Y, Niiya H, Azuma T et al (2005) Direct recognition and lysis of leukemia cells by WT1-specific CD4+ T lymphocytes in an HLA class II-restricted manner. *Blood* 106:1415–1418
- Hammer J, Bono E, Gallazzi F et al (1994) Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning. *J Exp Med* 180:2353–2358
- Honeyman MC, Brusica V, Harrison LC (1997) Strategies for identifying and predicting islet autoantigen T-cell epitopes in insulin-dependent diabetes mellitus. *Ann Med* 29:401–404
- Hural JA, Friedman RS, McNabb A et al (2002) Identification of naturally processed CD4 T cell epitopes from the prostate-specific

- antigen kallikrein 4 using peptide based in vitro stimulation. *J Immunol* 169:557–565
- Inaba H, Martin W, De Groot AS et al (2006) Thyrotropin receptor epitopes and their relation to histocompatibility leukocyte antigen-DR molecules in Graves' disease. *J Clin Endocrinol Metab* 91:2286–2294
- Iwai LK, Yoshida M, Sidney J et al (2003) In silico prediction of peptides binding to multiple HLA-DR molecules accurately identifies immunodominant epitopes from gp43 of *Paracoccidioides brasiliensis* frequently recognized in primary peripheral blood mononuclear cell responses from sensitized individuals. *Mol Med* 9:209–219
- Iwai LK, Yoshida M, Sadahiro A et al (2007) T-cell recognition of *Paracoccidioides brasiliensis* gp43-derived peptides in patients with paracoccidioidomycosis and healthy individuals. *Clin Vaccine Immunol* 14:474–476
- Khan AM, Miotto O, Heiny AT et al (2006) A systematic bioinformatics approach for selection of epitope-based vaccine targets. *Cell Immunol* 244:141–147
- Khanolkar A, Fuller MJ, Zajac AJ (2004) CD4 T cell-dependent CD8 T cell maturation. *J Immunol* 172:2834–2844
- Khanolkar A, Badovinac VP, Harty JT (2007) CD8 T cell memory development: CD4 T cell help is appreciated. *Immunol Res* 39:94–104
- Kim D, Hoory T, Monie A et al (2008) Enhancement of DNA vaccine potency through coadministration of CIITA DNA with DNA vaccines via gene gun. *J Immunol* 180:7019–7027
- Kobayashi H, Song Y, Hoon DS et al (2001) Tumor-reactive T helper lymphocytes recognize a promiscuous MAGE-A3 epitope presented by various major histocompatibility complex class II alleles. *Cancer Res* 61:4773–4778
- Kobayashi H, Omiya R, Ruiz M et al (2002) Identification of an antigenic epitope for helper T lymphocytes from carcinoembryonic antigen. *Clin Cancer Res* 8:3219–3225
- Koren E, De Groot AS, Jawa V et al (2007) Clinical validation of the “in silico” prediction of immunogenicity of a human recombinant therapeutic protein. *Clin Immunol* 124:26–32
- Lanzavecchia A, Sallusto F (2001) Antigen decoding by T lymphocytes: from synapses to fate determination. *Nat Immunol* 2:487–492
- Letvin NL, Mascola JR, Sun Y et al (2006) Preserved CD4+ central memory T cells and survival in vaccinated SIV-challenged monkeys. *Science* 312:1530–1533
- Li H, Xu CF, Blais S et al (2009) Proximal glycans outside of the epitopes regulate the presentation of HIV-1 envelope gp120 helper epitopes. *J Immunol* 182:6369–6378
- Lin HH, Zhang GL, Tongchusak S et al (2008) Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. *BMC Bioinformatics* 9(suppl 12):S22
- Livingstone AM, Wilson EB, Ontiveros F, Wang JC (2009) Unravelling the mechanisms of help for CD8+ T cell responses. *Immunol Res* 45(2–3):209–217
- Mallone R, Nepom GT (2004) MHC Class II tetramers and the pursuit of antigen-specific T cells: define, deviate, delete. *Clin Immunol* 110:232–242
- McNeel DG, Nguyen LD, Disis ML (2001) Identification of T helper epitopes from prostatic acid phosphatase. *Cancer Res* 61:5161–5167
- Mitra-Kaushik S, Cruz J, Stern LJ et al (2007) Human cytotoxic CD4+ T cells recognize HLA-DR1-restricted epitopes on vaccinia virus proteins A24R and D1R conserved among poxviruses. *J Immunol* 179:1303–1312
- Moise L, De Groot AS (2006) Putting immunoinformatics to the test. *Nat Biotechnol* 24:791–792
- Moudgil KD, Deng H, Nanda NK et al (1996) Antigen processing and T cell repertoires as crucial aleatory features in induction of autoimmunity. *J Autoimmun* 9:223–227
- Mustafa AS (2009) Th1 cell reactivity and HLA-DR binding prediction for promiscuous recognition of MPT63 (Rv1926c), a major secreted protein of *Mycobacterium tuberculosis*. *Scand J Immunol* 69:213–222
- Nielsen M, Lundegaard C, Blicher T et al (2008) Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. *PLoS Comput Biol* 4:e1000107
- Novy P, Quigley M, Huang X et al (2007) CD4 T cells are required for CD8 T cell survival during both primary and memory recall responses. *J Immunol* 179:8243–8251
- Ohkuri T, Wakita D, Chamoto K et al (2009) Identification of novel helper epitopes of MAGE-A4 tumour antigen: useful tool for the propagation of Th1 cells. *Br J Cancer* 100:1135–1143
- Panigada M, Sturniolo T, Besozzi G et al (2002) Identification of a promiscuous T-cell epitope in *Mycobacterium tuberculosis* Mce proteins. *Infect Immun* 70:79–85
- Pizza M, Scarlato V, Masignani V et al (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 287:1816–1820
- Rajapakse M, Schmidt B, Feng L et al (2007) Predicting peptides binding to MHC class II molecules using multi-objective evolutionary algorithms. *BMC Bioinformatics* 8:459
- Rajasagi NK, Kassim SH, Kollias CM et al (2009) CD4+ T cells are required for the priming of CD8+ T cells following infection with herpes simplex virus type 1. *J Virol* 83:5256–5268
- Rammensee H, Bachmann J, Emmerich NP et al (1999) SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 50:213–219
- Rappuoli R (2000) Reverse vaccinology. *Curr Opin Microbiol* 3:445–450
- Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol* 63:701–709
- Ridge JP, Di Rosa F, Matzinger P (1998) A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. *Nature* 393:474–478
- Robinson J, Waller MJ, Fail SC et al (2009) The IMGT/HLA database. *Nucleic Acids Res* 37(Database issue):D1013–D1017
- Rosa DS, Tzelepis F, Cunha MG et al (2004) The pan HLA DR-binding epitope improves adjuvant-assisted immunization with a recombinant protein containing a malaria vaccine candidate. *Immunol Lett* 92:259–268
- Rosa DS, Iwai LK, Tzelepis F et al (2006) Immunogenicity of a recombinant protein containing the *Plasmodium vivax* vaccine candidate MSP1(19) and two human CD4+ T-cell epitopes administered to non-human primates (*Callithrix jacchus jacchus*). *Microbes Infect* 8:2130–2137
- Sant AJ, Chaves FA, Jenks SA et al (2005) The relationship between immunodominance, DM editing, and the kinetic stability of MHC class II:peptide complexes. *Immunol Rev* 207:261–278
- Saravia C, Martinez P, Granados DS et al (2008) Identification and evaluation of universal epitopes in *Plasmodium vivax* Duffy binding protein. *Biochem Biophys Res Commun* 377:1279–1283
- Schoenberger SP, Toes RE, van der Voort EI et al (1998) T-cell help for cytotoxic T lymphocytes is mediated by CD40-CD40L interactions. *Nature* 393:480–483
- Serruto D, Rappuoli R (2006) Post-genomic vaccine development. *FEBS Lett* 580:2985–2992
- Sette A, Fikes J (2003) Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Curr Opin Immunol* 15:461–470
- Singh H, Raghava GP (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 19:1009–1014
- Smith CM, Wilson NS, Waithman J et al (2004) Cognate CD4(+) T cell licensing of dendritic cells in CD8(+) T cell immunity. *Nat Immunol* 5:1143–1148

- Stern LJ, Brown JH, Jardetzky TS et al (1994) Crystal structure of the human class II MHC protein MHC-DR1 complexed with an influenza virus peptide. *Nature* 368:215–221
- Sturniolo T, Bono E, Ding J et al (1999) Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol* 17:555–561
- Tatsumi T, Kierstead LS, Ranieri E et al (2003) MAGE-6 encodes HLA-DRbeta1*0401-presented epitopes recognized by CD4+ T cells from patients with melanoma or renal cell carcinoma. *Clin Cancer Res* 9:947–954
- Valentino M, Frelinger J (2009) An approach to the identification of T cell epitopes in the genomic era: application to *Francisella tularensis*. *Immunol Res* 45(2–3):218–228
- Wan XF, Chen G, Luo F et al (2007) A quantitative genotype algorithm reflecting H5N1 Avian influenza niches. *Bioinformatics* 23:2368–2375
- Wang P, Sidney J, Dow C et al (2008) A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol* 4:e1000048
- Watkins DI, Burton DR, Kallas EG et al (2008) Nonhuman primate models and the failure of the Merck HIV-1 vaccine in humans. *Nat Med* 14:617–621
- Wilson CC, Palmer B, Southwood S et al (2001) Identification and antigenicity of broadly cross-reactive and conserved human immunodeficiency virus type 1-derived helper T-lymphocyte epitopes. *J Virol* 75:4195–4207
- Wodarz D, Jansen VA (2001) The role of T cell help for anti-viral CTL responses. *J Theor Biol* 211:419–432
- Xu H, Wipasa J, Yan H et al (2002) The mechanism and significance of deletion of parasite-specific CD4(+) T cells in malaria infection. *J Exp Med* 195:881–892
- Yang X, Yu X (2009) An introduction to epitope prediction methods and software. *Rev Med Virol* 19:77–96
- Zhang GL, Srinivasan KN, Veeramani A et al (2005) PREDBALB/c: a system for the prediction of peptide binding to H2d molecules, a haplotype of the BALB/c mouse. *Nucleic Acids Res* 33(Web Server issue):W180–W183
- Zhang GL, Khan AM, Srinivasan KN et al (2007) Neural models for predicting viral vaccine targets. *J Bioinform Comput Biol* 3:1207–1225
- Zhang S, Zhang H, Zhao J (2009) The role of CD4 T cell help for CD8 CTL activation. *Biochem Biophys Res Commun* 384:405–408